Proceedings of the
1st Conference on Production Systems and Logistics

CPSL 2020

17 March 2020 - 20 March 2020
Stellenbosch, South Africa

Sponsored by
Institute of Production Systems and Logistics (IFA)
German Academic Association for Production Technology (WGP)
Leibniz Information Centre for Science and Technology and University Library (TIB)
Stellenbosch Institute for Advanced Study (STIAS)
All papers of the proceedings are made available to the public under the following Creative Commons license for an unlimited period of time:

**Creative Commons Attribution 3.0 Germany**

Link to summary and binding version of the license text:
[http://creativecommons.org/licenses/by/3.0/de/](http://creativecommons.org/licenses/by/3.0/de/)

If contents of the proceedings are used for further work, these are to be clearly referenced in accordance with good scientific practice. The recommended citation is:

Foreword

It is our great pleasure to publish the proceedings of the first Conference on Production Systems and Logistics (CPSL), which took place at the Stellenbosch Institute for Advanced Study (STIAS) in South Africa in March 2020.

The idea of the CPSL was born when “Fake Science” was hitting the media two years ago. At that time we sat down and investigated conferences all over the world and discovered questionable practices and standards, which ultimately led to the creation of a platform that unites scientists in a newly designed conference format that has a solid and transparent review process and is also attainable. With all of this in mind we wanted to create a conference that evolves with the way of science. Open access publishing, magnifying influence on the published papers by our participants and peers as well as working closely with renowned research institutions to establish a general standard are the key elements for the creation of the CPSL.

For the CPSL 2020 we received great interest from participants across the globe, especially from institutions that engage in the German Academic Association for Production Technology (WGP). Due to the current crisis of COVID-19, our conference and especially the travels of our participants were sadly impacted. Only a quarter of the initially registered participants could attend in person. With these new circumstances, we reached the agreement to publish these proceedings with all the accepted and carefully reviewed papers despite the crisis of COVID-19. Nevertheless, the CPSL 2020 turned out as a very inspiring and instructive conference. Due to the reduced number of participants, the conference sessions developed as intensive discussions between the authors, in which both praising and critical feedback helped to jointly develop the presented ideas and gain new perspectives on the research topics.

We would like to express deepest appreciation and gratitude to our sponsors, the research associates of the Institute of Production Systems and Logistics (IFA) of the Leibniz University Hannover, the Leibniz Information Centre for Science and Technology University Library (TIB) and the German Academic Association for Production Technology (WGP). We also would like to thank STIAS for providing a wonderful venue and all the support throughout at the conference. A special thanks goes to all the reviewers for engaging in this demanding but also encouraging review process. Last but not least, we would like to thank all of our participants and fellow researchers that have been willing to share their research knowledge and experience to all of us.

We look forward to meeting you again at the next CPSL in Vancouver 2021.

Prof. Dr.-Ing. habil. Peter Nyhuis
Scientific Partner

Dr.-Ing. David Herberger
Conference Chair

M. Sc. Marco Hübner
Conference Chair
Review Process

The Conference on Production Systems and Logistics (CPSL) is an international forum for the scientific exchange on current findings in the field of production engineering.

For the submission of a paper, an abstract of no more than 300 words had to be uploaded to the CPSL platform within the deadline. Abstracts on the following main topics were considered:

- Industry 4.0
- Modeling and Simulation
- Supply Chain Networks
- Lean Manufacturing
- Factory Planning
- Production Planning and Control
- Production Management
- Logistics
- Production Systems
- Knowledge Management
- Sustainability
- Ergonomics

The submitted abstracts were evaluated in an internal review process, whereby successful submissions were notified and invited to upload a Full Paper. Full papers had to be no longer than 10 pages (justified exceptions were partially accepted) and had to adhere to a specific template and format provided on the CPSL website.

Subsequently, the submitted full papers were reviewed in a two-stage peer review process by experienced scientists from renowned research institutions as well as authors of other submitted papers. This ensured a constant and high quality as well as the influence of all participants on the papers and reviews. Consequently, each paper submitted was sent to at least two reviewers, with a third reviewer being requested in case of non-consensus between the first two reviewers.

The reviewers were asked to review the submitted papers on the basis of a provided evaluation template and give detailed comments and suggestions for improvement. The following key questions, among others, were considered:

- Does the title reflect the contents of the paper?
- Do you consider the work a proof of a thorough research and knowledge of the latest literature in the field of research?
- Are the conclusions clear and valid?

After completion of the reviews, all authors were given sufficient time to adapt their papers and submit a revised paper.
Acknowledgements

Our sincere thanks go to our outstanding supporters who made this great and interesting conference possible.
# Table of contents

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regeneration Supply Chain Model and Pool Stock Dimensioning</td>
<td>1</td>
</tr>
<tr>
<td>Tammo Heuer, Benedict Wildmann, Torben Lucht, Peter Nyhuis</td>
<td></td>
</tr>
<tr>
<td>Disassembly sequencing in the regeneration of complex capital goods</td>
<td>12</td>
</tr>
<tr>
<td>Torben Lucht, Tammo Heuer, Peter Nyhuis</td>
<td></td>
</tr>
<tr>
<td>Contribution to the industry-specific identification and selection of a business model in machinery and equipment industry</td>
<td>21</td>
</tr>
<tr>
<td>David Maximilian Doerr, Thomas Bauernhansl</td>
<td></td>
</tr>
<tr>
<td>Development of the Supply Chain Management 2040 – Opportunities and Challenges</td>
<td>29</td>
</tr>
<tr>
<td>Martina Schiffer, David Maximillian Dörr</td>
<td></td>
</tr>
<tr>
<td>Innovative logistics concepts for a versatile and flexible manufacturing of lot size one</td>
<td>37</td>
</tr>
<tr>
<td>Ali Bozkurt, Manuel Hagg, Robert Schulz</td>
<td></td>
</tr>
<tr>
<td>Dynamic criticality assessment as a supporting tool for knowledge retention to increase the efficiency and effectiveness of maintenance</td>
<td>48</td>
</tr>
<tr>
<td>Theresa Passath, Cornelia Huber, Hubert Biedermann</td>
<td></td>
</tr>
<tr>
<td>Introduction of Traceability into the Continuous Improvement Process of SMEs</td>
<td>58</td>
</tr>
<tr>
<td>Nicolas Wittine, Sigrid Wenzel, Christian Kern, Robert Refflinghaus, Tim Trostmann</td>
<td></td>
</tr>
<tr>
<td>Data-based identification of knowledge transfer needs in global production networks</td>
<td>69</td>
</tr>
<tr>
<td>Günther Schuh, Jan-Philipp Prote, Andreas Gützlaff, Katharina Thomas, Moritz Johannes Volk</td>
<td></td>
</tr>
<tr>
<td>Evolution of a Lean Smart Maintenance Maturity Model towards the new Age of Industry 4.0</td>
<td>78</td>
</tr>
<tr>
<td>Oliver Schmiedbauer, Hans Thomas Maier, Hubert Biedermann</td>
<td></td>
</tr>
<tr>
<td>Development of a Concept for Real-Time Control of Manual Assembly Systems</td>
<td>92</td>
</tr>
<tr>
<td>Martin Sudhoff, Matthias Linsinger, Daniel Schulte, Bernd Kuhlenkötter</td>
<td></td>
</tr>
<tr>
<td>Concept for the cost prognosis in the industrialization of highly iteratively developed physical products</td>
<td>99</td>
</tr>
<tr>
<td>Günther Schuh, Jan-Philipp Prote, Andreas Gützlaff, Sven Cremer, Shari Wlecke</td>
<td></td>
</tr>
<tr>
<td>Methodology And Use Of Variant Fields In Factory Planning</td>
<td>110</td>
</tr>
<tr>
<td>Matthias Dannapfel, Sebastian Patrick Vierschilling, Florian Becker</td>
<td></td>
</tr>
</tbody>
</table>
# Table of contents

Value Stream Mapping and Process Mining: A Lean Method Supported by Data Analytics  
*Katharina Mertens, Robert Bernerstätter, Hubert Biedermann*  
119

A Framework for Data Integration and Analysis in Radial-Axial Ring Rolling  
*Simon Fahle, Bernd Kuhlenkötter*  
127

Optimization approach for the combined planning and control of an agile assembly system for electric vehicles  
*Peter Burggräf, Matthias Dannapfel, Tobias Adlon, Aaron Riegauf, Esben Schukat, Felix Schuster*  
137

Generative Design In Factory Layout Planning: An Application Of Evolutionary Computing Within The Creation Of Production Logistic Concepts  
*Matthias Dannapfel, Sebastian Patrick ViERSchilling, Stephan Losse, Oliver Matzke*  
147

Self-organization and autonomous control of intralogistics systems in line with versatile production at Werk150  
*Jan Schuhmacher, Vera Hummel*  
157

Socio-technical requirements for production planning and control systems  
*Günther Schuh, Andreas Gützlaff, Frederick Sauermann, Matthias Schmidhuber*  
167

Analytical model for determining the manual consolidation time for large equipment manufacturers  
*Jakob Schygga, Philipp Braun, Svenja Hilbrich, Johannes Hinckeldeyn, Jochen Kreutzfeldt*  
177

Implementation of Machine Learning to Improve the Decision- Making Process of End-of-Usage Products in the Circular Economy  
*Michael Diem, Anja Braun, Louis Louw*  
188

Deriving of Sequencing Strategies for Multi-Stage Productions Supported by Logistic Models and Software Tools  
*Alexander Mütze, Peter Nyhuis*  
198

Current Knowledge Management in Manual Assembly – Further Development by the Analytical Hierarchy Process, Incentive and Cognitive Assistance Systems  
*Robin Sochor, Teresa Sofie Schick, Lukas Merkel, Stefan Braunreuther, Gunther Reinhart*  
209
# Table of contents

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>An Upgradable Cyber-Physical System Enabling Smart Maintenance of</td>
<td>220</td>
</tr>
<tr>
<td>UV Lamps in Industrial Applications</td>
<td></td>
</tr>
<tr>
<td><strong>Max Böhm, Dominik Lucke, Konrad von Leipzig</strong></td>
<td></td>
</tr>
<tr>
<td>Part Based Mold Quotation With Methods Of Machine Learning</td>
<td>229</td>
</tr>
<tr>
<td><strong>Sebastian Brede, Benjamin Küster, Malte Stonis, Mike Mücke, Ludger Overmeyer</strong></td>
<td></td>
</tr>
<tr>
<td>Data-based identification of throughput time potentials in production departments</td>
<td>239</td>
</tr>
<tr>
<td><strong>Lasse Härtel, Peter Nyhus</strong></td>
<td></td>
</tr>
<tr>
<td>Sizing electric storage system for atypical grid usage of industrial consumers</td>
<td>249</td>
</tr>
<tr>
<td><strong>Fabian Zimmermann, Annika Wurster, Alexander Sauer</strong></td>
<td></td>
</tr>
<tr>
<td>Activity-based shop floor management – A concept to enhance flexibility</td>
<td>259</td>
</tr>
<tr>
<td><strong>Barbara Tropschuh, Svenja Korder, Fabian Dillinger, Harald Bauer, Gunther Reinhart</strong></td>
<td></td>
</tr>
<tr>
<td>Creating a Worker-Individual Physical Ability Profile Using a Low-Cost Depth Camera</td>
<td>269</td>
</tr>
<tr>
<td><strong>Justus Brosche, Hannes Wackerle, Hermann Lödding</strong></td>
<td></td>
</tr>
<tr>
<td>A Vision of Digitalization in Supply Chain Management and Logistics</td>
<td>277</td>
</tr>
<tr>
<td><strong>Michael Henke, Christoph Besenfelder, Sandra Kaczmarek, Michael Fiolka</strong></td>
<td></td>
</tr>
<tr>
<td>Incentive System Framework for Information Sharing in Value-Adding Networks</td>
<td>287</td>
</tr>
<tr>
<td><strong>Alexander Zipfel, Konstantin Wink, Gunther Reinhart</strong></td>
<td></td>
</tr>
<tr>
<td>Efficient Use of Human-robot Collaboration in Packaging through Systematic Task Assignment</td>
<td>297</td>
</tr>
<tr>
<td><strong>Sebastian Blankemeyer, Rolf Wiemann, Uwe-Karsten Vett, Tobias Recker, Dennis Pischke, Annika Raatz</strong></td>
<td></td>
</tr>
<tr>
<td>Virtual Asset Representation for enabling Adaptive Assembly at the Example of Electric Vehicle Production</td>
<td>307</td>
</tr>
<tr>
<td><strong>Peter Burggräf, Matthias Dannapfel, Tobias Adlon, Viviane Hahn, Aaron Riegauf, Patricia Casla, Angelo Marguglio, Izaskun Fernandez</strong></td>
<td></td>
</tr>
<tr>
<td>Blockchain-based Services in the Machinery and Plant Engineering Industry</td>
<td>315</td>
</tr>
<tr>
<td><strong>Themo Voswinckel, Günther Schuh, Volker Stich, Jan Reschke</strong></td>
<td></td>
</tr>
<tr>
<td>Method to Automatically Create an Initial Layout for a Production System</td>
<td>324</td>
</tr>
<tr>
<td><strong>Matthias Bartelt, Fabian Katzwinkel, Bernd Kuhlenkötter</strong></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>A methodology for simulation production systems considering the degree of autonomy</td>
<td>332</td>
</tr>
<tr>
<td>Patrick Voit, Marie Beller, Gunther Reinhart</td>
<td></td>
</tr>
<tr>
<td>Sheet-Metal Production Scheduling Using AlphaGo Zero</td>
<td>342</td>
</tr>
<tr>
<td>Alexandru Rinciog, Carina Mieth, Paul Maria Scheikl, Anne Meyer</td>
<td></td>
</tr>
<tr>
<td>Conceptual Approach of Robustness in Logistical Control</td>
<td>353</td>
</tr>
<tr>
<td>Alexander Bader, Ann-Kathrin Onken, Kirsten Tracht</td>
<td></td>
</tr>
<tr>
<td>Managing Disruptions in Production with Machine Learning</td>
<td>360</td>
</tr>
<tr>
<td>Vladimir Samsonov, Chrismarie Enslin, Ben Luetkehoff, Felix Steinlein, Daniel Lütticke, Volker Stich</td>
<td></td>
</tr>
<tr>
<td>Do-it-Together Concept for Production Ecosystems</td>
<td>369</td>
</tr>
<tr>
<td>Andreas Külschbach, Tobias Leiting, Volker Stich</td>
<td></td>
</tr>
<tr>
<td>Evaluation of (De-)Centralized IT technologies in the fields of Cyber-Physical Production Systems</td>
<td>377</td>
</tr>
<tr>
<td>Nick Große, David Leisen, Tan Gürpinar, Robert Schulze Forsthövel, Michael Henke, Michael ten Hompel</td>
<td></td>
</tr>
<tr>
<td>A Framework for Online Detection and Reaction to Disturbances on the Shop Floor Using Process Mining and Machine Learning</td>
<td>387</td>
</tr>
<tr>
<td>Markus Fischer, Mahsa Pourbafrani, Marco Kemmerling, Volker Stich</td>
<td></td>
</tr>
<tr>
<td>Efficient Task Realizations in Networked Production Infrastructures</td>
<td>397</td>
</tr>
<tr>
<td>Marcus Grum, Benedict Bender, Norbert Gronau, Attahiru S. Alfa</td>
<td></td>
</tr>
<tr>
<td>Case study on technological applications for production planning and control in the context of industry 4.0</td>
<td>408</td>
</tr>
<tr>
<td>Günther Schuh, Patrick Scholz</td>
<td></td>
</tr>
<tr>
<td>Development of a Model for describing, evaluating and designing</td>
<td>418</td>
</tr>
<tr>
<td>Communication Concepts in Factories in the Context of Industry 4.0</td>
<td></td>
</tr>
<tr>
<td>Yeong-Bae Park, Lennart Hingst, Peter Nyhuis, Michael Kelber, Jörg Rainer Noennig</td>
<td></td>
</tr>
<tr>
<td>Bio-inspired Factories of the Future</td>
<td>426</td>
</tr>
<tr>
<td>Arvid Hellmich, Brandon Sai, Marian Süße, Martin Schreiber, Torben Wiese, Steffen Ihlenfeldt, Thomas Bauernhansl, Matthias Putz, Gunther Reinhart</td>
<td></td>
</tr>
</tbody>
</table>
Table of contents

Development of a decision support app for short term production control to improve the adherence to delivery dates
Felix Steinlein, Yuan Liu, Volker Stich

Tactical Production Planning for Customer Individual Products in Changeable Production Networks
Johannes be Isa, Thomas Zimmermann, Philipp Scherwitz, Gunther Reinhart

Rapid and Long-term Measures for Prevention and Mitigation of Communication Barriers in Production Networks
Michael Kelber, Jörg Rainer Noennig, Peter Nyhuis

Assessing product portfolios from a production logistics perspective
Tim Kämpfer, Torben Lucht, Philipp Schäfers, Peter Nyhuis
1st Conference on Production Systems and Logistics

Regeneration Supply Chain Model and Pool Stock Dimensioning

Tammo Heuer¹, Benedict Wildmann¹, Torben Lucht¹, Peter Nyhuis¹  
¹Institute of Production Systems and Logistics (IFA), Leibniz University Hannover, Germany

Abstract

When regenerating complex capital goods, on-time delivery as the most valuable customers’ requirement is crucial. Schedule reliability and throughput times are being trimmed to meet their targets as precisely as possible while keeping the logistic costs in-check. After disassembly, a significant number of components face relatively long repair times and need to be reassembled on a timed schedule. The configuration of internal supply chains offers the potential to improve schedule reliability. Pooling strategies are developed and discussed to achieve higher flexibility and positive effects on logistical performance. Pools help with reducing throughput times and short-term capacity allocation to satisfy an optimal on-time schedule. Serviceable (SA) components that have already been repaired, are stored in SA-pools and, if necessary, are allocated to the reassembly. The focus of this paper is the non-serviceable pool (nSA-pool), which provides repairable components to the repair stage. The nSA-pool helps to streamline the workflow before components reach several repair shops and has a direct impact on the repair process. Therefore, a model that allows the comprehension of interactions within the internal supply chain was developed and expedient pooling strategies were derived. Furthermore, the related pool stock dimensioning of preceding pools (nSA) before the repair stage and following pools (SA) are put into perspective.

Keywords

Complex capital good; regeneration supply chain model; pool stock dimensioning

1. Logistical issues in regeneration supply chains

Complex capital goods are characterized by a high number of components and high investment costs [1]. Examples are aircraft turbines, machine tools, rail vehicles, transformers, wind turbines and aircraft. The regeneration of complex capital goods has gained importance in many industries due to their high value and associated high downtime costs. Thus, after a usage cycle, the regeneration process has the objective to regenerate the used equipment. Regeneration includes the comprehensive preparation of the complex capital goods to its original delivery condition and may go further than maintenance improving its functionality [2]. Due to the high investment value of capital goods, long throughput times and delivery delays are not acceptable [3]. Quick and predictable re-use of complex capital goods is necessary. In addition to high delivery reliability and short throughput times, logistics costs must be kept low [4]. Therefore, stable regeneration processes are becoming a factor of success in the competition [5,6].

The time-consuming repair is confronted with uncertain capacity requirements, as damage to components of the complex capital goods can rarely be predicted [7]. Internal pooling stations are added to the regeneration supply chain to ensure flexibility and high delivery performance. They are a proven instrument for making short-term adjustments in the event of under- or over-utilization of repair capacities and help improve the on-time supply of components. [8,9] This paper discusses two variants of pooling solutions. Both try to
improve the logistical performance of the preceding and following stages in the regeneration supply chain. The serviceable (SA) pool stores repaired components only. It is situated post-repair in the regeneration process. Contrary to that, the non-serviceable (nSA) pool temporarily stores damaged components coming out of disassembly. In order to make logistic problems within the internal regeneration supply chain transparent, connections between stages of the regeneration path need to be made visible. This paper also integrates the proposed pooling capabilities to develop an overview. Due to volatile order volumes and problematic utilization planning, pooling strategies require a sophisticated pool dimensioning approach to ensure high logistical efficiency, which is characterized by a high ratio of logistical performance and logistics costs [8,10].

2. State of the art of regeneration supply chains

Configuration of the regeneration supply chain plays a vital role within the regeneration process of complex capital goods [11]. This chapter allows an overview of the regeneration process and clarifies how the internal supply chain works with different pooling configurations. The regeneration process is visualized in Figure 1. The regeneration process of the service provider is marked in anthracite and can be separated into five main stages: disassembly, inspection, repair, reassembly and quality control [8]. When specific indicators are met, the regeneration process is triggered. For instance, regeneration triggers can be sudden damage or legal requirements, such as regeneration after a certain period of use [12]. Even before shipment and disassembly, the regeneration service provider estimates a probable throughput time. The following disassembly is the first physical contact of the complex capital good with the regeneration service provider. The complex capital good is getting disassembled into its subassemblies and components. For a more precise analysis, the following inspection is essential. The inspections findings enable precise planning of the work scope of all following work stations [7,13].

![Figure 1: Regeneration supply chain](image_url)

Optionally, the nSA-pool could be placed in between inspection and repair. If components do not need to be repaired and provided to reassembly immediately, they get stored in the nSA-pool. Otherwise, the components are forwarded to a repair stage. Here a high share of the throughput time is being used. The variability of the different components leads to a substantive difference in resulting repair times [1]. In order to meet the tight schedule for reassembly, all components must arrive on time [13]. Differences in arrival time are mainly influenced and rooted in the repair process [14]. After repair, single components are converging in a buffer uncoupling previous processes and reassembly. Between repair and reassembly, the SA-pool is located. Repaired components get stored in the SA-pool until they are needed for reassembly. Thus, the SA-pool allows on-time delivery by providing the required components to reassembly. Furthermore, new components can be procured externally to converge in reassembly. In other instances,
Components must be replaced by new components because they are not reusable or because the particular component has been irreparably destroyed during its usage cycle. However, ordering new components is often associated with costs and other difficulties and does not fit the maintenance providers' intent [15]. Components for both the nSA- and SA-pool can be procured externally. After the successful convergence of component streams and accomplished reassembly, the final test and quality control make the complex capital good ready for shipment to the customer. Subsequently, another usage cycle starts and after a defined amount of usage time, a new regeneration process should be started. [16] Following, the two pools and their interactions within the internal supply chain will be described with a more detailed look.

2.1 No integrated pooling

A regeneration supply chain without SA- and nSA-pooling has a simple structure. It is advantageous that no nSA- and SA-stocks need to be maintained. As a result, less storage and logistics space is required, and management of the pools is not needed. Restructuring processes to adapt to a pooling solution in the regeneration business can be an extensive project for a regeneration company. With no pooling, the procurement of new components is required for every replacement of scrap components [17]. However, the costs of the regeneration process are increased when purchasing new components. Therefore, it has to be determined whether a new component procurement is carried out only for scrap components. If there are any problems with the repair time of components, deadline shifts, and maybe penalty costs have to be accepted. Hence, without pooling, flexibility to react to problems within the repair is limited [7]. With no integrated pooling configurations, there tends to be a high amount of work in progress (WIP) taking up space in workstations of the repair and tying up capital. Due to e.g. problems in the repair, components can be overdue for reassembly. This results in a disrupted WIP in the reassembly, which contains all components that have to wait for delayed components to get reassembled to the complex capital good. Especially the SA-pool can reduce this issue. [8] But also, the use of an nSA-pool enables a purposeful work in the repair and, further, allows a suitable delivery of components and thus less disrupted WIP in reassembly. Additionally, it is difficult to compensate for utilization fluctuations without pooling. Hence, the repair must be accurately timed and scheduled in order to uphold punctuality of the entire regeneration supply chain. [14]

2.2 SA-pool

Implementation of an SA-pool has been discussed by other authors before. Often they are focusing on regeneration processes of aircraft components [18]. Another focus area that differs from considerations in this paper is the discussion about external pooling solutions that do not emphasize the direct repair process itself [19]. Instead, external pooling comprises the storage of spare components in a central warehouse to which several companies have access to and which considers the regeneration supply chain at a network level. FRITZSCHE investigates material availability of aircraft components with dynamic failure rate for external pooling, which makes pool dimensioning more precise and reduces storage costs for the network [19]. The original model used for this purpose is based on the Metric for Recoverable Item-Control (METRIC) model developed by SHERBROOKE [20]. The METRIC model and its numerous extensions concentrate on external pooling sectors and cross-company multi-level regeneration supply chains, where the single regeneration service provider cannot adapt the external pooling to its specific requirements. Thus, this paper focuses on a universal implementation of the SA-pool to the internal regeneration supply chain.

Based on the positioning before repair, the SA-pool primarily serves the timely provision of components in order to complete customer-specific orders on time. With the extension of the system by an SA-pool, increased on-time delivery to reassembly and thus an improved missing parts situation in reassembly can be expected [9]. This improvement reduces the costs for previous and subsequent stations of the SA-pool by offsetting costs for maintaining an SA-pool. The improved missing parts situation is caused by the fact that
the SA-pool synchronizes the completion of individual orders at the convergence point and consequently reduces disrupted WIP. Additionally, the SA-pool can generate a job-specific safety time by bringing forward requirement dates. [8,10] The repair achieves an improvement of flexibility. Preparatory repairs and existing material capacity in the SA-pool prior to reassembly means that the repair does not have to focus on providing components for reassembly only. Instead, components that are not currently needed can be repaired, pooled in the SA-pool, and added to the process later to ensure long-term availability of predicted component demands and leveling the repair workload.

Various strategic objectives are required for spare components pooling. Depending on the pursued strategy, different targets for regeneration logistics are primarily influenced. KUPRAT differentiates five different pooling strategies: [8]

1. The timely improvement of the already-existing situation in reassembly that positively influences punctuality and a positive due date deviation.
2. The reduction of regeneration costs and manufacturing costs.
3. Reducing or leveling the repair throughput time focusing on the throughput time.
5. Residue-oriented deployment of pool components which try to improve punctuality.

The focal point of the SA-pool is to support punctuality and on-time completion in the reassembly process. The combination of strategies 1 and 3 is the most accommodating way for an SA-pooling strategy. This can be described as a "reassembly-oriented pooling strategy" [8]. A major problem of supply chain configurations without pooling solutions described above is that in situations where a component is missing for reassembly, only the procurement of new components is an option [21]. Here the SA-pool offers a simpler and cheaper possibility. The decision of whether to take a new component or an SA component requires definitions and individual decisions. Even with SA-pooling, workload fluctuations in the repair are still to be expected [8]. For the SA-pool, a stock of components has to be defined. This stock takes up area and logistic capacities as a separate entity in the overall logistics planning. Due to more complex material flow, logistics planning becomes more extensive, and the picking effort increases, but with an SA-pool, the logistical performance for the customer can be significantly increased, and on-time delivery is ensured.

2.3 nSA-pool

The focus of research so far has largely been on the SA-pool; here, the additional nSA-pool is discussed. With only an SA-pool, since the repair only has a certain capacity, components from the disassembly must be stored in a buffer of the repair stage [8]. In order to gain a central distribution point and further confront logistic issues in front of the repair, an nSA-pool is added to the model. Then, components requiring repair are managed centrally by the nSA-pool. Therefore, there is a lower WIP in the repair. [10] The nSA-pool aims to achieve a balanced capacity utilization in the repair. The nSA-pool can enhance the utilization in repair by releasing nSA-components for repair when the current workload in the repair is low. A well utilized repair also results in a balanced load in reassembly and, thus, in conjunction with the SA-pool, increases schedule reliability and on-time delivery. Considering the descriptions mentioned above of pooling strategies for the SA-pool, the nSA-pool’s focus is on the utilization in the repair stage. The most suitable strategy is strategy 4, as mentioned in chapter 2.3. As a result, the nSA-pool strives for a "utilization-oriented strategy", especially for the repair stage. Due to central importance for the regeneration processes and the great expenditure of time in the repair, streamlining the logistic processes in or before repair also has a significant effect on the other stages. Compared to the models described above, the combination of both pool stages has the highest logistical space requirement. Hence, logistics control is becoming more complex. However, capacity utilization and logistical performance can be improved considerably.
3. **Pool Interactions**

Hereafter, interactions of the pools and effects on logistic parameters are described and summarized in an interaction model depicted in Figure 2. Since the focus of this paper is on the management of single components, interactions and logistic parameters between disassembly as the divergence point of the complex capital good and reassembly as convergence point are considered.

From disassembly, components can either go into the nSA-pool or repair. The output of the nSA-pool also provides access to the repair. In repair, an inventory is built up that includes both repair stock and WIP, since not all components can be processed immediately due to limited repair capacity. However, the stock of components in repair also influences capacity utilization, since an appropriate stock means a good relation of high utilization with suitable throughput times for components waiting for their repair [10]. After the components have been processed in repair, they are forwarded as the output of repair. Subsequently, components are ready for the time-critical reassembly and can be assigned directly to the access of the reassembly. Reassembly depends on a large number of very different components that converge to form the complex capital good. The direct path to reassembly is mainly used when a component has been repaired.

![Interaction model of the regeneration supply chain](image-url)
order-specific. Another possibility is that the repaired component gets stored in the SA-pool since the component has been repaired order anonymous.

In addition to the output of the disassembly, access to the nSA-pool is also impacted by the external procurement of components. An external procurement of nSA-components is worthwhile for the company if its repair is well developed, enabling low repair costs or if components are required for replacement. Access of the nSA-pool increases the nSA-stock. The higher the nSA-stock, the more the correlating nSA-pool costs. Though, the nSA-stock enables an increased nSA-service level for repair. All components taken from the nSA-pool are summarized in the output of the nSA-pool. Components from the nSA-pool can be sold or disposed of if an internal repair would not be feasible. The nSA-pool output passes into the access of the repair.

External procurement obtains new components. The access to the new components (SAN) pool affects SAN-stock in the new components pool. A higher SAN-stock increases current SAN costs and ties up capital due to high prices for new components. However, the SAN-stock has a positive impact on the SAN service level and, therefore, a positive effect on the schedule reliability of the reassembly. New components can be used immediately, mostly without compromise, and with full certification as a replacement in reassembly. However, external procurement is associated with procurement lead times and effort, which delays access to the SAN-stock. The output of the SAN-pool always goes directly into the access of the reassembly.

The SA-pool is integrated into the components flow to enable on-time delivery for reassembly. The output of repair feeds the access of the SA-pool. The difference between input and output of the SA-pool results in the SA-stock. On the one hand, stock means stockholding costs; on the other hand, it supports the SA-pool’s service level to repair, decreasing disrupted WIP in reassembly. For this reason, the SA-pool is crucial to increase overall logistical performance in this model. Due to its excellent cost-efficiency, compared to the new components pool, the SA-pool is to be preferred as far as it is possible under consideration of legal and customer requirements. If the product range no longer requires them, components can also be sold or disposed of via the actual output of the SA-pool.

Moreover, the regulatory authority must plan and monitor all material flows. A central regeneration planning and control (RPS) is located above the nSA- and SA-pool, as well as the new components pool, and keeps track of repair to achieve optimized performance for reassembly. [22] The choice of the material disposition path and sequencing in the repair have to be coordinated centrally and - due to complexity and the high number of possible plans - simultaneously [13]. Deviations in the scheduled sequences delay components for reassembly and delivery delays may result in penalties. In contrast, the use of the pool system can help to maintain delivery date reliability when such issues occur. For this reason, the pools’ stocks need to be planned. The RPS is intended to ensure that defined stocks in the pools do not fall below a certain threshold but do not generate a surplus. Using PUSH and PULL mechanisms, RPS can control the pools [23]. If, for example, the SA-pool reaches an average stock level, but utilization in repair shops decreases, the PUSH mechanism takes effect, and the nSA-pool provides the repair shop with components for processing at short notice. However, if the SA-pool stock decreases, a PULL mechanism requests repaired components of the repair stage. Thus, central control variables are the stocks in all stages.

4. Pool Dimensioning

After discussing the interaction model, one can conclude that pool stocks play an essential role in the regeneration supply chain. In addition to individual pool stocks, the relation of pool stocks is elementary. Works by KUPRAT and GEORGIADIS are concerned with the selection of components to be pooled. There are also approaches for calculating the optimal stock level of an SA-pool. Nevertheless, the relation between SA- and nSA-pool stocks and utilization of the regeneration supply chain is little discussed. [24,8] For this purpose, a qualitative relation is derived from the considerations of chapter 3 and displayed in Figure 3. In
the diagram, axis labels $S_{SA}$ and $S_{nSA}$ represent the individual stock of the pools. The stock of the SA-pool is placed on the ordinate axis, that of the nSA-pool on the abscissa axis. Furthermore, a safety stock in the SA-pool ($SSL_{nSA}$) and a safety stock in the nSA-pool ($SSL_{nSA}$) are marked. Three possible exemplary function courses of stock relations are shown, depending on future utilization in the repair and divided from a low to a higher future utilization level.

A high utilization means an increased volume of work in the repair stage, but especially an increased number of complex capital goods to be reassembled. In order to ensure timely completion of required components in reassembly, despite the unpredictable and increased throughput times due to high load, an increased SA-pool stock must be available to avoid delayed components and, finally, delays of the complex capital goods. In this case, only the SA-pool can contribute to punctuality. Availability of nSA-components in the nSA-pool does not significantly influence the required stock of the SA-pool since nSA-components cannot be processed quickly enough in repair. Hence, the course function for high utilization is flatter than the course function for low utilization. For the same reason the maximum amount of nSA-stock is limited on the abscissa axis as well. In order to enable a higher SA-stock, an increased external procurement, or an increased output of the repair is necessary. With low actual utilization, so-called pre-production is possible, which increases the output of the nSA-pool and thus the input and output of the repair. During high utilization periods, the repair stage is supplied earlier by the nSA-pool with time-critical components, which are significant for the punctuality of the entire regeneration supply chain. The stock of the nSA-pool is decreasing as the provision of time-critical components for repair can be executed even before disassembly is finished.

In expectation of lower future utilization, to decrease stockholding cost, the SA-pool stock is reduced. Therefore, the course function of low utilization is lower. On the contrary, nSA-stock is built up because stockholding costs are lower, and due to lower future utilization, nSA-components can be repaired as required. Preferably, components stored in the nSA-pool can be used again and sent to repair when a high utilization is predicted. Due to the dynamic order situation, a safety stock in the SA-pool is essential. Safety stock in the SA-pool maintains flexibility and enables it to react as capacity utilization changes to ensure a high service level and, accordingly, high delivery reliability [15]. Besides, the nSA-pool maintains a safety stock, which makes it possible to increase utilization in the repair if the workload in the repair decreases.

A constant adjustment of the related stocks of pools depending on future utilization is necessary. Accordingly, a space between the minimum and maximum utilization is put up. With concrete calculations for the use case, a family of curves must represent a dynamic transition for potential cases between extremes. For the first consideration, the stock in Figure 3 is assumed for an exemplary component type.
types of components, different course functions have to be developed. In addition, an assessment of costs must be carried out through the course of the curves for individual cases.

One problem in the interaction of SA- and nSA-pooling solutions is especially the matching of stocks of both pools [24]. Due to the mentioned close, indirect coupling of nSA- and SA-pool, corresponding stock fluctuations are expected. This, in turn, can counteract the original intention if components are not selected correctly. Too many pooling components lead to high logistical costs. Lower stocks diminish balancing effects of utilization and punctuality on the regeneration supply chain through the pools. On top of that, it will not be economical to build up pool stock for all components. For each use case analyses must be carried out to define components that are worth to be pooled [8].

5. Conclusion

This paper describes the possibilities of a regeneration supply chain with different pooling solutions, displaying the regeneration supply chain and its internal interactions, as well as pointing out the important aspect of pool stock dimensioning. Here, the regeneration supply chain is considered on an internal company level. First, SA- and nSA-pool are described, and their position in a regeneration supply chain is explained. Different effects on the regeneration supply chain without pooling, with an SA-pool or with an SA- and nSA-pool are presented. Additionally, already known approaches of regeneration supply chain pooling are explained. A solution with SA- and nSA-pool betters the logistical performance of the entire regeneration supply chain. In particular, the nSA-pool helps with a “utilization-oriented strategy”, which balances the utilization in the repair shops, and the SA-pool helps with a “reassembly-oriented pooling strategy”, which improves the punctuality of the reassembly. Also, expensive new components procurement is substituted as the only option for time-critical tasks. However, more logistics capacity and space is used, and more logistic planning is required.

In the presented regeneration supply chain interaction model, interactions of individual stages are explained. For pools, the pool stock is put into perspective with correlating pool costs and pool service levels. In the repair stage, repair inventory influences the throughput time and utilization. Inventory of the reassembly raises disrupted WIP and influences the throughput time in reassembly. Crucial interactions for the pooling layout are the output of the nSA-pool, which supplies the access of repair in order to influence utilization of the repair, and the access of reassembly, which is supported by the output of repair, SA-pool and SAN-pool, in order to enhance punctuality of the reassembly. These results can be used to align the regeneration supply chain design with the company's targets, improving logistical efficiency.

Moreover, the central regeneration planning and control influences schedule reliability and total costs by using the pooling infrastructure. Interactions within the regeneration supply chain for complex capital goods play an essential role in the dimensioning of pooling solutions. The pool stocks are related to each other and future utilization. In case of high future utilization, a high stock of the SA-pool is needed to enable quick support for the reassembly. In contrast, with low future utilization, the nSA-pool stock is higher, and the more capital binding SA-stock lower. Thus, low utilization allows a lower total stock level. The derived relations help the regeneration service provider to manage pool stocks according to forecasted utilization. For this purpose, relations of stocks need to be put into perspective for individual components, and a comprehensive strategy for possible situations needs to be developed. Here, more research is needed, and the next step would be to formulate specifications for regeneration planning and control to make the needed adjustments of pool stocks visible on the component level. The pooling design will be significantly specific to the considered application. Therefore, scaling is complicated. Finally, pooling solutions will need to be implemented and tested with data of a regeneration service provider, and especially the effect of the nSA-pool will require evaluation.
Acknowledgments

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - SFB 871/3 - 119193472.

References


**Biography**

**Tammo Heuer, M.Sc.** (*1992*) studied industrial engineering at the Leibniz University Hannover and has been working as a research associate at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover in the field of production management since 2018.

**Benedict Wildmann, B. Sc.** (*1994*) studies mechatronics engineering and has been working as a research assistant at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover since 2018.

**Torben Lucht, M.Sc.** (*1991*) studied industrial engineering with the focus on production technology at RWTH Aachen University. Since 2018, he works as a research associate in the field of production management at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover.
Prof. Dr.-Ing. habil. Peter Nyhuis (*1957) studied mechanical engineering at Leibniz University Hannover and subsequently worked as a research assistant at the Institute of Production Systems and Logistics (IFA). After completing his doctorate in engineering, he received his habilitation before working as a manager in the field of supply chain management in the electronics and mechanical engineering industry. He is heading the IFA since 2003. In 2008 he became managing partner of the IPH - Institut für Integrierte Produktion Hannover gGmbH.
Disassembly sequencing in the regeneration of complex capital goods

Torben Lucht¹, Tammo Heuer¹, Peter Nyhuis¹

¹ Leibniz University Hannover, Institute of Production Systems and Logistics (IFA), Germany

Abstract

Despite constantly increasing condition monitoring parallel to operation, maintenance, repair & overhaul (MRO) processes in the regeneration of complex capital goods are still characterised by a high degree of uncertainty regarding the capacity and material demands to be expected from a regeneration order. In order to meet the committed delivery times and dates, the disassembly at the beginning of the regeneration supply chain is of particular importance for the performance of the entire downstream regeneration process. High potential for improving logistical performance lies in an intelligent and logistics-oriented sequencing strategy in disassembly. In addition to technical-physical boundary conditions, the interaction of the disassembly process with downstream process steps and additional other control measures must also be taken into account. The logistical description and evaluation of the sequencing-oriented measures for improving the logistical performance of disassembly in the context of the regeneration of complex capital goods makes their modelling a necessary prerequisite and basis. This paper presents a basic logistical design and modelling approach.

Keywords

Disassembly; MRO; regeneration; complex capital goods; condition monitoring

1. Introduction

Complex capital goods like aircraft engines or wind turbines still show a high monetary residual value at the end of an use phase. The regeneration pursues the aim to keep this residual monetary value through maintenance, repair & overhaul (MRO) measures by transferring the capital goods into a new use phase. [1–3] According to the generic process model of the regeneration the procedure includes the main process steps diagnosis, disassembly, cleaning and inspection, repair, reassembly and quality assurance. [4] The starting point of the regeneration is an input inspection to identify first information that form the planning base for the regeneration process. Despite increasing condition monitoring parallel to operation [5,6] the regeneration is still characterized by a high uncertainty of information regarding the actual damage patterns. [7]. To gain access to single components for a detailed inspection, most of the complex capital goods need to be disassembled. After detailed inspection, reliable information about the repair measures required for regeneration and the spare parts to be procured is available. Once all components are ready for installation, the capital goods are reassembled and quality assurance is completed. While the information blur is constantly decreasing during the regeneration process, there is still almost complete uncertainty about the actual damage pattern during disassembly. [8–10] The increasing availability of data as well as the advancement of algorithms and systems for their analysis and interpretation bear the potential to allow
precise statements about the component conditions and thus about the expected capacity loads and material demands due to regeneration even before induction and diagnosis (see Figure 1) [1,11].

Due to the uncertainty about the exact damage pattern, the depth of disassembly required for the regeneration is rarely known prior to the physical arrival at the regeneration service provider. Consequently, the resulting workload of the regeneration orders and the delivery dates of the disassembled components at the downstream process stages are difficult or impossible to predict and plan. [7] However, in order to comply with agreed delivery times and dates, it is essential to take these interactions into account. [12] In order to achieve a high level of logistics performance despite remaining uncertainty of information, a highly responsive as well as an intelligent supply chain configuration is required. Besides the pooling of particularly time-critical or repair intensive components, especially sequencing and prioritization strategies for disassembly offer the possibility to positively affect the logistics performance of the entire regeneration by reducing the total throughput time of the regeneration $TTP_{tot}$ [13].

2. Boundary conditions for disassembly sequence formation

The flexibility in designing and controlling the disassembly process is limited by a variety of boundary conditions. On the one hand, the freedom in controlling the disassembly process is essentially defined by the order-specific design and selection of the manufacturing principle to be applied. On the other hand, additional processual constraints result from the regeneration process.

2.1 Manufacturing principles for disassembly in the regeneration of complex capital goods

In the regeneration of complex capital goods such as aircraft engines, the construction site principle and the flow principle are primarily used. [14,15] In the case of disassembly according to the construction site principle, the capital goods are positioned at a fixed workplace. From here the disassembled components are transferred directly to the repair workshops or to downstream module and individual parts disassembly workshops. This stationary disassembly is opposed by the application of the flow principle. In this principle, the capital goods are transferred to the subsequent work station after a defined scope of work. The sequence of the work stations to be passed is fixed and either hard or elastically linked via buffering storage. [16] Here, the disassembled components are also transferred to the respective specialized repair workshops or module and component disassembly workshops. In analogy to the application of these manufacturing principles in the manufacturing industry, the application of the flow principle generally offers advantages in terms of efficiency for recurring work, whereas the application of the construction site principle allows more flexible reaction to varying work content. [15,17] Changes of sequences and prioritizations for each
manufacturing principle are limited by constructive boundary conditions. When applying the flow principle, the freedom for design is additionally limited to disassembly sequences to be carried out within a work system due to the fixed linkage of the work systems. [14] Thus, comprehensive and extensive manipulations of sequences as well as prioritizations of parts to be disassembled are usually only logistically advantageous if the construction site principle is applied for disassembly.

2.2 Procedural boundary conditions for the disassembly sequence formation

For disassembly sequencing, a process-related special characteristic results from the regeneration process described in the introduction. Assuming a planned disassembly sequence that is initially ordered according to the installation depth of the components (see Figure 2a), the components that are disassembled last correspond to the lowest level of disassembly of the capital good. Under the assumption that reassembly is always performed in a defined, most efficient order with ascending assembly level, the parts must be available in reverse chronological order to the initial disassembly sequence for reassembly after the repair. In the initial state, this results in a "Last Out - First In" sequence across all processes (see Figure 2b). [18]
This usually causes long slack times TSTi for the components disassembled first. On the other hand, in particular components disassembled later within the disassembly process must be available for reassembly as early as possible in order to enable further reassembly. Looking at the example of an aero engine these components also are subject to the highest mechanical and thermic loads during operation and thus often require more repair work then components on higher assembly-levels [19].

3. Logistics-oriented disassembly sequencing

When designing an intelligent approach for disassembly sequencing particularly the boundary conditions described before have to be taken into account. The goal is to reduce the overall throughput time of the regeneration order as well as to avoid backlogs due to long repair throughput times that exceed the available repair time between completion of disassembly and the start of reassembly. Therefore components with negative slack times between disassembly and assembly need to be prioritized in disassembly. Under the framework conditions of conventional production processes, Tryzna demonstrated that fast-track orders up to a share of 30% do not compete significantly with each other and thus experience extended lead times themselves [20]. While making these components available to downstream repair processes earlier, it always has to be ensured non prioritized components do not become time critical by these prioritization measures.

The first step is to define the product architecture of the complex capital good to be disassembled and regenerated. In particular, structural or otherwise induced predecessor-successors relations have to be considered. For this purpose, a priority graph can be set up on the basis of work plans for disassembly and reassembly or construction drawings. Figure 2 shows an exemplary priority graph for a regenerative good, consisting of seven components distributed over three assembly levels. The figure shows that, for example, components 3 and 4 can only be disassembled after the component at the higher structural level (component 1) has been disassembled. (see Figure 2a). Starting from a delivery date agreed with the customer, a sequential backward scheduling of the reassembly steps can now be carried out. The planned start and completion dates of the upstream process steps of repair and disassembly are then determined on the basis of the results of the initial inspection. This scheduling can be used to determine the time-based criticality of the components to be disassembled and to make prioritization decisions in favor of the most time-critical components. The interactions with other (sub-) disassembly processes of the same disassembly order or other disassembly orders processed in parallel as well as the competition for limited resources must be taken into account. In order to achieve the highest possible lead time potential in disassembly, prioritizations should be carried out until no further reduction of the regeneration lead time can be achieved by changes in sequence resulting from the prioritizations. This should also include component damage that is only detected during dismantling and therefore requires a greater dismantling depth. Due to this circumstance, the prioritization decisions must be reviewed each time the extent of disassembly is adjusted.

If the capacity load of the process elements along the entire regeneration supply chain is significantly changed as a result, a reassessment of the criticality and adjustment of the prioritization decision may become necessary as well. While the primary logistic target pursued through prioritization is a shortening of the regeneration throughput time, the schedule deviation induced by changes in the sequence of disassembly and thus in the availability for the downstream repair must always also be taken into account. If components are accelerated or decelerated during disassembly without coordination with the downstream process steps, there is a risk of a negative influence on the overall logistics performance. If all these framework conditions are taken into account, the logistics-oriented prioritization in disassembly offers the potential for a reduction of the total throughput time of the regeneration order (see Figure 2, bottom).
4. Modelling of logistics processes in disassembly

In order to be able to describe the effects of controlling measures on the logistics performance of the disassembly as well as corresponding interactions with other processes or orders, a proper description and modelling of the processes in the disassembly as well as their interfaces to upstream and downstream processes is required. While the downstream inspection and repair workshops can be modelled as conventional working systems, a logistical description of the disassembly is required. Dombrowski et al. have already developed a disassembly throughput diagram for this purpose. [21] However, this diagram only shows the incoming and outgoing orders in the disassembly process and the subsequent inspection. This allows the description and evaluation of the work in progress (WIP) and load levels in these two process steps. Interactions with the subsequent repair and reassembly processes are not shown. Especially the demand dates of the disassembled components in reassembly and repair as well as the resulting workload in repair are necessary for the logistic description and evaluation of the effects of prioritization measures. In general, different perspectives can be used to describe logistics processes. [16,22] For the description of disassembly processes three perspectives may be used (see Figure 3):

- The order/product perspective,
- the component perspective and
- the resource perspective.

The order perspective considers the individual disassembly order, which consists of various sub-disassembly operations. It focuses on adherence to delivery dates, lead times and individual process costs. The component perspective is used to analyze stocks of individual sub-assemblies, modules and individual parts that flow from various disassembly orders into downstream stocks. In contrast, the resource perspective describes the logistical system behavior of work systems (in this case: disassembly systems). Essentially, the workload and the worklist of the resource as well as the adherence to delivery dates and lead times of all orders in the system during an examination period are analyzed. [22]

![Figure 3: Different perspectives on the disassembly process](image)

In general, disassembly represents a divergence point in regeneration supply chains, regardless of the perspective used. These divergence points have not been made accessible to any logistical description, yet. [23,24]. Depending on the perspective taken, the divergence refers to different subjects. From an order perspective, the divergence lies in the product structure, since the capital good is disassembled into several subordinate assemblies and individual parts. From a resource perspective, divergence in the material flow...
following the disassembly is considered. The component perspective, due to its focus, is not suitable for describing divergence. The order perspective is therefore used below to describe the order-specific design options in the disassembly process. As logistical description and coupling parameter, the lateness can be used [25–27]. In disassembly, the lateness describes the difference between the actual disassembly completion time \( EDO_{act} \) and the originally planned disassembly completion time \( EDO_{plan} \) of the overall disassembly order [28,29]. If an order is completed too early (\( EDO_{act} < EDO_{plan} \)), this results in a negative lateness value. Vice versa a positive lateness occurs, if a disassembly order is completed later than the planned disassembly completion date (\( EDO_{act} > EDO_{plan} \)). This does not exclusively need to be related to process errors, but may also be the result of backward scheduling from the agreed regeneration completion date \( EAO_{tot} \). Positive lateness can also occur if the disassembly completion date determined by backward scheduling is impossible to realize or is already in the past.

The lateness in the completion of disassembly \( L_{D,out,i} \) of a regeneration order basically consist of the lateness in the input of the disassembly system \( L_{D,in,i} \) as well as the relative lateness caused by the disassembly process \( L_{D,rel,i} \). From an order perspective, the lateness of the disassembly completion of the overall order can also be described as the difference between the actual disassembly throughput time \( TTP_{D,act,i} \) and the planned disassembly throughput time \( TTP_{D,plan,i} \), based on the input lateness \( L_{in,i} \). [30]. In this context, \( i \in I \) represent the set of all regeneration orders that are processed on a disassembly work system while \( j \in J_i \) do represent the set of sub-disassembly orders within a regeneration order \( i \). Consequently, the lateness of an individual disassembly order can be described as:

\[
L_{out} = EDO_{act} - EDO_{plan}
\]  

The lateness in the completion of disassembly \( L_{D,out,i} \) of a regeneration order basically consist of the lateness in the input of the disassembly system \( L_{D,in,i} \) as well as the relative lateness caused by the disassembly process \( L_{D,rel,i} \). From an order perspective, the lateness of the disassembly completion of the overall order can also be described as the difference between the actual disassembly throughput time \( TTP_{D,act,i} \) and the planned disassembly throughput time \( TTP_{D,plan,i} \), based on the input lateness \( L_{in,i} \). [30]. In this context, \( i \in I \) represent the set of all regeneration orders that are processed on a disassembly work system while \( j \in J_i \) do represent the set of sub-disassembly orders within a regeneration order \( i \). Consequently, the lateness of an individual disassembly order can be described as:

\[
L_{D,out,i} = L_{D,in,i} + L_{D,rel,i} = L_{D,in,i} + (TTP_{D,act,i} - TTP_{D,plan,i}) \quad \forall \ i \in I
\]  

with:

\[
TTP_{D,act,i} = \max(EDO_{act,i,j}) - \min(SDO_{act,i,j}) \quad \forall \ j \in J_i
\]  

Lödding et al. also mathematically described the lateness in the output of a work system as the sum of the backlog-related and sequence-related lateness [31]. This mathematical description of backlog-related and sequence-related components is not directly adaptable to the disassembly process. On the one hand, the statistical independence of the disassembly orders from each other as well as from sub-disassembly orders is not given, because of structural interdependencies between the components. On the other hand, the prioritization of individual components can influence the workload related to the (sub-) disassembly orders, whereby prioritization decisions have a mutual influence on each other as well as on the applied sequencing strategy. Both represent an essentially requirement for the separated, mathematical description of both components [32]. However, the relative lateness can also be attributed to residue-related and sequence-related influencing factors for disassembly. As a result of deviations between the work content assumed for the planning of a disassembly order and the actual workload, there is a backlog. For instance, this can result from component damage that is detected or even caused during the disassembly process and requires a subsequent extension of the disassembly level. A sequence-related lateness results from the prioritization of an entire order \( i \in I \) in relation to other regeneration orders as well as the prioritization of individual sub-disassembly orders \( j \in J_i \) within a regeneration order. The prioritization of entire disassembly orders can be considered equivalent to sequencing in conventional production processes, since the orders can be assumed to be stochastically independent of each other. Deviations from the generic disassembly sequence by prioritizing sub-disassembly orders within a regeneration order, on the other hand, influence both the lateness of the sub-disassembly orders themselves as well as the actual lead time and thus the output lateness of the overall disassembly and regeneration process. In addition to the direct prioritization of individual sub-disassembly orders, an additional indirect prioritization can result from constructive or other dependencies or links between components under consideration and directly prioritized components. Their effects on the
component-specific and regeneration order-specific lateness and lead time must be taken into account when making prioritization decisions as well as the effects on the logistics performance of the entire regeneration system.

5. Conclusion

The increasing availability and quality of status data from the operating phase even before or at the beginning of the regeneration process is opening up more and more possibilities for optimizing the regeneration process. A damage pattern based prioritization of disassembly operations offers the possibility to make heavily damaged components available to the downstream repair processes earlier. This bears potential for shortening throughput times by providing repaired components for reassembly more punctually as well as by shortening the order-specific, critical path. When applying these measures, numerous interactions and influencing factors must be taken into account which cannot yet be described with existing logistic models. Thus, modelling the effects of these measures on the adherence to delivery dates and delivery times of the individual orders as well as on the logistical performance of the regeneration system is an indispensable requirement for a successful application of the measures.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - SFB 871/3 - 119193472.

References


Biography

Torben Lucht, M.Sc. (*1991) studied industrial engineering with the focus on production technology at RWTH Aachen University. Since 2018, he works as a research associate in the field of production management at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover.

Tammo Heuer, M.Sc. (*1992) studied industrial engineering at the Leibniz University Hannover and has been working as a research associate at the Institute of Production Facilities and Logistics (IFA) at the Leibniz University Hannover in the field of production management since 2018.

Prof. Dr.-Ing. habil. Peter Nyhuis (*1957) studied mechanical engineering at Leibniz University Hannover and subsequently worked as a research assistant at IFA. After completing his doctorate in engineering, he received his habilitation before working as a manager in the field of supply chain management in the electronics and mechanical engineering industry. He is heading the IFA since 2003. In 2008 he became managing partner of the IPH – Institute for Integrated Production Hannover gGmbH.
Abstract

Industry 4.0 introduces a paradigm shift that will lead to changes of business models in many e. Today, industrial companies are gradually transforming their traditional and transaction-based business models into new business models made possible by cyber-physical systems. New business models such as as-a-service or platform-based business models emerge. This change brings enormous opportunities, but also many risks for the manufacturing industry. Many companies are faced with the problem of choosing from the multitude of new business models. The business model development to be found in literature primarily follows the needs of the customer. Machinery and equipment industry is a particularly interesting sector since more than 50% of the customers in machinery and equipment industry come from the same sector.

This paper develops a process model for the industry-specific selection of business models. The model includes the following questions: Which business models can be selected in the field of machinery and equipment industry? Which possible goals can be pursued with the respective business models? Which criteria are useful for deciding on the respective business model?

Keywords

Industry 4.0; business model; machinery and equipment industry

1. Introduction

New business models will be increasingly important for manufacturing companies in the future. Nevertheless, traditional companies, for instance in machinery and equipment industry, are reluctant to question and transform their own classic business models. However, due to the technological change in the course of Industry 4.0, it is necessary to question one's own business model. Shorter innovation and product cycles open up traditional business fields to players from outside. The systematic consideration of business model innovations has shown that there are specific patterns [1]. These business model patterns are generic and can in principle be applied to any industry. This paper examines how machinery and equipment industry can use these patterns and select a suitable pattern.

2. Problem definition

Machinery and equipment industry is an important economic sector. In terms of turnover, machinery and equipment industry is the second most important sector in Germany and by far the largest sector in terms of number of employees [2]. In contrast to the automotive or process industry, machinery and equipment industry is a SME sector. Approximately 85% of the companies are small or medium-sized [3]. In addition,
machinery and equipment industry is divided into many different branches. This and the fact that there are so many SMEs leads to a strongly heterarchical customer-supplier relationship in machinery and equipment industry [4] [5]. This means that for many companies it is not easy to define who the customer is and what market power the customer or supplier has.

In addition to current challenges for machinery and equipment industry from outside, such as international trade conflicts and the associated gloomy global economy, there are challenges inherent to the industry. The most important of these problems for companies is adherence to delivery dates, followed by manufacturing and development costs. Quality is currently only a minor challenge. [6].

In addition to these challenges, the fourth industrial revolution has created new challenges for machinery and equipment industry in order to defend its position as a world leader. Industry 4.0 is defined as "the intelligent networking of machines and processes in industry with the aid of information and communication technology". This definition is strongly based on technologies for Industry 4.0 cyber-physical systems, cloud computing and smart factory enablers for industry 4.0 [7]. Nevertheless, in Industry 4.0 it is not a technology push that matters but a customer and market-specific solution. In order to carry out a benefit-driven transformation, a suitable business model must be selected, so that Industry 4.0 can be successfully introduced [8].

The strategic problem is that these new technologies make classic business models obsolete. Nevertheless, companies in the machinery and equipment industry are sticking strongly to their long-established transaction-based business models with long life cycles.

Business model development in machinery and equipment industry can be divided into an evolutionary business model innovation that maintains the industry logic and a disruptive business model innovation that changes the industry logic. Previous business model development has often been evolutionary and driven by the customer. Thus, customer tasks were taken over by the company and a transformation from a part to a system supplier was carried out with the associated complexity [7].

Disruptive business model innovations that contradict the logic of the respective industry are currently also being pursued by companies in the machinery and equipment industry sector. The advantage here is that machinery and equipment industry traditionally thinks very customer-oriented and also takes them into account when developing business models. However, current studies show that these business model innovations are often just an idea and do not move towards realization. Companies that are already active in the market beyond the idea phase are confronted with high initial investments and cannot yet realize potential returns, which is why many companies shy away from this step [9] [10].

Although there is a clear advantage of investing in business model innovation [11], companies hesitate. There are two reasons, which are related to each other. First, companies see the relevance of investing in business model innovation, but don’t know how to start and where to invest. Second, companies which have invested in a broad field of business model innovations could not profitable returns from it, only focused investments were profitable by now [10].

Thus, it is necessary to provide assistance and clearly point out the possibilities and advantages of new business models.

3. **Business model and pattern development**

There is no generally accepted definition of the term business model in literature. However, the common definitions all point in similar directions. A business model describes the way in which a company creates value for its customers and earns money in the process. Similar definitions can be found in Osterwalder [12], Gassmann [1] or Nagl [13]. Business models take place at various company hierarchy levels. There are
generic business models up for entire industries or specific business models down to products or services [14]. The framework for business models within a company is formed by the strategy, within which the business model is described and selected [15].

A business model is usually described using a methodology. The methodologies are similar and differ only in detail. For example, Osterwalder uses 9 building blocks to describe a business model, Gassmann 4. When it comes to business model development or innovation, these methodologies usually start from the customer or the value proposition. The development of new business models and the observation and description of business models showed, that many business models are not new and have great similarities [1]. They mostly function according to the same pattern. Examples of such business model patterns are multi-sided platforms, a business model in which the value is not created by the company itself, but the value is created by the exchange of two parties or long-tail business model, in which not the core customers but many niches are served. Table 1 shows a selection of business model patterns that can be found in literature

<table>
<thead>
<tr>
<th>Author</th>
<th>Patterns described by the author</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gassmann</td>
<td>55</td>
<td>generic, transferable to all industries, claim to almost complete mapping of all possible business models</td>
</tr>
<tr>
<td>Osterwalder</td>
<td>5</td>
<td>only exemplary, no claim to completeness</td>
</tr>
<tr>
<td>Kinkel et al.</td>
<td>5</td>
<td>describes only approaches</td>
</tr>
</tbody>
</table>

The approaches presented in Table 1 show a selection of business model patterns. They, however, only describe possible advantages for companies within the patterns by example or are so generic that a sector-specific consideration is not possible. Also a process for the systematic selection of the business models is not used.

4. Framework architecture

In the previous chapters, the problem of machinery and equipment industry with regard to new business models was explained. The state of the art was presented with the deficit that there is no industry-specific approach for machine and plant construction. In this chapter, the basic procedure of the model described in this paper is explained.

The methodology follows three steps. In the first step, "Identification of key business models", the aim is to form a set from which one or more business models can be selected in further steps. In the second step, criteria are compiled on the basis of which a business model is selected from step 1 or on the basis of which a business model can be discarded. After successful selection, the third step addresses implementation. This step is not discussed in this paper.
The aim of the first step is to form a tangible number of business model samples in order to design a selection methodology. The identification of relevant business models begins with the basic set of generic business models (1.). Then, a filter is applied (2.), which screens industry-specific business models that are relevant for machinery and equipment industry. In the next step (3), unobserved business models are added, that are based on existing businesses and immanent in industry or emergent due to new technologies. In the last step (4.), similar business models are merged from the collected business models. Thus, a collection is compiled on the basis of which possible business models at the industry level are then led to possible business models at the enterprise level. In the second step, each of the business models is then checked for consistency using various criteria and a top-down approach. The criteria are as follows: cultural fit, strategic fit, market fit, operational and technological fit.

The objective is not to determine whether a business model works or not. This has to happen in practice or in the next step of implementation. It is examined here, in which pattern difficulties arise, so that the best functioning business model can be selected.

5. Selection process

As described, the selection process is divided into two parts. First, the business models possible for the regarded industry are selected and then, in a second step, those that are possible for a company in the respective industry.

5.1 Identification of key business models

The first step is divided into four sub steps. They consist of the basic set, filters, other business models, and summarizations.
5.1.1 Basic set of generic business model patterns

Here, a basic set is used as the basis for possible business model patterns. The compilation of Gassmann’s 55 business model patterns is suitable. Since 90% of all generic business model patterns are already represented here, this is a good starting point [1]. The transfer of the generic compilation and any other business models that are not yet included in this collection can be carried out in the next steps.

5.1.2 Filter

From the 55 business models, a qualitative filter is now used to select the suitable business models. This filter throws out inadequate business models to select those that are possible for manufacturing companies and known from practice. This results in the following business models: Digitalization Sensor as a Service, Two-Sided Market, Guaranteed Availability, Mass Customization, Pay-per-Use, Performance-based Contracting, Rent Instead of Buy and Subscription.

5.1.3 Adding further business models

The eight business models mentioned above will, in the next step, be expanded to include others that do not count among the basic set, but nevertheless already appear in the regarded industry. For instance, the Mass Customization business model only carries out individualizations [1]. In industry, the additional business model of a Configurator has been developed, due to increasing complexity. This makes it possible to map complex product designs using standardized interfaces. Production and the customer largely automate this process, since the customer configures the product himself and can thus trigger a production order.

5.1.4 Merging similar business models

Some of the nine business model patterns, that are similar in the regarded industry can now be merged. On the one hand, Two-Sided Market and Digitalization as a service can be combined to form platform business model patterns. The similarities occurring here are on the one hand a marketplace for physical goods and on the other hand a marketplace for data. In both cases, the company acts as a broker between suppliers and buyers. A further pooling can be carried out with guaranteed availability, pay-per-use and performance-based contracting to service-oriented business models. All three business models combine a transformation of transaction-based sale of machines towards service orientation. On the one hand, the input (availability), the throughput (pay-per-use) and the output (performance) are sold service-oriented, i.e. as a service.

6. Selection process for a suitable business model for a particular company

The selection of the respective possible business model patterns at company level must be made individually for each company. This is why a top-down approach is used here.

6.1 Company Culture Fit

First, it is checked whether the cultural fit of a company or the attitude towards business model innovations is given. This is particularly decisive for a transformation to a platform-oriented business model. This is because a rethink must take place throughout the entire company [16]. Since here money is no longer earned through the transaction of self-produced goods, but the company acts as an intermediary between suppliers and customers. Particularly critical is the entry into a cooperation with former competitors and thus the thinking in business eco systems.
6.2 Company Strategy Fit

At the strategic level, a decision must be made whether and to what extent investments in new business models should be made. It must also be evaluated at the strategic level, how external pressure will affect the company in a medium-term perspective, and how foreign players will penetrate the existing business model and disrupt it with new business models [7].

6.3 Market Position Fit

At the market level, the environment, development and the company's past and future position must be taken into account. The decisive factor here is whether a company possesses the market power to approach its customers with new business models. It is also decisive, how the market environment will change in the future. For example, new suppliers, mainly from emerging markets, who offer more competitive prices at a reasonable quality, can pose a threat. Service-oriented business models thus can offer the advantage of higher quality and an advantage in the investment decision like a transformation from capital expenditure to operational expenditure.

6.4 Operative and Technical Fit

Decisions must be made at operational and technological level, how a suitable new business model can be introduced. In particular, a configurator and the associated automatic or automated production and assembly of products require the technical possibility and a high degree of operational excellence. But also other business models need competences in a company due to their new high demand of ICT technology.

7. Summary and Outlook

In this paper, a methodology has been developed that describes an industry-specific approach for machinery and equipment manufacturers that can operate new business models as a result of the changes within Industry 4.0. First, the current and future problems were presented on the basis of current literature, in order to then present the state of the art. It was shown that so far, industry-specific consideration of business models do not exist. Furthermore, a methodology for selecting a model was presented and described in detail. First, a branch-wide compilation of possible business models was described, followed by a company-specific selection.

Further research should be on the implementation and integration process for the selected business model. This would then focus on organizational change and address issues mentioned in chapter 6. Also, further research in similar industries like process or automotive industry makes sense, since industry specific research on business models is relatively scarce by now.

Acknowledgements

The Work presented in the paper has been supported by EU Horizon 2020 grants No. 664404 “EPIC—Centre of Excellence in Production Informatics and Control.”
References


[11] Lindgardt, Reeves, Stalk, Deimler 2013: Business model innovation, When the game gets tough, change the game, in: Own the future, 50 ways to win from The Boston Consulting Group, Wiley, pp. 291–298


Biography

David Maximilian Doerr (*1989) is a research associate at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA. He studied industrial engineering at the RWTH Aachen University. Since 2017 he is coordinator for business models at the Fraunhofer IPA. Since 2019 he has been working in the Corporate Strategy and Development Department and is Deputy Manager of the Fraunhofer Project Centre for Smart Manufacturing in Shanghai.
Prof. Dr.-Ing. Thomas Bauernhansl has been director of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart since September 2011 and is director of the Institute for Industrial Manufacturing and Operation IFF at the University of Stuttgart.

Bauernhansl studied mechanical engineering at the RWTH Aachen University. After receiving his doctorate with distinction, he was employed by the Freudenberg company from 2003, most recently as director of the Technology Center at Freudenberg Sealing Technologies. His focus was on location planning, factory planning, production relocation, technology management and lean production. Bauernhansl was most recently responsible for the Europe, North and South America regions and optimized the global footprint of the production network, which at that time comprised over 50 sites.

Thomas Bauernhansl is involved in numerous advisory boards and executive committees in industry, associations, research and politics and is a member of the WGP, the Scientific Society for Production Technology, as well as a member of the Strategy Circle of the Platform Industry 4.0 of the Federal Government and Deputy Chairman of the Steering Committee of the Alliance Industry 4.0 BW. He is the author and editor of numerous books on topics such as transformability in production, Industry 4.0 and management in production.
Abstract
Logistics and supply chain management have undergone significant change due to technological changes in the recent years. The classic transport, handling and storage processes with a strict functional orientation have been transformed into a global, network-oriented task-field. The future challenges are individual customer requirements, shorter delivery times and increasing cost pressure. Due to these challenges and the increasing globalization, companies are confronted with ever more complex supply chain networks. The digital transformation is intended to remedy this situation. New technologies, comprehensive real-time information availability and agile value creation networks are just examples to meet these challenges.

This paper provides an overview of the expected developments in supply chain management over the next 20 years. Based on ten future megatrends, four main topics (technology, control tower, value adding and green logistics) were derived. The focus of this paper is on OEMs and Tier 1/n suppliers. Both are undergoing a major change on the customer and supplier side due to their central position within the supply chain.

Keywords
Supply Chain Management; Supply Chain Networks; Digitalization, Industry 4.0; Sustainability; Value Adding

1. Introduction
There are several development phases in logistics: In its beginnings, it was limited to the temporal and spatial optimization of transport processes. The first technologies and approaches for standardized transport and for company-wide coordination of material flow, storage and transport systems were developed. In the following years, this developed into today's understanding of logistics as a holistic management theory. Supply chain management, as well as logistics today, make a decisive contribution to the success of a company and are now even deeply anchored in secondary processes such as financing or insurance of goods [1]. The future challenges for manufacturing companies as well as for the logistics industry lie above all in individual customer requirements, shorter delivery times and increasing cost pressure. Based on these challenges and the increasing globalization as well as the required comprehensive environmental protection, companies today have to cope with increasingly complex supply chain networks [2]. In the course of the fourth industrial revolution, the next development phase also begins for logistics. The previously rigid value chains are developing into increasingly complex, intelligent networks in which goods and information are exchanged not only between individual players, but also between all of them. In concrete terms, digital transformation describes "the transformation of value-added processes through the further development of existing and the implementation of new digital technologies, the adaptation of corporate strategies on the basis of new digitalized business models and the acquisition of the necessary skills and qualifications" [3]. Logistics and
supply chain management will increasingly be affected by these changes. However, the extent to which companies and the associated processes, products and business models will change is still relatively unclear.

2. Methodic procedure

In order to be able to make as precise a statement as possible on the development of the supply chain management in the next 20 years, a multi-stage approach was chosen.

In the first step, the conventional supply chain was separated from a smart supply chain. In addition, framework conditions and assumptions for the year 2040 were subjected to closer examination from an economic, political and social point of view.

In a second step, an online survey was conducted along the entire supply chain. The aim of the survey was to identify current research topics and technology projects of the participants as well as future challenges and needs. Based on the online survey, expert interviews were conducted to identify further insights into future investment needs, challenges and customer requirements of OEMs, suppliers and logistics providers.

In total, the participants are assigned to the individual stages of the supply chain as follows (Figure 1):

3. Results

By the year 2040, politics, economy and society will change dramatically and place new demands on manufacturing companies. The nature will also change dramatically and offer new challenges for the physical supply chain. These changes can already be seen today and can be described in the form of megatrends. For decades, megatrends have brought about a slow, far-reaching change in society as a whole and have influenced the framework conditions for economic activity. New technologies and innovations contribute to this, as do cultural and political developments. Based on literature research and expert assessments from industry and research, the following 10 megatrends could be identified [4–7].

1. Individualization
2. Digitization / Connectivity
3. Demographic change
4. Urbanization
5. Globalization
6. Sustainability and social responsibility
7. Mobility
8. Data security and ownership
9. Servitization
10. Knowledge culture and information society
Based on these megatrends, four main topics were derived for this study. These form the basis for the detailed analysis of the results, which will be discussed in more detail below.

### 3.1 Technology

The global transport volume will continue to rise in the short and medium term. On the one hand there will be an increasing need for individualization in the consumer goods sector and a growing demand for such products in emerging markets such as Asia or Africa. On the other hand, there will be increased global sourcing in the industrial goods sector in order to further cut costs and secure competitiveness. According to 65% of respondents, this leads to a more complex and growing supply chain. Contrary, 30% of those surveyed assume that there will be a decline, since individualization can be localized using additive manufacturing processes or individualization can be mapped using software. However, until these trends have become established, there will be an increase.

The increase leads to the need for a supply chain aligned to the latest technologies. A large number of technologies are currently available for this purpose. In the study, the participants were asked about their assessment of these technologies.

![Figure 2: Opportunities for logistics improvement](image)

Figure 2 shows which technologies the respondents see as the most promising for coping with supply chain challenges. A clear dichotomy between technologies can be seen. Thus, a group of the first five technologies can be formed, which most respondents (> 60%) regard as trend-setting. Only a small number of the interviewees (< 40%) believe that the other technologies have a chance.

The greatest chances of success are attributed to artificial intelligence technologies. Fields of application include the analysis of large amounts of data in a short period of time. Fields of application are the reduction of freight costs, the improvement of delivery performance and machine learning in collaborative supply chain networks [8]. However, large amounts of data are required for this as described. Therefore, the following technologies from the survey cannot be considered separately. Sensors, for example, generate data automatically and thus more frequently and more cost-effectively than manual human input. Big Data stores, manages and makes this data available.
Another technology that will have a major impact on the supply chain of the future is autonomous driving. While previous technologies will support the human work, activities will be completely taken over here. Large changes are often attributed to the autonomous driving of passenger cars in a media-effective way. However, the respondents assume that there is great potential in the autonomisation of aircraft such as drones. The networking of transport vehicles, such as trucks, which can be combined into a network through “platooning” [9], also has great potential for future supply chains. Despite its disruptive character, the impact of autonomous driving is still very small at present, as the technology is still under development and many companies shy away from high development costs, see Figure 3.

![Figure 3: Participation in Autonomous Driving](image)

While additive manufacturing is only slowly gaining acceptance in the consumer sector, it has great potential in the B2B environment. Additive manufacturing processes can be used, for example, to produce moulds that are not possible using ablative manufacturing processes. Thus, they are not seen as a substitute for classical manufacturing processes. They show further advantages through decentralized service-oriented business models. For example, railway and aircraft operators and service companies use 3D printers for repairs in order to manufacture spare parts on site and thus save transport costs [10].

In summary, the study on technology shows that the most rapidly changing parameter will be the lead time. However, there is disagreement among the respondents as to whether the duration will increase or decrease. 66 % of the respondents see a shortening of the cycle time through the technologies described above, see Figure 4.

![Figure 4: Development Lead Time](image)

The flow of goods and information will continue to increase as a result of automation and digitization. However, 33 % of those surveyed see an increase in throughput time, as increased environmental restrictions will have a far greater impact than technological innovations.

### 3.2 Value Adding

Technical innovations in communication and information processing, permanent organizational change, the internationalization and digitization of business systems and processes, and the concentration on core competencies right through to virtual companies are leading to a new competitive landscape and suggest that logistics will continue to play a key strategic role as a company-specific value creation factor in the future.
Consequently, logistics is no longer seen as a pure cause of costs, but as a means of differentiation from the competition. The first idea behind the term "value adding" is the refinement of products in the course of the value creation process. But value adding is also possible in the context of supply chain management. The magic word for logistics service providers is: Value Added Services. This is not about increasing the value of the actual physical product, but about offering the customer an additional benefit/value or optimizing the cost-benefit ratio along the supply chain. This increases both customer satisfaction and customer loyalty. It would be conceivable, for example, to offer additional products related to logistics services. The value creation options today range from the configuration of e-commerce kits to the insertion of coupons or brochures in packages. It is also conceivable that the logistics service provider could reduce the number of suppliers by taking over upstream and downstream processes. Production or assembly during the transport process will also be possible in the future. The respondents see product monitoring with the aid of sensors in and on products in the supply chain as the most widespread in 2040, see Figure 5. Real-time tracking of products and processes as well as integrated condition monitoring (e. g. temperature, gases, vibrations, etc.) are conceivable areas of application here.

![Figure 5: Value Adding](image)

Product monitoring also has a high potential in the trade sector. Examples include real-time monitoring of ripening processes or full-time monitoring of the cold chain. On the one hand, this increase or ensure the quality of the products, facilitate documentation processes and reduce product rejects. The data obtained from product and process monitoring can also have an influence on the planning and control of processes. In the course of the expert interviews, the thesis of best-before date-controlled logistics was expressed. The products independently determine their flow rate in the supply chain on the basis of previously defined characteristics. Such a paradigm shift in the supply chain requires not only sensors on products, but also the adaptation of numerous processes and IT systems. Dynamic pricing depending on the condition of the product is also conceivable, e.g. in the food sector. The own control of the products through the supply chain and the pricing based on the condition is mainly made possible by IoT chips in/at the food product. An extension of the application cases to the pharmaceutical or feed industry is realistic.

A further aspect in the area of value adding is increasing safety. This includes threats that can be influenced (e. g. supplier failures, mistakes) as well as threats that cannot be influenced (e. g. natural disasters, environmental influences) [12]. Theft of goods also poses a threat to companies. In 2016 alone, goods worth 1.3 billion euros were stolen from trucks in Germany [13]. By using IoT or GPS data, customers can, for example, be offered real-time goods traceability down to package level across the entire supply chain and at the same time minimize susceptibility to theft. In addition, it is also conceivable that the logistics service provider might offer its customers additional insurance policies in order to financially cushion any
disruptions, e.g. in the process. Premiums could be determined dynamically based on the condition and position of goods.

The automatic picking of goods for delivery is another option in the area of value adding. Possible applications include automatic palletizing, depalletizing and container stacking as well as container destacking. On the one hand, this shortens throughput times and also avoids errors during order picking. Logistics employees are relieved of heavy, but also sometimes monotonous work and can thus devote themselves to other tasks. The unattended delivery of goods is also conceivable, e.g. by means of drones. Logistics companies can dynamically offer unloading locations and thus enable a higher delivery rate for the customer.

The parallelization of production or assembly with the transport process (e.g. on trains or container ships) has been discussed in practice and research for some time. On the one hand, parallelization would significantly shorten delivery times and also enable companies to increase their internal production capacities. The respondents are rather skeptical about the possibility of parallelization. On the one hand, it is unclear which safety requirements would be necessary for this, and on the other hand, the necessary requirements for stability during the transport process are difficult to achieve. Another important aspect that is difficult to overcome is the imbalance of the transport streams and thus the imbalance of the machine utilization. According to experts, parallelization would be desirable, but not feasible.

3.3 Control Tower

The trade with services and data is already almost more lucrative in many places than with goods. Due to the growing importance of data and its analysis, digital services will become increasingly important in the future. For example, a survey of experts from the forwarding industry shows that the pure transport of physical goods will almost no longer be profitable in the future [14]. The recording of all relevant transport parameters, their analysis, as well as the passing on of data/information to third parties is developing into a decisive competitive advantage. Another possibility is the networking of global trading and transport companies. Companies are increasingly developing into complex value creation networks in which information must be exchanged not only between two instances, but between all. 72% of those surveyed also see growing complexity and gave the following reasons for this, among others:

- Developing and emerging countries as further players in the supply chain
- Supply chain disruptions due to political and environmental influences
- Increasing transport volumes due to improved transport processes and transparency
- Expansion of the customer and supplier network leads to disproportionate increase in the number of players
- Shorter product life cycles increase transport quantity and speed; both in delivery and return logistics

A further factor is the difficulty of efficiently managing the entire supply chain due to growing transport volumes and the constant demand for shorter lead times. In the past, for example, communication and transport organization with the freight forwarder and, if necessary, an agent used to take place. Nowadays, portal systems are in use at every transport break and startups support every actor in the transport chain with web platform solutions.
In order to meet these future challenges, a central planning instance must be set up. The Control Tower is a central, independent and self-optimizing platform for the comprehensive regulation of the information flow and exchange of all data. In a conventional network the information is exchanged directly between the actors. The final stage of the Control Tower is a network in which all data flows with a 1:1 connection between the Control Tower and the respective actor. The overriding goal is an overall optimum of all actors within the supply chain network through secure and efficient data exchange.

3.4 Green Logistics

The following premise applies predominantly to the ideas presented for future logistics: Logistics must become much more environmentally friendly. This is not only demanded by politicians and environmental associations, end consumers, but also by companies. For example, Daimler is aiming for CO₂-neutral plants in Germany by 2022 [15]. It is highly probable that logistics in 2040 will be significantly more environmentally friendly, but not completely emission-free. There are many trends in this direction. As more and more goods are transported around the world, not only the transport volume but also the number of participants in the transport chain will increase. As a consequence, any environmentally harmful empty transports and relocation processes will increase proportionally [16]. The positive trend towards 100 % recycling of materials (circular economy) means that further transports are caused. All in all, this means that more transport vehicles have to be produced - even if they are driven emission-free. Finally, it must be realistically estimated that aircraft and container ships, among other things, will probably not be able to operate completely free of kerosene or heavy fuel oil in the next 20 years. In the future, customers will show more and more willingness to pay for a green and sustainable (complete) supply chain. A surcharge for sustainably produced and transported products is also conceivable. According to the experts surveyed, it will be a challenge in the future to get away from "greenwashing" or "marketing sustainability".

Reverse logistics, a kind of recycling economy, will be one of the main components of green logistics in the future. Around one third of the experts surveyed believe that in 2040 there will be a comprehensive use of return logistics and that almost every product can be reused / recycled. 54 % of the respondents, on the other hand, are of the opinion that a partial use of return logistics will take place in 2040.
The aim of recycling management is the sustainable and efficient use of resources. The driving forces behind this are legislation, corporate responsibility and potential economic benefits.

4. Conclusion

The year 2040 will reach us faster than we can imagine. Looking back, it is about 20 years ago that we have dealt with the Millennium Problem. Likewise, it is less than 20 years since the agile manifesto was defined. The possibilities for innovation in the supply chain are virtually limitless thanks to modern technology. But they will always be limited by the speed of change and adaptability of the users. In addition, external conditions and natural events will continue to restrict or at least slow down the further development of the supply chain.

Against this background, how must companies set up their supply chains in order not only to survive in 2040, but also to be able to offer first-class products and even better services? In the future, profits will not be made by the cheapest, but by the most innovative and flexible companies. Likewise, there is a need for a major rethink from looking at a flow of (raw) goods to the entire value chain.

Based on the identified focal points, the next step is to identify fundamental conflicts of objectives (e.g. reusable packaging increases the number of empty trips) as well as fields of action. Based on the fields of action, recommendations for action are to be developed for the manufacturing industry as well as for the logistics sector.

Acknowledgements

The work presented in the paper has been supported by EU Horizon 2020 grants No. 664404 “EPIC—Centre of Excellence in Production Informatics and Control.

References


**Biography**

**Martina Schiffer** (*1984) is a research associate at the Fraunhofer Institute for Manufacturing Engineering and Automation (IPA). Her research focuses on supply chain management and production planning and control. Before joining the Fraunhofer IPA, she worked 3,5 years as a Process Engineer for a plant manufacturer for metallurgical and rolling mill technology.

**David Maximilian Doerr** (*1989) is a research associate at the Fraunhofer Institute for Manufacturing Engineering and Automation (IPA). He studied industrial engineering at the RWTH Aachen University. Since 2017 he is coordinator for business models at the Fraunhofer IPA. Since 2019 he has been working in the Corporate Strategy and Development Department and is Deputy Manager of the Fraunhofer Project Centre for Smart Manufacturing in Shanghai.
Innovative logistics concepts for a versatile and flexible manufacturing of lot size one

Ali Bozkurt¹, Manuel Hagg¹, Robert Schulz¹
¹Institute of Mechanical Handling and Logistics, University of Stuttgart, Stuttgart, Germany

Abstract

Due to the high amount of car models and the increasing number of variations in the premium car segment, production logistics in the automobile industry will face new challenges. Conventional logistics concepts such as Henry Ford and Frederick Taylor’s conveyor technique will become rare. While conveyors are best-practice for high output and economical production, they are inflexible with high variations. Moreover, an assortment of different parts have to be provided at the line, which results in space constraints and inefficiency in increasing product variety. Hence, it is necessary to focus on contemporary logistics concepts and equipment in order to cope with customer demand when producing in lot size one. Therefore, various logistics concepts and components have been developed by the Institute of Mechanical Handling and Logistics (IFT) of the University of Stuttgart over recent years. The new logistics concepts are more versatile and enable a more flexible and changeable production logistics for a wide range of different products. The aim of this study is to provide an overview of the current situation of the innovative production logistics concepts and describe possibilities for further developments.

Keywords

Production logistics control; economic efficiency; automobile production; material handling, versatile logistics

1. Introduction

Since 1913, automobile production has been characterized by Henry Ford and Frederick Taylor’s moving assembly line system, which was a great invention for production methods at that time [1]. The assembly line system was an economic method to produce homogenous products in high quantity and contributed to a great development in the automotive industry. Because of today’s social trends like the growing demand of individualization in the premium car segment and the development of alternative drive units e.g. electric or hydrogen drives, manufacturers have to deal with much more variety and complexity [2–4]. For example, the latest Audi A3 was available in $1.1 \times 19^{38}$ possible configurations, which leads to a smaller amount per variant and up to lot size one [4]. In addition to other key issues, like volatile sales figures, the automobile industry will go through a transformation process in the next few years. Therefore, manufacturers have to consider new ways of production to overcome these challenges and allow a new level of freedom and the simultaneous production of different models and variations within the same factory [3–5].

In consideration of those challenges, new production concepts like the matrix production, flexible cell and fluid production have received much more attention over recent years [4,6]. In this way, the new production approaches imply a demand for flexibility and economic production with a high amount of variants [4,6].
However, the introduction of new production approaches has also led to a change from conventional logistics concepts to contemporary innovative logistics concepts. In the automotive production process, material supply takes place at the point of assembly. Small parts are supplied in small load carriers or in boxes. Larger and valuable parts are supplied in large universal bins. In this concept, the assembly worker picks all the necessary parts from the defined storage racks [7]. In most cases, gravity flow top-up racks are used. Each bin in this case is sequenced with the production cycle and, if the bins are empty, new carriers with parts will be requested by the consumption-driven KANBAN concept [7,8]. Inspired by Japanese car-makers’ success, production logistics in the automotive industry is characterized by the logistics concepts just in time (JIT) and just in sequence (JIS) [8]. JIT is an approach to minimize the inventory by delivering parts exactly when they are required. JIS focuses on the management of high numbers of different parts by sorting and providing them as previously scheduled [9]. However, this concept is mainly characterized by space constraints and inefficiency by increasing the amount of different parts [9].

There is no denying that the logistics concepts JIT and JIS are good methods for reducing costs and the amount of storage, but the expected increase of several variations will make them more difficult to manage with conventional logistics concepts [10,11]. For this purpose, the IFT is working with partners from industry on the development and validation of new production logistics concepts and components [12–18]. Thereby, the focus is on automotive production, but some of the concepts can also be transferred to other industry sectors.

The remainder of this paper is organized as follows: Section 2 provides an overview of the state of the art and presents developed logistics concepts and components; Section 3 deals with strategies for production logistics control as well as possibilities for novel economic efficiency methods; Section 4 concludes the key points of this paper and provides an outlook.

2. State of the art

The increasing amount of different parts at the point of assembly and new production concepts such as those introduced in section 1 have brought the established concepts to their limit. Therefore, new logistics concepts and components have evolved. Some of the concepts and components have already been implemented in automotive production processes and others are being researched and have to be tested under real circumstances [19,20].

2.1 Production logistics concepts

Set concept

The first concept is the set concept or kitting concept in which specific sets are bundled for each vehicle and supplied with a shopping cart at the point of assembly. In this case, the parts are within the reach of the assembly worker. The set concept is especially suitable for all small and medium sized parts which are needed for a specific number of assembly stations [7,8,15]. After the set building process, each set follows the allocated vehicle in shopping carts as the vehicle gets assembled at a defined number of assembly stations. Bigger parts will thereby be transported in the traditional way by larger bins or power and free conveyor systems. When all the materials are taken out, the empty shopping cart will be transported back to storage or to the supplier for further refills [3,7,15]. The main advantage of the concept is that all the necessary parts are supplied in the right order of the assembly and there is no necessary additional time for searching and identifying material or walking, which results in a better cycle time and efficiency in the production process [14,16,18].
**Single Autonomous Guided Vehicles (AGV) concept**

The next concept is the use of small, single and economic AGVs for the material transport. Each bin carrier will be transported by a single AGV to the assembly stations. Thereby, the paths for the AGVs are planned by using navigation software [16,17]. After the removal of the load carrier, the AGV will be routed back to storage for a new transportation order [17].

**Mobile supermarket**

The basic idea of the “mobile supermarket” production logistics concept was developed by Wehking and Popp [15] and is based on several components, designed by Hofmann [12]. Each component can be used separately, but can also be ad-hoc synthesized and build in cooperation the mobile supermarket concept. The mobile supermarket concept includes the following components: rack units for the storage of load carriers and a transportable automatic rack feeder for the loading or unloading of the load carriers [12].

A single AGV will be applied to transport the mobile racks to the assembly stations [14]. Due to the well-defined storage locations and the coordinated automatic material supply, the assembly worker will only be supplied with the necessary parts, which reduces the possibility of errors. Another benefit with this approach is that a large number of different parts will be supplied at the exact moment of the assembly, which is a further advantage compared with the JIT and JIS concepts. Moreover, just-in-real-time logistics (JIR) enables the sequencing of the next required parts at the very last moment without a fixed delivery time and exact defined storage location. The differences between JIT, JIS and JIR are shown in Figure 1 [12,14].

![Figure 1: Just-In-Real-Time (JIR) logistics](image)

**2.2 Logistics components**

**AGV**

The AGV, which is shown in Figure 2 is a special conception of the IFT and can be applied for several transportation processes. The modular design of the AGV allows the transportation of large parts, work pieces, assembled parts as well as the transportation of racks (Figure 2). The modular design of the AGV is therefore an approach to reduce the amount of different AGVs in production [12].

**Mobile assembly platform**

For a better value-adding logistics process in automotive production, Hofmann [12] has developed a disruptive innovation called the “mobile assembly platform”. The mobile assembly platform hosts an area for assembly workers as well as a docking system for mobile robots [15]. Due to this, production logistics will undertake tasks that are associated with being part of the assembly sector, using a traditional understanding of assembly and production logistics. As a result, the synchronization and integration of
logistics and assembly processes leads to value-added production logistics [12]. In contrast to conventional line technologies, the mobile platform vehicle can be discharged out of the assembly process in the case of a quality issue or other production issues such as supply bottlenecks [12]. After solving the issue, the mobile platform can be incorporated back into the assembly process. The mobile assembly platform can also be applied for the production of different models and variations from different manufacturers [12].

**Double skid system**

The Double skid system, which was designed by Weber and Wehking [17], can be used for the transportation of pallets. Due to the special design, the two forks drive in parallel without being connected, can drive under each standardized Europalet and can lift the pallets up. The double skid system is appropriate for the transportation of large and heavy parts with a maximum weight of 1000 kg. An integrated spindle drive is applied in this case for handling the necessary lift, drive and steering motions [18].

Figure 2 shows all the described components developed at the IFT. The prototypes of the components can be visited at the research campus Active Research Environment for the Next Generation of Automobile (ARENA2036) at the University of Stuttgart. The components for the supermarket concept including the racks and the automatic rack feeder are shown on the left side. The mobile assembly platform is in the middle and the single AGVs and the double skid system are shown on the right side.

![Figure 2: Innovative logistics components][18]

### 3. Further development of production logistics concepts

#### 3.1 Decentral production logistics control

Within versatile and flexible logistics and production systems, the redesign of assembly stations, process planning or control as well as target-oriented decision-making is challenging [21]. Thereby, a system can be defined as groups or sets of connected, interacting or interdependent elements with certain relationships that form collective entities. Usually, a system consists of elements, interconnections and a goal [22].
elements like a single AGV can be defined as the smallest system units. Interconnections describes the relationship between elements and determines the system structure [23].

The complexity of controlling and planning a modern production and logistics system is increasing significantly. The key drivers of complexity are based on external aspects like the volatility of the procurement market or the high number of companies involved in the supply chain structure as well as internal aspects such as an increased degree of freedom in production logistics processes [24]. In the first instance, using modular intralogistics components as described previously and shown in Figure 2 will further increase system complexity. One reason for that is that the higher number of interconnections and interfaces between system elements leads to a higher coordination effort because of the rising number of options available in fast-changing production logistics environment. However, these approaches offer opportunities for handling the challenges that logistics and manufacturing systems are facing such as manufacturing in lot size one. However, the impact of decision-making is difficult to estimate because of existing uncertainties within production logistics systems. For example, uncertainty can be traced back to the existence of interfaces and interdependencies between the system elements. An important system in the context of versatile and flexible manufacturing is the AGV system that consists of several elements [25]:

- One or several AGVs
- Guidance control system
- Devices for position determination and localization
- Data transmission equipment
- Infrastructure and various peripheral installations

Peripherals are built and installed particularly for the usage of AGVs, whereas infrastructure is usually already present in the environment and needs to be considered when operating the AGVs [26]. The AGV guidance control systems are responsible for the coordination of the system’s vehicles and the integration of the AGVs into internal processes [27]. In this case, coordination means balancing the individual activities of the single elements within the system in respect to an overall objective [28].

However, the tasks assigned to the guidance control system are context-sensitive and depend on factors like the competency of the AGVs, interaction with a higher level material flow control system or the degree of decentralization. In centralized control approaches, guidance control systems comprise all the relevant tasks like transport order management, vehicle dispatching, travel order processing or further service functions. This also includes coordination and communication with peripherals or mechanical interfaces (e.g. interaction between vehicle and loading aids) and functional interfaces (e.g. information flow such as the data exchange of loading conditions) [26].

For several years, these central control approaches were sufficient. However, with the increasing complexity of production logistics systems, these hierarchically structured concepts have reached their limits. One reason is the availability of different types of AGVs, for example, regarding their navigation principles. In addition, several AGVs from different manufacturers operate in the same environment, which cannot be managed by a single guidance control system due to missing compatibility. Another reason for this is that these approaches are rigid and inflexible due to a central decision-making instance. Therefore, different decentralized control concepts have been developed with the goal of shifting decision-making authority to individual entities [29]. Thereby, system complexity is reduced by breaking down a complex task into several sub-tasks for finding an adequate problem solution. For example, a group of closely spaced AGVs in a production layout independently negotiates who executes the next transport order without interacting with a central instance. A prerequisite for this is that the vehicles or infrastructure has the ability to interact with the required entities being part of the environment or sub-system. Consequently, system components are enabled for target-oriented simultaneous and locally concentrated decision-making.
However, since production and production logistics activities are becoming more and more interlocked, it would not be successful to optimize these sub-systems independently as standalone solutions in the near future. In particular, many current decentralized control approaches address and optimize the objectives of isolated sub-systems such as production or logistics. As a result, the control strategies of these approaches are often based on the minimization of travelling distance and energy consumption of AGVs or optimizing the workload of the transport or manufacturing system [30–32]. One reason for this is the architecture of software products available on the market that often has a modular structure, allowing it to operate independently as a self-contained system. Following this, it is questionable whether decision-making by these control approaches goes along with the strategic goals of the company. Rather, production and logistics need to be perceived as an integrated system with the result of preventing local functional optimization and increasing added value for the company. Thereby, it is conceivable that a sub-system (e.g. fleet of AGVs) accepts a short-term disadvantage (e.g. higher transportation costs) to prioritize the manufacture of a car of an important customer that helps to achieve the company’s strategic goals, more than producing cars of other customers, for example, due to a higher contribution margin.

Within a research project, the IFT of the University of Stuttgart intends to break the cycle of isolated and locally concentrated decision-making by developing a dynamic control approach. The decision-making mechanism of this decentral concept focuses upon a company’s individual strategic goals. Following this, the decisions made by the production and logistics components go along with the long-term goal of the companies. Since companies have different strategic goals, such as return on investment or revenue maximization, the control strategies developed within the research project will face several possible strategic goals. In the first instance, revenue, operation income, economic added value as well as customer satisfaction will be addressed as company goals. This development is required since the production and logistics system, and therefore the underlying software structures, executes more and more tasks due to increasing system complexity. However, this implies that the decision-making of production or logistics managers and production logistics components should be balanced and address the same goals. For instance, if the quality of managerial decision-making is measured based on achieving company goals, such as a defined economic added value, the decision-making process of the production logistics system should follow that logic as well.

At the moment, different strategies are tested and investigated by using an agent-based simulation approach. A critical success factor is how to deal with data underlying the decision-making process. Here, the short-term action alternatives at the shop floor level should be reconciled with the long-term success of the company. This requires a close interplay of the long-term-oriented Enterprise Resource Planning (ERP) and the short-term-oriented Manufacturing Execution System as well as an integrated consideration of the AGV systems.

### 3.2 Economic efficiency and decision-making

New logistics concepts are bringing economic benefits compared to the conventional concepts like the conveyor system. For example, conveyor systems are planned and calculated according to the production planning for a product cycle. After the product lifecycle, the logistics components are scrapped and in most cases a further use is not possible. Innovative logistics concepts, in contrast, provide the opportunity to reuse the components. A component such as the single AGV, as shown in 2.1, is designed for the transportation of a wide range of different parts and for several transport processes. In addition, due to the new concepts, there is less storage use for material supply than in conventional supply strategies like top-up racks. Another point to consider is that, due to business model innovations of the logistics component manufacturers in recent years, there are new opportunities for components acquisition.

The logistics components can, for example, be acquired via a leasing model. This opens the possibility to make short-term decisions and react to market demands.
Especially in the case of workload spikes or volatile sales numbers, leasing models should be considered. There are two options for leasing models [33]:

1) The leasing company is an external partner and provides the components
2) The leasing company is an internal business unit or cost center

In most of the concepts, no pre-commissioning process is necessary, which also reduces the handling steps. Moreover, the new concepts also benefit the production process. Due to the material supply at the point of assembly with the JIR concept, the assembly worker can thereby focus on the value-adding processes.

However, for better comparability, not only quantitative factors like investments costs should be viewed. Furthermore, the concepts also show qualitative factors, which are difficult to measure but are assessable for each proposed system [34].

- Scalability: Each concept can be scaled to different lot sizes.
- Universality: Most of the concepts can be used for a variety of products.
- Modularity: The modularity of the components allows the combined use and single use of the components.
- Compatibility: The interfaces of the components enable connections with each other [34].

Most research results show that there are only strategic considerations like investment decisions for logistics components or the outsourcing of logistics processes. However, for most of the logistics concepts in the automobile industry, there is no approach to evaluate and compare the different logistics concepts regarding their costs and characteristics. Therefore, the IFT is working with different partners on the development of new methods to evaluate and select innovative logistics concepts, which will promote a solution for short-term as well as long-term decision-making.

4. Conclusion and outlook

The purpose of this paper is to describe the current situation of the production logistics concepts and components that are needed for a versatile and flexible production of lot size one. One of these topics is the development of a decentralized production logistics control concept that places a decision-making mechanism that focus upon companies’ individual strategic and long-term goals. Another topic is the development of new methods that generate more economic efficiency like leasing concepts for production logistics components. Moreover, the next important steps are taking place to establish those concepts and components and test them under real circumstances.

Acknowledgements

The research presented in this paper has received funding by Ministry of Economic Affairs, Labour and Housing of Baden-Württemberg, Germany (Az: 33-4220.21/51: DezLog (Decentralized Logistics Control Concept) and FlexProLog (Flexible Production Logistics) (Az: 7-433262-IFT/3).

References


Biography

Ali Bozkurt (*1991) has been working for the Institute of Material Handling and Logistics at the University of Stuttgart since 2018. He studied Industrial Engineering and Management at the University of Erlangen-Nuremberg.
**Manuel Hagg** (*1990) has been working for the Institute of Material Handling and Logistics at the University of Stuttgart since 2017. He studied Technically oriented Business Administration at the University of Stuttgart.

**Robert Schulz** (*1970) is holder of the Chair of Mechanical Handling, Intralogistics and Technical Logistics of the Institute of Mechanical Handling and Logistics (IFT) at the University of Stuttgart. His main research areas range from classic mechanical handling technology to adaptable logistics systems. Schulz studied Mechanical Engineering at the University of Stuttgart and received his doctorate in 2002. He began his industrial career as a project manager for simulation and digital factory planning at Dürr Systems AG and then he moved to AUDI AG.
Dynamic criticality assessment as a supporting tool for knowledge retention to increase the efficiency and effectiveness of maintenance

Theresa Passath¹, Cornelia Huber¹, Hubert Biedermann¹
¹Department of Economic- and Business-Management/Montanuniversitaet Leoben, Leoben, Austria

Abstract

Digitalisation offers industrial companies a multitude of opportunities and new technologies (e.g. Big Data Analytics, Cloud Computing, Internet of Things), but it still poses a great challenge for them. Especially the choice of the maintenance strategy, the increasing complexity and level of automation of assets and asset components have a decisive influence. Due to technological progress and the new possibilities offered by industry 4.0, the interaction of different systems and assets is essential to increase the efficiency of the maintenance processes within the value-added chain and to guarantee flexibility permanently.

These factors lead to an increased importance of a process methodology for a dynamic evaluation of the asset’s condition over the entire life cycle and under changing framework and production conditions. Therefore, legal and environmentally relevant requirements are considered, based on the procedure of HAZOP (IEC 61882), and ensure the traceability of the results and systematically record the asset’s knowledge gained this way so that it is not tied to individual employees, as it is currently the case.

The criticality assessment as a basic component of Lean Smart Maintenance, the dynamic learning, and knowledge-oriented maintenance, offers such a holistic, value-added oriented approach. A targeted optimization of the maintenance strategy is possible through automated evaluation of the assets and identification of the most critical ones based on company-specific criteria derived from the success factors of the company and considering all three management levels normative, strategic and operational. By considering the resource knowledge in the maintenance-strategy optimization process and using suitable methodologies of knowledge management based on the prevailing data quality for it, the efficiency and effectiveness are permanently guaranteed in the sense of continuous improvement.

Keywords

Criticality assessment; Knowledge Management; Lean Smart Maintenance; Digitalisation

1. Introduction

Today, digitalisation and industry 4.0 are propagated everywhere and these topics have a significant influence on industrial companies. A multitude of opportunities and new technologies of the digital transformation exists and offers new challenges, advantages and disadvantages for the asset management itself. Because of these reasons also the requirements for maintenance have changed considerably in recent years.

To be able to assert oneself in the automated production economy also in the future, a dynamical adaption of the maintenance strategy due to the changing environmental as well as production conditions is necessary [1].

DOI: https://doi.org/10.15488/9646
Therefore, it’s very important to include the life-cycle-aspect in the strategy adaption process, which is already anchored in international and national standards such as ISO 55000, DIN EN 13306 and DIN EN 16646 [1]-[4]. The overall goal is to increase the asset’s availability and to minimize maintenance costs over the entire life of an asset; to start with the investment phase up to the disposal phase [5].

A suitable tool for this purpose is the criticality assessment. It is based on the idea of the Lean Smart Maintenance (LSM) concept, which is a process and value-orientated model with the focus on the Lean part (cost-efficient and continuous adaption of the maintenance strategy to changing environmental conditions) and the Smart part (learning and knowledge-orientated approach). Especially the knowledge retention plays an essential part in the strategy adaption process. The fact that the criticality assessment is carried out as automatically as possible, based on the data quality of a company, and has a uniform and standardized documentation involved, makes decision making and strategy adaption as easy and comprehensible as possible [1].

This systematic approach makes it possible to ensure the required objectivity of a qualified risk assessment. Various tools are available for the quantitative consolidation of risk factors. Some of the best-known approaches are HAZOP (Hazard and Operability Analysis), FTA (Fault Tree Analysis), ETA (Event Tree Analysis) and FMEA (Failure Modes and Effects Analysis) [6], [7].

In this article, the HAZOP method is presented in combination with the criticality analysis based on the LSM philosophy. The first step of this criticality analysis, the uniform plant evaluation, lays the foundation for an asset-specific dynamic strategy optimization based on a targeted risk analysis. The standardized and, if possible, automated asset valuation identifies cost- and risk-critical assets without much effort by using a standardized procedure. This uniform documentation, which is part of the HAZOP methodology also lays the foundation for long-term knowledge assurance and prevents loss of knowledge. Once the cost-critical and risky assets have been identified, they are analysed in detail using the HAZOP approach, which is defined in the standards IEC 61882, ISO 31010 [8]. Measures for improving the asset are derived and evaluated in terms of cost and risk reduction, and the most profitable measures are implemented with the best cost/benefit ratio and the maintenance strategy is simulated and subsequently adopted based on the previous validation [9].

The aim of this criticality assessment, the combination of the HAZOP approach with the criticality analysis, is to optimize the efficiency and effectiveness of the assets over the entire life cycle in the long term by dynamically adapting the maintenance strategy [10].

2. Dynamic criticality assessment

The dynamic criticality assessment (Figure 1) is a standardized methodology, split into 4 phases (planning, preparatory-, implementation phase, and result evaluation) to identify critical assets and dynamically adapt their maintenance strategy. With this method information can be generated out of the data used for the identification of most critical assets and subsequently generate knowledge out of it because of the standardized documentation of the results. This knowledge can be secured in long-term basis to fulfil the company’s and maintenance goals permanently [11].
At the beginning of an asset criticality assessment in the planning phase, the project framework must be defined, which is also part of the HAZOP approach. In the first steps, a project or study manager is nominated, who is responsible for the execution of the study and subsequently controls the implementation of the defined measures based on the criticality evaluation. Besides, the project objective must be defined which, in the case of the asset evaluation, is the identification of cost-critical or risky assets. Building on this, a multidisciplinary team composed of experts from different departments such as maintenance, production, IT, finance, etc. is created to fulfil the requirements of the assessment. At the same time, the consistency of the plant master data, as well as the data quality should be analysed in detail, as this forms the basis for an automated asset valuation. [12].

During the preparatory-phase for the criticality analysis, facilitated workshops must be held to determine the evaluation level and to select and define criteria used for the asset evaluation, based on the success factors of a company. Part of the workshops is the splitting of the systems, the process or the procedure into smaller...
elements or sub-systems [8]. In the special case of an asset, the sub-elements are defined as asset, sub-asset, assembly and components [13]. In the best way, the asset structure is defined correctly in an ERP (Enterprise Resource Planning) system, where the data can be transferred in another system or a data file. Depending on a company’s data quality the asset evaluation can be either dynamic or static [1]. If the master data for the assets are formulated and checked, a criteria catalogue for the asset evaluation is generated in the next step [1].

2.1 Criteria-definition

One of the most important steps in the criticality assessment is the criteria definition. In this step, qualitative and/or quantitative criteria are chosen based on a company’s data quality. For the criteria selection, it is important not only to focus on the area of maintenance and its key figures but also to try to derive criteria from all three management levels (normative, operational and strategic) to obtain an overall picture of the company and the most important influencing factors [2], [14].

To identify the most important success factors of a company or an industry, a morphological box is used, which forms the basis for the criteria selection. For this catalogue, categories have been defined which have a direct influence on maintenance, e.g. the production, the asset, the digitalisation, the company’s characteristics as well as maintenance itself (Figure 2). Gradations were defined for each category, consisting of individual elements [15]. For each element, characteristics were specified, based on benchmark surveys, project acquisition and extensive literature research. Based on these fundamentals, a profile was created for each industry branch, which should already give the company a direction for the criteria definition workshop. Figure 2 represents an example of such a morphological box for the steel industry. Especially the degree of digitalisation must be queried on a company-specific basis, as it is fundamental for the asset evaluation as well as for the criteria used for it.

<table>
<thead>
<tr>
<th>Category</th>
<th>Characteristic</th>
<th>Markedness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company</td>
<td>Industry branch</td>
<td>Automotive</td>
</tr>
<tr>
<td></td>
<td>Company size</td>
<td>&gt;500 employees</td>
</tr>
<tr>
<td>Production</td>
<td>Production type</td>
<td>one-of-a-kind production</td>
</tr>
<tr>
<td>Maintenance</td>
<td>Degree of automation</td>
<td>no</td>
</tr>
<tr>
<td>Maintenance</td>
<td>Budgeting</td>
<td>static</td>
</tr>
<tr>
<td>Maintenance</td>
<td>Maintenance strategy</td>
<td>reactive</td>
</tr>
<tr>
<td>Asset</td>
<td>Average asset age</td>
<td>&gt;5 years</td>
</tr>
<tr>
<td>Asset</td>
<td>Asset availability</td>
<td>&gt;95%</td>
</tr>
<tr>
<td></td>
<td>Degree of utilisation</td>
<td>&gt;95%</td>
</tr>
<tr>
<td>Digitalisation</td>
<td>Degree of digitalisation</td>
<td>high</td>
</tr>
<tr>
<td>Digitalisation</td>
<td>Data quality</td>
<td>very good</td>
</tr>
</tbody>
</table>

Figure 2: Example of the morphological box for criteria definition

If the data quality is very good, a dynamic asset evaluation can be used. If the data quality is barely sufficient or bad or if no data exists, only a static asset evaluation must be applied. The consequence of this is that only qualitative criteria can be used for the evaluation, whereby the human influence and the assessment’s effort are higher and therefore a stricter approach for the evaluation must be used to reduce human failures to a minimum. Qualitative criteria can be described subjectively, because these are undetermined numerical values and can only be measured to a limited extent [16]. Quantitative criteria are numerical values and additional data from measurements are used, which in turn contribute to the production of information. In the best case, these criteria are drawn from the ERP system by real-time interfaces, which are used to
dynamically adapt the maintenance strategy and to reduce the human influence and the related faults (Figure 3) [2].

For the criteria selection a maturity assessment, based on a questionnaire is used. As a result of it, a company gets its maturity level for the criticality analysis as well as a list of criteria that are suitable for the asset evaluation. The higher a company’s data quality is, the less qualitative criteria are used. In this case, the assessment helps to replace necessary qualitative criteria by quantitative ones. As a result of the questionnaire response, a list of criteria that represent the success factors of a company as well as the biggest influencing factors on the asset’s availability is generated.

In the next step, a definition for each criterion as well as respective gradations (low, medium, high) to provide a uniform understanding of the criteria and the weighting of the criteria to complete the preparatory phase for the criticality analysis must be set during a workshop. For a better understanding, if one criterion is the availability of the asset, a definition of high, medium and low availability in terms of moderate numbers have to be considered. A numerical value is then stored for each of these states of expression, which, after a successful evaluation, delivers the asset index, the evaluation’s result [17]. Once the assessment preparations have been completed, the assets can be rated based on the previously defined criteria. In the best case, if the data quality is high, this step is carried out automatically [2].

### 2.2 Criticality-analysis

The criticality analysis includes three steps: the criteria evaluation, the creation of an asset portfolio for the identification of cost-critical or risky assets and the detailed asset analysis of critical ones [2]. In the first step of the asset evaluation, the plants are evaluated based on the criteria defined in the previous steps, cost and critical ones are identified and measures to reduce cost or/and risks are deduced [14].
Asset-evaluation

In the step of the asset evaluation, an integral part of the implementation phase of the HAZOP approach, the assets are rated based on the predefined criteria. In the best case, this step is carried out automatically. The result of the asset valuation is the asset index for each of the assets valued. The higher the asset index, the more critical an asset is classified [18].

Identification of most critical assets

For the visualisation of the most critical assets an asset priority portfolio is used, in which the asset index is compared to the direct maintenance costs (Figure 4) [2], [16]. The portfolio is divided into 4 quadrants. Depending on the situation of the assets in the portfolio detailed cost or/and risk analysis must be performed. Assets above the cost line are classified as cost critical. For these assets, cost reduction should be achieved with little to no risk increase. Assets to the right of the index line are classified as risky. The most critical assets are placed at the top right, as they are classified as critical in terms of both costs and risk. In contrast, the assets on the top left or bottom right have only one critical influencing factor, either cost or the risk. The aim should be to shift all critical assets to the left bottom, the uncritical quadrant, by deriving the appropriate maintenance measures and adjusting the maintenance strategy [11], [18].

![Figure 4: Example of an asset priority portfolio](modified 13)

The next step is to examine the assets identified as critical in detail and try to adjust the maintenance strategy by taking the defined measures into account for the risk and cost reduction, starting with those that have two critical influencing factors [18].

Cost- and risk analysis

In this step either cost or critical assets are analysed in detail by looking at each criterion and its deviation from the target value. Afterwards, the reasons for this deviation are analysed in detail and the consequences of such a deviation are considered. Therefore, all deviations, as well as current and potential risks of the
critical assets, are assessed based on three dimensions and their predefined gradations: the extent of damage, the occurrence frequency and the probability of detection (Table 1).

Table 1: Example of definition of keywords

<table>
<thead>
<tr>
<th>Extend of damage</th>
<th>Occurrence frequency</th>
<th>probability of detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>unlikely</td>
<td>very high</td>
</tr>
<tr>
<td>noticeable</td>
<td>rare</td>
<td>high</td>
</tr>
<tr>
<td>high</td>
<td>occasionally</td>
<td>low</td>
</tr>
<tr>
<td>critical</td>
<td>Several times</td>
<td>unlikely</td>
</tr>
<tr>
<td>catastrophic</td>
<td>frequently</td>
<td>highly unlikely</td>
</tr>
</tbody>
</table>

For the gradation of each dimension, a qualitative description has to be created and concrete examples are used to show its plausibility (Table 2).

Table 2: Example of a qualitative description of keywords

<table>
<thead>
<tr>
<th>Appraisal</th>
<th>Qualitative description</th>
<th>Example</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>very high</td>
<td>Errors are almost always detected in time</td>
<td>Sensor technology detects errors in the system/asset</td>
<td>1</td>
</tr>
<tr>
<td>low</td>
<td>Errors are rarely detected in time</td>
<td>Errors detected by a machine operator</td>
<td>3</td>
</tr>
</tbody>
</table>

In the next, the Risk Mode and Effect Analysis (RMEA) is used for the risk evaluation. As a result, the Risk Priority Number (RPN) is formed as a product of the previously defined three dimensions (extent of damage, occurrence frequency, probability of detection), which represents the degree of risk of the asset [18]-[20]. The aim is to reduce the risk per asset to a minimum and to increase the value-added contribution of the company by defining measures for its reduction [2], [16].

For the detailed cost analysis, the costs are divided into their cost types and the largest cost drivers are identified. For both analysis-types, cost or risk-reduction measures are derived and the savings potential is considered. This regular automated system evaluation as the basis for detailed analysis enables companies to react more quickly to deviations, derive measures and thus permanently reduce costs and increase the availability of the systems. [17].

The derived tasks form the basis for the maintenance strategy adjustment, which in the best case should be carried out dynamically in the same way as the criticality analysis.

2.3 Adaption of the maintenance strategy

Based on the detailed analysis, the asset-specific maintenance strategy mix can be adapted [1]. By combining different maintenance strategies, it is possible to obtain a balanced mix that determines the optimum for the company concerned.

The following maintenance strategies are used to create the optimum strategy mix: reactive (failure-oriented), preventive, predictive (condition-based) and perfective (asset improving) [1], [2], [18]. By optimizing the maintenance strategy mix, both technical and cost-intensive weak points as well as organisational ones, which have been identified in the analysis phase, are eliminated. By adapting the maintenance strategy, changes can occur in the areas of maintenance resources, which can be differentiated.
into personal, structural and relational capital [18]. These changes can include the elimination of inefficiencies through adaptation of the maintenance process, the introduction or expansion of autonomous maintenance, the introduction of new technologies like Big Data Analytics, Simulation, etc. Furthermore, the qualification of the employees is very important and is essential to invest in facilities or to replace some of the old ones. An optimization of the spare part management is also part of the changes, as well as changes in the allocation of external services and also in the structural and/or process organisation [1], [18].

A cost/benefit analysis checks whether the costs incurred by the affected asset are justified in relation to the benefit. If unfounded, highly preventive and/or reactive costs become apparent, either the underlying activities or the biggest cost drivers must be identified. Afterwards, organisational measures can be derived like e.g. adjustment of the maintenance and inspection intervals, outsourcing of activities, employee trainings [1], [2].

In the sense of digitalisation and industry 4.0 the qualification of employees will become more important. The complexity of the assets will increase and the automation, as well as the heterogeneity of the asset park will rise. As a result, the working environment and requirement profiles of maintenance personnel will change in the future. To prevent loss of knowledge and to live a learning and knowledge-oriented maintenance, which is part of the LSM aspect, early competence development is indispensable as well as standardized approaches.

In order to see how the strategy optimization affects the cost- and risk level a simulation by optimizing the asset data used for the evaluation can be done and the shift in the portfolio due to it can be made visible. The simulation is used for trend analysis and as a decision support tool for several adjustment possibilities and helps to identify the best asset specific strategy.

### 2.4 Performance review

In order to check whether the strategy adjustment has been successful and whether the simulation has made the correct prediction, it is important to perform the asset valuation continuously and check the shift of the assets in the portfolio. If there has been no improvement in the risk or cost-level of a critical asset, the assessment should be carried out one more time and the derived measures must be questioned and further measures derived in order to permanently improve the state of indictment [18], [17].

The fact that many of the steps in the criticality assessment can be carried out automatically based on good data quality means that the resources required for this are minimal. Furthermore, this standardized procedure and standardized documentation ensure traceability throughout the entire assessment.

### 3. Conclusion

The more complex and diverse assets are, the more important it is to have a standard to dynamically adapt the maintenance strategy, due to the changing environmental and production conditions and to secure knowledge by using a standardized approach and documentation. Therefore, the criticality analysis can on the one hand help to significantly reduce the effort and create a standard independent of the data quality and, on the other hand, reduce the human factors influencing such maintenance decisions and to create a more objective and comparable evaluation by automating the evaluation step of the assessment. The good thing about this approach is, that the higher the data quality the more automated it can be carried out and the fewer human failures will occur. Another advantage of good data quality is that the optimization of the strategy can be simulated and thus forms the basis for trend analysis.
Therefore, the criticality assessment, which is based on the LSM concept, offers a good opportunity to optimize the efficiency and effectiveness over the entire life cycle of an asset through targeted plant-specific cost- and risk-reduction and, accordingly to lower the cost duration and increase availability through continuous improvement.

References


Biography

**Theresa Passath (*1990)** is a research assistant in the department of Economic- and Business Management (WBW) at the Montanuniversitaet Leoben since 2017. Dipl.-Ing. Theresa Passath is also the deputy general manager of the ÖVIA since 2018. Her research focus is asset-, risk- and quality management.

**Cornelia Huber (*1983)** is a research assistant in the department of Economic- and Business Management (WBW) at the Montanuniversitaet Leoben since 2018. Her research focus is asset management, knowledge management and business informatics.

**Hubert Biedermann (*1953)** is professor at the chair of Economics- and Business Management and head of the department of Economic- and Business Management at Montanuniversitaet Leoben. His main research areas are asset management - lean smart maintenance, production management, quality- and risk management. After his habilitation in industrial economics in 1989, he worked for several years in the industry and as a visiting professor. Appointed Montanuniversitaet Leoben in 1995, he held the position as a vice rector for finance and controlling for more than 12 years.
Introduction of Traceability into the Continuous Improvement Process of SMEs

Nicolas Wittine¹, Sigrid Wenzel¹, Christian Kern², Robert Refflinghaus², Tim Trostmann²

¹University of Kassel; Departement of Organization of Production and Factory Planning
²University of Kassel; Departement of Quality and Process Management

Abstract

The digitization in the wake of Industry 4.0 offers small and medium-sized enterprises (SME) the opportunity to improve processes and products [1]. In this regard, gapless traceability represents a crucial element but is usually introduced by SMEs only due to extrinsic motivation [2]. Insufficient funding, lack of expertise and a poor market overview hinder implementation [3]. In order to improve realization, SMEs need to gain insight into the advantages offered by a traceability system [4]. Especially the potential regarding the usage of collected data within the continuous improvement process (CIP) provides the opportunity to implement product and process optimizations more effectively and efficiently. Consequently, this paper presents a concept, which shows how traceability can support and supplement the CIP. In this context, the granularity of information in a traceability system is relevant since the amount of data required for tracking and tracing a uniquely identifiable unit scales with the level of detail [5] [6].

The paper is structured as follows: After an introduction a summary of the state of the art comprising features of a traceability system, a definition of traceability granularity and commonly used Auto-ID systems is described. Section 3 matches the features of a traceability system with stages of the PDCA-cycle (Plan – Do – Check – Act) via waste sources and point out how the traceability system can be advantageous for each of its individual phases. How the granularity of traceability information influences the performance and the benefits of the CIP is demonstrated in Section 4. In addition, benefits of a traceability system in a production context are highlighted. Section 5 specifies the preferences of commonly used automatic identification systems and their typical use case regarding derivable traceability information in relation to the granularity of a system. Finally, future developments are discussed.

Keywords
Digitization; SME; PDCA; traceability; CIP; granularity; Auto-ID; case study

1. Introduction

The 4th industrial revolution offers a broad variety of digital tools to improve and enhance the production capabilities in the industry. The digitization of production and logistics chains and the implementation of a gapless traceability system are key in advanced data driven methods. Especially, the lacking implementation of traceability systems in small and medium sized enterprises (SME) entails the risk to lose the connection to key technologies and being in a disadvantageous situation in the global competition. Furthermore, it can lead to a disadvantageous position in the global competition. A conducted survey regarding the application of traceability systems in the industry shows that already 76% of the participants count themselves as users...
of a traceability system [2]. However, when the analysis is limited to SMEs, the application rate decreases to 33% which depicts a scenario that urges action. Furthermore, the survey provides reasoning behind the usage of such systems and shows that SMEs mainly implement a traceability system to fulfil a customer requirement or to obtain a certain certificate. Identifying faulty components or using the collected data to improve products or processes represent potential not utilized by the recipients of the survey. Concerns regarding the implementation of a traceability system result from the lack of qualified personnel, the non-transparent market situation for traceability systems and the technical challenge. [2] [7]

This paper aims to increase the application rate of traceability systems by demonstrating the benefits of implementing such a system beyond the sole fulfilling of customer requirements. Besides this extrinsic motivation, the intrinsic motivation for implementation needs to be stimulated. Demonstrating the benefits of using the data, that even a basic traceability system generates to support the continuous improvement process (CIP), leads to a better understanding and thus has the potential to increase the realization rate.

2. State of the art

The following subsections contain definitions and explanations of the relevant topics and principles this paper contains. Different aspects of traceability and the differentiation of tracking and tracing are explained in 2.1. In subsection 2.2 the term “granularity” in the context of traceability systems is defined and we concurrently offer a definition for future publications. The last subsection discusses the commonly used Auto-ID systems.

2.1 Traceability

The DIN EN ISO 9001:2015 [8] describes the requirements for quality management systems and gives a definition for the term traceability. Traceability in its core includes the identification of a unit or service and the availability of information connected to it throughout the whole production chain [8] (see also Figure 1).

![Figure 1: Core elements of a traceability system [9]](image)

Ideally, a state-of-the-art traceability system contains the gapless history of a continuously identifiable production unit. In complex assembly groups, each unique identifier of the assembled parts is registered and attached to the finished product. Furthermore, the unique identifier is used to connect quality data, production parameters or settings of assembly machines with the final product. Ultimately, the gathered information is communicated along the supply chain and made available to the value-adding partners. [9] [10]

Traceability is no unidirectional ability. The corresponding flow of information can be interrupted and divided into tracking and tracing. Tracking can be used to locate the destination of an item and may be used to inform the successor in the production chain in case of a previously undiscovered production error. Tracing allows finding the origin of an inquired item as well as to identify the root cause regarding quality
issues or reclamations (see Figure 2). This allows fast responses to deviations in product quality and quick reporting to the relevant recipient. Moreover, preemptive measures can be determined in order to avoid subsequent reclamations. [9]

Figure 2: Traceability information exchange in case of an interruption in the production chain, according to [7]

Especially when combining quality and production parameters complex interdependencies of a system can be uncovered. Therefore, in the event of a quality issue outside of the influence of a specific actor in the supply chain, the size of recalls or compensation payments can be significantly reduced. Especially for the often used 8D reports it becomes easier to formulate an effective immediate action and to provide the root cause by reviewing the gathered production and quality data. [11]

The need to implement a traceability system can arise unexpectedly, e. g. when national laws change to protect customers. Thus, the food sector became closely monitored after food scares like bovine spongiform encephalopathy (BSE) or dioxin threatened consumer safety. Another motivation to use traceability systems is to avoid high compensation payments in case of product reclamations. [10] [11]

2.2 Granularity

Granularity is a broad term and is used in many different contexts (e. g. quantity of items inside a traceable unit, the accuracy of data, the level of detail in a supply chain, and software engineering [12] [13]) and therefore it is necessary to define the term. In the context of traceability systems, the term traceability granularity describes the level of detail concerning the gathered information for a uniquely identifiable unit or tracking unit [6]. In addition, it is important to point out that a tracking unit may not be one single item but may be a batch or bulk cargo which is unambiguously distinct regarding the connected information. The higher the level of the traceability granularity, the more specific data is acquired and needs to be managed. Furthermore, the requirements related to the identification of each level changes. [13]

For a detailed description of a traceability system the terms breadth, depth, precision and access are used. Breadth is used to quantify the amount of information and therefore data a traceability system has to record. Depth describes how far tracking and tracing is possible in a product life cycle. Precision measures the certainty of the location or the characteristics of a unit. The speed at which information is communicated to the stakeholders is described by access. [7]

To fully understand granularity in the context of a traceability system it is important to understand the differentiation between the amount of data and their quality. A traceability system with a high level of granularity not only has breadth but also precision. The data available is highly specific and connected to only few if not even a single distinct unit. Vice versa a traceability system with a low granularity level connects fewer data to more units, like batches or bulk cargo (see Figure 3).
2.3 Auto-ID

Auto-ID systems are used to mark and identify a track and traceable unit. The key component of those systems is a unique identifier for each marked item. The previously mentioned traceability granularity determines the quantity inside a unit. [14]

The most common technology worldwide is the 1D-barcode. Within this system, every item in a store is identifiable by its product specific code. An evolutionary advancement is the 2D-barcode in form of quick response codes (QR-Code) or Datamatrix-Codes (DMC) which can store additional information. Thus, these codes can be used to not only identify a general product but also apply a unique serial number containing batch information or machine data. The reading of the labels is done by scanners or camera systems which need line of sight and enough contrast. [14] Marking via laser, printer, engraving or applying a label are common ways of applying the code to the item [15] [16].

Radio frequency identification (RFID) can store large amounts of data and can be read across long distances with a high speed. Depending on the design three types can be differentiated. Passive RFID labels consist of a chip containing the stored data and an antenna for transmission. Rewriting data is possible but in terms of industrial application, they are mostly read only. Apart from that, they need the energy transmitted by the readers signal in order to receive or send information. Active RIFD labels have an additional power source and can emit a long-range signal with just a trigger signal from a nearby reader. Semi-passive RFID systems also use a power source but lack an own antenna. Transmitting data works only by modulating the back scatter. [17]

The magnetic identification uses magnetic stripes attached to the traceable unit and carries a unique number to identify the item. A reader needs to be brought close to the magnetic surface to pick up the contained data. [14] Biometric identification or fingerprint technology uses randomly formed aspects of an item for an unambiguous identification. The tagging of an item is not needed but the reader needs to be configured for the specific item in order to identify it reliably. [18]

3. Mapping of Traceability and CIP

The CIP as well as traceability systems are both customer-driven [19]. To gain a better understanding of the benefits of using a traceability system beyond the mere fulfillment of customer or legal requirements the added utility of such a system needs to be emphasized. One way to use traceability data from products or
production facilities is to integrate the traceability system into CIP. The CIP, as part of a lean approach, aims to reduce waste in a production system by utilizing the PDCA-cycle. [20] In this context, a functioning traceability system is a key factor for the identification and elimination of the main factors influencing the generation of waste.

The most expensive waste in production is a defective product or scrapping of a finished product [21] [22] [23]. In some cases, rework is feasible and recovers some of the value which otherwise would be lost [19] [21] [23] [24]. The origin of faulty products is often hard to find and can be a result of design discrepancies or inadequate measurements to uncover quality issues. An incoherent database filled with redundant, non-uniform or even useless data can be a result from a lack of deep understanding of the processes or products involved [22] [23]. Over processing in form of converting or formatting data due to different systems used or a non-standardized interface for data exchange [21] [22]. This results in needless movement of information or redundant staff meetings [21] [22] [23]. Furthermore, a lack of convenient access to information wastes resources via transporting the information to the corresponding recipient [22] [23]. An additional issue of receiving and delivering information are communication barriers; whether cultural, physical or digital, they hinder or prevent the direct transfer of valuable knowledge and important information [21] [25]. Thus, inconvenient access to data combined with tedious formatting leads to extensive waiting times when providing relevant information to decision makers and participants in the production chain [19] [21] [22] [23] [25].

Other types of waste do exist, but a traceability system does not aid in their elimination. Since its features do not provide the necessary information as they are situated in a different organizational section in a company. Those are overproduction, non-standardized processes, outdated technology, useless information, lacking customer orientation and unused resources in development or scale effects [19] [21] [22] [23] [25].

The eliminable types of waste can each be assigned to a phase of the PDCA-cycle which in turn can be mapped with a core element of traceability systems. Consequently, the Plan-phase consists of checking the examined system for wasted resources, like defective or reworked products. The phase greatly benefits from the capability to identify parts unambiguously which assists in defining measures and identifying key performance indicators. Regarding the Do-Phase where the measures are implemented a proper recording of the identified key parameters is essential and a first impression of the effectiveness of the measure can be gained. A coherent database avoids over processing of data and assists in locating the root cause. In case of a severe problem or a customer reclamation the collected data from the traceability system can be used for an immediate response. Needless movements of data, information or employees obstruct the Check-phase and it becomes challenging to ensure whether the taken measures are robust. Being able to link the performance indicators to each identifiable unit allows to evaluate the actual outcome. In the Act-phase an optimized standard is implemented and needs to be communicated among all stakeholders across potential communication barriers. Especially customer reporting greatly benefits from the reduced waiting times provided by a traceability system. In addition, the change is easily monitored continuously and tracks the long-term effectiveness of the measures and their sustainability.

To visualize the previously described method the known PDCA-cycle is extended by the mentioned core elements of a traceability system as shown in Figure 4. After the successful completion of the PDCA-cycle a new cycle can be started. The gained knowledge can act as a foundation for the next iteration by identifying improvable potential in the product life cycle.
4. Benefits of using a traceability system

This section is dedicated to highlight the benefits of using a traceability system via a few published use cases. Furthermore, it is outlined how different approaches can lead to different results. Special attention is given to the traceability granularity and the associated expenses to implement and operate a traceability system.

In general, not only the location where the tracking and tracing happens but especially the entrances and exits of closed systems like a factory or a chained production system are important aspects for the introduction of a traceability system [5] [26]. In order to select the best traceability system for a given task, investment and expected results, simulation can assist in comparing different solutions [27]. Resources spent on more granular data in traceability systems will be wasted if the return is not a more precise traceability [5]. Therefore, the technology and expected results regarding the introduction of a traceability system need to be evaluated. In this context, not only the investment for the system is relevant but also the added value for the customer can act as a unique selling point and increase sales. In order to estimate the value of implementing a traceability system, the failure mode and effect analysis can be used as a guidance to assess risks and responsibilities regarding faulty products. Furthermore, it has to be clarified that even a traceability system designed with a financially restricted budget can still be beneficial to the company itself and increase customer’s satisfaction [27]. With the introduction of a traceability system, the production chain becomes transparent and data of all relevant processing factors (e.g. order number, machine and tooling used, assigned personnel or quality measurements) can be tied to the finished product [10].

The food sector in the US offers examples regarding fresh produce, grain and livestock. Producers with a low packaging size transition easily to a higher level of granularity. When handling grain the elevator as a bottleneck serves as a batch size and allows traceability on a satisfying level. Regarding cattle, traceability systems allow for supply management and production planning as well as securing food safety. [28]

The Arizona State University demonstrated that the PDCA-cycle can be used as a tool to reduce energy consumption of the campus. However, to effectively plan an improvement they had to monitor and measure energy usage data for each of the buildings on the campus. Even though the available data had low level
granularity it was enough to start the PDCA-cycle and begin with the improvement process [29]. Another example for the usage of low level and fuzzy data in combination with statistic modelling of product flow was showcased in factories processing fruits as well as dairy [30]. Especially when bulk cargo like apples in large bins is involved, changing packaging and reducing batch size can lead to better traceability even with documentation gaps in the product flow [5] [31].

Another example comes from a production line in an automotive plant where the quality team was unable to detect the root cause of a quality issue even though a traceability system was implemented. The result were extensive reworks of each suspected batch. A temporary increase of traceability granularity in the context of a PDCA-cycle was able to help identify the root cause by linking the parts with machine data and quality measurements. After the successful identification a new standard for the production was formulated. Additionally, the gained insight into the interrelations of different parameters offered new approaches for further improvements and can be used as a foundation for the start of the next PDCA-cycle. [32]

The above-mentioned cases demonstrate how traceability systems can improve product quality, assist in driving change in a production environment, serve as a foundation for the CIP and increase customer satisfaction. A functioning traceability system even with low level granularity can already help identify sources of waste, substantiate business decisions or serve as a database for simulations. The improved product quality and deeper process knowledge are beneficial for all parts of the value chain and can create an advantage in global competition. Since the presented use cases are insufficient for a general approach, they at least show the potential of implementing a traceability system. Furthermore, additional case studies could lead to a conclusive analysis of the benefits.

5. Auto-ID usage

The previous sections have demonstrated the benefits of using a traceability system and being able to uniquely identify a unit. Therefore, the use of Auto-ID systems becomes almost obligatory. Whether this unit is a single part, a whole batch or bulk cargo needs to be decided by the stakeholders of the product life cycle. In order to allow SMEs to focus their resources on selecting only the relevant systems available on the market, a small survey is conducted to narrow down the selection. The survey illustrates tendencies for common use cases of Auto-ID solutions regarding the identification of single products, batches or bulk cargo.

The interview partners are sales experts of different sensory system suppliers currently active in the market for traceability systems. The combined revenue of the companies in question surpasses €7bn. The first task in the survey is a paired comparison of the four Auto-ID systems mentioned in 2.3. Each system is compared with each other and the one more important gets a point. The cumulated result in form of a total score is illustrated in Figure 5. The experts emphasize that generalization in this regard is only possible to a very limited scope and highly dependent on a specific application. Nevertheless, the results allow a tendential assessment regarding the significance of the different ID systems. Especially the frequently predicted prevalence of RFID as the main ID system [18] [33] may not be true and should be further investigated via a wider survey. In the context of this survey the 1D / 2D-Barcode or QR-Code systems are most likely used for Auto-ID purposes. This does not mean that RFID is irrelevant since it also scores high on the 2nd place and should be considered when selecting a traceability system. Regarding biometric ID systems the experts also see a huge potential for identification purposes when mass marketability is achieved. The lowest score is awarded for magnetic ID systems and may have a use in certain applications.

Related to the common use case of each Auto-ID solution, the five interview partners are asked to rank the Auto-ID systems by their favorable application regarding traceability (see Figure 6). For each given use case the achievable score for the most relevant technology would be four points whereas one point would be awarded for the least relevant. For the tracking of single parts, the barcode is the preferred technology.
followed by RFID. The same applies to the traceability of batches. Regarding tracking and tracing larger quantities of units, the bulk case points out RFID technology as the preferred system of identification.

Conclusively, a SME interested in implementing a traceability system needs to first check which level of traceability granularity is needed. Subsequently, it must be further investigated whether barcodes, RFID solutions or a combination of both should be applied. Additionally, restrictions like reflective surfaces, electrical shielding or the involved processes need to be considered when deciding for a traceability system on a hardware level. Barcodes and scanners seem to provide a flexible solution in terms of application and achievable granularity. Thus, represent an ideal foundation for SMEs intending to gain first-hand experience.

6. Summary and outlook

The motivation behind this paper is to show the potential users of traceability systems the benefits of implementation beyond the fulfillment of external requirements. In this context the state of the art for the most important aspects and components is outlined, comprising the core of traceability systems, a definition of traceability granularity and the most relevant Auto-ID systems.

In order to motivate members of a supply chain to introduce a traceability system, the internal benefits are highlighted by combining the PDCA-cycle with the capabilities of a traceability system. This allows companies to eliminate waste sources identified by data collected via a traceability system. For a more convenient transition from a theoretical approach to practical implementation, use cases are presented that clarify the strategies for using a traceability system to improve processes.

Since the introduction of a traceability system requires specific traceability hardware, this paper further offers some guidance for choosing relevant technologies. The five experts from companies operating in the field of sensory hardware do not confirm the often-proclaimed widespread application of traceability systems via RFID. Rather, barcode technologies like QR-codes are currently the technology of choice. Nevertheless, each specific use case needs to be evaluated regarding its necessary traceability granularity. Tracking and tracing a single item is associated with a large investment, both financially and in terms of knowledge required, but also offers the most information. However, compared to bulk level traceability, realizing a traceability system for batches can already offer insights into processes and may act as a proof of concept for further traceability endeavors. Considering the limited scope of the survey further research is needed.

Especially SMEs are at risk of falling behind when it comes to the industrial developments in the wake of the 4th industrial revolution. Therefore, future research should seek to further lower the hurdles for implementations of traceability systems by emphasizing internal benefits and promoting a change within the companies. Even though the hardware aspect of traceability systems was briefly covered, each individual case must to be treated appropriately. Furthermore, the use of a suitable reference model would be advisable for planning purposes. The software side of traceability is missing in this paper because a first proof of
concept or inhouse test should be kept at a manageable scope. For a holistic traceability approach including the whole production chain the used hardware and software must be considered in accordance with the concrete application.

Acknowledgements

The results presented in this paper originated in the joint research project "Entwicklung einer Methodik zur Auswahl und Einbindung von Traceability Systemen in den kontinuierlichen Verbesserungsprozess von KMU" of the Department of Quality and Process Management as well as the Department of Organization of Production and Factory Planning at the University of Kassel. The Industrial Collective Research (IGF) Operation 20205 of the FQS-Forschungsgemeinschaft Qualität e.V. was funded via the Industrial Research Associations (AiF) under the program for the promotion of the IGF by the German Federal Ministry for Economic Affairs and Energy (BMWi) based on a resolution by the German Bundestag.

References


**Biography**

Nicolas Wittine has been a research assistant at the Department of Organization of Production and Factory Planning at the University of Kassel since 2019. After obtaining his Master of Science in 2016, he worked as a development engineer in the automotive industry and gained first-hand experience in the development of safety-critical components, project management and industrialization. Currently, his work focuses on traceability and production planning.

Sigrid Wenzel is a professor and head of the Department of Organization of Production and Factory Planning, University of Kassel. In addition to this, she is a board director of the Arbeitsgemeinschaft Simulation (ASIM), spokesperson for the ASIM Section Simulation in Production and Logistics, member of the advisory board of the Association of German Engineers Society of Production and Logistics (VDI-GPL), and head of the Committee Modelling and Simulation of the VDI-GPL.

Christian Kern has been a research assistant at the Institute for Research and Transfer in Dortmund at the Department of Quality Management since 2009. Since 2019 he has additionally been employed as a research assistant at the University of Kassel in the Department of Quality and Process Management. His work focuses on the areas of process optimization, reliability analysis and requirements engineering.

Robert Refflinghaus has been professor and head of the Department of Quality and Process Management at the University of Kassel since 2012. His research focuses on process optimization as well as methods and techniques of quality management. In addition, he is a member of the board of the German Quality Function Deployment Institute.

Tim Trostmann has been a research assistant at the University of Kassel at the Department of Quality and Process Management since 2017. He completed his master's degree in 2016 and gained practical experience in the automotive industry, the implementation of Six Sigma projects and supplier development. Currently, his work focuses on the areas of process optimization, reliability analysis and workplace ergonomics.
1st Conference on Production Systems and Logistics

Data-based identification of knowledge transfer needs in global production networks

Günther Schuh¹, Jan-Philipp Prote¹, Andreas Gütlaff¹, Katharina Thomas¹, Moritz Johannes Volk¹

¹Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, 52074 Aachen, Germany

Abstract

Manufacturing companies’ value chains are increasingly distributed globally, which presents companies with the challenge of coordinating complex production networks. In general, these production networks grew historically rather than having been continuously planned, leading to heterogeneous production structures with many tangible and intangible flows to be coordinated. Thereby, many authors claim that the knowledge flow is one of the most important flows and the source of competitive advantage. However, today’s managers face major challenges in transferring production knowledge, especially across globally distributed production sites. The first obstacle to a successful knowledge transfer is to identify what kind of knowledge should be transferred between whom and at what time. This process can take months of information collection and evaluation and is often too time-consuming and costly. Thus, this paper presents an approach to automatically identify at what point knowledge should be transferred. In order to achieve this, the company's raw data is being used to identify which employees work on similar production processes and how these processes perform. Therefore, production processes, which can be compared with each other, need to be formed, even though these processes may be performed at different production sites. Still, not every defined cluster of production processes necessarily requires the initiation of knowledge transfer since performing a knowledge transfer always entails considerable effort and some processes might already be aligned with each other. Consequently, in a next step it is analyzed how these comparable production processes differ from each other by taking into account their performances by means of feedback data. As a result, trigger points for knowledge transfer initiation can be determined.

Keywords

Knowledge transfer; Production network; Network coordination

1. Introduction

In the last decades, production has spread globally and value chains are becoming more and more fragmented leading to growing and complex production networks [1]. One consequence of this development is the often rapid and partially unstructured growth of the globally distributed production sites. The coordination and synchronization of these sites among each other is a major challenge. [2] As a result, global production networks are considered among the most complex man-made systems with many tangible and intangible flows to be coordinated [3]. One of the most important flows is probably the knowledge flow, yet very difficult to manage [4]. Many authors agree that the cross-site knowledge transfer of methods and best practices is a major challenge and an unsolved problem in practice [4,5]. However, at the same time the
transfer of globally distributed knowledge in a production network is considered as a significant competitive advantage [6–9]. Knowledge transfer enables the alignment of performance differences between existing processes and supports employees to learn from each other by bringing them together in a context-specific way [7].

In general, knowledge transfer is defined as a process in which one organizational unit passes experiences, information and skills to another [10,11]. Thus, the elements of a knowledge transfer are the sender, the receiver, the transferred knowledge and the organizational context as for the relationship between sender and receiver (cf. Figure 1) [12]. In terms of knowledge transfer in production networks, production knowledge is particularly important, including manufacturing technologies as well as the operational knowledge as a recipe for action [4,5].

![Diagram of knowledge transfer elements](image)

Figure 1: Elements of knowledge transfer [12]

Considering these different elements of knowledge transfer and the complex structure of production networks, an almost infinite number of potential knowledge transfer opportunities exists. Hence, it is difficult to identify where knowledge transfer is necessary in production networks and how the elements of a transfer need to be defined. Some companies try to address this challenge by operating a knowledge data bank like a knowledge “yellow pages”. However, such knowledge data banks serves rather as provisioning for operational information than connecting the right employees for an actual transfer of production know-how. [5] In order to address this problem, this paper aims to present an approach to systematically identify the need for knowledge transfer with the corresponding transfer participants within a production network.

2. State of the Art

In recent years, many researchers have increasingly focused on the topic of knowledge transfer in inter-organizational networks. Thereby, a lot of scientific work is based on the work of TSAI from 2001. [13] TSAI studies how the network position as well as the absorptive capacity effect the ability to create and capture knowledge. He argues that social interaction can foster knowledge transfer across business units. For stressing out his statement, he analyzed data from two companies with a total of 60 business units with the result that by obtaining a central network position, an organizational unit can benefit more from knowledge transfer and sharing. With his work, TSAI presents first impulses to strengthen knowledge transfer. [9] However, a more detailed approach with a guidance for improving knowledge transfer in production networks is delivered by FERDOWS. He presents a framework for choosing the appropriate transfer mechanism depending on the type of knowledge. Therefore, FERDOWS differentiates between tacit and codified as well as slow and fast changing knowledge, leading to four different knowledge transfer mechanisms. Thus, he distinguishes between transferring production know-how via manuals and systems, joint developments, projects and moving people. [5]
CHENG et al. go one step further by not only discussing how knowledge should be transferred, but also where and when the transfer should take place. In this context, they propose a time-place matrix for coordinating knowledge transfer. The authors point out that, for an efficient knowledge transfer, the right sequence between know-where, know-which, know-when and know-how is important. Hence, CHENG et al. point out that, before performing a knowledge transfer, the transfer initiation with defining the knowledge that should be transferred and the transfer participants is crucial. [14] Based on the approach of CHENG et al., FRIEDLI et al. develop a framework for managing knowledge flows with an orientation guide to enabling knowledge transfer. They discuss as well what kind of knowledge should be transferred between whom, at what time and with what kind of transfer mechanism. Furthermore, FRIEDLI et al. analyze different exchange structures and the degree of transparency for improving knowledge transfer. [7]

As described above, some approaches exist on the topic of knowledge transfer in production networks. Most research has been done in developing frameworks for coordinating knowledge transfer and defining the correct transfer mechanism. In addition, CHENG et al. and FRIEDLI et al. point out that determining the right time and the right transfer participants is important for an efficient knowledge transfer. Still, an approach to systematically identify knowledge transfer needs and the required transfer participants is missing. At the same time, researchers such as SZULANSKI point out that the procedure of identifying knowledge transfer needs could take months of information collection and evaluation [15]. This is especially critical with regards to global production networks with a wide range and distribution of different production processes and experts. Thus, a high number of potential knowledge transfer opportunities exist. In this context, LEYER et al. claim that connecting employees for a knowledge transfer should be based on indicating their process-related areas of expertise [16]. Consequently, determining what production processes are comparable within a production network serves for identifying which employees work in similar fields and could potentially learn from each other. Therefore, approaches for analyzing similarities of production processes are relevant for this topic, although these approaches are not directly categorized in the research area of knowledge transfer.

For example, approaches that identify comparable production processes to improve production planning and manufacturing process design can be used for orientation. Therefore, LI et al. use publicly available and general information about manufacturing processes and technologies to identify similarities based on the process capabilities such as achievable tolerances and machinable materials. A pairwise comparison of each process based on its capabilities serves as a basis for the subsequent application of a hierarchical cluster algorithm and graphical evaluation. The systematic approach and the data-based analysis of process characteristics and capabilities serve as a good orientation for identifying comparable production processes. [17] In a similar way, AHN AND CHANG and ZHANG et al. analyze production processes based on two aspects: the attributes of involved machines and other resources as well as the order of process steps. AHN AND CHANG focus on using graphical modelling methods like the BPMN (business process model and notation) to compare the processes based on a standardized metric [18]. Next to that, ZHANG et al.’s approach is based on the description and subsequent comparison of production processes using process graphs [19]. The individual units of such a process graph, which are displayed in the processing sequence, contain information about the type of operation, the position of the individual process in the process chain and characteristics of the operation. Hence, the comparison of the sequence and properties of the individual process units gives a detailed picture of the composition and comparability of production processes.

Another approach for analyzing comparable production processes is to identify similar products. In this field, LENZ et al. and BRUNO analyze the manufacturing processes involved in the various products in order to identify comparable products. BRUNO uses a manufacturing process ontology and identifies similar products by comparing the manufacturing technologies that are used for the products. [20] LENZ et al. go into more detail by focusing on the analysis of the data from CNC machines with the G-code containing the various
attributes of machine movements. Thus, the G-code serves as the basis for the analysis of the manufacturing processes and subsequent identification of similar products. [21]

Examining these approaches shows that analyzing product characteristics as well as resource characteristics can be used to identify comparable production processes. Furthermore, data-driven analysis supports this process and helps to automatically identify comparable production processes. This can be helpful to identify knowledge transfer needs based on comparable production processes, especially considering the high number of different potential knowledge transfer opportunities in production networks. However, data-driven approaches are currently missing in the research area of knowledge transfer in global production networks and should be focused more intensively in the future.

3. Approach

Based on the analyzed requirements for the identification of knowledge transfer needs and the examination of existing approaches in this field of research, a new approach is presented in Figure 2. This approach is divided into three steps beginning with the characterization of production processes in production networks, followed by the identification of comparable production processes and the identification of knowledge transfer needs.

![Diagram](image)

In the first step of the presented approach, production processes are characterized in order to build a framework to automatically identify comparable production processes out of the diversity of production processes in global production networks. As analyzed in chapter 2, the description of production processes can be based on product and resource characteristics. This is consistent with the framework according to STEINWASSER who describes production processes as a combination of product and resource factors [22]. By distinguishing between these two areas, there is a clear separation between different production processes with simultaneously low complexity for further analysis. For defining both product and resource factors, a differentiation between different aggregation levels is necessary, since characteristic features differ depending on the structural level. At the product level, a distinction can be made between final products,
assemblies and components [23]. The characteristic features used to describe a final product are for example the product structure, the product range and the production scale in terms of batch sizes (cf. Figure 3). In contrast, on the component level relevant characteristic features are for example the raw part geometry, the material used and requested tolerances. At the resource level, in a production network a distinction can be made between the factory, the production line and the workstation at the lowest level [24]. At the factory level, information on the factory type, the logistics structure and the number of workstations and machines are relevant. Whereas at the workstation level, other characteristic features, such as the manufacturing technology and the specification of tools are focused. Depending on the structural level of the product and resource type, the level of the production processes differs with focusing either on the value stream analyzing the process flow between different process steps or rather the technology within a process step.

<table>
<thead>
<tr>
<th>Product characteristics</th>
<th>Description</th>
<th>Scale level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural level</td>
<td>Bill of material level of the product</td>
<td>Metric scale</td>
</tr>
<tr>
<td>Strategic positioning</td>
<td>Amount of the profit margin</td>
<td>Metric scale</td>
</tr>
<tr>
<td>Production costs</td>
<td>Costs for manufacturing a product</td>
<td>Metric scale</td>
</tr>
<tr>
<td>Direct production cost share</td>
<td>Percentage of costs of goods manufactured</td>
<td>Metric scale</td>
</tr>
<tr>
<td>Material cost share</td>
<td>Percentage of costs of goods manufactured</td>
<td>Metric scale</td>
</tr>
<tr>
<td>Product structure</td>
<td>Number of structural levels and items</td>
<td>Metric scale</td>
</tr>
<tr>
<td>Product range</td>
<td>Degree of standardization of product design</td>
<td>Ordinal scale</td>
</tr>
<tr>
<td>Production scale</td>
<td>Lot size in relation to repeatability</td>
<td>Metric scale</td>
</tr>
<tr>
<td>Stocking strategy</td>
<td>Level of the stockholding requirement item</td>
<td>Ordinal scale</td>
</tr>
</tbody>
</table>

Figure 3: Exemplary product characteristics on the level of a final product

In order to characterize production processes as described above, different data is needed to automatically identify comparable production processes in the following. Therefore, an UML data model is displayed in Figure 4 summarizing the information needs from different data sources. The product and resource information are typically stored in different information systems such as an ERP system (enterprise resource planning) or MDC system (machine data collection). However, these different kinds of data are typically connected with each other in a work plan for a specific process step. Thus, the work plan is the description of the process steps with an operation description and an associated process owner. This information primarily serves to identify comparable production processes and to identify which employees are working on these processes. Next to this information, performance characteristics by means of production feedback data are important in order to assess the production processes and to identify deviations. This information is needed to evaluate how comparable production processes differ from each other and in what dimensions employees should learn from each other. Feedback data is typically connected to a specific work order and in this way to a specific product and resource type.
Figure 4: Information needs to characterize production processes

After characterizing the different production processes within a production network, the second step of the presented approach is the application of clustering and similarity analysis based on the described data model to identify comparable production processes. When selecting suitable data-based analysis methods, several criteria must be taken into account. On the one hand, the analysis of large amounts of data should be carried out as quickly and largely automatically as possible, so that manual steps for the user are only necessary for a qualitative evaluation of the results. In addition, it must be possible to analyze both metrically and non-metrically scaled values since some product and resource characteristics are categorized in a nominal or ordinal scale (cf. Figure 3). Furthermore, the methods used should already have been tested in practice and thus be able to show a high user acceptance. Considering these prerequisites, the choice for the analysis of metrically scaled characteristics is a cluster analysis. The affiliation to a cluster is essentially determined by the distance of the characteristics between the individual objects. Objects that can be assigned to a cluster should therefore have a small distance to each other. This describes the homogeneity within the clusters. On the other hand, the different clusters should have objects with a high distance to each other, which describes the heterogeneity between the clusters. [25] Next to the cluster analysis, the analysis of comparability of non-metrically product and resource characteristics is determined using a similarity analysis. The pairwise calculated similarity values of two processes are the basis for the decision and are compared with a limit value defined by the user. If the similarity value is above the specified limit value, it can be assumed that the two processes are comparable. [26] The final decision on the comparability of the identified processes is made by the user within the framework of a qualitative evaluation of the results and a plausibility check.

The third step of the approach is finally the identification of knowledge transfer needs based on the determined cluster of comparable production processes in a production network. A cluster of comparable production processes does not automatically lead to a knowledge transfer need since these processes might already be aligned with each other. A knowledge transfer always entails considerable effort, which is why it should only be performed if a need exists. Hence, production feedback data to the production processes can be used to analyze the process performance as described in the data model in Figure 4. For automatically identifying process deviations, statistical process control (SPC) can be used. This is a widely used method for monitoring processes and detecting deviations. For this purpose, statistical control charts are the tools to implement SPC by systematically analyzing the output of processes. Therefore, upper and lower control limits (UCL, LCL) are defined depending on the mean value of the performances within a process cluster.
For the purpose of process monitoring in production networks, an adaptive control chart is necessary since the design parameters need to vary over time. Thus, the width of the control limits needs to be adjusted company-specific depending on the processes’ sensitivity. As a result of the SPC, process deviations within a cluster of comparable production processes can be determined and the upper and lower control limits can be used as trigger points for a knowledge transfer need.

4. Discussion and future research

In this paper, the main challenges for knowledge transfer in production networks and a three-step approach to systematically identify knowledge transfer needs are presented. Due to the high variety and amount of knowledge transfer opportunities in production networks, a data-based approach is chosen. In the first step, a framework has been developed to determine which production processes can be compared with each other. This step is necessary to identify which employees work on comparable production processes where a potential knowledge transfer need exists. For this purpose, production processes are described as a combination of product and resource factors with varying characteristics. In this context, an UML data model has been developed to show the information needs to characterize production processes. Based on this data model, the second step describes the application of clustering and similarity analysis to automatically identify clusters of comparable production processes. The third step of the presented approach focuses on the identification of knowledge transfer needs based on production feedback data. Therefore, statistical process control can be used by analyzing deviations within a cluster of comparable production processes and applying upper and lower control limits as trigger points for a knowledge transfer initiation.

Further research is required by detailing the needed feedback data for the statistical process control. Depending on the aggregation level of the cluster of comparable production processes, the target system for these production processes differs. Consequently, an adaptive target system with the corresponding production feedback data is needed. Moreover, further research aims to validate the developed approach and ensure its applicability in practices.

Acknowledgment

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References


Biography

Prof. Dr.-Ing. Dipl.-Wirt. Ing. Günther Schuh (*1958) holds the Chair of Production Systems at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University, is a member of the board of directors of the Fraunhofer Institute for Production Technology IPT and director of the Research Institute for Rationalization (FIR) at the RWTH Aachen. He is the founder of Schuh & Co. Group and the e.GO Mobile AG. Prof. Schuh is a member of several supervisory boards and boards of directors.

Dr.-Ing. Jan-Philipp Prote (*1986) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen and Industrial Engineering at the Tsinghua University in Beijing. He is Chief Engineer and Head of the Production Management Department at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University.

Andreas Gützlaff (*1989) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University. He is Chief Engineer and Head of the Production Management Department at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University.

Katharina Thomas (*1991) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University. She is a Research Assistant at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University and Group Leader of the Global Production group in the Production Management department.

Moritz Johannes Volk (*1994) studies Business Administration and Engineering specializing in Mechanical Engineering the RWTH Aachen. He is a Student Research Assistant at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University in the Global Production group in the Production Management department.
Evolution of a Lean Smart Maintenance Maturity Model towards the new Age of Industry 4.0

Oliver Schmiedbauer¹, Hans Thomas Maier¹, Hubert Biedermann¹
¹Chair of Economic- and Business Management, Montanuniversitaet Leoben, Leoben, Austria

Abstract

Over the last few years, the complexity of asset and maintenance management of industrial plants and machinery in the producing industry has risen due to higher competition and volatile environments. Smart factories, Internet of Things (IoT) and the underlying digitisation of a significant number of processes are changing the way we have to think and work in terms of asset management. Existing Lean Smart Maintenance (LSM) philosophy, which focuses on the cost-efficient (lean) and the learning organisation (smart) perspectives enables a value-oriented, dynamic, and smart maintenance/asset management. The associated LSM maturity model is the evaluation tool that contains the normative, strategic, and operational aspects of industrial asset management, based on which numerous reorganisation projects have already been carried out in industrial companies. However, due to the ever-increasing development of Industry 4.0 (I4.0), it is necessary to extend the model by selected aspects of digitisation and digitalisation. Based on a structured literature review (SLR) of state of the art I4.0 maturity models, we were able to investigate the essential maturity items for I4.0. To restructure and expand the existing LSM maturity model, the principle of design science research (DSR) was used. The architecture of the LSM maturity model was based on the structure of the Capability Maturity Model Integration (CMMI). Further development of a Lean Smart Maintenance maturity model thus covers the future requirements of I4.0 and data science. It was possible to enhance existing categories with new artefacts from the I4.0 range to represent the influence of cyber-physical systems (CPS), (big) data and information management, condition monitoring (CM) and more. Furthermore, the originally defined LSM-Model was restructured for a more simplified application in industrial use cases.

Keywords

Capability Maturity Model; Industry 4.0; Smart Factory; Maintenance; Asset Management; Lean Smart Maintenance; Digitalisation; Digitisation

1. Introduction

‘How do we adapt our production line and the associated maintenance strategy for the asset to the ever-increasing complexity of industrial plants and processes?’ - ‘How is it possible to achieve operational excellence with the help of the latest technological approaches of the fourth industrial revolution (Industry 4.0).’ These are only examples of the most frequently discussed questions in the industry, regardless of the sector in which the company (steel, paper, mobility, etc.) is located. Practitioners and academics all over the world are discussing how to implement new technologies, e.g. big data analytics, IoT, machine learning and deep learning applications in predictive maintenance, and therefore, a scientifically founded and tested concept has to be developed. The LSM concept, which is discussed in detail in chapter

DOI: https://doi.org/10.15488/9649
4, unites two approaches, the smart and the lean-approach. In the Lean Management concept, which is based on the Lean Production System by the Toyota Motor Company [1], a learning orientation, as well as risk and resource-oriented alignment, are crucial. The second part, smart maintenance management, attempts to put a focus on value-oriented, intelligent, and learning maintenance. In order to harmonise the LSM concept with the technical and organisational prerequisites of I4.0, the authors modify and expand the current LSM MM with selected categories and items from existing I4.0 MMs. In chapters two and three, the authors give a summary of I4.0, and its different key elements, as well as a systematic overview of MMs. Subsequently, the methodology of how this research was conducted is explained in chapter 5, which is followed by a deep dive into the Lean Smart Maintenance Maturity Model (LSM) and the main categories it contains, such as, e.g. ‘Corporate Culture’, ‘Spare Parts Management’, ‘Process Organisation’. This work concludes with Chapter 7, where a summary is given, and further research needs are identified.

2. Industry 4.0

The term Industry 4.0 is based on the German wording ‘Industrie 4.0’, which first appeared at the Hannover Messe in 2011. A group of experts from Germany defined this term in order to achieve industry leadership and excellence. [2] Although it is a German expression, and it is not well-known outside of the DACH-countries (Germany, Austria, Switzerland) [3], there are several similar approaches. The concept of Industry 4.0 is comparable to the ‘Industrial Internet’ [4], ‘Advanced Manufacturing’ [5] and also ‘Integrated Industry’ [6] or ‘Smart Manufacturing’, and ‘Smart Industry’ [7],[8],[9]. Industry 4.0 is characterised by the ever-increasing complexity of technical systems as well as by the holistic view of production systems [10]. Through the use of cyber-physical systems (CPS), advanced data analytics applications, cloud storage and computing and other technologies, the increasing connectivity between machines (Machine to Machines or M2M) or also known as the Internet of Things (IoT), Internet of Services (IoS) or machine to people (M2P) and technology-assisted people-to-people (P2P) interaction is resulting in the overall connectivity with unlimited data transfer and transparency on the Internet of Everything (IoE) [10],[11]. The company CISCO defines the IoE as ‘the intelligent connection of people, process, data, and things’ [12]. Related to the manufacturing and logistics industry, the term ‘Smart Factory’ combines all the aspects mentioned before [13].

3. Maturity Models

MMs are defined as artefacts with elements, that are arranged in an evolutionary scale with measurable transitions from one level to another and which are used for benchmarking, self-assessment and continuous improvement [14], [15]. MMs can be classified in different ways, one being the division of MMs into three different categories by Caralli [15]: (i) Capability Maturity Models, where the capabilities & processes of the organisation are the focus; (ii) Progression models, meaning a simple evolution of elements; and (iii) hybrid models, which combine the progression model architecture attributes, characteristics, patterns, etc., but the transitions between the maturity levels are defined by the capability maturity hierarchy. De Bruin describes that Maturity models can be descriptive, prescriptive, or comparative [16]. Fraser describes different possible components of MMs, which might be present in a model [17]: (i) a number of levels; (ii) a name of each level; (iii) a generic description of each level; (iv) a number of dimensions; (v) a number of elements for each dimension; (vi) a description for each of these elements. A maturity model architecture is defined as the fundamental organisation of a system embodied in its components, their relationships to each other, and to the environment, and the principles guiding its design and evolution [18]. The architecture shared by SW-CMM and CMMI is the most common MM architecture throughout the different domains, in which MMs are used [19]. Depending on the representation and CMMI version, CMMI uses five to six maturity levels, from Incomplete to Optimizing [20], [21]. Different authors have developed systematic
approaches for developing maturity models [22]. Most notably, de Bruin has developed a generic MM development framework, consisting of six development phases: (i) Determining the Scope of the desired model; (ii) Choosing the Design, or architecture of the MM; (iii) Populating the model with different components, or sub-components of the domain; (iv) Testing for validity, reliability and generalisability; (v) Deploying the model and (vi) Maintaining the model. According to these definitions, the LSM MM is currently in a development loop from phase iv to phase vi.

4. Lean Smart Maintenance

Smart factories, as part of I4.0, are getting more complex in the way of technological equipment, maintenance, and controlling of the overall manufacturing processes. The dynamic environment leads to an adaption of the maintenance processes and an adaptation of nearly every organisational aspect. As a result, the Lean Smart Maintenance (LSM) concept, a holistic management concept which increases the effectiveness and improves the efficiency of asset management was developed. Its foundation contains all aspects of planning, controlling and implementing of maintenance with strategic tools like performance measurement systems (KPI’s), Management by Objectives (MbO as target system), and also physical tools, e.g., IT-Systems, IoT-Applications, etc.[23] Each of the two pillars has its main focus, whereby one concentrates on the concept of the LEAN – cost-efficient and no waste, and the other one on the SMART – intelligent concept. Key concepts of the lean part are the efficiency perspective, where continuous reduction of waste, minimising non-value adding processes, and an overall sustainable resource orientation. Characteristic for the smart segment are the fundamental methods of problem-solving and performance enhancement, and the change from cost-controlled, input-oriented, across to output control, which focuses on reliability, availability, maintainability, and safety (RAMS).[23]–[25] As a result of the combination of the lean philosophy, and the smart maintenance approach, the LSM concept represents a holistic, sustainable smart, dynamic, and value-orientated maintenance management system [24].

5. Methodology

The methodological basis for the creation of the new MM is Design Science Research (DSR), as proposed by Hevner [26]. A further basis for this scientific work is formed by a structured literature review (SLR), as described by Tranfield [27]. The three stages and nine phases of an SLR are shown in Figure 1. The following sections outline the creation process of the new artefact, proposed by this scientific paper, following the three stages of a systematic review.

![Figure 1: Stages and phases of a Systematic Literature Review](image)

The impetus for this scientific contribution was the identification of a gap between the application of the LSM concept and the integration of I4.0 aspects. This gap was confirmed by LSM integration projects carried out by the Chair of Economics and Business Administration at the Montanuniversitaet Leoben. Therefore, it was decided to conduct a literature review on the basis of past literature reviews of I4.0 MMs and merge the collected data with LSM. Throughout this process, a review protocol was used to document all necessary information around the review. In a first step, literature searches were conducted using the abstract and citation data banks Scopus and Web of Science to identify past literature reviews on I4.0 and
digitalisation. This turned up 230 results in Scopus and 93 in Web of Science. After the elimination of duplicates and a title and abstract screening, six reviews on I4.0 MMs using Scopus and one further review using Web of Science were identified. A list of the seven review papers can be found in Table 2 in the Appendix. Next, a list of I4.0 MMs was created using the MMs reviewed in the aforementioned papers. After a title screening for relevance, this process resulted in 47 and 36 papers after an abstract screening. During the quality assessment, inclusion criteria were an industry focus of the paper, which eliminated a number of papers around software development, the presence of maturity levels and either a general description of the levels, or descriptions of the category levels and a number of maturity dimensions, or categories. The quality assessment resulted in a final number of 17 MMs for further investigation. A list of those 17 I4.0 MMs can be found in Table 1 in the Appendix. On this basis, the extracted I4.0 MM categories were compared to the 18 LSM categories. A few categories could not be allocated to LSM categories; these can be found under the category ‘Business Model and Service Strategy’. For this reason and because past projects have shown that a too high number of MM categories can result in problems when presenting the results and deriving measures, it was decided to restructure the maturity model into nine categories with a number of sub-categories each.

6. Lean Smart Maintenance Maturity Model

In this chapter, the new LSM categories are defined, and selected digitalisation and digitisation aspects are presented. The number of maturity levels, the names and generic descriptions for the levels are based on the CMMI architecture, it can, however, be classified as a hybrid and prescriptive model.

The St. Gallen Management model differentiates between three managerial levels: normative, strategic and operative management. While the goal of normative management is to develop a mutual understanding and reach social acceptance and societal legitimisation, strategic management refers to the development of a sustainable and competitive advantage and operative management strives to control efficient processes for daily problem solving and the day-to-day business.[28] Figure 2 depicts the different maturity categories ordered according to the managerial levels as presented in the St. Gallen Management Model from normative to operative management. Furthermore, the dependencies between the different categories are represented via white arrows. At the top of this LSM cube, the category ‘Philosophy & Target System’ can be found, which forms the basis for the categories ‘Corporate Culture’ and ‘Business Model & Service Strategy’, the latter also being influenced by the former. ‘Asset Strategy’ interacts with ‘Business Model & Service
Strategy’, and forms the basis for ‘Organisational Structure’, ‘Controlling & Budget’ and ‘Process Organisation’. ‘Controlling & Budget’ represents the basis for an adjustment of ‘Philosophy & Target System’ as well as ‘Asset Strategy’. ‘Data & Technology’ and ‘Knowledge Management’ form a second dimension that affects all LSM maturity categories. In total, the maturity model consists of nine main categories, where each of these main categories has a distinct number of sub-categories. Figure 3 is the graphical representation of the model and represents the ‘keyboard’ of the LSM concept, with the respective digital level of extension.

In the following sections, the LSM maturity categories will be described starting with the normative categories and ending with the second dimension depicted in Figure 2.

### 6.1 Philosophy & Target System

A defined corporate philosophy and the management philosophy derived from it, provide all employees with a direction for their behaviour, which is concretised by a vision [29]. A maintenance mission statement is used to communicate one’s requirements internally and to present and appreciate the importance of maintenance in the company [30]. The target system of the asset management, the maintenance and asset goals and their relationships are brought together and structured according to their target dimension [31]. For effective I4.0 implementation, maturity items like recognising digital challenges, addressing legal risk with collaboration partners and optimising the value chain network for legal & tax, security and compliance have to be considered [32]. Furthermore, a quality management system needs to be implemented throughout the organisation, including asset management [33].

### 6.2 Corporate Culture

One of the most critical aspects of management systems is how they handle corporate culture [34]. In the new LSM MM, corporate culture is composed of three sub-categories, ‘Culture & Motivation’, ‘Leadership & Change Management’, and ‘Communication’. Corporate culture is a very complex and multidimensional property of a group of people who learned to work together and achieve common goals [35]. To understand organisational behaviour, and transform it in a sustainable manner, it is necessary to know how to analyse, assess and change each element of its corporate culture [34]. Therefore, every category of the LSM concept is designed in such a way that it is deposited with the three cultural levels: artefacts, exposed values, and underlying assumptions. To increase the efficiency of transforming an organisation from a lower level of maturity to a higher one, the developed MM helps by identifying the right points to start the change process from top-down and bottom-up. Industry 4.0 and the digitalisation changes corporate culture in a way that the ‘willingness to change’ and also ‘social collaboration’ via the Internet becomes crucial for companies to survive. [36] The capability of each organisation to handle the data can be classified as ‘Innovation Culture’.
whereas Klötzer & Pflaum stagger this type of culture from the lowest level, ‘openness for service’ up to the highest level, which is represented by ‘digital enterprise thinking established’. [37]

6.3 Business Model and Service Strategy

Nowadays, companies are challenged with an increasingly complex environment and most business models, and service strategies have to undergo a radical change. Therefore, management has to become aware of new technologies, methods, and possibilities for this constant change through I4.0. [38] Geissbauer et al. identify the influence of I4.0 as disruptive for the whole supply chain, due to the change from mass production to single-item production. Through digitalisation and digitisation, a tremendous improvement of processes is possible and changes business as manufacturers become more and more permanent product owners who sell a service and not a product.[32], [39] The newly introduced LSM category ‘Business Model and Service Strategy’ extends the MM and enables enlargement of the understanding and analysis of maintenance as a service inside of companies.

6.4 Asset Strategy

The asset strategy is defined as the framework for how the operations around the asset are designed to reach a specific target and encompasses the sub-categories ‘maintenance strategy’, ‘maintenance prevention’, ‘outsourcing’ and ‘spare parts management’. The aspects of Maintenance Strategy is defined as a set of procedures (rules) that specify, in relation to the asset, which individual maintenance measures are to be carried out in terms of content, method and scope in a specific chronological sequence [40]. To guarantee the necessary agility in the future, a more frequent adjustment of the maintenance strategy is necessary. In order to be able to react to changes as quickly and flexibly as possible, processes for maintenance strategy adjustment must be defined [41]. Within the category ‘maintenance prevention’, continuous improvement and the extensive analysis of failures and occurring problems are used to transfer the gained knowledge into the acquisition of new equipment and assets. Through high connectivity between different types of assets, more in-depth insight into the behaviour of industrial facilities is possible.[42] This sub-category entails the influence of the present technological innovations on the lowest maturity level, where no maintenance prevention is given, up to the highest level, where a holistic life cycle assessment and life cycle thinking are implemented.[43]–[45] In the category ‘asset strategy’, ‘outsourcing’ refers to the procurement or outsourcing of maintenance services [46]. An overall outsourcing strategy has to be defined, and it is essential, that no core competencies are outsourced. The profitability has to be considered when outsourcing and service bundles should be outsourced together. Outsourcing should be used to cushion peak loads on the one hand while striving for long-term partnerships on the other. One-sided economic dependencies in external service relationships should be avoided, and continuous improvement in maintenance achieved. [47] The primary task of spare parts management is to provide the required quantity, type and quality of spare parts for the maintenance of plant and equipment to the appropriate user at the right time at minimum cost [48]. Modern spare parts management systems call for the use of tools for analysis of spare parts and a dynamic spare parts management strategy. New technologies, like mobile devices and 3D printing, have a high potential to revolutionise spare parts management [41].

6.5 Maintenance Budget & Controlling

Maintenance Controlling is used for planning, controlling and monitoring activities within Plant Maintenance as well as for coordinating and reconciling activities with Production, Materials Management and Cost Accounting [49]. In relation to plant maintenance, the budget is the cost-oriented limit within which plant maintenance has performance-related leeway [44]. A risk and decision-oriented budgeting approach is seen as the optimum for maintenance budgeting since it meets the requirement for dynamic and future-oriented budgeting and adjustment in the current period [50]. In addition, LSM calls for a key performance indicator system derived from the various success factors, and a cause-related recording of maintenance
costs [41], [46], [51]. In the I4.0 MMs factors like productivity, quality, cost, lead-time, safety, environment and the connectivity between Key Performance Indicators are mentioned [52], [53].

### 6.6 Organisational Structure

The formal division of job tasks, how they are grouped and coordinated is defined by the organisational structure [54]. In maintenance and asset management, design of the organisational structure is substantially influenced by choice of the organisational structural principle, the definition of the structuring option and the degree of decentralisation [46]. The LSM category includes ‘autonomous maintenance’, ‘object orientation’, the ‘location of workshops’ and ‘information around the organisational structure’. For successful digitisation and digitalisation of the company, flexible communities and decision rights management are identified as integral to the development of a company [36]. Collaboration is seen as necessary to become a digital champion [32]. Promoters, the adaption of the R&D department and the internal IT organisation, the implementation of an internal smart service organisation and a service ecosystem and finally (Big) data analytics are further aspects that have to be considered when adapting the organisational structure [37].

### 6.7 Process Organisation

Process organisation in asset management encompasses the combination of individual work steps to complex processes as well as the harmonisation within and across processes in terms of time and space [55]. The category focuses on ‘planning’, ‘information in process organisation’, ‘continuous improvement’ and ‘weak-point-analysis’. In the context of I4.0, an evolution to agile & networked optimisation of processes is necessary [56]. The goal of digitalisation is autonomous process planning and control [39], while an integration across the vertical and horizontal dimensions of the value chains, which includes real-time availability of all data, support by augmented reality and optimisation in an integrated network, is promoted [32], [42]. Real-time planning, the level of automation operation process traceability and simulation visualisation technologies are further aspects to be considered [57]. The overall digitalisation of business processes and if service system management and data lifecycle management processes are defined and established are other aspects to consider when assessing maturity [37].

### 6.8 Knowledge Management

As digitalisation increases, a tremendous amount of data and information are gathered, and therefore the management aspect of data, information and the resulting knowledge has to store and be accessible. To create a highly efficient maintenance organisation, processed knowledge has to be highly accurate, interdisciplinary and above all; it must be seen as a key production factor [58]. Klötzer and Pflaum have created a new perspective on knowledge by bringing together innovation and knowledge in the single category ‘cooperation’. For the future implementation of the digitalisation in the industry, the LSM concept entails knowledge management as a central component of employee training, qualifications, continuous improvement and in particular the accompanying error culture as several other authors have already mentioned in their publications.[37], [59]

### 6.9 Data & Technology

Since the LSM concept is based on the targeted use of technologies and the reasonable use of data, it is necessary to identify the intersections of technology, data and knowledge as well as the particular management aspects [41]. As it comes to the use of new technologies as handheld devices, wearables, Internet of Things (IoT), (Big) Data Analytics (BDA), to mention a few examples, aspects of LSM are useful for integration and acceptance in the company. In a management model, as it is represented by the LSM MM, every level of each category is influenced by the appearance of new technologies. For example, the
qualification and training of employees will massively change due to the availability and use of virtual reality, blended-learning concepts and the permanent availability of Information and Communication Technology (ICT).[60]

7. Summary & Outlook

In this research paper, the authors aimed to identify I4.0 elements to expand the LSM MM. The goal was to fill the gaps between the LSM concept and state of the art I4.0 MMs concerning maturity items and categories for digitalisation and digitisation. The methodology used is based on a structured literature review and design science research. After analysing literature reviews on MMs for I4.0 and Smart Factory, a comprehensive list of research papers regarding the aforementioned domain was created, the different maturity models were compared, and maturity categories and items extracted. Based on this analysis, missing digitalisation and digitisation properties of each maturity category and level were integrated into the LSM MM. In the course of this process, the need to restructure the LSM MM was identified, and the number of categories was restructured from eighteen to nine categories with several sub-categories and maturity items. The review of the I4.0 MMs has shown that the authors of the MMs do not consider ‘Outsourcing’ and ‘Asset Strategy’ sufficiently in their models. However, due to the rising integration of IT systems around aspects of these LSM categories, they were checked for I4.0 readiness. One reason for the lack of these maturity dimensions in the I4.0 MMs is that the reviewed MMs take a general perspective on the topic of I4.0 and do not focus on the domain of industrial asset management. This research only considered MM papers found in the review papers published from 2017 to 2019, listed Table 2 in the Appendix, which means, that I4.0 MMs published after the newest review paper were not considered. Furthermore, all MMs focused on software development were excluded after the abstract screening. The following steps include a field test, further development and improvement based on the newly structured LSM MM. The critical point is to prove the practical applicability of the new model. Due to the complexity and the advancement of technical developments, the need for further research in the different categories has been identified. Furthermore, further research is needed to identify the dependencies between the different sub-categories and maturity items of the model.
## Appendix

### Table 1: I4.0 MMs

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Title</th>
<th>Categories</th>
<th>Levels</th>
<th>Annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akdil et al.</td>
<td>2018</td>
<td>Maturity and Readiness Model for Industry 4.0 Strategy</td>
<td>3 categories, one of the categories has 3 sub-categories, 13 items</td>
<td>4 levels: Absence - Existence - Survival - Maturity</td>
<td>Detailed level descriptions for each category, questionnaire</td>
</tr>
<tr>
<td>Anderl et al.</td>
<td>2016</td>
<td>Guideline Industrie 4.0 - Guiding principles for the implementation of Industrie 4.0 in small and medium sized businesses</td>
<td>6 categories each for production and products</td>
<td>5 levels for each category</td>
<td>Description for each category and the implementation process</td>
</tr>
<tr>
<td>Berghaus et al.</td>
<td>2017</td>
<td>Digital Maturity &amp; Transformation Report 2017</td>
<td>9 categories</td>
<td>5 levels which are defined according to the difficulty levels of indicators</td>
<td>A comprehensive survey among Swiss companies</td>
</tr>
<tr>
<td>Anderl et al.</td>
<td>2016</td>
<td>Guideline Industrie 4.0 - Guiding principles for the implementation of Industrie 4.0 in small and medium sized businesses</td>
<td>6 categories each for production and products</td>
<td>5 levels for each category</td>
<td>Description for each category and the implementation process</td>
</tr>
<tr>
<td>De Carolis et al.</td>
<td>2017</td>
<td>A Maturity Model for Assessing the Digital Readiness of Manufacturing Companies</td>
<td>5 categories, 18 sub-categories</td>
<td>5 levels: Initial - Managed - Defined - Integrated and Interoperable - Digital-Oriented</td>
<td>3 case studies in a second publication with potentials for improvements</td>
</tr>
<tr>
<td>Ganzarain &amp; Errasti</td>
<td>2016</td>
<td>Three stage maturity model in SME’s towards Industry 4.0</td>
<td>3 categories</td>
<td>5 levels</td>
<td>process model</td>
</tr>
<tr>
<td>Gökalp et al.</td>
<td>2017</td>
<td>Development of an assessment model for industry 4.0: Industry 4.0-MM</td>
<td>17 items, that are assigned to the different maturity levels</td>
<td>6 levels: Incomplete - Performed - Managed - Established - Predictable - Optimizing</td>
<td>-</td>
</tr>
<tr>
<td>Geissbauer et al.</td>
<td>2016</td>
<td>Industry 4.0: Building the digital enterprise</td>
<td>7 categories</td>
<td>4 levels: Digital novice - Vertical integrator - Horizontal collaborator - Digital champion</td>
<td>List of possible pilot projects given</td>
</tr>
<tr>
<td>Halse et al.</td>
<td>2016</td>
<td>IoT Technological Maturity Model and assessment of Norwegian manufacturing companies</td>
<td>No categories, but 25 items, that are assigned to the different maturity levels</td>
<td>8 levels: 3.0 Maturity - Initial - Connected - Enhanced - Innovating - Integrated - Extensive - 4.0 Maturity</td>
<td>Assessment of four companies, very detailed level descriptions</td>
</tr>
<tr>
<td>Jung et al.</td>
<td>2016</td>
<td>An Overview of a Smart Manufacturing System Readiness Assessment</td>
<td>4 categories, 6 sub-categories</td>
<td>5 levels: Initial - Managed - Defined - Qualitative - Optimizing</td>
<td>-</td>
</tr>
<tr>
<td>Katsma et al.</td>
<td>2011</td>
<td>Supply Chain Systems Maturing Towards the Internet-of-Things: A Framework</td>
<td>4 categories</td>
<td>4 levels: ERP - ERP 2.0 - SOA/SAAS - Internet of Things</td>
<td>4 case studies and cross-case analysis</td>
</tr>
<tr>
<td>Klötzer &amp; Pflaum</td>
<td>2017</td>
<td>Toward the Development of a Maturity Model for Digitalization within the Manufacturing Industry’s Supply Chain</td>
<td>9 categories each for ‘Smart product realization’ and ‘Smart product application.’</td>
<td>5 levels: Digitalization awareness - Smart networked products - The service-oriented enterprise - Thinking in service systems - The data-driven enterprise</td>
<td>-</td>
</tr>
<tr>
<td>Lee et al.</td>
<td>2017</td>
<td>A Smartness Assessment Framework for Smart Factories Using Analytic Network Process</td>
<td>4 categories, 10 sub-categories, 46 assessment items</td>
<td>5 levels: Checking - Monitoring - Control - Optimization - Autonomy</td>
<td>A case study with 20 SMEs</td>
</tr>
<tr>
<td>Leyh et al.</td>
<td>2017</td>
<td>Assessing the IT and Software Landscapes of Industry 4.0-Enterprises: The Maturity Model SIMMI 4.0</td>
<td>4 categories</td>
<td>5 levels: Basic digitization level - Cross-departmental digitization - Horizontal and vertical digitization - Full digitization - Optimized full digitization</td>
<td>Activities for each level defined</td>
</tr>
<tr>
<td>Lichtblau et al.</td>
<td>2015</td>
<td>Industrie 4.0-Readiness</td>
<td>6 categories, 18 sub-categories</td>
<td>6 levels: Outsider - Beginner - Advanced - Experienced - Expert - Excellence</td>
<td>Study on I4.0 readiness of German companies</td>
</tr>
<tr>
<td>Schuh et al.</td>
<td>2017</td>
<td>Industrie 4.0 Maturity Index</td>
<td>4 categories</td>
<td>5 levels: Computerisation - Connectivity - Visibility - Transparency - Adaptability</td>
<td>Investigation of capabilities in different functional areas</td>
</tr>
<tr>
<td>Westermann et al.</td>
<td>2016</td>
<td>Reference Architecture and Maturity Levels for Cyber-Physical Systems in the Mechanical Engineering Industry</td>
<td>7 categories, 12 sub-categories</td>
<td>5 levels: Monitoring - Communication and Analysis - Interpretation and Services - Adaption and Optimization - Cooperation</td>
<td>Morphological Box for evaluation</td>
</tr>
</tbody>
</table>

86
Table 2: I4.0 MM Reviews

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Title</th>
<th>Focus</th>
<th>#MMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Felch et al.</td>
<td>2019</td>
<td>Maturity Models in the Age of Industry 4.0 – Do the Available Models Correspond to the Needs of Business Practice?</td>
<td>Scientific and consultancy I4.0 MMs</td>
<td>20</td>
</tr>
<tr>
<td>Unterhofer et al.</td>
<td>2018</td>
<td>Investigation of Assessment and Maturity Stage Models for Assessing the Implementation of Industry 4.0</td>
<td>Broad-spectrum of I4.0 MMs from different domains</td>
<td>60</td>
</tr>
<tr>
<td>Blatz et al.</td>
<td>2018</td>
<td>Maturity model of digitization for SMEs</td>
<td>Focus on a few MMs for the as-is analysis of the level of digitisation of SMEs</td>
<td>4</td>
</tr>
<tr>
<td>Mittal et al.</td>
<td>2018</td>
<td>A critical review of smart manufacturing &amp; Industry 4.0 maturity models: Implications for small and medium-sized enterprises (SMEs)</td>
<td>Analysis of I4.0 and smart manufacturing MMs with a focus on their categories and items, on research gaps and the adaption for SMEs</td>
<td>15</td>
</tr>
<tr>
<td>Akdil et al.</td>
<td>2018</td>
<td>Maturity and Readiness Model for Industry 4.0 Strategy</td>
<td>Industry 4.0 MMs</td>
<td>4</td>
</tr>
<tr>
<td>Gökalp et al.</td>
<td>2017</td>
<td>Development of an Assessment Model for Industry 4.0: Industry 4.0-MM</td>
<td>Industry 4.0 MMs</td>
<td>7</td>
</tr>
</tbody>
</table>

References


Biographies

Hubert Biedermann (*1953) is professor at the chair of Economics- and Business Management and head of the Department of Economic and Business Management at Montanuniversitaet Leoben. His main research areas are asset management - lean smart maintenance, production management, quality and risk management. After his habilitation in industrial economics in 1989, he worked for several years in the industry and as a visiting professor. Appointed Montanuniversitaet Leoben in 1995, he held the position as vice-rector for finance and controlling for more than 12 years. ORCID-ID: https://orcid.org/0000-0003-4678-4392

Hans Thomas Maier (*1989) started his academic career in the International Study Program in Petroleum Engineering at Montanuniversitaet Leoben. Since 2018 he is working at the Chair of Economic- and Business Management at Montanuniversitaet Leoben and focuses on research in the fields of industrial asset management and change management. ORCID-ID: https://orcid.org/0000-0001-5714-3959

Oliver Schmiedbauer (*1990) studied Mechanical Engineering and Business Economics at Graz University of Technology. Since 2018 he is working at the Chair of Economic- and Business Management at Montanuniversitaet Leoben and focuses on research in the fields of industrial asset management and business excellence. ORCID-ID: https://orcid.org/0000-0003-0051-9614
Development of a Concept for Real-Time Control of Manual Assembly Systems

Martin Sudhoff¹, Matthias Linsinger¹, Daniel Schulte¹, Bernd Kuhlenkötter¹
¹Ruhr-Universität Bochum, Chair of Production Systems, Universitätsstr. 150, 44801 Bochum, Germany

Abstract
In contrast to automated machines and installations, manual assembly still lacks real-time process monitoring and possibilities for short-term control and adaptation of assembly systems. This article describes an approach for a concept of real-time control of manual assembly systems. For this purpose, KPIs that can be determined predictively are considered. These indicators enable a standardized and objective process data acquisition and a local process optimization for a higher flexibility and adaptability. In addition to the key figures developed, an approach for the automated acquisition of appropriate process data in manual assembly is described. The further usage of the KPIs and the validation within a real production environment is finally presented.

Keywords
Production Planning; Manual Assembly; Process Control; Key Performance Indicator; Standards

1. Introduction
Shorter product life cycles and smaller batch sizes impact increasingly on industrial production. In addition, there is a diversification of the product range and increasing cost pressure. Corresponding to these conditions, future innovations are more and more subject to the targets of adaptability and flexibility [1]. One way to ensure this flexibility in assembly systems is to continuously monitor process parameters and adjust the configuration of the system. This is already frequently used in automated assembly. In manual assembly, however, this process monitoring and control is missing [2]. The reason for this is the absence of process data acquisition. The integration of sensors and specific evaluation is rarely used in manual assembly [3]. Due to the non-existent objective data acquisition, the current state in manual assembly does not enable real-time process monitoring and thus also no short-term production control. Furthermore, there is no standard concerning KPIs (Key Performance Indicators) for real time control in manual assembly.

In order to implement real-time process monitoring in manual assembly systems, KPIs commonly used in other fields of production are first analysed. For the development of the KPIs existing systems of predetermined times (e.g. Work Factor or MTM (Methods-Time Measurement)) on the one hand and already existing KPIs used for automated machine monitoring (such as Overall Equipment Effectiveness (OEE)) on the other hand are purposefully derived. After that, the application of the developed indicators within an industrial environment at the LPS is described and a concept is presented which allows real-time data acquisition with Message Queuing Telemetry Transport (MQTT) within manual assembly systems.

DOI: https://doi.org/10.15488/9650
2. State of the art

2.1 KPIs for production systems

Key Performance Indicators are indicators that enable organisations to measure progress or the degree to which key objectives have been met. A business key figure serves as a basis for making decisions, for control (target vs. actual) as well as for the documentation and coordination of important processes. [4] The standard VDMA 66412-1 summarizes common KPIs in industry. Based on the classification of the standard for use in Manufacturing Execution Systems (MES), an application focus of KPI for processing machines, operating personnel and other automated processes can be identified [4]. This is also reflected in the KPIs mentioned in the standard, such as “Overall Equipment Effectiveness” (OEE) to describe the efficiency of machines and systems, or “availability” as an indicator of machine utilization.

It can be concluded that many KPIs have been developed in particular for automated production processes and cannot be applied to predominantly manual and hybrid assembly without additional adaptation. However, some of the indicators described in the mentioned VDMA-standard, such as lead time, idle time or employee productivity, are of equal importance for manual and hybrid assembly systems.

2.2 Value stream analysis

Within value stream analysis, key figures such as lead time, waiting time, stocks or processing times are recorded in order to obtain an indication of the degree of flow of a process. As a result, the process behaviour can be displayed via cycle time diagrams and the bottleneck of a process can be determined very quickly. At this point, no distinction is made between manual and automated processes. [5,6] Hence, the value stream method uses some process analysis tools which are suitable to measure manual or hybrid assembly processes.

2.3 Assembly process planning

There are numerous methods for medium to long-term assembly process planning. The purpose of assembly planning is to minimize the costs per unit. It contains the methods for designing the work content of an assembly system [7]. A distinction is made between assembly system planning and assembly process planning. REFA presents a general planning procedure for production systems in six stages, which can be applied for the planning of assembly systems [8]. Another example is an approach by Lotter. He has developed a planning system specially designed for assembly systems. It consists of 11 steps that can be adapted to the assembly requirements depending on the characteristics and complexity of the product [9]. For process planning, the MTM procedure can be mentioned as a system of predetermined times for manual assembly [10]. To describe an assembly process in terms of sequence and restrictions, a precedence graph or assembly priority chart can be used. It shows the individual assembly steps in a technically and organizationally predefined sequence. In addition, processing times and required resources are documented. [11,12,7] Line balancing is a key factor in assembly process planning. The more balanced a process, the shorter the idle times and waiting times and the better the utilization of personnel and assembly technology [11,13]. One key figure used to determine line balancing, for example, is the flow rate [5]. Both, assembly system planning and assembly process planning methods are always applied before the implementation and execution of an assembly process. Consequently, they are not suitable for monitoring an ongoing assembly process.

In the short-term planning horizon, assembly control deals with the activities necessary for the fulfilment of assembly orders on the basis of the results of assembly planning. This includes the definition of the workload, the provision of materials, the provision of information, personnel deployment planning as well as the monitoring of the assembly progress and reaction to malfunctions. [14,15] However, there are no methods found in literature that measure KPIs in real time within a manual or hybrid assembly process and use them for systematic in-process assembly control such as situational line balancing for instance.
As a summary of the state of the art, it can be stated that for manual and hybrid assembly processes no KPIs are measured in real time and used for process control. Thus, there is a specific need for action in the definition of suitable KPIs for manual and hybrid assembly processes as well as their technical real-time elevation in an ongoing assembly process in order to be able to take suitable control measures at short notice.

3. Indicators for manual assembly systems

As a first step of the presented approach, suitable KPIs for assembly process monitoring have been identified. A crucial criterion is that KPIs can be determined predictively in order to be able to make short term forecasts for imminent assembly processes. For a first pilot application, the following KPIs have been identified for process prediction and monitoring:

- lead time,
- throughput,
- process ratio and
- cycle time deviation.

The calculation of the lead time is usually described as the sum of all processing times, waiting times and transport times (see [5]). These are summarized under the process time (PT). In order to determine the lead time of an entire lot in an assembly system with several workstations, the bottleneck process with PT\text{max} must also be taken into account. For this purpose, Linsinger and Stecken et al. have presented a formula (1) [16]:

\[ L_{Lot} = \sum_{i=1}^{n} PT_i + PT_{max} \cdot (m - 1) \]  

(1)

Formula (1) shows that the lead time of an entire lot L\text{lot} over several process steps results from two summands: The first addend describes the lead time of the first part, which can run through the work steps of an empty system without waiting times as there are no other parts within the assembly system. Its lead time results from the sum of all process times PT\text{n}. The second addend determines the lead time of all following parts of the lot. The first part that has already been determined with the first addend must be subtracted from lot size m. Since these parts always have a preceding part in the process, they can only move through the process depending on the longest work step PT\text{max} as PT\text{max} represents the achievable cycle time of the current process.

The throughput specifies the product quantity that can be produced in a time unit. It is calculated by dividing the lot size by the lead time.

The process ratio gives an indication about the so called process density. The higher the ratio, the more efficient is the process. It is calculated by dividing the value adding process steps (in time unit) by the lead time.

The cycle time deviation compares the highest process time PT\text{max} within an assembly station (this gives the actual possible cycle time) with the average ideal cycle time T for an optimized balanced assembly system. It is calculated according to formula (2):

\[ D_T = \frac{PT_{max} - T}{T} \cdot 100\ [%] \]  

(2)
4. Industrial Application

The suitability of the identified and adapted KPIs for assembly process monitoring has been demonstrated using the assembly line for terminal strips at the LPS learning factory. The motivation to operate assembly line production for industrial customers at LPS in cooperation with Phoenix Contact GmbH & Co. KG. is the development and testing of new assembly processes and technologies under industrial conditions. The u-shaped assembly line consists of six stations. After cutting the strips, terminals are mounted onto the strip at station 1. Then the labelling takes place at station 2. Based on this, circuit bridges and other additional components are assembled at station 3 before the quality with regard to deviations is checked at station 4. Finally, pre-cabling (Station 5) and packaging for shipping the terminal strip is carried out (Station 6). [17] In order to support the employee, a mobile robot is either used for terminal (Station 1A) or circuit bridge assembly (Station 3A). [16]

We developed a software tool, which uses planned process values based on MTM in order to calculate the KPIs. Based on the input of the current customer order, the program displays the predicted lead time with the current configuration of the assembly system by means of a KPI cockpit. In addition to the presented KPIs, a cycle time diagram of the assembly process is given. It displays the processing time of each assembly station. The KPI cockpit is shown in Figure 1 by means of an exemplary real customer order of 15 terminal strips.

Using the given configuration of the assembly system, a cycle time of more than 15 working hours is predicted for two employees. With a process time of approximately 20 minutes the cycle time diagram visualizes a bottleneck at station 3 where the processing time is almost two times as high as at the remaining station. Therefore, the cycle time diagram immediately provides an explanation for the low process ratio of only 28 % and the high cycle time deviation of 132 %.

Based on the information of the cockpit, the need of an immediate countermeasure to reduce processing time at station 3 can be derived. An organisational measure could be to double station three (reduced lead time by 36 %). Another possibility is moving the mobile robot from the first to the third station (reduced lead time by 25 %). The example shows the importance of monitoring assembly processes. Using MTM planning values, KPIs can be calculated to optimize the process for a given order. However, in order to conduct an in-
process monitoring for manual and hybrid assembly processes, technical solutions have to be implemented and methods as well as countermeasures for situational process adaptation have to be developed.

5. Objective data acquisition for real time monitoring

It has been demonstrated that individually developed KPIs on the basis of planning values are enabling efficient monitoring in manual assembly. In order to further improve this process prognosis, the static basic data must be supplemented with dynamic process data. These have to be collected and processed automatically. A concept for automatically capturing and processing this data is presented in the following.

In our assembly system the relevant data are both process data and customer order data. The relevant customer order data consists of the product identification number, the ordered lot size and the required delivery date. The dynamic process data to be collected includes the processing times, waiting times and transport times (see chapter 3). In addition, the throughput must be measured.

In order to collect these data, the terminal strips are provided with an individual QR-code so that they can be located within the assembly system. Furthermore, as a first step, each assembly station will be equipped with a scanner integrated into the workpiece fixture, so that an automatic detection can take place without causing additional work for the employee. The scanner is connected to a microcontroller and integrated in a network using MQTT. [18] This protocol is designed to connect embedded devices with applications and middleware and is therefore an optimal protocol for machine to machine communication. [19,20] Hence, the sensors act together with the microcontroller as publishers and send the information to the broker (Figure 2). In addition, the Order Client receives and processes the customer order data for publishing it. The broker provides this data to various subscribers like the actuators and dashboards within the assembly system on the one hand. On the other hand, an OPC-Router acts as a subscriber and protocol gateway for providing the data in OPC UA protocol for the integration into the already existing IT infrastructure in the learning factory. Thus, the data from the manual assembly as well as the data from the manufacturing area are transferred to the SQL database where they can be further analyzed and evaluated.

This implementation for data acquisition and processing enables real-time calculation of process and waiting times. In addition, the stocks between the stations and the throughput can be determined. This makes the prediction, which was previously based on static data, real-time-capable and allows flexible reaction to unforeseen changes like rework or technical faults.
6. Summary and Outlook

The article introduces a concept implementing a real-time control of manual assembly systems. For this purpose, existing KPIs are analysed and adapted to manual assembly. These indicators were afterwards tested in a real production environment and it is shown that an increase in productivity is possible with process control. In order to further improve this process prognosis, real-time data will be included. For this reason, an approach for real-time process data acquisition and analysis is presented. This concept will be integrated into the existing IT infrastructure of the LPS in 2020. The KPIs, validated with static data to date, will be validated and enhanced on the basis of real-time data taken from the manual assembly system. Furthermore, the sensitive use of personal data will be considered within this application. The following tests will show to what extent process control can be improved with the help of real-time data.

References


Biography

Martin Sudhoff (*1991) is a member of the Lehrstuhl für Produktionssysteme (LPS) at the Ruhr-University of Bochum since 2017. He earned a bachelor’s and master’s degree in mechanical engineering at the Ruhr-University of Bochum. His primary research topics are the digitalization and automation of assembly systems.

Matthias Linsinger (*1979) is a member of the Lehrstuhl für Produktionssysteme (LPS) at the Ruhr-University of Bochum since 2016. He earned a diploma in industrial engineering at the Cooperative State University Stuttgart and a master’s degree in manufacturing engineering at the University of Nottingham. His primary research topics are hybrid assembly systems and human-robot interaction in assembly.

Daniel Schulte (*1990) is a member of the Lehrstuhl für Produktionssysteme (LPS) at the Ruhr-University of Bochum since 2018. He earned a bachelor’s degree in electrical engineering and master’s in industrial informatics at the Hochschule Emden/Leer. His primary research topics are the industrial communication and digitalization of manufacturing facilities.

Till 2009 Bernd Kuhlenkötter was responsible for product management and technology at ABB Robotics Germany. In 2009 Bernd Kuhlenkötter took over the Professorship for "Industrial Robotics and Production Automation" at the Technical University of Dortmund. Since 2015 he holds the professorship of “Production Systems” at the Ruhr-University of Bochum.
Concept for the cost prognosis in the industrialization of highly iteratively developed physical products

Günther Schuh¹, Jan-Philipp Prote¹, Andreas Gützlaff¹, Sven Cremer¹, Shari Wlecke¹
¹Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Germany

Abstract

With the ongoing technological progress and increasing global competition, companies are facing a continuously changing market environment. Due to the volatility of the market, rapid product adjustments and shorter product life cycles are required. Changing customer requirements are rarely taken into account, leading to inventions that do not make the transition to innovations. Highly iterative product development poses a possibility to integrate the customer voice into the development process and thus shorten the time-to-market and enable companies to respond to changes in requirements.

Within the scope of highly iterative product development methods, cost analysis remains one of the main challenges for companies. Since the scope of development is not known at the beginning of a project, neither development nor industrialization costs can be specified. This, however, is essential for product and process development to meet cost-related customer requirements and for forecasting the production and investment budgets. With existing methods, it is either possible to agree to a fixed development budget and target price or to enable the customer to make changes during development.

The concept presented in this paper aims to counteract this challenge. Therefore, existing approaches are analyzed with regard to derived requirements for the transfer from highly iterative and integrated product and process development to agile cost analysis. Influencing factors on product and production process costs are identified based on findings from literature. By aligning the influencing factors and requirements, dependencies between target costs of a product and degrees of freedom of highly iterative product and production process development can be derived and used for the development of a framework for iterative cost analysis. In conclusion, a concept for an agile cost prognosis for the industrialization of highly iterative developed physical products is presented.

Keywords

Highly iterative product development; Process development; Industrialization; Cost analysis; Cost prognosis

1. Introduction

In order to remain successful in a continuously changing market environment, manufacturing companies must be able to react to changing customer requirements and market influences throughout the entire development process. Main reasons for this change are increasing volatility of markets, reduced product life cycles and increasing product complexity. [1] Due to these challenges, the importance of the industrialization regarding the time and cost required for an invention to become an innovation is crucial for the product’s success. Thus, companies must not only reduce the time to market but also remain adjustable during product development. [2]
Highly iterative product development (HIPD) offers the possibility to include changing customer requirements continuously in every project phase. [3,4] The approach is based on the fact that the results of each iteration cycle can be realized as an independent prototype. These prototypes enable a fast integration of customer requirements even in later project phases, an improved product quality and the shortening of the time-to-market. [5] However, the use of HIPD results in high planning uncertainties for downstream production process development due to continuous and late changes. [4] Whereas with HIPD a short time horizon drives development through short-term sprints, production and assembly planning is usually carried out on a long-term basis up to the start of production. [3,6] This excludes the possibility to make early investment decisions in HIPD, which contradicts with production process development since these pose the basis of the development progress. Therefore, an integrated product and production process development for the industrialization of physical products is needed to create a decision-making basis for long-term decisions in an iterative development approach. [4,7]

With the highly iterative and integrated development, cost prognosis poses a main challenge and is difficult to estimate due the uncertainties of development and industrialization costs. The purpose of cost calculation is to allocate all costs that occur in the production process according to their origin. Therefore, different unit-related cost information, purchase prices and manufacturing costs are complied. On this basis, market and internal transfer prices are defined and process decisions such as in-house production, external production and manufacturing processes can be made. [8] Cost forecasting is important to place a new product not only on the market at the right time, but also at market-driven costs. It is crucial to have sufficient transparency in the entire product and production process development as to which costs are associated with the product. A sufficient cost prognosis must be as highly iterative as the considered development process. Current cost prognoses do not meet this challenge, since they are based on definitions and decisions at the beginning of development. [9]

In current cost calculation the development department is of central importance. Even though only 9% of the costs are caused in this development, 70% of the costs are determined in this phase. In contrast, production is one of the main cost drivers, accounting for 28% of the costs incurred. However, during the development phase most of the costs are not considered so that it is not possible to take targeted countermeasures to ensure the market price. [10] To be able to estimate the costs arising during the design phase, calculations are currently based on similarity comparisons with already manufactured and recalculated products. [9] However, changing customer requirements lead to significant cost differences. Apart from purchasing, considerable costs are incurred in the production process during product manufacturing, hence special focus is placed on potential savings. In contrast to development costs, the costs incurred can be allocated directly to the product as direct costs. To identify potential savings, precise cost breakdowns must be available. [10] Therefore, the concept presented in this paper allows the inclusion of the cost prognosis in the highly iterative and integrated product and production process development.

In conclusion, the imbalance between cost generation and cost causation arises due to the department-specific approach, as opposed to an interdisciplinary approach. In addition, the lack of transparency in product cost accounting evolves due to calculations based on similarity comparisons. Downstream process costs are rarely included in current cost calculation. Production costs are rationalized after the product cost calculation has been completed, thus synergies between product and production costs are not considered. This leads to a lack of transparency in cost prognosis throughout the development progress and is not compatible with highly iterative, customer-oriented approaches.

2. Prior research in cost analysis and prognosis for highly iterative

For an agile cost prognosis based on highly iterative and integrated product and production development, costs need to be regarded completely from the beginning. Therefore, production related costs need to be
integrated. For this integration information exchange must be possible throughout the development progress with a responsive and transparent cost prognosis that enables interaction between departments. A literature review was conducted using a forward and backward analysis according to WEBSTER AND WATSON. [11] Since cost calculation is the foundation of this work, this term was used as the origin for this research, which has been extended by agile approaches for cost analysis. In the following, the focus lays on cost prognosis for highly iterative product and production process development.

“Target costing” is a concept for market-oriented target cost management and is applied in the early phases of product development. The aim is to increase competitiveness by means of market-oriented requirements. The focus lays on the derivation of permitted market costs using market prices as the basis. The costs allowed by the market are compared with the production costs, so that target costs can be derived for individual product groups. [12,13] However, the focus lays on determining the maximum production price using a backward calculation. Thus, the calculation is neither complete nor are the production costs integrated. In addition, a fixed market price must exist for a retrograde cost calculation.

“advANTAGE: A Fair Pricing Model for Agile Software Development Contracting” by BOOK ET AL. provides a pricing model to share the risk of agile software development between customers and suppliers. The development process is divided into four phases, in which requirements are recorded, divided and implemented in successive cycles, as in Scrum approaches. For cost prognosis, the first and last phase are relevant. While in the first phase upcoming costs of each cycle are estimated, in the last phase costs for the development are flexibly adjusted according to the fulfillment of the task within the previously defined resources. If fewer resources are consumed than specified, the charges are reduced by the rate of underspend resources. If the consumption is too high, the charges are reduced by a certain percentage, which disadvantages the supplier. [14] The model and subsequent adjustments are based on previously estimated values, making the entire process highly dependent on the accuracy of the first estimate. Also, it is used in Scrum approaches, so that it is mostly suitable for software development.

“Process costing” is an approach to increase the cost transparency of indirect performance areas. It does not constitute an independent cost accounting system, but complements the traditional systems with an improved allocation of overheads. [15] Resource-oriented cost accounting focuses on indirect costs through variant diversity aiming to improve product and development-related cost information in complex series production. [16] In practice, however, process costing is not used to calculate total costs. The aim is to model only those areas of the company that are directly related to the variety of variants. In addition, it is hardly possible to flexibly react to changes in the product and production process.

“Total cost of ownership (TCO)” approaches are mostly used in the field of mid- and long-term purchasing decisions. [17,18] They are used as a purchasing tool aimed at understanding the actual cost of buying and using a product or service from a supplier. TCO are the sum of purchase, financing, running, and infrastructure costs and highly sensitive to assumptions. Most TCO approaches are used to support the purchasing case in new technology fields with high uncertainties like electro mobility. [19] The high level of customer focus ensures the integration of customer requirements. However, the focus on improving the product from a technical side of view neglects the advantages and requirements of integrating production. Further on, the transparency on costs is missing which is why permanently changing costs cannot be analyzed within a highly iterative development.

The “functional cost analysis” is a method to determine the value of a product’s function. Functional costs include all planned or incurred resources that are necessary to provide a function. Functional cost estimates are made at the value creation stage, i.e. before the design or development and realization of the product. It is possible to allocate component or assembly costs to the functions they perform or in the performance of which they are involved. The approach is applicable to physical as well as non-physical products and integrates purchasing, production and sales as well as suppliers and customers are included in the analysis.
However, the approach is not able to react and adapt to constantly changing development, nor is it an approach for a stand-alone total cost accounting.

**Conclusion.** The evaluation of the different approaches shows that current cost calculations are not able to meet the requirements of integrated product and production process development, as shown in Figure 1. Especially when estimating the total costs at the beginning of the development, the production process costs are usually not considered differentiated. Calculations are not based on the bill of material (BOM) and therefore not product- and process-oriented. Without a complete cost calculation, it is not possible to proactively react to costs that exceed the target costs. Thus, the responsiveness of the calculations to changes within the development process is hardly given, provided that a calculation accompanying the development is possible at all. The transparency of information about the cost development in different departments during the development progress is lacking. During development it cannot be ensured that target costs will be met on the market. The consideration of changing costs during HIPD in relation to an integrated product and production process development is not given. Therefore, an approach for agile cost prognosis for the industrialization of highly iteratively developed physical products is required.

### 3. Influencing factors and requirements for the transfer from highly iterative and integrated product and process development to agile cost analysis

To overcome the conflict that classical cost analysis poses to highly iterative and integrated development, a concept for an agile cost prognosis is derived. Hence, it is crucial to identify influencing factors for product and production process-related costs to transfer the integrated development approach to an agile cost prognosis. Figure 2 illustrates the integration of HIPD and highly iterative production process development on the left side. The integrated product and production process development is divided into five parallelized phases. Each phase decreases the freedom of the production process and increases the maturity of the product. The central cost drivers of each phase were derived by literature and are allocated on product, production process and integrated side. The derived cost drivers are analyzed and their interdependencies are illustrated on the right side of the figure. The related cost drivers are listed on the different axles. The spanned planes create tables, in which the correlation of two cost drivers is shown. If cost drivers do not influence each other, or the correlation is so low that it is not noteworthy, the cell stays empty. If they do, a “+” marks the influencing correlation.
In the first development phase the user stories, the production targets and the key performance indicators are defined. User stories are short descriptions regarding product functions in specific use cases. The main cost drivers in this section are the number of requirements, which need to be defined on product and production process development side, since complex systems cannot be handled by existing requirement engineering methods and up to 50% of all errors are based on incorrect requirements. Therefore, it is urgent to combine the traditional and agile way of handling requirements – reducing the number of requirement errors and detecting errors immediately. The cost driver regarding the integration of product and production process development is the number of development steps from the product idea to the successful placement of the product on the market to have a full overview on the end-to-end-process. This includes processes such as homologation and certification, which are cost-intensive but are not considered in current cost accounting approaches.

In the second phase design prototypes are developed on product side while in process development the value-added depth on module level is defined. Cost drivers are the number of prototypes on product side, the number of strategic partnerships on production process side and marketing on an integrated perspective. Marketing is necessary to get in contact with possible partners. The cost drivers influence each other, because the number of needed strategic partnerships affect the amount of necessary marketing activity, which influences the number and kind of design prototypes that are needed.

In the third phase functional prototypes are developed and the automation level is specified. Main cost drivers are the change fees of already manufactured prototypes as well as the number and impact of the change requests. The number of change requests is influenced by the requirements. The less requirements exist and the better these are defined, the less change requests will occur. Likewise the change fees and the impact of the change requests are effected by the number of requirements.

In the fourth phase technical prototypes and the value-added depth on parts level are defined. The technical validation is a main cost driver because it can include complex testing scenarios that take a long time and can destruct one or more technical prototypes. The dimension of the infrastructure includes i.e. the investment in new buildings and factories and is effected by the number of process and product requirements, the number of strategic partnerships on module and on parts level. Tooling costs are a critical investment and affected by product requirements.

In the fifth phase preproduction prototypes are manufactured and production is set up. Cost drivers are manufacturing costs, the ramp-up and the initial operation, which are strongly correlated. The ramp-up is effected by the number of strategic partnerships, since the risk of a supply shortfall increases with a smaller number of partnerships. Also, the manufacturing costs are dependent on the infrastructure and the tooling costs and the number of strategic partnerships. The initial operation has beside the named ones mostly correlations to other process cost drivers.
With the identified cost drivers and their interdependencies, the foundation is built for a concept for an agile cost prognosis. To transfer the characteristics of highly iterative and integrated product and process development to agile cost analysis, requirements are derived. The continuous integration of the customer is crucial to ensure the target orientation of the development task. Bidirectional information asymmetries that occur with an integrated development approach need to be overcome with a complete cost calculation from the beginning of the development. Through the collaboration of the development tasks synergies in both directions are identified and proactively addressed to either optimize the costs or to ensure that the target costs are not surpassed. The cost prognosis needs to be based on interdisciplinary teams that involve the production and related departments from the beginning. Especially including all emerging costs in an end-to-end consideration of the processes needs to be ensured. With this, the customer focus is integrated in the cost prognosis. To be able to track the cost status quo continuously in the development progress the costs need to be defined completely based on the BOM in the beginning.

4. Concept

To successfully place a new product on the market, the speed with which an invention is developed into an innovation and its target costs on the market are decisive. [7] To do so, a concept for an agile cost prognosis using data based and integrated collaboration in the development is derived. During the integrated development of the product and production process a continuous cost tracking poses a central lever to enable the highly iterative development approach including the customer. Using the infrastructure of the Internet of Production (IoP) the databased collaboration between the user, development and production cycle is enabled. [25] In the user cycle customer requirements are continuously identified and translated into User Stories. With these, target costs are derived using the product structure that states the input for the development cycle. The total costs are defined highly iteratively in a loop with the production cycle in which process, procurement and plant regarding costs are specified in an end-to-end consideration. Therefore, a holistic but also agile cost calculation approach is enabled. Changing customer requirements and product specifications are continuously included through the infrastructure of the IoP. Complex and bidirectional information asymmetries coming from the collaboration of the parallel product and production process development can be cured and transparency of the cost analysis is ensured.

In the following, the concept for an agile cost prognosis is presented as a framework within the integrated and highly iterative product and production process development, as shown in Figure 3.
Figure 3: Concept for an agile cost prognosis in the industrialization of highly iteratively developed physical products

The concept is derived to be used in the scope of the integrated and iterative product and production process development, which is characterized by complex dependencies and bidirectional information flows between both development strands. At the same time the voice of the customer is of particular impact to become successful on the market. This is linked to high uncertainties within the development of products including their industrialization. Therefore, the continuous cost analysis and prognosis is crucial in order to create transparency on the possibilities and decide on changes within development. The previous identification of product and production process cost drivers as well as their interdependencies is a supporting tool throughout the progress within the concept.

In the first step the total costs are calculated. Therefore, user stories are created using the data from the user cycle. In these, the required functions coming from the market are determined. To be able to place the innovation successful on the market, the price the customer is willing to pay needs to be calculated. Therefore, the target costs are calculated using the functional cost analysis top down in the first place. The functions relate to the functional structure of the product and enable the product and production development to think in new structures. This is especially relevant when developing high technology innovations that are accompanied by uncertainties. Thus, the functional structure of a product is directly linked to the product structure (cf. step 2).

Through the highly iterative and integrated product and production process development functions can change throughout the development process depending on changing customer requirements or production processes. Therefore, the functional cost analysis is specified in each iteration by getting deeper in the product structure including all production process related costs. After the first prognosis of the target costs is carried out, the target costs for the product are completely calculated. Further on, it is necessary to continuously prognosticate all costs that occur, not only focusing on functional costs. For this, an end-to-end approach using process cost accounting and TCO analysis is executed. Hence, all emerging processes, i.e. certifications of processes and products or testing of products, can be integrated from the beginning of the analysis. The total cost table is completely filled in each iteration at different levels of maturity. The target costs are extended with all process relevant costs as well as use-related costs. At any time the total cost table needs to be filled out completely in every position. In order to continuously monitor the degree of maturity of the cost prognosis and thus make decisions for or against changes, each cost estimate is linked to a Harvey Ball evaluation. An empty ball represents a simple estimate, a quarter ball is a researched value from a
desktop study. Half a ball gets cost statements, which could be obtained by an indicative price offer, a three-quarter ball by a customer specified offer. A full ball is attainable at a negotiated or paid price.

The full cost table is specified with each iteration and is dependent on the second step, the complete determination of the BOM. Starting with the product structure that is derived by the functional structure (cf. step 1), the highest level of the BOM is determined. To enable the integrated development, the progress relies on a hypothesis-based approach. [7] In this step only the currently prioritized BOM is considered. It is crucial to first ensure the completeness of the BOM on a high level than specify parts that are not relevant in this early stage of development. Therefore, all integrated functions in the user story need to be fulfilled by the product structure. Having a complete product structure that is fully priced (cf. step 1), the foundation for the target driven development is defined. After the completeness on a high level is achieved, all positions are specified within the development progress. To do so, the product structure needs to be broken down into modules, assemblies, components and parts and developed bottom up. Since from the beginning of development the BOM is completed on level 0 and combined with the total cost calculation from step 1, a complete priced BOM represents the reference for the further development. After this, all changes are treated as deviations (cf. step 3).

Having the reference BOM that is complete and priced at all times with different maturity degrees, the actual costs during the complete development task can be tracked and compared with the planned costs in the third step. The continuous cost tracking makes the bidirectional connections between product and production process development transparent. These connections occur due to the integrated development. [7] Affected product parts are identified in the event of changes and linked to the respective production processes.

In addition, the evolving structure is continuously evaluated with regard to change effort. It is assumed that a product consists of several modules (M) and individual parts (P), which are manufactured in production processes (PP). With regard to prospective digitalization, the dependencies are represented by matrices. First, the dependencies between the modules need to be determined and marked with an "x" in the matrix. The number of dependencies is a measure for the possible change effort. The more dependencies a module has, the higher the potential change effort. A change of the module can lead to the compatibility to other modules no longer being given, which leads to additional changes. Furthermore, the planned and actual costs are assigned to the individual modules from the total cost table (cf. step 1). The use of colors indicates whether there is a need for change or not. In Figure 3, the total actual costs are lower than the total planned costs. This implies that, fundamentally, there is no need for change. However, the current total costs are displayed in light grey, which shows that at least one module is more expensive than planned. To find the cause the next matrix shows which parts the module consists of and with which production processes it is planned to be manufactured. Actual and planned costs are assigned to the individual parts dependent on BOM levels and cost maturities from the first two steps. Since the entire module exceeds the plan costs, the total actual costs are colored black. It also shows that P2 is responsible for the cost overrun (colored black). On this basis, it can be decided whether the geometry of P2 has to be changed, alternative production processes have to be used or the part has to be purchased from an external supplier. This decision is supported by the cost drivers on product, production process and integrated side (cf. Figure 2). By analyzing the interactions of the cost drivers assigned to specific development cycles, it is possible to identify previously unrecognized cost reduction potentials. Therefore, through the continuous cost tracking, the actual costs of all product and production process related costs are monitored and compared to the target costs that the customer is willing to pay.

Having identified the opportunities for changes, a proactive counter management process is included as the fourth step. If target costs are exceeded or cost potentials are identified due to the improved transparency, changes must be initiated in a different way. The opportunities of change are therefore clustered in four different management processes with different responsibilities. According to the continuous cost tracking, opportunities for cost driven changes are identified. These need to be simulated within the four steps to
validate if with the planned change the target costs can be achieved. If not another hypothesis-based product
or production process related BOM needs to be considered. Changes that allow target costs to be achieved
must be approved by the respective process owner. The process owner is dependent on the maturity of the
product and the freedom of the process. In the first phases in which the product maturity is low and the
process freedom high, the engineering department represents the process owner to decide on the changes.
This is due to the high freedoms in the production process development. After a certain maturity degree of
the product is reached and the minimum viable product is defined by the functions, the process owner shifts
to the production department. Therefore, all change requests need to be negotiated between the inquiring
entity and the process owner. The change requests are permitted if either the target costs or the customer
benefit is improved, i.e. with lower TCO. If the actual costs surpass the target costs, this concept enables an
active counter management. With the continuous tracking a specific i.e. module, part or production process
that is too high in cost is shown and can be directly linked to the BOM and functional structure that is the
foundation of the target cost calculation. Therefore, the process owner knows which change has to be done
in order to fulfill the customer requirements at any time. Besides the cost driven change opportunities,
product or process driven changes can occur. These change requests origin in the interdisciplinary
development approach, which integrated the product and production process development. Since all
dependencies are transparent and connected to the cost drivers, also potential driven changes can occur.
These changes indicate proactive improvements in costs, even if the target costs have not been exceeded.
This is the result of interdisciplinary cooperation from the integrated development approach.

5. Summary and outlook

In highly iterative product and production process development, continuous changes are used to adapt the
product to the changing requirements of the customer. For an efficient implementation of these changes, an
agile cost prognosis is necessary. To integrate an agile cost prognosis concept into the highly iterative and
integrated product and production process development, cost drivers were derived and their
interdependencies analyzed. This is the foundation for the developed concept presented in this paper.
Through a complete and fully priced BOM the actual costs are continuously monitored. With this, changes
are identified proactively to achieve the customer driven target prices. This concept is applied at the moment
for the industrialization of a fuel cell range extender regarding the commercial proof of concept. The BOM
is completed on level 3 and fully priced with different maturity degrees of costs. The agile cost tracking
enables the integrated development to identify changes that need to be considered quickly to ensure the right
direction of development even though high uncertainties characterize the development progress. Further
research will be needed to specify the concept and to develop methods for the data based implementation.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s

References

AMM 794, 532–539.


**Biography**

**Prof. Dr.-Ing. Dipl.-Wirt. Ing. Günther Schuh** (*1958) holds the Chair of Production Systems at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University, is a member of the board of directors of the Fraunhofer Institute for Production Technology IPT in Aachen and director of the Research Institute for Rationalization (FIR) at the RWTH Aachen. He is the founder of Schuh & Co. Group and the e.GO Mobile AG. Prof. Schuh is a member of several supervisory boards and boards of directors.

**Dr.-Ing. Jan-Philipp Prote** (*1986) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen and Industrial Engineering at the Tsinghua University in Beijing. He is Chief Engineer and Head of the Production Management Department at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University.

**Andreas Gütlaff** (*1989) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University. He is Chief Engineer and Head of the Production Management Department at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University.

**Sven Cremer** (*1988) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University. He is a Research Assistant at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University and Group Leader of the Process Management group in the Production Management department.

**Shari Wlecke** (*1993) studied Business Administration and Engineering specializing in Mechanical Engineering the RWTH Aachen and Industrial Engineering at the Tsinghua University in Beijing. She is a Research Assistant at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University in the Process Management group in the Production Management department.
Methodology And Use Of Variant Fields In Factory Planning

Matthias Dannapfel¹, Sebastian Patrick Vierschilling¹, Florian Becker¹
¹WZL of RWTH Aachen University, Aachen Germany

Abstract
Anticipation capability, responsiveness as well as the ability to deal with changes have become key factors of manufacturing companies. These developments directly affect factory planning. Factories need to provide a constant and responsive adaptability. This paper contributes to the adaption of the planning process with regard to the key factors. Modularization of the production system is one aspect of addressing the described challenges, the planning process itself has to be adapted to modular factory structures.

The question to answer is which variation of production system modules is most appropriate in specific conditions. Today’s planners lack a tool that enables them to make the right decision about the configuration of the factory. This paper proposes a configuration logic for the modularized production system. The logic intends to generate an optimally configured production system. Furthermore the developed Variant Fields offer the possibility to apply the configuration logic within the planning process in practice.

In a Variant Field features of a module of the production system (e.g. room utilisation factor) are related with one another, which span a two-dimensional field. Each axis is divided by different characteristic values of variants of a module. According to the characteristic values along the axles different areas within the Variant Field are determined. These areas are thus linked to the variants of a module. For the configuration of the production system, its requirements in the two features of the Variant Field then are marked in it. The variant is selected according to the area in which the marking falls. In the paper the methodology and the use of Variant Fields will be explained. In order to do so, the configuration of logistic systems is used as an example for illustration.

Keywords
Modular factory; factory configurator; requirement-orientated design; productions systems; logistics systems, logistics planning

1. Introduction
Planning processes are characterized by a high number of challenges and requirements. Efficient and productive processes are a precondition for successful planning. This paper proposes a method to better fulfil this precondition for a specific planning domain through the use of the modular factory concept in planning. This domain is factory planning that is facing the challenges mentioned before. However, the proposed method is designed in a way that it can be applied to various planning activities.

Factory planning is strongly influenced by an increasing dynamic and complex environment. Today, manufacturing companies are confronted with challenges like shorter product lifecycles, higher variance, innovation pressure and dynamic change processes [1], [2]. Thus, anticipation capabilities, responsiveness as well as the ability to deal with unexpected changes have become key factors for success [3]. Factories

DOI: https://doi.org/10.15488/9652
need to provide a constant and responsive adaptability [4]. Therefore, the time needed to plan the factory is a strategic success factor for manufacturing companies [6]. The time needed to plan the factory is influenced by the complexity of the planning project. In order to reduce the complexity and create mutable structures within the factory a modular factory structure is already developed by researchers [7], [3]. This leads to a significant reduction in complexity, which in turn reduces the planning and implementation costs of planning projects, shortens planning and implementation cycles and improves the general adaptability of the factory. The benefits of such a modularisation of the factory design can, however, be substantially increased. Today, the planning process itself does not comply with the challenges and requirements factory planning is facing. Thus, time goals cannot be achieved. [8] Therefore, the planning process itself needs to be adapted in order to keep up with the standardization and modularization of the factory. This paper contributes to the necessary adaption of the planning process.

Currently, there is a lack of methodology in order to decide how to configure a modular factory. The question to answer here is which variant of production system modules is most appropriate in specific conditions. This paper proposes a configuration logic as a part of a modularized planning process. As an exemplification the configuration of logistic systems is used as an example. The developed methodology for the planning process enables optimal results by the existence of defined decision guidelines. The configuration logic intends to produce an optimally configured system or subsystem for an individual application based on given input information. The developed application of the methodology is called Variant Field, which is a two dimensional graph, and represents its final result. They will be described in Chapter 2.2.

The Variant Fields and the corresponding methodology propose a solution to the above described challenges. By usage of Variant Fields, the planning efficiency and the planning effectiveness should be increased. While a specific area of factory planning is chosen for exemplification, the above described challenges are valid for all kinds of industries where planning processes take place. Variant Fields can be used whenever there is the need to choose between different variants based on different factors. Thus, the proposed method not only contributes to the efficiency and effectiveness of the factory planning process but also to efficiency and effectiveness of planning processes in general.

2. Methodology and use of Variant Fields

In order to introduce the methodology and use of Variant Fields, an overview of the configuration logic is given first, with logistic systems as an example. The Variant Field and its structure is introduced afterwards. The second part of the chapter deals with the development and application of the Variant Field.

2.1 Configuration logic

The Variant Field is a part and the result of the followed explained configuration logic. The configuration logic enables the configuration of factory elements in terms of their variants with regard to the requirements and operating conditions of the factory. To develop the configuration logic, all features of a factory element are identified and evaluated, in order to find out which features have to be changed by configuration to meet the requirements. The variants of a factory element are then identified. Since a variant of a factory element may not be compatible with or require combination with a particular variant of another factory element, the creation of combination rules is required. These show mandatory and optional relationships that must be taken into account during the configuration.

A factory element is a subsystem of a factory itself. A factory element is characterized by different features. Additionally, a factory element can be realized through different variants and extension modules. The term variant is used for each design or alternative solution of a factory element, which differs in at least one qualitative or quantitative parameter. An extension module is a self-contained unit that provides the factory...
element with additional or better features or functions. It can also influence the combinability of a factory element with other elements (see Figure 1).

The main difference between an extension module and a variant is that variants have to replace each other, whereas also several extension modules can be used in parallel. Moreover, two types of extension modules exist. General extension modules, on the one hand, are independent of the variant and can therefore generally be used for a factory element. These modules are regarded as cross-variant. Variant-specific extension modules, on the other hand, can only be used for one or more specific variants. Figure 1 shows a schematic variant structure of a factory element. Variants and extension modules are characterized by specific characteristic values of the features of factory elements.

![Figure 1: Schematic structure of a factory element](image)

The above described terms are illustrated in the following example. Typical features that characterize a storage equipment are, among others, their height, their room utilisation factor and their assortment size. The values for the room utilisation factor, for example, are reaching from low to high. Possible variants of storage equipment are a floor storage, a pallet rack and a high level rack. Examples for extension modules are an automated booking system for entry and exit of products or an anti-theft protection for very valuable parts. It is also possible to quantify the characteristics. To keep the example in this paper simple, qualitative values are used to explain the methodology. So a floor storage has a low room utilisation factor, it can be characterised as middle for a pallet rack and as high for a high level rack.

The basis of the configuration logic are different factory elements. Important factory elements can be identified using a literature analysis. The identified elements are then clustered into subsystems in order to increase the transparency. As an example, Figure 2 shows an overview as part of the results for logistic systems.

![Figure 2: Overview of factory elements of logistic systems](image)
The elements can then be categorised in order to specify how the particular element is treated within the configuration logic. The factory elements are assigned to one out of three classes. These classes are individualization, modularization or standardization. A factory element in the class individualization is planned and designed individually. It is therefore not suitable for a modular system or a configurator and will not be considered further. Factory elements in the class modularization or standardization are relevant for the configuration logic. Factory elements in the class modularization can be partially standardized. Due to different demands on the factory element, different variants of the element should be used in the factory. So they are configured according to specific requirements for each individual area of the factory in which the element is to be used.

Factory elements which are categorized as standardization can be fully standardized across the entire factory. This elements exist in the entire factory in only one variant. They are rigidly dimensioned for the maximum design requirement with a view to its long-term development.

In order to assign the factory elements to one of the classes, the factory elements are generally assessed in terms of the interactions with the product, the influence on the Return on Planning (ROP) [9] and the module suitability. The assessments are then transferred to a portfolio from which the classification can be derived. To evaluate the interactions with the product and the influence on the ROP different key figures are calculated. A standardized factory element should have a high influence on the ROP with the least possible interaction with the product. The suitability examines whether a factory element is principally suitable for standardization. This is done by using module drivers for the evaluation.

The underlying assumption when using the configuration logic is that there is an optimal configuration for each factory element for a certain operating point in the factory. This optimal configuration will be identified by using the configuration logic and the Variant Field. In terms of the above described example, the configuration logic and the Variant Field will answer the question which storage equipment shall be used in the factory.

The configuration logic is generated by using a four-step approach. First, relevant features of the factory element are identified. However, the most important aspect is to filter the most relevant features from the total set of identified features. Afterwards, these features are collected in a feature record for each factory element. The selected features should express the performance of the factory element, meaning that the requirements to the factory element can be expressed directly by its features. Features that only indirectly describe performance are not included in the feature record. This also includes features from which other features can be derived. In addition, the selected features must be communal so they can be applied to all variants and extension modules of the respective factory element. For example, the feature room utilisation factor is relevant for all of the above described variants (floor storage, pallet rack and high rack).

In a second step, the variants and extension modules of that element are defined. At first, all identifiable Variants are generated by analysing the state of the art. Those variants that are suitable for use in the existing factory are selected from the variant pool, based on the specifications of the products manufactured in the factory and the structural conditions of the factory.

The third step deals with combination rules of variants and extension modules. As a variant of a factory element may not be compatible with or require combination with a particular variant of another factory element, the creation of combination rules of variants and extension modules are required. Combination rules define mandatory and non-relationships or bids and prohibitions when combining certain variants and extension modules of different factory elements. For example, only a narrow aisle high-bay warehouse can be served with a suitable forklift, which does not fit to any other rack variant. Once these steps have been terminated, the Variant Field can be set up as the fourth step of the methodology, which is described in more detail below.
2.2 The Variant Field

The Variant Field is the key element of the configuration logic that, once set up, enables the fast and optimal planning of a factory element, self-sufficient of location and requirements. It visualizes the configuration logic and ensures its applicability. The goal of the Variant Field is to decide which variant of an element shall be chosen based on specific boundary conditions. Before describing how the Variant Field is developed its structure will be introduced first.

In the Variant Field two features are related with one another in a two-dimensional field. Each axis represents one feature and is divided by different characteristic values of the feature, defined by the variants of the factory element. These characteristic values can either be qualitative or quantitative (which will be further detailed in chapter 2.3). Moreover, the relevant extension modules with the areas along the range in which they shall be used are also marked along the axes. Depending on whether it is a cross-variant or a variant-specific extension module the nomenclature is different (see Figure 1). While a cross-variant extension module is called EM 1, a variant-specific extension module contains the variant in its name, for example EM A.1, for an extension module used for variant A (see Figure 3).

![Figure 3: Structure of the Variant Field](image)

The area in which an extension module is to be used is marked by two arrows. The arrow pointing to the inside of the field indicates the point from which the module is used, the arrow pointing out of the field indicates the point to which the module is used. Whereas there does not have to be an arrow pointing out. According to the characteristic values along the axis different areas within the Variant Field are determined. These areas are linked to different variants of an element. In order to determine which variant should be chosen, the requirements for the element needs to be defined related to the features. Once this is done, the defined requirements for both features are marked along the axis. The point of intersection marks the operating point. The operating point is located in the area of a certain variant and is used to identify the best variant and the necessary extension modules for the given requirements.

2.3 Development of the Variant Field

The central challenge in developing the Variant Field is the consistent integration of elements, their variants and characteristics. Therefore, two mutual complementary or competing features have to be examined. Two features can be indifferent, complementary or competing. Indifferent features do not influence each other. For complementary features, the increase of one feature leads to the increase of the other feature, while for
competing features it is the other way around; the increase of one feature leads to the decrease of the other feature. In case of a complementary relationship, the relation can be either symmetrically complementary or asymmetrically complementary. Two features are symmetrically complementary if the increase of feature A leads to an increase of feature B and the increase of feature B leads to an increase of feature A as well. Symmetrical complementary features can be eliminated, when this relationship is valid over the entire value range, thereby reducing the number of features.

Once this is done, characteristic values of a feature are determined which is a decisive step for the configuration logic. To configure the factory elements, the characteristic values for each variant and extension module must be known in order to be able to adapt the factory element to the given requirements. In this step, variants and extension modules are assigned to a characteristic value of each feature. In this way, the different characteristic values indicate the limits to which or up to which a variant or extension module can be scaled.

As already mentioned above, the range of the characteristic values can either be quantitative or qualitative. A qualitative range can be developed by using eligibility criteria or expert interviews. As an example, a qualitative range can contain the values low, middle and high, as used in the example before. Moreover, a qualitative range can be used when indicating the presence or absence of a certain characteristic. The range then contains the values yes and no. In general, qualitative scales are used as guidelines for the configuration. A quantitative range allow an objective configuration of the factory element. Therefore, quantitative scales are preferred. The scale can be linear as well as logarithmical. An example for a quantitative range is the storage volume per storage area, which is measured in cubic meters per square meter.

Figure 4 shows examples for both, a qualitative and quantitative range of characteristic values. Moreover, extension modules are marked with regards to their range for each feature. This is realised by arrows. For cross-variant extension modules (see EM 1 in Figure 4) both, the starting and ending point of the area in which this module is to be used, needs to be determined. Starting points are marked with an arrow pointing upwards, while ending points are marked with an arrow pointing downwards. Whereas for variant specific extension modules (see EM B.1 in Figure 4) their range never extends beyond the range of the corresponding variant and therefore, only the specification from which the module is to be used must be determined. Finally, these results are visualized, i.e. the Variant Field is generated.

For complementary features, the values along the axes have to be arranged in an increasing way from the intersection of both axis. For competing features, one axis needs to be increasing while the other needs to be decreasing. For indifferent features, there is no general rule nor guarantee, that an arrangement of axes can be found to generate a consistent Variant Field why indifferent features should be avoided.
3. Application and benefits

A fictional example for a specific Variant Field can be seen in Figure 5. It was developed in order to configure the storage equipment mentioned above. The identified relevant features are the assortment size and the room utilisation factor [4], [10], [11]. Both are described with qualitative ranges. The relevant variants are floor storage, a pallet rack and a high level rack. As seen in the Variant Field a high assortment size and a high room utilisation factor are realized through a high level rack storage, whereas a low assortment size and a low room utilisation factor are realized through a floor storage.

In this example the company for which this Variant Field has been generated, has grown over decades and has no capacity reserves in terms of area and space. Due to the lack of space, the requirements to the space utilization level of the warehouse are very high. However, the requirements for the product range that can be stored are medium to high. The optimum configuration of the storage is derived from the Variant Field based on the described requirements. The operating point is located in the area of the variant high level rack storage. For the company it is therefore recommended to use this variant.

The underlying goal of the configuration logic and the Variant Field, as described above, is to increase the efficiency and effectiveness of the planning process in order to increase flexibility and responsiveness. The variant field is particularly suitable for rapidly changing boundary conditions for the selection of variants in the factory, whereby the possible variants remain largely the same. Thus, the new requirements to the features only have to be transferred to the variant field and a new variant has to be read, instead of having to carry out a new technology and variant search with the new requirements.

The rise in efficiency or in effectiveness result in a greater ROP value. Moreover the Variant Field offers the possibility to practically apply the generated configuration within the planning process. The visualization of different options and their coherence increases the understanding of complex decision aspects. The presentation is clear and easy to understand. The Variant Field is universally applicable. It can be used in various industries and planning scenarios, since the displayed variants as well as the characteristics of a feature can be adjusted. Moreover, the Variant Field can be used both in initial planning processes as well as adjustment processes.

4. Discussion

Variant Fields allow planners to derive the optimal variant of a factory element directly from the requirements to the element, described by its variable features. The Variant Field as a key result of the
application-oriented configuration logic enables the simple and qualitative or quantitative configuration of factory elements. It thus speeds up the planning process and increases the ROP through reduction of the planning costs. Moreover, it increases the planning reliability. The Variant Field supports the change from a one-time, project related planning to a continuous planning throughout the factory lifecycle. By using the Variant Field, elements can be planned with less effort and easily be adapted to changing boundary conditions. In addition, the Variant Field contributes to the modularization of the planning process by identifying elements suitable for modularization and standardization. The knowledge gained and the generated Variant Fields can be used across projects to further reduce the planning effort. Moreover, the intended decision support for the planner is fully realized through the Variant Field. Finding the optimal configuration with the Variant Field requires no knowledge about the background of the configurator or the factory element. Thus, short training is needed for employees before using the Variant Field. This makes the tool interesting for a variety of applications.

Despite these benefits there are also some limitations when using Variant Fields.

An underlying assumption when using the Variant Field is that an optimum configuration exists for each element for a specific operating point. By using the Variant Field only the characteristics of the elements but not their required number are determined. Practical application has shown that it is difficult to make a clear distinction between features of a factory element and to define influences on an element and its boundary conditions. In addition, the development of a Variant Field requires a great amount of information that is not necessarily available immediately. The evaluation of the features remains a subjective process and is thus subject to uncertainties. The selection of variants and their assignment to a characteristic value can lead to ambiguities or contradictions in complex application examples, which contain, for example, more versatile requirements in the factory or overlapping characteristics of variants. In addition, the exact determination of the characteristic values, on both quantitative and qualitative scales, requires high effort. Furthermore, combination rules can only be utilized if all variants and extension modules of all factory elements, used in the configurator, are known. This circumstance causes high efforts for the complete development of a factory configurator.

5. Outlook

Despite the various benefits, further research needs to be done in the future. The design of the configuration logic is to be expanded by defining the rules to select the variants used in the configurator from the state of the art. Furthermore, the different areas of the Variant Field, which are linked to the variants and extension modules, are separated by non-linear boundaries in the case of complex interactions of the features. Their definition must be examined. The dependency of features also lead to non-linear functions describing the areas of the Variant Field. The relationships between the axes need to be analysed in more detail in order to be able to set up the Variant Field. Moreover, a clear distinction from and up to which point a variant shall be used should be worked out. Nevertheless, the Variant Field is a strong tool that fulfils the needs of today’s planning processes. Once their development has been described in more detail, nothing stands in the way of their widespread application.

References


Biography

**Florian Becker** (*1992) studied mechanical engineering and production technology at RWTH Aachen and received an award for his master's degree and is member of the Dean’s list. In 2017 he started his doctorate at the WZL Aachen. Currently, he works as a member of the research group ‘Digital Factory Planning’ researching and developing integrated factory modelling processes as well as automatized planning procedures. His main areas of research are automation of planning with the help of AI, regarding the matching between planning tasks and automation algorithms.

**Dr. Matthias Dannapfel** (*1986) is Chief Engineer of the Department of Factory Planning at the Chair of Production Engineering within the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University. During his research he focuses on the topics Factory of the Future as well as automated and integrated factory planning processes. Furthermore, he is CEO of the ONEFACTORY GmbH.

**Sebastian Patrick Vierschilling** (*1990) is research associate at the Department of Factory Planning at the Chair of Production Engineering within the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University. Currently, he works as group leader of the research group ‘Digital Factory Planning’ researching and developing integrated factory modelling processes as well as automatized planning procedures. Beforehand, he finished a Double Master Degree Program within RWTH Aachen University and Tsinghua University, China.
Abstract
The analysis and further the reorganization of an order-production is a typical scope of application of a value stream mapping. Value stream mapping is a lean-management method to map the current state of a series of processes that are necessary to manufacture a product or provide a service. The ever-increasing digitalization emphasizes the importance of the information flow. Besides the material flow, the information flow is the second focus of value stream mapping. It is needed to get valuable insights into the production process by applying state-of-the-art data analytics methods. Process Mining is a possible method to analyse (business) process and process sequences based on event logs. This paper illustrates a method of combining conventional value stream mapping and Process Mining. While value stream mapping shows the material and information flow as the people think it is, Process Mining shows how they are, based on the data fingerprint. The comparison of the two outcomes allows conclusions for the following value stream design with a special focus on the use of data for Industry 4.0 applications. The application of both methods to an order-production enables to present results and compare both as well as present advantages and disadvantages.

Keywords
Value stream mapping (VSM); Process Mining; Lean 4.0

1. Introduction
The improvement of processes to increase efficiency is an ongoing goal in companies. Lean Management is a well-known and established philosophy to achieve that. For this purpose, value stream analysis is one of the best-known methods.

Industry 4.0 pursues, among other goals, the same goal of increasing efficiency and consequently value competitiveness. In the last years, there was an increased focus on how Industry 4.0 can help to achieve these goals. The major question is how the principles of Lean Management and Industry 4.0 support and substitute one and another. Consequently, one may be able to use Industry 4.0 and its tools to enhance the tools of Lean Management.[1]

The flow principal and one-piece flow are fundamentals of Lean Management. To enable them the value stream mapping is an important tool, to identify, analyse and improve the value stream. It is done by hand and is therefore rather time-consuming to be done on a regular basis.[2] Using principles of Industry 4.0 it was tried to streamline the process of the analysis and mapping of the value stream.
Agility and flexibility are besides increasing of efficiency a major factor to push Industry 4.0. It uses the possibilities of digitalization to satisfy these goals. The increased demand for custom-made products to low prices encouraged companies to further diversify their portfolio with sheer endless variants. This confronts the planning and production with the challenges of lot-size-one production and one-piece flow. The horizontal and vertical integration of IT-Systems and the followed increased amount of data enables new possibilities to make a lot-size-one production cost-efficient. A wide variety of data analysis methods are the tools that help to tackle the tasks at hand in the future.[3] Process mining is a relatively new method to map and improve processes based on the traces of data [4].

In this paper, a method is presented to combine Process Mining, as a method possible through the advances of Industry 4.0 and value stream mapping, as a method out of the toolbox of Lean Management. The case study is the production of a contract manufacturer. They are faced with lot size 1 manufacturing ever since. A highly variable product mix leads to a constantly changing main material flow. To analyse and map the flow to improve upon it or use it for better planning is a near-impossible task to perform by hand. Therefore, the combination of Process Mining and value stream mapping is presented to remedy the disadvantages of the classical value stream mapping.

2. Theoretical Background

2.1 Value stream mapping in Order-Production

The idea of lean manufacturing, i.e. structuring along the value stream and minimizing waste, has found its way into many companies. Value stream mapping (VSM) is a widely used method to support these actions. Material and information flows are represented in a standardized way, the potential for improvement can be derived, and the target state can be developed.[5] Originally, the conventional VSM was designed by Toyota for series-production[6] and the focus is on the series production representing logistical targets, e.g. the execution time, process time, set-up time, cycle time, machine reliability, and inventory.[5] The timeliness is neglected, although most users see their main goals when applying the method in the recording of the lead time and the timeliness.[6], [7] Since in order-production customers are directly influenced by delay in delivery, timeliness is an essential logistical performance indicator.[6] Another difference is the order of process steps. While a series-production is rigid, an order-production can be flexible and versatile. Therefore, KOCH developed a VSM method, especially for order-production. Timeliness, as well as the order of production, are assessed next to the logistical performance indicator known from the conventional VSM. The method is going to be conducted in 4 steps, shown in Table 1.[6]

<table>
<thead>
<tr>
<th>Step</th>
<th>Result</th>
</tr>
</thead>
</table>
| 1 Preparation | – Selection of an area  
– Work schedules and performance data of the work systems are determined |
| 2 Production tour | – Logistical performance indicators and influencing variables determined  
– Information about production planning is recorded |
| 3 Data integration | – Material flow is shown  
– Information flow is shown |
| 4 Analysis of fields of action | – Logistical field of actions are revealed |
In step 1 “Preparation”, the production area in which the analysis is to take place is defined. Production data of completed orders are needed to determine the average performance in the number of orders per operating calendar day and the average performance in the number of hours per operating calendar day. Furthermore, a product-work system-matrix matching workplaces with products is going to be set up. While the other steps are location-independent, step 2 consists of a production tour. The data collected during the tour can be divided into three categories: (1) Information about work system, (2) logistical performance indicators with their influencing variables, and (3) planning and control information. These include for each work station e.g. the number of machines (1), the used shift system (1), inventory (2), lead time (2), estimated schedule variance due to backlog (2), sequencing (3), and capacity planning (3). Order acceptance, as well as order release, is information, that is captured production wide. Step 3 uses the information gathered in steps 1 and 2 to provide a value stream model. A product family is selected to represent the material and information flows, which are almost like a conventional VSM. The conventional VSM is enhanced with the logistical performance indicators, so timeliness, sequencing, and estimated schedule variance due to backlog are represented. In step 4, the newly created value stream card is examined, as well as order release, for possible fields of action. To find the field of improvement, KOCH states 6 questions regarding order release, inventory, sequencing, order control, and capacity planning.[6]

2.2 Process Mining

Process mining builds a bridge between traditional model-based process analysis, like the visualisation of business processes, and data-centric analysis techniques, like data mining/machine learning. Process mining utilizes event data to be able to provide insight, such as identifying bottlenecks, anticipating problems, and streamlining processes. Based on the event log and additional information three types can be carried out: [8]

- Discovery: A model is build based on the event log without a priori information.
- Conformance: A model of an already existing process is compared with the recorded event log of this process. Conformance checking is used to verify, if the recorded event log, reflecting the reality, coincides with the process model.
- Enhancement: The aim of this third type is to change or extending an a priori model.

Events include “life events”, “machine events” or “organization events”. The application of Process Mining assumes that the events considered can be recorded sequentially and that each event refers to a specific process activity and a specific process instance or case. For the use, the following data is needed or optional for each event:[8]

- Case ID – mandatory: identifies an order through the data.
- Event ID – mandatory: identifies the consecutive production steps in the data. They are linked together by the Case ID.
- Timestamp – mandatory: Gives information about the start, end and hence length of a production step.
- Activity – optional: identifies the work in each activity.
- Resource – optional: identifies the used resource in a production step (e.g. machines or employees).

3. VSM with Process Mining in Order-Production

3.1 Approach

As previously described, VSM is a standard tool to improve one’s production process, but it is also quite time-consuming and in an order-production tricky to execute. Based on the functionalities Process Mining offers, the intent of this paper is to show that Process Mining is able to conduct a VSM based on historical data containing production information of completed orders (case ID, event ID, resource, timestamps).
Furthermore, both models of the production shall be comparable to see shortcomings between reality recorded by VSM and by IT/Process Mining. Two types of Process Mining are considered to be used, to generate a model comparable to VSM. As a first step, “Discovery” ensures to visualize a model of the production steps and their sequence. After that, the “Enhancement” gives two possible perspectives. The first one is the possibility of extension, like adding KPIs to the discovered model e.g. lead time per process step. The second perspective is repairing, which helps to modify the model to reflect the reality of the data.[4] In Table 2 the already presented Table 1 is enhanced by Process Mining. In each step, Process Mining is able to support a VSM as well as to conduct to a certain extent. Subsequently, the conformance check shall be done manually since it’s a comparison of a VSM model and the Process Mining model.

Table 2: VSM in order-production enhanced by Process Mining

<table>
<thead>
<tr>
<th>Step</th>
<th>VSM in order-production</th>
<th>Process Mining</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Preparation</td>
<td>Selection of an area.</td>
<td>Done by the selection of the historical data.</td>
</tr>
<tr>
<td></td>
<td>Work schedules and performance data of the work systems are determined.</td>
<td></td>
</tr>
<tr>
<td>2 Production tour</td>
<td>Logistical performance indicators and influencing variables determined.</td>
<td>Logistical performance indicators can determine based on historical data.</td>
</tr>
<tr>
<td></td>
<td>Information about production planning is recorded.</td>
<td></td>
</tr>
<tr>
<td>3 Data integration</td>
<td>Material flow is shown.</td>
<td>Material flow can be generated by using Process Mining.</td>
</tr>
<tr>
<td></td>
<td>Information flow is shown.</td>
<td>Most frequently used material flows are shown independently from the product family.</td>
</tr>
<tr>
<td>4 Analysis of fields of action</td>
<td>Logistical fields of action are revealed.</td>
<td>Logistical fields of action are shown due to the interpretation and key figures of the model.</td>
</tr>
</tbody>
</table>

3.2 Case Study

The case study was performed in a company that produces different types of test facilities for combustion engines. Due to a rapid increase in demand for test facilities in the past, certain issues in the process stability became obvious. Its most apparent manifestation has been shown in the declining of the delivery reliability. Since prognoses for future demand show an increase as well, it was decided to improve the production process. Based on these facts, the decision was made, to conduct a VSM. Since additionally, digital event logs were available, they should be analysed as well.

Starting with the preparation phase of the VSM for order-production, historical production data was analysed to extract KPIs. The provided data was organized as an event log that is used in Process Mining. Therefore, it was obvious to try to use Process Mining, to shorten the time to do the on-site VSM and to gain additional insight.

The following data were available:

- Order ID which represents the test bench. It matches the intent of the case ID in Process Mining
- Production ID is to identify the different production orders to produce the given test bench. It corresponds to the event ID in process Mining.
- The Position ID sorts the activities of the production order in accordance with the indented production flow. Combined with the Production ID, it makes a unique identifiable event ID.
- Start date and end date of the activity in the event. They are used for the timestamp information.
- The work description is equivalent to the activity in Process Mining and informs about the performed work in the event step of question.
- The workplace ID and the workplace name go hand in hand with the resource ID in Process Mining.

As mentioned above the original intent was to use Process Mining in the sense of enhancement and extent to gather the necessary KPIs. Unfortunately, the data was not suitable for a detailed analysis. As mentioned in section 2.2, timestamps are necessary. In this case, case ID (a combination of the order number and production ID) and resource (workplace ID) were given. The timestamps referenced only to the day of the start and end of the mentioned event, not the actual starting and end time in minutes and hours. Additionally, the processing time was recorded. Therefore, it was not possible to bring the records in a definite chronological order. So, it is not possible to extract viable KPIs concerning lead times or execution time. Since defined production processes and flows were not available, conformance checking was out of the question as well.

As described in Table 2, process mining was used in the data integration phase to determine the most common material flow. The event log was coherent enough to show, which workplaces where frequently used and in which sequence. Process mining allowed to identify all predecessors and successors of each workplace from the perspective of the resources (workplaces). As a result, the process map was enormous and after filtering the most essential resources and cases, the basis for the most common material flow shown in Figure 1 was created. In at least 80% of the cases the black marked rectangles/workplaces are a part of the most common material flow. How often an event took that path is shown by the thickness of the arrows. In this case, it was distinguished between often (black) and seldom (grey). Grey rectangles are workplaces that are not that common in the material flow and were deemed unnecessary for the VSM.

![Figure 1: Most common material flow](image-url)
The used sequence was extracted by analysing the process flow map of individual cases. So, it was possible to determine the repeated execution of the work at the assembly and the sequence of loops.

Using Process Mining to extract the most common material flow, was extremely helpful because there are no records of standard products. So, it would be impossible to define which order is produced on which workplace and create product families. We circumvented the issue by using Process Mining and determine the most common material flow that way.

Further detailed analysis revealed that the intended sequence of events and hence activities do not always correspond with the flow of time according to the timestamps. The possible reasons were narrowed down to two different aspects. Firstly, it could be possible that the feedback information of the timestamp was incorrect or secondly, the process steps can be switched. At the on-site VSM, it was realized that the second issue applied in most cases. Therefore, the way of how the process can be executed depends on either the respective supervisor in charge or the responsible worker. Based on his/her experience, they know how the process can be performed in a more efficient way. The identification of this issue is a typical example of enhancement by Process Mining by repairing the process itself.

4. Conclusion and Further Research

The approach of combining and replacing the VSM with Process Mining in this case study didn’t lead to the expected result. But using Process Mining in order manufacturing to get the most common material flow is a considerably useful case of Process Mining applications. Without it would be impossible to define product families and start the VSM, especially in highly diversified make to order or engineer to order-productions. The application of Process Mining also showed that it is not necessarily helpful to form product families. Further research is needed to find evidence to make workflow families instead and how to assign the orders to them. In highly flexible production environments, a classification by product attributes, like already done in, for example, product families, might not be the best thing to do in order to unify the value stream. Table 3 discusses the advantages and disadvantages shown in this paper based on the case study. A combination of both methods leads to an efficient and holistic value stream model of order-production.

| Table 3: Summary of advantages and disadvantages of VSM in order-production and Process Mining |
|---|---|
| **Advantages** | **Disadvantages** |
| **VSM in order-production** | Production tour supports understanding of the production | No insight into the digital footprint |
|  | Contexts that cannot be mapped in the data can be captured | One standardized product (family) is necessary |
| **Process Mining** | The on-site recording is more efficient due to preliminary results | The result depends on data quality |
|  | First results quickly visible | converging material flows are only possible to a limited extent |
|  | No reference product (family) is needed |

Furthermore, the intentional approach should be verified in the future. Studies show that a high level of digitalization correlates with a lean organization. [1], [9] By improving data acquisition, it will be possible to improve the results of the VSM in the future. It is quite common to repeat a VSM every few years. With high-quality data, it is viable to do it automatically with Process Mining, to shorten the effort and therefore
the intervals between analyses. Well-defined and recorded timestamps will make it possible to analyse the different parts of waste in the process and improve further on them.

References


Biography

Katharina Mertens (*1989) studied Industrial Logistics at Montanuniversitaet Leoben, Austria and has been a research assistant at the Chair of Economic and Business Management at Montanuniversitaet Leoben since 2017. Her research areas are production and industrial asset management, combined with data analytics.

Robert Bernerstättler (*1984) was a research assistant at the chair of Economic and Business Management at Montanuniversitaet Leoben, Austria for five years. His research area was the application of data analytics in production and asset management with focus of the effects of data quality. His doctoral thesis has the title “Maturity Model for the Evaluation of the Input Factors for Data Analytic Applications - Conceptualization on the Example of Weak Point Analysis“. Since January 2020 he is working as data scientist for voestalpine Signaling Zeltweg.
Hubert Biedermann (*1953) is professor at the chair of Economic and Business Management and head of the Department of Economic and Business Management at Montanuniversitaet Leoben. His main research areas are asset management - lean smart maintenance, production management, quality and risk management. After his habilitation in industrial economics in 1989, he worked for several years in the industry and as a visiting professor. Appointed Montanuniversitaet Leoben in 1995, he held the position as a vice-rector for finance and controlling for more than 12 years.
A Framework for Data Integration and Analysis in Radial-Axial Ring Rolling

Simon Fahle¹, Bernd Kuhlenkötter¹
¹Chair of Production Systems, Ruhr-University of Bochum, Bochum, Germany

Abstract

Data-driven analytical approaches such as machine learning bear great potential for increasing productivity in industrial applications. The primary requirement for using those approaches is data. The challenge is to not only have any kind of data but data which has been transformed into an analytically useful form. Building upon this initial requirement, this paper presents the current state concerning data analysis and data integration in the industrial branch of hot forming, specifically focusing on radial-axial ring rolling. The state of the art is represented by the results of a data survey which was completed by six of Germany’s representing radial-axial ring rolling companies. The survey’s centre of interest focuses on how data is currently stored and analysed and how it gets depicted into eight different statements. Based on the results of the survey a framework is proposed to integrate data of the whole production process of ring rolling (furnace, punch, ring rolling machine, heat treatment and quality inspection) so that data-driven techniques can be applied to reduce form and process errors. The proposed framework takes into account that a generalized standard is hard to set because of already grown structures and a huge variety of analytical methods. Therefore, the framework focuses on data integration issues commonly found in an industrial setting as opposed to controlled research environments. The paper proposes methodologies on how to utilize the potential of each company’s data. As a result, the proposed framework creates awareness for saving the data in a standardized and thoughtful manner as well as building a data-driven culture within the company.

Keywords

Machine learning; Data integration; Framework; Radial-axial ring rolling

1. Introduction

In times of 33 Zettabytes of digital data in 2018 and a predicted growth to 175 Zettabytes by 2025 referring to a white paper of the International Data Corporation, there seems to be a high demand for analytical approaches to face those enormous amounts of information [1]. In recent years, the usage of machine learning especially supervised methods are heavily researched and used in several industrial solutions. One of the biggest restraints to use and benefit from these techniques is the demand for a lot of data in an analytically useful forms. Today, many companies are already saving a lot of their information and storing it, for example in a huge data lake, but only a few have defined production goals in mind while storing this data. A survey carried out by MicroStrategy in 2018 shows that the majority of the participants state the need to drive strategy and change as their primary focus followed by industrial/manufacturing improvements for data analytics in Germany [2]. This leads to a status quo where data is accessible but often not useful because of several striking issues concerning the form in which the data is stored. In this paper the status quo within the hot forming sector, specifically the radial-axial ring rolling (RARR) is being investigated by displaying...
the results of a data survey which was carried out in 2019 and six of Germany’s representing ring rolling companies which participated in the survey. The results are discussed and a framework on how to store and integrate data the right way for further analysis it is proposed. This framework will be used to analyse process and form errors of the rolled ring (e.g. climbing, non-circularity, fish-tail [4]) using methods of machine learning for time series and sequence data. The high complexity of the ring rolling process and thereby high amount of influencing variables suggest the approach of using machine learning.

The paper highlights common mistakes made during the process of storing data and proposes ways to overcome these issues while presenting a best practice advice for companies in the hot forming industry. The biggest problem concerning this issue are grown structures in each company which makes it hard to propose a standardized framework for every industry. Further, the different data types (structured, semi-structured and unstructured [3]) that arise in the different areas of industry aggravate the challenge of coming up with one standardized way to tackle this challenge. This is the reason why to the best of the author’s knowledge there is no such thing as a sole standard for data integration in ring rolling or similar processes.

The paper starts by focusing on related work in RARR and other fields of production engineering, machine learning for time series and data integration frameworks. It follows up with the introduction to the industrial process of RARR and presents the carried out survey. Based upon this, the survey is discussed and used as a starting point for the proposed framework for data integration and analysis in RARR. The paper closes with a section regarding the desired deployment of the proposed framework in two industrial companies.

2. Related Work

2.1 RARR and production engineering

First approaches of using process data to evaluate the ring rolling process regarding its process and form errors were made by the Lehrstuhl für Produktionssysteme. The potential to classify process-failures, such as ring-climbing, using data mining techniques (fuzzy-logic) has already been shown in 2004 by LAGAO [5]. HUSMANN ET AL. used process-data in form of structured log-files from a ring rolling machine to avoid ovality of the ring [6]. Additionally, HUSMANN ET AL. analysed the ring rolling process further by using image data as well as thermography data in different ways, producing unstructured data regarding the ring rolling process [7,8]. Similar to those approaches for RARR, WANG ET AL. use support vector machines to reach up to 95 percent classification accuracy in equipment failure classification for optical networks [9].

Besides those approaches for process failure prediction, machine learning is used in a variety of applications throughout other fields of production engineering and not all can be addressed during this paper. The paper focuses only on current implementations of predictive maintenance and optical quality inspection in production engineering environments. For example, LEE ET AL. use support vector machines as well as deep learning networks to model the remaining useful lifetime of a machine tool system [10]. In contrast, WÖSTMAN ET AL. present an implementation of a predictive maintenance system using neural networks, decision trees, random forests and other algorithms for a retrofit robotic cell [11].

Another field of operation is optical quality inspection using machine learning models. MAYR ET AL. use machine learning algorithms to classify welding defects for a quality monitoring system [12]. A different approach to analyse computed tomography images using a pre-trained neural network (AlexNet) is proposed by HALWAS ET AL. to validate the layer structure of windings in the automotive sector [13].

2.2 Machine Learning for time series

Regarding current research in the area of time series classification, FAWAZ ET AL. make use of different approaches using Deep Neural Networks [14]. A more specific approach to time series classification is reviewed by SANTOS AND KERN as well as RUBWURM ET AL. with regards to early classification time series
(ECTS), which enable the algorithms to classify process data on-line [15,16]. Regarding the RARR process, faults and failures are heavily under sampled. As a result, He et al. propose an ensemble of classifiers to work on the challenge of imbalanced datasets for ECTS [17].

2.3 Frameworks

To the best of the author’s knowledge regarding modern frameworks for data integration and analysis in the area of RARR no prior research was found which implies the assumption that the presented approach is the first in RARR. Regarding other fields of the industry, there is a lot of related work that has been done on the issue at hand. Much research can be found in the field of medicine and healthcare, for example Peral et al. propose an ontology based architecture to deal with the unique requirements of data in combination with Telemedicine Systems [18]. El Aboudi and Benhlima reviewed popular healthcare monitoring systems and propose an approach which is focused more on the big data aspect of healthcare [19]. Further, Samourkasidis et al. propose a framework for environmental data and discuss different approaches to solve the problem of data heterogeneity [20].

3. Motivation

3.1 Radial-axial ring rolling

Many machines and systems in the aerospace and energy industry need seamless formed rings made of steel, aluminum or nickel, which additionally need to satisfy high standards e.g. highly dynamic load capacity and high variability. A majority of those high-performance rings are being formed through the hot forming process RARR. This process can form seamless rings ranging from 100 mm to 16 m in diameter to a maximum wall height of up to 4 m. The weight can vary between a few kilogram to up to 300 t [21]. Applications of ring rolled rings can be found in the automotive, the aerospace and rail traffic sectors. The importance of RARR within Germany is depicted in an Euroforge-survey from 2016, which states that with 156,000 t half of all open die forged rings were made using RARR [22]. The whole process of RARR consists of five main production steps. A blank is cut from raw material and is usually heated up to forging temperatures ranging from 1000 °C up to 1300 °C depending on the materials that are being used. The hot blank is then pierced by a punch to create the preform of the ring. During the rolling process, the ring is formed simultaneously in two opposing located radial and axial rolling gaps. The radial rolling gap consists of a driven main roll in collaboration with a non-driven mandrel. The mandrel’s purpose is to continuously move towards the main roll translational to lower the rings wall thickness by pressing the ring against the main roll. On the opposing side, the axial rolling gap consists of two driven conical rolls. A downwards movement of the upper axial roll towards the lower roll realises the reduction of the ring’s height. Due to the constancy of material volume, the ring is continuously and primarily growing in its diameter. The axial rolling gap is then steadily moving away from the radial rolling gap to ensure an optimal rolling position synchronised to the ring’s diameter growth. To increase stability during the whole process, two guide rolls are located besides the radial rolling gap to stabilize the ring. The formed ring is then cooled off and a first quality inspection is being run. Depending on the use-case of the ring, it might undergo another production step of heat treatment (e.g. stress-relief annealing and tempering). The last step in the process flow is a final quality inspection. This can range from automated ring measuring to crack testing, using ultrasonic methods. The whole production flow is depicted in Figure 1.
Regarding all steps of the whole production chain, this paper looks at the current state of the art of data management and data analysis in the forming step (step 4 in Figure 1). This is done by presenting the results of a survey made in 2019 in which six representing ring rolling companies situated in Germany were interviewed regarding data storage and data analysis in the ring rolling process. The results of the survey are examined and a standardized framework of how to store and merge data in the ring rolling process is proposed in this paper.

### 3.2 Data Survey

The carried out data survey consists of eight different statements:

1. Process data is measured and stored
2. Process data is used and analyzed
3. Data is saved in a database
4. Data is (additionally) stored in its raw-format
5. Data is stored regarding common standards
6. Data is stored regarding internal standards
7. Operating data (e.g. tool replacements) are digitally tracked and stored
8. The whole production chain is connected and produced data is relatable

The results of the whole survey are shown in Figure 2. Possible answers to the statement were “yes”, ”no” and “no information”.

![Data survey radial-axial ring rolling 2019](image)

Figure 2: Data survey results 2019 showing all data
The overall result of the survey shows that six out of eight statements were answered with different responses, which indicates a mixed state of the art throughout all participants. Taking a closer look at groups of statements and their responses gives additional useful insights. To do so, the responses of statement one and two as well as the responses of statement three and four are depicted in Figure 2.

As seen in Figure 3a, statements one and two were particularly asked to elaborate the state of the art concerning data acquisition in RARR. The responses show that all companies do measure and store the process data but only half of them are actually taking a look at the data and are using it for analysis. That means that at least 50 percent of the companies do not recognize whether the stored data is in the right format, has all the necessary features or might eventually be measured in an accuracy which is too low to represent the meant information. Especially the last aspect is a common mistake that, once data is being saved, is never revised and validated. Regarding the responses of statement three and statement four in Figure 3b, a positive aspect is, that in 66.7 percent the data is already stored in a database and not in files on local servers all over the production environment. A downside of the responses is that only 16.7 percent store the data in a database in its raw-format, which means that the data is not aggregated or modified in any way. The storage of data in its raw-format will also be discussed later on as it is one key aspect for a convenient data integration framework. The last aspect of the survey, statement seven, needs to be discussed separately, which was entirely responded to with “no”. This is a drawback, because the information loss linked to this is immense. For various applications the information, whether a tool was replaced or a machine was under maintenance is at least very useful if not critically important.

The survey discussed above shows that in the area of RARR, there is a need to propose a framework to use process data for analytical approaches and to state common mischiefs concerning data storage and data integration. This is why in the following, a framework is presented and common mistakes and lessons learned are presented to address those issues and emphasize the need to deal with the data that is being stored.

The majority of the companies surveyed do not use existing production data to analyze their production process, yet current research in machine learning shows big improvements for several applications mentioned in section 2. Further, LAGAO managed to proof classification results up to 88.4 percent using data-mining methods in 2004. This indicates the usefulness of machine learning methods in the topic of quality prediction and process-failure classification. Moreover, new algorithms in form of deep learning tools, such as convolutional neural networks or long short-term memory cells are used for sequence problems and show great improvements over former algorithms [23,14]. All those improvements contribute to the fact...
that machine learning bears great potential in RARR and other production engineering sectors, which is underlined by a paper of the Wissenschaftlichen Gesellschaft für Produktionstechnik in 2019 [24]. Lastly, referring to a survey of IDG, more than 34 percent of the companies mention data quality as the biggest problem for machine learning applications [25]. This is why the following framework is proposed to set a standard for data acquisition in RARR to enable future applications using machine learning.

4. Framework

The proposed framework is depicted in Figure 4 and can be divided into three layers visualized by three central grey boxes. The framework should not be seen as the one and only standard but rather as a good example for the forming industry to start working with data analytics.

The first layer of the framework represents every step in the production process of RARR. On this layer raw sensor data is produced. This is an important step, because the sensor data needs to be saved in its raw format, meaning it should not be aggregated, truncated or varied by taking the mean, max or min. Once raw data is transformed in anyway before saving it, the raw dataset is lost even though the raw sample-frequency, aggregation-level or dimension-depth might be needed sometime in the future [26]. By saving raw data, future data analysis is enabled without the exhaustive and time consuming task of providing sufficient data once again. Additionally, the decentralized storage of raw data will be beneficial in a later layer of the framework as well. As for data storage, there are plenty of suggestions for database architectures depending on the type of data. For RARR it is mainly sensor and machine data, which can be stored in a relational database, for example PostGres, SQLITE, MYSQL or time series databases such as InfluxDB and TimescaleDB [27]. A summary of all currently produced data of the production chain can be seen in Table 1 and represents the state of the art in ring rolling. Keeping in mind that there are grown structures in industry and relational databases are the most used [28], it is a considerable trade-off to use an already existing relational database instead of deploying a whole new database architecture. Yet, these data storages should not be used for analysis.

<table>
<thead>
<tr>
<th>Production Step</th>
<th>Structured data</th>
<th>Unstructured data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furnace</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Press</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>RARR</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Quality inspection</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Heat treatment</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>External data</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Internal data</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

The actual storage of analytical data is depicted in the second layer. The raw data is then syntactically and semantically transformed by an Extract-Transform-Load-Process in addition to a validation-check. This validation check $f_{val}(x, t)$ checks the sensor data from the raw databases on state and time before transforming them. This step is inspired by a framework proposed by SEJA [29]. Additional external and internal data is also validated and transformed to be stored in non-relational databases such as MongoDB or Cassandra [30]. By adding external data the main dataset from the production chain is enhanced by further metadata. This process is called data enrichment, where the additional data offers further information to the
dataset [31]. For example, external data delivered to the company could be metallurgic inspection data or geospatial data from the supplier referring to the blanks origin.

The distributed databases are then managed by a framework such as Apache Spark or similar frameworks to orchestrate all decentralized databases and to manage all necessary steps for the final data integration and preprocessing steps. It also offers potential to be run in the cloud or on a standalone cluster [32]. This hybrid approach makes use of a combination of centralized and decentralized structures and thereby uses all advantages of fog and cloud computing altogether [33].

The third layer is called the analysis layer. To analyze the production data, the only requirement is to connect to either Apache Spark’s build in machine learning tools or use the preprocessed data for quality prediction in an external data analysis software.

The simple structure of this framework aids in the process of deploying it into grown structures as well as making troubleshooting a lot easier, because fewer nodes need to be looked at. In addition, the raw data storage can help fixing the framework, because a possible error can be traced back to the raw data causing problems in the $f_{val}(x,t)$ function and thereby the issue can be fix more quickly. Keeping in mind that some of the requirements today might change, with regards to the upcoming possibilities in research, meaning that the framework has to adapt as well. This task is made easier if the framework is more generalized and simpler rather than it being with a highly specific and restrictive framework.

One thing, especially for the use of machine learning, in fact supervised methods, is, to keep the target that needs to be predicted in mind and always making sure to relate the targets to its data and to store the targets as well. Moreover, it is common practice in many areas to split the storage and computation of such data. In addition, the framework and analysis should also be split up as shown by the dark bounding boxes, see Figure 4. This ensures a higher adaptability for a data-driven company to changes in data (-formats), computational hardware and analysis. [26]
5. Conclusion and key takeaways

Concluding, the data survey shows that the ring rolling industry already has good prerequisites in terms of produced data. The next step is to make this data available in an analytically useful form and embed this data in a framework as presented above. This should guide a company towards an overall data-driven culture, enabling all kinds of approaches towards data analysis to enhance and optimize present and future industrial applications and production processes.

As key takeaways for the hot forming sector, the following guidelines should be taken into consideration when starting data analysis tasks or establishing a data driven culture inside a company. The key takeaways are summarizations of general proposed guidelines from recent research as well as individually developed best-practices while establishing a data analysis framework in the hot forming sector:

- Data should be saved in its raw-format [29,34],
- Data dimensions should be validated before saving (e.g. frequency, accuracy),
- supervised machine learning methods always need a target related to the data [35],
- timestamps or other identifications should be used[34],
- a standard for measurement units should be set [36],
- a standard for missing values should be set (e.g. “Nan” or “/”) [36],
- it is highly recommended to keep up data conformity (throughout the whole company) [24],
- simple is better than complex [37].

6. Framework validation

As for a practical implementation the deployment of the proposed framework is currently being put to the test in a company to enable future data analysis. A great challenge is the interoperability with existing structures and undefined standards. The aim is to deploy all key takeaways and the proposed structure in the future.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 404517758

References


[22] Euroforge, 2016. EUROFORGE Production per specialities.


[26] Dunning, T., Friedman, E., 2019. AI and Analytics in Production: How to make it work.


Biography

Simon Fahle (*1994) is a member of the Lehrstuhl für Produktionssysteme (LPS) at the Ruhr-University of Bochum since 2019. He earned a bachelor’s and master’s degree in mechanical engineering at the Ruhr-Universität Bochum. His primary research topics are machine learning, time series and radial-axial ring rolling.

Till 2009 Bernd Kuhlenkoetter was responsible for product management and technology at ABB Robotics Germany. In 2009 Bernd Kuhlenkötter took over the Professorship for "Industrial Robotics and Production Automation" at the Technical University of Dortmund. Since 2015 he holds the professorship of “Production Systems” at the Ruhr-Universität Bochum.
Optimization approach for the combined planning and control of an agile assembly system for electric vehicles

Peter Burggräf1,2, Matthias Dannapfel1, Tobias Adlon1, Aaron Riegauß3, Esben Schukat1, Felix Schuster3

1Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Campus Boulevard 30, 52074 Aachen, Germany
2Chair of International Production Engineering and Management (IPEM), University of Siegen, Paul-Bonatz-Straße 9-11, 57076 Siegen, Germany
3Chair of Production Engineering of E-Mobility Components (PEM), RWTH Aachen University, Campus Boulevard 30, 52074 Aachen, Germany

Abstract

For some years now, the automotive industry has been challenged by growing market dynamics, shorter product lifecycles and customers' increasing demands for individualization. In order to cope with this development, the automotive assembly needs to adapt quickly to changing demands with a low level of investment in the future. Under the current circumstances, the traditional line assembly for high volume production is reaching its limits in terms of adaptability and scalability. A promising solution to address the current challenges is the concept of the agile assembly. The concept of agile assembly breaks up the rigid linkage of assembly stations and, thus, enables full flexibility in the sequence of assembly operations only limited by the precedence graph. Therefore, the routing of electric vehicles in the agile assembly is based on the availability of resources such as assembly stations and automated guided vehicles that handle the material supply. Further, by transferring the transport function to the vehicle itself, investments for convey or systems are eliminated. This research work presents an optimization approach for the machine scheduling and transportation planning, which derives instructions for electric vehicles, assembly stations as well as automated guided vehicles. For each electric vehicle, an optimized route is calculated, taking into account product-specific precedence graphs and minimizing the overall makespan. In addition, the machine scheduling and transportation planning is integrated into a combined planning and control concept which covers the allocation of resources and the assignment of capabilities of the entire assembly system. The approach is implemented and applied to a practical case of a compact electric vehicle. Thus, the work contributes to the evaluation of agile assembly systems in automotive production.

Keywords

Automotive assembly; Agile assembly; Assembly control; E-mobility production; Factory planning; Assembly planning

1. Introduction

Shorter product life cycles, an increasing variety of variants, growing uncertainty about sales forecasts and rising demand for electric vehicles (EVs) are the main challenges of today’s automotive production [1,2]. Automotive manufacturers are thus in constant conflict between short product life cycles and low investment costs for the production. To address these challenges, production systems in the automotive industry must
be able to adapt quickly and with low investment within their existing infrastructure in order to respond to changing market and demand conditions [3,4].

Until now, automotive assembly has mainly been characterized by line assembly. Line assembly is particularly efficient in the production of large quantities due to specialization advantages and continuous flow. However, it has low flexibility regarding changes in production volume and is rather limited in terms of line balancing, which makes it challenging to integrate new variants or products. Large differences in cycle times in the mixed model production lead to significant inefficiencies due to lower utilization of assembly stations (ASs). These disadvantages are particularly evident in production systems for EVs. Future market demands for EVs are uncertain and difficult to predict. Therefore, established automotive manufacturers are reluctant to make major investments in the production of EVs. In order to minimize the investment risk, the manufactures tend to integrate the assembly of EVs in existing lines that were initially designed for vehicles with an internal combustion engine. This leads to increasing planning efforts and cycle time losses due to fundamental differences in the assembly processes [5,6].

The Chair of Production Engineering of E-Mobility Components (PEM) and the Laboratory of Machine Tools and Production Engineering (WZL) at RWTH Aachen University have developed the concept of the agile assembly to address the above stated challenges in order to enable the economic small batch production of EVs. One of the main pillars of the concept is the transfer of the transport function within production from the conveyor system to the EV itself. Shifting the mounting and the commissioning of the powertrain into the early phase of assembly, the EVs become self-driving. The breakup of the rigid line structure enables a flexible sequence of assembly processes, which is only determined by restrictions of the product-specific precedence graph. Thus, agile assembly achieves scalability in terms of quantity and variance as well as significantly lower structural investments. With the flexible flow of EVs in the production, however, the requirements for the assembly control and its complexity increase. The elementary premise for the operation of the agile assembly is the creation of a network that links all participating resources within an assembly ecosystem. Automated guided vehicles (AGVs) in intralogistics, assembly stations and self-driving EVs need to act as cyber physical systems.

This work aims at the conception and demonstration of a combined assembly planning and control framework for the agile assembly. The assembly planning and control includes a machine scheduling approach that determines the assembly sequence of the EVs considering the product-specific precedence graph and material supply by AGVs.

## 2. State of the art

In order to lay the foundation for the assembly control system for the agile assembly of EVs, existing approaches for flexible assembly systems were examined. The primary object of consideration were principles for decision-making in the flexible assembly in the form of algorithms and constraints for the vehicle routing. In addition, the material supply by AGVs was taken into account. Attention was also paid to the classification of approaches in the overall context of production planning. Production planning in general involves preserving and enhancing all production processes while adapting to changes of the product portfolio or demand [7]. The projects freeMoVe, SmartFace and the approach of Bochmann were particularly relevant for this scientific work and are therefore explained in the following.

The research project freeMoVe covers the development of an ideal concept of flexible assembly as a new assembly organization form and includes the design of a control system, material supply and layout. It focuses on the development of a control system architecture for machine scheduling [8,9]. However, the scheduling of assembly tasks is not integrated into the overall planning process of the production. Furthermore, no concrete algorithms for implementing decision-making in the control system are presented. The developed concept of the control system architecture does not take any AGVs into account.
The SmartFace project develops a production planning concept for the final assembly of EVs according to the principles of flexible assembly. The goal of the project is the breakup of the fixed cycle times in the assembly and a transfer into a self-organizing cyber-physical system. As an interface between production program planning and production control, the concept of the volume cycle is presented. Instead of defining a specific production time, a production volume for a certain period of time is planned [10]. Thus, the material supply can continue to be provided just-in-time. The production control translates the processing jobs of the volume cycle into assembly orders and transport requests. The control is no longer carried out by a central planning and control system, but by a multi-agent system, which is composed of independent software agents. Each software agent performs its activities on the basis of limited information, such as its environment and its own capabilities [10,11,12]. The developed principle of a combined production planning and control remains conceptual while no specific algorithms for decision making are presented. Furthermore, the decentral approach based on a multi-agent system cannot ensure that a globally optimal solution is found [13].

The approach by Bochmann focuses on layout planning, machine scheduling and transport route planning of a flexible assembly system. The goal of the layout planning is to minimize transport costs within the assembly system. For this purpose, the statistical transport intensity between ASs is derived on the basis of the capability profiles of each AS. The assignment of assembly tasks to stations is based on a machine scheduling model and solved with a Tabu-Search heuristic. Eventually, Bochmann concludes that the investment costs for the investigated concept of agile assembly may exceed those of a classic line assembly [14]. However, this approach does not provide any integration of the presented planning tasks into a broader concept of assembly planning. Furthermore, the machine scheduling does not take the utilization of AGVs into account, which restrict solid cost analysis.

The discussed assembly control approaches lack in setting the developed methods in the context of the overall assembly planning. Therefore, no clear planning horizons and interrelations for layout planning, required production resources or production program planning are provided. Moreover, neither of the approaches presents specific methods of decision-making and algorithms or further considerations of AGVs in the scheduling approaches. In order to address the deficits of the existing approaches, a framework for combined assembly planning and control is presented in the following. This framework lays the fundament for the later implementation of a comprehensive machine scheduling approach for the agile assembly that includes the planning of EVs, ASs and AGVs.

3. Combined assembly planning and control

In a monolithic approach, increased flexibility requirements in agile assembly lead to a significant increase in complexity. To ensure that the overall assembly planning and control can be solved, the concept of hierarchical planning from operations management is applied. Hierarchical planning separates all process tasks and decisions into subtasks with defined interfaces between the subtasks. As the subtask level decreases, the level of detail increases while the observation horizon decreases. Decisions and solutions on higher planning layers are passed down as instructions for more detailed planning. Feedback is returned to upper layers in order to enhance the planning and decision quality in future iterations [15].

Two main goals of an assembly control are optimized efficiency and real-time performance, which contradict each other. Thus, specific time intervals for the transmission of instructions and feedback are defined for a rolling planning. It features a temporal disaggregation when passing down tasks. Upper planning layers have long planning horizons with rough resolutions of quarters or years. Once their planning is completed, decisions are fixed and passed down as instructions. The receiving layer disaggregates the planning interval into smaller units with higher resolutions. In general, planning horizons are shorter in agile assembly than in
classic assembly due to higher flexibility. The developed model for a combined assembly planning and control is shown in Figure 1.

The developed framework consists of a planning and control system as well as an object system. While the object system represents the physical production resources through cyber-physical instances, the planning and control system include more abstract planning layers.

**Capacity planning** is the top-level task and is executed first. The purpose of capacity planning is the strategic procurement of production resources within time horizons of quarters to years. It is based on long-term sales forecasts and the feedback from lower planning levels such as occupation rates from the machine scheduling. As a result, the layout planning layer may be instructed to introduce, remove or reallocate resources. The **layout planning** consists of the allocation of resources within the physical factory layout as well as the optimization of transportation routes and material flows. Its output are the availability of physical resources and capability profiles for each AS. To achieve this, further feedback is gathered from the object system containing information about transport intensity and time. The next layer is the **production program planning**. Its goal is to determine a profit-optimized material demand and delivery schedule for the products. Hence, outputs are both just-in-time material orders and the volume cycles, which contain all information about type and variation of the order in a specific timeframe. On this layer, planning horizons equal weeks and each production sequence is determined for time frames of shifts or working days. The process is similar for line and agile assembly concepts and relies on input data regarding capacity, orders and short-term prognoses. **Machine scheduling** then distributes the orders from the production program planning to the available production resources. Timed output commands comprise operations to ASs as well as material delivery assignments to AGVs. Key goal is a short makespan for the volume cycle. This planning step is unique to agile assembly as the classic line assembly features an already predetermined order of operations. The scheduling is discussed in detail in the following chapter. Top-down inputs are the planned volume cycle, available resources and capability profiles. The actual transport times are fed back from the object system to the upper levels. Interruptions such as defect stations or missing parts are countered through a fallback strategy, which is an interface between the scheduling and lower decision layers. Short-term rescheduling should only take minutes, thus ensuring a continuous production until the scheduling level has recalculated the entire volume cycle.
The final process step is the **transport route planning** for the EVs and AGVs. Transport instructions for both AGVs and the EVs are given to the system along with the current position of the physical objects. The route planning process is divided into rough and detailed path planning. Collisions with static and dynamic objects, buffer areas and avoidance of congestion must be considered within the time intervals of several seconds. The calculated trajectories are forwarded to the instances of the object system as the final result of the planning process. In the object system, the physical trajectories of the EVs and AGVs as well as the assembly operations are executed.

As one of the major challenges in the transition from a classical line to an agile assembly, the layer of the machine scheduling and its decision making algorithms will be detailed in the following section.

### 4. Machine scheduling for the agile assembly

In the concept of agile assembly the machine scheduling will replace the planning tasks of line balancing and the sequence planning of classic line assembly. While the classical line assembly offers limited possibilities for the resequencing of ASs and operations, the agile assembly enables a flexible flow of the assembly objects. The holistic concept for the control system is visualized in Figure 2 and describes the relationship between production program planning, machine scheduling and the object system including the flexible flow of the assembly objects.

![Figure 2: Interconnection of the machine scheduling in the concept of agile assembly](image)

The production program planning manages the existing orders and groups them to volume cycles which represent a set of orders to be produced in a specific time horizon. The machine scheduling receives a specific volume cycle from the production program planning. Each order within the volume cycle is described by a precedence graph as well as a set of required assembly operations. In addition, the machine scheduling planning requires deterministic information about the object system. This includes the capability profile, the availability of resources as well as the transport times. The capability profiles describe which assembly operations can be performed by which ASs. This ensures, for example, that a lifting platform is available at
the respective AS for underfloor work. Resources are defined as ASs and AGVs. Transport times quantify the distance between two ASs for EVs or between the supermarket and an AS for AGVs, respectively. The machine scheduling then derives the instructions for the ASs, AGVs and EVs. These instructions represent the control of the physical object systems and determine which EV is at which station at which time. The same applies to AGVs.

This implies first of all that the machine scheduling includes a combination of a scheduling and a routing problem. The scheduling problem is the creation of a sequence in the number of required operations (assembly tasks) to process EVs while the routing problem consists of the assignment of these operations to the ASs. In order to solve the combination of this scheduling and routing problem, a method for model-based decision making is required. Mönch (2006) classifies the methods of model-based decision making into priority rules, simulation approaches and deterministic machine scheduling, which will be used in this work [13]. Such machine scheduling problems with jobs consisting of multiple operations performed on different machines (such as ASs) are named Shop-Problem [16]. An extension of the general Shop-Problem is the Flexible Job Shop Problem (FJSP). Each job has its own individual material flow path and is not transferred to every machine [17]. Özgüven et al. (2010) extended the FJSP to the FJSP-PPF by introducing full routing and process plan flexibility (PPF) [18]. Routing flexibility means that multiple or redundant machines may be available for the execution of assembly operations. Process plan flexibility means that different process sequences can exist for different machining jobs. This is required as we assume that a few ASs will have an overlap in their capability profiles. The FJSP-PPF comes closest to the requirements of the control system in agile assembly. A major advantage of the FJSP class is the availability of heuristics for solving [19,20].

However, the control system is also supposed to generate instructions for the AGVs and EVs. AGVs are major cost drivers in the assembly and, thus, their utilization must be considered in the assembly control. For this reason, the formulation of the FJSP-PPF will be extended to include the consideration of transport times (TT) to the FJSP-PPF-TT. The extension of the FJSP-PPF to the FJSP-PPF-TT does not only imply additional constraints, but also requires an extension of the indexing and an introduction of new variables. However, since no research results exist yet on this subject, a novel approach has been developed. This approach suggests a specific distinction between the transport processes of materials (AGV transport) and the transport of the EV itself. This distinction is necessary, since the two transport procedures differ from each other. Furthermore, they do not reflect regular set-up procedures. Figure 3 visualizes the transport processes.

![Figure 3: Distinction AGV transport and EV transport](image)

Before an assembly operation can be performed at a certain AS, material may need to be transported from the supermarket to the AS. In this example, operation "o04" is executed on AS "s04". The necessary material supply by an AGV has to finish before the operation starts. This requirement cannot be represented by corresponding entries in the precedence graph since the transport operation and the assembly operation may overlap under certain circumstances. Instead, new constraints must be defined in order to take the transport of material into account. It must be considered that the AGV will return to the supermarket after the material
has been delivered and is blocked for the full transport time from the supermarket to the AS and back. Therefore, the operation "o04" can start after half of the transport time assuming a constant speed of the AGV. It should be noted that the transport time does not depend on the type of operation, but on the destination, which is related to the location of the AS (here “s04”). The example for the EV transport includes two consecutive operations ("o04" and "o06") to be executed at two different ASs ("s04" and "s01"). The EV ("v1") thus needs to move from one AS to another. As the used ASs are not occupied during EV transports they can be used to execute operations on other EVs. However, the EV is locked for the duration of the transport and cannot be processed. In contrast to AGV transport, the transport time of the EV depends both on the current AS as well as the destination AS.

The developed model was validated with a reference data set developed by Özbakır (2004) which has been adapted to the formulation of the FJSP-PPF-TT [21]. To show the practical relevance and the possibility of exploiting flexibility potentials, the machine scheduling model will be applied to a concrete example below.

5. Practical Application

To further demonstrate the capabilities of the developed model, a relevant scenario configuration and the assembly description for the agile assembly of a compact EV have been designed. Figure 4 shows the resources of the scenario. The setup consists of five different types of ASs (A)-(F) as well as two different types of AGVs (α)-(β). Figure 5 presents the precedence graph for the EV. The assembly process of the EV consists of 13 operations (“o01” to “o13”).

![Figure 4: Scenario resources](image1)

![Figure 5: Scenario precedence graph](image2)

The necessary assembly operations of the EV are detailed in Table 1. For each operation, a specific amount of AGV transportations is given. Additionally, process times were assigned to each operation as well as ASs that are capable of performing the operation, marked by the dots in Table 1.

<table>
<thead>
<tr>
<th>Number</th>
<th>Description (α) (β)</th>
<th>Process time [s]</th>
<th>(A)</th>
<th>(B)</th>
<th>(C)</th>
<th>(D)</th>
<th>(E)</th>
<th>(F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>o01</td>
<td>Assembly Chassis, Spaceframe &amp; Backend</td>
<td>1</td>
<td>1</td>
<td>1513</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>o02</td>
<td>Wire Harness &amp; Control Unit Assembly</td>
<td>1</td>
<td>0</td>
<td>3915</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>o03</td>
<td>Dashboard Assembly</td>
<td>1</td>
<td>0</td>
<td>1527</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o04</td>
<td>Side Covering Assembly</td>
<td>0</td>
<td>1</td>
<td>437</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>o05</td>
<td>Side Windows &amp; Roof Assembly</td>
<td>1</td>
<td>1</td>
<td>713</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>o06</td>
<td>Windshield Assembly</td>
<td>0</td>
<td>1</td>
<td>728</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o07</td>
<td>Interior &amp; Safety Belt Assembly</td>
<td>1</td>
<td>1</td>
<td>1121</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>o08</td>
<td>Rear Body Assembly</td>
<td>1</td>
<td>0</td>
<td>1137</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>o09</td>
<td>Seats Assembly</td>
<td>0</td>
<td>2</td>
<td>1039</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o10</td>
<td>Frontend &amp; Front Hood Assembly</td>
<td>0</td>
<td>1</td>
<td>2446</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o11</td>
<td>Doors Assembly</td>
<td>0</td>
<td>1</td>
<td>789</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o12</td>
<td>Underbody Assembly</td>
<td>0</td>
<td>1</td>
<td>981</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o13</td>
<td>Fluid Filling</td>
<td>0</td>
<td>0</td>
<td>910</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Operations and resources needed to assemble the compact EV
In order to solve the introduced scenario, the machine scheduling algorithm was implemented in Python using the high performance MIP-Solver Gurobi. On an Intel Core i7-3612QM 2.1 GHz CPU and 8 GB RAM, the calculation time for the scheduling of one EV was 435.66 s, whereby the global optimum was already found after 25 s. Thus, an explicit solution for more than one EV is not feasible within relevant time. To overcome this constraint, the problem is solved sequentially. At the beginning, the scheduling is solved for the first EV and the results are stored. Then, the second EV is added to the scheduling problem with the previously generated results as additional constrains. This procedure is repeated for each newly added EV. The downside of this technique is that reaching the global optimum cannot be ensured. The resulting schedule of the exemplary application for four EVs is shown in the Gantt-Chart in Figure 6.

The rows represent the schedule for each AS (A)-(F) and each EV represents an assembly job. For instance, job j001 starts at AS (F) where 4 assembly operations are executed (separated by vertical lines). Then, job j001 performs a transport operation to AS (D) where another 2 assembly operations are executed that are followed by another transport operation back at AS (F). This transfer is required due to varying capability profiles of AS (D) and (F).

The result of the machine scheduling exposes that the assumed quantity of AGVs is significantly too high and hence no limiting resource. Therefore, the schedule for the AGVs is not visualized here. As a result of the sequential scheduling, the most versatile station (F) has the highest utilization rate and, thus, represents a critical resource. Embedded in the combined assembly planning and control framework, this information can be used to re-evaluate the allocated resources and the layout. Moreover, the missing global optimum can be seen in the schedule. Job j001 is mainly processed at station (F), since this station (F) has a far-reaching capability profile and, thus, transport times can be minimized. However, the utilization of station (F) increases with each additionally dispatched EV whereas station (B) has an overall low utilization. Therefore, bottleneck stations and recommendations for the duplication or removal of certain assembly stations as well as required capability profiles can be derived from the schedule.

6. Conclusion and outlook

This research work presented a combined planning and control concept for the flexible and scalable assembly of electric vehicles: the agile assembly. In agile assembly, self-driving electric vehicles get routed to assembly stations depending on their order-specific equipment features. Automated guided vehicles provide the material required for an individual assembly operation to the assembly stations. The flow of electric vehicles and automated guided vehicles as well as the sequence of assembly tasks in the agile assembly is determined by machine scheduling. For this purpose, an enhanced Job Shop Problem taking into account
transport times of electric vehicles and automated guided vehicles was formulated and applied to a practical use case. The application has shown that valid schedules and routes can be generated for electric vehicles and automated guided vehicles. Furthermore, statements can be made about resource utilization and bottlenecks. However, further research is needed to develop heuristics as a solving method to improve the solvability for larger volume cycles as the current computing time does not yet meet future requirements of a real-time control. In addition, the related planning activities described in the combined planning and control concept such as capacity planning, layout planning and production program planning must be algorithmically detailed in order to quantify the entire system behaviour of agile assembly. This scientific work constitutes a significant added value by presenting a machine scheduling approach for the agile assembly, which enables the economic evaluation of agile assembly.

Acknowledgement

The research work has been conducted by the Laboratory of Machine Tools and Production Engineering (WZL) and the chair of Production Engineering of E-Mobility Components (PEM) at RWTH Aachen University within the research project LoCoMo, funded by the German Federal Ministry for Economic Affairs and Energy (BMWi); Grant Agreement No. 01MX15008A

References


Biography

Peter Burggräf (*1980) has been head of the Institute of International Production Engineering and Management at the University of Siegen since 2017. Prof. Dr.-Ing. Peter Burggräf is a member of the Supervisory Board of the electric vehicle manufacturer e.GO Mobile AG and Managing Director of StreetScooter Research GmbH.

Matthias Dannapfel (*1986) has been head of department of Factory Planning at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2017.

Tobias Adlon (*1989) has been research associate at department of Factory Planning at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2015.

Aaron Rieglauf (*1990) has been research associate at the chair of Production Engineering of E-Mobility Components (PEM) of RWTH Aachen RWTH Aachen University since 2017.

Esben Schukat (*1993) has been research associate at department of Factory Planning at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2019.

Felix Schuster (*1992) was a master thesis student at the chair of Production Engineering of E-Mobility Components (PEM) of RWTH Aachen RWTH Aachen University and has been a Global Trainee at ZF Friedrichshafen AG since 2019.
Generative Design In Factory Layout Planning: 
An Application Of Evolutionary Computing Within The Creation Of Production Logistic Concepts 

Matthias Dannapfel\textsuperscript{1}, Sebastian Patrick Vierschilling\textsuperscript{1}, Stephan Losse\textsuperscript{2}, Oliver Matzke\textsuperscript{1} 
\textsuperscript{1}WZL RWTH Aachen Germany 
\textsuperscript{2}StreetScooter Research GmbH, Aachen Germany 

Abstract 

This paper describes the creation and application of a generative design approach in the production logistics layout concept creation as part of factory planning projects. Production systems evermore are influenced by an increase of product variants during the planning stages as well as shorter replanning cycles due to higher agility requirements to the production system. Thus, requiring the planner to more frequently conduct the highly complex planning procedure of creating layout concepts for the material supply within the assembly line. Currently, mathematical or graphical assignment methods are executed but are often used disjunct and are not used jointly. Furthermore, today’s planning methods are mainly based on manual planning and assignment activities. To address the aforementioned issues, this paper elaborates the application and usability of generative design methods for production logistics planning. As first step the scope and requirements definition for the new production logistics layout application is conducted. Afterwards, generative design, including a multi-objective genetic algorithm, is used to serve as a solution to compile and search through the high-dimensional solution space of all possible logistic layout concepts. Here, layout restrictions and production goals, such as cost and time savings, are reconciled. After the design creation and evaluation by the algorithm, the planner overviews the results and enhances the design parameters until a final concept is reached. This paper concludes with a SWOT analysis of the new planning approach to investigate the used methods, evaluate the impact of the approach on planner’s work and identify additional research potentials of using the generative design for other factory planning domains. 

Keywords 
Generative design; Multi-objective optimization; Factory planning; Production logistics planning 

1. Motivation 

In the United States (U.S.) during the second half of the 20th century, an estimation by TOMPKINS and WHITE has ascertained that 8% of the US gross national product has been spend on new facilitates annually [1]. Also in Germany, despite the ongoing challenges in the global economy, the number of factory and workshop buildings completed between 2001 and 2018 shows a constant development. Those numbers, although not even including the modification of existing factories, illustrate the fundamental existence of factory planning activities. [2] 

In the field of planning manufactories, layout planning and therefore solving the facility layout problem (FLP) stands out as it represents not only one of the utmost essential processing steps in the overall process
of factory planning but also interrelates with many other components of factory planning [3]. As major part of factory planning and the Aachen Factory Planning Framework, layout planning handles the arrangement of the operational functional units [4]. This spatial assignment is carried out several times during the full factory design (e.g. for the coarse and detail planning or the ideal and real planning) and at different levels [5]. Its importance is shown, since an effective layout helps to improve the business performance and can reduce up to 50% of total operating expenses [1]. On the other hand, an ineffective layout can result in up to 36% more material handling costs [6]. A further important part of factory planning and very closely linked to the layout planning is the design of the production logistics. As interface between the warehouse logistics and the assembly, the related processes are required to handle the highest amount of diversified and individualized parts and discuss their transportation [7]. In recent years, the planning frequency is increasing due to changes in the market and shortened product life cycles [8].

This research starts by looking at the current status of designing factory layouts and solving the FLP. Then, it develops a new production logistic layout planning approach and tool using prevailed software engineering methods that combine a suitable genetic algorithm with user willingness. The goal of the application is to be able to interact more closely with the practice and to conduct future research directly application-oriented.

2. Literature review FLP and Production Logistics Planning

The target criterion solving the FLP is the optimal arrangement of a variety of objects within a specified area [4]. The various applications of layout planning within the entire planning process are defined by PFOHL. He regards the outer level of layout planning as the internal locations for production segments. Below, he names the determination of the locations of the production units within the production segments. This means the arrangement of the production units such as machines and working stations, but also storage areas. Here, the interrelation between the layout planning and the production logistics becomes apparent. Once discussing the concept of the material supply one uses arrangements for the storage areas and vice versa. [9] In order to narrow down the subject area of factory planning, focus in this paper is laid on depth and not the degree of the dispute with the research field. Notably, the arrangement of objects within the material supply area, therefore the provision area in front of the assembly station, are looked at.

KOOPMANS and BECKMAN were first to define the problem in research and science. Here, under the name of FLP, it is classified a classic computer science problem. Its task is the arrangement of objects within a given area following the material flow between them while certain constrains are fulfilled and objectives are given. [4] Therefore, for finding the best solution, often the principle of form follows flow is used [10]. KOOPMANS and BECKMAN formulated the problem as a quadratic assignment problem (QAP) [4]. Others have formulated the problem as a quadratic set covering problem, a linear integer programming problem, a mixed integer programming problem and a graph theoretic problem [11]. The literature has added in defining it as a NP-Problem [12]. The FLP is either applied for equal areas or unequal areas and extensions were made to add more than one target criteria and therefore have multiple objective functions as this better reflects the reality [13]. From the beginning of research, many algorithms have been developed to solve the problem and various papers were published surveying the existing algorithms. If one wants to access more literature regarding this topic, one can refer to Hosseini or Kusiak [14, 11].

To allow a concurrent engineering approach, in conjunction with layout planning, this parts looks at production logistics. Following LUCZAK, the literature subdivides the logistics of a company into procurement logistics, production logistics and distribution logistics [15]. According to the flow of goods within the enterprise logistics, the productions logistics is incorporated between the procurement and the distribution logistics and connects them with each other [9]. The production logistics planning includes the design of the logistic system within the factory starting at the incoming goods area, going through the internal material supply to the outgoing goods area. This includes the required logistical areas, storage and transport.
means as well as load carriers, supply principles and buffer sizes. [16, 17, 9] Those aspects all aim to achieve a supply of material at their designated areas. Depending on the size of the cooperation, it can be useful to divide the production logistics after spatial measures. Here, the areas covered can be divided into an intra-apparative (inside individual production plants), an in-house, an inter-company and an inter-plant logistics.

In the course of this work it is decided to focus only on the intra-apparative area. [18] LOTTER AND WIENDAHL describe the supply of materials as essential for the delivery time, delivery reliability and thus as a logistical success factor [19]. The goal is the optimization of the internal material and information flow to and between the production sites [7, 20].

Despite increasing digitalization and extensive computer support also in factory planning, currently, the process of newly designing or modifying layout concepts including the production logistic aspects still requires manual handling, which makes it a very unpleasant task [21].

To face the circumstances, an analysis of artificial intelligence approaches and paradigms has taken place and evolutionary algorithms and especially genetic algorithms have turned out to be most matching to the FLP. Especially, generative design (GD) as application-oriented tool using genetic algorithms in engineering shows great potential in factory planning. Therefore, in the following GD is further presented.

### 3. Introduction: Generative Design

AGKATHIDIS points out the fact that there is no clear definition of the term „generative design“. He describes it „[…] as a design method where generation of form is based on rules or algorithms […]“. [22] More engineering related, MCKNIGHT defines GD as „[…] the process […] mimicking natures evolutionary approach… to quickly explore thousands of design variants to find the best solution[...]“ [23].

MCKNIGHT describes the GD process in four phases. During the first phase the person in charge is required to define the input parameters and additively the goals. Possible parameters include constraints or properties of the materials. Often goals contain the mass and volume, which can in many cases be linked to the costs. The data can be of qualitative and of quantitative nature. The second phase uses computational power, sometimes extended by cloud computing, and various algorithms to generate a vast number of designs. This phase can be broadened by adding performance analysis for each design. The third phase includes an iteration process by the engineer. After obtaining the results from phase two, the various concepts are studied and possibly parameters and goals are adapted. In addition with AI, the human iterates new concepts until a most relevant solution is identified. The last phase describes the aftermath to the GD process. Combined with additional tools and methods the final design is used to solve the general problem. [23] The above described phases slightly amended by VILLAGI ET AL. are presented in Figure 2.
Here, next to the GD aspects, also possible engagements and levers of virtual reality and the Internet of Things are shown. Nowadays, research of GD has ventured forward into the architectural field [25 – 27].

Next, a methodology has been developed which creates a GD tool for the application within the production logistics layout planning.

### 4. Methodology

This paragraph describes the methodology which makes use of software engineering and in particular of the stages of the waterfall model in a modified form which are seen in the Figure 2. In order to reach the aim of a GD tool for layout planning, it is necessary to specify the requirements. Within this first stage of the waterfall model, the requirements were analyzed and listed in a threefold way. First, the planner as user sets certain standards for the tool. The tool is supposed to fit to the current knowledge of the planner and respects its abilities of using software and computer support. Second, the integration of the genetic algorithm needs to be considered. The functionality of the tool in relation to the algorithm requires to take a look on how the genetic algorithm works in detail and how it then can be implemented within the software. The system itself is required to run a genetic algorithm for its evolvement of the layout concepts. By changing the inputs, calculating specified goals, evaluating and then evolving various concepts and performing the optimization given by the genetic algorithm, a high number of layouts is created and optimizations are driven. The layout creation should also include a randomization of possible layouts and therefore have the possibility to not use any optimization algorithm. For the ability of running the genetic algorithm, the system needs input parameters, constrains and fixed goals. Since this tool applies in the creation for the layouts of the assembly station provision area, a detailed analysis of the layout related elements including the aspects from the production logistics topic is necessary. As the purpose of the instrument is to create a number of possible layout concepts for factories considering production logistic aspects, which are depending on specific rules and objectives, also those have to be defined clearly and scientifically derived from appropriate literature. Most important, the tool needs to be compatible with the requirements from the layout concept creation process.
For this reason, following elements were assessed more in detail: potential layout contour, number of components, buffers within the storage, production matrix on a time axis, material supply strategy, means of conveyance for the transportation, conveyance aids and area of storage. These are not mentioned as part of this publication due to these areas being common for every production and have been investigated using experience and standard literature.

Next, for combining the mentioned elements to the layouts, multiple objectives and restrictions are found in the factory planning literature. Here, the two criteria of material flow and costs are further presented, while it should be feasible for the tool to be expandable on more aspects. KETTNER ET AL. analyze the process of arranging the operational tools within the factory and they conclude that the production flow must be observed. They assess the four main components of the production flow as material, personnel, information and energy flow. Because the material flow is determined as the main cost driver, most methods focus on the minimization of transportation cost. Inter alia, in the internal location planning, the main goal is to minimize the transport costs [28]. Therefore, as most methods only allow one optimization criterion, the material flow is the chosen one [5]. During the layout planning of a material flow system, the goal is to find favorable spatial arrangements for the material flow relationships up to finding the minimum cost configuration. To measure and rate the costs for transport ARNOLD and FURMANS propose using the distance and the throughputs [29]. Since transport is the moving of objects and transportation referring to TAKEDA this is a way of waste, the goal is to minimize the distance covered by the loading carriers [30]. Those discussions deduce to having a minimized transport path as target. In addition, the required space is a highly important commodity as it links directly with area costs. In this case, either the space is given, or it is to be calculated. Therefore, another objective is the degree of used space and the goal is for it to be kept low [31]. On the one hand there is a given space which is then claimed by several entities. On the other hand, the space is to be estimated and then directly linked with the price for purchase or rent of a piece of land. In either occasion, the target is to require the smallest space possible. Another target for the economic efficiency is the goal of keeping the investment low [32, 19]. Also, a low tied capital can be targeted with a stock as little as possible [33]. Therefore, the goal of the GD instrument is to lower the tied capital or the investment costs. Additionally, the target is to keep the labor costs down by reducing operational times per operation.

Creating the product requirement document, it is indispensable to consider the system itself, but also its influence and connection with and from the outside of its boundary. For this reason, this second step discusses the design of the overall system and the software itself. Here, the GD tool on a high and on a lower level must be represented and the software architecture acts as outcome. For the high level of designing the tool, as a system within the factory planners task and delineated within the modules from the Aachen Factor Framework, the use case diagram as one of fourteen UML diagrams is well suited. Since the conception of
the tool orientates towards GD, the model mentioned above is taken for the method design of the tool. Here, as aforesaid, currently the planner collects necessary qualitative and quantitative data and then creates suitable layouts per hand or with diverse but not sufficiently automated tools. Possible layouts with positive features are then further looked at and analyzed mathematically, before a final layout concept is decided on. Thus, then getting implemented until modifications come up. Using the GD approach from architecture, the phase where the planner creates suitable layouts per hand, the planner now is supposed to be using a GD tool.

First, the factory planner collects the qualitative and quantitative data and uploads the data into the system. Inside the system, the tool automatically generates a high number of layouts via iteratively creating, evaluating and evolving the layouts applying a genetic algorithm. Finally, a number of various designs is presented. Thus, the factory planner has to select one or more layouts which are then exported and can be used within other tools or can directly be implemented. In the beginning, the user simultaneously defines the targets of the layout and inputs the qualitative and quantitative data. This data for the above explained application consists of the layout, the possible material supply strategies, the possible means of conveyance, the possible conveyance aids, the possible provision technology as well as all restrictions and constraints and last the production numbers. The next step is the creation of the layout. Here, a possible selection of inputs is collected and a random layout is designed based on the production numbers and the restrictions and constrains and directly evaluated by the defined goals. This layout is then saved including its collected and calculated data.

In the next step the genetic algorithm changes the inputs and the loop runs again. When a specified amount of loops is done and different layouts are developed, the loop stops and the selection is presented to the factory planner.

In the third step the software was constructed and coded according to the requirements and the software architecture from the proceeding step. Before programming, the exact implementation of the design was broken down to various modules and those were then being described and later programmed. In the fourth step, the modules were combined and linked to each other. Also, within this step the explanation of how to get the tool running and where it fits in the overall planning cycle of the layout creation was given. This included the explanation of the general usage of the software for the planner. Some of the other activities within this step involved testing the system as a whole on compliance with the requirements and restrictions. As last part of the waterfall model, the system was then put into practical use. A first use case of the application is explained in the following chapter.

5. Use case / Application

Since there are many existing programs on the market which enable to not fully rewrite a new software and still fulfill the requirements, this work rather extends existing VPL tools. In the course of this work, Dynamo as part of the Revit toolkit is used. Here, Revit 2020 and the included Dynamo 2.1 serve as visualization, displaying and visual programming tool and Project Refinery Beta enables the GD ability.

As use case, this work used the workshop production layout problem delineated by WÄSCHER, which is then applied within the visual programming code. This problem uses a delimited area (layout boundary), a set of arrangement objects (AO), material flow relationships between the AOs, additional conditions such as relative or absolute installation conditions. The spatial arrangements can then be valued using a set of destinations such as transport distances and transport costs. [34]

In the next step, the above deployed process is running a randomization creation phase resulting in 300 different possible layouts. For applying a genetic algorithm for an optimization of the concepts, tests were done, but show that the time required to optimize all outputs takes comparatively long. Therefore, within this use case, the objectives are randomized and not archived through the application of a genetic algorithm.
Nevertheless, this optimization is possible and doable with small modifications and enough time resources and within existing software, which enables the establishment of the above created software methodology. Figure 3 shows part of the results.

Figure 3: Extract of possible layout concepts

The application and the results show, that for creating a variety of different possible layouts, the developed tool serves as a feasible solution. The GD approach from architecture therefore is applicable and suitable for factory planning, since all the requirements are fulfilled while all constrains are satisfied. Thus, the factory planner is now able to view the data and decide on the best possible layout. Nevertheless, the designs show the possible layouts but are not equipped with a visualization of the output figures. So in the first step the output figures need to be generally designed but also determined and weighted for every specific planning case. This could be improved in future simulations. A qualitative evaluation by the planner is furthermore possible and subsequent arrangements can be done to the constraints and restrictions. Thereafter, a new running of the algorithm or a new randomization can be started and even more concepts can be created and improved. With this, the results from RIPON ET AL. are successfully integrated into a tool and an iterative solution for the planner to create layouts exists [35]. The step of applying visual programming to create the use case and all the above explained modules so far was not as easy as expected. Many restrictions within the software made it necessary to work with simplifications and work arounds. Thus, creating high efforts for the planner. Especially, as this use case is quite simple, still much work was required and further extensions of the case probably come with even higher effort. In general, it turns out that, although the software is used in the field of architecture, there is still a lot to be done in the much more complex area of factory planning. Yet, for this specific use case, the implementation of the tool application in the workshop case was successful.

6. Conclusion

At the end of this study, a SWOT analysis of the new planning approach investigates the created methodology and lists additional research potentials of using the GD in other factory planning domains.

The whole dispute was started with the idea of applying GD in the area of factory planning. In retrospect, this was quite dangerous as it does not confirm with the scientific approach. Normally, a research gap is deduced in the beginning and then possible solutions are searched and created. In this case, it was the prerequisite that GD is applied and the whole construct of this work was then build around this idea. The risk in this attempt is, that the research can come to the result, that GD is not even the best solution or any possible solution. Luckily, the results do not confirm this risk, whereby an influence cannot be excluded by the sequence followed here. The business demand of creating a tool which helps the factory planner on
choosing the suitable material strategy and its production logistic concepts on the one hand is comprehensible, although more research is required to fully examine the whole spectrum of possible solutions.

One of the weaknesses of this work is clearly the amount of simplifications. Here the area should have been narrowed down even further in advance. Furthermore, the dispute of AI is far not deep enough and the holistic view requires a further debate with more literature and experts in the field of AI. The use case shows, that many weaknesses exist also in the existing software. The disadvantages of the used software were listed and either the software requires many modifications or updates, or a whole new software is necessary.

Despite those negative aspects, the work gave a first impression on GD within factory planning and therefore starts an interesting subject of research. With modifications and improvements, the tool in the future could help the planner to actually design layouts, which were before unthinkable and furthermore assists in the evaluation phase. This research pictures the feasibility of a software tool to assist the planner and with the use of AI algorithms improve the planner’s tasks. Using existing software from another area such as architecture, was helpful in that respect, that it can be built on top of this research instead of requiring to completely redeveloping all the knowledge.

Those strengths create many opportunities for further research as the time is currently right with cloud computing and an advanced level of engagement within this field of research. Many new topics are found which can be worked out to independent works. A study on current papers on the facility layout problem algorithms is required, since the last one is a couple of years old. This work has only touched the topic of genetic algorithms as FLP solutions, this needs a detailed analysis. Apart from this, a detailed analysis why GD from architecture is suited for factory planner (similarity of the two jobs) would be interesting. This can be archived in levels above the planning of the provision area by doing the choice of location, the general construction of building, the arrangement of operating departments or the arrangement of operational functional units. Below, it can be done within the material supply strategy, e.g. material handling and supply strategies are planned automatically or decision of which supply vehicle/ tool to use for which task) or the workplace structure.

Consequently, this work forms a basis for further research work in the entire field of future factory layout design applications. This fundamental new approach to generative design might support the vision of the future factory to become reality faster.

References


Authors:

**Dr. Matthias Dannapfel** (*1986) is Chief Engineer of the Department of Factory Planning at the Chair of Production Engineering at the WZL of RWTH Aachen University. During his research he focuses on the topics Factory of the Future as well as automated and integrated factory planning processes. Furthermore, he is CEO of the ONEFACTORY GmbH.

**Sebastian Patrick Vierschilling** (*1990) is research associate at the Department of Factory Planning at the Chair of Production Engineering at the WZL of RWTH Aachen University. Currently, he works as group leader of the research group ‘Digital Factory Planning’ researching integrated factory modelling as well as automatized planning procedures.

**Stephan Losse** (*1981) has been a Senior Consulting Engineer at StreetScooter Research GmbH since 2017. During his research time at the Chair of Production Metrology and Quality Management at the WZL of RWTH Aachen University, he worked in the field of holistic quality management systems for knowledge-based organizations.

**Oliver Matzke** (*1993) is participant of the CDI MBA program. Beforehand, he researched as a master student at the Chair of Production Engineering at the WZL of RWTH Aachen University.
Self-organization and autonomous control of intralogistics systems in line with versatile production at Werk150

Jan Schuhmacher¹, Vera Hummel¹
¹ESB Business School, Reutlingen University, Reutlingen, Germany

Abstract

The planning and control of intralogistics systems in line with versatile production systems of smart factories requires new approaches and methods to cope with changing requirements within future factories. The planning of intralogistics can no longer follow a static, sequential approach as in the past since the planning assumptions are going to change in a high frequency. Reasons for these constant changes are amongst others external turbulences like rapidly changing market conditions, decreasing batch sizes down to customer-specific products with a batch size of one and on the other hand internal turbulences (like production and logistic resource breakdowns) affecting the production system. This paper gives an insight into research approaches and results how capabilities of intelligent logistical objects (intelligent bins, autonomous transport systems etc.) can be used to achieve a self-organized, cost and performance optimized intralogistics system with autonomously controlled process execution within versatile production environments. A first consistent method has been developed which has been validated and implemented within a scenario at the pilot factory Werk150 at the ESB Business School (Reutlingen University). Based on the incoming production orders, the method of the Extended Profitability Appraisal (EPA) covering the work system value to define the most effective work system for order fulfilment is applied. To derive the appropriate intralogistics processes, an autonomous control method involving principles of decentralized and target-oriented decision-making (e.g. intelligent bins are interacting with autonomously controlled transport systems to fulfil material orders of assembly workstations) has been developed and applied to achieve a target-optimized process execution. The results of the first stage research using predefined material sources and sinks described in this paper is going to set the basis for the further development of a self-organized and autonomously controlled method for intralogistics systems considering dynamic source and sink relations. By allowing dynamic shifts of production orders in the sense of dynamic source and sink relations the cost and performance aims of the intralogistics system can be directly aligned with the aims of the entire versatile production system in the sense of self-organized and autonomously controlled systems.

Keywords

Intralogistics; Autonomous Control; Self-organization; Decentralized decision-making;

1. Introduction

Companies are increasingly confronted with growing international competitive pressure, decreasing batch sizes due to a rising demand for individualized products as well as the demand for the shortest possible delivery time [1,2]. In order to enable the factories to produce personalized products in small batch sizes down to a batch size of "1" under the performance and cost conditions of mass production, logistics systems must be able to adapt flexibly to changing requirements of the material flow [3,4]. Therefore a versatile
behavior of the production and logistics system in the sense of an early or proactive adjustment of objects, structures and processes of value creation is crucial. The trigger of a required adaption of production systems are internal and external turbulences. The term “turbulence” stands for the effects of mostly unexpected changes which are acting from the outside or within a company [5]. A distinction is made between internal and external turbulences. Internal turbulences are having their origin in changes within the production system, like machine breakdowns or new product variants, whereas external turbulences are arising due to reasons of the outside of the company like changing market conditions or delayed deliveries of suppliers [6]. As each product in a customized production with various potential turbulences differs from the previous products in terms of the required manufacturing and assembly processes as well as the required components and their flow through the factory, the real-time configuration, control and decision making is a central challenge. Self-organized, autonomous controlled material flow systems will distribute the necessary decision-making and control processes among intelligent logistics units [7]. Machines and other objects in production will jointly decide on the used tools and machines in close cooperation with autonomous transport systems deciding on the transport of components and (semi-finished) products from their current location to the next production step.

2. Research design

To investigate the potential of self-organization and autonomous control for an improved target achievement within intralogistics systems the research methodology of a reasoning cycle has been chosen. The reasoning cycle starts with the hypothesis formulation and continues with the deduction of predictions, testing and observation of predictions and induction/feedback into the initial hypothesis [8]. The main hypothesis to be proven is that the application of self-organization and autonomous control leads to an improved achievement of cost and performance goals within versatile production systems. To prove this hypothesis a two-step approach is followed. In the first step, a first approximation based on fixed material sources and sinks in the work system is investigated. The second step will be the investigation of the application of self-organization and autonomous control within production systems with flexible material sources and sinks for an integrated and simultaneous cost and performance optimization of all production and logistics resources.

3. Self-organization and autonomous control

In general, the concept of self-organization deals with the explanation of the autonomous emergence of ordered structures in open, interacting, non-deterministic, dynamic-complex systems. In addition, the approach of self-organized systems is focused on how a system designs its processes and systematic structures in an autonomous manner. The concept of autonomous control generally describes processes of decentralized decision-making in non-hierarchical (heterarchical) structures based on interacting elements in non-deterministic systems with autonomous decision-making capabilities. Autonomous controls aims on achieving a higher robustness and positive emergence of the overall system through a distributed, flexible management of dynamics and complexity arising in the system. [9]

Although the concepts of self-organization and autonomous control have many common characteristic features, the approach of self-organization is more focused to the management and organizational level of holistic systems whereas the approach of autonomous control is more evident at the execution level and single object level of systems. For a more detailed characteristics-based differentiation of autonomous control and self-organization, please also see Windt [9] and Schuhmacher [10]. The ability of self-organized systems to change the system’s structure or processes requires in particular the timely recognition of a need for change (e.g. due to external or internal turbulences) and the rapid planning and implementation of the necessary change in industrial management. Therefore, it is essential to master the planning complexity, which can be reduced by decentralized control systems. A major task of the company’s management is to
define guidelines and decision corridors for the decentralized initiation and control of change within the company. In this way, autonomous control can evolve from the local level such as such as the intralogistics system to the target-oriented self-organization of the entire company [11]. Based on the self-organization capabilities of the production system in conjunction with an appropriate visualization application, the plant manager or corresponding specialist worker gains a real-time overview of production for decision support and can react quickly to complications in order to make well-founded decisions for adjustments within the production system. [12]

In addition decision-making within technical systems will no longer be possible through hierarchical structures following the classical automation pyramid, since a large number of decisions for the control of the material flow must be made in near-realtime and a predefined, target size-optimized solution cannot be predefined centrally for every eventuality. Instead, these tasks will be performed in a decentralized manner by the intelligent objects in the material flow system, such as the transport units, transport vehicles and software agents. Thus, according to Hompel [11] it can be stated that with an increasing complexity of logistics systems the degree of decentralization and self-organization must increase in order to be able to control them. [11][7]

4. First approximation for versatile logistics

As a first step before the development of an entirely self-organized and autonomous control method for a holistic consideration of logistics and production goals within versatile production environments, a first approximation using a two-step approach for the sequential optimization of the value creation system and the autonomously controlled intralogistics system has been used (see Figure 1). Therefore two separate equation systems for the optimization of the value creation system (consisting of assembly, production and work system-related storage resources) and for the autonomously controlled, versatile logistics system (consisting of transport systems) have been set up covering the cost and performance measures of these systems. To derive the most effective work system to fulfill the incoming small batch size production orders in line with the set targets (cost, performance and qualitative goals) of the work system, the method of the Extended Profitability Appraisal (EPA) has been applied at the pilot factory Werk150. Based on the cost calculations and determination of the work system values of different work system alternatives which are covered by the EPA, the value creation system with an optimized target achievement for changing production system requirements can be determined. Building blocks for the definition and calculations of the work system alternatives, further referred as value creation system to distinguish it from the logistic transport systems, are the locations of value creation (LVC) which are representing value adding production resources as workstations or machines and the locations of storage (LOS) for buffer storages in the production system. The result of the EPA is the definition of the target-optimized value creation system including the required material flow relations. After the determination of the most target-oriented work system configuration, the work system will be implemented and the developed autonomous control method will be adjusted to the new optimized layout and the defined material sources and sinks of the value creation system. The transport order allocation is than entirely done in an autonomous, target oriented manner following the defined optimization equation for cost and performance calculations of all available transport system (TS) of the logistics system to achieve a target system-optimized logistical process execution also in case of arising turbulences. In case of target-system changes or deviations from the defined target values (e.g. due to changed products or production numbers), the procedure of the EPA can be (manually) initiated again by the production manager to change the structural formation of the value creation system. This separate consideration of production and logistics targets and manual triggering of the restructuring processes is a pre-stage to a fully autonomous self-organized work system behavior, in which the need for a restructuring would have to be detected and initiated autonomously and the targets of production and logistics are considered simultaneously based on
adaptable source and sink relations in the production system (second step following the chosen research design).

**First approximation for versatile logistics**

![First approximation for versatile logistics with fixed sources and sinks](image)

**Legend:**
- EPA: Extended Profitability Appraisal
- LVC: Location of value creation
- LOS: Location of storage
- TS: Transport system
- Material flow

Figure 1: First approximation for versatile logistics with fixed sources and sinks

### 4.1 Extended Profitability Appraisal

The EPA procedure has been developed and tested during the Federal German government’s program for Humanization of Working Life (HdA - Humanisierung des Arbeitslebens) in collaboration of employers’ associations, trade unions and researchers in the 1980s [13]. The conventional methods of profitability and investment calculation are limited to the measurement of the profitability of technical investments and directly quantifiable monetary data as expenditures for technology, operating resources and personnel as well as estimated revenues. Important indirect monetary data such as the reduction of machine downtimes, increased flexibility of the work system, absenteeism and scrap costs for the monetarization and operationalization of these objectives is not taken into account in conventional of profitability and investment calculation methods [14,13]. For the development and comparison of different work system alternatives and the assurance of a high planning quality and acceptance of work system redesign activities, the integration and contribution of experienced employees with specific knowledge from different functional areas is of decisive importance. The EPA procedure differs significantly from conventional methods of profitability calculation by a holistic consideration of economic factors such as costs and performance as well as technical, organizational and personnel factors (work system value). The EPA is therefore divided into the two subsections of the economic efficiency comparison and the work system value determination. The determination of the work system value covers factors and aspects which can hardly or not at all be assessed monetary, whereas the economic efficiency comparison section covers purely monetary factors. At the end of the EPA both sub-ratings are summarized in a joint presentation of the results [14]. An overview of various diagnosis-oriented and decision-oriented EPA procedures can be found in [15].

Besides the method of the work system determination, the method of an argumentative balance sheet can be applied to assess monetary hard to quantify evaluation criteria. This balance sheet is used to list advantages and disadvantages in the spheres of “Effects on the production system itself” and “Effects of a system implementation regarding customer markets, customers and suppliers”. The argumentative balance sheet is a pure collection of arguments on advantages and disadvantages without any ranking possibilities of
alternative work system alternatives, therefore it can be only seen as an addition to the work system value
determination. The method of work system value determination uses the method of cost-benefit analysis and
is particularly well suited for the evaluation of work system alternatives [14]. The result of the work system
value determination also provides starting points regarding strengths and weaknesses of different work
system configurations. By a combination of advantages of single solutions from different investigated work
system alternatives, a target-oriented work system solution can be determined iteratively via various runs of
the work system value determination procedure.

For the developed and applied method the procedure of Bullinger [14] is applied to determine the work
system values (WSV) of different possible work system alternatives (also see Figure 2). The WSV
determination procedure starts with the selection and definition of the evaluation criteria based on the non-
monetary and/or hard to quantify targets which have to be fulfilled by the work system. The weighting of
the evaluation criteria in the next step is done by a pairwise comparison to calculate the weighting factors of
each criteria. The third step covers the determination of the fulfilment factors for every criterion and work
system alternative. The determination of the fulfilment values is done with a table matrix containing all work
system planning alternatives and evaluation criteria. The fulfilment of each criteria is estimated by an expert
team based on a point scale from 0 (Criteria is not fulfilled) to 10 (Criteria is fulfilled) for each work system
alternative after the other. The calculation of the work system value for each alternative is done by
multiplying all the fulfilment factors with the (normalized) weighting factors for every criterion and adding
up all the sub values resulting in the work system value for each planning alternative. The last step of the
WSV determination is the evaluation, identification of the work system with the highest WSV and
presentation of the results, e.g. via bar chart visualizations for a simplified comparison and interpretation of
the analysis results.

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Selection and definition of evaluation criteria</td>
</tr>
<tr>
<td>2</td>
<td>Weighting of the evaluation criteria</td>
</tr>
<tr>
<td>3</td>
<td>Determination of fulfilment factors per criterion and alternative</td>
</tr>
<tr>
<td>4</td>
<td>Work system value calculation</td>
</tr>
<tr>
<td>5</td>
<td>Evaluation and presentation of results</td>
</tr>
</tbody>
</table>

Figure 2: Procedure for work system value determination (cf. [14])

4.2 Autonomous control method

Autonomous control of logistical processes is given when the logistical object itself processes information,
makes decisions and executes them [9] [16]. The autonomous decision making and behavior in general of
intelligent logistical objects interacting with each other has to follow defined goals to achieve the desired
effectiveness and efficiency of the production system. The general goals of production logistics following
Wiendahl [17] are set by a maximization of logistics performance in form of a close adherence to delivery
dates and short throughput respectively delivery times as well as a minimization of logistics cost represented
by capital commitment costs and a high utilization. The cost and performance goals of specific intelligent
logistic objects as transport systems, transport units and services have to be defined and matched with their
(logistic) function to achieve the required target-oriented object behavior within the production system. In
the sense of self-organized systems, these targets as well as the systems elements and structure will change
over the time to maintain a viable, effective and efficient (production) system [18,19]. Based on the selected
work system/value creation system configuration determined by the EPA, the target systems (e.g. targets and
prioritization) of the intelligent objects have to be adjusted to the respective production system needs which
are linked to the evaluation criteria used for the WSV. For example, the goal of a higher logistics performance
can be dynamically prioritized higher than logistics costs in the target systems of the logistics objects in
production scenarios involving peaks in customer orders to achieve shorter lead times in the production
system. The results of the implementation and validation of the developed first approximation method within
a scenario at the pilot factory Werk150 for a target-oriented configuration of work systems in combination
with an autonomous control method for a versatile intralogistics system behavior will be described in the
following.

5. Validation Scenario at pilot factory Werk150

Pilot factory Werk150 is a close-to-reality research, education and training environment at ESB Business
School (Reutlingen University). Learning factories, such as the Werk150, covering a real value chain and
product have proven to be an ideal environment for the development and demonstration of future production
scenarios [20]. Holistic learning factories, like Werk150, are especially suitable for complex research topics
like the industry-oriented development of self-organization and autonomous control methods since state-of-
the-art industry infrastructure is available and at the same time, production downtimes within the learning
factories do not lead to any financial losses. So for testing and validation of the developed first approximation
for a cost and performance optimized value creation system configuration and target-system oriented
autonomously controlled logistics process execution, a versatile production scenario involving different
product-mixes, production numbers as well as new products has been defined.

5.1 Selection of target-optimized work system setup

Based on the incoming production orders and defined strategic goals of the pilot factory Werk150, the EPA
method covering economic efficiency comparisons and work system value (WSV) determinations for
qualitative factors which can hardly be assessed monetary has been applied to define the most effective work
system for order fulfilment in a versatile production scenario. For the WSV determination the procedure
described in section 4.1 has been followed. In the first step, the evaluation criteria for the work system
alternatives has been selected and defined to cover the requirements of versatile production systems with
changing production volumes, models mixes and turbulences like production and logistics resource
breakdowns as well as rush orders. Amongst others, the adaptability to volume fluctuations, sensitivity to
resource failure and flexibility regarding process changes have been considered. Afterwards the weighting
factors of each criterion has been calculated by applying the method of pairwise comparison.

In the next step different work system alternatives defining the assembly system (e.g. individual workstations
doing all required assembly steps, assembly lines with a division of tasks,…) as well as the logistics system
(e.g. pre-picking of customer individual product parts and delivery to dynamically selected assembly station
vs. provision of all parts at the workstation) have been developed and the fulfillment factors per weighted
criterion and alternative has been determined by a group of students and researchers. Next, the WSV and
cost and performance values of the different work system alternatives have been calculated. The process
steps of defining work system alternatives and calculating the WSV and cost and performance values based
on the developed equation system of the value creation system has been repeated based on the results of the
first runs to generate the target-optimized work system solution which is shown in Figure 3.
The assembly system consists of six assembly work stations of which two workstations can each carry out the same assembly steps (but with different tools or collaborative robots) to achieve a high flexibility regarding changing product variants as well workstation failures. The workstations are on wheels and therefore movable to change the assembly system layout e.g. into an assembly line to maximize output for higher batch production batch sizes. At all assembly workstations, only the required standard components or C-parts are kept in large quantities, all customer-specific components are either delivered pre-picked or delivered directly to the corresponding assembly workstation. The pre-picking is done at human-robot-collaboration workstations (see right below in Figure 3) for the scooter which is the multi-variant base product assembled at Werk150. The pre-picked components are directly placed by the robots as well as the human pickers on fixtures which are used for transport as well as for assembly purposes at the workstations. The pre-picked components are then transported from the picking workstations by a modular, decentrally controlled roller conveyor to a transfer point at the end of the roller conveyor system from where the scooter components are transported to the next workstation with free work capacity based on an autonomous control method for a target-oriented selection of the transport system (manual, semi-automated/collaborative, fully automated transport system).

### 5.2 Autonomous control of transport orders

As a first preliminary stage before a completely self-organized intralogistics scenario, the transport orders in the scenario described above are coming with predefined material sources and sinks (workstation with least remaining work) to test and validate the developed autonomous control method for a cost and performance optimized transport order allocation involving various alternative transport systems. The target measures to achieve economic efficiency by the selection of the appropriate transport systems within this method have been derived from [17] [21] and operationalized for the desired transport order allocation application. The logistics performance is measured by the target value of the adherence to schedule (deviation to due date of delivery) and transportation and waiting times. The logistics cost dimension is measured based on the transport system-related process costs aiming on a high capacity utilization. These performance and cost target dimensions are considered by the transport systems as well as the (intelligent) bins and transport units for close-to-real-time decision-making and execution of material transports. For example, if a bin with c-parts or a fixture with pre-picked scooter components has to be transported from its current location to a specific workstation, the intelligent transport unit (or in case of non-intelligent transport units the transport client) communicates directly with the transport systems of the work system at Werk150. The transport systems are then responding to this enquiry with their specific cost and performance values to fulfill the transport. In case of manual transport systems involving a human worker the enquiry is handled...
via a self-developed logistics worker client app, which keeps track of open, denied and accepted transport orders and also gives the worker the possibility to manually accept and reject transport orders. The transport unit or intelligent bin as the customer within this processes, then decides on the most favorable transport offer according to its target measures and communicates the decision to the transport systems. In addition to conventional automated guided vehicles, semi-automated transport systems (e.g. electric pallet trucks) as well as manual transport systems (e.g. human workers, handcarts) this autonomous control method is also applied for a collaborative tugger train system. In contrast to conventional tugger train systems which are driven by a human worker and also the manipulation of the goods is done by a human, this tugger train system is able to drive and manipulate goods autonomously without the need for predefined transport routes or transport schedules. The tugger train system consists of an autonomous robot platform which is towing the trailers of the tugger train. On top of the robot platform, a collaborative robot is mounted to pick small load carriers from shelves, place them on the trailers and unload them at the respective work stations. So the benefits of tugger train systems, like the possibility to fulfill high volume transport orders, can be combined with the potentials of autonomous controlled material flow systems.

A simulation study of this autonomous control method has already proven an increased target achievement of transport system related cost and performance goals for versatile production environments with turbulences (like breakdowns of transport systems). The simulation showed amongst others a lead time reduction of up to 30 % as well as significant improvements of the utilization of the transport systems and the adherence to schedule (also see Grosse-Erdmann [22 – to be published]). First practical tests of the autonomous control method described above involving an autonomous robot transport system and an autonomous collaborative tugger train system of the project “Collaborative tugger train 4.0” have already shown similar results in combination with the EPA of the first approximation for versatile logistics. Although the benefits considering costs and performance have shown to be especially significant in factory scenarios with arising internal turbulences like transport system breakdowns or unplanned rush orders which have to be fulfilled.

When the monitored cost and performance measures of the logistics system and/or the assembly system are falling below defined limits, the EPA procedure is restarted to restore a target-oriented work system configuration as described in section 4.1 following. This process starts again with the economic efficiency comparisons and the WSV determinations for different value creation system alternatives based on the incoming production orders and defined work system goals. After the determination of the work system alternative with the highest target fulfillment, the target systems of intelligent logistical objects are adjusted to the modified production system needs coming from the evaluation criteria for the WSV. Next, the defined work system is executed and the target achievement of the work system is monitored to maintain a target-oriented work system structure and behavior following the concept of self-organized systems.

6. Conclusion and outlook

The first approximation for a target-oriented work system design and autonomous execution of intralogistics processes for the changeable factory environment of the pilot factory Werk150 described in this paper has shown a significant potential to cope with changing production requirements of versatile production systems. One of the next steps will be covered in a second approximation which will be the further development of the described autonomous control method focusing on logistics goals towards a self-organized and autonomously controlled method covering cost and performance goals as well as the WSV of logistics and assembly/production in a joint equation system enabling target-oriented decisions with dynamic source and sink relations. Also the autonomous initiation of the work system (re-)configuration will become a system inherent feature following the approach of self-organized systems anticipating a target-oriented system (re-)structuring in an autonomous manner. Through this widening of the scope of self-organization and
autonomous control to the entire production system including the logistics processes, a target-oriented system behavior might be achieved also in case of arising internal and external turbulences in the versatile production system. The hypothesis to be proven at the end is that the extension of the scope of self-organization to production and logistics improves the overall goal achievement (costs and performance) of versatile factories.

References


Biography

Jan Schuhmacher, M.Sc, MBE, is a research associate at ESB Business School and part of the research, training and consultant group of the Werk150. His major field of research is the design and optimization of changeable intralogistics systems.

Vera Hummel, Prof. Dr.-Ing. has been a professor at the ESB Business School, Reutlingen University since 2010. She is a founding member of the “International Association of learning factories IALF” and the initiator of “Werk150 – the Factory on Campus of ESB Business School” for research, education and industry training.
Socio-technical requirements for production planning and control systems

Günther Schuh¹, Andreas Gützlaff¹, Frederick Sauermann¹, Matthias Schmidhuber¹
¹Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Germany

Abstract
Due to increasing customer requirements and intensifying competition, manufacturing companies are facing growing challenges in the successful order handling. As a result, employees are forced to make increasingly complex decisions in the shortest possible time. At this, the tasks of production planning and control (PPC) are particularly affected. In response to the increasing complexity of tasks, companies rely more than ever on the potential of socio-technical systems, rendered possible by the integration of information systems (IS) in the daily decision-making process. However, due to the increasing complexity of systems used, many users are not capable to raise the potential of information systems acquired, which is why the benefits of IS implementation often fall short of expectations.

The following paper thus analyses and structures potential decisive factors causing the lack of problem-solving capability in context of using PPC systems. Based on findings from acceptance research, socio-technical influencing factors for the targeted handling of information systems are determined. The developed requirement framework is furthermore compared with current IS implementation strategies to derive future research needs.

Keywords
Information systems; production planning and control; user acceptance

1. Introduction
Today's production environment is characterized by high fluctuations in sales, rising numbers of variants and the increased need to react flexibly to short-term disruptions [1,2]. Particularly affected by these developments are the task areas of PPC, which through their decisions and actions have an instantaneous influence on the logistic target achievement [3]. In order to cope with the increasing complexity, companies increasingly rely on the support of IS [1]. The PPC therefore increasingly develops into a socio-technical field of activity, which is characterized by system factors human factors and interface factors [4].

Despite the existing performance potentials and expectations, collaboration between IS and their user faces considerable problems in terms of potential utilization. True to Ashby's law, which argues that the variety of a control system must be as great as the variety of disturbances occurring to the system to be controlled, commercial IS themselves are characterised by increasing functionalities and complexity [5]. The rising complexity of advanced IS in return is contradicted by human factors striving for simple solutions and therefore complicates the targeted use by employees [3,6]. As a consequence of this mismatch, in practice decision support of IS oftentimes remains ignored, leading to manually adapted order sequences or work in progress (WIP) levels being actively held high [3].

DOI: https://doi.org/10.15488/9658
In reaction to the experienced difficulties with IS collaboration, several research endeavours focussed on identifying decisive influencing factors for a successful IS implementation, looking beyond the oftentimes solely technical oriented approaches [7]. At this, critical success factors (CSF) were identified using empirical research methods [8]. However, given persistent reports of lacking performance improvements throughout the IS lifecycle, criticism on consisting IS implementation strategies and the usability of CSF remains [9][10]. In addition to increasing the applicability of implementation strategies, the need for research is seen in particular in the identification and handling of human factors [4]. Addressing these deficits, this paper is structured as follows: In section 2, the results of a literature review illustrating the current research deficit regarding the use of IS are presented. In section 3, existing models describing influencing factors addressing general technology acceptance are analysed, in order to develop a framework of socio-technical requirements for PPC systems. In section 4 this framework will be matched to existing IS implementation strategies in order to derive future needs of action. The summary and an outlook for further research will be given in section 5.

2. Literature Review

2.1 Human system partnership in PPC

The performance of manufacturing companies is, in addition to the products and services offered, primarily determined by the degree to which PPC is capable of meeting logistical targets [3]. Due to the mutual dependencies of the logistic target values, clear and transparent decision-making situations are rarely given, which severely increases the complexity to the decision maker [3]. Given the limited capabilities of dealing with complexity and uncertainty, however the decision-making process often solely takes place within an employee’s individual observation horizon and thus under pursuit of local rather than global, company-wide optima [1,3].

The desire of organizations to make decisions that take into account all aspects of the business is a major reason for the high degree of penetration of IS [11]. According to Alter (2008), IS can be defined as work systems focussing on the processing of information, thus providing support for an organization’s network in terms of information creation, gathering, processing or storing [12]. As a subclass of IS, PPC systems act as central logistical control mechanism, matching a company’s output to customer demands. The basic tasks of PPC systems thus include the planning, releasing and controlling of production orders as well as the monitoring and readjusting of production orders and production plans in case of unforeseen disturbances and deviations. [6] Enterprise resource planning (ERP) systems are among the most widespread and researched IS in manufacturing companies. The primary functions of ERP systems include enabling and processing of business transactions, furthermore addressing the problem of fragmentation by integrating internal processes throughout the entire company. [9] Based on the fundamental functions of ERP, the goal of manufacturing execution systems (MES) is to bridge planning and management systems at enterprise level to control systems on the shop floor, thus enabling an intermediate level of passing information in real time [13],[14,15]. In accordance with the diversity of the business units involved in the order processing, several additional IS systems have been established under the system classifications of operational systems, planning systems, management information systems and cross section systems [16].

The most influential shared ability of IS, regarding the collaboration between user and system, is automation. Automation can be defined as a technology performing tasks, which were previously done by humans [17]. Automation thus bares the potential to extend an employee’s physical and cognitive capacities, enabling a joint achievement of objectives [18]. As described by Bainbridge as “Ironies of Automation”, the role of the user at this becomes, due to unavoidable supervisory functions left with the user, more, rather than less, important within this collaboration [19]. Crucial for an efficient interplay is an holistic automation design which considers characteristics of the joint cognitive system emerging from the combination of humans and
automation [18]. Depending on the fulfilment of expectations and requirements from the support operators, determining the individual levels of trust and acceptance towards automation, collaboration with automation can be differentiated between purposeful use, misuse and disuse [20]. At this, disuse due to limited trust describes the case, if employees do not use IS or decline decision support offered by IS. On the contrary, misuse related to an excessive level of trust describes the case, if employees follow automation despite obvious malfunction, due to a missing challenging of results. [20]

2.2 Adoption strategies of information systems

In order to address the complexities accompanied by IS collaboration, several research endeavours focus on adoption and management strategies. In accordance to the life cycle of IS, a common distinction is to be found between research approaches focussing on the pre- or post-implementation phase of IS [21]. The pre-implementation comprises the life cycle phases of adoption decision, acquisition and implementation, whereas the post-implementation phase consists of use and maintenance, evolution as well as retirement [22–24].

Major problems related to the pre-implementation phase can be related to mismatches between IS functionalities and company requirements [25]. At this, requirements management is an essential task of the pre-implementation phase [26]. A requirement at this can be defined as prerequisite or ability a system must fulfil [27]. However, existing approaches to requirements management often face criticism due to primarily technical or cost-oriented perspectives [7]. It is also for the neglect of socio-technical aspects, that around 40 % of the efforts in the IS development or configuration process are related to the implementation of changes [28]. In response to IS performance criticism, several research efforts on pre-implementation strategies accompanied and analysed actual IS implementations projects in order identify CSF [8]. CSF can be understood as those conditions, which must be met in order for the implementation process to work successfully [25].

A considerable proportion of manufacturing companies face problems with IS utilization especially in the post-implementation phase. According to industry reports, 57 % of industrial organisations report considerable process stoppages due to IS problems in the post-implementation phase [29] and 67 % report on missed performance expectations [30]. These failures are among other things attributed to a lack of organizational attention, after overcoming a resource and time intensive pre-implementation phase, and the consequent lack of working on the establishment of processes [31]. A similar shift of awareness also prevails the research agenda, as Esteves et al. (2007) showed, that of all ERP-related articles, 47 % address pre-implementation strategies whereas only 15 % address the post-implementation phase [32]. It is therefore, that a growing number of research articles start focussing on success-related factors influencing the post-implementation phase. At this, DeLone and McLean’s information systems success model is the research framework primarily used to determine influencing factors in post-implementation [33] [9,21,31,34–37].

2.3 Research Deficit

Given cause to a persistent research deficit regarding IS adoption and usage is found in lasting company reports of unsuccessful IS implementation and user cycles phases. Cases such as Nike’s ERP-failure which cost $100 million in sales and resulted in a drop of stocks in 20 % [38] are thus no rarity, given evaluations such as from Deloitte, which state that of Fortune 500 companies, 25 % are struggling with ERP adoption and performance [36].

Reasons for the persistent problems in dealing with IS match criticism on the IS adoption research, facing inadequate interpretation as well as missing generalizability and applicability of influencing factors throughout the life cycle [10]. Facing interpretation issues, several research contributions state, that sufficient effort was made to address system factors in influencing models, not so however to analyse human factors, consequently biasing conclusions on performance constellations [39]. A further issue of extant adoption
literature is the biasing influence of the respective point of view, induced by case related empirical research methods used and therefore hindering the generalisability and thus operational application of influencing factors [8]. This point of criticism is reaffirmed by the strong predominant system focus of ERP systems. Very few research endeavours analyse influencing success factors on ERP system supplements such as MES or advanced planning systems (APS) [40]. Due to the continuing failures and criticisms, the question of the requirements for a successful collaboration between users and PPS systems, especially with regard to socio-technical interactions, continues to arise.

3. Socio-technical requirements for user acceptance

In response to the persistent criticism, the following paper will shift the focus of IS adoption from the company perspective to the individual user perspective, analysing factors influencing the collaboration between user and IS. This is in line with research and practice insights, according to which the employee adoption and use rate are still major barriers for the success of IS [41].

3.1 User acceptance models

Research on IS has long studied the rationale of how and why individuals adopt new technologies [42]. One of the early but still widespread model is the Innovation Diffusion Theory (IDT) by Rogers (1995) [43]. Diffusion in this context can be defined as the process by which technology spreads across a population of organizations [44]. The IDT thus represents an empirical construct from the field of sociology that enables the evaluation of user acceptance of technical innovations and thus investigates the question of how, why and at what rate innovative ideas and technologies prevail in social systems. [44][42]. As relevant influencing variables on innovation acceptance Rogers identified: relative advantage, compatibility, complexity, trialability and observability.

As another representative of the user acceptance research area, Goodhue and Thompson (1995) developed the Task-Technology Fit model (TTF) [45]. The objective of the TTF is to predict technology acceptance by means of increasing the individual performance of its user [45]. For this purpose, the theoretical model draws from the supplementary research insights, using user attitude as predictor of utilization and task-technology fit as a predictor of performance [45]. As relevant influencing variables on task-technology fit the TTF identified: quality, locatability, authorization, compatibility, ease of use, production timeliness, systems reliability and relationship with users.

Within the framework of the Unified Theory of Acceptance and Use of Technology (UTAUT), Venkatesh and Davis (2003) worked towards the aggregation of several user acceptance work streams. The resulting theoretical model merges individual, organizational as well as work related perspectives on user acceptance [42]. As relevant influencing variables being suitable to aggregate in the UTAUT Venkatesh and Davis identified: performance expectancy, effort expectancy, attitude toward using technology, social influence, facilitating conditions, self-efficacy and anxiety.

Being among the most applied theoretical models explaining user acceptance is the Technology Acceptance Model (TAM), being initially developed by Davis in 1989 [46]. Drawing from the theory of reasoned action (TRA), Davis acknowledged every behaviour being preceded by a behavioural intention which is influenced by attitudes and subjective norms [46]. Transferred to his research subject, Davis identified the variables of perceived usefulness and perceived ease of use as prerequisites for user acceptance and hence technology usage [41]. In continuation of the research results, the TAM was extended by variables explaining the influence on the construct of perceived usefulness (TAM 2) [47]: computer self efficacy, perception of external control, computer anxiety, computer playfulness, perceived enjoyment and objective usability, as well as the on the construct of perceived ease of use (TAM 3) [41]: perceived ease of use, subjective norm, image, job relevance, output quality and result demonstrability.
3.2 Synthesis of relevant requirements for user acceptance

Analysing the results of the user acceptance models shown in section 3.1, a total of 32 empirical constructs influencing the user acceptance on the individual level were identified. After the consolidation and exclusion of the redundant influencing variables across the models, the following 18 socio-technical influencing variables on user acceptance were derived. The result Table 1 lists all requirements, including a model assignment and a description aggregated from the models.

Table 1: Socio-technical acceptance requirements

<table>
<thead>
<tr>
<th>#</th>
<th>TAM 3</th>
<th>TFF</th>
<th>IDT</th>
<th>UTAUT</th>
<th>Requirement</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>Anxiety</td>
<td>User are not exposed to apprehension or fear when facing the possibility of using the system and thus the risk of incorrect use.</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>Appropriate Complexity</td>
<td>Using and understanding an innovation is perceived as an appropriately difficult task.</td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>Compatibility</td>
<td>The innovation is perceived as being consistent with the existing values, needs and past experiences of potential adopters.</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Data Locatability</td>
<td>User are easily capable of determining what a data element on a report or file means, or what is excluded or included in calculating it.</td>
</tr>
<tr>
<td>5</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>Data Quality</td>
<td>Data meets user needs in terms of currency, maintaining the right level of data as well as the right level of detail.</td>
</tr>
<tr>
<td>6</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>External Control</td>
<td>User have access to organizational and technical resources to support the use of the system.</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>Intrinsic Motivation</td>
<td>User have an intrinsic, positive attitude towards using technology.</td>
</tr>
<tr>
<td>8</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>Job-Fit</td>
<td>The system fits the users’ job/activity requirements and corporate goals.</td>
</tr>
<tr>
<td>9</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>Output Quality</td>
<td>The system performs its tasks according to the job requirements.</td>
</tr>
<tr>
<td>10</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>Production Timeliness</td>
<td>The system meets pre-defined production turnaround schedules.</td>
</tr>
<tr>
<td>11</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>Relative Advantage</td>
<td>Using an innovation is perceived as being more advantageous as using its precursor.</td>
</tr>
<tr>
<td>12</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>Result Demonstrability</td>
<td>The results of system usage are tangible, observable, and communicable.</td>
</tr>
<tr>
<td>13</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>Social Influence</td>
<td>The variables determining the social variables are aligned to the acceptance and use of the systems. This includes subjective norm (addressing people being important to the user) and image (addressing the social status within the company).</td>
</tr>
<tr>
<td>14</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>System Self Efficacy</td>
<td>The supporting functions of the system allows users to complete a system job or task in case no external support is available.</td>
</tr>
<tr>
<td>15</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>System Reliability</td>
<td>The system is dependable in terms of access and uptime.</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>Trialability</td>
<td>Innovations can be carried out on a limited basis prior to adoption.</td>
</tr>
<tr>
<td>17</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>User Authorization</td>
<td>User obtain the authorization to access data necessary to do their job.</td>
</tr>
<tr>
<td>18</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>Usability</td>
<td>Daily usage of a system is associated with little effort and learning a skillful handling of the system is easy.</td>
</tr>
</tbody>
</table>

4. Comparison of socio-technical acceptance requirements with IS implementation strategies

In order to compare the identified socio-technical acceptance requirements with IS implementation strategies, a literature research was conducted to identify current pre- and post-implementation approaches. The literature research was carried out using the common search engines and literature databases: google scholar, researchgate, sciencedirect, and excluding contributions with a publication date prior to 2010. A
combination of the following terms was used as keywords for the search: IS OR information system OR ERP OR MES OR APS AND implementation OR pre-implementation OR post-implementation OR success factors OR performance assessment.

4.1 Research contributions on pre- and post-implementation strategies

Addressing the pre-implementation process, Hailu and Rahman (2012) reviewed industry and academic literature in order to evaluate key success factors influencing ERP implementation [48]. Dezdar (2012) used the survey approach to collect data throughout Iranian companies in order identify tactical and strategical factors being crucial for a successful ERP implementation [49]. Behehsti et al. (2015) used a qualitative research method to study six diverse manufacturing companies in the US in order to identify CSF for ERP implementations [50]. Chatzoglou et al. (2016) used empirical methods in order to test a conceptual framework including factors enabling a successful ERP implementation especially for Small and Medium Enterprises [51]. To the best of our knowledge, Lee et al. (2012) contributed the only research paper focussing on critical success factors for the MES implementation process, therefore conducting a survey throughout 163 manufacturing companies in South Korea [14].

Addressing the post-implementation process, Ha and Ahn (2014) evaluated factors affecting the post-implementation performance of ERP systems using pilot studies throughout Korean companies [36]. Hecht et al. (2013) conducted an extensive literature research in order to identify factors influencing the post-implementation success of ERP systems and to derive requirements and capabilities for ERP maintenance [21]. In their research contribution, Hsu et al. (2015) analysed how different qualities of ERP system affect the post-implementation success, especially regarding the user perspective [31]. Ifinedo et al. (2010) investigated the relationship among six models of ERP post-implementation success measurement models from an organizational level [37]. Using a fuzzy analytic network process, Moalagh and Ravasan (2013) developed a practical framework for assessing companies’ ERP post-implementation success [9].

4.2 Comparison of socio-technical requirements with implementation strategies

The results of the comparison of the socio-technical acceptance requirements derived in section 3.2 with the implementation strategies as outlined in section 4.1 are shown in Table 2. Harvey balls following the semantic depicted in Figure 1 were used in order to assess the comparison.

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requirement is directly addressed through a corresponding requirement in the implementation strategy.</td>
</tr>
<tr>
<td>Requirement is directly addressed through a corresponding action described in the implementation strategy. (anxiety - user training and education)</td>
</tr>
<tr>
<td>Requirement is indirectly addressed through a corresponding action or requirement in the implementation strategy. (intrinsic motivation - user involvement)</td>
</tr>
<tr>
<td>Requirement is loosely addressed through a generic action or requirement in the implementation strategy. (social influence - top management support)</td>
</tr>
<tr>
<td>Requirement is not addressed by the implementation strategy.</td>
</tr>
</tbody>
</table>

Figure 1: Semantic evaluation model

The results of the comparison show, that the implementation strategies mostly cover the basic technical aspects identified in the acceptance requirements, such as data quality, output quality, usability, system reliability and trialability. Furthermore directly addressed through corresponding actions and requirements, details in parenthesis, were the organizational requirements of external control (IT & vendor support) and job-fit (user involvement & user requirements) as well as the user related requirements of anxiety (training), intrinsic motivation (job-enhancement & user satisfaction) and production timeliness (timeliness &
performance). Not or mainly generically addressed were the mainly organizational and user related requirements of relative advantage, compatibility, appropriate complexity, result demonstrability, data locatability, user authorization, system self efficacy and social influence.

Table 2: Evaluation of IS implementation strategies

<table>
<thead>
<tr>
<th>#</th>
<th>Requirement</th>
<th>Pre-Implementation</th>
<th>Post-Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Relative Advantage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Compatibility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Approp. Complexity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Trialability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Res. Demonstrability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Data Quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Data Locatability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>User Authorization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Usability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Job-Fit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Prod. Timeliness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>System Self Efficacy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>External Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Output Quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Anxiety</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Intrinsic Motivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Social Influence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>System Reliability</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. Summary and Outlook

Employee adoption and user rates of IS are still today major barriers for IS success [41]. This study thus examined the most well-known user acceptance models and derived an aggregated framework of socio-technical influencing factors on IS user acceptance. In order to examine the existing criticism of IS implementation strategies, the study furthermore checked the identified acceptance requirements for conformity with the findings and recommendations of current research endeavours on implementation strategies. The results lead to the following three conclusions for current and future IS research: (1) Implementation approaches only take into account a section of the relevant influencing factors (missing completeness); (2) Implementation approaches lack concrete feasibility for companies (missing operationalisability); (3) IS implementation approaches in the PPC environment neglect advanced systems like APS, MES, etc. (ERP focus).
As shown in the matching overview in Table 2, the aggregated implementation strategies together, cover entirely only seven of the 18 identified acceptance requirements. Accordingly, none of the individual approaches achieves a degree of coverage of more than 33 %. As has already been the subject of criticism, the focus here is mainly on the technical requirements of user acceptance [4]. Acceptance requirements of the user- and organization-oriented perspective remain for the most part neglected. Thinking in terms of IS utilisation (use, disuse and misuse), aspects such as IS rejecting social structures or authorities (social influence), restricted IS authorization (user authorization) or non-transparent relative advantage of IS compared with existing customised solutions (relative advantage) pose a major risk for disuse of introduced systems. Requirements such as a self-explanatory support functions within IS (system self-efficacy) or user knowledge about the design of the operational information structure and its operating principles (data locatability) furthermore are fundamental in avoidance of system misuse.

Restrictions of the operationalisability of current implementation strategies become transparent, especially in view of their application to the entire IS life cycle. While many IS success models of post-implementation approaches feature a comparatively high level of detail in the sense of requirements management, CSFs of pre-implementation approaches are characterised by a high level of abstraction, which is why factors such as top management support or vendor support remain intangible. Future IS research should therefore be directed at holistic implementation strategies covering the entire IS life cycle (pre- and post-implementation) as well as the concretisation and operationalisation of CSF in the sense of controllable requirements.

The results of the literature research on existing IS implementation strategies further indicate a dominant focus on ERP systems. As for the search criteria used, this study only found one research paper focusing on success factors for MES implementation and no research contributions addressing success factors relating to APS. The question of PPC-stakeholder requirements for dealing with IS, especially including blue collar user close to the shopfloor, therefore remains largely unaddressed and demands further research focus.

Acknowledgement

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References


[23] Achargui, ZaouiaHosted, cloud and SaaS, off-premises ERP systems adoption by Moroccan SMEs, A focus group study, 344-348.


Analytical model for determining the manual consolidation time for large equipment manufacturers

Jakob Schyga\textsuperscript{1}, Philipp Braun\textsuperscript{1}, Svenja Hilbrich\textsuperscript{1}, Johannes Hinckeldeyn\textsuperscript{1}, Jochen Kreutzfeldt\textsuperscript{1}

\textsuperscript{1}Institute of Technical Logistics, Hamburg University of Technology, Hamburg Germany

Abstract

Large and complex products such as aircrafts or wind energy plants are usually produced in small batches. This leads to particular requirements for logistics with regard to process efficiency, flexibility and reliability as well as coordination and control of the processes. In practice, the consolidation of parts from different storage areas is a labour-intensive process, which must be considered in the planning or remodelling of warehouses. This work presents an analytical model to determine the consolidation time within a single formula at an early planning stage. Within this formula transport, handling, base and allowance times are considered. Finally, the analytical model is applied in an industrial project which deals with the new planning of warehouses for a large equipment manufacturer. The presented work connects academic approaches for calculating the commissioning time with practical experiences gained in industrial projects. Researchers benefit from this work, as practical considerations and corresponding solutions are pointed out and insights into practical projects are given, while logistics planners could benefit by applying the developed model.

Keywords

Logistics planning; Consolidation; Analytical model; Large equipment manufacturer

1. Introduction

Personnel costs often dominate the proportion within the cost structures of a warehouse. The calculation of logistics personnel is therefore an essential component in the planning of warehouses or the remodelling of logistics processes. The commissioning process, i.e. the combination of subsets from a total quantity of goods on the basis of requirements [1] is a labour-intensive process, due to the high complexity in material handling. In particular for large equipment manufacturers, manual handling and transport of goods is common, due to low production rates and high requirements for flexibility [2]. Consolidation describes a part of the commissioning process that serves to merge goods, picked from several storage systems. Especially at an early planning stage, the planning data base is often insufficient for a precise calculation of the required consolidation personnel [3]. During the course of a warehouse planning project for a large equipment manufacturer, the need for a model to calculate the effort in consolidation became apparent. Hence, an analytical model of manual consolidation processes was built and subsequently applied in order to compare different storage scenarios as part of the planning project. The model extends the established considerations for commissioning, as formulated by ten Hompel [4], among others. The particularity of the
model consists in process-specific adaptions and simplifications for the consolidation process of large equipment manufacturers, resulting in a single equation for the calculation of the consolidation time.

The paper is structured as following. The requirements on the commissioning process and in particular on the consolidation process for large equipment manufacturers are described in section 2. Subsequently, section 3 briefly describes the state of the art, puts the model into context and points out gaps in the present literature. The analytical model is presented in section 4 and the exemplary application during the planning project is described in section 5. The paper closes with a conclusion and an outlook on future works.

2. Requirements and description of the commissioning process

The production schedules of large equipment manufacturers are determined by a large variety of variables, which are often hardly predictable and difficult to combine in the operative production and logistics processes [5]. The dependency on delivery schedules, the availability of labour and technical resources as well as multiple production stations with diversified functions require efficient and responsive logistics processes [6]. Due to production structures demanding for geometrical diversified parts in small batch sizes, the logistics department must offer high flexibility in its processes and storage systems [7]. Consequently, manual labour is common.

As exemplary depicted in Figure 1, the process of commissioning is initiated by the placement of work orders based on the production schedule containing articles from several article families stored in different storage systems. First, the required parts for a work order are picked from a storage system. Therefore, several storage systems containing different article families, such as shelving racks for standardized storage boxes or honeycomb racks for long goods, need to be addressed by specific picking orders. Inside the storage systems, the parts are picked and aggregated to collection units. During this process, labour is required to pick the requested parts from the storage system and to perform additional handling steps like repacking and labelling of the parts [4]. After completing the steps of the picking process, the finalized collection units must be transferred to their designated consolidation buffer. Subsequently, the ordered articles need to be consolidated as shipping units in predefined load carriers and delivered to the corresponding work station in the production [8].

The consolidation process of the collection units is initiated when all units required for a specific work order are finalized and available at the respective consolidation buffers. The outcome of the considered consolidation process is a shipping unit which merges all collection units from the addressed storage systems.

Figure 1: Consolidation of shipping units according to work orders.
to a single load carrier per work order. The appropriate load carrier is predefined depending on the geometric structures of the collection units in order to optimize material transport or handling. After setting up the load carrier, an employee in the consolidation area starts walking to all consolidation buffers storing collection units for the current work order. At every consolidation buffer the required collection units must be searched and placed safely and efficiently on the load carrier. After gathering all collection units, the employee must complete the load carrier by performing additional steps like printing labels for the work station in the production and finally bringing the shipping unit to the dispatch area. At this point the consolidation employee is available to process the next work order.

3. State of the art and revealed gaps

The calculation of the required personnel is part of every production and logistics planning method. Established planning methods by Wiendahl [8], Kettner [9], Grundig [10] and VDI guideline 5200 - part 1 [11] exist among others. The methods can be clustered into four phases: project setup, structuring, system design and realization. The aim of the structuring phase is to create a holistic concept of the planning object [3]. An important part of this planning phase consists in the determination of personnel requirements, which are further detailed in the structuring phase. A commonly applied way of calculating the personnel requirements consists in the determination of the overall work effort (time requirements per work unit) of all manual activities [10]. Therefore, the manual activities need to be defined and their time requirements quantified. Established analytical methods for calculating commissioning efforts are given by Martin [12], Gudehus [13] and ten Hompel [4]. The named authors split the commissioning time into base time, transport time, handling time and dead time as illustrated in Figure 2. In VDI 4481 [14] the component allowance time is included to consider performance lowering factors such as work interruption for distractions or disturbances.

Ten Hompel [4] describes the processes and the analytic calculation of the commissioning time for different processes such as conventional person to goods commissioning, commissioning on the high rack or the commissioning on an Automated Storage and Retrieval System (ASRS). The calculations are based on average working times per part position for all named components of the commissioning time. The process of conventional person to goods commissioning is partly analogous to the considered consolidation process described in section 2. However, rather than walking along the aisles of a shelving system, the employee moves along the transport routes in the warehouse to merge collection units from the consolidation buffers instead of the picking units from a storage location.

From the particularities of the process described in section 2 arise special requirements for the calculation of the commissioning time, which cannot be fully satisfied by the given methods. First, the considered process is not analytically described in the common literature. The existing methods therefore need to be adapted to the described consolidation process. Second, for large equipment manufacturers the determination of average values per collection unit is difficult. The planning database is often inadequate and characterized by many uncertainties in particular at an early planning stage [3]. A detailed calculation of the

<table>
<thead>
<tr>
<th>Commissioning time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base time</td>
</tr>
<tr>
<td>• Administrative activities</td>
</tr>
<tr>
<td>• Takeover/handover of handling unit/load carrier</td>
</tr>
<tr>
<td>Transport time</td>
</tr>
<tr>
<td>• Walking, standing, driving (with lifting)</td>
</tr>
<tr>
<td>• Positioning effort</td>
</tr>
<tr>
<td>Handling time</td>
</tr>
<tr>
<td>• Grabbing</td>
</tr>
<tr>
<td>• Taking up</td>
</tr>
<tr>
<td>• Carrying</td>
</tr>
<tr>
<td>• Putting down</td>
</tr>
<tr>
<td>Dead time</td>
</tr>
<tr>
<td>• Searching</td>
</tr>
<tr>
<td>• Counting</td>
</tr>
<tr>
<td>• Labelling</td>
</tr>
<tr>
<td>• Weighing</td>
</tr>
</tbody>
</table>

Figure 2: Division of commissioning time according to Martin [13], Gudehus [14] and ten Hompel [4].
commissioning time is often not feasible or necessary at an early stage of planning. In order to calculate the approximate consolidation time, assumptions and simplifications have to be included into the model. As an example, the transport time for acceleration and deceleration due to manual transports over long distances is very low for the given process and can therefore be neglected. As the process times of the components of the consolidation time are highly varying for large equipment manufacturers, the use of average workloads per collection unit limits the comprehensibility of the calculation and therefore the potential for identifying optimization measures. Furthermore, the division of the handling time or dead time into process steps, such as labelling or carrying, require a detailed description of the processes, which are in practice usually not defined at an early planning stage. This detailing rather takes place later as part of the system planning. A fine-grained calculation model is therefore not appropriate at an early planning stage.

4. Analytical model for calculating the consolidation time

For the analytical model the process described in section 2 is considered. Consolidation orders are not triggered until the required collection units are placed into the consolidation buffers. In order to calculate the distances for transportation the layout has to be previously defined. Furthermore, the approximated work order data are required for a representative time period. The larger the chosen period the more accurate is the model. The recommended minimal time period is depending on the dynamics of the production structure. At this stage of planning it is also possible to use estimated values in order to receive first results as an approximation. To determine the time components of the respective processes, time recordings from already existing commissioning systems can be used. A table with a brief description of all the used variables for the analytical model can be found in the appendix.

To reduce the complexity of the model the consolidation time can be calculated per individually defined time interval. Therefore, the calculation of average values per position for the components of the consolidation time can be avoided. The calculations for the components are simplified according to the considered process and compromised in a single formula. The consolidation time ($t_{consat}$) is the result of the sum of the total transport time ($t_{transport}$), handling time ($t_{handling}$), base time ($t_{base}$) and allowance time ($t_{allowance}$) for the defined time interval. The handling time thereby includes the dead time. The allowance time is usually given in percent by the efficiency factor ($x_e$) of the system. The equation for the consolidation time is

$$t_{consat} = t_{transport} + t_{handling} + t_{base} + efficiency$$

whereby the individual components of the equation are described below.

The transport time for each transport can be calculated as the distance travelled ($d$) divided by the average velocity ($v$). To calculate the total transport time per time interval, the obtained value needs to be multiplied with the amount of transports ($i$) per given time interval. As previously described the considered consolidation process consists of the manual merging of collection units from several consolidation buffers. The consolidation buffers are indicated from 1 to $n$. The starting point to pick up the empty load carrier is indicated by 0 and the handover point for the shipping unit by $n + 1$. For each route from point $i$ to point $j$
the travel distance \((d_{ij})\) and the transport intensity \((i_{ij})\) are entered into the separate matrices \(D\) and \(I\), resulting in

\[
t_{\text{transport}} = \frac{1}{v} \sum_{j=0}^{n+1} \sum_{i=0}^{n+1} d_{ij} \cdot i_{ij},
\]

with \(D = \begin{pmatrix} d_{00} & \ldots & d_{0n} \\ \vdots & \ddots & \vdots \\ d_{01} & \ldots & d_{nn} \end{pmatrix}\) and \(I = \begin{pmatrix} i_{00} & \ldots & i_{0n} \\ \vdots & \ddots & \vdots \\ i_{1} & \ldots & i_{mn} \end{pmatrix}\).

Exemplary, the transport process with its parameters and the derived matrices are illustrated in Figure 3. The distance \(D\) can be inferred from the layout, while \(I\) can either be estimated or concluded from existing work order data. Since only manual transport is considered, the velocity of the picker is assumed to be constant regardless of the route section. Acceleration and deceleration can be neglected for manual transport. Commonly a value between 0.8 m/s and 1.1 m/s is assumed for walking velocity.

In the presented model the dead time is included into the handling time per time interval. The selected time interval has to correspond to the considered unit for the transport time. The time for handling mainly differs depending on the physical time components such as for grabbing or putting down of the collection units. As described in section 2, an article family, e.g. small parts, long or pallet goods, is commonly stored, each in an appropriate storage system. Therefore, consolidation buffers usually contain similar collection units. As a reasonable assumption at an early planning stage, the handling time for the merging of a collection unit \((h_i)\) is mainly depending on the consolidation buffer \(1\) to \(n\). The handling time for merging all collection units during the chosen time interval from a certain consolidation buffer onto different load carriers is given by the amount of collection units to be merged from each consolidation buffer \((c_i)\) times the handling time per collection unit \((h_i)\), resulting in

\[
t_{\text{handling}} = \sum_{i=1}^{n} c_i h_i
\]
Figure 4 exemplary illustrates the handling process and the components of the given equation. To merge the collection units of work order 2 onto a common load carrier four collection units need to be handled, whereby two of the collection units are merged from consolidation buffer 1. To calculate the total handling time for both work orders 1 and 2 the single handling times $h_i$ need to be summed up as

$$ t_{\text{handling}} = \sum_{i=1}^{n} c_i h_i = 3h_1 + 2h_2 + 2h_3. $$

The base time consists of administrative and manual activities related to take over or hand over of the load carrier. Since a consolidated load carrier builds the work order in the considered process, the total base time $t_{\text{base}}$ per time interval is the sum of the base times $b_{wo}$ for all work orders ($wo = 1 \ldots m$). In a production providing consolidation process as described in section 2 different types of load carriers, e.g. pallets, roll cages or special load carriers are used. As a simplification it can be assumed that $b_{wo}$ is identical for the same type of load carrier. The formula for the base time is therefore given by

$$ t_{\text{base}} = \sum_{wo=1}^{m} b_{wo} \quad \text{with} \quad \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}. $$

The efficiency factor $x_e$ can either be chosen by standard values or on the basis of time measurements of existing logistic processes.

5. Application of the developed model in a warehouse planning project

The developed model was applied as part of a warehouse planning project for a large equipment manufacturer. The aim of the project was to store highly varying types of goods in a centralized warehouse in order to supply the production efficiently with a high flexibility, reliability and responsiveness. The total warehouse area is about 40,000 m². The considered parts from the warehouse are picked from several storage systems, consolidated on the appropriate load carriers and pushed to the production. In the production the load carriers are buffered and pulled to 20 different work stations on the assembly line.
For the evaluation of different storage scenarios, the total investment and operational costs had to be calculated as part of the concept planning under a high uncertainty of the planning data. The scenarios were consisting of different combination of storage systems for storing small part containers, long goods and pallet goods including the processes for put away, picking and consolidation. Depending on the scenario the personnel costs make up a dominant part of 70 % to 90 % of the total costs for 10 years. The calculation of the consolidation time according to the developed analytical model is exemplary stated for the following scenario: Small goods are stored into an ASRS, long goods in storage lifts and pallet goods in pallet racks. Figure 5 shows a reduced section of the proposed layout and depicts the affiliated consolidation process. Goods from the ASRS are picked into containers. Subsequently these containers are transported back automatically into the ASRS to be temporarily stored for buffering. The containers can be retrieved from the system when needed and automatically be transported to consolidation buffer 1. The goods from the pallet racks are picked by forklift and transported to consolidation buffer 2. The long goods are picked directly from the storage lifts to the shipping unit load carrier. Long goods can therefore already be seen as a collection unit and the storage lifts as consolidation buffer 3. Table 1 summarizes the scenario by affiliating the article families with their dimensions to the storage system and their number of storage locations.

<table>
<thead>
<tr>
<th>Article family</th>
<th>Dimensions</th>
<th>Storage System</th>
<th>Storage Locations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small parts</td>
<td>Container $&lt; 600 \times 400 \times 220 \text{ mm}$</td>
<td>ASRS</td>
<td>40,000</td>
</tr>
<tr>
<td>Long goods</td>
<td>$1200 \text{ mm} &lt; \text{Length} &lt; 1800\text{ mm}$</td>
<td>Storage lift</td>
<td>5,000</td>
</tr>
<tr>
<td>Pallet goods</td>
<td>Packages $&lt; 800 \times 600 \times 500 \text{ mm}$</td>
<td>Pallet rack</td>
<td>3,000</td>
</tr>
</tbody>
</table>

For each product 57 differently composited shipping units are needed, which comprise one or more types of article families. The transport order and distances were concluded from the layout. As all transports are carried out manually by walking, the transport velocity is assumed to be constant with $0,9 \ \frac{m}{s}$. In order to set up matrix $I$ the composition of the 57 load carriers was considered. The exact composition varies due to customization of the product, but the contained article families and therefore consolidation buffers, that need to be visited are consistent. Each month around 160 products are completed, thus 8 per working day. The time interval for the calculation of the consolidation time was chosen to be one working day. The values for calculating $t_{transport}$ are given as
In average there are about 700 small parts containers, 32 long goods and 147 pallet goods required from the planned warehouse for each product. The handling times $h_1$, $h_2$ and $h_3$ per collection unit for the consolidation buffers 1, 2 and 3 were determined from time measurements in similar processes. The handling time $h_2$ is particularly high, due to the waiting time for the plateaus from the storage lift. The components of formula (3) are given by

$$\hat{c} = 8 \cdot \left( \begin{array}{l} 700 \\ 32 \\ 147 \end{array} \right)$$ and $$\hat{h} = \left( \begin{array}{l} 10 \\ 70 \\ 40 \end{array} \right) \text{s.}$$

The four different types of load carriers (roll cages, special pipe carriers, special load carriers and pallets) are used in the consolidation process. The base time per load carrier was determined by tests while the efficiency factor $x_e$ of 20% was taken from an analysis according to REFA [15] conducted in another warehouse. The components of formula (4) are given by

$$t_{base} = 8 \cdot \sum_{i=1}^{57} b_{wo} \quad \text{with} \quad b = \begin{pmatrix} b_{cage} \\ b_{long} \\ \vdots \\ b_{pallet} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} b_{cage} \\ b_{pipe} \\ b_{special} \\ b_{pallet} \end{pmatrix} = \begin{pmatrix} 30 \text{s.} \\ 50 \text{s.} \\ 70 \text{s.} \\ 50 \text{s.} \end{pmatrix}.$$  (7)

The consolidation time per day was calculated by inserting the components from the equations (5), (6) and (7) into (1). Figure 6 shows the time shares according to the components of the consolidation time. The resulting consolidation time is 60 h. Considering seven hours work shifts, 9 workers per day are required as the operational team for consolidation. The biggest share of the consolidation time consists in the handling time. The base time, at 2%, represents a comparatively small proportion. It is therefore expected that the greatest potential for optimization lies in a reduction of the handling time. Consequently, the application of modular load carriers according to Sliwinski [6] in order to reduce repacking efforts and work order optimized put away strategies for storage lifts according to Nicolas [16] could be identified as the main optimization measures.

### 6. Conclusion and outlook

The consolidation of work orders for large equipment manufacturers is a highly manual and labour-intensive process, which is characterized by high requirements on flexibility and responsiveness. Experiences from industrial projects revealed the demand for a mathematical model to calculate the consolidation time regarding the process-specific particularities and conditions of logistics planning projects at an early planning stage. An analytical model based on the calculation of transport, handling, base and allowance times was presented and subsequently applied during an industrial planning project for a large equipment manufacturer in order to evaluate different storage scenarios. The determination of the required planning
data was based on analyses of the warehouse and work order structures as well as time recordings of already existing commissioning and test systems. The applicability of the model was successfully proven. In order to validate the presented model, simulations of calculated commissioning process are currently being carried out. Subsequently, empirical data will be recorded once the warehouse is in operation. The model will be applied in future projects in order to further validate it.

![Figure 6: Shares of time components of the calculated consolidation time.](image)

**Appendix**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\vec{b}$</td>
<td>Vector containing base times required per work order 1 to m</td>
<td>-</td>
</tr>
<tr>
<td>$b_{wo}$</td>
<td>The base time required per work order</td>
<td>s</td>
</tr>
<tr>
<td>$\vec{c}$</td>
<td>Vector containing amounts of collection units to be merged from consolidation buffer 1 to n</td>
<td>-</td>
</tr>
<tr>
<td>$c_i$</td>
<td>Amount of collection units to be merged from each consolidation buffer per chosen time interval</td>
<td>-</td>
</tr>
<tr>
<td>$d_{ij}$</td>
<td>Distance from point $i$ to point $j$</td>
<td>m</td>
</tr>
<tr>
<td>$\vec{h}$</td>
<td>Vector containing handling times per collection unit from consolidation buffer 1 to n</td>
<td>-</td>
</tr>
<tr>
<td>$h_i$</td>
<td>Handling time required per collection unit</td>
<td>s</td>
</tr>
<tr>
<td>$i_{ij}$</td>
<td>Transports from point $i$ to point $j$ per chosen time interval</td>
<td>-</td>
</tr>
<tr>
<td>$m$</td>
<td>Amount of work orders per chosen time interval</td>
<td>-</td>
</tr>
<tr>
<td>$n$</td>
<td>Amount of consolidation buffers</td>
<td>-</td>
</tr>
<tr>
<td>$t_{allowance}$</td>
<td>Allowance time for work interruptions per chosen time interval</td>
<td>s</td>
</tr>
<tr>
<td>$t_{base}$</td>
<td>Base time for preparing the consolidation process for a chosen time interval</td>
<td>s</td>
</tr>
<tr>
<td>$t_{consol}$</td>
<td>Consolidation time for the consolidation process per chosen time interval</td>
<td>s</td>
</tr>
<tr>
<td>$t_{handling}$</td>
<td>Handling time for the merging of collection units per chosen time interval</td>
<td>s</td>
</tr>
<tr>
<td>$t_{transport}$</td>
<td>Transport time during the consolidation process per chosen time interval</td>
<td>s</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Average walking velocity of the consolidation employees</td>
<td>m/s</td>
</tr>
<tr>
<td>$x_e$</td>
<td>Performance factor to consider factors such as work interruption for distractions or disturbances</td>
<td>%</td>
</tr>
</tbody>
</table>
References


Biography

Jakob Schyga, M.Sc. (*1992) is a research associate at the Institute of Technical Logistics at Hamburg University of Technology. Jakob Schyga completed his studies in mechanical engineering with a focus on production at Hamburg University of Technology in 2018.

Philipp Braun, M.Sc. (*1992) is a research associate at the Institute of Technical Logistics at Hamburg University of Technology. Philipp Braun completed his studies in industrial engineering with a focus on information technologies and logistics at Hamburg University of Technology in 2020.

Svenja Hilbrich, M.Sc. in Physik (*1989) is a research associate at the Institute of Technical Logistics at Hamburg University of Technology. Svenja Hilbrich completed her studies of physics at the TU Dortmund University in 2015.

Johannes Hinckeldeyn, Dr. (*1979) is Chief Engineer at the Institute of Technical Logistics at Hamburg University of Technology followed his PhD in the UK, Johannes Hinckeldeyn served as Chief Operating Officer for a manufacturer of measurement and laboratory technology for battery research. Johannes Hinckeldeyn studied business engineering, production engineering and management in Hamburg and Münster.
Jochen Kreutzfeldt, Prof. Dr.-Ing. (*1959) is Professor and Head of the Institute for Technical Logistics at Hamburg University of Technology. After studying mechanical engineering with a focus on production engineering, Jochen Kreutzfeldt held various executive positions within a company specializing in automotive safety technology. Subsequently, Jochen Kreutzfeldt took over a professorship for logistics at the Hamburg University of Applied Sciences and became head of the Institute for Product and Production Management.
Implementation of Machine Learning to Improve the Decision-Making Process of End-of-Usage Products in the Circular Economy

Michael Diem1,2, Anja Braun1, Louis Louw2
1 Reutlingen University, Alteburgstrasse 150, 72762 Reutlingen, Germany
2 Stellenbosch University, Banghoek Road, 7600 Stellenbosch South Africa

Abstract

Rising consumption due to a growing world population and increasing prosperity, combined with a linear economic system have led to a sharp increase in garbage collection, general pollution of the environment and the threat of resource scarcity. At the same time, the perception of environmental protection becomes more sensitive as the consequences of neglecting sustainable business and eco-efficiency become more visible. The Circular Economy (CE) could reduce waste production and is able to decouple economic growth from resource consumption, but most of the products currently in use are not designed for the reuse-forms of the CE. In addition, the decision-making process of the End-of-Usage (EoU) products regarding the following steps has further weaknesses in terms of economic attractiveness for the participants, which leads to low return rates and thus the disposal is often the only alternative.

This paper proposes a model of the decision-making process, which uses machine learning. For this purpose, a Machine Learning (ML) classification is created, by applying the waterfall model. An artificial neural network (ANN) uses information about the model, use phase and the obvious symptoms of the product to predict the condition of individual components. The resulting information can be used in a downstream economic and ecological evaluation to assess the possible next steps. To test this process comprehensive training data is simulated to train the ANN. The decentralized implementation, cost savings and the possibility of an incentive system for the return of an end-of-usage product could lead to increased return rates. Since electronic devices in particular are attractive for the CE, laptops are the reference object of this work. However, the obtained findings are easily applicable to other electronic devices.

Keywords

Artificial Neural Network; Circular Economy; Classification; Decision-making process; Machine Learning; Remanufacturing; Reuse; Sustainability;

1. Introduction

Worldwide, one person generates on average 0.74 kilogram per day and the world generates 2.01 billion tonnes of municipal solid waste annually [1]. Furthermore, experts expect global waste to grow to 3.4 billion tonnes by 2050. Reasons for this enormous amount of waste are the increasing production and consumption of goods, caused by a growing population and increasing wealth as well as the low proportion of reused resources [2][3]. Consequently, the produced waste pollutes the environment and emissions lead to climate change. A way to reduce waste streams is reusing products or components and recycling of resources. The amount of resources needed to produce products increased significantly during the last years and is mainly
covered by newly extracted resources [4]. During the last century, the focus has been on the production of goods from virgin materials [5]. Major improvements have been made in increasing resource efficiency instead of systematically designing out material leakage and disposal [6]. However, the awareness of society is shifting to environmental problems and lack of resources [5]. Rising standard of living with a simultaneously growing population can only be ensured in the long term through sustainable economic growth [7]. An opportunity to recover and reuse resources in order to reduce waste and improve sustainable use of resources is the CE.

1.1 Circular Economy and decision-making process

The CE replaces the ‘end-of-life’ concept of a common linear supply chain (SC). It supports reuse and use of renewable energy and tries to eliminate waste through superior design of the products and business models [6]. To this end, the concept goes beyond resource efficiency and productivity approaches and positions itself between pure efficiency approaches and schools of thought of sufficiency and post growth economics. Figure 1 illustrates the structure of the CE.

In comparison to the linear economic system, there are more options than just disposal for an EoU product. It can be returned to the CE by different recovery options. An EoU product is a product that can be operative or inoperative. The owner can remove it or lead it to a decision-making process, which is represented by the black bar. It decides whether it is possible to lead the EoU product back to the SC and which is the preferred option. The inner circle is preferred of four main recovery options [6]. Recycling describes the least preferred, since it only recovers resources, which need energy to build a new product [9][10]. Remanufacturing is the process to disassemble the product to replace the defective components and worn out parts in order to achieve at least as good as new condition. The customer gets a guarantee on the entire product [11][12][13]. Repair is similar to remanufacturing, except only the defective parts are replaced and the customer only gets a guarantee on this component [14][15]. Reuse is the preferred option, since the product is only reconditioned. This requires the least energy and the most resources are reused [6][16].

In the decision-making process various factors are considered. These can be divided into economic, environmental and social factors [17][18]. The economic factors are subdivided into recovery value and process costs. In order to compare the value of the recovered product to the costs to enable the recovery, it is crucial to evaluate the current condition of the EoU product. The total costs are mainly determined by purchasing costs and processing costs of the EoU product and the costs for the spare parts.
1.2 Problem and objective

A major problem regarding CE are the low return rates of the EoU products [8]. This superior problem has further partial problems as cause, which are listed in Table 1. Low resource prices compared to high labour costs resulting from manual examination of the EoU product, make new production more attractive than recovery. Short lifetime cycles and many different products lead to a complex decision-making process with frequent updates, which are difficult to implement at the different collection points. Economic success continues to be the driving force behind decision and investment. The recovery options are associated with a high degree of uncertainty, which is why many products are not returned to the SC and no investments are made in the circular infrastructure.

Table 1: Problems of the decision-making process

<table>
<thead>
<tr>
<th>Problem</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low resource prices compared to labour costs and low automation level</td>
<td>[6][19]</td>
</tr>
<tr>
<td>Short lifetime cycles and many different products</td>
<td>[8][20][22]</td>
</tr>
<tr>
<td>Uncertainties of return</td>
<td>[3][14][22]</td>
</tr>
<tr>
<td>Complex decision-making process. Need of know-how with regard to disassembly</td>
<td>[11][14][22]</td>
</tr>
<tr>
<td>Missing preselection and infrastructure</td>
<td>[6][22][19][20]</td>
</tr>
<tr>
<td>Products are not designed for the CE</td>
<td>[9][24]</td>
</tr>
<tr>
<td>Missing attractiveness for owner of EoU product to lead it back to SC</td>
<td>[3][6]</td>
</tr>
</tbody>
</table>

In addition, the selection of the EoU products takes mainly place in central collection points and not upstream at the owner of the EoU product. This leads to unnecessary transports. Furthermore, most products currently in use are not designed for the CE. It is difficult to determine the condition and to refurbish these products. Most product owners choose the easiest way for themselves when a product reaches the end of its lifetime, which is in the most cases disposal. An incentive system and easier process could change this so that it is led to the decision-making process [23].

ML is already applied in various areas of CE. It is used for strategic decisions such as the product design, but also operational decisions, as estimating recovery costs. One example is the method developed by Goodall et. al. in 2015, which is a cost estimation for remanufacture with limited and uncertain information [20]. Jun et. al. have developed a method which determines the best possible recovery option by means of an evolutionary algorithm [25]. However, most methods assume that the condition of the EoU product is known, because the product is physically present at the decision-maker or using incomplete data. The focus is not on determining the condition of components of the EoU product in order to carry out a preselection.

The objective of this work is to develop a new decision-making process for the CE in order to increase the returning rates of laptops. The requirements of the new decision-making process are derived from the problems mentioned before. The decision-making process should be carried out as preselection at the customer in a decentralised form, but the expertise is to be developed and provided as a central system. In addition, an incentive system should be created for the EoU product owner to increase the motivation to return the EoU product. For this purpose, this paper develops a model to decentral determine the condition of a laptop. This can serve as a basis for an economic and ecological evaluation to select the right recovery option or for a cost estimation or selection method of best recovery option mentioned before. The model classifies selected components of the EoU laptop as defective or not defective by applying a ML based mixture of prognosis and diagnosis. Therefore, information is identified to determine the laptop condition and a possibility of capturing is defined to use them in the ML classification.
1.3 Methodology

The structure and methodology of this paper is based on the waterfall model, which serves software development. Figure 2 represents the steps to develop the laptop condition determination for the decision-making process. In contrast to the classic waterfall model, iterative interactions between various phases are possible here. After the problems to be dealt with and the objectives derived from them have already been worked out, this information is used to develop the decision-making process, which defines the system requirements. From these, the software requirements are derived in order to subsequently select a suitable one. The software selection is not covered in this paper. In the next step, the input variables and the output variables of the ANN are defined, as well as the structure and functionality. The information, which is defined for the determination of the condition of the component, forms the input. The output is the calculated condition of the component. The results are implemented in Microsoft Azure ML Studio. In order to finally test and evaluate the developed model, data is simulated beforehand.

2. Development of the decision-making process

The problems and objectives defined in chapter 1.2 are used to model the decision-making process illustrated in Figure 3. In order to make it possible to carry out the process by the owner without sending the laptop to a collection point, only information is used which is easy to capture. Therefore, it has to be gathered without expertise, diagnostic tools and disassembly. Information of the laptop model and information about the use phase and external appearance are combined in order to determine the condition of defined components of the laptop. There is only a difference made between defective and functional. However, components are also assigned to the class defective if they should be replaced, since they only have a short service life. The condition of the components can then be used to conduct the economic and ecological evaluation of the recovery options in order to select the best.
As part of this work, 10 components are selected, of which the condition will be determined. These are susceptible to wear or have a high value and thus have a high relevance for the evaluation of the laptop. The components are battery, CPU, GPU, display, keyboard, mousepad, motherboard, outer casing, hard drive and cooling system. First, the stress factors are determined, which have an impact on the condition of laptops. Figure 4 shows the stressing factors in the left table. Most stressing factors cannot be detected without installation of sensors and the use of software. The laptop should not be upgraded before usage, since the model should also work for laptops, which are not designed for the CE. Therefore, information is selected which is easy to gather at the moment of evaluation, but still provides information about the stressing factors during the use phase. Figure 4 shows the information, which will be used as input for the ANN, in the right table. It is not possible to record information concerning all stress factors. The focus is on usage intensity and temperature.

<table>
<thead>
<tr>
<th>Stress factors [27]</th>
<th>Defined information for determining condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>Brand</td>
</tr>
<tr>
<td>Humidity</td>
<td>Model</td>
</tr>
<tr>
<td>Dust and dirt</td>
<td>Price</td>
</tr>
<tr>
<td>Radiation</td>
<td>Age</td>
</tr>
<tr>
<td>Acceleration</td>
<td>Type of usage (categories)</td>
</tr>
<tr>
<td>Bending and pressure</td>
<td>Mobile usage</td>
</tr>
<tr>
<td>Current</td>
<td>Usage of laptop bag</td>
</tr>
<tr>
<td>Voltage</td>
<td>External damages of 6 external components (categories)</td>
</tr>
</tbody>
</table>

Figure 4: Stressing factors and defined information for determining conditions

The brand, model and price provides information about the quality of the components and the construction of the laptop. This has an influence on wear, fall damage and the ability to dissipate heat. Age is related to usage time and components installed. The type of usage can provide information about how intensively the laptop in general and certain components were used. Therefore, the categories surfing, business, computational intensive, gaming and videos are defined. A laptop in mobile use is exposed to more dust and higher risk of falls. Accumulated dust in the ventilation reduces the ability to dissipate heat and therefore causes higher temperatures. A case protects the laptop from dust and fall damage and can also be an indication of how carefully the laptop was handled. External damages to the laptop can also be easily captured by the owner. Therefore, these are recorded and assigned to one of six defined locations of the laptop. There are different damage classes for each of the six locations. Damages to different locations can have different influences on individual components of the laptop. The information described (see right table of Figure 4) is used as independent variables of the classification process and serves as input to the ANN. The 10 components are the dependent variables and form the output.

The ANN is chosen as the ML method as it is very powerful in classification applications. It can uncover many complex correlations in data sets and process different inputs. Moreover, the hardware costs incurred for the supply of computing power have fallen significantly in recent years. Since the application considered here is a mixture of prognosis and diagnosis, a feed forward ANN is applied.

### 3. Data simulation

In order to train and evaluate an ANN, comprehensive data is required. This is simulated in RStudio following the procedure shown in Figure 5. Sixteen exemplary models are defined and 250,000 objects are created from these. The model clearly defines the brand and price. In addition, each object is assigned an age by means of a normal distribution, which has a mean value of 36 months and a standard deviation of 8 months. Each model has different probabilities that the object will be assigned to one of five defined usage
types. In order to simulate the dichotomous variable of mobile use, the Monte Carlo approach is applied. The probability of mobile use of the laptop is calculated based on the model and the type of use. This is then compared to a random number (between 0 and 1) to determine the value. The calculation of the variable “use of case” is analogous to this. The value of the six variables, which describe the external condition, is then simulated for each object. This is done by only using probabilities for the error classes and is independent of the values of the other variables. Finally, the dependent variables are calculated for each object. For each of the 250,000 objects and for each dependent variable a probability is calculated that the regarding component is defective. This is done by means of logistic regression function, as it provides values between 0 and 1. Here the correlations between input and output are implemented by defining the weighting factors, which determine the influence of independent variables and characteristics of variables. For example, a cheap laptop ages faster than an expensive one or a gaming laptop stresses the CPU and GPU more intensive. By applying the Monte Carlo approach again, it is determined whether the component is has to be replaced or not.

![Figure 5: Simulation of independent variables](image)

### 4. Implementation of neural network

After the requirements for the neural network have been defined and the data simulated, the neural network is now implemented in Microsoft Azure ML Studio. Figure 6 illustrates the process. Since a separate ANN must be trained for each of the ten components, this process is conducted ten times.

![Figure 6: Process of creating the artificial neural network](image)

After importing the simulated data, the independent variables and the respective dependent variables are selected. It is determined which variables are used to determine the condition of which component. The 250,000 data sets are subsequently divided into training and test data. Literature recommends a value of 80% training data [28]. Furthermore, the desired Two-Class Neural Network is selected as ML method. In addition, the structure is defined. The hidden layer specification is set to fully connected and the number of hidden nodes to 100. This means that each of the input neurons (independent variables) is linked to all
neurons of the hidden layer, which are connected to the output neurons (defect or not defect). The next step is to train the created neural network with the training data. For this it calculates an output based on the input. This is then compared with the correct solution, which is why the method applied here is called supervised learning. By the determined error value, the configurations in the ANN are adapted in order to reduce the error value. This process is repeated for all data sets. In the next step, the trained ANN calculates the solutions for the test data. In the last step, the results obtained are compared with the actual solutions in order to evaluate the model using different metrics.

5. Results

Microsoft Azure ML Studio provides metrics, which evaluate the developed models. Figure 7 illustrates in the left hand side exemplary the receiver operating characteristics (ROC) curve of the battery model and in the right hand side metrics for all models.

![ROC Curve for battery](image)

The ROC is used to evaluate the performance of classification models in ML. It provides the area under the curve (AUC), which describes the ability of the model to separate the calculated values of the two output options. Figure 7 shows that the AUC is well above 0.90 for all models except the hard drive, which is a very good value. Furthermore, the accuracy is listed, which shows how many of the test data sets are determined correctly. The recall is the fraction of all defective data sets, which are labelled as defective. The precision is the proportion of true results of all as defective labelled data sets. A distinction of the errors is meaningful, since these have different consequences. For the system considered here, recall is the most important metric because the error that a defective component is considered as a functional one is the most serious error. This would lead to a laptop being purchased at a higher value than it actually has and would lead directly to a loss of money. The error that a functional battery is considered to be defective is not so bad. This would lead to the laptop being assigned a lower value than it actually has. Only possible turnover is lost as a result but not the own money. Therefore, the precision and accuracy is subordinated to recall in this case. The recall value can be increased by lowering the threshold value. However, this leads to a decrease of accuracy and precision. The threshold is set so that it does not fall below a value of 0.25, that not too many objects are considered as defective. It is also chosen as high as possible while ensuring a recall of 0.85, in order to purchase many EoU laptops. The recall values show that the defective objects of some components are detected very well. However, for those below 0.80, the model should be further improved by using more independent variables or implementing a manual check for objects of which the calculated value cannot be clearly assigned to defective or not defective.
6. Conclusion

The developed ML classification for determining the condition of EoU laptops can significantly improve the decision-making process in the CE, especially for products, which are not developed for the CE. The developed procedure, which collects information on stress factors via indirect relations to process it in an ANN, offers great potential for products that cannot be equipped with sensors. Its decentralized application saves costs for infrastructure and unnecessary transports. In addition, the return of EoU products becomes significantly more attractive, since the information obtained can be used to carry out an economic evaluation. This means that the owner can be offered a price for the EoU product. Since the condition of the product is determined at component level, the costs incurred for repair can be determined more precisely. Without the laptop being physically present, it is possible to determine to a certain extent which repairs are required. This increases the safety of the purchases of EoU products and thus also the number of recovered products. In addition, an ecological evaluation is made possible. The emissions generated during the production of a new laptop can be compared with the emissions generated during the production of the required spare parts for the recovery. In addition, the information gained offers further potential. Conspicuous features in the data sets can be used to improve product development. Furthermore, information regarding the number and models of returning products can be used for sales forecasts.

Because the classification is centrally hosted, it is much more flexible and cost-effective. New product generations or product types can easily be integrated and significantly more data is generated and collected. This is important, since the more complex the ANN gets, the more data is required to train it. Therefore, further possibilities must be investigated in order to implement more products and information in the process without increasing the complexity too much. This includes the generation of information categories, which would significantly reduce the number of input neurons. An example would be to only record the brand instead of model or summarizing models with a similar type of construction or price segment.

By applying ANNs, it is possible to uncover relations between the describing variables and the condition of the laptop, which one does not suspect. Therefore as much data as possible should be collected. The increasing number of sensors in devices in the progression of the Internet of Things will have a supporting effect. In future, further sensors and software should be implemented in the development phase of electronic devices in order to be able to better determine the condition of the product.

As the results show for certain evaluation values, considerations must be given to adapting the model or checking a component manually by first sending in the laptop before signing a sales contract, if the calculated value of the ANN is close to the defined threshold and so it could be defective or functional. Where possible, the existing temperature sensors and the already implemented components utilization (CPU and GPU) should be used. This requires a pre-installation of software and would make the model more complex, but could significantly improve the results of the classification. In addition, it must be noted that in this work the data are simulated on the basis of own knowledge and estimations. The reality is much more complex, which can lead to change in the results.

References


Biography

Michael Diem (*1994) studies Digital Industrial Management and Engineering at Reutlingen University and Stellenbosch University. This is a research master double degree program. Previously Michael Diem studied Industrial Engineering at University of Applied Science in Karlsruhe.

Anja Braun (*1985) has been a professor at Reutlingen University since 2017. Previously Prof. Dr.-Ing. Anja Braun spent five years at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA as a project manager. Her research interests are in the field of Circular Economy.

Louis Louw (*1974) has been an associate professor at Stellenbosch University since 2018. Previously Prof. Dr.-Ing. Louis Louw spent twelve years in industry as a business consultant and senior manager, before joining the university in 2014. His research interests are in the field of Digital Operations & Supply Chain Management, and he is a member of the International Association of Learning Factories.
1st Conference on Production Systems and Logistics

Deriving of Sequencing Strategies for Multi-Stage Productions Supported by Logistic Models and Software Tools

Alexander Mütze¹, Peter Nyhuis¹

¹ Institute of Production Systems and Logistics (IFA), Leibniz University Hannover, Germany

Abstract

Sequencing as a core task of production control has a significant influence on the logistical performance and efficiency of a single work system. Particularly in the presence of sequence dependent setup times, systematic sequencing can increase the productivity of a work system by saving them. This, however, leads to a decreasing schedule reliability of the work system, which creates an area of conflict. In recent years, mathematical models have been developed at the Institute of Production Systems and Logistics (IFA) that describe the influence of different sequencing rules on the schedule reliability and productivity of a work system. In a further step, these single so-called partial models can be linked with each other. This allows a calculation of the lateness behaviour of a multi-stage production in dependency of the sequencing rules assigned to the individual work system and thus of the overall sequencing strategy.

This paper presents the possibilities of linking different logistic models in order to quantify the influence of sequencing on logistic target values as well as two software tools by which the impact and combination of various sequencing rules can be examined based on production feedback data or by means of a generic supply chain. As a result, it is possible to assess different sequence configurations of a multi-stage production and thus strategically align the production in the area of conflict between productivity and schedule reliability.

Keywords

Production Planning and Control; Sequencing; Supply Chain Configuration; Software Tool; Demonstrator

1. Introduction

Manufacturing companies today are increasingly faced with the challenge of offering an ever-larger number of product variants and reacting to individual customer requirements while at the same time satisfying the demands for short delivery times and high due date compliance [1]. For companies in the manufacturing industry, the best possible fulfilment of these requirements at the lowest possible cost represents a main logistical challenge [2]. In order to meet these market requirements, high productivity in production is necessary, which can be achieved by eliminating processes that do not add value [3]. In addition to factors affecting the entire production process, such as unnecessary transports or multiple handling of products, scientific research and industrial optimization projects also focus non-value-adding activities on individual work systems. This includes in particular the handling and management of unavoidable machine setup processes [4–7].

In general, setup processes are required in case of switching between two orders. For example, a milling machine usually requires the tool table to be rebuild, while manual workstations may require fixtures to be
set up or tools to be changed. The setup time linked to the processes, during which no physical processing of work pieces takes place, can be subdivided into two components: a sequence dependent component, whose amount depends on which order was previously processed, and a sequence independent component, whose amount does not depend on the predecessor process. In case that parts of the setup time are sequence dependent and technologically (e.g. through externalization) or organizationally unavoidable, the setup time at a work system can be reduced by a setup time optimized sequence in order to achieve productivity gains [7,8]. This further means that products or orders that have similar setup requirements, e.g. require the same clamping system or a very similar tool set, can thus be assigned to a so-called setup family and have to be processed successively one after the other in order to decrease the total setup efforts at the work system [5,8].

However, the productivity increases due to a reduction of incurring setup times by setup-optimized sequencing are accompanied by a negative impact on the lateness behaviour of work systems, since such a prioritization of orders lead to an increase in the spread of the lateness in the output of the work systems [6,8,9]. As a result, the important logistic objective of single work systems as well as of the whole production schedule reliability is influenced directly and in a negative way [10].

In comparison to setup time optimizing sequencing rules, due date-oriented sequencing rules have a positive impact on the schedule reliability of a work system, as they reduce the spread of the lateness in the output compared to the spread in the input of the system [11,12]. Nevertheless, prioritization by due-dates results in no systematic sequencing in the sense of a setup time optimization takes place, because reductions in setup time only occur coincidentally and the intensity is thus significantly lower. In addition, it should be noted that sequencing and thus prioritising orders by due-dates provokes a more or less strong deviation from the incoming sequence of the orders, depending on the lateness in the input. Consequently, a more or less strong re-sorting of the orders is required, leading to rising searching and sorting efforts. As a result, a reduction in productivity may occur at the work systems, whose severity depends for instance on the ratio of necessary efforts to the average operation time.

The order sequencing according to a First-in-First-Out (FIFO) logic on the contrary does not lead to any change in the order sequence between the input and the output at a work system [13]. Nonetheless, sequencing according to FIFO leads also to randomly occurring setup time savings [5]. An actively intended impact on the productivity as well as on the lateness behaviour and thus on the schedule reliability of a work system is therefore not derivable.

Consequently, it can be stated that order sequencing, as part of the major task production control, takes place in a field of tension between the cost-oriented objective productivity and the performance-oriented objective schedule reliability [5,9]. Due to the high relevance of the adherence to delivery dates, as one of the most important indicators for assessing logistical performance [14], and the influence of productivity on economic efficiency a target-oriented positioning in this field of tension through the selection of suitable sequencing rules in the planning and control of production systems is of great importance.

From the perspective of the practical user, however, the question arises as to how the various sequencing rules and possible configurations influence the schedule reliability and productivity on an explicit work system and, in particular, how they influence further the entire production consisting of several work systems and stages. Only by the possibility of building-up a calculable chain of action, it is possible to configure a closed multi-stage production in such a way that based on the general preconditions, e.g. requirements for throughput and delivery times, and strategic specifications, the best possible positioning in the area of conflict between productivity and schedule reliability can be found.

At this point, this contribution applies and presents the linking logic between the single logistic models developed at IFA for the description of the impact of sequencing rules on the logistic objectives as well as two software tools with whose help an exact calculation or a generic and didactic discourse is made possible.
2. **Sequencing at work systems and its influence on logistic objectives**

The following section gives a brief overview of order sequencing at work systems and subsequently deals with the influence of different sequencing rules on the lateness behaviour and productivity of work systems. However, this chapter can only give a general overview about the individual models and their derivation. Therefore, reference is made to the authors and publications at the specific passages.

### 2.1 Order sequencing at work systems

Sequencing rules for the formation of order sequences at work systems determine the priority with which orders are to be processed [15]. Through this prioritization according to specific criteria, the processing sequence of the orders is determined. On the one hand, this can lead to deviations from the planned order sequence created by detailed production planning. On the other hand, existing deviations from the planned sequence caused by predecessor processes can be partially compensated [16]. According to various studies, the most frequently used sequencing rules in industrial practice are the First-in-First-Out (FIFO) rule, followed by schedule-oriented and setup time-optimizing sequencing rules. [17,18]

If FIFO is used, orders are processed in accordance to the actual sequence in the input of a work system. Consequently, the order with the highest waiting time in the waiting queue is given the highest priority. Schedule-oriented sequencing rules are considering the planned completion or planned dispatching dates of orders determined by detailed production planning when prioritizing orders. One of the most used and known schedule-oriented sequencing rule is the earliest operation due-date (EODD) rule. Out of all waiting orders, it assigns the highest priority to the order with the earliest planned operation due-date. [12,13,15]

In contrast, setup-optimizing (SO) rules aim to reduce setup times and setup costs. Consequently, this is only achievable if the setup times at a work system have a distinct sequence dependent component for the orders of the current order spectrum and thus a reduction of the total setup time at the work system is possible due to reprioritization of the orders. For a more detailed description of various rules for the creation of setup time optimizing order sequences, reference is made here to [4–6,8–10].

### 2.2 Influence of sequencing rules on the lateness behaviour of work systems

For sequencing in accordance to the FIFO, EODD as well as to the SO sequencing rule, BERTSCH developed logistic impact models [11,19]. These models enable a description of the lateness behaviour of work systems under consideration of structure-relevant variables and are based on two fundamental principles.

The first fundamental modelling principle is the differentiation between backlog-related and sequence related lateness at a work system. The backlog-related lateness, as the name suggests, is influenced in particular by the backlog of a work system compared to the production plan. The sequence related lateness, which is in the focus of this contribution, can be derived directly from the prioritization of work orders and the related sequence changes compared to the planned sequence. The exact derivation of these two components can be found in [11]. The second fundamental modelling basis is that the output lateness behaviour of a work system can be described by an output lateness distribution and by the distribution key figures mean output lateness and standard deviation of output lateness. Further investigations have shown that sequencing, if it is independent of work content and any productivity gains are compensated by balancing capacity and load, only leads to a change in the standard deviation but not to a shift of the mean value of output lateness. The reason for this is that any acceleration of an order as a result of a sequence reversal also leads to a corresponding deceleration of one or more other orders. [11,20]
In recent research activities at IFA, logistic models have been developed based on the findings of BERTSCH using Little's Law [22] to quantitatively describe the relationship between the work in process (WIP) at a work system and the resulting standard deviation of the sequence dependent output lateness in dependence on the sequencing rule used [5,9,11,21]. The mathematical partial models, valid in case of constant planned throughput times, for the sequencing rules FIFO, EODD and SO, verified by means of a parameter study, are shown in Figure 1 in the form of ISO-curves for the standard deviation of the sequence dependent input lateness of 1, 3 and 5 shop calendar days (SCD). Below the illustrated ISO-curves, the approximated calculation formulas for the calculation of the standard deviation of the sequence dependent output lateness of a work system are given for each sequencing rule. The WIP of a work system is modelled in accordance with LITTLE [22] by the mean virtual throughput time (TTP\textsubscript{vir}) as a parameter of time, which is calculated by dividing the mean work in process (WIPO\textsubscript{m}) measured in number of orders by the unweighted mean output rate (ROUTO\textsubscript{m}) measured in number of completed orders per SCD. Thus, it describes how long an order stays on average at a work system [19]. The standard deviation of the sequence dependent output lateness is an external or given influencing variable in the models and can be influenced, for example, by predecessor processes or the selected method and parameterization of order release, which influences the input sequence of orders at a work system [16].

The models confirm the assertions made in chapters 1 and 2.1. Thus, given constant planned throughput times, it can be recognised that the sequence dependent standard deviation of the lateness in the output of a work system is independent of the virtual throughput time in the of FIFO sequencing and depends exclusively on the sequence dependent standard deviation of the lateness in the input. The reason for this is that sequencing according to the FIFO logic does not result in sequence reversals and therefore the sequence in the output corresponds to the sequence in the input of a work system. However, if sequences are made in accordance to the EODD rule, sequence reversals between input and output can be observed in order to re-establish the planned sequence as far as possible resulting in a decreasing standard deviation of the sequence dependent lateness in the output. Hereby, it applies that the larger the virtual throughput time at a work system
system is the more deviations between the input and the originally planned sequence can be compensated. Furthermore, the model shows that the marginal utility of increasing the virtual throughput time decreases and the standard deviation of the sequence dependent lateness in the output asymptotically approaches the x-axis (see Figure 1). This result is explainable by the assumption of a normal distribution for the lateness distribution.

If the processing sequence at a work system is created using a setup-optimizing (SO) logic, the orders are placed in a random sequence from a scheduling point of view. The reason for this is that the distribution of setup time relevant characteristics like the affiliation to setup families is independent of the due-dates of the orders. As Figure 1 and the related formula show, converges the standard deviation of the sequence dependent lateness in the output asymptotically against the virtual throughput time with their rise. The prerequisite for the application of the partial model for setup-optimized sequencing is that the productivity gained, which will be discussed in the next chapter, is compensated to the extent that the virtual throughput time and thus the operating point of the work system remain fixed on the ISO-curve.

For a more detailed description of the influence of the described sequencing rules on the standard deviation of the sequence dependent lateness in the output of a work system as well as the detailed derivation of causal relationships between the discussed influencing variables, reference is made here to [5,9–11,21].

2.3 Influence of sequencing on productivity

The productivity gain on a work system that can be achieved by the creation of setup-optimized order sequences are describable according to various research activities of NYHUIS and MAYER [5,10]. A basic assumption of NYHUIS and MAYER is that the order spectrum on a work system can be divided into setup families. Within such a setup family, an order change at a work system does not cause any setup time. However, the setup time for changing between orders of two different setup classes depends on the predecessor/successor relationship of these setup families, which is often not symmetrical. For the modelling of setup-optimized sequencing, NYHUIS and MAYER are using a rule that chooses the next setup family according to the Minimum Marginal Setup Time [5,23]. According to this rule, a new family is chosen by the lowest quotient of setup time and the number of jobs belonging to that family waiting in the waiting queue at this moment, if the current setup family is empty.

In general, the productivity gain of a work system achievable through setup-optimized sequencing can result from two effects illustrated in Figure 2. On the one hand, there may be a setup time saving resulting from the fact that an order is produced directly after an order of the same setup family (Figure 2 case "II"). This saves the setup time of the order completely and the processing can take place directly after completion of the previous order. Another possibility is a setup time reduction when switching between the different setup families, which results from a sophisticated, setup time reducing sequence creation due to the sequence dependency of the setup times (Figure 2 case "I"). As an example, it can be seen in the figure that the sequence formation C-B-A-(C) results in a total setup time of 57 units per cycle, while another order sequence leads to significantly longer setup times. It follows that sequencing in the order A-B-C-(A) would result in a setup time of 75 units per cycle.

In general, it applies that, given the same virtual throughput time at the work system is considered, the effect of setup time savings dominates with few setup families, while with many setup families the setup time reduction is dominant. This is explainable by the constant length of the queue in front of the work system and the accompanying probability of having more than one order of a setup family in the waiting queue [5].

At this point, reference is made to [5,10] for the detailed derivation of the cause-effect relationships and the calculation formulas which would go beyond the scope of this paper due to their complexity and the need for explanation.
In summary, it can be concluded that productivity gains can be achieved with setup-optimized sequencing, while these are accompanied by an increase in the standard deviation of the sequence dependent lateness in the output. Thus, there is a conflict of objectives between productivity and schedule reliability and there is a need for the strategic positioning of a company in this area of tension [5,9,10].

3. The logic for linking the described models to a chain of action

The partial models presented in chapter 2 can be linked with each other via different coupling sizes in such a way that it is possible to calculate the impact of sequencing strategies on schedule reliability and productivity for a multi-stage production and thus for a production area. Therefore, the developed partial models for the calculation of the impact of sequencing rules on the lateness behaviour of work systems can be linked via the coupling size standard deviation of sequence dependent lateness. The models developed by NYHUIS and MAYER to calculate possible productivity gains can be linked via the virtual throughput time with the partial model to describe the impact of a setup-optimized sequencing on the lateness behaviour of a work system.

Figure 3 shows as an example a production area consisting of three linearly linked work systems. The logic for linking the different models is explained using this example. The production area, and thus the first work system, receives the orders with a standard deviation of sequence dependent lateness of 3 SCDs. The first work system processes the incoming orders according to the FIFO logic, so that the orders leave the first work system also with a standard deviation of sequence dependent lateness of 3 SCD and reach the second work system. A setup-optimized sequencing is selected for the second work system, as the aim is to increase the productivity of the work system. Reasons for this could be, for example, that additional shifts on this system should be reduced in the future. On the basis of the virtual throughput time, which has an average amount of 5 SCD on this work system at the time under consideration, it follows that the orders leave the work system with a distribution with a standard deviation of sequence dependent lateness of approx. 6 SCD and reach the last work system of the production area. In return, a productivity gain of 20% can be expected. At the last work system of the production area EODD is chosen as sequencing rule in order to reduce the standard deviation of sequence dependent lateness in the output of the work system and thus of the entire production area. It follows, given a virtual throughput time of 3 SCD, that the orders leave the production area with a lateness distribution according to a standard deviation of sequence dependent lateness of 3 SCD.
As the example shows, the impact of the sequencing rules on the production area can be calculated by linking the models by means of the coupling variables standard deviation of sequence dependent lateness and virtual throughput time. The models can, however, be used further to calculate various different configuration-cases and help to develop a sequencing strategy for a production area and thus a multi-stage production. For example, the spread of lateness in the output of the production area could be further reduced by an increase of the virtual throughput time through planning measures at work system 3. In addition, the productivity gain at work system 2 could be further increased with the consequences for the standard deviation of sequence dependent lateness in the output of the system already discussed. To ensure a constant total virtual throughput time, a reduction at work system 1 equivalent to the possible increases could be carried out objective neutral.

To enable strategic positioning based on actual production feedback data and to support users in designing their production, two intuitive software tools for use in practice were developed at IFA providing the findings and research results described above. Both tools are presented in the following.

### 4. Software tools for analysing & describing the influence of sequencing rules in production

The developed software tools were both programmed in Microsoft Excel® using the programming language Visual Basic for Applications (VBA). This ensured that the tools can be used in as many companies as possible and that no software licenses have to be purchased to use the tools. In the following, the presentation of both tools is limited to their general functionality, the adjustable parameters as well as the necessary input variables. The first developed tool gives the user the possibility to carry out a what-if analysis based on production feedback data for a real multi-stage production and to determine the impact of sequencing strategies on a production area. In this tool, the main emphasis is on the lateness behaviour of the single work systems as well as on the whole production. The second programmed tool is comparable to a demonstrator and contains a generic reference supply chain enabling studies of the influence of sequencing strategies on a multi-stage production in an easy way. It also includes the possibility to calculate the possible productivity gain through setup-optimized sequencing according to NYHUIS and MAYER [5,10].
4.1 Software tool for analysis on the basis of production feedback data

The developed software tool enables the user to calculate the impact of individual sequencing strategies based on production feedback data for a production area of up to 80 work systems. As input data, the tool requires, on the one hand, the production feedback data for the actual input and output at the work systems, sorted according to order and process number and assigned to a work system of the area, as well as the corresponding planned values and the work content of the orders at every work system. Furthermore, a capacity table can be inserted, that is used by the tool to determine all SCDs per work system in the time period provided by the inserted production feedback data.

Using the entered input variables, the tool automatically calculates the predecessor/successor relationships and thus the material flow relationships between the individual work systems and establishes an impact chain between the work systems. Here, a specially developed algorithm is applied which eliminates circular references between work systems and thus enables a clear direction of action from the entrance to the exit of the production area, so that the calculation of the impact of sequencing rules is made possible.

In the output, the tool can display distribution diagrams and key figures for each work system related to the theory of the presented partial models for the examination of the impact of sequencing rules on the sequence dependent lateness behaviour. In addition, the most frequent successors of a work system are determined. Based on the presented values and diagrams and the made calculations in the backend, the user can change the sequencing rule (FIFO, EODD or SO) for every work system, resulting in an automated presentation of the changed distribution and key figures by the tool considering the inserted production feedback data. Furthermore, the user can set any sequencing rule in order to define a sequencing strategy, which is used to recalculate the distribution and key figures in the input of every work system and the entire production area. This allows the user to set up a sequencing strategy for his production area. For comparison purposes, the values for the case of a "pure" sequencing strategy (all work systems create the order sequence according to the same logic) are also compared at the observation level of the entire production.

To use the software tool, it is necessary to provide well-prepared production feedback data. In order to enable the practical user to derive initial findings for the configuration of sequencing in his own production, even without the available or sufficiently consistent feedback data, the second software tool was developed.

4.2 Software tool based on a generic supply chain for instant use

The second software tool, as mentioned above, comprises a generic reference supply chain consisting of one input, six work systems and a buffer towards the output (see Figure 4). In the process of creating the software tool care was especially taken while designing the reference supply chain to cover as many order flow relationships as possible but also providing an easy and understandable application. For this reason, one spreading point (behind work system 1) and one merging point (in front of work system 6) are established.

The user can set the individual parameters in the software tool by simply entering numerical values or using logical connected buttons. The configuration options range from the “external” standard deviation of the sequence dependent lateness in the input, the configuration of the individual work systems (virtual throughput time, WIP, output rate, sequencing rule) to the proportion of the order flow between the upper (work system 2 & 4) and the lower branch (work system 3 & 5). In addition, the expected productivity gain according to NYHUIS and MAYER under variable configurations can be calculated. Therefore the number of setup families as well as the distribution form of the setup time (Erlang distribution, equal distribution, and others), which have a direct impact on the possible productivity gain, are included as configurable values.

An especially noteworthy element is the buffer, which has been integrated as an additional supply chain element and makes it possible to integrate a safety time into the consideration of the lateness behaviour of a multi-stage production. This allows to auxiliary integrate the adherence to delivery dates into considerations.
5. Conclusion

In this paper, the theoretical foundations for describing the impact of sequencing rules in relation to the virtual throughput time on the lateness behaviour and the productivity for single work systems were discussed and the possibilities for linking the models in an easy way via coupling sizes were presented. Thereby it is made possible to calculate the influence of sequencing rules on the logistic objectives schedule reliability and productivity in a multi-stage production and to derive an individual elaborated sequencing strategy.

The considerations made in this paper always assume that any productivity gains are compensated by measures of production planning and control in the interest of having a stable system (stock and throughput time). The same also applies to productivity losses due to, for example, sorting efforts in case of sequencing in accordance to the EODD logic, which are not included in the models and were therefore not discussed further in this contribution, but should not be neglected in practice.

In order to support the practical user in deriving a sequencing strategy for his production area, two lightweight software tools with different approaches have been developed at the Institute of Production Systems and Logistics, which can be downloaded from the website http://go.lu-h.de/sequencing.

Acknowledgements

The contribution was created within the knowledge transfer project "Logistics-Oriented Sequencing strategies for multi-stage production", funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 409759709. In addition, the authors would like to thank the Koenig & Bauer Industrial AG & Co. KG as transfer partner for the valuable input and contribution to the project.

For the correctness of the calculations in the software tools, no liability is taken over by the authors nor the Institute of Production Systems and Logistics belonging to Leibniz University Hannover.

Figure 4: Schematic representation of the elements and configuration possibilities within the software tool
References


Biography

**Alexander Mütze, M.Sc. (*1994)** studied industrial engineering at Leibniz University Hannover and has been working as a research assistant at the Institute of Production Systems and Logistics (IFA) since 2018 in the field of production management with a focus on production planning and control.

**Prof. Dr.-Ing. habil. Peter Nyhuis (*1957)** studied mechanical engineering at the Leibniz University Hanover and subsequently worked as a research associate at the Institute of Production Systems and Logistics (IFA). After obtaining his Dr.-Ing. doctorate, he was habilitated before working as an executive in the field of supply chain management in the electronics and mechanical engineering industry. He has been head of IFA since 2003.
1st Conference on Production Systems and Logistics

Current Knowledge Management in Manual Assembly – Further Development by the Analytical Hierarchy Process, Incentive and Cognitive Assistance Systems

Robin Sochor¹, Teresa Sofie Schick², Lukas Merkel¹, Stefan Braunreuther¹, Gunther Reinhart¹

¹ Fraunhofer Research Institution for Casting, Composite and Processing Technology IGCV, Provinstr. 52, D-86153 Augsburg, Germany
² Technische Universität München, Arcisstraße 21, D-80333 München, Germany

Abstract

The complexity of manual assembly is continuously increasing due to a large variety of products, multi-product assembly or a batch size of one. To stay ahead in competency and competition, and to ensure adaptability and flexibility in today’s dynamic production environment, awareness of knowledge as the 4th factor of production, as well as the effective management of knowledge, are crucial. The present research therefore aimed at further advancing knowledge management in manual assembly by (1) assessing cognitive assistance systems and organisational incentive systems by use of an online survey distributed to German production companies, and by (2) applying the Analytical Hierarchy Process (AHP) as a transparent decision-making tool for knowledge-based improvements in the manual assembly process and workplace design. By employing an exemplary case of two feasible assembly alternatives, the AHP was applied as a method of knowledge measurement in a specific use case revealing priorities for knowledge-based ideas. To properly compute a final priority ranking of workers’ knowledge ideas, an algorithm written in Python programming language in accordance with the problem-solving framework previously published by Thomas L. Saaty (Decision Sciences, 18: 157-177, 1987). The performance of the algorithm shows that the rating process can be standardised and automated to a high level, and that the AHP may thus provide supportive evidence for assembly optimisation. The AHP-derived results can be used as a suitable basis for a bonus-point incentive system, which should contain both material and immaterial incentives. To operationalise this, it is therefore recommended to integrate the AHP rating process into a knowledge management application of hand-held devices, such as tablets, which are widely used in the production environment.

Keywords

Knowledge Management; Manual Assembly; Analytical Hierarchy Process; Incentive Systems; Cognitive Assistance Systems

1. Introduction

“Progress depends on the exchange of knowledge” [1]. This quote by the physicist, Albert Einstein, may well be applied to the present dynamic production environment driven by ongoing digitalisation and challenged by a demographical shift within the workforce. Multiple changeovers and adjustments during the individual steps of a given production cycle, smaller batch sizes and a growing variety of products significantly contribute to increased complexity in manual assembly [2, 3, 8]. Manual assembly operations
put a high demand on assembly workers with the speed, precision and quality of their work being essential requirements [3]. The optimal fulfilment of these requirements is strongly related to the workers’ comprehensive cognitive abilities, distinct sense of touch and individual experience and capabilities. Thus, experienced workers fulfil the role of knowledge-carriers and problem-solvers in production companies [4]. 

On the basis of their detailed knowledge concerning various assembly processes and their implementation, workers might be able to suggest alternative assembly sequences for products, different from those determined by the work instructions, thereby potentially increasing efficiency. To systematically benefit from workers’ prevailing implicit knowledge, a platform is needed [6] by which workers may share their knowledge and participate in decision-making processes with regard to product assembly, workplace design or product development [5]. Whilst it is widely accepted that an adequate knowledge management system does represent such a suitable platform [11], the issue of ‘how to assess knowledge’ remains an unsolved problem of knowledge management, as has been consistently documented by previous research [11, 12, 13]. Up to now, available tools for the assessment of knowledge have predominantly been of a quantitative nature [12]. These tools, however, are considered rather inadequate, or even counterproductive, for the purpose and, therefore, qualitative aspects also need to be included into an assessment process [11]. 

The present research was therefore designed (1) to determine relevant aspects of knowledge management in manual assembly by a web-based survey and a literature review, (2) to examine the suitability of the Analytical Hierarchy Process as an assessment tool for workers’ knowledge-based ideas towards the advancement of manual assembly, and (3) to propose a knowledge management system employing both incentive systems and cognitive assistance systems.

2. Web-Based Survey – Results and Theoretical Framework

300 German production companies from various industrial sectors, using manual assembly, were approached for participation. Information on knowledge management and knowledge types, knowledge transfer, incentive systems, and cognitive assistance systems in manual assembly was obtained via a web-based survey from participating firms from 1st May 2019 to 30th June 2019. Out of the 300 companies, a total of 11 completed all sections of the survey, resulting in a response rate of 4%. This relatively low response rate should be kept in mind when interpreting the results which are expressed as percentages in Table 1.

<table>
<thead>
<tr>
<th>Industrial Sector</th>
<th>Department of Respondents</th>
<th>Annual Turnover in Million Euros</th>
<th>Number of Employees worldwide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mechanical Engineering</td>
<td>73% Assembly</td>
<td>&lt;100</td>
<td>46% &lt;100 18%</td>
</tr>
<tr>
<td>Automotive</td>
<td>45% Management</td>
<td>&gt;100</td>
<td>27% &gt;100 37%</td>
</tr>
<tr>
<td>Electronics</td>
<td>18% Sales</td>
<td>&gt;10,000</td>
<td>18% &gt;1,000 18%</td>
</tr>
<tr>
<td>Assembly Technology</td>
<td>18% HR</td>
<td>Not available</td>
<td>9% &gt;10,000 27%</td>
</tr>
<tr>
<td>Medical Technology</td>
<td>18% Production</td>
<td>9%</td>
<td></td>
</tr>
</tbody>
</table>

2.1 Opportunities of knowledge management in manual assembly

All survey participants reported a high to very high estimate of their companies’ expected opportunities secondary to the implementation of knowledge management in manual assembly. Benefit was particularly expected towards product quality and reduced error rate (100%), but also increased efficiency and customer satisfaction (91%), time expenditure (82%) and cost-savings (73%). Some companies also considered knowledge enhancement (45%) and workers’ satisfaction (36%) as benefits of knowledge management.
2.2 Knowledge categories in manual assembly

To rate the importance of different categories of workers’ knowledge relevant for continuous improvements in manual assembly, respondents of the survey were able to choose between five options, ranking from most important to least important. The relevant categories of knowledge employed were those previously defined by Berger et al. [4]. Our survey revealed that the most important categories of assembly workers’ knowledge are knowledge of production means (91%), joining technologies (82%), product characteristics (82%), assembly procedures (73%) and of problem-solving in engineering processes (55%). Moreover, detailed knowledge of materials (45%), technologies (36%), product use (27%) and suppliers (9%) was also considered very important by some respondents.

2.3 Knowledge transfer

Knowledge transfer is not an end in itself. Especially in today’s dynamic production environment where agility is a key for success, knowledge transfer across hierarchies, departments, business entities, organisations and generations of employees plays a decisive role [17]. However, in practice, knowledge transfer in manual assembly is often hindered [12], which is why this issue was also addressed in our survey. 73% of the responding companies reported that knowledge transfer in manual assembly is compromised because there is not enough time to share knowledge. Furthermore, the workers’ lack of motivation to share knowledge, as well as the (lack of) a clear channel of information and communication, were considered as transfer barriers by 45%. The unwillingness to use somebody else’s knowledge (27%), lack of trust (9%), but also the workers’ insufficient language skills (9%) may be relevant barriers, as reported by some of the participating companies. To enhance the exchange of knowledge in manual assembly, the majority of companies use process modelling (73%) and knowledge maps (55%).

2.4 Knowledge management systems

A knowledge management system can be defined as a dynamic application system that contains tools to identify knowledge, to assist in the active search and transfer of knowledge, and to advance the distribution and usage of existing and development of new knowledge [16]. For our survey, items of existing knowledge management systems were used as previously published [1,16]. Our results show that content-orientated systems, such as document and content management systems or learning systems [16] are the most frequently (73%) used systems by the survey participants for knowledge management in manual assembly. Authoring systems, a sub-group of classified-as-content-orientated systems, are currently applied by 36% of the responding companies, which is in line with previous research suggesting suitability and the great potential of authoring systems for knowledge management in manual assembly [5]. Groupware systems, such as communication, collaboration and co-ordination systems, are currently used by 45% of the responding companies in our survey, while executive information systems, such as data warehouse or data mining systems, are used by 36%, and systems of artificial intelligence, e.g. text-mining systems or expert and agent systems, are not yet widely used with only a 9% reporting rate.

2.5 Incentive systems

Incentives, material or immaterial, have the potential to activate individuals’ motivation and satisfy their needs. Thus, the implementation of incentive systems aims at motivating workers to share their knowledge within their company [6]. The present study revealed that the impact of material incentives on the motivation of workers to share their knowledge is strong for 30%, medium for 45%, and weak for 18% of the companies. Immaterial incentives, however, such as the personal recognition of workers (82%), corporate culture (73%), promotion opportunities (64%), work content and workplace design (55%), safety (45%) and playful incentives (36%) have a strong to very strong impact on employees’ motivation to share knowledge. A variety of requirements have previously been identified as relevant for the design of incentive systems [7],
and were employed as survey items in our study. The responding companies reported transparency as the single most relevant requirement (100%). This is in accord with previously reported findings by Bullinger et al., who similarly reported transparency to be the most important design feature of an incentive system based on a survey in both producing and non-producing companies (87.2%) [7]. In our study, the following requirements were also considered of high relevance: motivational stimulation (82%), economic efficiency (73%), flexibility (73%), performance orientation (73%), rewarding effects (64%), individual and group applicability (64%) and fairness (55%). The incentive systems currently in use by the majority of the production companies responding to our survey are point-based incentive systems (55%), balanced scorecard-based systems (27%), gamification (9%) and management by knowledge objective (9%).

2.6 Cognitive Assistance Systems in Manual Assembly

In view of the complexity and product variety in manual assembly [3, 8], an exchange of knowledge will undoubtedly benefit from support by cognitive assistance systems. 64% of the production companies in manual assembly in our study currently use cognitive assistance systems in manual assembly while 18% use none. Cognitive assistance systems, such as mobile devices (wearables, handhelds) or immersive technologies (Augmented Reality, Virtual Reality), can support assembly workers by providing, processing or collecting information [8]. Moreover, they might motivate workers to share their knowledge and support the externalisation of knowledge, from tacit knowledge (i.e. experience), to explicit knowledge (i.e. information) [9]. In our survey, the majority of companies use tablets as cognitive assistance systems (62%), while wearables, such as data glasses (26%) or smart watches (12%), are used to a lesser degree. The major area of application of such cognitive assistance systems is the one related to training of workers (86%). The higher acceptance of mobile devices, such as smartphones or tablets, as compared to smart watches or data glasses, has previously been reported [10, 3] and may be attributed to the widespread familiarity of these devices in the consumer sector [10]. On the basis of our survey, and in line with findings recently published, we propose handheld mobile devices as the most suitable cognitive assistance tool for purposes of knowledge management.

3. The Analytic Hierarchy Process (AHP)

The assessment of workers’ knowledge-based ideas presents a major challenge in the field of knowledge management, as no validated decision-making tools have been agreed upon so far [11, 13]. In this paper, we suggest the Analytic Hierarchy Process (AHP) as a suitable tool for such decision-making.

The AHP, first described by Thomas L. Saaty, represents a mathematical method designed to assist in multi-criteria decision-making [15]. In this context, decision-making has the objective of finding the best alternative or set of alternatives by considering a number of criteria. Based on experts’ judgements, a pairwise comparison of multiple criteria with the alternatives under consideration is mathematically performed and yields a ranking and value of alternatives which by the decision-maker is expected to be correct [14, 15]. The AHP method is particularly advantageous when both quantitative and qualitative criteria have to be considered simultaneously [18].

Up to now, the AHP method has been employed in quite a number of decision-making situations. Thus, the AHP has been proven useful to assess the risk of losing knowledge [13], to evaluate knowledge management tools [19] and to support innovation processes in organizations [20]. The AHP has also been applied in the field of operations management (e.g. choice of technology, product design or plant layout) [21], in the evaluation or selection of new product ideas [22, 23] or as one of the most commonly multi-criteria decision-making methods in the area of transportation research [18].

To our knowledge, however, in the field of manual assembly the AHP has not yet been tested to assess the superiority of workers’ knowledge-based ideas versus established assembly procedures. Therefore, we
attempted to determine the suitability of the AHP as an assessment tool in an exemplary decision-making case of two manual assembly alternatives.

For this purpose, an algorithm, following the previously published computation rules by Saaty [14, 15], was written in the Python programming language which allowed for an automated and standardised mathematical prioritisation of a particular worker’s knowledge-based idea.

The AHP is carried out by the following fundamental steps according to Saaty T.L [15] and Saaty R.W. [14]: (1) A specific decision-making problem has to be structured into a multi-level hierarchy, descending from an overall objective at the top-level, to criteria at the intermediate level, and to alternatives at the lowest level. (2) Subsequently, comparative judgements in terms of pairwise comparisons between defined criteria are conducted. (3) Based on the resulting judgement matrix, local priorities can be derived by solving for the principal eigenvector of the matrix and normalising the result. (4) To determine the inconsistency of each matrix of pairwise comparisons, and also for the entire hierarchy, a consistency ratio (CR) has to be computed. Inconsistency is tolerated up to a level of 10%.

### 3.1 Application of the AHP in an exemplary decision-making case of a knowledge-based idea

#### 3.1.1 Case description

For the manual assembly of a specific component, an assembly worker has to grab five springs individually. Based on his/her work experience, the worker is convinced that this assembly step could be optimised if he/she was allowed to grab five springs in one single movement. He/she therefore shares his/her knowledge-based idea with the company.

#### 3.1.2 AHP Decision-Making Hierarchy

A company’s overall objective is to continuously improve manual assembly processes, which may vary from company to company and therefore need to be specifically defined. In the exemplary decision-making case (3.1.1), a judgement has to be made which one of two alternatives should be preserved or implemented, respectively (Figure 1). For this purpose, seven criteria were considered relevant: (1) Time Advantage, (2) Cost Advantage, (3) Quality Improvement, (4) Employee Satisfaction, (5) Technical Feasibility, (6) Ergonomic Design, (7) Reduction of Waste (Figure 1).

![Figure 1](image)

Figure 1: AHP decision-making hierarchy for the proposed exemplary manual assembly case.
3.1.3 Pairwise comparisons of assessment criteria

Based on the fundamental scale by Saaty [14, 18] ranging from 1 to 9, with 1 representing the equal importance of two criteria, and 9 the highest superiority of one criterion over the other, all seven criteria were weighted pairwise against each other with respect to the overall focus, the continuous improvement of manual assembly (Table 2). Based on work experience and professional judgement, decision-makers have to ask themselves the question: “How much more is quality improvement preferred over time advantage?” In our exemplary case, quality improvement in manual assembly is considered slightly more important than time advantage. Thus, the fundamental scale value of 3 is entered in the (3.1) position. The reciprocal value 1/3 is automatically entered for the transpose position (1.3).

Table 2: Pairwise comparisons of defined assembly assessment criteria.

<table>
<thead>
<tr>
<th>Defined Criteria</th>
<th>TA</th>
<th>CA</th>
<th>QI</th>
<th>ES</th>
<th>TF</th>
<th>ED</th>
<th>RW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Advantage (TA)</td>
<td>1</td>
<td>0.33</td>
<td>0.33</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Cost Advantage (CA)</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Quality Improvement (QI)</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Employee Satisfaction (ES)</td>
<td>0.25</td>
<td>0.33</td>
<td>0.33</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Technical Feasibility (TF)</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Ergonomic Design (ED)</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.33</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Reduction of Waste (RW)</td>
<td>0.14</td>
<td>0.2</td>
<td>0.14</td>
<td>0.33</td>
<td>0.5</td>
<td>0.2</td>
<td>1</td>
</tr>
</tbody>
</table>

3.1.4 Local priorities (or weights) of assessment criteria and alternatives

Entry of weighted assessment criteria into the Python-written Saaty algorithm [14] of the AHP yielded the local priorities of each of the criteria. The seven local priorities constitute the principal eigenvector of all seven criteria together (Table 3, Column 2). In addition, matrices of paired comparisons were set up for the two assembly alternatives compared with respect to each of the seven criteria. There were seven matrices in total. The local priorities for the alternatives (Table 3, Column 3) were also computed by the Python written AHP-algorithm, yielding the principal eigenvector.

Table 3: Local priorities of criteria and alternatives in the proposed exemplary decision-making case.

<table>
<thead>
<tr>
<th>Defined Criteria</th>
<th>Local Priorities of Criteria</th>
<th>Local Priorities of Alternatives (Alternative 1, Alternative 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Advantage</td>
<td>0.27</td>
<td>(0.100, 0.900)</td>
</tr>
<tr>
<td>Cost Advantage</td>
<td>0.26</td>
<td>(0.125, 0.875)</td>
</tr>
<tr>
<td>Quality Improvement</td>
<td>0.19</td>
<td>(0.833, 0.166)</td>
</tr>
<tr>
<td>Employee Satisfaction</td>
<td>0.11</td>
<td>(0.166, 0.833)</td>
</tr>
<tr>
<td>Technical Feasibility</td>
<td>0.08</td>
<td>(0.875, 0.125)</td>
</tr>
<tr>
<td>Ergonomic Design</td>
<td>0.04</td>
<td>(0.166, 0.833)</td>
</tr>
<tr>
<td>Reduction of Waste</td>
<td>0.02</td>
<td>(0.875, 0.125)</td>
</tr>
</tbody>
</table>

The consistency ratio (CR) calculation was also performed by the Python algorithm. The computed CR of the criteria matrix (Table 2) is 0.097 (i.e. 9.7%) and therefore below the acceptable inconsistency tolerance level of 10%. Thus, inconsistency has no adverse effect on the overall result. However, if a better CR was required, the pairwise comparisons would have to be redone with alternative ratings. Professional software packages, such as Expert [14] and others, will directly display inconsistency assessments in terms of critical judgement input [14]. A consistency check for the alternative matrices is unnecessary, as inconsistency cannot arise by comparing only two elements.
3.1.5 Global priorities (weights) of assembly alternatives

To finally derive an arithmetic basis for the ultimate decision on the best alternative from the computational results, the local priorities for the two assembly alternatives (Alt. 1, Alt. 2) are combined to an overall 2x7 matrix and are multiplied with the principal eigenvector, the 7x1 matrix of the seven predefined criteria, as follows:

\[
\begin{pmatrix}
0.100 & 0.125 & 0.833 & 0.166 & 0.875 & 0.166 & 0.875 \\
0.900 & 0.875 & 0.166 & 0.833 & 0.125 & 0.833 & 0.125 \\
\end{pmatrix}
\times
\begin{pmatrix}
0.27 \\
0.26 \\
0.19 \\
0.11 \\
0.08 \\
0.04 \\
0.02 \\
\end{pmatrix}
= 
\begin{pmatrix}
0.343 \\
0.657 \\
\end{pmatrix}
\]

Figure 2: Matrix multiplication of both alternatives (Alt. 1, Alt. 2) and criteria ratings (A) yields the global priorities which are automatically displayed as a bar graph (B) by the Python-written AHP algorithm.

In our exemplary decision-making case, alternative 2 received a global priority of 0.657 and was, therefore, superior to alternative 1, which had a global priority of only 0.343 (Figure 2, A). The global priorities of both alternatives are automatically generated by the Python-programmed algorithm and displayed in a bar graph (Figure 2, B). Before a final decision with regard to the implementation of alternative 2 as the future assembly process is made, a sensitivity analysis should be considered. Hereby, the criteria weights are systematically altered to determine the inflection point where the ranking of the alternatives is reversed, i.e. alternative 1 would then display a higher global priority than alternative 2. In the case of strategic decisions in manual assembly, a sensitivity analysis is generally recommended to test whether the calculated priorities are stable [18].

4. Proposition of a Knowledge Management System Employing both Incentive Systems and Cognitive Assistance Systems

The present study revealed that mobile assistance systems, i.e. tablets, are most widely used as cognitive assistance systems in manual assembly (62%). Since tablets provide user-friendly display options [10], and may be flexibly used at multiple work stations [2], we recommend tablets as cognitive assistance systems for applications in knowledge management. For this purpose, the development of a knowledge management application software (app) is recommended which should (1) enable workers to share their knowledge via text messages, photos or videos, (2) enable supervisors to assess workers’ knowledge-based ideas by the AHP method, and (3) to motivate workers to share their knowledge by a point-based incentive system allowing for both material and immaterial benefits. Since the application of AHP in our exemplary case has proven to be effective, we suggest the employment of the AHP as a suitable method to evaluate workers’ knowledge-based ideas and to set up the incentive system on the basis of the AHP results obtained. In analogy to the fundamental scale published by Saaty, a point-based scale ranging from 1 to 9 could be implemented as proposed in Table 4. The worker in our exemplary decision making case would therefore receive six incentive points for sharing his/her knowledge by suggesting alternative 2 with its global priority of 0.657 (Table 4). Incentive points might subsequently be redeemed in real time in a reward shop. This proposed incentive system however shows certain constraints.

It should be noted, though, that such an incentive system also bears potential inherent risks. If more than two knowledge-based ideas were to be considered, the maximal achievable number of incentive points for a
knowledge-based idea, derived from global priority values obtained by the AHP, would decrease, potentially leading to workers’ disappointment and/or perhaps even resignation. Furthermore, such incentive systems need to be made resistant to manipulation. These issues will have to be examined in more detail in follow-on studies.

Table 4: Suggestion of an incentive system analogous to Saaty’s fundamental scale [14] relating incentive points to global priority values obtained by the Analytical Hierarchy Process (AHP).

<table>
<thead>
<tr>
<th>Global Priority AHP Result</th>
<th>Receiving Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 0.1</td>
<td>0</td>
</tr>
<tr>
<td>0.1 to 0.2</td>
<td>1</td>
</tr>
<tr>
<td>0.2 to 0.3</td>
<td>2</td>
</tr>
<tr>
<td>0.3 to 0.4</td>
<td>3</td>
</tr>
<tr>
<td>0.4 to 0.5</td>
<td>4</td>
</tr>
<tr>
<td>0.5 to 0.6</td>
<td>5</td>
</tr>
<tr>
<td>0.6 to 0.7</td>
<td>6</td>
</tr>
<tr>
<td>0.7 to 0.8</td>
<td>7</td>
</tr>
<tr>
<td>0.8 to 0.9</td>
<td>8</td>
</tr>
<tr>
<td>0.9 to 1.0</td>
<td>9</td>
</tr>
</tbody>
</table>

5. Summary and Suggestions for Further Research

Awareness of knowledge as the 4th factor of production, as well as the effective management of knowledge, are crucial in today's dynamic production environment. In particular, knowledge transfer, and the assessment of workers’ knowledge-based ideas, pose major challenges in the field of knowledge management. The empirical results of this present research reveal that production companies expect high-quality improvements in manual assembly by means of effective knowledge management. Moreover, individual recognition of workers by their supervisors was identified as the most motivating immaterial incentive to share knowledge. The Analytical Hierarchy Process (AHP) as applied in our exemplary case displayed great suitability as a tool to assess workers’ knowledge-based ideas. A knowledge management software application running on tablets is recommended as the preferred cognitive assistance system to enhance the intra-organisational knowledge transfer and also, by enabling the AHP assessment of workers’ knowledge-based ideas, to reward the workers based on the AHP assessment of worker’s knowledge based ideas.

The applicability of the AHP should therefore be examined in the production environment to obtain a larger number of results and to arrive at a consolidated judgement of its suitability in this particular setting. For the implementation of an incentive system, it is recommended to assess already existing company-specific systems in respect of their usefulness and, if needed, to adjust, expand or redesign them towards a purposeful knowledge transfer. For the design of a knowledge management app, as suggested in this paper, the technical implementation, as well as the user-interface of the application, need to be precisely defined and the integration into higher-level software applications needs to be studied in more detail.

Acknowledgements

This publication forms part of the ‘SynDiQuAss’ project for excellence in assembly in the context of Industry 4.0. It is funded by the German Ministry of Education and Research (BMBF) and the European Social Fund (ESF) (grant no. 02L15A281).
References


Biography

Robin Sochor (*1992) obtained his Dipl.-Ing. (FH) degree in Mechanical Engineering and his M.Eng. degree in Industrial Engineering at the University of Applied Sciences at Kempten. Since 2017, he has been conducting research on cognitive assistance systems at Fraunhofer IGCV in Augsburg.

Teresa S. Schick (*1995) holds a Master’s degree (M.Sc.) in Management & Technology from the Technical University of Munich with majors in Finance and Accounting and Mechanical Engineering. She currently serves as Scientific Officer to the Director of the Fraunhofer Institute for Systems and Innovation Research, Karlsruhe.

Lukas Merkel (*1990) obtained his M.Sc. degree in Mechanical Engineering and Management from the Technical University of Munich. Since 2015, he has been conducting research on cognitive assistance systems at Fraunhofer IGCV.

Prof. Dr.-Ing. Stefan Braunreuther serves as professor at the Faculty of Mechanical and Process Engineering at the Augsburg University of Applied Sciences, and as head of the department of production planning and control at Fraunhofer IGCV.
Prof. Dr.-Ing. Gunter Reinhart serves as director of the iwb at the Technical University of Munich and as head of the Fraunhofer IGCV. He conducts research into assembly technology, automation and robotics, as well as production management and logistics.
1st Conference on Production Systems and Logistics

An Upgradable Cyber-Physical System Enabling Smart Maintenance of UV Lamps in Industrial Applications

Max Böhm1, 2, Dominik Lucke1, Konrad von Leipzig2
1 Reutlingen University, Alteburgstrasse 150, 72762 Reutlingen, Germany
2 Stellenbosch University, Banghoek Road, 7600 Stellenbosch South Africa

Abstract

The supply of customer-specific products is leading to the increasing technical complexity of machines and plants in the manufacturing process. In order to ensure the availability of the machines and plants, maintenance is considered as an essential key. The application of cyber-physical systems enables the complexity to be mastered by improving the availability of information, implementing predictive maintenance strategies and the provision of all relevant information in real-time. The present research project deals with the development of a cost-effective and retrofittable smart maintenance system for the application of ultraviolet (UV) lamps. UV lamps are used in a variety of applications such as curing of materials and water disinfection, where UV lamps are still used instead of UV LED due to their higher effectiveness. The smart maintenance system enables continuous condition monitoring of the UV lamp through the integration of sensors. The data obtained are compared with data from existing lifetime models of UV lamps to provide information about the remaining useful lifetime of the UV lamp. This ensures needs-based maintenance measures and more efficient use of UV lamps. Furthermore, it is important to have accurate information on the remaining useful lifetime of a UV lamp, as the unplanned breakdown of a UV lamp can have far-reaching consequences. The key element is the functional model of the envisioned cyber-physical system, describing the dependencies between the sensors and actuator, the condition monitoring system as well as the IoT platform. Based on the requirements developed and the functional model, the necessary hardware and software are selected. Finally, the system is developed and retrofitted to a simulated curing process of a 3D printer to validate its functional capability. The developed system leads to improved information availability of the condition of UV lamps, predictive maintenance measures and context-related provision of information.

Keywords

Smart Maintenance; Cyber-physical System; Smart Factory; UV lamps

1. Introduction and motivation

UV lamps are used in a wide range of applications of machines and equipment. One application area comprises disinfection and sterilization processes such as water treatment. Here, UV lamps are a safe and environmentally friendly solution. Another application area of UV lamps is in the hardening of polymers in coating equipment or 3D printers. Traditional UV lamps provide a broad light spectrum and high energy, and even though a short lifetime compared to LED-based UV lamps, these are still required in the mentioned application areas. From an economic perspective, the UV lamps are often a big cost driver in the operational cost of machines and equipment. Unplanned failure of an UV lamp in a machine or equipment can have far-reaching consequences, resulting in high failure cost. Here, objectives of maintenance are to safeguard a high safety, availability, reliability and value preservation at a minimum cost [1]. The value-adding potential
of maintenance, therefore, results from the three to five times higher follow-up costs avoided as a result of inadequate or neglected maintenance [2]. Preventive, condition-based and predictive maintenance strategies avoid unplanned breakdowns. So far, there are only a limited number of established methods to monitor and predict the condition of components in machines and equipment, such as vibration analysis, ultrasonic analysis, oil analysis or the analyses of electrical parameters [3]. However, for a large number of critical components, there is still no established method for condition monitoring which is the foundation of predictive maintenance strategies. In particular, there is no method described in the literature for the upgradable condition monitoring of UV lamps, although UV lamps are used in a wide range of industrial applications and have often a leading role in the respective processes. For this reason, approaches must be found to improve the maintenance of UV lamps. UV lamps are often part of machines and equipment and do not have an individualised or extra condition monitoring system. This motivates the following main requirements for the system to be developed:

- simple adaption of the system to the intended application for upgrading existing machines and equipment;
- using reliable low-cost components in order to be also affordable in developing countries;
- improving the availability of information about the condition of the UV lamp;
- enabling predictive maintenance strategies.

One possibility to address the differing objectives and challenges is to apply information and communication technologies and use the approach of cyber-physical systems (CPS) [4] [5]. CPS can be applied on basically all applications in various areas of the factory, such as development, production planning or maintenance. A value-adding network including a factory formed by CPS is then called a Smart Factory, for which several concepts already exist [6] [7] [8]. Smart maintenance in this context is defined as “learning-oriented, self-regulated, intelligent maintenance to maximise the technical and economic effectiveness of maintenance measures through the use of digital applications, taking into consideration the respective existing production system” [9]. This paper presents an upgradeable low-cost CPS focused on the purposes of maintenance and is in this paper referred as a smart maintenance system. The goal of this smart maintenance system is to improve the maintenance of UV lamps in machines and equipment.

2. Related work

For the development of the envisioned smart maintenance system approaches and architectures of existing smart maintenance systems, as well as the lifetime estimation of UV lamps, are considered.

2.1 Smart maintenance systems

In both literature and in the market there are several existing approaches for smart maintenance systems. The basis of a smart maintenance system is a planning system that manages maintenance objects, schedules, and controls maintenance tasks. In addition, other functions can also be integrated with the planning system, such as spare parts and ordering, maintenance controlling or maintenance personnel management. The IT system thus supports the maintenance staff in the planning and execution of maintenance activities. The overall goal of such smart maintenance systems is to dynamically adapt the maintenance strategy based on the current condition and remaining lifetime of components and machines as well as planning information such as the production plan or the shift calendar of maintenance workers. Therefore, a smart maintenance system integrates modules for maintenance management and planning, condition monitoring, deterioration and remaining useful lifetime calculation [2] [10] [11]. On a technical side, the so-called Industrial Internet of Things (IoT) architectures can be used as a foundation. One example is the IoT reference architecture by Guth et al. [12]. Also, there are mostly domain agnostic commercial solutions such as the ActiveCockpit of Bosch Rexroth AG available in the market already [13]. Commercially available systems in the market for
maintenance are made either for a specific solution, for instance as bearing monitoring such as ABB ability [14] or I-care [15]. Other solution covering only a part of a smart maintenance system such as a condition monitoring system or platforms for predictive maintenance such as ABB ability, General Electric Predix [16] or Mathworks [17]. Further solutions are based on traditional computerized maintenance management systems for planning.

2.2 Influencing factors of the lifetime of UV lamps

In the literature, there are only a few rough lifetime models of UV lamps by manufacturers. Typically, a UV lamp with mercury spectrum has a lifetime of 1500 hours specified by the manufacturer. The emitted UV radiation of a UV lamp decreases with increasing operating time due to physical effects. As soon as 75% of the initial UV intensity is measured, the useful lifetime is exhausted. The lamp can still be operated, but it can no longer be used for curing. Also, the lifetime of the UV lamp is stated under the premise that a maximum of three starts per day are permitted. Each additional switching on and off of the UV lamp leads to a reduction of the lifetime by 0.5 hours. Furthermore, UV lamps have to be cooled, because the glass temperature is about 600 °C – 900 °C and from temperatures above 1000 °C the quartz softens and the UV lamp inflates or bends. For optimum cooling, an air volume of 100 m³/h applies per kW UV lamp output. In addition, the resulting hot air cushions are aspirated above the UV lamp [18] [19]. In highly accelerating lifetime tests with discharge lamps, the vibration experienced by the lamp needs to be considered, as vibration has an influence on the lifetime [20]. In summary, it can be said, that the burning time, which is the time the lamp is being switched on, is the key factor leading to wear – through further factors which reduce the lifetime of a UV lamp are switching on and off, increased temperature and vibrations during operation.

3. Concept

The main idea in order to improve the maintenance of UV lamps as critical and costly components in machines and equipment is to use the approach of a cyber-physical system. Here, the key element is the functional model, describing functions and the relations between the different modules (Figure 1). Main modules of the smart maintenance system for UV lamps are the condition monitoring system and the smart maintenance planning system presented in the following sections.

3.1 Condition monitoring system

Sensors and actuators build the foundation of each data-driven application. In this case, as relevant measures for estimating the remaining useful lifetime of a UV lamp the temperature, acceleration and UV intensity, as well as the electrical power consumption, have been identified. The corresponding sensors are connected to the signal processing of the condition monitoring system. The measurements are carried out automatically and at defined time intervals. An actuator is provided as optional, in order to control the UV lamp directly. In the present case, the device forms the condition monitoring system and aims to monitor the condition of the UV lamp and can intervene if required by the actuator. All functions for the acquisition, processing and storing of the collected data is performed locally on-site, with the goal to reduce the data to prediction relevant information. Also, the device acts as a connection component to facilitate the connection of different sensors with the smart maintenance planning module. Driver software integrated on the device allows
uniform access to the different sensors. The device has communication standards like I2C, Onewire, IP, Ethernet, WiFi, Zigbee or a transport protocol like HTTP or MQTT and has a compatible payload format, which means that no further gateway is needed for communication to other IT Systems. The measured values are temporarily stored for a short time on the device and then sent to the IoT platform via a transport protocol.

### 3.2 Smart maintenance planning

The main functions of the smart maintenance planning module are an analysis and categorisation of information received from the condition monitoring system as well as a determination of a remaining useful lifetime of the UV lamps. Further functions are related to context-related maintenance information provision and visualisation of the communication with the human planner or supervisors. The smart maintenance planning module is implemented on an IoT platform, which can run as a local or cloud-based solution. The IoT platform receives the protocol with all measured values sent by the device, adds a timestamp and stores it in a database. The individual measured values of the sensors are clearly displayed in diagrams in the dashboard of the IoT platform. The live values are presented here, but historical values can also be viewed. Furthermore, the remaining useful lifetime of the UV lamp is calculated as well as on and off switching operations. As a result, the maintenance staff can get an overview of the live status of the UV lamp and at
the same time have access to important information about the UV lamp, both for the currently running process as well as about past processes.

The maintenance technician has a decisive role in maintenance. It is therefore suggested that context-based maintenance measures are used. In this case, the specific place in the user manual where the UV lamp replacement is described is mentioned. The maintenance staff is informed by their preferred output medium such as a computer, a tablet or a smartphone, which maintenance measures he should carry out in order to preserve or restore the operating condition of the UV lamp. The stated information is automatically sent by SMS or email with a PDF attachment to the maintenance staff. The maintenance staff can now access all relevant information from a mobile device and perform the suggested maintenance measure based on this. He can intervene in the UV lamp in different ways. For the actuator, power measurement values can be defined from which the UV lamp can be switched off. The UV lamp can also be switched off and on directly via the actuator.

4. Implementation

Based on the requirements and the functional model, the hardware and software are selected and implemented. The selection criteria for the hardware and software components are a low cost, wide availability and preferably open source. Based on these premises for the implementation of the smart maintenance system the following components have been selected:

- Sensors and Actuators:
  - Adafruit MCP980 temperature sensor
  - Adafruit MMA845 accelerometer
  - Adafruit VEML6075 UV sensor
  - HomeMatic radio switching actuator with power measurement

- Condition monitoring system:
  - Raspberry Pi 3 Model B
  - HomeMatic radio gateway

- Smart maintenance planning module:
  - Cloud-based Losant IoT Platform

The smart maintenance system monitors the influencing factors for the lifetime of UV lamps. For this purpose, the burning time, the number of the switch-on and switch-off operations per day, the temperature in the direct vicinity to the UV lamp as well as the vibration that the UV lamp experiences is captured with the mentioned sensors. Furthermore, the relative UV intensity of the UV lamp is determined since the UV intensity influences the process result of the present UV application. For the implementation, the measurands are temperature, acceleration, relative UV intensity as well as power-related measurands such as current and voltage. The measured values are categorised into an acceptable, critical and rejected range on the basis of prescribed limit values in order to preserve the operating condition of the UV lamp and adhere to the required process factors.

The starting point for the calculation of the lifetime is the useful lifetime of 600 hours specified by the manufacturer. In addition, the existing lifetime model is integrated into the calculation. The model describes additional wear of the UV lamp of half an hour per further start of the UV lamp, which is performed for the fourth time or higher within 24 hours. A function is generated to compute the remaining useful lifetime that uses the measured current and a timestamp to determine the burning time as well as the switch-on and switch-off operations within 24 hours. Subsequently, the remaining lifetime is analysed and categorised. As an example, the acceptance range is defined from 600 to 50 hours of remaining useful lifetime. The critical range is stated between 50 and ten hours. The dashboard shows the message that the critical remaining
lifetime has been reached. It also states that the stock of replacement lamps must be checked and ordered as required. This ensures that a replacement lamp is available when the remaining useful lifetime is over, and no unplanned machine downtime occurs. In addition, emails and SMSes will be sent to the maintenance staff with the above information. The email could also be sent directly to the spare parts supplier who provides a new UV lamp. The rejection range is accordingly from ten hours to the end of the useful lifetime. A red highlighted alarm message appears that the remaining useful life is reached.

The visualisation of the measured and categorised values are displayed in a dashboard. The Losant IoT platform is used for this purpose. An individual function for each temperature sensor categorises the values into an accepted, critical and rejected range. In the dashboard, accepted values are highlighted in green background and no further information is required. Values in the critical range are highlighted in a yellow background, and a warning message appears. In addition, the specific sensor location where the criticality occurred is displayed. Also, the values of the acceleration sensor are given in a diagram in the dashboard. The acceleration is given in x-, y- and z-axes and is expressed in the unit m/s². The values are stored, and it is checked afterwards how significant the influence of vibration is on the lifetime of the UV lamp. Also, for analysis purposes, the power measurement captures the values such as voltage, current, effective power, frequency and energy consumption. All values are displayed separately in a diagram. The values of the UV intensity in the UVA and UVB spectrum are also visualised in a diagram. The dashboard features live and historical values. The initial values of the UV intensity in the UVA and UVB spectrum measured with the UV sensor are used as the reference values. All following measured values in the UVA and UVB spectrum are set in relation to the defined initial values. The acceptance range is defined as 80 % to 100 % of the initial UV intensity which corresponds to a remaining buffer of 20 % until insufficient UV intensity is reached. The critical range is therefore from 75 % to 80 % of the initial UV intensity.

In a second step, an email and SMS is sent to the maintenance staff informing them where measured values have been exceeding the critical range. As soon as a value is measured in the critical range, the dashboard highlights in yellow or red and an alarm message pops up with additional recommendations e.g. that the system must be switched off immediately. The message sent by SMS and email also contains an attachment with a report and suggested maintenance measures. In this report not only measurement values and the maintenance object is stored, but also the relevant section in the user manual is named to provide the appropriate maintenance measure for the present case. In this way, search times for required information are reduced through improved user guidance, as context-related maintenance measures are suggested to the maintenance staff.

5. Validation of the smart maintenance system

The developed smart maintenance system is validated on a system, which simulates the application of a UV lamp in an additive manufacturing process (see Figure 2). The validation process is carried out according to a defined test cycle, which is based on three different factors. The printing mode can be set to high-speed or high-quality mode. The size of a simulated printed object is preset to either 50 mm or 100 mm, which is stimulated by the linear working spindle. The complexity of an object to be printed is determined as low or high, which leads to a simulated printing time twice as long with a high complexity. A combination of the three factors results in a total of eight different printing settings. The performed validation reveals that the functional capability of the smart maintenance system is achieved. All previously defined requirements for smart maintenance are fulfilled. The system is able to provide information about the condition and the remaining useful lifetime of the UV lamp by using an existing lifetime model. As soon as measured values are in critical or rejected ranges, measures are initiated to preserve the operating condition of a UV lamp such as the provision of context-related information. For the application of the smart maintenance system for other UV lamps, the influencing factors of the components to be investigated are already known.
lifetime can be determined on the basis of the existing lifetime model and adapted to the lifetime specified by the UV lamp manufacturer.

![Figure 2: Setup of the printing process simulation system](image)

The required sensors are retrofitted to the printing process simulation system (see Figure 3).

![Figure 3: UV sensor under the UV lamp with reflector unit, temperature sensor next to the UV lamp, acceleration and temperature sensor in the upper part of the printing chamber (from left to right)](image)

If a lifetime model of the UV lamp applied is available, this model can be used as a basis. Furthermore, the factor which has an influence on the process result of the application must be known. For an application with a certainly required UV intensity, a reference value must first be defined. The reference measurement is carried out with the UV sensor at the same distance to the UV lamp as the sensor is integrated within the later application. The defined reference value is set in relation to the values determined during the operation of the application. Subsequently, the respective limit values for the application have to be determined, such as the required UV intensity or external temperature.

6. Conclusion and outlook

This paper presented the concept and implementation of an upgradable low-cost CPS, to enable smart maintenance of UV lamps in machines. Further research is needed to provide more accurate predictions about the remaining useful lifetime of the UV lamp. Also, with regard to the influence of vibration on the UV lamp, further testing needs to be done. The approach of the statistical design of experiments serves as the basis for the creation of a more accurate lifetime model. More specifically the developed additive manufacturing process simulation system can be used for the execution of further experiments. The statistical experiments have to be carried out to obtain the precise wear of the UV lamp depending on the load which is derived from the size and complexity of the objects to be printed as well as the printing mode setting. The objective is to develop a model that calculates the remaining useful lifetime taking into account the experienced, current and expected load of the UV lamp. The experienced load reflects the wear caused by the prints already performed, whereas the current load is the live wear during the printing process. The
expected load represents the wear of the UV lamp, which is predicted due to the characteristics of the object to be printed and the printing mode settings. In addition, the expected load on the wear of the UV lamp of the following objects to be printed can be considered. In this way, the objects to be printed can be individually evaluated according to their expected load on the wear of the UV lamp and the objects can be selected or prioritised depending on whether the remaining lifetime is still sufficient.

References


Biography


Prof. Dr.-Ing Dominik Lucke (*1980) is a professor at the ESB for production technology, automation and digitization of production and project leader at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart in the department „Sustainable Production and Quality“. He studied mechanical engineering from 2001 - 2007 and finished his doctorate in 2013 at the University of Stuttgart. From 09/2007 - 09/2017 he worked at the Institute of Industrial Manufacturing and Management IFF of the University of Stuttgart as well as at the Fraunhofer IPA in the topics digital factory, smart factory, factory planning and maintenance management. His focus topics are apart from the development of industry 4.0 applications, the optimization of maintenance.

Konrad von Leipzig (*1958) is a senior lecturer at the Stellenbosch University and director of the Institute for Industrial Engineering. He studied Industrial Engineering (M.Eng.) at the Stellenbosch University from 1991 – 1993. From 1986 – 1990 he completed his Bachelor of Commerce at the University of South Africa. He studied Industrial Engineering (B.Eng.) at the Stellenbosch University from 1978 – 1981. His focus topics include the application of smart technologies in a developing and skills-scarce environment.
Part Based Mold Quotation With Methods Of Machine Learning

Sebastian Brede¹, Benjamin Küster¹, Malte Stonis¹, Mike Mücke², Ludger Overmeyer¹

¹Institut für Integrierte Produktion Hannover gGmbH, Hannover, Deutschland
²Phoenix Contact GmbH & Co. KG, Blomberg, Deutschland

Abstract

The dominating mass-manufacturing process of today is plastic injection molding. This production process uses economies of scale because parts are produced in seconds at marginal cost of plastics. However, upfront investment costs for the tooling of molds are the basis for deciding if a mold is tooled and hence if a part is viable for mass-production. If tooling costs are too high, a product may not viable for production. If tooling costs are estimated too low by the tool shop, contract implications may arise. Because injection molds differ in their complexity, price estimations for the tooling of molds are an ongoing quest. There are various methods for estimating the costs of injection molds such as rule based, analytical or data driven approaches. The advantage of data driven approaches is the ability of adjusting to historical production data as well as readjusting while training on new batches of recent data.

The focus of our research was to support the quotation process of tool shops. To this end, we studied a data driven machine learning approach.

The goal of this research is to develop a method with humanlike quotation accuracy, achieve standardization, factor in historic quotation data and shorten quotation process times. The machine learning approach developed is based on geometry data of parts and additional meta-information.

Within this research, a system was developed to interact with live production systems of an electronic part producing tool shop.

The method developed was trained and validated on production data in a case study. To enhance the quotation process, the method developed was embedded into a server-based application with a web user interface and interfaces to live production systems for the automation of processes.

Keywords

Machine Learning; Injection Molding; Cost Estimation

1. Introduction

1.1 Production by injection molding – quotation of molds

Injection molding of thermoplastics is a mass manufacturing process which enables an economic production of goods, i.e. a fast production process which needs little to no postprocessing of parts. To ensure an economic efficiency of production, costs must be considered as early as possible. With injection molding, there always needs to be a distinct mold for each part produced.
Production costs may be divided into the two categories of fixed costs and variable costs. While variable costs depend on production quantity such as material, energy and machine costs (as a product of machine rate and required process time), fixed costs are often created before the actual production of plastic parts begins. While the retrieval of variable costs is trivial, fixed costs may be unpredictable due to the manufacturing processes of molds and can create a great amount of debt that needs to be compensated during production.

This research is focussed on the estimation of fixed costs that result from the tooling of molds. The upfront costs of mold tooling not only consist of machining costs such as milling, sanding, drilling or electrical discharge machining but also of manual labour-intensive assembly and inspection. Because thermoplastic materials have a high shrinkage ranging from $s_l = 0.015 \frac{cm}{cm}$ to $s_h = 0.060 \frac{cm}{cm}$ [1], machining molds of tightly tolerated parts poses a challenge. To face the resulting difficulties, molds are often quoted by workers with a high level of experience in the field of tooling and quoting of molds. This experience-based cost-method is subject to a worker’s individual abilities and therefore prone to errors. The aim of this research is to evaluate a method that supports the quotation process of injection molds and accounts for the characteristics of given historic data.

1.2 State of the art

During the last decades, various attempts at estimating mold costs have been made. The research presented in this section therefore focuses on methods that establish machine learning techniques for cost estimation.

In 2003, Wang et al. researched a case-based reasoning (CBR) system to aid workers’ quotation of injection molds [2]. For retrieving similar parts, they used artificial neural networks (ANN) to calculate similarities between parts. The input parameters consisted of meta-parameters such as injection part type, tolerance requirements and parting plane complexity as well as direct properties such as part size and number of internal / external undercuts. The validation performance of the CBR-system Wang et al.’ proposed was not disclosed.

In 2005, Xu et al. researched a Fuzzy ANN for estimating injection mold design time [3]. Their dataset consisted of 72 samples where 60 samples were used for training. Besides the direct parameters of part size, part precision and minimum thickness, they input the meta-parameters part structure and part shape. They report mean absolute percentage errors of 9 - 12 % for the models they researched.

Wang et al. researched the cost estimation of injection molded parts multiple times [4],[5],[6]. However, they did not focus their research on the mold’s tooling but rather calculated the whole design to market costs, including the injection molding production process, of a given part. In 2007, they researched a back-propagation ANN with 1-2 hidden layers with 34-36 hidden neurons [4]. The input parameters were part- and mold-specific; material, weight, surface area, parting line’s projection surface area, bounding box dimensions, volume, wall-thickness and cavity quantity. They used an industry-related dataset with 1070 samples.

In 2010, Che from the same research institution as Wang et. al. evaluated a particle swarm optimization (PSO) based ANN on a bigger dataset containing 2100 samples using the same parameters as Wang et al. [5]. Wang et al. picked the PSO-ANN up again in their 2013 research [6]. The ANN studies mentioned above have a remarkably low relative error ("cost percentage error") of maximum 1.95 % [6], 0.72 % [5] and 1.1 % [6] in their 20 validation samples from a total dataset of 2100 samples although they incorporate only basic part features and a small ANN.

Florjanic et al. undertook expert’s opinion to decide upon 22 input variables which they used for the estimation of total machining hours spent per part via ANN [7]. Besides part specific input-variables such as part material, surface area and complexity, they also considered incorporating mold variables such as
mold dimensions, parting line complexity, slider- and lifter-count. These mold variables created the need to include expert’s opinion into the quotation process. Their final ANN consisted of one hidden layer with four hidden neurons. Their dataset consisted of 105 samples with 20 % used for evaluation. They encountered maximum relative errors of 38 % during their testing and advised a safety factor of 25 % to be applied to their model for performing predictions.

Eilert et al. researched an ANN method to predict manufacturing costs of sheet metal forming dies [8]. They automated the construction of dies from a given part and estimated tooling costs given on automatically constructed die features, such as ejector and stage count. They used a total of 30 samples as dataset for training and testing.

Kočov et al. developed an expert system for assistance with injection mold’s cost estimation [9]. They designed their expert system with focus on guiding an expert worker through a quick, basic design of a possible injection mold for a given part. They then infer an estimated cost from the given mold design.

Börzel and Frochte investigated a method to estimate die-casting molds’ tooling costs from small datasets [10]. They reduced geometry information from 3D computer aided design (CAD) data to scalar features. They considered common geometric properties as well as ratios of these and expert parameters (e.g. surface area influencing demolding). They performed a principal component analysis and found the first three principal components describing 96 % of the data variance. These principal components incorporate six parameters, namely: Part volume, surface area, projected bounding box surface area, bounding box volume and two specified bounding box side lengths. It is noteworthy that no expert parameter nor any of their defined ratios was part of the main principal components. With the mapped input parameters, they evaluated various machine learning techniques such as regression models trained by ANN, K-nearest neighbor regression and random forests regression (RFR). They grouped their 700 samples dataset into three groups of mold designs and trained distinct models for these. After this they performed evaluation on 20 % of their dataset of 700 samples. They identified outliers with more than 50 % error and achieved a mean relative uncertainty of 10 % to 14 %.

2. Conception of quotation method and model architecture

2.2 Conception of quotation method

With the methods researched by the authors mentioned above, parameters were used that describe only one single property of the part, e.g. part surface area and part volume. However, this poses a problem, because these parameters have a one-to-many relation: While, for example, a part’s volume is dependent on a part’s geometry and there exists exactly one volume for a given geometry, there may be infinite geometries possible for a given volume. This means, these parameters can only represent parts to a certain degree and will fail when parts with the same scalar parameters but different underlying geometries are considered.

Some authors added complexity parameters such as part complexity and parting line complexity [7]. But these parameters bear two problems: On one hand, these input parameters are a dimensionality reduction of the actual features which cannot be reconstructed by the input parameter’s description. On the other hand, an expert’s opinion is needed to decide on a part’s level of complexity. Therefore, a method which accounts for the actual geometric features and their spatial relations is needed.

The seemingly natural way for inputting the part’s geometric three-dimensional (3D) data into neural networks would be to input vertices’ normalized position data into 3D convolution layers. However, this poses several challenges: On the one hand, 3D CAD data is used for part and mold construction. While 3D CAD data is unit driven, its surface-topology is secondary. It is defined / visualized during the tessellation stage by means of algorithms as marching cubes or marching tetrahedrons. That means, the resulting
topology which defines the vertices’ quantities and positioning, is dependent on the algorithms used for exporting the part data.

On the other hand, 3D data can easily be corrupted, i.e. non-manifold surfaces or bad distribution of vertices. To overcome these difficulties, a visual representation of a part’s 3D-data is proposed.

Because a 3D part has more than one possible side, a multi-view convolutional model, based on multi-view projections of the input part, is chosen for the processing of 3D data. These kinds of models were researched by Su et al. to classify various 3D models with a high precision compared to other methods such as 3D-voxel grids, 3D-shape descriptors or pointclouds [11].

To guarantee every part an equal image proportion, the part needs to be normalized to fit in set input layer boundaries. To account for meta-information not representable within a part’s geometry - such as planned mold features – we propose an extra, parallel neural network.

Both model architectures are then merged to output the quotation estimate. The choice of meta-parameters to be included in the meta-information is made by examining expert’s opinion, and by performing a statistical effect analysis on the data. The meta-parameters chosen are number of part-cavities in molds, mold-type, project type and runner-type.

### 2.3 Model architecture

The underlying model is separated into two parts, one convolutional part for image processing as well as one multi-perceptron part for meta-information processing. A schematic of the used model architecture is displayed in Figure 1. The convolutional part of the model (Figure 1, upper neural network part) is an adoption of Su et al.’s multi-view convolutional neural network, which in turn is an adoption of AlexNet, an early image processing model used for image recognition [12] with an additional view pooling layer [11].
That means, each of the derived 20 images is fed into an individual convolutional neural network with a first layer of 512 px \( \times \) 512 px as input layer. After input layer, a batch normalization layer to readjust the submitted images is inserted. After the batch normalization, various convolutional layers with differing filter sizes (as depicted in Figure 1) and differing strides, this means the distance step during filter convolution [13], follow. Between each filter layer, a maximum pooling layer with set pooling size of 3 and stride of 2 is inserted after each convolution layer except for layers 3 and 4. A spatial drop-out layer, to randomly drop filters and therefore counteract possible overfitting is placed between every convolution layer. To combine the activations of all 20 convolutional neural networks a flat, dense layer is used as a view pooling layer.

For additional meta-information, a flat, dense network architecture is chosen. This architecture is assigned once per submitted part. The flat layers start with a varying, dependent on chosen parameters, scaled input layer. After input layer, a normalization is made and several flat layers with architecture as depicted in Figure 1 follow. Between flat layers, only dropout layers are used.

To combine the view pooling layer as result of the convolutional neural network and the last layer of the flat, dense network for meta-information, both are concatenated. After this concatenation layer, a series of flat, dense layers follow. The last layer is the output layer, consisting of the number of classes created by mapping the costs to classes.

For activations, scaled exponential linear units, which are proposed as adding robustness during training by Klambauer et al. are implemented [14]. As cost function, cross entropy as well as absolute difference with equal weight is applied. As optimizing function, Adam optimizer, known for its ability to cope with sparse and noisy gradients, is utilized [15].

3. Case Study: Evaluation of proposed part cost quotation method

3.1 Software architecture overview

To accommodate for architectural difficulties arising from different operating systems and varying states of installed programs, the software to be developed is conceptualized as a web service. This enables to provide for good usability, rapid software rollouts, identical software used by every user and centralized data storage as well as user access-management. Because of the centralized data storage, it becomes possible to gather further data samples to retrain the quotation model on more recent data.

Because the proposed quotation method is to be tested on real data in a real production environment, certain requirements, as a link to enterprise resource planning software, construction database and CAD software must be established as well as good usability. The schematic of the base softwares’ architecture is displayed in Figure 2. The software is divided into a back-end server task and a front end. While the back end is designed as two separate, different server instances (creating separate server tasks), the front end is designed within web technologies HTML5 and JavaScript. The back end is realized with the Python programming language in its 3rd version and a micro framework flask for routing and managing user requests. Model serving and deriving of projections is coordinated via task queue module huey to decouple long lasting calculations from back end processes and to enable an enhanced user experience. Projections and measurements are either made with a call to proprietary CAD software Creo Elements Direct 3D Access executing a Common-Lisp script for files submitted in proprietary Creo file-format or with a call to open source software blender.

For gathering of training data, a link for retrieving historic records of part-data is made via web-interface of the construction data system. To ensure restricted access to model maintenance tasks and system/user database management a user access-rights system is implemented. The ANN is programmed with open source deep learning framework tensorflow in its version 1.13.
3.2 Data pre-processing / model training

700 samples are created by gathering of 3D CAD data of injection mold parts and their corresponding meta-information. Each sample contains unique, historic part quotation data. For alignment of part’s 3D CAD data, each part is normalized to a length of longest part side of $l_{part} := 1$ according to equations 1 and 2:

$$l_{max} = \max(l_{part_x}, l_{part_y}, l_{part_z})$$ (1)

$$l_{part_i} \Rightarrow \frac{l_{part_i}}{l_{max}}$$ (2)

There are 20 image projections derived via a rendering of the part’s 3D CAD data as shown in Figure 3. The virtual camera alignment is adopted from Su et al.’s second camera setup; 20 cameras are placed on the vertices of a sphere around the part [11]. Hereby the images are rendered in 8-bit grayscale with black background and a resolution of 512 px · 512 px. Meta-information is prepared either by a categorization through grouping with a k-means algorithm for part properties with more than 3 levels or by normalizing to range of $\{i \in \mathbb{R} | 0 \leq i \leq 1\}$. 
Each sample therefore consists of the two-dimensional projections and the normalized meta-information. Manufacturing cost data is used as label data. The cost data is rounded and mapped to the closest cost class with an equidistant cost distribution with 5,000 EUR steps. This means, the cost regression task is mapped onto a classification task. Before the conversion to a binary format for enhancing the data feeding pipeline, sample data is shuffled randomly and split into a training and testing dataset by a split of 80:20. The testing dataset is kept in a separate binary file reserved for model evaluation.

The model training is performed on gpu-hardware and is stopped after one week and after the convergence of the train-loss curve. By this time, $172 \cdot 10^5$ steps have been trained.

4. Evaluation of quotation method

4.1 Results of model evaluation

The trained model is evaluated on $n_{\text{eval}} = 20\%$ of total samples. For reference, a model trained on the same dataset but with only the convolutional part of the proposed model is compared to the model trained on parts’ projections and meta-information. The results are displayed in Figure 4. The share of classes as a function of relative errors are shown. The relative error is calculated by equation 3 and grouped into classes for ease of display.

Figure 4: Display of model evaluation. Shown are grouped shares of relative errors of total evaluation dataset
The grouping is done in steps of 10% with the final step being a group for all relative errors with $|\varepsilon_{rel}| > 100\%$.

$$|\varepsilon_{rel}| = \frac{|c_{pred} - c_{real}|}{c_{real}}$$  \hspace{1cm} (3)

It can be noticed, that overall, the extended model scored twice as good as the model solely relying on image data. The model trained with additional meta-information scores twice as many parts without prediction error than the model solely relying on image data. Predictions with errors in class $\varepsilon_{rel} > 100\%$ by the model with additional meta-information are at just the tenth of the other models. The mean relative error is calculated as displayed in equation 4.

$$\bar{\varepsilon}_{rel} = \frac{\sum |\varepsilon_{rel}|}{n_{samples}}$$  \hspace{1cm} (4)

For the model trained solely on images, the resulting mean relative error is calculated as $\bar{\varepsilon}_{rel} = 112\%$. The extended model with additional meta-information scores a mean relative error of $\bar{\varepsilon}_{rel} = 51\%$.

4.2 Discussion of results

As can be seen in the results, the mean relative error could be lowered to less than half of the model’s error without the extra meta-information NN. Comparing the results to the models of Wang et al., their mean relative errors being $\varepsilon_{rel} < 2\%$, their achieved errors are comparably low [4], [6]. It is important to underline, Wang et al. not only estimated the fixed upfront mold costs, but also the production costs to produce a part via injection molding. Xu et al.’s evaluation’s mean relative error is as low as $\varepsilon_{rel} = 12\%$ [3]. Florjanic et al.’s work is closest to the results developed during this research with $\varepsilon_{rel} = 38\%$ [7]. For comparison, the total size of the datasets used and the validation sample size must be considered. For example, for Wang et al., the validation sample size is at 1.9% of their total dataset [6]. On the other hand, NN models are data-driven, and comparing these models without access to underlying datasets is questionable.

In contrast to the cited authors’ methods, the system we developed is designed not to need any mold specific parameters such as mold dimensions or parting line complexity, and its inputs do not rely on experts’ opinion as in determining part complexity. To further enhance the model’s performance, the server software system developed is designed to be extensible by adding further samples to the dataset.

5. Conclusion

During this research, a method to support the cost quotation process for injection molds was developed. The underlying ANN-model was trained on real world injection molding data. To evaluate the researched model, it was implemented as a stand-alone server-side application. This server-side application was interfaced to various production systems to have the possibility of gathering real data. The developed server-side application allows for continuous and automatic extension of the underlying dataset. Possibilities to further enhance the model’s performance could be achieved by adding further data to the dataset or extending parameters in the dataset as well as further aligning historic data records.

Acknowledgements

The authors would like to thank the Phoenix Contact Foundation for supporting this research.
References


Biography

Sebastian Brede (*1988) is project engineer at the IPH - Institute of Integrated Production Hannover (non-profit limited company)

Benjamin Küster (*1988) leads the department of production automation at the IPH

Dr.-Ing. Malte Stonis (*1979) is managing director of the IPH
Dr. Mike Mücke (*1982) is technology development manager for software solutions at Phoenix Contact

Prof. Dr.-Ing. Ludger Overmeyer (*1964) heads the Institute of Transport and Automation Technology (ITA) at the Leibniz Universität Hannover and is member of the management board of the IPH
Data-based identification of throughput time potentials in production departments

Lasse Härtel¹, Peter Nyhuis¹

¹Institute of Production Systems and Logistics, Leibniz University Hannover, Germany

Abstract

Logistics performance becomes an ever more important strategic factor for manufacturing companies to obtain a competitive advantage. Yet, numerous companies fail to meet their own corporate goals or customer requirements. One of the most important objectives in logistics is speed in terms of short delivery times which are mainly determined by the production throughput times. Derivation of effective improvement measures requires a profound understanding of logistic cause-effect relationships. At a time of increasing digitalization, an increasing amount of feedback data is available that offers great potentials to discover novel insights. Yet, the vast amount of data can also be overwhelming and result in unsystematic and ineffective analysis of less meaningful data. Therefore, in this paper a systematic procedure is presented that allows data-based identification of throughput time potentials in production departments. The quantitative analysis framework is based on a generic driver tree structuring the influencing factors on throughput time. The approach will boost the understanding about logistics relations and will particularly help SMEs to focus on the most relevant influencing factors and data. Furthermore, it provides a basis for future more advanced information systems that will help companies to continuously improve their logistics performance and adapt their supply chains to ever-changing conditions.

Keywords

Throughput time; Logistics; Production controlling; Data analysis

1. Introduction

Logistics performance plays a highly important strategic role for manufacturing companies to successfully compete in today’s difficult economic environment [1]. Studies show that companies striving towards consistent optimization of their supply chain regarding logistic key performance indicators (KPIs) can verifiably increase market success [2]. Besides on-time delivery particularly short delivery times are considered an essential target affecting customer satisfaction [2,3]. In order to continuously improve logistics performance, a systematic production controlling is required continuously collecting, analyzing and interpreting relevant feedback data within the closed loop of production planning and control (PPC) [4]. Digitalization of production processes and an increasing data availability offer tremendous improvement potentials regarding decision support systems in the context of production controlling. Yet, production controlling still is a mostly manual activity as companies are not willing to rely on automatically generated planning results [5]. At the same time, companies often lack the understanding of the manifold and multi-causal logistical interactions in supply chains [6,7]. Hence, there is a high risk of unsystematic data analyses and ineffective improvement measures. Current trends therefore aim at applying simulation, artificial intelligence (AI) and machine learning algorithms to identify cause-effect relations [8,9,10]. But due to the
low transparency of these methods and the chance of identifying pseudo-correlations, the risk of misleading data interpretation remains. A profound understanding of the most relevant cause-effect relationships in logistics thus is the necessary basis to effectively detect improvement potentials in logistics. Therefore, the bottleneck-oriented analysis approach has been developed applying logistic models [11]. Yet, the approach is limited to the application of production operating curves and requires a lot of expert knowledge as guidelines are missing which influencing factors to include in the analysis.

This paper shows that if the primary drivers of the logistics KPIs have been identified, comprehensive analyses can be conducted only based on a limited amount of meaningful data. Using the example of production throughput time, it is presented how generally valid driver trees can be derived based on logistic models in section 2. Subsequently, it is presented how the bottleneck-oriented analysis approach can be integrated in a systematic procedure and how the driver tree helps structuring the analysis for non-experts in section 3. An industrial case study demonstrates how the most significant levers for throughput time reduction can be identified using generally valid logistic models and simple analysis methods.

2. Influencing factors on throughput times

In the following, a brief overview about the different definitions and time shares of throughput times is provided before the main influencing factors are derived based on the theory of the logistics operating curves. Based on these fundamentals, a driver tree for order throughput times is derived setting the guidelines for effective data analysis.

2.1 Throughput time definition

Throughput time is a KPI that is measurable in the production stages along the supply chain. It can be recorded on different levels (see Figure 1). In general, it is distinguished between two types of throughput times: order throughput times and throughput times at operation level. The order throughput time comprises the time span from order release to the end of production of a production order (PO). Depending on the order fulfillment strategy, one PO may consist of several manufacturing orders in a pre-production stage supplying components for a subsequent assembly order in the end-production stage. Within each production stage, POs are being processed in different operations. According to the throughput element [12], the throughput time on operation level consists of an inter-operation time between the completion of the predecessor operation and the beginning of the next operation and the operation time. The inter-operation time comprises a share of waiting time post-processing, the time required for transportation to the next work system, and a waiting time pre-processing. Particularly in shop production the share of productive operation time is mostly marginal compared to interoperation times.

![Figure 1: The throughput element (in accordance with [12] and [13])](image-url)
2.2 Production operating curves

The most important control variable determining throughput times is the work in process (WIP) in production departments and thus the WIP level at the particular work systems. The higher the WIP level the longer the waiting queues and consequently the resulting throughput times. However, a low WIP level may cause losses in capacity utilization. This planning dilemma can be resolved by the theory of the production operating curves (see Figure 2), quantifying the relation between WIP, capacity utilization (or output rate) and throughput time of a work system [14]. In an idealized state the output rate, respectively capacity utilization, of a work system increases proportionally with an increasing mean WIP level until it reaches the maximum utilization. The WIP level at which the utilization is just the maximum without queues forming is called ideal minimum WIP (iWIP\textsubscript{min}). That means that neither a PO waits for processing nor the work system waits for its next PO. Up to this point there are no queues and throughput times hence equal the sum of the minimum transition time and the operation time. A further increase in the mean WIP does not lead to a further increase in capacity utilization, but rather causes longer queues and thus longer throughput times. In reality work systems are subject to varying workloads or capacities and other forms of disturbances (e.g. lateness of incoming orders). Hence, real operating curves are curve-shaped with the ideal characteristic curves as natural thresholds. For every work system there is an operating state where utilization losses are marginal at a reasonable WIP level (intermediate state). Higher WIP levels result in an overload state without notable utilization increases but rising throughput times. Throughput times at these work systems could easily be reduced by reducing the WIP level. In contrast, lower WIP levels cause an underload state with significant utilization losses but short throughput times.

![Figure 2: Exemplary production operating curves of a work system](image)

The shapes of the operating curves are specific for each work system. Besides the available maximum capacity, the main influencing factor is the position of iWIP\textsubscript{min}, which in turn is determined by the order workload (in target hours) and their variance in particular. Furthermore, especially the spread of workload and the available capacity flexibility have an impact on the gradient of the curves. In order to be able to compare the WIP levels of several work systems, the relative WIP level, as the ration between the absolute mean WIP level of each work system and its ideal minimum WIP level is used. The relative WIP is a suitable indicator to evaluate the operating point of work systems. While work systems with a very high relative WIP (>>300%) are likely to operate in an overload state with long waiting queues and long throughput times, very low relative WIP levels (<200%) indicate underload states. [11]
2.3 Derivation and structuring of the main influencing factors on order throughput times

Based on the throughput element and the theory of the production operating curves presented above, generally valid influencing factors on the mean throughput time of POs can be derived and structured in a logical driver tree (see Figure 3). In accordance with section 2.1, on a first level there are two main drivers on order throughput times: the general production and product structure, and the throughput times of the single operations during order processing. Regarding the production / product structure, particularly the number of process steps and / or the number of production stages define the shortest possible throughput time of a PO. The number of operations and production stages in turn, are determined by product complexity, the in-house production depth, and the position of possible order decoupling points (storage stages). Another influencing factor on the mean order throughput time in a production department concerning production / product structure is the variant-creation point. A late variant-creation for instance allows producing in bigger lot sizes upstream the variant creation point and thus minimizes necessary set-up processes. A further subdivision of the identified drivers is not made, since this is hardly possible on a generic level and requires company-specific analyses.

According to the throughput element, throughput times generally comprise two time shares: operation time, and inter-operation time. Operation times define how long a work system is blocked with an order. Firstly, this depends on the workload of the orders (measured in target hours), which is determined by the setup time, the lot size and the process time per unit. How fast order workloads can be processed further depends on the available capacity (measured in hours per shop calendar day). The capacity of a work system results from the number of parallel and substitutable work stations or employees, the working hours per shop calendar day (SCD) and idle times due to technical or organizational downtimes (e.g. errors or maintenance).

Inter-operation times arise from required minimum transition times due to transportation or technological-induced idle times (e.g. cool down after oven processes) and queuing times caused by the WIP level. As shown in the production operation curves, the WIP level is directly related to the planned operating point and therefore to the planned throughput time. The planned operating point should be the result of a logistic positioning in accordance with the underlying logistics targets. Viable throughput times at the desired operating point depend on the shape of the throughput time curve, which is affected by the maximum output rate, the mean order workload and the variance of order workloads, the input variance and the available capacity flexibility, as well as the capacity structure (e.g. two workstations and single-shift operation vs. one workstation and two-shift operation). If the actual throughput times deviate from the planned throughput times, backlog occurs that causes increasing (or decreasing) waiting queues. Furthermore, the WIP level is significantly influenced by the choice of production control methods. Workload-oriented order release procedures for instance, such as the ConWIP procedure, can be applied to keep the WIP and thus the throughput times at a constant level [15].
3. Analysis approach

The general driver tree helps structuring throughput time analyses as well as focusing on the most relevant parameters and data required to identify the main levers for throughput time reduction. The comprehensive analysis procedure consists of four steps as illustrated in Figure 4. In a first step, a high level throughput analysis is conducted based on order generation, order release, and order completion dates to identify systematic planning errors and to initially localize problem areas. Furthermore, the system state is examined mainly using throughput diagrams.

The second analysis step focuses the production structure to account for the upper branch of the driver tree. A main task here is to map the actual order throughput for the most relevant product families, which can oftentimes be obtained from operation event logs and simple predecessor-successor analyses. From there, for instance, information about the amount of operations per PO can be obtained which indicates if reduction of process steps offers significant throughput time potentials. This could also be the starting point for additional analyses (value stream analysis, variant trees, etc.) not solely using feedback data in order to create more detailed insights.

After having evaluated the general production structure, data analyses on work system level are performed in step three mostly following the general steps of a bottleneck-oriented logistics analysis (see [11]) in order to identify throughput time potentials for single process steps. The general ideal is that the distribution of order throughput times usually follows the Pareto principle, meaning that a limited number of work systems are responsible for the most significant percentage of the total order throughput time. Consequently, in a first
step, analysis focuses on the throughput time determining bottlenecks. A very helpful tool to identify the most critical work systems and the main throughput time drivers is the logistical resource portfolio [16].

A detailed process and operation analysis follows in step 4 in order to detect the real root-causes for high operation or inter-operation times. From a logistical point of view, this primarily comprises identification of potentials for $i\text{WIP}_{\text{min}}$ reduction to decrease inter-operation times applying production operating curves. However, further company and process specific data and KPIs can be analyzed to examine process stability or quality for instance decreasing the maximum output rate. The developed driver tree helps deciding which further influencing factors might be of high relevance and should therefore be evaluated. Once, the main throughput time drivers have been identified, applicable measures can be derived. The general procedure is demonstrated in an industrial case study after briefly deriving data requirements in the following.

![Figure 4: Throughput time analysis procedure](image)

### 3.1 Data requirements

Knowledge about which relations to analyze and which analyses to perform significantly reduces analysis and interpretation effort as most analysis steps can be conducted using only a limited amount of data. Contrary to the trend towards big data analyses, only well-founded correlations need to be examined already allowing a valid delimitation of the essential potentials. Table 1 provides an overview of the minimum data requirements for each analysis step. It turns out that only a relatively small amount of data is required that can usually be easily retrieved from ERP systems. This data basis should be the starting point for each throughput time analysis. Nevertheless, use-case specific extensions of the data basis by adding further process or product parameters (e.g. product or order characteristics) is possible and may provide additional insights through correlation analyses or other analysis methods. As statistical analyses are performed based on mean values and standard deviations, it is crucial that the applied data basis represents a typical period of operation and includes a sufficient amount of observations in order to derive statistically valid conclusions. Furthermore, validity checks are required as otherwise implausible or faulty feedback data might be evaluated. Therefore, process experts should always be consulted to discuss available data.
Table 1: Minimum data-requirements for each analysis step

<table>
<thead>
<tr>
<th>Data requirements</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>order no.</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>date order generation</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>date order release</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>date order completion</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>order processing time</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>date order procurement</td>
<td>x</td>
<td></td>
<td></td>
<td>(MtO only)</td>
</tr>
<tr>
<td>delivery date procurement order</td>
<td>x</td>
<td></td>
<td></td>
<td>(MtO only)</td>
</tr>
<tr>
<td>delivery date to customer</td>
<td>x</td>
<td></td>
<td></td>
<td>(MtO only)</td>
</tr>
<tr>
<td>operation no.</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>work system no.</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>start of operation</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>end of operation</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>setup time</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>no. parallel work stations</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>capacity / max. output rate</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>further order / product data</td>
<td></td>
<td></td>
<td></td>
<td>use-case specific</td>
</tr>
</tbody>
</table>

3.2 Analysis procedure – Case study

In the following, the general analysis procedure is demonstrated based on an industrial case study conducted at a mechanical engineering company producing customer specific machines following an MtO-strategy. Data of a representative six month period have been analyzed comprising 3,114 POs. The overall throughput analysis revealed that a large amount of the delivery time is caused by a certain manufacturing department (mean order throughput time: 25 SCD) as to why this production department has been chosen for further throughput time analysis. Ready to sell parts as well as components for subsequent assembly stages are produced in a shop production. An analysis of the order throughput revealed that the manufacturing stage is operating in a steady state and the demand rate and output rate are harmonized (see Figure 5). However, the mean WIP is rather high and in average the proportion of operation time accounts for only about 6% of the total order throughput time. This indicates that WIP reduction offers the highest potential to reduce throughput times. In average, the planned throughput time further match the actual throughput time and no systematic backlog occurs.

![Figure 5: Throughput diagram of the manufacturing stage under examination](image)

In step 2 the process structure was examined without questioning product design. The average number of operations per order was seven, also including storage and retrieval as well as transport operations. Focusing on the relevant value-adding operations, the number of operations decreases to approximately two per order. Hence, further analyses addressed the single work systems (WS) within the manufacturing department. In a first step, the logistical resource portfolio has been developed as illustrated in Figure 6. The portfolio indicates that WS 1 accounts for about 12% of the total throughput time within the examined period. As WS
1 further shows a very high relative WIP of about 1,050% it is very likely to operate in an overload state and to offer significant WIP reduction potential. For WS 2 in contrast, also showing a significant share of throughput time, the relative WIP equals only about 350% and throughput times probably cannot be decreased by reducing the WIP level without having to expect significant utilization losses.

In step 4 WS 1 and WS 2 have been analyzed in detail to identify appropriate levers for throughput time reduction using the theory of the production operating curves (see Figure 7). For WS 1 the detailed analysis revealed that a reduction of the relative WIP to about 500% would still allow a mean capacity utilization of more than 99% while reducing the mean throughput time from 10 SCD to 5 SCD. Further WIP reduction (and thus reduction of the planned throughput time) would theoretically be possible, but is not considered reasonable as relatively high setup time shares require optimization of setup processes parallel to main production time. For WS 2 a WIP reduction to 250% would result in a reduction of the mean throughput time from almost 8 SCD to 5 SCD. However, losses of capacity utilization of almost 2% would be expected. Additionally, WS 2 also requires setup optimization and thus a steady amount of waiting orders which is why a rather high WIP level is aspired. As minimum transition times are negligible for WS 2, according to the driver tree, another influencing factor causing a high WIP level and thus high throughput times is the order workload and the variance of the order workload in particular. Analysis of the order workload for WS 2 showed a relatively high variance with a variant coefficient of 1.12 resulting in an iWIP\textsubscript{min} of 42h. If it is possible to reduce the variance coefficient of the order workload (e.g. by splitting large order lot sizes) to only 0.9% the mean throughput time could be reduced to 5 SCD without any losses of capacity utilization. For further information regarding the underlying calculation rules and basic principles see [11]. Having a look at the driver tree, throughput times could also be decreased if the maximum output rate could be increased. Yet for both, WS 1 and WS 2 no significant losses of the maximum output rate due to unstable processes could be identified. Hence, the output rate could only be increased by increasing the capacity by investing in new machines, which was not considered feasible.
Concluding, as WS 1 and WS 2 account for about 22% of the total throughput time in the production department, implementation of only the described two measures (WIP reduction at WS 1 and harmonization of the order workload at WS 2) would already result in a reduction of mean order throughput time of about 10%. The logistical resource portfolio indicates further critical work systems (WS 4, WS 3, and WS 8) that should be analyzed similarly. This allows a structured analysis procedure to identify potential levers for significant throughput time reduction only using the required data pointed out in Table 1. For more complex processes and identification of sources for failures for instance, further analysis of process and sensor data could create additional insights.

4. Conclusion

In this paper a systematic approach for data-based throughput time analysis has been presented. Based on well-established logistic models and generally valid cause-effect relations, a throughput time driver tree has been derived which helps companies focusing on the most relevant influencing factors on throughput times and sets the guidelines for the analysis procedure. It can be seen as an extension and systematization of the bottleneck-oriented logistics analysis enabling non-experts conducting the most relevant analyses own their own and adding further meaningful data available in their company. It has further been demonstrated that the most significant throughput time potentials can already be detected using only a limited amount of data. Similarly, detailed analysis procedures for the most relevant logistics KPI have been developed based on generally valid cause-effect relationships, which set the basis for the development of effective production controlling systems ideally supporting in identifying logistic weak points (find detailed analysis guidelines at www.quantilope-ifas.de). Increasing digitalization and available sensor data can create additional insights if used for further, use-case specific analyses. Therefore future research could focus on combining the generally valid analyses applying logistic models with further data-mining and AI approaches in order to include company specifics and previously unknown relations in the analysis of logistics performance.

References


Biography

**Lasse Härtel, M.Sc.** (*1988*) studied industrial engineering at RWTH Aachen University and has been working as a research associate at the production management department of the Institute of Production Systems and Logistics (IFA) since 2016.

**Prof. Dr.-Ing. habil. Peter Nyhuis** (*1957*) studied mechanical engineering at the Leibniz University Hanover and subsequently worked as a research associate at the Institute of Production Systems and Logistics (IFA). After obtaining his Dr.-Ing. doctorate, he was habilitated before working as an executive in the field of supply chain management in the electronics and mechanical engineering industry. He has been head of IFA since 2003.
Sizing electric storage system for atypical grid usage of industrial consumers

Fabian Zimmermann\textsuperscript{1}, Annika Wurster\textsuperscript{2}, Alexander Sauer\textsuperscript{1}

\textsuperscript{1}University of Stuttgart, Institute of Energy Efficiency in Production EEP, Nobelstraße 12, 70569 Stuttgart
\textsuperscript{2}University of Applied Sciences Mannheim, Paul-Wittsack-Straße 10, 68163 Mannheim

Abstract

There are many applications for electric storage systems (ESS) in manufacturing systems. While applications for maintaining production in case of a blackout are already established and economical, applications for optimizing energy supply are becoming increasingly interesting for manufacturing companies. Atypical grid usage is one application for optimizing the energy supply which has the potential to reduce the grid fee of industrial consumers. The grid fee for industrial consumers depends on the characteristics of the energy consumption. The smoother the power is drawn from the grid, the less grid fee has to be paid. This goal can be achieved by integrating an electric storage system. Electric storage systems offer high power and capacity, making them the ideal solution for this application. The challenge is the sizing of the electric storage system and the resulting economic efficiency. In this article a sizing methodology for electric storage systems, aiming for atypical grid usage, is presented.

Keywords

Energy storage; energy flexibility; sizing methodology; atypical grid usage

1. Introduction

Electric storage systems (ESS) offer a wide range of applications within industrial companies. ESS have been established to ensure an uninterruptible power supply (USP) \cite{1}. Overall, the applications for a short- to medium-term storage period can be categorized as shown in Table 1.

Table 1: Applications for ESS in manufacturing companies according to entrepreneurial benefit \cite{2}.

<table>
<thead>
<tr>
<th>Maintaining production</th>
<th>Optimization of energy supply</th>
<th>Provision of system services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security of supply</td>
<td>Self-consumption optimization</td>
<td></td>
</tr>
<tr>
<td>Quality of supply</td>
<td>Recuperation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trading on power exchange market</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grid fee reduction</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Switchable loads</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Provision of balancing energy</td>
<td></td>
</tr>
</tbody>
</table>

The applications for maintaining production include security of supply as well as the maintenance of supply quality. ESS are available on the market for these applications and already in use \cite{1}. The overriding benefit of these applications is the avoidance of production disruptions or rejects caused by voltage fluctuations or blackouts and the resulting quality deviations of the product \cite{2}.
The applications in the category optimization of energy supply are becoming more important for companies due to the required increase in energy efficiency [2]. These applications include optimization of self-consumption, recuperation, peak load reduction in order to reduce grid fees and load shifting through trading on the power exchange market.

The overriding benefit of these applications is to reduce energy costs either by reducing energy consumption (increasing energy efficiency) or by adjusting power consumption in order to reduce grid fees [2].

In the third category of applications, the provision of system services, ESS are used to serve the grid. This means, a ESS is used in such a way that it contributes to the stabilization of the primary energy system [3]. The provision and actual retrieval of capacities is remunerated differently by the associated transmission system operator depending on the product. The system services include the provision of control energy, divided into primary control, secondary control and tertiary control, as well as the switchable loads. The overriding benefit of these applications is to generate revenue by providing system services. Another application in this category would be the provision of the current reserve, but there is currently no payment for this application [3].

According to a survey conducted by Zimmermann et al., manufacturing companies see a high potential in the applications of the second category optimization of energy supply [2]. 17 % of survey participants consider the grid fee reduction to be one of the most important applications. In the future, this potential will increase continuously, as average grid fees for industrial consumers in Germany rose by around 6.5 % per year between 2011 and 2018 [4]. According to Consentec and Fraunhofer ISI, the future grid fees for industrial customers are expected to rise by up to 71 % till 2030 [5]. Therefore, this paper describes the atypical grid usage as a part of grid fee reduction and shows a method for sizing an ESS.

2. Atypical grid usage

The goal of atypical grid usage is to reduce the grid fee ($GF$). In Germany, the framework for this is regulated by the electricity grid fees ordinance ($StromNEV$) [6] and the energy industry act ($EnWG$) [7]. The grid fee

$$GF = DR \cdot P_{con,max} + ER \cdot E_{con}$$  \hspace{1cm} (1)

is calculated using the maximum annual peak load $P_{con,max}$, the corresponding demand rate ($DR$) and the annual energy consumption $E_{con}$ with the corresponding energy rate ($ER$). Atypical grid usage is not the only option to reduce the grid fee. According to Rothacher et al., there are four options [8]:

- Reduction of the annual peak load
- Change of the annual utilization hours
- Atypical grid usage
- Power-intensive final consumer

Due to the mutual influence of atypical grid usage and the other three options, only atypical grid usage is considered in this paper. Atypical grid usage means that less energy is consumed when all other consumers require a lot of energy from the grid [9]. This time period is called peak load time window ($PLTW$). All grid operators annually publish the $PLTW$. An ESS offers the possibility of pushing the peak loads out of the $PLTW$ [8]. However, two requirements must be fulfilled for this. On the one hand, a specific load transfer potential ($LTP$) and on the other hand, the materiality threshold ($MT$) have to be satisfied. For the calculation two peak loads are used, the maximal annual peak load $P_{con,max}$ and the maximum peak load in $PLTW$ $P_{PLTW,max}$. The $LTP$ must be at least 100 kW and is calculated by the difference between $P_{con,max}$ and $P_{PLTW,max}$ as described in formula 2 and 3 [8].

$$LTP = P_{con,max} - P_{PLTW,max}$$  \hspace{1cm} (2)
The MT is calculated by the ratio of \( LTP \) and \( P_{con,max} \) (see formula 4) \[8\]. For industrial consumers at low-and medium-voltage level a MT of 30% must be achieved (see formula 5) \[10\].

\[
MT = \frac{LTP}{P_{con,max}}
\]

\[
MT \geq 30\%
\]

If these requirements are met, the calculation of \( GF \) is not based on the \( P_{con,max} \), but on the \( P_{PLTW,max} \) (see formula 6) \[10\].

\[
GF = DR \cdot P_{PLTW,max} + ER \cdot E_{con,new}
\]

Since the \( GF \) can be reduced to a maximum of 20% of the original \( GF \), according to §19 Para. 2 p. 1 StromNEV, a maximum cost reduction of 80% is possible \[6\].

3. State of the art

In order to assess the utility of an ESS a load profile analysis of the consumer is required \[8, 11\]. The load profile, which represents the power consumption of a consumer over a year in 15 minutes average values, can either be recorded by measuring the power or can be estimated using a standard load profile \[8\]. The individual measurement of the load profile is required for consumers with an energy consumption over 100 MWh according to § 20 StromNEV \[6\]. Depending on the application, static, dynamic or optimization models are used for sizing an ESS \[12\]. Optimization models for different applications have already been implemented in \[12–16\]. In these optimization models, a business key figure is always defined as target function. In \[12, 17\] revenue maximization is used. The net present value (NPV) is evaluated in \[18\] and the biggest cost savings in \[19\]. General optimization models are suitable for applications that have a business focus \[12\]. This also includes the reduction of grid fees through atypical grid usage. This paper describes an optimization approach using the \( NPV \) as target function for sizing the ESS for atypical grid usage with a charging and discharging strategy.

4. Sizing an ESS for atypical grid usage

The proposed sizing method provides the optimized rated capacity and power of an ESS for the application of grid fee reduction. The method consists of four steps as shown in Figure 1.

![Figure 1: Four steps of sizing methodology](image)

The four steps are explained in the following sections.

4.1 Load profile analysis

The aim of the load profile analysis is to identify key figures from the load profile \( P_{con} \) that are relevant for the application described above and for sizing the ESS. The annual peak load
is the maximum of the 15 minutes average values for one year. An average power
\[ P_{\text{con.avg}} = \frac{1}{35040} \sum_{k=1}^{35040} P_{\text{con}}(k) \]  
(8)
can also be calculated from the 15 minutes average values of one year. The integral of the load profile shows the annual energy consumption
\[ E_{\text{con}} = \int_{k=1}^{35040} P_{\text{con}}(k) \]  
(9)
\( P_{\text{con,max}} \) and \( E_{\text{con}} \) can be used to calculate the original \( GF \) (see formula 1). For atypical grid usage the maximum peak load in the time period of \( PLTW \)
\[ P_{PLTW,\text{max}} := \max\{P_{PLTW}(k)\} \text{ for } k \in PLTW \]  
(10)
must be identified. By using this key figures in the next step the required power of the \( ESS \) can be limited.

### 4.2 Determination of required power of the \( ESS \)

The identified key figures from chapter 4.1 can be used to check the requirements \( LTP \) and \( MT \) for atypical grid usage. If these are not fulfilled, \( P_{PLTW,\text{max}} \) is reduced step by step (see formula 11).
\[ P_{PLTW,\text{max},ESS,x} = P_{PLTW,\text{max}} - x \text{ for } x = [0, P_{\text{con.avg}}] \]  
(11)
For each \( P_{PLTW,\text{max},ESS,x} \) an \( ESS \) is sized and economically valued. First, the requirements are checked (see formula 2 to 5). If the requirements are not yet fulfilled, the new \( GF_{ESS,x} \) is calculated as shown in formula 1, otherwise \( GF_{ESS,x} \) is calculated as shown in formula 6. For the calculation of the \( ESS \) power \( P_{ESS,x} \), the following applies for the new \( P_{PLTW,\text{max},ESS,x} \) (see formula 12):
\[ P_{ESS,x} = P_{PLTW,\text{max}} - P_{PLTW,\text{max},ESS,x} \]  
(12)
In formula 13 the required power \( P_{ESS,x} \) is extended by the efficiency factor for the interfacing AC converter \( \eta_{AC} \) and DC rectifier \( \eta_{DC} \) [1].
\[ P_{ESS,\text{real},x} = \frac{P_{ESS,x}}{\eta_{AC} \cdot \eta_{DC}} \]  
(13)
The following third step for the determination of the \( ESS \) capacity is performed for each \( P_{ESS,\text{real},x} \).

### 4.3 Determination of the required capacity of the \( ESS \)

A requirement for determining the capacity is that the \( ESS \) has full forecasting capability. First, the energy demand per time step
\[ \Delta E_x(k) = (P_{\text{con}}(k) - P_{PLTW,\text{max},ESS,x}) \cdot 0.25h \text{ for } k \in PLTW \]  
(14)
is calculated. \( \Delta E_x \) is negative if the energy storage can be charged. The following applies for the discharging capacity \( E_{ESS,dch,x} \). If energy is required at the next time step \((k+1)\), meaning \( \Delta E_x(k+1) > 0 \), \( E_{ESS,dch,x} \) is increased by \( |\Delta E_x(k+1)| \) plus the efficiency factors \( (\eta_{AC}, \eta_{DC}, \eta_{ESS}) \) at time step \( k \) (see formula 15). The
efficiency factor $\eta_{\text{ESS}}$ depends on the storage technology. $E_{\text{ESS},dch,x}$ remains the same if there is no energy demand.

$$E_{\text{ESS},dch,x}(k) = \frac{|\Delta E_x(k + 1)|}{\eta_{AC} \cdot \eta_{DC} \cdot \eta_{\text{ESS}}}$$  \hspace{1cm} (15)

In this paper, the charging strategy "charge as much energy as necessary as late as possible" according to Kaschub is used [20]. For this purpose it is checked, whether energy is needed (see formula 16).

$$E_{\text{ESS},dch,x}(k + 1) > 0$$  \hspace{1cm} (16)

In the next step the charging energy $E_{\text{ESS},ch,x}$ must be defined. If the maximum possible charging energy within a quarter hour ($P_{\text{ESS,real},x} \cdot 0.25h$) is less than the energy demand $\Delta E_x(k + 1)$ reduced by the efficiency factors of the AC converter and DC rectifier, the maximum possible charging energy in a quarter hour limits the $E_{\text{ESS},ch,x}$ (see formula 17 and 18).

$$P_{\text{ESS,real},x} \cdot 0.25h < |(\Delta E_x(k + 1) \cdot \eta_{AC} \cdot \eta_{DC})|$$  \hspace{1cm} (17)

$$E_{\text{ESS},ch,x}(k) = P_{\text{ESS,real},x} \cdot 0.25h$$  \hspace{1cm} (18)

If this is not the case, the charging energy $E_{\text{ESS},ch,x}$ is limited by $\Delta E_x$ itself. Then $E_{\text{ESS},ch,x}$ is calculated as shown in formula 19.

$$E_{\text{ESS},ch,x}(k) = |\Delta E(k + 1) \cdot \eta_{AC} \cdot \eta_{DC}|$$  \hspace{1cm} (19)

Subsequently, it is checked, whether the charging energy $E_{\text{ESS},ch,x}$ is sufficient for $E_{\text{ESS},dch,x}$. If not, $E_{\text{ESS},ch,x}(k - y)$ for $y = [1,35040]$ is accumulated, with the steps of formula 14 to 19, until the discharging energy $E_{\text{ESS},dch,x}(k + 1)$ is covered. The required capacity $E_{\text{ESS},max,x}$ is determined by using the maximum

$$E_{\text{ESS},max,x} := \max\{|E_{\text{ESS},dch,x}(k)|, |E_{\text{ESS},ch,x}(k)|\}$$  \hspace{1cm} (20)

This capacity $E_{\text{ESS},max,x}$ is increased according to Köhler et al. by the depth of discharge DOD which is different for each storage technology, ageing surcharges $\eta_{EOL}$ and a general reserve capacity $\eta_{\text{res}}$ (see formula 21) [1].

$$E_{\text{ESS,real},x} = \frac{E_{\text{ESS},max,x}}{\frac{1}{(1000)} + \eta_{EOL} + \eta_{\text{res}}}$$  \hspace{1cm} (21)

With the completion of this step, for each possible $P_{\text{ESS,real},x}$ an associated $E_{\text{ESS,real},x}$ is identified. The optimum ESS size can be identified on the basis of the calculated economic efficiency.

### 4.4 Economic evaluation of the ESS

An economic evaluation is performed by calculating the NPV. Several input parameters are required to calculate these two key figures. The lifetime of the ESS depends on the number of full cycles. For this reason, according to Fuchs et al. equivalent full cycles per year

$$FC_x = \frac{\sum_{k=1}^{35,040} |E_{\text{ESS},dch,x}(k)| - \sum_{k=1}^{35,040} |E_{\text{ESS},ch,x}(k)|}{2 \cdot E_{\text{ESS,real},x}}$$  \hspace{1cm} (22)

are calculated [21]. Since each ESS, depending on the technology, has both a calendar $T_{cal}$ and cyclic lifetime $T_{cyc}$, it must be determined which one is reached first. If the quotient of $T_{cyc}$ and $FC_x$ is greater than or equal to $T_{cal}$, then $T_{cal}$ is the lifetime $t_{\text{ESS}}$ of the ESS, otherwise $T_{cyc}$ is equal to $t_{\text{ESS}}$. The lifetime $t_{\text{ESS}}$ is also the planning horizon for the economic evaluation. The new grid fee $GF_{\text{ESS,x}}$ per year is equally relevant for the
economic evaluation. This depends on the new \( P_{\text{PLTW,max,ESS,x}} \) as shown in formula 11. Compared to the original load profile \( P_{\text{con}} \), the new load profile

\[
P_{\text{con,new,x}}(k) = P_{\text{con}}(k) + \frac{(E_{\text{ESS,dch},x}(k) - E_{\text{ESS,dch},x}(k - 1)) \cdot \eta_{\text{AC}} \cdot \eta_{\text{DC}} \cdot \eta_{\text{ESS}}}{0.25h}
\]

(23)

\[
P_{\text{con,new,x}}(k) = P_{\text{con}}(k) + \frac{(E_{\text{ESS,ch},x}(k) - E_{\text{ESS,ch},x}(k - 1)) / (\eta_{\text{AC}} \cdot \eta_{\text{DC}})}{0.25h}
\]

(24)
decreases for discharging processes (see formula 23) and increases for charging processes (see formula 24). This means that \( E_{\text{con,new,x}} \) can also be calculated using formula 9. These parameters can be used to calculate \( GF_{\text{ESS,x}} \) per year (see formula 4). The revenues

\[
R_x = GF - GF_{\text{ESS,x}}
\]

(25)

per year are the savings from the difference between \( GF \) and \( GF_{\text{ESS,x}} \), including an annual increase of grid fee (see formula 25) [22]. The investment costs \( C_0 \) for the ESS depend on \( E_{\text{ESS,real,x}} \) and \( P_{\text{ESS,real,x}} \) [23, 24]. \( C_{0,x} \) is calculated as shown in formula 26. Specific investment costs \( c_p \) and \( c_E \) for power and capacity vary according to the storage technology.

\[
C_{0,x} = P_{\text{ESS,real,x}} \cdot c_p + E_{\text{ESS,real,x}} \cdot c_E
\]

(27)
The payments

\[
A_x = C_{0,x} \cdot C_B + (E_{\text{con,new,x}} - E_{\text{con}}) \cdot C_S
\]

(28)
consist of the operating costs \( C_B \) per year [24] and additional electricity costs \( C_S \) including an annual increase [25] (see formula 28). Using \( R_x \), \( C_{0,x} \) and \( A_x \), the \( NPV \)

\[
NPV_x = -C_{0,x} + \sum_{t=0}^{t_{\text{ESS,x}}} \frac{-A_x(t) + R_x(t)}{(1 + i)^t}
\]

(29)
can be calculated (see formula 29). For discounting an interest rate \( i \) is used. \( NPV \) can decide, whether the investment is economically viable. If the \( NPV \) is positive, the ESS is economical. This step is performed for each possible \( P_{\text{ESS,real,x}} \) and associated \( E_{\text{ESS,real,x}} \). The optimal ESS size can be identified from the maximum \( NPV \).

5. Case study

For validation, the load profile of an automobile plant is analyzed and \( DR \), \( ER \) and \( PLTW \) from Stuttgart Netze are used to calculate \( GF \) and \( GF_{\text{ESS,x}} \) [26]. An annual increase of 4% after BNetzA is chosen for \( GF \) [22]. \( C_S \) is assumed to be 0.1844 €/kWh with an annual increase of 3.3% [25]. \( C_B \) per year is 3% of \( C_0 \) according to Sterner and Stadler [24]. The efficiency factors of the power electronic devices \( \eta_{\text{AC}} \) and \( \eta_{\text{DC}} \) are fixed at 95% [1]. For \( \eta_{\text{EOL}} \) 20% and for \( \eta_{\text{res}} \) 10% of the required capacity are assumed [1]. An interest rate \( i \) of 3% is set for the calculation of the \( NPV \) [1]. For the analysis three storage technologies are considered: a lead-acid battery (LAB), a lithium-ion battery (LIB) and a redox-flow battery (RFB). The key figures taken into account are shown in Table 2.
If the requirements are not met, the analysis shows that none of the considered storage technologies can achieve a positive NPV. Figure 2 shows the NPV via $P_{\text{ESS,real,Lx}}$ from the time when the requirements are fulfilled.

As soon as the requirements are fulfilled, only LIB achieves a positive NPV for the analyzed load profile. No economic result can be achieved for LAB and RFB. For LIB the smallest size that fulfills the requirements can achieve a positive NPV. The technical and economical key figures of the most economical LIB are shown in Table 3.

### Table 3: Key figures for the most economical LIB.

<table>
<thead>
<tr>
<th>Key Figures</th>
<th>Unity</th>
<th>LIB</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{\text{ESS,real}}$</td>
<td>[kW]</td>
<td>7,357</td>
</tr>
<tr>
<td>$E_{\text{ESS,real}}$</td>
<td>[kWh]</td>
<td>30,014</td>
</tr>
<tr>
<td>$t_{\text{ESS}}$</td>
<td>[a]</td>
<td>15</td>
</tr>
<tr>
<td>$C_0$</td>
<td>[€]</td>
<td>14,387,994</td>
</tr>
<tr>
<td>NPV</td>
<td>[€]</td>
<td>916,506</td>
</tr>
<tr>
<td>Grid fee reduction</td>
<td>[%]</td>
<td>18,85</td>
</tr>
</tbody>
</table>

A positive NPV can be achieved over the lifetime of 15 years. However, $C_0$ are very high compared to the NPV. But a reduction of 18 % for $GF$ could be achieved. $E_{\text{ESS,real}}$ of LIB is large for the analyzed load profile, since $P_{\text{PLTW,max}}$ is drastically reduced to meet the requirements that the LIB has to discharge over the
6. Conclusion

It can be seen that the atypical grid usage has great economic potential. However, the economic efficiency of an ESS is strongly dependent on the individual load profile. LIB can be economically sized for the analyzed load profile. For the economic efficiency of LAB and RFB, other load profiles should be analyzed. The method could also be transferred to other applications and thus offers the possibility to combine applications, since a multifunctional storage operation ensures a higher utilization of the ESS and an associated increase in revenue [12]. In addition, the interest rate $i$ has a significant impact on the NPV. The influence can be checked by a sensitivity analysis.

Acknowledgements

This work is funded by the German Federal Ministry of Education and Research (BMBF) within the Research Program “Kopernikus-Projekte für die Energiewende” within “Synchronisierte und energieadaptive Produktionstechnik zur flexiblen Ausrichtung von Industrieprozessen auf eine fluktuierende Energieversorgung (SynErgie)” (fund number 03SFK3K1) and managed by the Projekträger Jülich (PTJ). The authors would like to thank the BMBF and PTJ for supporting and all partners of EEP within this project.

References


Biography

**Fabian Zimmermann** (*1991)* studied industrial engineering at the University of Applied Sciences Würzburg-Schweinfurt and at the University of Stuttgart. Since 2015, he has been working as a research associate at the Institute of Energy Efficiency in Production EEP of the University Stuttgart since 2015. The main topics of his research are applications of energy storages and the integration of them in production sites.

**Annika Wurster** (*1992)* is a student at the university of applied sciences in Mannheim. She studied industrial engineering in Mannheim with a bachelor’s and master’s degree. She wrote her thesis at the Fraunhofer IPA and the EEP of the University of Stuttgart.

**Alexander Sauer** (*1976)* is the Executive Director of the Institute for Energy Efficiency in Production (EEP) at the University of Stuttgart, as well as the Divisional Director of Resource-Efficient Manufacturing of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA. His research focus is on energy and resource efficiency in manufacturing. He is author of numerous publications and active in a number of initiatives and on jurys.
Activity-based shop floor management –
A concept to enhance flexibility

Barbara Tropschuh¹, Svenja Korder¹, Fabian Dillinger¹, Harald Bauer¹, Gunther Reinhart¹
¹Institute for Machine Tools and Industrial Management (iwb), Technical University of Munich, 85748 Garching b. Munich, Germany

Abstract
Volatile markets, an increasing shortage of skilled workers and individual customer requirements as well as the growing desire for a fulfilling work-life balance among employees are influencing the production environment. Flexibility and adaptability are possible key factors to enable companies to meet these challenges. A concept for an activity-based shop floor management has been developed to make work organization more flexible in the area of production.

In this concept, first of all activity packages are defined and evaluated, for example with regard to their requirements. Furthermore, it is necessary to define competence levels into which the employees are clustered according to their abilities. Thus, a competence-oriented matching of activity packages and employees is possible. The employee can choose tasks from the pre-grouped activity packages. The result is a generally valid concept for flexible staff deployment planning, which is evaluated in the automotive industry. This concept allows a change from role-based to activity-based task assignment with gamified incentive system, whereby specialists can be deployed more efficiently according to their qualifications.

Keywords
Flexibility; shop floor management; competence orientation; work organization

1. Introduction
In today's turbulent market environment, manufacturing companies and their employees face numerous challenges. Demographic development means that the total number of skilled people will decrease in the coming years [1]. In addition, the average age of the workforce in companies is rising significantly. This leads to a shortage of skilled labour as well as the diversification of competences and performance [2]. Therefore, the efficient use of existing skilled workers and a competence-based allocation of activities to the individual worker are becoming more and more important. Due to the increasing amount of individualized products in small batch sizes (mass customization and mass personalization) as well as shorter product and technology life cycles, workers are confronted with low repetition rates and frequent changes in the activity sequences [3]. As a result, the workplace gets more and more complex and the operator has to adapt quickly to these changing requirements. Conventional learning curve effects cannot be achieved, which are predominantly based on practice through repetition (as, for example, the experience curves of [4] show).

Hence, instead of long training methods shorter activity-oriented tutorials are considered that enable a new way of interacting on smart devices. As part of the digital transformation, smart technologies are being introduced into factories. This allows a flexible, personalized and automated communication with each worker [5]. In addition, playful incentive systems can be applied to increase the individual employee motivation as well as the group spirit. Digital technologies can also be used to create new opportunities for...
workforce support within value creation. At the same time, employees desire a well-adjusted work-life balance due to social changes [6]. Employees would like to have flexible working hours depending on their current life situation [6]. In order to remain attractive as a company, especially in the context of the increasing shortage of skilled workers, more flexible working time models must be taken into account. Figure 1 shows the main development drivers and their impact on manufacturing companies. Flexibility and digitalization are possible key factors to meet these challenges in order to remain competitive in the volatile market.

Figure 1: Development drivers and their impact on companies based on [1-6]

In this paper, a concept for an activity-based shop floor management with an integrated incentive system has been developed, to meet the listed challenges. As a result, the work organisation becomes more flexible and companies remain competitive on the market. Moreover, the concept can be combined with new working time models to give more flexibility to both the company and the workers.

The paper is structured as follows: Chapter two gives a short overview of the current state of research and the pursued research goal. Afterwards the concept of an activity-based shop floor management system is presented that can be extended by different topics as for example gamification (chapter three). One application of the approach is described in chapter four, followed by an interpretation and critical reflection (chapter five). Finally, section six summarizes the paper and outlines future research work.

2. Literature review

Within the existing literature on work scheduling, especially aspects on activity- and competence-based approaches have been analysed. The focus in the following analysis is the production area of manufacturing companies. The projects and approaches introduced hereafter are meeting the focus and given topic.

2.1 Existing approaches

In the KapaflexCy project [7], smartphones in combination with a cyber-physical system (CPS) are used as a central capacity provider in order to allow employees to coordinate upcoming work assignments on their own responsibility. The project aims on shortening the reaction time of companies in times of fluctuating customer demand and volatile markets on the one hand and wants to motivate employees through flexible and self-determined working hours on the other hand. Employees who work shifts receive deployment requests on their mobile devices and decide together and decentral whether they want to accept or reject requests. Requests for deployment are made according to priority rules. The required competencies are the top priority, followed by legal framework conditions. If these are met, the system takes into account the employees’ time accounts and, lastly, their personal preferences [8]. In this way, employees are actively involved in personnel requirements planning. The focus of this approach is on HR-flexibility for shift workers.

Another approach for competence-based personnel scheduling was published by Denkena et al. [9]. This involves collecting production data that can be used to draw conclusions about the performance and skills of employees. Production data could be e. g. the number of good parts, the cycle time or rework. The collected data is then analysed and interpreted. This information is used by an algorithm for personnel planning in order to further develop individual employee competences. In this way, competences can be developed ideally while utilizing production capacity at the same time [9].
Dollinger et al. [10] focus on the use of so-called ‘jumpers’ or ‘floaters’ in production. Among others, they want to achieve the development of competences. In order to deploy employees as jumpers in the line, the requirements at the individual workstations were first recorded. The requirements were then transferred to the required competences of the employees and stored in a workplace-competence matrix. Moreover, the existing competences of all employees working in this line were recorded and also systematized in a matrix. Three deployment scenarios for jumpers have been defined: A coaching scenario, an assistance scenario and a replacement scenario. Especially by using the coaching scenario, competences of workers can steadily be improved [10].

The approach of Korder et al. [11] deals with the deployment of employees in reconfigurable manufacturing systems (RMS). When RMS are used in production, employees may have different requirements for fulfilling a work task at a given workstation. In order to enable companies to quickly and easily check the correspondence between the requirements of the machines and the competences of the employees, the publication developed an approach for formalizing employee skills comparable to the formalization of machine skills. By comparing the two, competence gaps or deficits can be identified. These can be closed by targeted further personnel development measures [11].

In the article of Arena et al. [12], ontologies for human resources optimization are used. Ontologies are suitable for knowledge-intensive applications. Therefore, they were used in the mentioned approach to create information models about the shop floor. Subsequently, employee groups and their experiences (in three levels) were defined. Afterwards, the employee groups were assigned skills to differentiate the groups. Based on the requirements of tasks, the right person for the right job could be suggested in real time [12].

2.2 Research demand

The literature review shows that there are no concepts that deal with future challenges in work organization by combining a competence-based personnel planning with flexible working models and a gamification approach for employee satisfaction. Especially the efficient allocation of worker skills and task requirements to deal with the growing lack of skilled workers is hard to address. The extension of this organizational approach with flexible working time models, employee qualification and incentive systems also has a novelty value. With the combination of the mentioned topics, the challenges of both, company and employee, can be met. On the one hand, companies can cope with the lack of skilled workers, and fluctuations in customer demand can be met by higher flexibility in personnel planning. On the other hand, employees get more flexibility in working time and transparency with work tasks and achieved goals in individual qualification. Transparency, in particular, must be emphasized, as it is not yet possible for employees to compare their performance with that of their colleagues. This is especially relevant when employees have the feeling of an unfair distribution of tasks or strongly varying work performance. Therefore, the aim of this paper is to develop a universal approach where all described challenges can be met.

3. Activity-based shop floor management – approach of a new work organization

This chapter presents a holistic concept for activity-based shop floor management. The concept was developed for the production area. In the first step, the ideal match between skilled workers and given tasks should be enabled for more flexibility in distribution of tasks. This is the foundation for adding the matching with a flexible working time model. By embedding the results in a gamification approach, an incentive system for employees was created, which can also be completed with employee qualification in the last step.

3.1 Competence-based matching between tasks and workers

This section presents an instrument for matching activities with operational employees of production. For matching, activities and employees must be specified with selected attributes. For employees, the goal of
matching is to independently select activities according to their competencies and time availability from a dynamic and individual task pool. A generally valid procedure is required for attribute specification. A characteristics value is rated using a four-level scale ranging from not available (value 0) to low (value 1), medium (value 2), and high (value 3).

**Activity specification:**
First, the activities, which are derived from the production orders, and thus the work tasks must be classified. Therefore, each activity is characterized by six relevant characteristics: requirement degree, frequency, availability, planned activity duration, basic condition as well as priority (see Figure 2). The requirement degree represents the level of physical and psychological demand of an activity on the employee. The frequency is defined by the number of repetitions of the activity within a defined time period, whereas the availability describes how many employees in the respective company can perform the activity. The planned activity duration requires a time analysis for the corresponding activity and defines the planned processing time. The attribute basic conditions comprises topics from occupational health and safety such as safety instructions or basic training, which the employee must present for the maintenance of the system. It must be fulfilled for the execution of the activity. The priority is primarily determined by the time criticality, e.g. urgent parts have a high priority (value 3). In addition, the prioritization also takes into account whether an activity is connected to a bottleneck system or a particularly important customer order.

**Employee specification:**
Analogous to the activities, the employees are characterized with the characteristics degree of competence, availability, basic prerequisite as well as the selected working time model and the competence development goals. The degree of competence shows to which extent an employee fulfils the physical and psychological prerequisites for the execution of a required task. This makes it possible to compare each activity with a competence level, see Figure 2. The evaluation of the characteristics is carried out exclusively on the basis of available competences, to ensure that tasks are assigned with the help of the competences actually available and not theoretically certified. Availability and basic prerequisites are defined analogously to the activity attributes of the same name. The availability of the employees refers to the activities. The characteristics of employee availability are therefore defined for each task.

**Matching:**
The specification of employees and activities forms the basis for matching. For this purpose, activities are proposed to the employees whose degree of requirement is less than or equal to the employee's level of competence. In addition, the basic prerequisite of an activity is taken into account, which the employee must fulfill in order to be able to carry out the task. Due to the prioritisation of time-sensitive tasks, activities with high priority (3) must be processed with priority. As far as prioritisation permits, the planned duration of the activities is taken into account. In order to minimise the division of labour across shifts, the remaining working time of the worker is taken into account in the distribution of tasks. In addition, a points system is introduced for matching. Activities are provided with additional points (AP), whose requirement degree (RD) exceeds the competence degree (AD) of the worker. This serves to qualify employees in a targeted manner. Points can also be awarded for activities that require an extraordinarily high level of physical or mental effort (E) or have a special significance (SO) for the organizational unit or the company. These three criteria are evaluated analogously to the activity and employee attributes on a four-point scale, from no value (value 0) to high (value 3). The additional points are calculated as follows (1):

\[
AP = (RD - AD) + E + SO
\]

Using this logic, employees can earn a maximum of 9 additional points for an activity, which they can redeem profitably for themselves or their team (see Section 3.4). The point system therefore offers an additional incentive to carry out activities. The following figure (Figure 2) illustrates the matching between
activities and employees. All activities are specified and collected in the activity-pool. Only the tasks, which are relevant to the employee, are displayed. The additional points are calculated according to the employee’s qualifications. The employee then selects an activity from his or her individual and dynamic activity pool, which he or she then processes.

### 3.2 Flexible working time models

In order to be able to react as flexibly as possible to the rapidly changing fluctuations in orders, companies are very interested in scheduling their employees as needed. Furthermore, employees wish to have the opportunity to decide flexibly and autonomously on their current working hours [6]. In this case, flexible working time models would be advantageous for both sides [13]. Of course, the entire workforce does not prefer more flexibility, which is why individual agreements should be made and different working time models offered. A working time model is to be regarded as flexible if at least one design parameter of the working time can be changed. The working time parameters include the location, duration and distribution of working time [14]. In addition, the company already benefits from knowing the current personnel requirements as precisely as possible. Overstaffing can only be identified and efficiency can be improved if the personnel requirement is known. [15, 16]

As part of the concept, numerous working time models and flexibilisation measures were analysed. There are numerous adaptable and agile measures that are suitable for classic shift work, like job sharing, group jobs, flexitime in shift work or manual/automatic/app-based resource scheduling. In the use case of our approach, described in section 4, recommendations for action were generated and discussed with the company respecting flexitime in shift work, where the worker can begin and end his daily work outside core working hours, whereby only the annual number of working hours must be covered. In addition, the model of job sharing was discussed in which, for example, two persons share a workplace in a self-organised way.

### 3.3 Incentive system combined with a gamification approach

Motivated employees are more productive and more satisfied at work [17]. As a result, more and more companies would like to integrate incentive systems into their daily work to motivate each individual and the group. Incentive systems comprise all material and immaterial incentives to directly or indirectly influence the willingness of employees to perform or to strengthen desired behavioral patterns [18].

---

**Figure 2:** Overview of the competence-based matching between tasks and workers
incentives, such as bonuses or a salary increase, only motivate in the short term. Intrinsic incentives such as assuming responsibility have a more long-term effect and increase motivation and performance [20,19]. Wanting to encourage the individual motivation of the workers as well as the group-spirit while maintaining the same levels of hierarchy and responsibility, approaches from the gaming industry are transferable. Using game-typical design elements and processes in a non-game context is referred to as gamification [21]. The aim is to change the behavior and increase the motivation of the users. Typical game elements are e. g. descriptions of goals and participants, point or price systems and comparisons [22]. Typical game processes include the accomplishment of tasks through individual or collaborative performance. [23]

In order to find a suitable incentive system, bonus points are combined with selected gamification elements (see Figure 3). The result is a scoring system, where company-points can be awarded for task-specific work or trainings on the one hand, as for example a video about the handling of a machine (learning nugget). On the other hand, bonus-points can be gained for task-unspecific qualifications, e. g. back workout. In the chosen approach, the bonus-points are only remunerated in non-cash benefits. The approach supports the team spirit, because the workers are in no competition with each other and can also collect team-points. Within the framework of the activity-based shop floor management, the individual tasks are scored on the basis of their requirements for the employee. Three categories of the scoring can be distinguished: 1-5 (low requirement level), 6-10 (medium requirement level), 11-15 (high requirement level). In order to stimulate the daily performance of the employees, a certain number of points can be declared as standard performance, which each worker should fulfil during the working day. Of course, the qualified employees should not only complete the tasks with the most points, but should also work through activities from each category in order to retain their connection to the machine through cleaning work, for example, but also to prevent the creation of a class-divided workforce. Playful learning in an informal setting shows better results in terms of learning success, as learning through reward (bonus-pints) is linked to positive experiences [24]. Figure 3 shows one view of the implemented mockup with gamification elements: in the upper part the worker has his own profile where he has a permanent overview of his current skill and performance status. In the lower part, performance, points and detailed skills as well as qualification goals are listed.

3.4 Operational employee development

Qualification concepts previously used in production no longer meet today’s qualification requirements. Through shorter product and production life cycles as well as rapidly changing technologies, employees must constantly acquire new knowledge [25]. Previous measures for the qualification of employees in manufacturing companies were mostly carried out off-the-job, e. g. in learning facilities away from the workplace [26]. There, employees are taught defined lessons on a specific topic within a few days. However,
it cannot be predicted with certainty that employees will need the knowledge they acquire for fulfilling the work task. One possibility to meet the changed requirements is the micro-learning approach of [27]. In this approach, employees are taught knowledge directly at the workplace by means of information-communication-technology (ICT) using short learning units. The on-the-job qualification can be carried out by the employee at the moment when the knowledge is needed. The learning strategy of companies thus changes from strategic to operative. In order to develop these short learning units, so-called learning nuggets, the company must be willing to keep the capacity of an interdisciplinary team of employees, team leaders and human resources partially free. The technical infrastructure for the presentation of video or text-based units for example, must be created. Last but not least, the corporate culture must be given to accept forms of learning close to the workplace and the employees must have the willingness to learn in a self-organised way [27]. The approach described in this publication can be combined with micro-learning to build the skills employees need. Thus, the range of tasks can be extended step by step. For example, individual short video sequences could be recorded in which longstanding, experienced employees explain facts or technologies. These contents are necessary for further development to the next higher competence level.

3.5 Roadmap for the implementation

With five consecutive stages, companies can apply the developed concept. Figure 4 shows the roadmap of the concept schematically: according to the classification of the work tasks and the employees (step 1), a competence-based matching of activities with their requirements to employees with their competences takes place within the scope of the activity-based shop floor management (step 2). On the basis of this agile organization of tasks, it is possible to combine it to flexible working time models (step 3). Furthermore the integration of individual, gamified incentive systems follows, which integrate a point system into the workflow (step 4). At the last step, a new learning approach is implemented, which enables operational further training through small learning units (learning nuggets).

4. Application of the approach

The described approach was applied in a manufacturing company in the automotive industry. A pilot production area with about 15 employees was selected for implementation. In a first step, all work tasks were categorised and evaluated according to their requirements and rated by the four levels described in chapter 3. Comparable with qualification matrices already used in many companies, the existing competencies of each employee in the pilot area were recorded and also grouped according to the four levels. Subsequently, the work organization logic underlying the approach was implemented in a software tool. On a smartwatch, the employees were able to control the assignment of tasks in real time. A smartphone as well as PC- or tablet-based information terminals in production are also possible. In particular in context of usability for older employees a tablet at the workstation would be recommended. The implementation of the approach in the pilot area showed several advantages. Time savings has to be mentioned first, as the employees get real-time information about task status and unscheduled tasks with high priority. Also the status of machines could be considered with the software tool, so that the employees don’t have to check the status of machines
regularly. By avoiding long ways on the shop floor the employees can use their time more productively. Furthermore, the decision support by ranking the single work tasks or describing them by given requirements was stated as an advantage. This also relieved the workload of the team leader so that he could increasingly take over the coaching and individual development of the employees. All in all, positive feedback from employees was recorded.

5. Interpretation and critical reflection

The presentation of the concept and the application of the approach is followed by a critical reflection of the success factors and the effort involved in the implementation. The success of an activity-based shop floor management system with an integrated incentive system depends essentially on the acceptance of the employees and their integration in the development and implementation phase. During the conception phase, numerous discussions were held with the employees on the shop floor as well as from related areas (e.g. work preparation, human resources, and works council) in order to allow early integration of the employees. The employees emphasized positively that the activities on the company side and the needs and wishes of the workers had been equally incorporated into the concept. Of course, the high initial effort for the implementation of such a system must not be disregarded: For the formation of activity groups all tasks must be considered in detail, whereby an analysis of the job descriptions is not sufficient. In addition, a classification of the products according to their degree of difficulty as well as the allocation of production times for an optimal production planning are required. This additional effort is incurred in the area of work preparation where the processing time of the activities is already estimated. The actual additional expenditure would thus be reduced to the evaluation of the difficulty. However, it should be noted that the difficulty of the task usually correlates with the processing time, therefore only a separate consideration of the work requirements is added. Finally, the employee's commitment to the machine changed drastically and must also be taken into account. Prior to the approach, one employee was responsible for one or two machines, supervised them completely and was responsible for their operation. Now, one employee is currently performing activities on all machines in the field of activity and has to adapt to the new system of general group responsibility for all machines.

6. Summary and outlook

The research demand and the literature review demonstrate that a flexible concept for an activity-based shop floor management, which realizes a competence-based assignment of activities to workers, needs to be developed. In this paper a concept was presented which supports the successful handling of future challenges in the production environment. In view of the increasing shortage of skilled workers, the focus lies in particular on the more efficient use of skilled workers. In addition, work organization is adapted to the changed environmental conditions and significantly more transparency is created for the company as well as the employees. The implementation takes place on the basis of a five-stage roadmap to an activity-based shop floor management with a gamified incentive system.

Since the implementation has only just begun, the concept needs to be tested for a longer period of time. In the context of a longer experimenting period, the satisfaction of the workers and the long-term increase in productivity has to be evaluated. In addition, long-term data evaluation can identify bottlenecks in the execution of activities, from which qualification demands can be derived. The presented concept was designed for the shop floor, but can transferred to other areas. The conceptual approach is even applicable in indirect areas. This work setting also necessitates an assignment of activities with different requirements and employees with different competencies. The applicability of the concept in this new use case should be evaluated in another test phase. Since the approach described contains the scoring of individual activities as well as the variable processing of these in flexible working time models, a performance-related remuneration system can be integrated in a further step.
Acknowledgements

We want to thank the German Research Foundation (DFG) for funding this work as part of the collaborative research centre SFB 768 “Managing cycles in innovation processes: Integrated development of product-service-systems based on technical products”.

References


Biography

Barbara Tropschuh (*1994) is currently a PhD candidate at the Institute for Machine Tools and Industrial Management (iwb) at the Technical University of Munich (TUM). She received her master’s degree in Mechanical Engineering from TUM. Her research activities focus mainly on competence- and strain-oriented workforce scheduling.

Svenja Korder (*1990) is currently a PhD candidate at the Institute for Machine Tools and Industrial Management (iwb) at Technical University of Munich (TUM). She received her master’s degree in Human Manufactures Engineering from TUM. Her research focuses mainly on improving satisfaction of employees.

Fabian Dillinger (*1991) is currently a PhD candidate at the Institute for Machine Tools and Industrial Management (iwb) at TUM. He received his master’s degree in Industrial Engineering from the Technical University of Applied Science Rosenheim. His research focuses mainly on human-centered Lean 4.0 Management.

Harald Bauer (*1994) is currently a PhD candidate at the Institute for Machine Tools and Industrial Management (iwb) at TUM. He received his master’s degree in Production and Automation from the Munich University of Applied Sciences. His research focuses mainly on manufacturing change management and Lean Management.

Prof. Dr.-Ing. Gunther Reinhart (*1956) holds the chair of Industrial Management and Assembly Technology at the Institute for Machine Tools and Industrial Management (iwb) at Technical University of Munich since 1993. Since July 2016, Prof. Reinhart has been additionally the Managing Director of Fraunhofer Institute for Casting, Composite and Processing Technology (IGCV).
Creating a Worker-Individual Physical Ability Profile Using a Low-Cost Depth Camera

Justus Brosche¹, Hannes Wackerle², Hermann Lödding¹

¹Institute of Production Management & Technology, Hamburg University of Technology, Hamburg, Germany
²Institute for Biomechanics, BG Unfallklinik Murnau, Murnau, Germany and Paracelsus Medical University Salzburg, Salzburg, Austria

Abstract

Assembly workers suffer from long-term damage performing physically intensive tasks due to workstations that are not ergonomically designed for the individual’s needs. Current approaches towards ergonomic improvements of workstations only assess the workstations themselves without taking the individual worker and abilities into account. Therefore, physical limitations, such as age-related loss of range of motion, are not addressed. Work-induced long-term damages result in employee absences, especially of workers close to their pension. Regarding the demographic change, this issue will be even more prevalent in the future. The current approaches, like the functional capacity evaluation, allow movement analysis of individuals, but are too time-consuming to be performed on all workers of a production site. This paper presents a method to assess the individual ability of a worker using a low-cost depth camera with full body tracking to determine the angles between body segments. A set of ergonomic exercises is used to demonstrate relevant abilities for assembly and commissioning tasks. By capturing the motion sequence of these exercises, a physical ability profile can be created with little effort.

Keywords

Ergonomics; Motion Analysis; Intel® RealSense™; 3D Depth Camera

1. Introduction

Demographic change has a major impact on the working population: the average age of working people in developed countries is increasing while the dependency ratio (the ratio of people of working age to those of non-working age) decreases. With an aging population, companies face a more and more challenging age structure which makes the efficient use of the workers’ abilities more important than ever [1]. With repetitive motions and sometimes heavy work loads, assembly tasks promote musculoskeletal disorders, especially among older workers [2]. Thus, in order to prevent work-induced injuries, companies need to take into account the design of the work environment and assignment of employees to workstations.

Different approaches exist which aim to improve the ergonomic design of workplaces by analysing the physical stress on the worker’s body, such as the Owako Working Posture Analysis System (OWAS) or the Rapid Upper Limb Assessment (RULA) [3,4]. All of these methods rely on motion analyses for the individual workplace without taking the worker’s individual abilities into account. In contrast, a functional capacity evaluation (FCE) can measure the worker’s physical capacity [5]. However, results from this evaluation show whether or not a worker can go back to work without taking the workplace into account.
Hence, workers with physical limitations or unusual body dimensions only benefit to a certain level from improvements derived with the OWAS, RULA or FCE.

This paper presents an approach to assess a worker’s individual ability using a 3D-camera. With the Intel RealSense D415’s full body tracking, the physical constraints of a worker can be determined by measuring the angles between body segments. The worker is recorded while performing a set of ergonomic exercises that demonstrate the range of motion (ROM) of joints typically strained in assembly. Based on an exemplary exercise, this paper shows how a physical ability profile can be created.

The development of an individual worker assessment is part of the research project ErgoTrack. The main research question is:

> How can each worker’s individual abilities be taken into account when designing workstations or allocating the workers among workstations?

In order to answer the main research question, the following questions need to be addressed:

- What is the individual physical capacity and flexibility of a worker?
- What are the workplace requirements for each worker?
- How can the physical ability profiles be matched with the workplace requirements?
- What is the individual musculoskeletal stress for each worker?

This paper only addresses the first research question partly.

## 2. Current state of research

### 2.1 Conventional ergonomics analysis

The conventional ergonomics analysis methods can be divided into methods that measure an individual’s physical capacity and methods that analyse the workplace with its work sequence. The Functional Capacity Evaluation and the ERGOS® work simulator assess the physical capacity. During the FCE, a set of 29 standardised functional performance tests is performed. If required, additional work-specific tests can be added to the evaluation for certain work activities. The ERGOS® work simulator is a test station with 5 available units that each assess different capabilities, such as full-body movement. The FCE is performed in 5-6 hours over two days and the ERGOS® tests take 4-5 hours, excluding the time for evaluations and reports [5]. Both methods are therefore typically too expensive to be used on every worker of a production site.

For the workplace ergonomics analysis, among others, the OWAS and RULA can be considered state of the art. The OWAS consists of two parts. The first part is a method to evaluate working postures by observation in the form of work sampling. The second part defines a set of criteria for the redesign of working methods and places [3]. RULA is a method that assesses the postures of the neck, trunk and upper limbs which can lead to upper limb disorders. As a result, an action list is generated that specifies how much intervention is needed at each workstation in order to reduce stress on the body and thus the risk of injuries [4].

### 2.2 Ergonomics analysis using a 3D camera

A variety of approaches exist for ergonomic analysis using a 3D camera. Most research in this field was conducted with the no longer produced Microsoft Kinect and the corresponding Microsoft SDK. While some studies concluded that the Kinect is a reliable technology to measure the shoulder ROM [6–8], Huber et al.’s study [9] revealed concerns about using the Kinect’s motion data if precise angles are required. Kitsunezaki et al. [10] and Fernández-Baena et al. [11] examined the Kinect’s applicability for rehabilitation support and showed that it can be used for training monitoring and effectiveness. Furthermore, several systems have been developed that use the Kinect to perform a RULA in real time and give immediate
feedback about the recorded person’s posture in the form of a reliable RULA score [12–14]. The same has been done for the OWAS [15,16].

Instead of the Microsoft Kinect, we use an Intel RealSense camera, for which Siena et al. [17] reviewed the technology’s technical capabilities and concluded that “the Intel RealSense system can be seen as a comparable if not superior alternative to the Microsoft Kinect”.

3. Creating a worker-individual ability profile with a 3D camera

3.1 Experimental setup

In order to create a physical ability profile, the angles between body extremities are measured. For this purpose, a set of ergonomic exercises has been derived at the Institute for Biomechanics at the BG Unfallklinik Murnau, Germany. The exercises focus on the body parts that are subject to stress during assembly and commissioning tasks, such as the shoulders, elbows and spine. For this paper, one exemplary exercise was chosen that allows the measurement of the ROM for shoulder extension and flexion. The exercise and the corresponding angles are described in the next subsection. Each exercise is performed twice in succession to obtain reliable values for the ROM while keeping the expenditure of time low. The exercise was performed under the instruction that the subject should move their arms as far as they could. The exercise performance was only corrected if it differed significantly from the instructed movement. Minor rotations in the arm and/ or shoulder were tolerated.

19 subjects performed the exercise and analysis determined whether or not a reasonable angle can be deducted automatically. To determine if an angle is reasonable, a low-resolution screenshot of each recorded frame was captured. This allows the user to manually assess if the tracking of the required joints was performed correctly. Figure 1 shows an example of correctly tracked elbow joints (a) and falsely tracked elbow joints (b). The camera’s technical accuracy is neglected in this test setup.

![Correctly and falsely tracked elbow joints](image)

**Figure 1**: Low-scale screenshots for data verification.

3.2 Ergonomics exercise

The ergonomics exercise for this paper aims to assess the ROM for shoulder extension and flexion in the sagittal plane. The sagittal plane divides the body into left and right parts and is represented by the y-z-plane, see Figure 2a). In the exercise, the subject completes a shoulder extension. This requires them to move their arms as far as possible behind the back and making sure the palms face each other. Afterwards, during shoulder flexion, the subject raises the extended arms forward and above the head, maintaining the palms
facing each other. The ROM is the maximum angle between spine and upper arm in the sagittal plane of each movement, see Figure 2b).

![Diagram of body planes and axes](image1)

![Diagram of shoulder extension/flexion](image2)

Figure 2: Body planes and axes [18] and shoulder extension/flexion [19]

### 3.3 Data recording

#### 3.3.1 Data Acquisition

For the data acquisition we use the Intel RealSense D415 and the 3D body skeletal tracking middleware Nuitrack SDK. Intel’s 3D-camera uses infrared technology in order to create a depth map which is used by the Nuitrack SDK to approximate the position and orientation of 19 body joints. The orientation of each joint is described by a local coordinate system. For the data acquisition, the subject stands in front of the camera facing the lens. During the recording, the camera takes data at a rate of 15 frames per second to allow the processing of all relevant information. The Nuitrack SDK generates a value that describes the confidence in joint identification for each frame and each joint. The joint confidence is only given on a nominal scale, hence it only says whether the joint data can be trusted or not. This value is used for prefiltering the data, but preliminary tests showed that this filtering is not sufficient to exclude all false data from the analysis.

#### 3.3.2 Angle Calculation

To create a physical ability profile we are interested in the extreme values of the angles taken during each ergonomics exercise. The exercise discussed in this paper requires the calculation of the shoulder extension and flexion angles shown in Figure 2b). Using the position and orientation data of the shoulder, elbow and torso joints, the angles can be determined for each frame recorded by the camera. Both angles are defined as the angle around the x-axis between the vector from the elbow to the shoulder joint, the upper arm vector, and the vector from the torso to the left collar joint, the upper spine vector (see Figure 3). In order to obtain only the angle around the body’s transverse axis, all 3D coordinates are transformed onto the left collar joint’s local y-z-plane, representing the sagittal plane, see Figure 2a).
4. Results and discussion

Overall, the results show that we can detect the shoulder ROM in the sagittal plane by recording only two repetitions of the exercise. 84.2% of the recordings’ first maxima gave valid results, meaning the joints were tracked correctly when a maximum angle occurred (see Figure 1). The recording time ranged from 11 to 26 seconds (mean: 19.5 seconds). If the joint is not reasonably tracked on the first try, it can be assumed that either a joint was falsely tracked resulting in high angles at arbitrary frames or that the tracking was not working properly in the area where the angles’ extreme values occur. The data showed that the latter was the case. Since the software displays the recorded frame when a maximum occurred, the observer can identify false recordings right after the exercise and let the examined worker repeat it. Given the fact that the success rate of a valid recording lies above 80% and the examinations only take about 20 seconds, the method is by far faster than conventional methods. Further, we believe that the tracking can be improved by optimizing the recording conditions, such as tight and white clothing and better lighting.

Figure 4 shows the cumulative distribution function of the measured angles for shoulder flexion and extension. The cumulative frequency that corresponds to an angle is the percentage of people whose ROM lies below that angle. The mean angles for shoulder flexion were higher than the expected values from literature, which are 180° for shoulder flexion and 45°-60° for shoulder extension [20]. One of the reasons for this is that the values from literature are determined in clinical ROM assessments, in which the examiner restricts the motion by controlling the subject’s movement in a given plane. The captured screenshots show that the subjects did not only rotate their shoulders in the sagittal plane but also abducted them to allow a greater ROM. Since the derived angle is not intended to be used in a medical way, this circumstance is not disadvantageous. In fact, in order to assess the physical capacity for assembly and commissioning tasks, it is beneficial to assess the ROM from a rather natural movement, as these can be matched more easily with the workplace requirements. However, the additional shoulder abduction might not be the only reason for the high values. Future testing needs to show whether the measured values are correct or subject to systematic bias.

Even though the experiment was only conducted with a small sample, it shows the importance of an individual ability assessment. Both flexion and extension angles vary strongly among the participants. Assuming that the measured values are correct and that workers need a shoulder flexion of at least 180° during a certain overhead work task, 37.5% of the test persons would not be able to perform the overhead work due to their limited ROM in the right shoulder (see Figure 4b).
Figure 4: Cumulative frequencies of the measured angles with valid tracking. $n$ is the number of measurements, $\bar{x}$ is the mean value, and $S_x$ is the standard deviation.
5. Summary and Outlook

This article presents an ergonomic analysis method in which a depth camera is used to assess the individual abilities of a worker within a short time. Based on the example of a shoulder ROM measurement, we were able to demonstrate that it is possible to assess a worker’s movement ability by letting him perform two repetitions of an ergonomics exercise with an average recording time of 19.5 seconds. The experiments showed that the shoulder ROM varies strongly among the participants. That underlines the need of an individual ability assessment in assembly and commissioning. Looking at the cumulative frequencies of the measured angles, the assessor can easily see how many people can perform certain movements. For example, 37.5% of the participants are not able to perform overhead work tasks that require a shoulder flexion of 180° or above. Further testing is required to expand the method onto other ergonomic exercises to create a complete physical ability profile that can be matched with workplace requirements. The workplace requirements are planned to be assessed with an Xsens motion capture suit since current 3D cameras show weaknesses with the assessment of workstations due to bad lighting conditions and occlusion of relevant joints.

The absolute angles that we measured were higher than the values we would have expected from literature. Even though these values can be at least partly explained by the different forms of measurement, the system’s accuracy needs to be investigated. While several studies examined the accuracy for angle measurements with the Microsoft Kinect camera, there have been no studies on the accuracy of the Intel RealSense D415 with the Nuitrack SDK.

Acknowledgements

The presented work was part of the project ErgoTrack (20465 N), funded by the German Federal Ministry for Economic Affairs and Energy due to a decision of the German Bundestag. The authors would like to thank the AiF-Forschungsvereinigung as well as the associated companies SCHOLZ MECHANI K GmbH, Getriebebau NORD GmbH & Co. KG, Seapack Verpackung und Transport GmbH, Hermes Fulfilment GmbH, REFA Nordwest e.V., Hagemann Consulting and MediaAnalyzer Software & Research GmbH.

References


[16] Horejši, P., Gorner, T., Kurkin, O., Polasek, P., Januska, M. e-mail: tucnak@kpv.zcu.cz.


Biography

**Justus Brosche** (*1993) studied Mechanical and Process Engineering with a focus on production management at the TU Darmstadt. He started working as a research associate at the Institute of Production Management and Technology at the Hamburg University of Technology in 2018.

**Hannes Wackerle** (*1993) has been research associate at the Institute for Biomechanics at the BG Unfallklinik Murnau, Germany, and Paracelsus Medical Private University Salzburg, Austria, since 2017. He has studied Sports Science and Ergonomics – Human Factors Engineering and is a member of scientific societies of sports science and human movement analysis.

**Hermann Lödding** (*1971) studied Industrial Engineering and Management at the University of Kaiserslautern and completed a PhD in Engineering at the University of Hannover. From 2004 to 2009, he was working for Robert Bosch GmbH. Since 2009, he has been working as a professor for production management at the Hamburg University of Technology.
1st Conference on Production Systems and Logistics

A Vision of Digitalization in Supply Chain Management and Logistics

Michael Henke¹, Christoph Besenfelder¹, Sandra Kaczmarek¹, Michael Fiolka²

¹Fraunhofer Institute for Material Flow and Logistics IML /Chair of Enterprise Logistics, TU Dortmund University, Dortmund, Germany
²Fraunhofer Institute for Material Flow and Logistics IML, Dortmund, Germany

Abstract

Digitalization requires a new form of management to master the transformation process of corporations and companies. The Dortmund Management Model structures the focus areas of the digital transformation along the management tasks goal, planning, decision, realization and monitoring as well as the common socio-technical subsystems technological, organizational and personnel - enriched by a fourth dimension: information. Additionally, the acceleration factors transformation, migration and change management are taken into account.

This paper embraces a vision for a persistent management of production and supply chain networks in order to achieve a holistic Management 4.0. The emerging developments of technology, methods, tools and models in production and supply chain research are connected and merged into a big picture of digital supply chain management and logistics. The interfaces between management tasks show specific characteristics of digital business processes in particular, which are hereinafter exemplarily outlined:

New business models and value-creation networks are based on adaption intelligent production systems, which are interconnected with digital models for continuous planning and reconfiguration. At the shop floor and between sites orders are completed by autonomous guided vehicles (AGV) with intelligent load carriers. Decentralized negotiations and decisions across company boundaries concluded with smart contracts are enabling reasonable and sustainable distribution of the value creation processes. Humans are still in the center of action – abilities are developed by integrated competence management, new learning approaches and human-centered assistance systems coupled with AI-based decision-making support. New types of organizations allow a synergetic collaboration of humans and machines. The benefit of integrating new production and transport technologies becomes assessable and accelerates the ongoing renewal of existing networks.

This paper provides an overview of possible potential and connecting factors by linking different technological developments towards supply chain, logistics, production and management research and shows further research demands.

Keywords
Plattform economy; Silicon Economy; Management 4.0; Individualization

1. Introduction

Individualized customer requirements are increasingly driving the diversity of variants in Supply Chain and Productions Systems – the ‘lot size one’ is a continuous tendency. Industry 4.0 holds the promise for highly customized processes while maintaining the productions costs of mass productions systems. How has a
vision of future value-chain networks, process chains and technologies have to look like, to fulfil this objective properly? There are certain requirements that are already clearly noticeable. And all companies and industries in logistics and supply chain management are obliged to meet them. Same-Day-Delivery without additional shipping costs in almost all product categories is already a unique selling point and a differentiating characteristic of large trading platforms. Processes and structures have to adapt to fulfil these promises, while using the current technological developments in a human-centred and economical modality. One key challenge of the management is the goal-orientated design of the transformation process in their company. This paper describes a vision of digitized logistics and supply chain management. Research results are combined to give a complete picture of future supply chains, from the shop floor to the control of operational processes, from human-machine collaboration to autonomous supply chains. Afterwards, the ‘Dortmund Management Model’ is briefly presented, which structures the transformation process holistically and the need for further training for the management of digital transformation is illustrated. Finally, the development towards platform economics is presented and limitations and further research are outlined.

2. A Vision of Digitalization in Supply Chain Management and Logistics

The image of distribution centers in Logistics and supply chain management is characterized by miles of conveyors and giant sorters guiding parcels through the system. These large, fixed units are designed for a steady and high capacity utilization with long-term forecasted standardized processes to operate economically. These two requirements are no longer present [1]. The prospective supply chain systems are obliged to handle quantities and processes economically, which are unknown figures at the point of investment. In conclusion, significant changes on the shop floor are necessary. Fixed steady conveyors, which run a preconfigured route are not adaptable enough. Especially warehouse-location decisions for distributions centers are not possible to conclude for a number of decades. Changes within the range of products are more frequent due to shorter product life cycles and are influencing the necessary and reasonable locations. These changes create new requirements for the supply chains of the future and their components.

2.1 Infrastructure-less logistics

Forthcoming logistics systems not only have to be mobile and adjustable. At best, they have to be without any infrastructure at all. That does not mean it has to be without technology, but characterizes a vision of a capacity, which is fast to set up and dismantle. Research projects already describe concepts for a distribution center, which is ready for operation within a few weeks [2,3]. Crucial parts of this concept are automated guided vehicles (AGV). They not only fulfil the tasks of traditional unsteady conveyors like forklifts, they also replace conveyors. AGV systems are already able to realise transport relations on the shopfloor without being dependent on determined routes. Furthermore, they are also able to react directly on changes. If products, load carriers or storage locations change, the vehicle calculates a new route to navigate efficiently through the system. Capacities do not have to be expanded by long-term investments, the need for upscaling to compensate a larger job load is obsolete. The modularity of the systems makes it possible to compensate increasing demands by complementing the system with additional AGVs temporarily or permanently. A proper example for this kind of transportation systems are cellular transport systems. These systems act autonomously and are capable of driving on the shop floor as well as in storage racks [4]. This said, a handover form storage and retrieval machine to carriers is rendered mandatory. The capacity for moving goods in and out of rack systems no longer has to be set at the point of investment, but can be controlled by the deployment of cellular transport systems. It is possible, that in future warehouses shelves are no longer present. Innovative AGVs have the ability to pick a single small load carrier out of a stack located on the floor. In conclusion, simple stacks of small load carriers are sufficient to gain direct access on the whole inventory [5].

278
2.2 Decentralized control in production and logistics

Planning and control of this kind of adaptable systems influenced by fluctuating capacities requires new concepts. So far, strictly hierarchically organized and centrally arranged production planning and control is possibly no longer suitable for volatile market conditions. Mathematical problems induced by capacity planning and scheduling are solvable in a reasonable period using heuristics, but using the flexibility of the system by decentralized decision-making may be the more favourable option. The uprising challenge is to proof the decision quality of decentralized planning and guaranteed on-time delivery for customers. The necessary advancements are aiming for a connection between flexibility and decentralized autonomous control and the obligation of centralized approaches. Also, the expansion of planning steps of autonomous control in upstream processes and operations across company boundaries have to be addressed. [6–8]

Planning in partnership-based collaborations of maintenance and production will be a critical requirement for smart-production systems [9]. By using sensory and operational data, impending machine failures can be predicted [9,10]. Spare parts can be produced right at the place of demand by the use of 3D-printing technology [11]. Decentralized intelligence of machines enables an autonomous reaction, for example rescheduling of jobs in context of critical system conditions [12]. These are just a few examples of intelligent assistance systems that will support logistics tasks in the years ahead.

2.3 Social Networked Industry

Prospectively, nearly every workplace will be characterized by digital technologies and interaction with intelligent systems will be everyday practice. The work still has to be human-centered to create an adequate working environment and a sustainable and competitive economy [13]. Intelligent assistance systems relieve humans from monotonous and physical tasks. For example, AGVs can follow the worker automatically [14] while exoskeletons are reducing the workers physical distress while handling heavy goods [15]. In the vision of the social networked industry, humans and machines are working hand in hand. Like in social networks, entities communicate with each other and keep interchanging [16]. But not only the information exchange is enhanced, physiological and psychological constitutions of the workers are taken into account. The intelligent shelf recognizes the workers condition, for example the exhaustion level and proposes breaks accordingly [17]. The AGV shows its condition or ‘mood’ by easy interpretable symbols [18]. Considering the technological evolution, the realisation of this synergetic relationship among human and machine requires not only the inclusion of physiological parameters, but also the involvement and active design of cognitive aspects [19]. The development of corporations in this direction additionally includes the aspect of the organisation. Individual and fast changing customer and market requirements are leading to different activities and in conclusion to inconstant team compositions. Including social factors in the change process is more important than ever [20]. Virtual planning environments will be used to increase the participation in change processes and results inevitably in the inclusion of the process expertise of all employees [21]. Additionally, the demands upon the employee’s competences are going to change. The comprehension of the overall process and problem solving becomes more relevant and has to be measured and developed by a consistent competence management [22]. Competence requirements of the workers are evaluated and deviations from the competence profile are going to be levelled out using individual qualification measures. In doing so, new learning concepts are utilized: Game based approaches like serious games in further training and gamification in training measures will be integrated to increase self-learning competence and motivation of employees [23]. Intelligent assistance systems will also support the learning progress of employees. VR- and AR-technologies enriched by eye-tracking and AI are able to trace the learning success of the employees [24]. The necessary digital support during training and the level of difficulty for competence measures will be set individually and in real-time. The described changes in different logistics systems also affects the cooperation in networks in which new coordination and cooperation procedures become necessary.
2.4 Autonomous Supply-Chain networks

Especially the transition in corporate networks poses the question for automated and even autonomous conclusions of contracts between intelligent systems. The technology for negotiations of rates, but also the trigger of payments or mandatory orders are smart contracts based on distributed ledger technology. The immanent immutability of datasets using this technology ensures not only the manipulation security, but also audit secure accounting. The legal basis for autonmization of the financial flows, which is besides the material and the information flow significant in supply chain management, is currently investigated by the ‘Recht-Testbed’-Project. [25].

Artificial Intelligence and logistic assistance systems are also supporting the strategic and operational decisions in supply chain networks. Nowadays, simulation software is utilized to predict disturbances on the transport routes and used to choose the best possible option [26,27]. While planning value chain networks, not only cost and performance are taken into account. Targets concerning energy efficiency, for example warehouse locations problems and means of transportation are taken into consideration, too [28]. Future productions systems will be fully modular and will be dynamically plan- and optimizable by using assistance systems [29]. Decentralized controlled supply chain networks of the future will integrate the necessary intelligence for control, assessment of alternative options and communication already in the handled objects itself. A container recognizes delays on its path, works out alternative means of transport, commissions a service provider and handles the payment on pick-up [30]. Logistics 4.0 realized in this way requires the consistent and near real-time digital image of the enterprise’s resources. This challenge is called digital twin or digital shadow. Due to a digitalized and reality-synced image of the company and the related processes, the transparency of the present state is increased and also renders effects of possible measures calculable or simulateable [31]. These options contribute to automated adjustable production and logistics systems - the first step to adaption intelligent systems [8]. The challenge of the coming years will be to design the change in the direction of the possible future vision outlined above. The change of companies will not happen overnight and technologies and research findings must be evaluated, selected and implemented step by step in a targeted manner.

3. Management 4.0 – shaping the transformation process

As for digitalisation in logistics and supply chain management, the is still a long way to go. It requires a company-specific objective and a systemic approach. The D Management Modell offers an overview for the necessary tasks of the needed management 4.0. In contrast to other transformation models or management models, the Dortmund Management Model focuses on a domain and explicitly on the change in industry 4.0 (for a comparison of the models, see [32]).

The Dortmund Management Model is a framework to support management in the digital transformation of companies and is to be understood as a universal model. The challenge lies in an integrative management process that analyses the corporate culture, the corporate organization and the value-added processes and at the same time promotes process and technology innovations and must be operationalized individually under consideration of the individual needs of people.

In order to advance the transformation of a company to industry 4.0, three acceleration factors are integrated into the model: transformation, migration and change management. Transformation focuses here on an agile organization, made possible by the use of suitable technologies and organizational learning to adapt to changed framework conditions [33]. Migration describes the step-by-step introduction of technologies in companies and their processes [34]. Change management covers all activities that aim to bring about far-reaching, cross-functional changes in the implementation of new strategies, structures, systems, processes or behaviours in a company [35,36].
In addition, the framework of the Dortmund Management Model structures the management of digital transformation in two dimensions: a management dimension that describes the relevant (standard) tasks of management ('goal', 'planning', 'decision', 'realization' and 'monitoring') [37,38], and an organizational design dimension in the form of four pillars.

The integrated view of the Dortmund Management Model is reflected in the four pillars 'technology', 'role of man', 'organization' and 'information', which are equally included: This comprehensive organisational design approach is based on the MTO concept according to Ulich and includes the characteristics 'human', 'technology' and 'organisation' [39]. The increasing quality and quantity of digital data available in real time offers new possibilities for companies within the framework of industry 4.0. By combining digital and real processes in the sense of digitization, companies automatically or autonomously generate amounts of data at their disposal. The information resulting from this data, for example from cross-company value creation networks, is a valuable asset for innovative data-supported business models and thus a central asset in industry 4.0. This leads to a self-optimizing organization that enables autonomous and timely adaptation to the changing conditions in the business environment [33]. Consequently, a fourth feature is added to the MTO model: Information (I) (MTO plus I), [36]

![Dortmund Management Model](image)

Figure 1: Dortmund Management Model

On the basis of the framework, necessary fields of action for the management of digital transformation can be described along the management tasks and their interrelations with the pillars of the organizational design dimension. Consequently, it enables management to merge isolated technological solutions into an integrated management model by using the design principles of supply chain management. The model can thus be used by companies to structure the digital transformation process and include all necessary perspectives.
For concrete transformation projects, the individual tasks must be operationalized with appropriate methods. Various company case studies in which industry 4.0 technologies have been implemented in transfer and implementation projects have shown that the challenges that have arisen are explicitly addressed by the management model [40]. However, the employees, especially those responsible for transformation, must design the process and consider the pillars integrated.

Thus, the management of the digital transformation has to be seen as an interdisciplinary field of activity across company boundaries, which is currently animating new responsibilities and in conclusion new job descriptions like ‘digital transformation manager’, ‘chief digital officer’ or ‘digital transformation architect’. The competence profiles of these new job descriptions are nearly the same. For 99% of the companies, features like the will to change, handling complexity, communication and especially lifelong learning are the field where most action is required. In contrast, more than half of the questioned companies have no actual strategy for the development of these special digital skills. Simultaneously, over 74% of the employees desire more further training opportunities in the field of digital transformation. [41,42]

Under the lead of the Fraunhofer Institute for Material Flow and Logistics IML the project associates of the Center of Excellence Logistics and IT [43] in Dortmund, Germany, and the Fraunhofer Academy are developing a modular training format for the management of the digital transformation on the basis of the Dortmund Management Modell.

4. Outlook – Silicon Economy

The necessary technology for the given examples already exists. Objects like load carriers with sensors and micro computers may work energy self-sufficient for several years, and not only track their own movement, but being localized through 5G and communicate without the need of further infrastructure [44]. They are able to post payments via smart contract platforms in digital currencies and book logistics services independently. The linkage of the intelligent physical objects (IoT-broker), blockchain based transactions (blockchain-broker) and logistics services (logistics-broker) generates a new platform economy for the B2B-sector – the silicon economy [45]. The continuous connection of all necessary resources for a reasonable movement of things enables totally new business models [46,43]. The crucial data exchange between objects and companies will be bound to the purpose of use, based on the principle of data sovereignty, and is therefore trustworthy and secure [47].

The Dortmund Management Model is a contribution to structuring the task of digital transformation. The model is constantly being further developed and various methods for operationalization are still the object of research [32]. The task of the companies will be to develop a vision of digitalized logistics and supply chain management for themselves. One success factor will be the qualification of the employees for the transformation.

References


282


Biography

Prof. Dr. Michael Henke (*1971) has been in the leadership of the Chair of Enterprise Logistics (LFO) and institute director of the Fraunhofer Institute of Material Flow and Logistics (IML) at the TU Dortmund University since 2013. He is also an adjunct professor for Supply Chain Management at the School of Business and Management at the Lappeenranta University of Technology in Finland since 2015.

Dr.-Ing. Christoph Besenfelder (*1984) has been a research manager at the Chair of Enterprise Logistics (LFO) at the TU Dortmund University since 2017 and at the Fraunhofer Institute for Material Flow and Logistics (IML) for the Enterprise Logistics Division.
Dipl.-Päd. Sandra Kaczmarek (*1986) studied educational science with a focus on vocational education at the TU Dortmund University and has worked as a research assistant at the Chair of Enterprise Logistics (LFO) and the Fraunhofer Institute for Material Flow and Logistics (IML) since 2014.

Michael Fiolka B. Sc. (*1989) studied Logistics and Supply Chain Management at the TU Dortmund University and has worked as a student assistant at the Fraunhofer Institute for Material Flow and Logistics (IML) since 2017.
Incentive System Framework for Information Sharing in Value-Adding Networks

Alexander Zipfel¹, Konstantin Wink¹, Gunther Reinhart¹,²
¹Fraunhofer IGCV, Fraunhofer Research Institution for Casting, Composite and Processing Technology IGCV, Augsburg, Germany
²Technical University of Munich, Institute for Machine Tools and Industrial Management (iwb), Garching b. München, Germany

Abstract

To meet current market requirements and improve their competitive position, companies cooperate with their partners in value-adding networks. To exploit potential performance improvements, it is essential for companies to increasingly share Production Planning and Control-related data. Financial and non-financial incentives are beneficial to foster the inter-company exchange of such data. This paper proposes an approach for designing an application-specific incentive system framework, forcing information sharing in value-adding networks. The framework is based on a requirements analysis towards both value-adding networks and Production Planning and Control. The outcome of the approach is the possibility of deriving concrete incentive systems. The approach developed was applied and verified in a use case.

Keywords

Value-adding networks; production planning and control; incentive systems

1. Introduction

The tightening of global competition, e. g. through new entrants disrupting entire industries with innovative business models, and the customer’s requirement for individualised products being delivered within short times are only some of the market trends manufacturing companies are facing in the 21st century [1, 2]. Focussing on core competencies and collaborating in inter-firm value-adding networks is a key enabler for enhancing performance and, hence, staying competitive within this volatile economy [3-5]. The evidence that supports pursuing inter-company functionalities may even be indispensable in the future to allow companies to remain partners in an existing value-adding network [6].

To meet the customer’s requirements, special significance needs to be assigned to the sharing of information related to Production Planning and Control (PPC). Information sharing can be seen as the heart of inter-company collaboration, and thus forms the basis for efficiency improvements in Supply Chain Management (SCM) [7]. A potential barrier to information sharing across companies is a lack of incentives caused by imbalances in benefiting [8]. Therefore, finding suitable incentives for partners, e. g. through gain sharing mechanisms, may be a solution to foster the sharing of data [9, 10]. This publication provides a conceptual model that supports designing a suitable incentive system for information sharing and herewith facilitates practical implementation. In order to validate its functionality, the approach was applied and verified in a use case.
2. Information sharing in value-adding networks

The concept of cooperating with legally independent partners in supply chains and value-adding networks respectively in order to achieve both individual and common goals is crucial for remaining competitive in customer-centric business ecosystems that include both suppliers and customers (see Figure 1) [11, 12]. Porter [13] established the concept of reducing operational inefficiencies during the value creation of a product, however, the goal of improving cross-company production processes has evolved significantly since Croom et al. [14]. Companies tend to focus on their core competencies in order to reduce their real net output ratios and offer specialised products and services in value-adding networks. Competition shifts from a firm to a network level [15]. Within supply chain management, emphasis is placed especially on the value added in every step, in contrast to the cost incurred [16, 17]. Collaborative-based SCM initiatives mainly include the sharing of information about demand and operations, joint planning and joint decision making, incentive alignment, resource sharing, collaborative communication and joint knowledge creation [18]. Examples of this are Vendor-Managed Inventory, Efficient Consumer Response and Collaborative Planning, Forecasting and Replenishment. Collaborative planning, knowledge transfer and replenishment activities between the entities enable operational efficiency improvements of the value-adding network performance, which can lead to financial and operational benefits for every firm in the network. Financial benefits include reduced costs or increased sales, while operational benefits are less tangible, such as higher customer satisfaction, greater flexibility in the value-adding processes, increased product availability, reduced batch sizes, lower lead and replenishment times or shorter reaction times to changed demand [11, 19].

Due to its importance for inter-company collaboration, a definition of information sharing is given that is highly appropriate for this publication:

“Information sharing refers to the extent to which a firm shares a variety of relevant, accurate, complete, and confidential information in a timely manner with its supply chain partners” [18, p.166].

In greater detail, downstream information sharing by a target company means vertical sharing of information with the organisation that is located at a later stage in the value-creating process, e. g. a retailer is located downstream of a manufacturer. Accordingly, upstream information sharing by the target company refers to vertical information sharing to an organisation that is producing a good prior to the value-creation process at the target company, e. g. the target company shares information with its supplier. Finally, horizontal information sharing refers to the sharing or receiving of information from an organisation that operates at the same stage of the production process, e. g. competitors. See Figure 2 for a detailed overview of the directions of information sharing.

Figure 1: (a) Supply chain; (b) Value-adding network [20]
Figure 2: Typology of information sharing types adapted from Barrat [21]
3. State of research into incentive systems fostering information sharing

Researchers have already been investigating the effects of incentive systems for information sharing in value-adding networks and revenue models. Based on the relevant work presented, the research motivation for this work is derived below.

As information sharing improves the performance of a value-adding network under certain conditions [10], incentives are needed for stable cooperation in terms of solving the issue of a fair distribution of gains [22], i.e. appropriate rewards. The prerequisite for being able to align incentives among partners is assessing the gains that are to be shared [18], as well as revenue models in terms of “incentive schemes such as pay-for-performance and pay-for-effort” [23, p. 264].

Feldmann and Müller [24] suggested an incentive scheme for companies to share true information within value-adding networks as they identified actively falsified information as a major deficit of information-sharing concepts. Their incentive scheme is based on financial rewards for sharing true information, and penalties for sharing falsified or manipulated information. No incentive schemes are presented that incentivise value-adding network partners to intensify information sharing in the first place.

Wang et al. [25] created an incentive model for information sharing in the context of value-adding networks to enhance coordination among partners and improve the overall network performance. To estimate the impact of information-sharing initiatives in value-adding networks, they introduced an index that is based on the four incentive mechanisms “value-adding network contracts”, “team trust”, “joint decision making” and “information technology coalition”, all of which can be calculated by simulating values for diverse indicators of each. However, concrete recommendations are missing that would indicate which form of the four incentive mechanisms should be specifically applied in a concrete setting, e.g. which value-adding network contract would be most suitable for which kind of information-sharing structure.

Simatupang and Sridharan [23] introduced the integrative framework for value-adding network collaboration suggesting taking into account the following five core dimensions to plan and operate inter-company collaboration initiatives: “collaborative performance system”, “decision synchronisation”, “integrated value-adding network processes”, “incentive alignment” and “information sharing”. Their framework allows collaborative initiatives to be evaluated towards a company’s effectiveness in improving its SC performance, as well as actions to be defined on how to design, implement and run value-adding network collaboration.

Gassmann et al. [26] provide an overview of 55 business model types, including the underlying revenue model, which can in part be used to design incentive schemes for value-adding networks. Amongst others, they introduce the business model types “add-on”, “crowdfunding”, “freemium” or “pay-per-use”.

Even though scientific studies exist that deal with how to distribute earnings among the value-adding network partners through gain-sharing mechanisms [27], an approach to design a specific incentive system has not been modelled yet. In fact, this work follows up on a publication by Zipfel et al. [20], who suggested developing an incentive system for fostering information sharing.

To the best of our knowledge, no prior research has been found that formulates a model especially suited to incentivise value-adding network partners via concrete, individual financial or non-financial rewards to boost information sharing.

4. Approach for an incentive system for information sharing in value-adding networks

The application-specific incentive system framework developed for information sharing in value-adding networks consists of five steps (see Figure 3). The first four steps lay out the requirements for PPC-related information sharing, whereas in step five, the incentive schemes are derived from the requirements. The five building blocks of the framework are presented in greater detail in the subsequent sections 4.1 - 4.5.
4.1 Defining the performance target system

To set the goal of each information-sharing initiative, in a first step, the application-specific performance target system needs to be defined. Originating from Wiendahl’s PPC target system [28] that includes the two dimensions logistics performance and logistics cost, the value-adding network performance target system is divided into the two dimensions of financial and non-financial targets [29]. The financial targets are further divided into sales increase and cost decrease, whereas the non-financial targets are operational improvements or of intangible impact. As operational improvements, smaller batch sizes, shorter lead times or a higher adherence to delivery times may be mentioned, for example. Performance targets with rather intangible impact include, for instance, greater customer satisfaction.

4.2 Describing the value-adding network characteristics

The framework is based on a value-adding network with multiple entities, including suppliers, manufacturers and retailers, to serve the final customers’ needs. To set up an application-specific incentive system, the number of partners participating in the sharing agreement and the direction of the information flows is highly relevant. For instance, if safety stocks cause high inventory costs at multiple retailers of one manufacturer, the incentive system to be set up should consider all relevant partners. Furthermore, existing data flows between value-adding partners and ERP/PPC systems have a direct influence on the design and implementation of incentive systems, as existing data streams facilitate enhanced information exchange. This characteristic becomes especially important if implementation effort is compared with the expected outcome of setting up an incentive system for information sharing.

4.3 Identifying the PPC information needed

The next step requires the potential of different types of information to be isolated in order to attain the desired performance improvement targets. This building block is, hence, based on two factors: the type of information shared and the direction of information sharing. Information types derived from prior research can be categorised as demand (e.g. order size, due date or deviations), forecast (e.g. planning of machine capacities and batch sizes) and event (e.g. machine breakdowns or rush orders). The direction of information sharing is divided into upstream, downstream and horizontal information sharing (see Chapter 2). See Table 1 for an exemplary overview of typical effects of information sharing on the performance targets.
Table 1: Effects of information sharing on the performance targets

<table>
<thead>
<tr>
<th>Performance target</th>
<th>Improving information-sharing mechanism</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales increase</td>
<td>- Upstream demand for information sharing (UDIS) weakens double marginalisation and avoids missed stock-outs</td>
<td>[9, 30-32]</td>
</tr>
<tr>
<td></td>
<td>- Downstream forecast of information sharing leads to better retail prices</td>
<td></td>
</tr>
<tr>
<td>Costs decrease</td>
<td>- UDIS weakens bullwhip effect, thus enabling inventory cost savings, reduces overproduction and leads to unit production cost savings</td>
<td>[19, 32-38]</td>
</tr>
<tr>
<td></td>
<td>- Downstream information sharing (DIS) of shipping data reduces safety stocks while the sharing of product-related event data reduces inventory costs</td>
<td></td>
</tr>
<tr>
<td>Operational</td>
<td>- UDIS sharing can lead to shorter lead times and smaller batch sizes</td>
<td>[19, 34, 37]</td>
</tr>
<tr>
<td>improvement</td>
<td>- DIS of shipping data can lead to a higher level of timeliness</td>
<td></td>
</tr>
<tr>
<td>Intangible impact</td>
<td>- UDIS can incentivise the manufacturer not to establish a direct selling channel and lead to higher customer satisfaction due to fewer stock-outs</td>
<td>[11, 32, 39, 40]</td>
</tr>
<tr>
<td></td>
<td>- Upstream and horizontal sharing of forecast information can lead to a higher capacity of production</td>
<td></td>
</tr>
</tbody>
</table>

4.4 Assessing further company-specific requirements

The fourth step for designing incentive systems covers the identification and assessment of company-specific requirements. This building block allows to identify specific conditions in a standardised, structured and comparable manner. The requirements are derived from four dimensions, namely network-, information-, operations- and business model-related requirements. Each dimension covers multiple indicators that take into account the specific needs of an enterprise towards an incentive system for information sharing. In order to classify the company-specific requirements, a five-level scoring system was developed. Table 2 exemplarily introduces the scoring system for the network-related requirement “degree of flexibility”.

Table 2: Scoring system for the indicator “degree of flexibility”

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Indicator</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network-related requirements</td>
<td>Degree of flexibility</td>
<td>Very low, when future adaptations are very unlikely</td>
<td>Low, when future adaptations are rather unlikely</td>
<td>Intermediate, when future adaptations are rather likely</td>
<td>High, when future adaptations are likely</td>
<td>Very high, when future adaptations are very likely</td>
</tr>
</tbody>
</table>

4.5 Selecting the incentive scheme

After having set all relevant prerequisites (see Chapters 4.1 - 4.4), the fifth step is to select an adequate incentive scheme for each value-adding partner. Based on a literature review, ten incentive schemes have been identified as particularly suitable for the purpose of information sharing [10, 19, 26, 29, 41-44]:

- **Add-on**: A value-adding network firm offers its partner basic information for a specific price with an add-on option for further information at an additional cost.
- **Barter**: Partners in the value-adding network exchange and trade equivalent information among themselves.
- **Cost sharing**: Two forms of cost sharing might be applicable to information exchange: On the one hand, the cost of gathering the additional data is shared among the partners and on the other hand, the reduced cost and the resulting benefits incurred through information exchange might be shared.
- **Freemium**: A firm shares basic data for free with its partner, however, a premium option is available for sharing additional, more detailed data.
- **Open data model**: Firms collaborate by sharing information with each other for free to use and analyse, which leverages the whole network’s competitive position against other networks.

- **Pay-per-use**: A firm offers its partner information under the payment mechanism pay-per-use when needed, which means costs accrue only after relevant data has been exchanged and used.

- **Pay-with-data**: A physical or intangible product or service offered by one partner is paid for by the other partner providing certain information instead of money.

- **Revenue sharing**: The additional revenue realised through information sharing is shared according to a specific percentage among the partners contributing to achieving such revenue.

- **Subscription**: The firm sharing information offers its partners the opportunity to subscribe to information flows, hence, to get access to relevant data by paying a weekly, monthly or yearly fee.

- **Trade credit**: By sharing information with upstream network partners, companies may be granted approval for a longer payment period and, hence, a so-called trade credit.

To select the best-fitting incentive scheme in accordance to a certain use case, two process steps are required. Firstly, the general applicability of the incentive schemes has to be analysed at a high level (see Chapters 4.1 - 4.3). For this purpose, all incentive schemes were evaluated with regard to the type and direction of information needed and the number of partners involved in the network setting. Table 3 gives an exemplary overview of the evaluation results for the incentive scheme pay-per-use. By comparing the evaluation results with the scenario conditions given, a pre-selection of potential incentive schemes can be made.

**Table 3: Evaluation results for the incentive scheme pay-per-use**

<table>
<thead>
<tr>
<th>Incentive scheme</th>
<th>Type of information shared</th>
<th>Direction of information sharing</th>
<th># partners involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pay-per-use</td>
<td>Demand Forecast Event Diverse</td>
<td>Downstream Upstream Horizontal</td>
<td>Single Multiple</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X X -</td>
<td>X - x x -</td>
</tr>
</tbody>
</table>

Secondly, the characteristics of the remaining incentive schemes need to be contrasted with the company-specific requirements identified (see Chapter 4.4). By analysing the incentive schemes’ capabilities, amongst others by conducting a pairwise comparison, the incentive schemes that are best suited to meet the requirements resulting from the indicators could be derived. The analysis results are shown in Table 4.

**Table 4: Overview of preferred incentive schemes for company-specific requirements**

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Indicator</th>
<th>Preferred incentive schemes for highly scored indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network-related requirements</td>
<td>Degree of flexibility</td>
<td>Add-on, freemium, open data model, pay-per-use</td>
</tr>
<tr>
<td></td>
<td>Dependency on trustworthiness</td>
<td>Add-on, subscription, pay-with-data, pay-per-use, barter</td>
</tr>
<tr>
<td></td>
<td>Strength of ties among the partners</td>
<td>Revenue or cost sharing, pay-with-data, barter</td>
</tr>
<tr>
<td>Information-related requirements</td>
<td>Acceptable uncertainty about the outcome</td>
<td>Revenue or cost sharing, pay-per-use, open data model</td>
</tr>
<tr>
<td></td>
<td>Importance of the value of information shared</td>
<td>Add-on, subscription, revenue or cost sharing, pay-with-data, barter</td>
</tr>
<tr>
<td>Operations-related requirements</td>
<td>Ease of implementation</td>
<td>Trade credit, open data model, freemium</td>
</tr>
<tr>
<td></td>
<td>Performance measurement</td>
<td>Add-on, subscription, revenue or cost sharing, pay-with-data, pay-per-use</td>
</tr>
<tr>
<td>Business model-related requirements</td>
<td>Degree of monetary incentives</td>
<td>Add-on, revenue or cost sharing, subscription, pay-per-use</td>
</tr>
<tr>
<td></td>
<td>Level of integration into the company’s own business model</td>
<td>Revenue sharing, cost sharing, pay-with-data</td>
</tr>
</tbody>
</table>
5. Use case

For evaluation purposes, the developed framework was applied in a specific use case. The use case deals with series production of an e-scooter manufacturer M, who produces a maximum daily output of 50 (see Figure 4). This size differs from the batch size of the wheel supplier S2, which produces in batches of 400 due to the economic use of its production machinery. This is because the order release is based on actual consumption by end customers, i.e. orders are placed when a current stock of final products falls below a certain limit. In this case, the next batch of 400 wheels is produced and shipped resulting in stocks becoming available at the manufacturer’s plant. Furthermore, so that the buyers’ needs can be serviced in a timely manner, high safety stocks of wheels are maintained along the value chain. The unfavourable high stocks therefore result not only in high processing times, but also in extra administration and capital commitment cost. Hence, the hypothesis formulated is that an increase in information sharing between the retailers, manufacturer and suppliers can facilitate planning and reduce stocks. The incentive system shall be implemented with R2 first and rolled out later to other critical steps in the production process. According to M, the outcome should be focussed on financial benefits, which means that decent effort in setting up the collaborative system is acceptable if the likelihood of financial success is high.

![Exemplary e-scooter value-adding network structure](image)

To find the most suitable incentive system that allows M to reduce its stocks, the approach developed was applied. Once the inventory cost to be reduced had been defined as the performance target and the value-adding network characteristics had been analysed, the upstream demand information from R2 could be identified as the PPC information needed to be able to reduce inefficient high safety stocks. Hence, the incentive schemes trade credit, open data model and pay-per-use could be excluded from the final incentive scheme list, due to the information type and direction of information sharing. As a fourth step, the company-specific requirements were assessed and scored as either low, intermediate or high. For example, as the degree of flexibility with regard to future internal or external changes is not of upmost importance for company M, this indicator was rated low. However, as M strives for short-term financial benefits, the degree of monetary incentives could be rated high. Once all indicators had been assessed and rated for this particular use case, the requirements implicitly imposed could be compared with the capabilities of the incentive mechanisms in the following fifth step to identify the one that fits best. As an outcome of the analysis the incentive scheme revenue sharing was proposed as the most suitable for incentivising R2 to share upstream demand information so that M could lower inventory levels, reduce searching and administration times and finally achieve a higher output. Therefore, as decreasing costs was one of the goals formulated by M, it is fair to say that the obvious incentive scheme – in this case, cost sharing – is not always the best-fitting one.
6. Conclusion and outlook

Sharing of PPC-relevant information among partners in value-adding networks is a key factor in remaining competitive in the 21st century, since specific PPC information can have a positive effect on the performance of the whole value chain. Incentive systems make it possible to influence the behaviour of third parties, and thus foster inter-organisational information exchange. However, to date, too little conceptual research has been conducted on incentive frameworks for information sharing in value-adding networks. This study provides an approach for designing application-specific incentive systems fostering the exchange of PPC-relevant information among partners: Firstly, the performance targets of the improvement initiative and, secondly, network characteristics need to be defined. Once the basis is set, the relevant PPC information required needs to be identified in a third step. Once further company-specific requirements have been assessed, the incentive scheme most suitable to the use case can be selected.

In order to improve the applicability of the approach presented, expert interviews and further use cases will be conducted in the near future. These will, in particular, enable more customised recommendations for action to be included, such as adding industry-specific assessment criteria or further innovative incentive schemes. Moreover, it would be interesting to broaden the scope of the framework towards more complex settings, such as conflicting interests of partners that are part of various value-adding networks. Finally, it would be highly beneficial to include a mechanism for assessing the monetary value of information shared among the network partners, in order to be able to estimate the amount of the different rewards.

Acknowledgements

The Mittelstand 4.0-Kompetenzzentrum Augsburg research and development project is funded by the German Federal Ministry for Economic Affairs and Energy (BMWi) within the “Mittelstand-Digital” framework concept.

References


Biography

Alexander Zipfel, M.Sc. (*1992) works as a Research Assistant at the Fraunhofer IGCV, focussing on Production Management. He graduated from the University of Stuttgart with a degree in Technology Management in 2014 and from the Technical University of Munich with a degree in Mechanical Engineering and Management in 2017.

Konstantin Wink, M.Sc. (*1993) has worked as a Consultant at McKinsey & Company since 2020, focussing on Manufacturing & Supply Chain Management. Prior to this, he wrote his Master’s Thesis at the Fraunhofer IGCV and graduated from the Technical University of Munich in 2019.

Prof. Dr.-Ing. Gunther Reinhart (*1956) holds the chair for Industrial Management and Assembly Technology of the iwb at the Technical University of Munich. Since 2016, he has also been Director of the Fraunhofer IGCV. Furthermore, Prof. Reinhart is a board and council member for several companies.
1st Conference on Production Systems and Logistics

Efficient Use of Human-robot Collaboration in Packaging through Systematic Task Assignment

Sebastian Blankemeyer¹, Rolf Wiemann², Uwe-Karsten Vett², Tobias Recker¹, Dennis Pischke³, Annika Raatz¹

¹Leibniz University Hannover, Institute of Assembly Technology, Garbsen, Germany
²Sennheiser electronic GmbH & Co. KG, Wedemark, Germany
³Leibniz University Hannover, Institute of Production Systems and Logistics, Garbsen, Germany

Abstract

The ageing workforce in Germany is a major challenge for many companies in the assembly and packaging of high-quality products. Particularly when individual processes require an increased amount of force or precision, the employees can be overstressed over a long period, depending on their physical constitution. One way of supporting employees in these processes is human-robot collaboration, because stressful process steps can be automated in a targeted manner. With conventional automation, this is currently not economically possible for many processes, as human capabilities are required. In order to achieve a balanced cooperation based on partnership, as well as to use additional potentials and to consider restrictions such as process times, it is necessary to ensure a good division of tasks between human and machine. The methodical procedure of allocation presented in this paper is based on the recreation of the process from basic process modules conducted by the process planner. Subsequently, these processes are divided according to the respective capabilities and the underlying process requirements. The company-specific target parameters, such as an improvement in ergonomics, are taken into account. The assignment procedure is described in a practical use case in the packaging of high-quality electronic consumer goods. Furthermore, the use case demonstrates the applicability of the approach. For these purposes, the parameters and requirements of the initial and result state of the workplace are described. The procedure and the decisions of the approach are shown with regard to the achievable goals.

Keywords

Human-robot collaboration; Design method; Task assignment

1. Introduction

In a globalized environment, manufacturing companies are confronted with a dynamic market that is characterized by increasing numbers of variants, simultaneously lower quantities and shorter product life cycles. These developments imply particular challenges to assembly systems in particular since assembly is responsible for up to 70 percent of product costs [1]. On the other hand, assembly processes are difficult to automate due to a high proportion of secondary work steps and peripheral equipment. Full automation can often only be implemented profitably in the area of large-series production since it requires high investments and is associated with a low level of flexibility and transferability to other applications. Considering this, assembly processes are often dominated by manual execution [1]. In addition, companies in the western industrial nations are faced with cost pressure due to high wages in this area. Furthermore, they must react
to a shortage of skilled workers and demographic changes in order to secure their international competitiveness [2]. These factors force companies to increase the degree of automation in assembly which leads to higher productivity.

Collaborative assembly systems connecting humans and robots are expected to be one way to meet these challenges. They offer a possibility to use automation potentials in the area between manual and fully automated assembly. However, when looking at the market for industrial applications of collaborative robots, it is noticeable that only a small proportion of industrial robots are collaborative and only a few real collaborative applications have been implemented [3]. The high potential that the human-robot collaboration (HRC) offers for modular, flexible and agile production processes is therefore obviously not exhausted yet. The reasons for this are manifold. Often companies state that they have limited experiences in planning and development as well as in the implementation of this comparatively new technology. The relatively long planning and development phases, which are connected with high personnel and financial efforts, represent a significant obstacle for many production companies [4].

In order to counteract the lack of experience and uncertainty, the “SafeMate” research project aims to develop a guideline for the introduction of collaborative systems. This guideline should support companies in planning and development as well as in commissioning of such systems and thus contribute to better exploit the potential of HRC [4]. In this paper, the assignment of tasks between humans and robots is described. The approach is explained using an application case as an example. In this way, the challenges in the implementation of the process become clear.

The article is divided into four sections. We present current approaches to systematic task assignment in HRC. Followed by this, the developed procedure for assigning work contents is explained in the example of a packaging process. In chapter four, the implemented solution and the results gained from it are presented. Finally, the article is briefly summarized.

2. Systematic Task Assignment

In order to simplify the introduction of HRC and thereby advance technical progress, researchers are increasingly working on developing methods that make it easier for companies to plan and implement such systems successfully and reliably. One of the essential steps in the introduction of collaborative systems is the evaluation of suitability for the execution of the manufacturing steps using a robot. This step is often carried out without any quantitative calculation because the decision which step is useful for the robot is based on previous experiences of the employee and process-specific aspects.

In order to systematically select a workplace for the implementation of HRC, several approaches have already been developed. A central decision in the partial automation of a process is to clarify which processes are suitable for HRC and what the task assignment looks like since the success of implementation can be significantly influenced by these facts. Thus, there are developments that try to determine a suitability for HRC based on capabilities [5] or standardized work descriptions [6]. With other approaches, the automation ability of each individual process is considered and transferred to the complete process [3, 7]. In addition, scientists have also developed procedures to allocate the jobs in human-robot interaction. One approach uses mathematical models in combination with a genetic algorithm to find a proper distribution [8]. Another assignment method estimates the time using Methods-Time Measurement to minimize the total production costs [9]. In Tsarouchi et al. the user starts by developing a layout in order to simulate the process [10]. Based on the results the assignment of the tasks is carried out. Pischke et al. use the specifications of the sub-processes to allocate the tasks to the robot and worker [11].

The approach developed in the “SafeMate”-project is based on a two-stage procedure for determining HRC potential. First, we ask a number of questions to determine whether there is general suitability in terms of
ergonomical improvements, rationalization possibilities or increased output [3]. The identified questions include topics such as component feeding, the workflow and the necessary cycle time. If this step fails, the user should continue with the evaluation of another workplace instead of going on with this one. Passing leads to the second step in which the user recreates the process with the help of basic process blocks. After that attributes are assigned to the process blocks so that the capability can be assessed. A list was created on the basis of an expert survey to identify capabilities. The expert survey rated different configurations. The scale ranges from 0 (more suitable for humans) to 1 (more suitable for robots). A value close to 0.5 implies that there is no preferred solution.

The sub-processes are then assigned with the help of the predefined capabilities, among other parameters. At first, the sub-processes that are particularly suitable for a resource are assigned [11]. For further assignments, the user has to evaluate if a rescheduling of sub-processes is possible. If it is the case, the user can consider whether the assignment allows building complete process blocks for a resource or an alternating work sequence between the collaborating partners. In this context, a block-building means that several consecutive sub-processes can be carried out one after the other by the same resource without dependence on other process steps. This has the advantage that complete task packages can be performed by one resource without disruption. The product is transferred to the other resource when the task is complete. If the complete blocks of both partners have similar cycle times, the waiting times of each resource can be avoided or at least minimized. Another approach is a sequence in which the partners can work alternately. This allows one partner to prepare components while the other partner works on the product. If it is well coordinated, this approach can also reduce the cycle time. A decision, which system is suitable for the process, can be made on the basis of the cycle time. In many processes, the cycle time is a target specification that can be determined with the help of real measurements or, for example, the MTM analysis.

3. Initial Situation

In addition to economic efficiency and productivity, the ergonomics of workstations are also critical challenges for manufacturing companies. Workplaces with a high potential for improvements in economics and ergonomics are therefore particularly suitable for partial automation with a collaborative robot. Sennheiser electronic GmbH & Co. KG is a company that has already solved such challenges at workplaces with the help of robots capable of collaboration. However, these workplaces can more likely be classified as the so-called coexistence [12]. This means that humans and robots work together without a separating safety fence but no real cooperation takes place. Instead, the human operator just takes over provisioning tasks.

As part of the "SafeMate" research project, a manual packaging workplace for microphones was selected by Sennheiser. The use cases in the project are supposed to implement real cooperation between humans and robots. The following application was selected in close consultation with the production managers and employees based on ergonomic aspects and economic reasons. This application is particularly suitable for HRC, since traditional, hard automation is very difficult to implement or very cost-intensive due to various complicated sub-processes, which are described in the next subsections.

3.1 Current Process

At the selected workstation, two workers pack microphones and the needed equipment manually in a carton box (see Figure 1). The equipment includes several parts, such as the storage bag for the microphone and the instruction manual. During the packaging process, additional cardboard parts are folded so that the microphone fits perfectly into the box. Finally, the carton must be closed. Most of these steps are pick-and-place processes, thus automation is possible. However, some of the remaining sub-processes are
difficult to automate due to the properties of the parts (e.g. elasticity, stiffness...) or the process requirements (e.g visual inspection...).

Table 1 shows the individual sub-processes in their sequence, priority and respective process times. The steps performed by worker 1 sum up to a total of 11 s, which the employee can perform in 10 s by parallelizing some steps using both hands. The second employee needs 18 s. A rearrangement of the assembly steps is only possible to a limited extent since the processes are based on each other and the sequence of steps is often predetermined.

Table 1: Process steps with priorities and execution times

<table>
<thead>
<tr>
<th>Priority</th>
<th>Step</th>
<th>Time [s]</th>
<th>Priority</th>
<th>Step</th>
<th>Time [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>take carton</td>
<td>1</td>
<td>1</td>
<td>take carton</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>insert bag</td>
<td>2</td>
<td>2</td>
<td>wide cuff</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>insert manual</td>
<td>2</td>
<td>3</td>
<td>fold cuff</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>fold cardboard bottom</td>
<td>1</td>
<td>4</td>
<td>visual inspection of microphone</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>insert clamp, close flap</td>
<td>2</td>
<td>5</td>
<td>joining cuff and microphone</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>fold and insert spacer</td>
<td>2</td>
<td>6</td>
<td>insert microphone into carton</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>supply carton</td>
<td>1</td>
<td>7</td>
<td>close carton</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>place carton on labeler</td>
<td>2</td>
</tr>
</tbody>
</table>

\[
\sum_{\text{due to parallelization}} 11 \quad \sum_{\text{due to parallelization}} 18
\]

Based on the manual work steps described in Table 1, an automation concept is being developed in which each sub-process is examined concerning its automatability using a collaborative robot. The main goal is the reduction of the workforce to one employee, maintaining the current cycle time.

### 3.2 Potential analysis for human-robot collaboration

As stated in the section before, the analysis of the use case presented in this article leads to the conclusion that there is a potential for rationalization and an improvement in ergonomics. There are two main reasons for this assumption. On the one hand, the process contains activities that are difficult to automate. For example, a visual inspection must be carried out where a worker checks the microphone basket for a firm fit. He also checks the microphone for small scratches or other irregularities so that the customer can be provided with a faultless product. This is difficult to achieve with today's automated systems because of varying ambient conditions and the uncertainties in the fault itself. Also, the packaging contains flexible parts so that full automation can only be achieved with considerable effort. Partial automation, on the other hand, makes sense because there are many sub-processes suitable for automation and some of which place huge physical stress on the employee. For example, the "wide cuff" process demands a high force applied...
by the worker. The openings of the cuff must be widened with a force of 80-100 N on a mandrel, which puts a considerable strain on thumb and shoulder at more than 700 times per shift (see Figure 1). Another goal, which is to be achieved with partial automation, is to increase productivity so that in future one person can carry out the packaging process alone. This is of additional relevance for the company because the demographic change (the proportion of the working population in Germany is continuously declining [13]) makes it difficult to replace employees leaving the company due to retirement.

In order to investigate the packaging process, the assessment process of Pischke et al. [11] is carried out. Therefore, all sub processes defined in Table 1 need to be analyzed. The results show that the processes "fold in inlet" and "wide and fold cuff" offer a high degree of eligibility for the robot (Figure 2). On the one hand, this is because high stresses and strains occur for humans, which must be prevented. On the other hand, only simple movements are required for the process that can easily be taken over by a robot. At the same time, other processes require capabilities that can only be provided poorly by an automated system like the visual inspection mentioned above. Also, there are many processes in which both partners are equally well qualified. Especially with these processes, it is essential to properly evaluate how these can be distributed among the partners.

Figure 2: Capabilities and task assignment of the developed workplace

Overall, the evaluation shows that the workstation is suitable for a HRC from a sub-process point of view because it contains processes that utilize the strength of humans and robots. If this is not the case, manual or fully automated assembly can be a better option. Based on this assessment, the entire, semi-automated process including the new workstation layout can then be developed.

### 3.3 Task Assignment

The division of the sub-processes was first carried out on the basis of the capabilities, whereby first the processes "fold in inlet" and "wide and fold cuff" were assigned to the robot and "insert clamp and close flap", "fold and insert spacer", "microphone visual inspection" and "close carton box" were allocated to the worker. After this assignment and the consideration of the cycle time, the time frame remaining for additional processes of the worker in each cycle is only 6 seconds. This corresponds to about two additional steps. Accordingly, the other processes must be allocated to the robot. After a rough estimation of the cycle times for the robot (e.g. via RTM [14]), it is decided that a second robot is necessary. This is because too many tasks have to be assigned to the robot, which cannot process the tasks in the given cycle time. With regard to the prioritization of the processes and the use of another robot, an alternating sequence is hardly possible.
in this case because waiting times for all partners would occur. Accordingly, three sub-process blocks are formed, whereby one is done by the worker and two by the robots (see Figure 2).

The first block deals with the preparation of the cardboard box. This block includes the sub-processes “insert bag”, “insert manual” and “fold cardboard bottom”. The second block contains the preparation of the cuff and is also carried out by a robot. The operator is supposed to do the processes “joining cuff and microphone”, “insert microphone into carton” and “place carton on labeler”. The whole process of the worker is as follows. At first, the worker puts the clamp in the carton box. After that he folds and inserts the spacer into the prepared carton. He then carries out a visual inspection of the microphone and, if there are no faults, fits it into the cuff. The microphone is then packed in a carton, the carton is closed and transferred to the labeler. This results in a cycle time of approx. 22 seconds. A cycle time of 18 seconds can be assumed by parallelizing the processes (e.g. inserting the clamp and folding the insert, since the human operator carries out these steps in parallel by working with both hands).

The resulting scheduling must then be investigated in more detail in experiments or simulations, as the times for the robot and the human parallelization are based on an estimate. However, with the example could be shown that a methodical procedure helps to generate an efficient sequence based on partnership.

4. Implemented Process

After a successful investigation of the cycle times, the implementation process of the workplace can proceed with the development of the workplace realization. Taking into account the derived job scheduling and cycle times, a layout needs to be developed. In addition, there are safety standards to be fulfilled, which can result in longer cycle times or even prevent an implementation of HRC workplaces. It is possible that the scheduling has to be adjusted due to safety reasons. In this case, the user needs to switch the relevant sub-processes. In the following part, the developed packaging process as a HRC workplace is described with the emphasis on process times and safety.

4.1 Collaborative Process

As described in the last chapter, the developed packaging process is divided into two automated and one manual processes (Figure 3). In the first process, a bag and an instruction manual are inserted into the carton, the inlet is folded and finally made available to the operator. This process is therefore decoupled from the worker and not in the worker's work area. For this reason, this process can be considered a coexistence. Nevertheless, the safety standards, in particular ISO/TS 15066 [16], must be obeyed. This means that the speed and acceleration of the robot must be limited in order to comply with the force and pressure limits defined in the safety standards. After conducting an experiment on the cycle time, it is found that the desired cycle time cannot be achieved with the intended hardware. Consequently, additional automation is needed in order to maintain a cycle of less than 18 seconds (Table 2). In this application, a friction feeder can carry out the insertion of the operating instructions to save manual process time. Once the instruction manual has been inserted, the robot removes the bag from a magazine and inserts it into the carton. The gripper is executed as a suction pad, as this makes it easy to separate and grasp the pockets. The gripper was specially designed and has rounded edges for the protection of the human collaborator. The gripper also serves as the device to fold in the inlet. Finally, pneumatic cylinders transport the carton into a buffer, thus making it available to the operator in the next step.
In the manual process, the worker removes the carton from the buffer and inserts a microphone clamp and an insert. The operator then picks up a microphone, performs a quality check and inserts it into the cuff provided by the second robot (Figure 4, left). Finally, he inserts the microphone with the cuff into the carton, seals it and places it on a labeling machine.

Table 2: Collaborative Packaging Process

<table>
<thead>
<tr>
<th>Step</th>
<th>Process 1 [s]</th>
<th>Worker [s]</th>
<th>Process 2 [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>insert bag + manual, fold carton bottom</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>convey in buffer</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>insert clamp and insert</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fold, wide and provide cuff</td>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>visual inspection of microphone, joining</td>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>with cuff, close carton, place on labeller</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>∑</td>
<td>20</td>
<td>23</td>
<td>16</td>
</tr>
<tr>
<td>due to parallelization</td>
<td>17</td>
<td>18</td>
<td>16</td>
</tr>
</tbody>
</table>

In the second automated process, the cuff is expanded, folded and made available to the operator. A self-developed fixture is used to widen the cuff. Two pneumatic cylinders are used to expand the cuff after the robot put it in the device (Figure 4, right). Since the robot works in the same working area and on the same product (cuff) with the worker, this is a collaborative process. The cuff is both folded and the openings widened by pneumatic cylinders. As for the first process, a suction pad was specially designed for this purpose. The magazine works on the same principle as the one in robot process 1. After folding and widening, the cuff is placed on a joining device so that the worker can reach it. An inductive sensor is installed in the joining device, which detects the microphone inserted by the worker. The robot waits until a
microphone is detected and then starts with the next cuff. In this way, the employee is controlling the cycle time, since a new cuff is prepared simultaneously by the robot.

4.2 Safety

In principle, DIN EN ISO 10218-1 [15] and ISO/TS 15066 [16] must be complied with for applications with collaborative robots in Germany. The safety of the person is guaranteed by safe hardware in the robot or by a safety-rated control system. Additional safety housing and distance monitoring can be used in this case. To ensure safe operation within the framework of ISO/TS 15066, the biomechanical force and pressure limits defined in the standard for the various body regions must be met. Therefore, force and pressure measurements are carried out at all potential pinching and contact points. If the limit values were exceeded, the acceleration, speed or force limitations of the robot would have had to be adjusted. These measurements were already carried out in the initial test setups in order to be able to draw conclusions about the expected cycle time at an early stage of development.

Particularly critical squeezing points in the force and pressure measurements have been evaluated between the robot gripper and the folding device in the second process. For this reason, additional padding was attached to the folding device. All in all, rounded edges were added at the grippers and the devices, so that the safety limits can be fulfilled at all potential contact points during the whole application. For the risk assessment, however, it is not only the robots that have to be considered but also the entire system. The system is controlled by a safety PLC, which ensures, for example, that the pneumatic cylinders can only extend at certain times to prevent injuries to the operator. Finally, a CE declaration of conformity is drawn up taken all the safety standards into account.

4.3 Economic Efficiency

In addition to the goal of ergonomic relief for the worker, the packaging process should be able to be carried out by just one person. Productivity should be increased and economic profitability is not least the prerequisite for automation. In production, the packaging station is operated in two shifts and in rare cases three shifts, so that four employees are planned for the workplace. The collaborative process allows this to be reduced to two employees. This saving is offset by the costs for the system, which are made up of the material costs and the working hours. The working hours account for a large part of the costs. This can be explained above all by the new technology of collaborative robots, which is very time-consuming due to the lack of experience in process development and, in particular, in risk assessment. In order to counteract precisely this problem, the SafeMate research project subsidizes working hours and, in some cases, material costs. This enables an ROI of fewer than two years, which is required internally at Sennheiser. The experience gained through the SafeMate project thus contributes to the fact that future applications can be implemented faster and thus more cost-effectively.
5. Conclusion

In this paper, an opportunity for assigning tasks between humans and robots in HRC is described using the example of a workplace for packaging microphones. Based on the process description, the work contents could be divided economically and capability-based. It turned out that for an efficient solution an employee should work together with two robots. In addition, a friction feeder was installed in order to meet the required cycle time.

Difficulties in the implementation were particularly evident in the handling of the parts. Some packaging materials made of paper behave differently depending on the prevailing environmental conditions (e.g., temperature and humidity). This influences the process stability if the robot cannot react to the corresponding uncertainties. In addition, the authors recommend a precise feeding of semi-finished products in order to provide appropriate conditions for automation. At the same time, the achievement of specified cycle times is sometimes only possible with increased effort due to compliance with the relevant safety standards. This is primarily because the speed of the robot is reduced in order to maintain the contact forces with the worker.

Acknowledgements

This research and development project was funded by the German Federal Ministry of Education and Research (BMBF) within the “Innovations for Tomorrow’s Production, Services, and Work” Program (02P15A080) and implemented by the Project Management Agency Karlsruhe (PTKA). The authors are responsible for the content of this publication.

References


Biography

Sebastian Blankemeyer studied mechanical engineering at the TU Braunschweig and has been working as a research associate at the Institute for Assembly Technology at Leibniz University Hannover since 2013. His work focuses on automation and human-robot collaboration.

Rolf Wiemann studied mechanical engineering at the Leibniz University Hannover and wrote his master thesis in the context of the SafeMate project. Since 2018 he has been working as a process developer at Sennheiser electronic GmbH & Co. KG and focuses on human-robot collaboration.

Uwe-Karsten Vett studied electrical engineering at the FH Hannover and has been working at Sennheiser electronic GmbH & Co. KG since 1984. He has many years of experience in production engineering and is currently focused on the introduction of human-robot collaboration.

Tobias Recker studied mechatronics at the Leibniz University Hannover and has been working as a research associate at the Institute for Assembly Technology at Leibniz University Hannover since 2015. His work focuses on human-robot collaboration and mobile robots.

Dennis Pischke studied mechanical engineering at the Leibniz University Hannover and has been working as a research associate at the Institute for Production Systems and Logistics at Leibniz University Hannover since 2014. His work focuses on human-robot collaboration.

Annika Raatz has been leading the Institute for Assembly Technology at Leibniz University Hannover since 2013. Her key research areas include Soft Material Robotic Systems, precision assembly and automated assembly and disassembly.
Virtual Asset Representation for enabling Adaptive Assembly at the Example of Electric Vehicle Production

Peter Burggräf¹,², Matthias Dannapfel³, Tobias Adlon³, Viviane Hahn³, Aaron Riegaut³, Patricia Casla⁴, Angelo Marguglio⁵, Izaskun Fernandez⁴

¹ IPEM of University of Siegen, Siegen, Germany
² Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University, Aachen, Germany
³ Chair of Production Engineering of E-Mobility Components PEM of RWTH Aachen University, Aachen, Germany
⁴ Fundación Tekniker, Eibar, Gipuzkoa, Spain
⁵ Engineering Ingegneria Informatica S.p.A, Rome, Italy

Abstract

Manufacturing companies are confronted with the challenge of adapting to ever-changing requirements of markets in order to remain competitive. Besides the rising number of product variants, increasingly frequent product changes require a continuous adaptation of assembly processes including its work instructions. Adaptive and highly connected agile assembly systems are designed to meet these challenges by enabling the interaction of various assets in assembly. A successful implementation of such Industry 4.0 (I4.0) solutions requires the development of a semantic oriented adaptive framework, which connects the physical with the virtual world. It enables interactive and situation-aware solutions such as Augmented Reality applications to adapt to worker capabilities and to improve worker satisfaction by providing information, based on individual experience, skills and personal preferences. A central part of the adaptive framework is the semantic representation of tangible and intangible assets through a Virtual Asset Representation containing all relevant asset information for adaptive assembly. This paper shows a three levels structure for adaptive assembly implementation, consisting of the adaptive framework level, the Virtual Asset Representation (VAR) ontology level and the use case level. The implementation of an adaptive assembly system is shown in the use case of a rear light assembly process of an electric vehicle in the context of the EU funded project A4BLUE. Based on the gained experiences a critical reflection on target fulfilment and user-friendliness of the VAR is given.

Keywords

Internet of Things; Industry 4.0; Agile Assembly; Adaptive Assembly; Virtual Asset Representation; Semantic Technology

1. Introduction

The globalisation, increasingly individualized products and complex production processes require flexible, responsive production systems and personnel. At the same time productivity and quality should be held on a constant high level. [1] Innovative solutions in production from the field of I4.0 are used to encounter these challenges. [2] Agile assembly systems enable the required flexibility and responsiveness offering the opportunity to consider individual and short-term customer wishes related to a product. Even small batches of lot size one can become profitable due to integrated digital engineering [3]. Dynamic process design through ad-hoc connectivity of cyber-physical production systems (CPPS) can lead to fast, flexible reaction
and decision-making ability and enable to react to changes in the surround. Furthermore, the productivity of many processes can be increased by using intelligent assembly assistance solutions like Computer Aided Assembly Systems, worker guidance tools or Augmented Reality (AR). [4] Using the described technologies and further intelligent assistance systems, a consistent quality in production can be ensured. [5] While connectivity is a prerequisite, these technologies enable to react adaptively to changes. This paper addresses the research question of how a VAR functions within a semantic oriented adaptive framework and how such a framework can be used to enable agile and adaptive assembly.

2. Frameworks for enabling Adaptive Assembly

2.1 Adaptive automation assembly

Assembly systems enhanced by digital technologies and intelligent automation need to be easily reconfigurable to achieve improvements in system adaptability and a minimisation of installation efforts. [6] For this, the I4.0 revolution is driving the development and introduction of new technologies and automation for adaptive assembly systems. [7] Adaptive systems are able to adapt to changes in the production environment as well as to different operators whereby the latter is focused in this paper. The system considers the user as well as the assembly task and provides optimal ways of interaction and assistance. [8] The ambition in this context is an ideal collaboration of humans and technology with the exploitation of their respective strengths in hybrid assembly systems. [9]

2.2 Industry 4.0 assembly devices

Various I4.0 technologies are designed to improve assembly processes while the solutions address different areas of manual assembly and have different levels of technological maturity. Applications can roughly be subdivided into two groups. Technologies to improve interactivity are, for example: [9]

- Electronic Lifting Aids/Exoskeletons (Lifting Aids)
- Collaborative Robots (Cobots)
- Driverless Transportation Systems (DTS)
- Optical Control Systems (Optical Control)

Technologies for improving adaptivity: [9]

- Interactive/Adaptive Interaction Mechanisms (Interaction)
- Augmented Reality/Assisted Reality/ Virtual Reality (AR)
- Adaptive/Self-Learning Production Control (Prod Control)

are focused in this paper.

2.3 VAR within A4BLUE Adaptive Framework

In order to enable the functioning of I4.0 elements, different systems must be connected to each other. The structure for the integration of systems like AR Guidance or a Decision Support System (cf. Figure 1) and the resulting interfaces can be defined in a semantic oriented framework. The development of the A4BLUE Adaptive Framework is based on well-known reference architectures for digital industry such as RAMI 4.0, IIRA and FIWARE for Industry. The Reference Architecture Model Industry 4.0 (RAMI 4.0) is a three-dimensional layer model in which the essential elements of Industry 4.0 are brought together. Industry 4.0 technologies can be systematically classified in this model. [10] Industrial Internet Reference Architecture (IIRA) is a standard-based open architecture for Industrial Internet Systems (IISs). The description and presentation of the architecture are generic and at a high level of abstraction to support the required broad industry applicability. [11] FIWARE for Industries is an open-source smart manufacturing platform based
on industry standards and open source components facilitating the development of apps for Smart Industry solutions. [12]

The Adaptive Framework was specified in A4BLUE in-line with recently introduced reference architectures for the manufacturing industry. It envisions automation and adaptation functionalities in combination with factory physical processes, including real-time operations and taking into account process, product as well as operator variability. One module of the framework is the Virtual Asset Representation (VAR) functioning as a central semantic repository. [9] The VAR is based on an ontology, which are generally used to represent knowledge. Domain-specific vocabulary is combined with statements about relations of the entities to which the vocabulary belongs. Natural language is replaced by a structure from which the information can be consistently retrieved. [13] The A4BLUE Adaptive Framework is built on the pillars: virtualisation, adaptation management, worker assistance support and monitoring. Adaptation management supports continuous data gathering, analysis and reaction to relevant events, which leads to the triggering of real time adaptation actions. Worker assistance support contributes to context aware work instruction via VR/AR hardware and software components. Also, the transfer of knowledge from skilled to less experienced workers should be facilitated. Monitoring enables the evaluation process by supporting the acquisition and visualisation of performance indicators (KPIs) to assess the impact from an economic and social perspective. [14]

3. VAR for Assembly Structure

The VAR ontology forms the basis for the actual semantic repository. It follows a modular approach to virtually represent production resources and thus enabling connectivity. To understand the VAR structure and its integration within the Adaptive Framework three hierarchical levels can be distinguished (see Figure 1). The first level consists of the Adaptive Framework, which enables event-based control of I4.0 solutions during assembly. The represented assets like AR devices or tools are defined in the ontology which is part of the assembly level. The conceptual foundations for designing such an ontology are described in the following (see chapter 3.1). The VAR for assembly can be divided into four modules: Manufacturing Assets, Plug&Produce; Traceability; and Interaction [15]. Individual classes of each module are detailed on the use case level. Exemplary classes of the Manufacturing Assets module are product, equipment and personnel (physical assets) as well as processes (intangible assets). Object properties define the relations between individual classes of different modules. The characteristics of individual classes are described by a set of datatype properties.

3.1 Conceptual Foundations for the VAR for Assembly

The semantic repository VAR enables the information exchange from and to diverse agents in the assembly. In order to enhance interoperability with external sources (e.g. legacy systems such as ERP or MES) its ontology’s design has been based on the B2MML standard. It supports the assembly functionality and is a central part of the modular and functional architecture of the A4BLUE Adaptive Framework for adaptive assembly systems.

Functional applications of the VAR ontology can be defined as:

- Adaption: The used assembly devices meet the level of experience/preferences of the worker.
- Interaction: The VAR is connected to other elements of the Adaptive Framework in order to enable information exchange as a basis for the management of events (initiation or reaction).

Non-functional applications of the VAR ontology also comply with general attributes such as:

- Standard-based to support interoperability and reusability.
- Flexibility and sustainability: The ontology is able to evolve along time to correctly represent reality when changes to the assembly system occur. It is able to be easily extended to fit new circumstances, e.g. by adding custom domain concepts.
- Worker assistance and optimal degree of automation: The VAR supports socio-economical evaluations for determining optimal automation configurations by including parameters for cost and worker satisfaction.
- Compatibility and adaptability: The VAR is compatible to various connected components accessing the contained information. The connected components adapt to changing circumstances and as a result, the VAR ontology is dynamic as well.

![VAR level structure](based on 14)

### 3.2 VAR Application Scenario in Electric Vehicle Assembly

To explain the use case level and the application of the framework in combination with the VAR, a few steps of a rear light assembly process of an electric vehicle are exemplary shown (cf. Figure 2). The production scenario of this laboratory use case is characterised by a high number of product variants and low production volumes as well as frequent production ramp-ups. The assembly takes place in the Ramp-Up Factory Aachen of RWTH Aachen University.

The operation starts by scanning a QR-code while the assembly operator wears an AR-device. For this step, the AR guidance needs an interface to the framework. The VAR instantiation includes the level of expertise for a person in a specific type of process, comparable to a skills matrix. Besides this, the operator can choose from a set of skill levels in an adaptable manner. For example, the operator obtains more detailed work instructions by the AR-device if a lower skill level is chosen (novice) than if a higher skill level (expert) is selected. The next assembly step comprises grasping of required rear light components. The AR-device visualises a picking list and the demanded part. The information for instructions are given by the job order and CAD data that are activated via the AR guidance of the framework. After this, the AR-device supports in finding the right position of the rivet nuts while illustrating how to use the rivet tool correctly. The use case shows that assembling the parts is made possible for novice as well as experienced users, enabled by the VAR embedded in the A4BLUE Adaptive Framework.
4. Functionality Evaluation and Critical Reflection

The following evaluation is a critical reflection regarding target fulfillment (cf. table 1) and user-friendliness of the VAR for assembly. The evaluation is based on the gained experience during the A4BLUE project execution at the example of the RWTH Aachen University use case from a production engineering perspective. The evaluation criteria, which can be derived from the acatech Industry 4.0 Maturity Index, are named and described in the following Table 1: [16]

<table>
<thead>
<tr>
<th>Evaluation criteria</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>Describes the operation and modification effort as well as the required skill set for adapting the instantiated VAR in assembly. The rating scale ranges from programming/modeling expert to an intuitive operation by the operator during assembly. The objective is to achieve the highest possible usability by allowing as many people as possible to use and adapt the ontology with lowest possible effort.</td>
</tr>
<tr>
<td>Adaptivity</td>
<td>Consideration of the adaptivity of the I4.0 solutions to the individual employee. Different skill levels should be adjustable. For example, an unskilled assembly operator should receive more detailed instructions than a skilled assembly operator. The criterion assesses whether the VAR does not meet these requirements, meets only partially or completely for all potential properties and characteristics.</td>
</tr>
<tr>
<td>Vertical integration</td>
<td>Considers the system integration capability by providing interfaces across multiple levels of the enterprise system structure. For example, an interface between the ERP and MES systems can be listed here.</td>
</tr>
<tr>
<td>Complexity ability</td>
<td>The criterion describes the ability to model a certain degree of complexity of an assembly system, i.e. the variety of system elements (equipment, assistance systems, processes, product variants etc.) and their characteristics.</td>
</tr>
</tbody>
</table>

The evaluation of the VAR was carried out qualitatively using a five-level item specific scale. It is derived from initial experience with the implementation of the ontology model. The evaluation serves as a basis to improve any identified deficiencies of the VAR and is shown in Figure 3. The prototypical implementation of the RWTH use case is evaluated within the five-digit scale while the best evaluation corresponds to a “5”.

Figure 2: AR-guided rear light assembly process of A4BLUE use case
Figure 3: VAR evaluation based on RWTH Aachen University Use Case

The usability is evaluated with “1”, because ontology instantiation changes have to be performed by means of an ontology editor (i.e. Protégé). As a result knowledge of the used program and access to the programming interface must be available, which is usually not possible for workers in production. Consequently, with regard to usability and adaptability, operation by an expert is required.

Moreover, the adaptivity assessment for the VAR is evaluated as a “4” which means "almost complete". Three levels of worker experience can be differentiated within the VAR (novice, trained, expert). The level selection can be done via an AR application and is user-friendly. However, new levels or changes to them have to be elaborately programmed and scripted.

Since no ERP/MES systems are used in the RWTH Aachen use case, no interfaces are triggered via the current VAR. Hence, the VAR is rated “2” in the evaluation of the vertical integration based on the example of electric vehicle production.

Furthermore, the ontology is arbitrarily expandable and therefore the complexity is scalable. In the beginning not every single assembly application is modeled in classes. However, only classes relevant for the use case are instantiated in the VAR. Because of these reasons, the complexity ability is evaluated as “3” which means medium.

Summarising, the VAR ontology enables adaptive assembly applications in the RWTH Aachen University use case. It hence fulfils the main objective to enable data exchange from and to diverse assembly components to reduce complexity in the assembly process. The operator support works perfectly for the applications implemented. However, to add further functions and devices the knowledge of a programming expert is required. While the vertical integration capability of the VAR is generally given, it was not implemented in this use case scenario. It has been proven in the other A4BLUE use cases though.

5. Summary and Outlook

The VAR embedded in the A4BLUE Adaptive Framework is used to enable intelligent, adaptive assembly systems. These are designed to improve worker satisfaction by providing information based on individual experience, skills and personal preferences. Furthermore, training times for new employees as well as during launch of new products and product variants can be reduced. Process efficiency can be improved by reducing errors in picking of variant parts, tools and connecting elements through interacting information provision systems. In this paper three levels of the VAR for assembly are described (adaptive framework level, VAR level and use case level). The VAR enables the integration of adaptive technological solutions to support workers during assembly. This was implemented and demonstrated in the A4BLUE project - amongst others - by means of AR solutions. In the context of the described use case in electric vehicle assembly, the VAR was critically evaluated. In the future, the VAR could be transferred to other areas within production, such as logistics. Already existing reference architectures as well as the presented adaptive framework need to be enhanced to further facilitate the introduction of I4.0 solutions throughout all areas of production. Further research is needed regarding future staff qualification. With the increasing use of adaptive information
technology, it is important to further examine the impact on working environments. An aspect in this context are the capabilities and skills of an assembly operator being required within a progressively adaptive information technology environment. With regard to the application of adaptive technologies and their underlying semantic oriented frameworks in industrial practice, the quantitative benefits need to be worked out more precisely.

Acknowledgements

This research work has been elaborated within the project A4BLUE (Adaptive Automation in Assembly For BLUE collar workers satisfaction in Evolvable context) which has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 723828.

References

Biography

Prof. Dr.-Ing. Peter Burggraf (*1980) is scientific head of the Department of Factory Planning at the Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University and has been head of the Institute of International Production Engineering and Management at the University of Siegen since 2017.

Matthias Dannapfel (*1986) has been head of department of Factory Planning at the Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2017.

Tobias Adlon (*1989) has been group leader Assembly Planning at the Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2017.

Viviane Hahn (*1991) has been research associate in the department of Factory Planning at the Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2018.

Aaron Riegauf, (*1991) has been research associate at the chair of Production Engineering of E-Mobility Components (PEM) of RWTH Aachen University since 2017.

Patricia Casla (*1968) is Health and Social Technology Coordinator as well as Project Manager at Industrial Management and Innovation Unit of TEKNIKER.

Angelo Marguglio (*1982) is Research Area Manager and Head of the "Smart Industry and Agrifood" Unit of the Research and Development division at Engineering Ingegneria Informatica Spa.

Izaskun Fernandez (*1980) is senior researcher in the Intelligent Information System Unit of TEKNIKER.
Blockchain-based Services in the Machinery and Plant Engineering Industry

Themo Voswinckel, Günther Schuh, Volker Stich, Jan Reschke

Institute for Industrial Management (FIR) at RWTH Aachen University Campus-Boulevard 55, 52074 Aachen, Germany

Abstract

A large number of product-accompanying services in the machinery and plant engineering industry is based on the cross-company exchange of data and information. By providing services, additional sales potential on the manufacturer side as well as far-reaching product and process advantages for applicators can be reached. However, the necessary cross-company exchange of information is nowadays limited due to a lack of trust in the interacting partner and the applicable existing technologies, which results in significant losses in the terms of business potential. The uncovering of this potential now seems to be made possible by the use of the Blockchain technology. Through the key factors security, immutability, transparency and decentralisation, it serves as an enabler for cross-company communication and product-accompanying services. The technological implementation of a Blockchain can take on a broad spectrum of attributes, which can lead to decisive restrictions for the execution of services. This justifies the necessity for a qualified and context-related assessment of service-types-individual specifications and the resulting requirements on the system. Within the scope of this paper, different types of product-accompanying services are identified and analysed regarding their requirements for a Blockchain-based machinery and plant connection. This can serve as a basis for a qualified and goal-oriented configuration of the Blockchain.

Keywords

Blockchain; Digital services; Blockchain-based Services, Machinery and Plant Engineering Industry

1. Introduction

The most efficient use of resources and capacities in production is a fundamental prerequisite for the competitiveness of companies. The machinery and plant engineering industry is under constant pressure to offer their internationally distributed customers reliable products. Increasing competition in the machinery and plant engineering market leads to declining margins, which can be achieved by simply selling the physical product.[1] The solution for creating business potential are product-accompanying services that generate additional value for the applicator based on data generated during the user cycle. For the realization of services, data and information necessarily flow from the machinery or plant operating company to the service providing company. Due to concerns about the misuse of this information, many operators block the corresponding monitoring and remote access to their machinery and plants. [2,3] Blockchain technology combines approaches from academic disciplines such as cryptography, network and database management, and can create technology-based control and trust and thereby function as an enabler for service realization [4,5]. However, the implementation or design of such a Blockchain-based system offers a complex and difficult to obtain prerequisite analysis due to the versatile technical freedom of design such as the structure of data storage, the proof mechanism, access rights or the ability of carrying out smart contracts. The design
of the Blockchain implementation is decisive for its possible success and must be defined in a qualified manner. [6]

These widely differing technical configurational opportunities require a qualified and detailed requirements analysis of the intended service implementation. Within the scope of this publication, different types of data-based services are identified and described with regard to their requirements of a Blockchain-based machinery and plant connection.

2. Identification of data-based services

Based on the classification of assets used in the literature, data-based services can be sorted in the field of intangible goods (see Figure 1). In the context of the machinery and plant industry, data-based services are product-accompanying value creators, which can be offered based on data and information, which are generated during the user cycle of the machinery or plant. [7]

![Figure 1: Classification of assets](image)

All services which are considered within this paper, require production and process data as a base for service execution. Depending on the offering portfolio of the machinery and plant manufacturing company, data-based services can be executed in a one-to-one relationship between the manufacturer and the applier, but as well with third parties as service providing companies.

The following chapter 2.1 describes the characteristics with an impact on the design of the Blockchain-based machinery and plant connection, for the differentiation of data-based services. In the chapter 2.2, different types are diverted.

2.1 Criteria for differentiating data-based services

In the following chapter, the characteristics and specifications for the differentiation of the service requirements are worked out. For the given research purpose, the high number of potential characteristics for the design of a Blockchain-based machinery and plant connection are reduced to those service characteristics, which are actually affecting technical properties. This restriction reduces the number of characteristics for differentiation to a total of five features in the context of Blockchain and three in terms of machinery and plant connection.

2.1.1 Characteristics of data-based services with an impact on the Blockchain

**Transmission content** – The kind of data necessary for executing service is a decisive factor for the implementation and can vary from simply event related notification to extensive control data sets. In the case of services with an extensive data analysis, large amounts of data are necessary for meaningful data analyses [9]. Nevertheless, since machines are generating several gigabytes of data per day, not all data can be shared in a network [10]. Due to this fact, only the data or information should be transferred, to keep the
necessary network traffic low. The unification for separating different quantity requirements is realized in event-related notification, protocols and data records up to extensive control data sets. Whereby a remote control of a machine requires full data records, for certain application only a data transmission in the event of an exceeded tolerance is sufficient.

**Access rights** (Reading right, writing right, identity assignability) – The specification of the service in terms of access rights require characteristics for the differentiation regarding reading rights, writing rights and the identity assignability. Last one describes the possibility, whether the identity of the stakeholder is publicly visible or anonymized/pseudonomized. The properties of the right to read and write are each subdivided into an approval or approval-free authorization right. This characteristic describes whether a third party may access the transaction and its content without prior clearance.

**Latency tolerance** – The latency tolerance describes the temporal offset of the information transfer from the time of creation to the arriving at the responding instance. The specifications vary from uncritical to real time. This characteristic is highly dependent on the degree of interaction which is required by the service. The implication of this characteristic is far-reaching for the design of the Blockchain, as the latency depends heavily on the respective consensus algorithm.

**Plausibility check with context data** – A technical opportunity which can be applied in dependence of the Blockchain realization, is the application of context data related plausibility checks. This characteristic is not a requirement originating from the data-based services, but might support services in their qualitative execution. Due to the historical immutable saving of information in the Blockchain, the service can relate to own historical data as well as external information. The benefit of using contextual data for avoiding wrong decision bases, can be specified into value adding and not value adding.

**Smart Contracts** – Another technical opportunity which is made available by the Blockchain are smart contracts. A smart contract is a software code, which represents a business logic and initiates automatically predefined actions between interacting parties. [11] Smart contracts are web-based computer protocols that map contracts and handle them automatically. Since the reason for service initiation as well as the scope are directly protocolled in the Blockchain, the application got useful aspects for numerous services. The specification of smart contracts are value adding or not value adding.

2.1.2 Characteristics of data-based services with an impact on machinery and plant connection

**Direction of communication** – For most data-based services, the communication direction from machine to network is sufficient, since the output of the service is physically disconnected. In this case, the data and information from the machinery or plant are applied as a decision base. Nevertheless for single services, a remote access is mandatory, which require access to actuators and corresponding control technology for by the service provider. [12,13]

**Measurement principle** – The measurement principle describes the information gathering during the user cycle, which can be specified into event-discrete, time-discrete or continuous. A continuous measurement process is created by an analogue signal, whereas a time-discrete measurement process measures values at regular time intervals. An event-discrete measurement only generates values when a certain event occurs, such as the exceeding of a threshold. [14] The measurement principle, which is related to the required kind of data of the service, restricts the range of possibilities in service execution.

**Type of data** – When designing a machinery and plant connection, it should be determined which data is required for the service and transmitted into the Blockchain. The possible data to be generated during the user cycle of the machinery or plant are process data (technical data as pressure, temperature or flow rate), machine data (Set-up times, production and downtimes, interruptions), order data (number of units or production times) or personnel data (hours worked and absences). [15]
2.2 Types of data-based services

The evaluation of system requirements regarding specific use-cases, four types of data-based services are derived. Based on the differentiation characteristics of data-based services described above in combination with the concrete examples identified by secondary literature, the following types of data-based services are defined. These service types are used to derive requirements for the system design of the Blockchain and the machinery and plant connection. The described characteristics are found in the morphological boxes to describe the differentiations.

2.2.1 Type 1 – Event-discrete services

The first type of services are the “event-discrete services”, which describe services that only require a single information notification for initiation and are executed in an exclusively bilateral relation. Based on a predefined event-discrete trigger, as for example reaching a threshold, a specific mileage or operating hours, process or machine information gets transferred to the service executing party. Examples for this type of services are reactive maintenance, preventive maintenance and object self-service. Reactive maintenance provides the service of the machinery or plant after its failure, whereas for preventive maintenance, operating data is used for precautionary service offers [16,17]. Object self-service describes needs-based operating material provision. The scope and details of the service which gets executed, must be predefined in a service level agreement.

![Figure 2: Characteristics of event-discrete-bilateral services](image)

Due to the predefined service conditions, the unilateral information exchange is exclusively visible for the two parties. The access rights are approved in advance and pseudonyms ensure the secure information access. The latency tolerance of these services is tendentially uncritical, since these services are reactive and the service content is not time critical on real-time or near-real-time level. Since only an information notification is transferred, the data quantity is not sufficient for a plausibility check with external data. Even if the scope of services is predefined, the use of smart contracts may bring added value under certain circumstances.

2.2.2 Type 2 – Operating-data-based services

In comparison to the first type of services, the “operating-data-based services” found on the exchange of more detailed and extensive data as protocols and data records. The broader data can be used for analytical
purposes, where background information is required. Examples for services of this type are predictive maintenance, remote monitoring, data-based component and machine optimization as well as data-based production optimization. The access rights are similar to those from the first type of services, since the services require as well predefined agreements, which follow in a bilateral exclusively data and information exchange. The latency tolerance is still not in the real-time or near-real-time range, but the data transfer should take place within minutes to hours or within one day.

Figure 3: Characteristics of operating data-based bilateral services

For some data based services of this type as those with an optimizing purpose, the context based plausibility check can be applied beneficially, for taking external data as additional data bases. Since the scope and effects of these services are far-reaching, the use of Smart Contracts may also be effective, if automated actions are to be carried out. The communication direction is furthermore unilateral from the machine to the network, since the service content is independent from the operation of the machinery or plant. For some services, event discrete measurement principles are no longer sufficient, as continuous information and detailed historical process-data is necessary. Furthermore, a broader type of data is required to generate the intended value.

2.2.3 Type 3 – Bidirectional services

The third type of services are bidirectional services. These services are characterised by the direct remote access to the machinery or plant by the service provider, for generating added value from remote. Examples for services of this type are remote control and software activation. Remote control describes the functionality to take over the complete control sovereignty for e.g. maintenance work or machinery or plant adjustments. The service of software activation describes the functionality to activate software features and programs, as special manufacturing programs, even after the machinery or plant had been handed over.
Based on this functionality, more detailed data in the form of control data sets must be transferred to the value creating company. These services require near real-time to real-time data interchange with a very low latency tolerance and a continuous measurement of process and machine data. The communication direction is bidirectional with exclusive access and control rights. Smart contracts and plausibility checks are not value adding, since those services take place immediately.

### 2.2.4 Type 4 – Operating data marketplace

The fourth type of service is the operational data marketplace, which is a special case that differentiates fundamentally from the previous types. In this case, the interacting companies are not necessarily known in advance. The concept provides that companies offer their process or machine data unilateral on a marketplace, where third parties can buy them for data analysis without a direct interconnection to the providing company.

---

**Figure 4: Characteristics of bidirectional services**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmission content</td>
<td>Event-related notification</td>
</tr>
<tr>
<td>Access rights</td>
<td>Approval required</td>
</tr>
<tr>
<td>Reading right</td>
<td>Approval required</td>
</tr>
<tr>
<td>Writing right</td>
<td>Anonymized/ Pseudonimized</td>
</tr>
<tr>
<td>Identity assignability</td>
<td>Uncritical</td>
</tr>
<tr>
<td>Latency tolerance</td>
<td>Not value adding</td>
</tr>
<tr>
<td>Plausibility check</td>
<td>Not value adding</td>
</tr>
<tr>
<td>Smart Contracts</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Communication direction</td>
<td>Machine to network</td>
</tr>
<tr>
<td>Measurement principle</td>
<td>Event discrete</td>
</tr>
<tr>
<td>Data type</td>
<td>Process data</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Blockchain</th>
<th>Machinery and plant connection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
<td>Specifications</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Transaction content</td>
<td>Event-related notification</td>
</tr>
<tr>
<td>Access rights</td>
<td>Approval required</td>
</tr>
<tr>
<td>Reading right</td>
<td>Approval required</td>
</tr>
<tr>
<td>Writing right</td>
<td>Anonymized/ Pseudonimized</td>
</tr>
<tr>
<td>Identity assignability</td>
<td>Uncritical</td>
</tr>
<tr>
<td>Latency tolerance</td>
<td>Not value adding</td>
</tr>
<tr>
<td>Plausibility check</td>
<td>Not value adding</td>
</tr>
<tr>
<td>Smart Contracts</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Communication direction</td>
<td>Machine to network</td>
</tr>
<tr>
<td>Measurement principle</td>
<td>Event discrete</td>
</tr>
<tr>
<td>Data type</td>
<td>Process data</td>
</tr>
</tbody>
</table>
The required data is relevant in form of protocols and data records. Since the information and data must be accessible to every interested party, the rights must differ to those of the other services. The reading and writing right must be approval free and the identity behind the information must be shown publicly. The latency tolerance of the marketplace is rather uncritical. Plausibility checks with contextual information can validate the quality and content and raise the market value. The usage of smart contracts enable automated billing in case of access to the shared data.

3. Conclusion and Summary

Against the background of the growing digitalization and networking of machinery and plants, data-based services will play an increasingly important role for companies in the future. For the implementation of data-based services, a wide variety of prerequisites and requirements for a qualified design of the Blockchain-based machinery and plant infrastructure have to be met. In the present paper, a typology of data-based services regarding their technical requirements is carried out. The varying requirements between the different types of services show, that the technical realization needs a precise and qualified analysis of the service properties. Only by a detailed preliminary work and a Use-case specific consideration of the requirements, the fulfilment of the necessary specifications for an efficient system design is realizable. The elaborated analysis is now to be further concretized with regard to the effects on the actual design of the Blockchain-based machinery and plant connection.

References


Biography

Themo Voswinckel, M.Sc. (*1992) is Project Manager and PhD-Candidate at the Institute for Industrial Management (FIR) at RWTH Aachen University. He works in the department of production management and is member of the task force ‘Blockchain for industrial applications’. In the context of his work, he is managing projects in the context of digitalized production and process management in research, consortia and consulting projects.
Prof. Dr.-Ing. Dipl.-Wirt.-Ing. Günther Schuh (*1958) is the director of the Institute for Industrial Management (FIR) at RWTH Aachen University, holds the Chair for Production Systems at the WZL and is a member of the board of directors of the Fraunhofer Institute for Production Technology IPT in Aachen. He is the founder of Schuh & Co. Group as well as founder and CEO of e.GO Mobile AG in Aachen. Professor Schuh is a member of several supervisory boards and boards of directors.

Prof. Dr.-Ing. Volker Stich (*1954) is head of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 1997. Prof. Dr.-Ing. Volker Stich worked for 10 years for the St. Gobain-Automotive Group and lead the management of European plant logistics. His research focuses on operations management, logistics and business applications systems.

Dr.-Ing Jan Reschke (*1984), studied industrial engineering at the RWTH Aachen University. Since 2012 he has been working as a research assistant at the RWTH Aachen and since 2015 as head of the production management department at the FIR. He is also head of the Enterprise Resource Planning Center.
Method to Automatically Create an Initial Layout for a Production System

Matthias Bartelt¹, Fabian Katzwinkel¹, Bernd Kuhlenkötter¹
¹Ruhr University Bochum, Chair of Production Systems, Bochum, Germany

Abstract
A well-proven means to automate processes are industrial robots. Nevertheless, there are still many processes that are not automated, especially in small and medium sized companies. A main reason is a missing automatism to create suitable solutions. To meet this challenge, the ROBOTOP research project tries to find a list of appropriate components and arrange them in a suitable layout. This paper addresses the second step and an algorithm is described that generates a suitable layout. Thereby, a main aim is to generate a layout that is plausible to the user. The problem relates to the facility layout problem and the proposed algorithm is also applicable to non-robotic related tasks. However, current methods do not yield to an appropriate, because due to the simplifications of the used models a manual effort is required when transforming the model to a 3D scene. The algorithms accepts a description of the process and identifies different kind of patterns. For each kind of pattern, layout rules are defined. With this, the model can be transformed to a plausible 3D setup.

Keywords
Facility layout problem; Design pattern; 3D layout

1. Introduction
The aim of the ROBOTOP research project is to develop a platform that enables non-experts to plan robot-based production systems. The user of the platform must not have any robotic knowledge. Instead, an online assistant acquires certain information about the user’s needs. Using a case-based reasoning method, the assistant derives a process description as well as a list of production units. The method also provides a layout hint, but the user may add, change, or remove units. Additionally, the user may start without using the assistant. Hence, the layout starts with a process definition and a list of production units in general. For this, the problem of how to find an initial layout automatically must be solved.

This problem of arranging production units with respect to each other is referred in literature as facility layout problem. A survey of different approaches are given in [1–3]. Because most layout problems are complex and generally NP-hard [4], heuristic methods are used to find a solution. An important sub-problem in facility layout planning is the adjacency problem: which unit should be placed next to which other unit?

A common approach to solve this problem is the method according to Schmigalla [5]. The method uses a triangular mesh to arrange units and yields to an initial solution that is not necessarily an optimum. Other methods try to improve the resulting solution, e.g. by utilizing genetic algorithms [6]. Graph theory is used by another popular approach: the adjacency problem can be solved by constructing a maximum planar weighted graph [7]. Thereby, a unit is described as vertex, material handling from one unit to another is described as edge between the corresponding vertices, and the handling effort defines the weight. With this,
the production can be formulated as a complete, weighted graph \( G = (V, E, w) \) with vertex set \( V \) and edge set \( E \). \( w \) assigns a weight to each edge. With this, a planar sub-graph with maximum weight has to be found. Further approaches are the quadratic assignment problem (cf. e.g. [8]), or slicing-tree-based methods (e.g. [9]).

These methods have usually a simplified view of the units: units are of the same size and the distance of the units is only treated partially within the handling effort. Additionally, transportation to and from a unit can be done in any direction. As a result, a transformation to the real physical environment is difficult. In [10], a software tool for layout modelling and optimization is used to take real sizes into account. Starting with a layout received by the Schmigalla method, the units are rearranged by using a material flow visualization and a distance-intensity graph. However, this rearrangement must be done manually and is not applicable to find a solution automatically.

![Figure 1: Exemplary layout with four units A, B, C, and D received by Schmigalla’s method (a). The arrows drawn in the units indicate the interconnection points, where material must be handled through. Manually changing the layout with respect to the units’ sizes and the interconnection points results in the optimized layout (b).](image)

Furthermore, a main drawback is the simplification of interconnection points, where material is passed to and from the unit. The most units – especially units to automate processes – have certain interconnection points or areas. An industrial robot, for instance, can often not move to its back. Hence, there is an interconnection area in front of and at the side of the robot. Conveyor belts commonly have interconnection points at the beginning and at the end of the belt. CNC machines have a small interconnection area in front of the machine, and so on. Consider an example of four units as depicted in Figure 1a. It shows a layout as received by using Schmigalla’s method. Obviously, the manual layout shown in Figure 1b is a better one, although the adjacencies are the same. Merely, the units are moved and rotated with respect to the interconnection points.

2. Manufacturing and Handling Units

In order to find an appropriate layout, it is essential to create an appropriate model of the units. Like most methods (cf. e.g. [11]), manufacturing units and handling units are handled differently. Consider a manufacturing unit \( m \) as \( m = \{s, I, O\} \) with size \( s \), ingoing interconnection set \( I \), and outgoing interconnection set \( O \). Any interconnection \( i \) can be both, part of set \( I \) and part of set \( O \). Figure 2 depicts a simple example. The unit’s area (8) is most important for the layout process, because other manufacturing units should not cover this area. The machine’s required space (5) is completely embedded within the unit’s area. However, it may be allowed for handling units to overlap with the unit's area but not with the machine’s area. Besides these areas, the algorithm must consider more regions. For instance safety regions, regions to place human machine interfaces, or regions for maintenance. Some of these regions may overlap, some may
not. An overview on different regions and how they may overlap is given in [12]. For now, the algorithm uses a simplified model that contains the unit’s area and the interconnection points, as depicted in Figure 3.

Figure 2: Sketch of the required space for a production unit for the example of an assembly unit. The unit’s area (8) has interconnection points to receive material at (1) and (2) as well as a point to provide material at (3). Each point requires area to place the materials (7). In addition, the machine (4) inside the unit requires a certain area (5) and has a safety region (6). (9) visualizes the overall size.

A handling unit $h$ connects two manufacturing units $m_1$ and $m_2$. In this context, a material buffer is also a manufacturing unit, albeit it manufactures nothing. With this, $h$ can be defined as $h = \{m_j \rightarrow m_k\}$. Handling units do not have a specific size at this point of planning, because their size highly depends on the position of the connecting manufacturing systems.

### 3. Algorithm

Figure 4: Example process with six manufacturing units.

Consider a manufacturing process with $n$ manufacturing units $m_i$. An adjacency matrix $A$ describes the connections by handling units. Figure 4 shows a simple example. The corresponding adjacency matrix is

$$ A = \begin{pmatrix}
0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}. \quad (1)
$$

A one in a row describes an outgoing connection to the corresponding unit. A one in a column describes an incoming connection to the corresponding unit. With row $j$ and column $k$, a handling unit $h = \{m_j \rightarrow m_k\}$ yields to $A_{j,k} = 1$. In order to find a suitable solution, a divide-and-conquer algorithm [13] is utilized. The
algorithm operates on $A$ and will search for units that define serial or parallel structures. Similar to an equivalent circuit (cf. e.g. [14]), a found structure replaces the corresponding units in $A$. The result is a new adjacency matrix $A'$ with a reduced complexity.

### 3.1 Auxiliary Units

In order to identify groups, two auxiliary units are introduced: distributor and merger units. A distributor is a virtual unit with one incoming and two outgoing connection. Thus, material can be distributed from one to two other units. The merger unit works vice versa. It has two incoming connections, but only a single outgoing one. Even though both type of units are used for handling, they are treated as (virtual) manufacturing units. The main purpose of the auxiliary units is to reduce the number of incoming and outgoing connections, which simplifies the structure identification.

![Figure 5: Example process with two added distributor units](image1)

![Figure 6: Alternative option to add the two distributor units](image2)

With this auxiliary units, the example given in Figure 4 can be modified by adding distributor units between $m_1$ and $m_2$ to $m_4$. As shown in Figure 5, the number of connections of $m_1$ reduces from three to one. However, there is not only a single way to add the auxiliary units. Figure 6 depicts another option to add the two distributor units. Obviously, the latter one yields to a structure that can be simplified easily.

![Figure 7: Simplification of the structure by identifying serial structures $s$ or parallel structures $p$.](image3)

As shown in Figure 7, $m_2$ and $m_3$ build a parallel structure and can be replaced by $p_1 = \{m_2, m_3\}$. Next, there is a serial structure with $d_2$, $p_1$, and $m_5$ that is replaced by $s_1 = \{d_2, p_1, m_5\}$. $s_1$ and $m_4$ are now parallel and are replaced by $p_2 = \{s_1, m_4\}$. Finally, there is only a serial structure left.

After adding auxiliary units, the algorithm operates on the adjacency matrix to find serial and parallel structures. In order to find a serial structure, the algorithm search for two units $m_j$ and $m_k$ that are linked with a single connection. Such a pair is found, if the conditions (2) to (5) are true:

\[
j \neq k, \quad (2)
\]

\[
A_{j,k} = 1, \quad (3)
\]
\( A_{j,x} = 0 \) with \( x = 1 \ldots n, x \neq k \), and
\( A_{x,k} = 0 \) with \( x = 1 \ldots n, x \neq j \).

Next the algorithm searches for a unit \( m_i \) that builds a serial structure with \( m_j \) or \( m_k \). If found, there is serial structure with three units and the algorithm searches for another unit that corresponds to the structure, and so on. A set of units that have the same single origin for an incoming connection and the same single target for an outgoing connection defines a parallel structure. Hence, there is a parallel structure between \( m_j \) and \( m_k \) if
\( A_{j,x} = A_{x,k} \) with \( x = 1 \ldots n \),
and furthermore \( \forall l \in \{x \mid A_{j,x} = 1\} \) it is
\( A_{l,x} = 0 \) with \( x = 1 \ldots n, x \neq k \), and
\( A_{x,l} = 0 \) with \( x = 1 \ldots n, x \neq j \).

3.2 Loops

There are two types of loops that may occur in a structure: bypassing and rework of materials. A bypassing structure is sketched in Figure 8a. Material is partially handled from one unit \( m_j \) directly to another unit \( m_k \), while other parts are treated by additional units between. The solution is to add a virtual unit into the bypassing, as depicted in Figure 8b. With this, the algorithm can treat the bypassing structure similar to the parallel structure as described above.

![Figure 8: Bypassing structure (a) and the transformation to a parallel structure (b) by adding a virtual unit \( u \).](image)

The other kind of loop is a rework of materials. If, for instance, a manufacturing unit polishes a product and a subsequent inspection unit verifies the process, the inspection unit may identify the need for a rework. Another example is the use of a carrier system, because at some point in the manufacturing process, the carriers return to the beginning. Figure 9a visualizes this structure, where a material is handled backwards from \( m_k \) to \( m_j \). The solution to resolve this loop is a substitution of the units involved from \( m_j \) to \( m_k \) by a single one, as depicted in Figure 9b. If there are non-serial connection between \( m_j \) and \( m_k \), they have to be reduced first.

![Figure 9: Rework structure (a) and the transformation to a single unit (b).](image)

3.3 Cross-Connections

Consider a structure as sketched in Figure 10. For such kind of structures, the algorithm can identify neither a serial nor a parallel structure. The algorithm must handle such structures in a special way. For this, a further
substitution unit is defined that substitutes the complete cross-connection structure consisting of \( m_1, m_2, m_3, \) and \( m_4 \). The algorithm treats the remaining structure again as described above.

![Diagram of cross-connection](image)

**Figure 10:** Example of a cross-connection.

### 4. Deriving a Layout

The previous section describes how to identify different kinds of structures, e.g. serial or parallel ones. In order to find an appropriate layout, a design pattern is defined for each possible structure. By applying these patterns, the algorithm creates the final layout.

As discussed above, a set of units linked by a single connection defines a serial structure. The interconnection points of the units may be located on different sides of the unit, i.e. on opposite sides, on adjacent sides, or on the same side. Figure 11 depicts design patterns to arrange these types.

![Diagram of serial structure](image)

**Figure 11:** Design pattern for serial structures. The grey regions indicates handling between outgoing connections \( o \) and incoming connections \( i \).

![Diagram of parallel structure](image)

**Figure 12:** Design pattern for parallel structures. The grey regions indicates handling between outgoing connections \( o \) and incoming connections \( i \).

![Diagram of bypassing structure](image)

**Figure 13:** Design pattern for bypassing structures. The grey regions indicates handling between outgoing connections \( o \) and incoming connections \( i \).

![Diagram of auxiliary unit](image)

**Figure 14:** Handling of auxiliary units. The grey regions indicates handling between outgoing connections \( o \) and incoming connections \( i \).

Similar to serial structures, a patterns for parallel structures is defined. Figure 12 sketches the required patterns for units with connections on opposite sides, adjacent sides, and the same side. Obviously, the layout is not
optimal, as the complexity for the handling-system can be reduced. However, as stated above, the main aim is to find a plausible layout, not an optimal.

Both bypassing and rework structures can be treated similar. Figure 13 shows the design pattern for a bypassing structure. The pattern for a rework structure looks similar, except that incoming and outgoing connections are swapped.

The main purpose of the auxiliary units is to identify serial and parallel structures. When designing the layout, they can be ignored usually. However, it may occur that the described pattern yield to an overlap of units. In such cases, the algorithm moves the units until the overlap is eliminated, as depicted in Figure 14.

5. Conclusion

A main problem with existing layout methods is the idealization of material handling. As stated, it is important to generate a layout that is plausible for the user automatically. The presented approach introduces auxiliary units that help to identify especially serial and parallel structures.

For the different structures, design patterns are defined. Each design pattern can be used to replace an identified structure with a suitable layout. After the top-down process of identifying structures, the described algorithm constructs the complete layout by combining the different patterns.

Although the presented method is able to create layouts, there are further improvements possible. By now, the interconnections are considered as a single point. Nevertheless, units usually have an interconnection area. Furthermore, the connections are unweighted. With an appropriate weight, the resulting layout may change. Next, the space between the units must be optimized with respect to e.g. safety and maintenance regions. Finally, the handling paths must be optimized by allowing them to overlap the units' areas.

The main aim was to derive a plausible layout. Thus, the received layout is not optimized. In future work, the algorithm must be improved in order to optimize the current results. Nevertheless, the method yields to a suitable layout that considers interconnection points of the manufacturing units.

Acknowledgements

The research project ROBOTOP is funded by the Federal Ministry of Economic Affairs and Energy (BMWi). It is part of the technology program “PAICE Digitale Technologien für die Wirtschaft” and is guided be the DLR Project Management Agency “Information Technologies / Electromobility”, Cologne.

References


Biography

Matthias Bartelt (*1978) obtained his PhD degree from Ruhr University Bochum in 2017. Currently, he is post-doctoral fellow at the Chair of Production Systems. His specific focus of research interest is on the application of cyber-physical systems, virtual commissioning as well as simulation of production systems.

Fabian Katzwinkel (*1990) worked at the Chair of Production Systems as a student assistant. While working at the Chair, he started his Master thesis with the title Automatic arrangement of automating components. In 2018, he received his Master degree.

Bernd Kuhlenkötter (*1971) was responsible for the product management and technology at ABB Robotics Germany until 2009. In 2009, he took over the Professorship for “Industrial Robotics and Production Automation” at the Technical University of Dortmund. Since 2015, he holds the professorship of “Production Systems” at Ruhr University Bochum. In parallel, he is the vice president of the Academic Society for Assembly, Handling and Industrial Robotics (MHI e.V.).
A methodology for simulation production systems considering the degree of autonomy

Patrick Voit¹, Marie Beller¹, Gunther Reinhart¹

¹Fraunhofer IGCV, Fraunhofer Research Institution for Casting, Composite and Processing Technology IGCV, Augsburg, Germany

Abstract

The increasing number of product varieties and declining product life cycles combined with individualised customer behaviour demand flexible and efficient production systems. A proper solution approach can be the use of intelligent technologies, capable of autonomous processing in order to react rapidly to changing requirements. However, production planners need a profound planning approach for the implementation of such technologies in production systems due to their cost intense investments. Therefore, simulation studies are suitable means for the analysis of a proper degree of autonomy in production systems. An appropriate methodology for the simulation of such systems is presented in this paper. The methodology is aligned with common guidelines on simulation studies and focuses on system analysis, formalisation and simulation. It is based on consistent methods – fact sheets and Value Stream Design for system analysis, Unified Modelling Language (UML) diagrams for formalisation and agent-based simulation. A central contribution to current research is the modular modelling of intelligence skills in production resources and parts in a simulation environment. Consequently, the developed methodology provides a basis for the implementation of simulation experiments in order to facilitate the evaluation of the economically efficient use of intelligent objects in production systems.

Keywords

Simulation Modelling Methodology; Autonomous Production Systems; Agent-based Simulation; Production System Planning

1. Introduction

Market developments like globalisation, mass customisation, declining product and technology life cycles are putting high demands on production [1]. It no longer suffices to produce solely large batch sizes cost efficiently. Moreover, high flexibility and changeability of production structures are necessary in order to cope with these trends and meet fluctuating demands [2]. Therefore, production systems need to be capable of reacting rapidly to external factors, e.g. demand fluctuations regarding product variants, as well as internal factors, e.g. machine capacities and breakdowns – and in the best case even anticipate them. The complexity of planning and scheduling production systems increases with these factors as efficient and at the same time flexible production chains are required [3]. The industry faces this challenge by introducing intelligent technologies such as automated guided vehicles to attain autonomous processes in production systems. Instead of using central planning and scheduling systems, self-reliant entities are applied, which resolve problems locally and thus contribute to the overall solutions [4].
Production planners can apply simulation studies in order to validate the efficient use of autonomous processes. Nevertheless, various factors, such as the individual behaviour of objects, have to be considered. This leads to an increased planning effort in implementation of simulation studies. Therefore, a modelling procedure is presented for simulation of production systems containing different degrees of autonomy in production processes. Beforehand, the paper gives an introduction to autonomous processes in production systems and existing procedures for simulation studies.

2. Autonomous processes in production systems

Within the context of production systems, a manufacturing or assembly system is subdivided into independent subsystems and modules with standardised interfaces. Thereby, it can have a given degree of autonomy to respond flexibly to changes by themselves and adapt to new requirements. Besides the independence from neighbour systems and its environment, it must also be fulfilled the ability of the system to control itself. Thus, it is capable of using these autonomous decision-making and action-performing processes. This self-control ability demands the decentralisation of the control system corresponding to the level of detail of the subsystems and modules. Hence, the degree of autonomy of the system elements depends on two characteristics. First, the given freedom of action by the superior system and second, the capability of the element to make use of this given freedom of action. [5]

The interlinking of these system elements can lead to autonomous processes, e.g. transport or assembly processes in the production system. For this, the objects like parts and resources as system elements “should be able to identify and locate themselves, to sense their environment, to communicate with other resources and parts […] and to control themselves autonomously by using their integrated control system” [5]. Approaches for such intelligent objects are, e.g., parts with integrated RFID devices, automated guided vehicles, intelligent robots or intelligent grippers [5]. Based on the independency and autonomous behaviour of such intelligent objects, flexibility increases in the production system. This can lead to higher capacity utilisation of installed equipment, less personnel costs or shorter throughput times.

Nevertheless, the increased degree of autonomy in the production system can imply higher overall costs. On the one hand, the investments in intelligent objects for autonomous processes requires higher initial costs [6,7]. On the other hand, a high degree of autonomy can lead to higher operating costs [8,9]. Therefore, a certain degree of autonomy has to be determined during the planning process of manufacturing systems in order to select appropriate production resources for the dimensioning of production systems.

3. Methods for simulation studies

3.1 General approach for simulation studies and simulation methods

In order to plan the introduction of intelligent objects on economic grounds, simulations can examine their behaviour and effects. However, there is a wide range of approaches for the simulation of production processes in production systems using different kinds of simulation methods or combining them. Two of the most commonly used simulation methods are discrete-event simulation (DES) and agent-based simulation (ABS). In DES, the system is modelled as a sequence of events, which occur at discrete points in time [10]. Hence, system changes are discontinuous. In comparison to DES, ABS is not an independent simulation type. It is based on existing types as parallel and distributed discrete-event simulation, object-oriented simulation and dynamic micro simulation [11]. It was established in order to study the unique behaviours of individual actors, map their decision-making processes, respectively resulting actions and make valid statements on the emerging global system [12,13]. Therefore, it is the preferred way of simulating autonomous processes as it has strengths regarding the representation of flexibility and autonomy [14].
However, there are a number of steps involved in simulation studies. A generally applicable approach for carrying out simulation studies gives the German guideline VDI 3633. The guideline only differs insignificantly from other procedures (cf. [15–17]). The general phase-oriented approach can be summarised in the phases: task definition, system analysis, model formalisation, implementation, experiments and analysis as well as the parallel-running phase’s data acquisition and data preparation. The steps and their outputs are depicted in Figure 1. The following capture outlines specialised methodologies for simulation of autonomous processes.

![General approach for simulation studies](image)

**3.2 Approaches for simulation studies of autonomous processes**

There is a wide range of approaches for the simulation of autonomous processes in production systems. In this context, simulations are mainly used for the investigation of control methods (cf. [4,19]) rather than in the context of planning production systems. Nevertheless, there are also appreciable simulation modelling approaches. In the following, an excerpt of relevant approaches is given.

For example, Scholz-Reiter et al. [20] present a modelling method for autonomous logistic processes. It contains a requirement analysis by specification of the system as well as methods for business processes and simulation modelling. Another approach delivers Köchling [21]. The presented solution contains a three-step methodology. It is used for integrative planning of self-optimisation behaviour in production systems. Based on the structuring and creation of the simulation model, the operative implementation of self-optimising behaviour is facilitated. Besides both methodologies, the procedure of Lass can also be mentioned [22]. The methodology uses a hybrid simulation approach in the form of a model factory. It consists of virtual and real elements. Thus, it is possible to model cyber-physical production systems and implement a scenario. Thereby, the scenario can be individually tested in the simulator to adapt cyber-physical systems.

From the reviewed simulation approaches, further potential can be identified. Existing methodologies do not suffice to replicate the characteristics of autonomous behaviour. Thus, it is difficult to draw conclusions on the system as a whole. Additionally, the approaches are economically inefficient in the early planning stage as their development is highly time consuming or an expensive investment is necessary for model factories. Therefore, the objective is to develop a simple yet effective method of simulating autonomous behaviour in the context of production system planning. Thus, it should provide a basis to facilitate the evaluation of the economically efficient use of intelligent objects in production. Furthermore, it should ensure a consistent combination with methods for describing and formalising production systems.
4. Simulation modelling methodology

4.1 Concept of the modelling methodology

The developed methodology focuses on the main steps of simulation studies according to the mentioned German guideline VDI 3633. It incorporates extended approaches regarding the modelling and simulation of production systems considering the degree of autonomy. Figure 2 assigns these specific methods to the respective steps in the simulation study approach.

Based on the task definition, which defines the tasks suitable to attain the goals regarding the planning of autonomous processes in production systems, the system analysis can be started. This analysis step comprises the definition of system boarders, the structure, relationships and characteristics of its elements, and the description of the process structure as well as its basic principles. For this purpose, the methods of Value Stream Design and fact sheets are used to describe the characteristics of production systems objects (resources and parts). The next step is the model formalisation by displaying the system elements and relationships in a formal representation. In this context, Unified Modelling Language (UML) diagrams as a static representation of the production system, its objects and relationships are used. Afterwards, the formalised model can be transformed into an executable simulation model as part of the model implementation phase. A multi-agent simulation serves as a simulation platform in order to represent the dynamic and autonomous behaviour of the production system and its elements. Finally, parameters and variable values can be implemented into the simulation model for purposes of experiments and analysis. The experiments can be executed with varying parameter values or structural adaptations of the model before the individual experiment can be analysed regarding economical aspects. In the following, the individual methods to the respective steps are presented and required extensions for autonomous processes are discussed.

4.2 System analysis

Based on the initial task definition, the production system planner can conduct the analysis of the production system. The system analysis consists of both: Value Stream Design and fact sheets. The method of the Value Stream Design can be used in order to visualise the process chains of the production system. The defined symbolism of the method helps to represent the logistical interactions of the system elements like inventory, supplier, customer, production processes as well as material and information flows [23]. However, the information flows are just demonstrated on an aggregated level. Therefore, an extension of the Value Stream Design is necessary to describe production systems containing different degrees of autonomy in processes.
Such a modified Value Stream Design is demonstrated by Theuer (cf. [23]). Next to Value Stream Design, fact sheets are applied for the textual, skill-based description of objects (resources and parts) in the production system. Thereby, it is possible to carry out the object’s ability to participate in autonomous processes based on their intelligence level.

Figure 3 depicts an exemplary resource fact sheet. It describes an intelligent cleaning machine for the powder removal of metal laser-melted parts. The fact sheets for resources contain basic information and a list of operating and intelligence skills. The characterisations of the operating skills are founded on guidelines like VDI 2860 or DIN 8580, whereas the seven intelligence skills are defined on the aggregation of existing classifications describing intelligent objects (cf. [24]). The classification is carried out according to an allocation matrix connecting the possession of certain architecture elements with feasible intelligence skills. Exemplarily, the skill taking Action requires the architecture elements energy supply, communication interface, data storage and processor, as well as actuator. Analogous to the fact sheets for resources, those for product orders comprise product features and intelligence skills. However, products do not possess actuators that are necessary for intelligence level 6 and 7. Hence, only intelligence level 1 to 5 can be attained.

![Figure 3: Fact sheet of an intelligent cleaning machine for powder removal](image)

### 4.3 Model formalisation

After the production system analysis, the system elements and their connections together can be structured into a formal model. This model forms the basis for the implementation of the system into an executable simulation model. The most commonly used formalisation methods for production systems are petri-nets and Unified Modelling Language (UML) diagrams. Both are suitable in the context of production systems planning [25,26]. They have a graphical presentation, formal notation and are capable of mapping hierarchies, (de)compositing elements and illustrating dynamic time behaviour. In addition to petri-nets, UML diagrams allow demonstrating the modular structure of production systems. Thereby, it facilitates the transformation of the structure model into an executable simulation model [27]. Therefore, UML is the preferred method to illustrate and analyse the objects, their structure and relationships in the production system graphically and formally.

Figure 4 shows the UML diagram for intelligent resources in general. Based on the fact sheets, the diagram comprises basic information, the basic structure as well as further architecture elements related to carrying
out skills. Organisational, environmental and economic information contain data for the parameters and variables of the simulation. The basic structure possibly in combination with energy supply, sensors and actuators enables operating skills. Furthermore, intelligence skills are enabled by the energy supply, identifier, information storage, information processor, communication interface, sensors and actuators. The more kinds of these architecture elements are given and intelligence skills are enabled, the higher is the intelligence level. The fact sheet for parts is similarly structured but doesn’t contain operating skills and actuators.

![UML class diagram of the generic architecture of an intelligent resource based on [28]](image)

Besides the formalisation model of objects, a formal model of the entire system has to be developed as only the interlinking of multiple objects enables autonomous processes in the production system. Hence, the UML class diagram of the production system demands process functions, besides objects and relations. The function are necessary to represent the autonomous processes which are enabled by the resource skills. Moreover, the resulting production costs can be allocated to the produced products and their features. The first figure in chapter 5 illustrated an exemplary formal model of a production system.

### 4.4 Model implementation and experiments

After the model formalisation, the model can be implemented into an agent-based simulation model for purposes of experiments and analysis (cf. chapter 3.1). Figure 5 illustrates exemplarily the schematic structure of such a simulation model for the software environment AnyLogic. The model is divided into three interconnected levels: the main, the database and the agent. The main level represents the process chains containing the process costs and the corresponding resource pools. In comparison, the database comprises the relevant parameters of the orders and their product features as well as the resources and their skills. The level agent is used to model different agent types of objects - orders and resources - including their specific behaviour realising autonomous processes.

In order to implement the objects’ behaviour based on their intelligence skills, the following modules of the software AnyLogic have to be used alone, in combination or extended:

- **Parameters** are used for agent (or system) information, which is constant during simulation time, e.g. resource capacities or hourly rates.
- **Variables** should be used for information that changes over time, e.g. an agent’s location.
- **State charts** allow the agent to monitor its current state by itself. State changes are triggered by message exchanges between levels or agents. Furthermore, state changes can cause another message exchanges initiating further state changes.
- **Statistic objects** monitor and save data on the system’s or the element’s behaviour. Therefore, it can be used for KPIs and represent the basis for decision-making processes, e.g. throughput.
- **Database** interfaces for example in form of the Excel spreadsheets serve as an external data storage similar to cloud solutions in real life.
- **Functions** enable modelling of complex agent-based decision-making processes by inserting customised code blocks and assigning them to certain agent types.

![Figure 5: Structure of an agent-based simulation model with the main level, agent level and database](image)

5. **Exemplary application of simulation modelling methodology**

The developed simulation modelling methodology was applied for the use case of a powder removal process for metal laser-melted parts. The simulation study had the goal to evaluate different degrees of autonomy in the powder removal processes. Therefore, several production resources were firstly investigated regarding their intelligence level as well as their interlinking with each other. Based on this system analysis, the production process was formalised into a UML class diagram (Figure 6). The cleaning machines (cf. Figure 3), AGVs and robots containing different intelligent skills are linked to an **autonomous production system** and interact via communication interfaces. The intelligent products also interact via communication with the system. Furthermore, workers can communicate with resources and products via the user interface. Therefore, a stationary terminal exists for the cleaning system and a mobile terminal is used for AGVs, robots and products. The post-process in the context of additive manufacturing uses the operating skills **swivel, position and convey** enabling the functions **cleaning, handling or transporting**. These functions change the **product’s state** regarding location or production status as well as allocate the incurring **costs** to the relevant order. The **parts’ features** determine the required operation skills of the resources.
In the next step, the UML diagram was implemented into the simulation environment AnyLogic. Figure 7 shows the visual representation of the simulation model. The cleaning process contains the machines for cleaning, transportation and handling as well as workers operating for manual tasks. In this context, experiments were applied by using parameter variation. Thus, different degrees of autonomy in the cleaning process could be examined in order to determine the optimal mix of intelligent resources. The evaluation of the experiments is based on average cost per unit considering the restriction of maximum throughput time.

6. Summary and Outlook

Production planners can apply simulation studies to validate the efficient use of autonomous processes. However, different aspects have to be considered in the studies implementation phase as the degree of autonomy depends on two factors: First, the given freedom of action by the system and second, the objects’ intelligence skills enabling the use of this freedom. In order to reduce the implementation effort, the paper presents an adequate methodology for the simulation of (semi-) autonomous production systems.

The methodology is based on the German guideline VDI 3633 and contains the four main steps: System analysis, model formalization, model implementation and experiments. At first, the method of Value Stream Design and fact sheets are used to analyse the system and its objects. Afterwards, the analysed production system should be formalized in UML diagrams. Thus, the resulting formalized model can be transformed into an agent-based simulation model, which is applied to evaluate the autonomous system behaviour.

In addition, further potential was identified. On the one hand, the focus on flexibility and changeability could be increased by including scenario analysis in order to investigate stochastically alternate developments and events. On the other hand, production scheduling has a high impact on the system’s performance. Therefore, researchers are advised to analyse the impact of differing scheduling methods.
References


Biography

Patrick Voit (*1990) is research assistant in the group planning of manufacturing systems at Fraunhofer Research Institution for Casting, Composite and Processing Technology IGCV. He has a M.Sc. in the field of Industrial Engineering and Management and is pursuing a Ph.D. degree in Mechanical Engineering at the Technical University of Munich.

Marie Beller (*1993) is a M.Sc. candidate in Management & Technology at the Technical University of Munich. She has a B.Sc. in the field of International Business Studies from the Otto-Friedrich University Bamberg and studied at the Clemson University in South Carolina.

Gunther Reinhart (*1956) is the holder of the chair for Industrial Management and Assembly Technology of the Institute for Machine Tools and Industrial Management (iwb) at the Technical University of Munich. Prof. Dr.-Ing. Reinhart is also the chairman of the Bavarian Cluster for Mechatronics and Automation e.V and director of Fraunhofer Research Institution for Casting, Composite and Processing Technology IGCV.
Sheet-Metal Production Scheduling Using AlphaGo Zero
Alexandru Rinciog\textsuperscript{1}, Carina Mieth\textsuperscript{2}, Paul Maria Scheikl\textsuperscript{3}, Anne Meyer\textsuperscript{1}

\textsuperscript{1}Chair of Enterprise Logistics (LFO), TU Dortmund University, Germany
\textsuperscript{2}TRUMPF GmbH + Co. KG, Ditzingen, Germany
\textsuperscript{3}Institute for Intelligent Process Automation and Robotics (IPR), Karlsruhe Institute of Technology, Germany

Abstract
This work investigates the applicability of a reinforcement learning (RL) approach, specifically AlphaGo Zero (AZ), for optimizing sheet-metal (SM) production schedules with respect to tardiness and material waste. SM production scheduling is a complex job shop scheduling problem (JSSP) with dynamic operation times, routing flexibility and supplementary constraints. SM production systems are capable of processing a large number of highly heterogeneous jobs simultaneously. While very large relative to the JSSP literature, the SM-JSSP instances investigated in this work are small relative to the SM production reality. Given the high dimensionality of the SM-JSSP, computation of an optimal schedule is not tractable. Simple heuristic solutions often deliver bad results. We use AZ to selectively search the solution space. To this end, a single player AZ version is pretrained using supervised learning on schedules generated by a heuristic, fine-tuned using RL and evaluated through comparison with a heuristic baseline and Monte Carlo Tree Search. It will be shown that AZ outperforms the other approaches. The work’s scientific contribution is twofold: On the one hand, a novel scheduling problem is formalized such that it can be tackled using RL approaches. On the other hand, it is proved that AZ can be successfully modified to provide a solution for the problem at hand, whereby a new line of research into real-world applications of AZ is opened.

Keywords
Production Scheduling; Sheet-Metal Production; Job Shop Scheduling Problem; Reinforcement Learning; Monte Carlo Tree Search; AlphaGo Zero

1. Introduction
1.1 Motivation
The reasons for focusing on sheet-metal (SM) production scheduling are twofold. On the one hand, SM products are ubiquitous. The spectrum of SM products is very large, ranging from industrial to household items. As such, successfully optimizing the production process would be of high impact. On the other hand, the problem is only summarily studied [1], [2] and of considerable difficulty.

The input to a sheet-metal production system is a stream of product specifications, where each product has a variable number of associated constituent parts, a monetary value and a deadline. To complete a product, parts are batched onto a metal sheet from which they are then separated, bent into three-dimensional shapes and assembled together. Parallel processing resources are available for each step. During batching, material waste occurs. If deadlines are missed, the product value is penalized. The goal of a production scheduler is to map operations to resources, such that both tardiness and material waste is minimized.

DOI: https://doi.org/10.15488/9676
SM production scheduling is most closely related to the job shop scheduling problem (JSSP) with routing [3], [4], albeit considerably more complex and presenting a unique challenge: Since cutting operations occur in dynamic batches, whose processing times are only available after batch definition, scheduling is interwoven with a two-dimensional packing problem. Additionally, SM production systems are capable of processing a large number of highly heterogeneous parts simultaneously. Mapping the SM production reality to a JSSP would lead to far larger instances than any available in benchmark data sets [5].

Commonly, high dimensional JSSPs, are solved by selectively searching the space of all possible schedules using mainly local search and genetic algorithms (GA) [6]–[10]. More recently, reinforcement learning (RL) approaches for solving JSSPs started to gain momentum. This is because RL presents some advantages over GA/search and other approaches: As per Waschneck et al. [11], global transparency and global optimization are among the advantages offered by RL solutions. Global transparency pertains to scheduling systems’ ability to monitor the production state as a whole, as opposed to merging several local optimization strategies. Global optimization describes the system’s capability to jointly optimize different goals. Given the interdependence between batching and scheduling, both properties are very desirable for SM production scheduling. AZ has the added advantage of incorporating search with RL, thereby alleviating the sample inefficiency problem deep RL approaches can suffer from [12].

1.2 Related Work

**Reinforcement Learning:** RL has been applied to small JSSPs with varying optimization targets and degrees of success since the mid-90s [13]–[19]. More recently, Reyna et al. [20] uses a value based Q-learning approach on a simple JSSP with makespan as the optimization target. The RL solutions are tested against optimal solutions from the OR-Library benchmark data set (up to 20 jobs and 5 machines). Qu et al. investigate the adequacy of multi-agent Q-learning on top of an ontology in the context of a multi-objective dynamic flow shop scheduling problem [21]. Waschneck et al. embed the Atari deep Q-learning algorithm [22] in a multi-agent system aimed at optimizing the uptime utilization in the semiconductor industry [11]. The trained system fails to outperform the heuristic baseline. In [23], [24], policy learning, namely Trust Region Policy Optimization [25] is used to solve the complex JSSP of the semiconductor industry. Resource utilization and lead time provide the concomitant optimization goal. The RL algorithm surpasses a heuristic approach on both optimization targets.

**Supervised Learning:** Another AI solution to complex scheduling problems is offered by supervised learning (SL) as investigated in [26]. The authors use a data set created by human demonstrators and a synthetic data set created using heuristics for binary classifier training. The binary classifiers are used to discriminate high from low priority scheduling tasks given the two tasks and the scheduling problem’s state. Using the trained models, tasks are ranked through pairwise comparison. The approach achieves a high task assignment accuracy of over 90%. The most recent supervised learning approach to solving the JSSP w.r.t. makespan was taken in [27]. Here, the authors generate optimal solutions for a 6 machines and 6 jobs JSSP. They then train a linear classifier to learn a dispatching rule given the production state. While the results obtained through SL are proved to be superior to heuristics, generating training data for large JSSPs is intractable.

**AlphaGo Zero:** The first version of AlphaGo, introduced in [28], combines SL, RL and a probabilistic search strategy, namely Monte Carlo Tree Search (MCTS) [29]. Through SL, human expertise is incorporated in a neural network (NN). Through self-play using RL and MCTS, the NN is further tuned until it outperforms human players at Go. The second version [30], named AlphaGo Zero (AZ), simplifies both the self-play algorithm and the training process considerably: Here, the algorithm relies solely on RL for training. In [31], the authors use a slightly modified version of AZ, to outperform the respective state of the art for Chess and Shogi, in addition to Go. This constitutes proof of AZ’s game agnosticism.
1.3 This Work

The SM scheduling problem can be formalized as a JSSP with routing flexibility, cutting constraints, assembly constraints, (dynamic) batch processing capabilities and an additional optimization target. Jobs correspond to SM parts and consist of 3 operations: cutting, bending and assembly. The set of jobs is divided into partitions corresponding to SM products. In SM-JSSPs, operations are fixed to a machine type, whereby routing flexibility ensues. Standard JSSPs limit machines to processing one operation at a time in the absence of preemption. This applies solely to bending machines in the SM-JSSP context. Cutting operations can be processed simultaneously, as long as the associated parts fit on the same sheet of metal (cutting constraints). Assembly operations of jobs from the same partition must be processed together (assembly constraints).

While the processing times for bending and assembly operations are known a priori, the duration of cutting operations is only known post the dynamic batch definition, i.e. during schedule computation. Figure 1 provides a representation of the main aspects of the SM-JSSP. Besides the common JSSP optimization targets (e.g. makespan, flow/lead time, tardiness [32]), the minimization of material waste incurred during cutting is additionally considered in SM-JSSPs (additional target).

![Figure 1: A representation of the SM-JSSP. A stream of products with variable number of parts, a monetary value and a deadline are the SM-JSSP input. Processing starts with batching parts onto a metal sheet. After cutting and extracting the parts, bending operations are executed. All parts from a product are assembled together in a final step. Bending and assembly times can be computed a priori per part and per product respectively. Cutting times vary depending on the batch content.](image)

This work focuses on jointly minimizing tardiness and material waste for medium-sized, offline SM-JSSP instances using AZ. Since this is a preliminary study, we defer some supplementary sources of complexity to future work. For now, we do not consider setup times, machine/worker availability and transportation times. Furthermore, we flatten the two dimensional packing problem into one dimension, i.e. the cutting constraints now simply state that the summed areas of all parts batched on a sheet must be smaller than the sheets’ area. We also limit ourselves to considering SM-JSSP instances, which, while large with respect to the JSSP literature, are not large enough to accommodate the larger SM production systems. The exact instance size is presented in Section 2.3.

In SM production, despite advanced planning systems (APS) being used for coarsely defining production plans, daily scheduling decisions on the shop floor are often done by human experts using simple heuristic solutions such as earliest due date (EDD) to prioritize operations. This is because APS plans can quickly become obsolete given unforeseeable events. The solution we investigate as an alternative involves modifying AZ to a single player version, training its integrated NN in a supervised fashion on heuristic solutions to offline SM-JSSP instances and fine-tuning the NN through self-play. Our scientific contribution is twofold. First, we design a new state formalism for a complex new scheduling problem and integrate it in an environment, which can be used with RL methods. Secondly, we show AZ to be able to incorporate and
outperform a heuristic approach dominated by the EDD. Priority ties are broken by supplementary criteria (see Section 2.3). MCTS was added as a supplementary comparison baseline.

This paper is structured as follows: Section 2 presents the inner workings of AZ in Subsection 2.1, followed by the RL design for SM-JSSP in Subsection 2.2, as well as our experiment setup in Subsection 2.3. After a brief discussion of our results in Section 3, we present our conclusion in Section 4.

2. Methods

2.1 AlphaGo Zero

The data structure central to AZ is a tree with nodes corresponding to game states $s$ and edges corresponding to actions $a$. Each edge in the tree stores the probability $P(s, a)$ of it being the best action in the state $s$ a visit count $N(s, a)$ and the average expected result $Q(s, a)$ of taking action $a$ in state $s$. Nodes store the expected result over all actions possible from state $s$.

Nodes store the expected result over all actions possible from state $s$ as $V(s)$. Below we describe how moves are selected using NN and MCTS, whereupon the NN training scheme is detailed.

**Neural MCTS:** Given a state $s$, AZ picks a move repeating four phases $m$ times followed by a final move selection. The four phases are selection, expansion, evaluation, and backup as shown in Figure 2. These correspond to the MCTS phases modified to be guided by a NN. During the selection phase (Figure 2a), nodes reached over edges maximizing

$$U(s, a) = P(s, a) CPI + \sum_b N(s, b) / N(s, a)$$

are selected until a dangling edge is encountered. In the Equation (1), $\sum b N(s, b)$ is the cumulative visit count of all outgoing edges from $s$ and $CPI$ is a tunable exploration parameter. At first, actions with high probability and low visit count are preferred. Asymptotically, high valued actions are preferred. By increasing $CPI$ this asymptotic transition is slowed down. Assuming the current iteration is $n$ and $s_L$ is the leaf node reached in iteration $i$ the current leaf node $s_L$ reachable over the selected dangling edge, as well as its egress edges are added to the tree. This constitutes the expansion phase. During the evaluation phase, the NN $f_n$ is used to evaluate the new node $s_L : P(s_L, \cdot), V(s_L) = f_n(s_L, \cdot)$. First $P(s_L, \cdot)$ is used to update the egress edges from $s_L$ (Figure 2b). Then the backup phase ensues (Figure 2c). Herein, all edges up the selection path are updated. This update implies incrementing the edge visit counts $N(s, a)$ and setting their values $Q(s, a)$ to the average accumulated value up to the current iteration. This is described by Equation (2), where $1(s, a, i)$ is 1 if the $s_L$ is reachable over the edge $(s, a)$ and 0 otherwise:

$$Q(s, a) = \frac{1}{N(s, a)} \sum_{j=1}^n Q(s, a, i) V(s_L).$$

After the predetermined number of MCTS iterations the final move selection is performed. The visit counts of the root edges are exponentiated with $1 / \tau$, where $\tau$ is an exploration parameter, and normalized to create a probability distribution over the legal actions relative to the root node. The final move is then selected by sampling from this distribution. Note that infinitesimal values of $\tau$ induce a Dirac distribution, while high values of $\tau$ induce an asymptotically uniform distribution.

**AZ Training Loop:** The NN needs to be trained to estimate the move probabilities $P(s, \cdot)$ and the expected value of a node $V(s)$ accurately. AZ training has two steps, which are repeated until the model weights saturate. These are self-play and neural-network training. During self-play (Figure 3a), the MCTS scheme introduced above is used to play games from start to finish, i.e. iterations 1 to $\tau$ by sampling from probability
distributions $\pi_t$ returned by the MCTS algorithm for states $s_t$. States and the corresponding MCTS action probabilities $(s_t, \pi_t)$ are stored for every performed move. When a terminal state $s_T$ is reached (i.e. end of the game), the reward $z := r(s_T)$ is used to form triples $(s_t, \pi_t, z)$. The values of $z$ depend on how the reward function $r$ is modelled, e.g. for chess, $z$ is either -1, 0 or 1 for loss, draw or win respectively, for Go $z = \pm 1$. The AZ network can now be trained on them using stochastic gradient descent to minimize the loss function $l = (z - \hat{v})^2 - \pi^T \log \hat{\pi}$ where $(p, \nu) = f_\theta(s)$ (Figure 3b).

Figure 2: Neural MCTS: a Nodes are selected recursively by traversing edges corresponding to an action $a = \arg \max_a Q(s, a) + U(s, a)$ until the egress edge of a leaf node; b A new node $s_k$ is added to the tree, $P(s.), V(s_k)) := f_\theta(s_k)$ is evaluated and its egress edges are updated with probabilities $P(s,.)$, c Action values $Q$ are updated up the tree path using the mean of all state values $V$ stored in the nodes. Source: [30]

Figure 3: (a) Games are played from start to finish, using neural MCTS to select a move. The move probabilities $\pi_t$ are stored together with the corresponding state $s_t$. When the end-state $s_T$ is reached, the reward $z$ is associated with every pair $(s_t, \pi_t)$. (b) The network $f_\theta$ is trained using stochastic gradient descent to minimize a combination of mean squared error on the value head and cross entropy on the policy head. Source: [30].

2.2 RL Design for SM-JSSP

To apply AZ to the SM-JSSP, the agent environment interaction needs to be redesigned. This is done by creating the SM-JSSP state and action space together with a reward signal corresponding to the joint objective function of material waste and tardiness. Additionally, the NN architecture is adapted.

The following state space was used for SM-JSSP instances of $m_1$ cutting, $m_2$ bending and $m_3$ assembly stations and up to $nk$ jobs. Processing stations are indexed using the set $M := \{1 \ldots m_1 + m_2 + m_3\}$, where 1 to $m_1$ correspond to cutting machines, $m_1 + 1$ to $m_1 + m_2$ correspond to bending machines and so on. Jobs are represented through the indices of a $nk \times k$ matrix, where every row corresponds to a job partition. To each job we associate 2 descriptors: Area and outline. Both are needed to compute the dynamic processing times for the cutting (and extraction) operations. Areas are also needed to enforce area constraints. The material waste is stored per cutting machine in a vector of length $m_1$. Operation processing times are stored per job -
To track remaining operations, for every job, the index $i \in M$ of the last machine a contained operation was assigned to, is noted down. Additionally, the state encodes whether the current machine $i$ is actively processing the corresponding operation (1) or whether the operation has finished (2) for every job. The state also encodes the remaining slack time and value for every job partition. Finally yet importantly, the idle machine onto which an operation needs to be scheduled next is encoded. This yields a state space of $5nk + 3n + 2m_1 + 1$ entries.

The AZ agent schedules by interacting with a deterministic event discrete simulation: Whenever a machine is idle, the agent is asked to schedule an operation to it. The environment maps the operation, selects the next free machine for a decision and updates the state representation accordingly. If no decision is possible at the current time, the state is rolled forward in time by marking the operation with least remaining processing time as finished, updating all the time variables and freeing the corresponding machine. Cutting batches are defined iteratively by requesting operations for the same machine until the agent produces a finish flag. Assembly operations are triggered by the agent outputting the index of the first job in a job partition. Bending operations are triggered by a job index. As such the action space is given by $nk + 1$.

To compute the SM-JSSP reward, the environment keeps track of wasted material and the tardiness for every job partition. After all operations were scheduled, $r_{abs} := -c(W) + \sum_{i=1}^{n} v_i - \lambda \max\{0, T_i\}$, where $c(W)$ is the cost of the total used material (including waste), $T_i$ and $v_i$ are the tardiness and value for the job partition $i$ respectively and $\lambda$ is parameter punishing tardiness. $r_{abs}$ reflects the sum of product values, discounted proportional to deadlines, minus the total material cost. The reward is scaled to $[0, 1]$ using the maximum possible score $r_{max}$ i.e. no tardiness and no waste, $r_{rel} := 1 - \frac{r_{abs}}{r_{max}}$.

Both move probabilities and expected value from a state are provided by a single NN as in [28]. As opposed to [28] however, a simple feed-forward NN is used for the SM-JSSP instead of ConvNets. This is because we do not stack the input states, and there is no geometric correlation in the SM-JSSP states. Note that since there is a strong temporal correlation between states, passing a stack of states to a ConvNet could be beneficial.

### 2.3 Experiment Setup

To validate our approach, we create 80 different offline scheduling instances and use AZ, the EDD heuristic, as well as an MCTS implementation as per [29] to find a scheduling sequence. Then we compare the respective results in three ways: First, we plot the relative scores achieved to get a rough assessment of the scheduling behavior. Secondly, we count the number of times a particular scheduling approach provided the best result among its peers. Lastly, we average the 80 obtained scores for each individual scheduling scheme. We chose the EDD baseline as it is the most intuitive solution for a scenario where tardiness is to be minimized. Since due dates can be the same for different parts, ties are being broken by higher product value, larger area and higher number of bending steps. Whenever a machine is free, the operations that can be assigned to it are prioritized as per EDD and the one with the highest priority is scheduled to it.

The SM-JSSP instances are created as follows. The set of jobs $j$ contains 10 job partitions. The number of jobs per partition is sampled from $U(1,7)$ Areas, outlines and number of bending steps are sampled from the uniform distributions $U(1dm^2, 164.26dm^2)$, $U(1dm, 122.83dm)$, and $U[0,30]$ respectively. The resulting one dimensional histograms are used to create a three dimensional histogram from which part descriptors are sampled uniformly at random for each job. These part descriptors are the independent variables needed to calculate the processing times for the different operations. The sheet area is fixed to 589.19$dm^2$. Processing times for cutting are calculated on batch definition using a nonlinear estimator with the summed part areas to sheet area and summed part outlines to sheet outline ratios as input. Bending times are a linear function
of the area and the number of bends per part. Assembly times are a linear combination of the average part area, the average number of bends and the total number of parts in a job partition.

In terms of the AZ implementation, we use a fully connected feed forward NN with 2 layers à 1024 and 512 nodes respectively, a dropout of 0.3 to avoid over-fitting, batch-normalization to speed up training, and Adam for learning rate optimization. During self-play, \( \tau \) is set to 1 for the first 15 decisions and then dropped to 0. During evaluation \( \tau = 0 \). \( c_{\text{puct}} \) is 1.5 throughout. Prior to any decision, the selection, evaluation, expansion and backup steps are repeated 6 times. The tree is kept until the end-state is reached.

The three scheduling techniques considered were implemented in python. The environment implementation is consistent with the openaiGym (https://gym.openai.com) API for RL. The AZ agent was implemented using Keras with a TensorFlow backend for the embedded NN and was trained using 12 CPU cores for parallel execution of self-play episodes and 2 GPUs for NN training.

Rather than training AZ from scratch, we pretrain it to mirror the EDD heuristic first. To that end, we use the random SM-JSSP instance generation scheme described above to generate \( 10^6 \) (state, decision, reward) triples on which we then train the NN until a validation accuracy of 90% is achieved for move probabilities. To check the learned behavior, we do a small evaluation round on 30 SM-JSSP instances at this point as well. Thereafter we run the AZ RL training pipeline for 254 iterations, which takes about 5 days. For each iteration, data from 80 self-play episodes is added to a replay buffer of the 800 most recent SM-JSSP games, on which the NN is then trained.

3. Evaluation

Scheduling a SM-JSSP instance to completion (circa 70 moves) takes about 150 seconds for both MCTS and AZ. Note that, while the self-play time is currently comparable, AZ should scale much better than MCTS with the size of the scheduling instance. For a game of depth \( n \), MCTS runs \( n \) simulations to completion for every move selection step, while AZ simply makes \( n \) calls to the network’s predict function. As \( n \) increases, so does the simulation time, while the call to the predict function stays constant.

Figure 4 shows the performance of AZ, EDD and MCTS before and after AZ RL fine-tuning side by side. We elaborate on the aspects showcased by the figure. As a first observation we note that the performance of MCTS alone is quite lacking, compared to the heuristic approach or AZ. It registers both the lowest average score and the lowest number of wins in both evaluation rounds. This is to be expected given the small number of rollouts for a game tree as vast as the one corresponding to the SM-JSSP. Nevertheless, in 3% and 7% of post pretraining and post RL finetuning cases respectively, MCTS does find the best schedule.

Secondly, the scheduling behavior of AZ and EDD seems to be quite similar, as can be seen in the lineplots in Figure 4. This is not at all surprising given that EDD is a fairly good heuristic and AZ’s network was pretrained on it. We notice that the overlap between the AZ and EDD curves is significantly higher immediately after pretraining than after RL fine-tuning. This suggests that AZ has developed strategies additional to EDD during self-play. The lineplots also reveal just how different the considered scheduling instances are, with best solutions achieving scores between 0.4 and 1 of the maximum producible value.

Thirdly, and most importantly, training AZ using RL leads to it outperforming its teacher, EDD. Immediately after pretraining however, the best scheduling approach, as ranked by both the number of wins and the average score, was EDD.
Figure 4: Comparison between AZ, MCTS and EDD using the relative score $r_{rel}$ defined in Section 2.2.

4. Conclusion

This work studied the applicability of AZ ran on modest hardware to the static SM-JSSP with a combined material waste and tardiness minimization target. We have shown, that AZ leads to better results than both EDD and MCTS, thereby providing the first successful application of AZ to the realm of production scheduling. On the way to our results, we formalized a novel production scheduling problem corresponding to sheet-metal production by extending the JSSP formalism and providing a RL design for it.

Our solution was tested with SM-JSSP instances of up to 70 jobs, 6 resources and 8 paths, leading to a combinatorial complexity higher than anything published in literature to date. For real world sheet-metal production however, the solution needs to scale to hundreds of jobs. To that end two aspects should be studied in the future. On the one hand, the investigation should be extended to larger SM-JSSP instances. On the other, it should be studied how AZ performs in an online setting. The two targets can be combined to provide the scalability required for real world applications.

Furthermore, it could pay off to offer some attention to the use of ConvNets within AZ for SM-JSSP to capture temporal correlations between states as well as relaxing the environment observability requirement. Currently, all the part descriptors involved in calculating the dynamic batch processing times are provided in the state model. If these times are to be computed accurately, the dimension of the agent’s input space will explode, since both cutting and extraction times are dependent on a large number of factors such as machine configuration, material properties, further geometric features and so on. As such, it should be investigated whether AZ can still perform as strongly if we eliminate the part descriptors completely, save for areas.
Currently, the validation of AZ for the SM-JSSP is limited to the comparison with a heuristic baseline. It is planned to test AZ against GAs and exact solutions on smaller SM-JSSP instances. Additionally, AZ should be tested against the state of the art for JSSPs on benchmark data sets.

References


Biography

Alexandru Rinciog obtained his M.Sc. in Computer Science from Karlsruhe Institute of Technology in 2019. He became a Ph.D. candidate at the Chair of Enterprise Logistics of the TU Dortmund University the same year. As part of the DFG project “Smart Production Control”, his focus lies with production optimization using machine learning.

Carina Mieth is a PhD candidate in mechanical engineering at the TU Dortmund University. She received her M.Sc. in electrical engineering from the Karlsruhe Institute of Technology in 2017. Currently, she works in the predevelopment networked systems at TRUMPF GmbH + Co. KG. She is an associated member in the DFG-funded research training group adaption intelligence of factories in a dynamic and complex environment. Her research focuses on cyber-physical production systems and simulation input modeling for manufacturing simulation.

Paul Maria Scheikl received his B.Sc. and M.Sc. in mechanical engineering from Karlsruhe Institute of Technology, Germany in 2015 and 2017, respectively. He is now pursuing his Ph.D. in robotics to continue
his research with learning systems and collaborative robots. His research interests include dynamic systems, computer vision, and machine learning.

Anne Meyer is a junior professor for enterprise digitalization and supply-chain management at the Chair of Enterprise Logistics of the TU Dortmund University. She obtained her Ph.D. from Karlsruhe Institute of Technology in 2015. The 10-year collaboration with the FZI Research Center for Information Technology, served to grow her hands-on expertise in the fields of logistics, machine learning and data analytics.
1st Conference on Production Systems and Logistics

Conceptual Approach of Robustness in Logistical Control

Alexander Bader¹, Ann-Kathrin Onken¹, Kirsten Tracht¹,²

¹University of Bremen, Bremen Institute of Mechanical Engineering (bime), Badgasteiner Str. 1, 28359 Bremen, Germany
²MAPEX Center for Materials and Processes, University of Bremen, Bibliothekstr. 1, 28359 Bremen, Germany

Abstract

The logistical control and order release of structurally dynamic systems represent a great challenge. In contrast to classical shop floor production, these systems can involve transitions between the two different organizational forms job shop production and flow production. This results in a system that is susceptible to turbulence and malfunctions. This paper presents an approach that improves the robustness of these systems and achieves a stable system state even under external influences. Decentralized stock control loops control the WIP of the individual stations in a closed loop. When organizational transitions occur, interlinking losses arise, which reduce the productivity of the system. Similar effects can cause disturbances and blockages. The introduction of a robustness controller adapts the input parameters of the order release to the current system status, so that an adapted reaction takes place in the event of strong disturbances.

Keywords

Production Planning; Robustness; Logistics; Manufacturing system

1. Introduction

The fulfilment of customer requests is one of the most important objectives in production. Changed market conditions and target audiences during the distribution of high-quality products lead to flexible mass production. As a result, a challenge arises to develop an adjusted system of production, which deals with a high variety of products whenever the quantity of individual variants declines [1]. Flexible manufacturing systems are used for instance, to operate the production system with a high capacity. The increasing use of flexible and scalable assembly lines leads to further development from flexible manufacturing systems to reconfigurable production systems [2]. In order to fulfill the customer requests, the flexible mass production is developed into a personalized production [3]. One aspect of these systems is the different working stations, which can be freely interlinked and adjusted to the individual product or variant. A similar system design can be found while looking at the high-throughput system used for the development of new structural materials [4]. The free interlinkage of the working systems found in these high throughput systems requires a change of the organizational structure of the different sections of the system. These so-called temporarily paced sequences describe a shift from job shop production to flow production [5]. Thus, the productivity advantages of the flow production are linked with a flexible job shop production. This results in a complex material flow, which in turn increases the overall complexity of a system. Due to this, the sturdy design of these systems presents a particular challenge, as they need to be especially resistant to external influences, so that the systems are still able to achieve a high productivity, despite complex and volatile surrounding conditions. In this context, other challenges are the order authorization and the logistical control. These systems are among those for which it applies that a high flexibility of the system is achieved by short lead
times with low circulating stocks. In this paper, an approach is presented, which is used to design an order release and a logistical control for dynamic production systems, which in turn will prevent capacity loss and performance loss in the case of malfunctions and ensure the robustness of the system. The presented multidimensional order release prevents an upward oscillation of the production system and can avoid capacity losses.

2. Robustness in Production Systems

2.1 Definition of Robustness

The term robustness can be found in various disciplines with different cases of application. No matter in what discipline it is used though, robustness is generally understood as something having resistance towards malfunctions or as a sensitivity of the quantity of results of input controls [6]. For instance, in biology, robustness encourages the development of organisms, as these can survive and evolve in unpredictable environments [7]. With regard to an assembly line, robustness can be understood as the ability to be able to reach the planned production output, despite possible malfunctions, changing working conditions, or faulty components [8]. Thus, the robustness of an entire supply chain stems from the robustness of individual assemblies, as well as from the communication and interaction between business partners. Each of these aspects has to be taken into account whenever a robust supply chain is established, as a single unreliable supply partner is able to reduce the robustness of every other following partner of the supply chain [9].

[9] describes robustness as the ability to deal with turbulence while maintaining system performance on a high leverage. This robustness is achieved by resistance to disturbances. If a major malfunction occurs, the system is adjusted, after which the system if returned to its original state. It is also possible for the system to learn from said malfunctions and to be permanently changed, so that similar malfunctions, that may occur later, are less serious [9]. Furthermore, it is possible to react to a change of circumstances adequately, in order to minimize the damage and to further handle customer requests satisfactorily. In literature, this is known as flexibility or variability, even though the differences are not always well defined. While flexibility is only concerned with the changes within a defined structure, variability refers to changes beyond that structure. An example for flexibility is the use of a replacement machine in the case of the malfunction of the machine that was originally intended to be used. This fixed structure is not present when looking at an agile system. This kind of system is able to adapt to unpredictable events without a prior defined procedure, by developing alternative actions whenever a malfunction occurs [10]. In contrast, according to [6], a stable system abides to the defined schedule, even in the event of a malfunction. This schedule does not only include the position of orders, but also other features, such as the regulation of delivery dates, times for the order authorization for suppliers, or the planning of equipment, as for example tools. Thus, the performance of a stable system depends on the quantity of planned orders [6].

Robustness is generally understood as a resistance against malfunctions. However, it is essential to define the term malfunction more concretely. In the VDMA 66412, interruptions that are caused by malfunctions are defined as periods, that occur during order processing and extend the occupancy time unintentionally [11]. Within a production process, this usually describes, among others, machine failures, changes of production structures, problems with quality, missing tools or parts, or the occurrence of urgent orders. In addition, during the execution of production plans, error occur, which tend to be caused by missing organizational discipline, and can lead to a lengthening of the occupancy time and a reduction of the performance of the production system [12,13,9].
Figure 1: Process characteristics of a robust system [14]

Therefore, it is possible to describe robustness as a combination of controlled and capable system characteristics in a production engineering context [6]. In both cases, the spread of the output value and spread of the result quantity of one process differs from one another, as it can be seen in the figure. In this context, while controlled processes are characterized by a consistent mean, capable processes are defined by a minimal process scatter. Thus, the combination of both can be understood as a robust system.

2.2 Robust Optimization

For the calculation of robustness, [15] use the correlation between robustness and the performance of the network with varying malfunctions. For this, the production system is modeled and the addressed malfunctions are simulated with the help of a discrete event simulation. During such a simulation, all resources and entities are simulated, while the system status only changes during defined events and defined times. Thus, an interaction between multiple entities at arbitrary times is impossible, which is why a distinction can be made between discrete event simulations and agent-based modelling. A significant advantage of the discrete event simulation is the comparably short processing time [16]. Furthermore, the performance of the network is measured with the help of a priori defined key performance indicators (KPI) [15]. Examples of such a KPI are the capacity of the machines, an average processing time, and an average delay of orders [17]. All malfunctions have a cause, which can be further analyzed. Moreover, it is possible to interfere with the system, in order prevent malfunctions or to reduce the extend of said malfunction. Each of these interferences significantly changes the production model. In order to measure the impact each interference has on robustness, an interference is integrated in the simulation model, so that the KPI can once again be captured. Due to the summation of all relevant and independent KPI to a single figure known as aKPI, it is possible to compare the robustness of both systems and to measure the increase of robustness caused by an interference [15]. According to [17], it is possible to generate robustness through an excess of operational capacities of equipment or through the creation of redundancies of equipment. Both options generate costs, which is why it is necessary to find a balance between robustness and generated costs for each individual system. The higher the robustness is, the higher are the costs and the efficiency of the system [17].

3. Logistical Control of Dynamic Systems

In production control, important information for processing orders, such as type, quantity, sequence, quality, and the time distribution of products, is defined. Thus, production control is the link between long-term production program planning and short-term operative execution [18]. In addition to controlling planned orders, production control must also react to random disturbances without compromising the fulfillment of
logistical targets. This is only possible if production control has been set up according to certain principles. [19] divides production control into tasks, manipulated variables, controlled variables, and target variables. The tasks are divided into order generation, order release, sequence planning, and capacity control. When implementing production control in a production system, it is essential to select suitable control variables. This includes the entries and exits, as well as the sequence of the orders. The planned values are defined in production planning and compared with the actual values from production control. Furthermore, the backlog and the stock, which is also to be seen as the logistical target variable, are used as control variables [19]. This special role of work in progress in a production system and its good controllability gives it a special status in order release procedures.

Order release for logistic control can be divided into the two different approaches of push and pull principles. According to [20], due to several specific differences, pull systems are generally preferable to push systems due to the consideration of the throughput instead of the stock when releasing the order. The main reason for the higher performance of pull systems compared to push systems is the several beneficial effects resulting from the limited work in-progress. This reduces capital commitment costs, lead times and the schedule performance with decreasing stocks. At the same time the reaction time of the production decreases, since changed release decisions have a faster impact on the rest of the production due to the shorter throughput times. In addition to reducing break in material flow, this leads to a smoothing of peak loads and thus helps to avoid bottlenecks. According to the authors, this significantly simplifies the control of pull systems compared to push systems. In [21] they make clear that pull systems are much more robust against a false estimation of the existing capacity, since they restrict the existing stock explicitly. In push systems, on the other hand, such a false assessment leads to an unlimited build-up of stock in the affected work system. This has correspondingly negative consequences for the lead time and thus the future adherence to schedules.

4. Multi Dimensional Extremum-Seeking Logistical Control

In production planning and control, the order release is the central method of controlling the load in a production system. In the event of disturbances and external influences from the production environment, the performance of the production system changes, for example, due to delays in order processing or logistical blockages within the system. In this way, critical system states can arise in which a significant decline in performance and thus a reduction in capacity utilization and productivity occurs. A specific order release for corresponding system areas shortens the effects of the critical system states or reduces them. The result is a system that is robust against breakdowns and that, despite its high complexity, has a high level of productivity. The occurrence of organizational transitions, for example in the form of temporarily paced sequences, results in a complex planning and control system that leaves the operating point of the system, for example due to breakdowns or material bottlenecks. A robust, logistic control system can identify critical system states and a control system can stabilize the system by means of a targeted order release. This prevents a continuation of the disturbance effects along a process chain and ensures the achievement of objectives.

The consideration of order release as a control engineering system leads to a differentiated decision in order release with regard to the load of individual stations or system sections. Additional order control elements are used to adapt controller parameters to the current system status. For example, a pacing controller can delay the processing release until the paced sequence is reset. In order to achieve a stable system state, it is essential to avoid turbulences caused by organizational transitions and to reduce the effects of disturbances. Studies show that an adapted method for order control reduces interlinking losses due to temporary pacing in particular. Despite this specific release of paced sequences, the method does not completely prevent blockages and does not achieve a high utilization level at all stations [22]. From these studies, it can be deduced that in particular the highly complex material flow is not completely controlled by the applied control system. The undirected material flow leads to an uneven load in the system. Many process chains
begin with processing at a selection of a few stations, but then have different processes and focuses. The inventory control loops selected there are not sufficient to counteract this effect.

![Closed loop control system for order release](image)

**Figure 2: Closed loop control system for order release**

The logistic control presented in this paper distinguishes three control loops for order release (see Figure 2). Decentralized allocation control loops keep the stock level of each workstation constant at the operating point. This ensures that capacity utilization and throughput times are achieved. The pacing control loop blocks the workstations when a paced sequence arrives. This enables continuous processing of the order over several stations without delay due to unplanned throughput times. The robustness control loop is the outermost control loop. Disturbances and blockages lead to oscillations in the system, which lead to an increase or decrease of the stock and thus also of the output. A low-pass filter identifies strong changes at the signal outputs for which readjustment is required.

### 5. Conclusion and Further Research

This paper presented a method for logistical control and order release that can increase the robustness of the system when used. The logistic control of production systems with a high number of variants for small quantities requires the use and development of reconfigurable production systems, which represent a highly dynamic system. The order-specific combination of manufacturing and assembly steps to a quasi-unique process chain results in a high material flow complexity with a high proportion of return flows. Organizational transitions can also occur, which further increase the overall complexity. This creates a high degree of sensitivity to internal and external influences. Particularly in the case of malfunctions, they can worsen the logistical target values, so that overall productivity drops sharply. The presented logistic control can stabilize the system with its small, decentralized loading control loops and the larger robustness control loops. The corresponding signal filters of the output signals of the individual working systems are used to address different areas of the control system, resulting in an integrated logistic control system. The presented procedure for order release and logistic control is suitable for the integration of a Manufacturing Execution System (MES) of a structurally dynamic system. After each processing step or sequence of a process chain, a repeated processing release or change of the order throughput takes place. Depending on the system-
specific control parameters, further releases can be withheld or accelerated in the event of faults in order to avoid critical system states. For example, if there is a risk of underutilization, additional orders can be released by adjusting the parameters accordingly.

The performance of the method should be investigated in future investigations. For controller selection and design, it makes sense to use an event-discrete simulation model. Different controller combinations and their characteristics can be implemented and tested in this model. It is also possible to adapt the model to different system characteristics.

Acknowledgements

Financial support of subproject P03 ‘Planning and Control of Defining and Testing Processes of Microstructures’ of the Collaborative Research Center SFB 1232 "Farbige Zustände" by the German Research Foundation (DFG) is gratefully acknowledged. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 276397488 – SFB 1232.

References


Biography

Alexander Bader, M.Sc is research assistant at the Bremen Institute for Mechanical Engineering (bime) at the University of Bremen since 2016. His field of research is production planning and scheduling of dynamic networks.

Ann-Kathrin Onken, M.Sc is research assistant at the Bremen Institute for Mechanical Engineering (bime) at the University of Bremen since 2016. Her research focuses on Design of production chains for linked micro parts.

Prof. Dr.-Ing. Kirsten Tracht is head of the Bremen Institute for Mechanical Engineering (bime) at the University of Bremen since 2008. She received her PhD from the University of Hanover in 2001. Prof. Dr.-Ing. Tracht is member of the MAPEX Executive Board.
Managing Disruptions in Production with Machine Learning

Vladimir Samsonov¹, Chrismarie Enslin¹, Ben Luetkehoff², Felix Steinlein², Daniel Lütticke¹, Volker Stich²
¹Cybernetics Lab IMA & IfU, RWTH Aachen University, Aachen, Germany
²Institute for Industrial Management (FIR) at RWTH Aachen University

Abstract

Changing customer demands lead to increasing product varieties and decreasing delivery times, which in turn pose great challenges for production companies. Combined with high market volatility, they lead to increasingly complex and diverse production processes. Thus, the susceptibility to disruptions in manufacturing rises, turning the task of Production Planning and Control (PPC) into a complex, dynamic and multidimensional problem. Addressing PPC challenges such as disruption management in an efficient and timely manner requires a high level of manual human intervention. In times of digitization and Industry 4.0, companies strive to find ways to guide their workers in this process of disruption management or automate it to eliminate human intervention altogether. This paper presents one possible application of Machine Learning (ML) in disruption management on a real-life use case in mixed model continuous production, specifically in the final assembly. The aim is to ensure high-quality online decision support for PPC tasks. This paper will therefore discuss the use of ML to anticipate production disruptions, solutions to efficiently highlight and convey the relevant information, as well as the generation of possible reaction strategies. Additionally, the necessary preparatory work and fundamentals are covered in the discussion, providing guidelines for production companies towards consistent and efficient disruption management.

Keywords

Disruption Management; Machine Learning; Production Control; Visual Analytics; Deviation Detection; Similarity Analysis; Decision Support; Assembly; Assistance Systems; Mixed-Model Assembly

1. Introduction

“Anything that can possibly go wrong, does”, is the so-called Murphy’s law. This means, that unwanted events are unavoidable and something is bound to go wrong at some point in time. These unwanted events, also called disruptions, have a negative impact on the productivity and, therefore, profitability of a production company [1,2].

In times of shrinking product lifecycles, delivery times and increasing market volatility, production companies are more prone to disruptions [3]. Decreasing product life cycles and growing customisability lead to more frequent process changes, which has a direct effect on the possibility of disruptions [4]. Adequate disruption management will help avoid stagnating or even decreasing productivity levels in the production processes in the future.

Efficient disruption management in modern production is a multidimensional problem. It requires perceiving the possible challenge, understanding it and generating a suitable, possibly novel, solution. Perception,
cognition and action are the three main abilities that define intelligence in general [5]. Within the context of disruption management, perception means creating a full picture of the current production state, based on the available information. Often, available information is incomplete because of time pressure and lack of knowledge or data [6]. Cognition is related to the current production status analysis, recognition of occurred and anticipated disruptions, as well as the assessment of their probable impact and severity. Action in the context of disruption management refers to the execution of a number of activities aimed at the prevention or minimisation of negative consequences. Humans, as intelligent beings, possess natural abilities in perception, cognition and action execution and traditionally constitute the main driving force behind disruption management. However, modern production imposes new challenges beyond human capabilities. Production managers are incapable of managing the growing flow of information and often have to make decisions based only on a fraction of available information, which can lead to a suboptimal decision. A logical way to deal with the growing complexity of disruption management is the involvement of larger human teams. Nevertheless, within a team, every member has to digest the relevant information and decide upon required actions, based mostly on their personal experience and judgement. A substantial time investment is required to form a well-coordinated team with clear communication and the required levels of domain experience to make sound decisions, keeping the company’s best practices in mind. Nowadays, achieving such a goal is greatly impeded by factors such as an ageing workforce, increasing staff turnover and constantly changing production environments. However, recent advances in Machine Learning (ML) provide an indication that Artificial Intelligence capable of tackling complex problems will soon be a reality [7]. This can be of great help for experts while dealing with production disruptions.

We believe that in the age of Digitization, Industry 4.0 and ML, humans will still be heavily involved in the decision-making process. However, it will consist of a hybrid workflow where part of the task will shift from a mainly manual effort to a semi-automatic or even completely autonomous approach. Smart experts and assistance systems will support the worker in making the right decisions by providing the relevant information at the perception stage, guiding the decision-making process or automating certain decisions at the cognition step and facilitating chosen actions to execute at the action stage.

In order to develop useful systems for this application area, several aspects have to be considered. In the perceptions stage, it is important to select, consolidate and transform relevant information from multiple sources to a form easily comprehensible by human experts. At this stage, ML can augment human intelligence with speed, consistency and precision in processing big amounts of data. Beneficial patterns in the data can be recognised that are invisible to the human eye [8]. At the cognition stage of disruption management, past and future disruptions need to be considered and evaluated. Disruptions should be classified based on their estimated negative impact. Possible mitigations, also known as reaction strategies, need to be matched to the various disruptions. Addressing different tasks at the cognition stage often requires intuitive capabilities, the ability to come up with the non-standard solutions as well as conducting complex planning. People are notably competent in this range of tasks. Nevertheless, people can still rely on ML for generation of data-driven predictions concerning the disruption risks and their severity. Additionally, human experts can be presented with recommendations to viable actions, based on similar events in the past or with the help of rule-based expert systems.

In the next chapters, we give an overview of the state of the art for both classical and ML-related approaches in the field of disruption management. Subsequently, our approach is presented, together with some related requirements and practical results based on a real use case. Finally, the conclusion and outlook finalises the paper.
2. State of the art

For the present topic, two research fields are of interest in the state of the art. Firstly, it is worthwhile to investigate the state of the art in disruption management to define its research deficits. Secondly, since disruption management is part of Production Planning and Control, the state of the art of ML in PPC is of interest.

2.1 Disruption Management

Disruptions are critical events that lead to a break in the workflow of a production process, such as machine failure, rework or missing personnel. Therefore, disruptions have a negative impact on the production system due to their substantial influence on logistical targets.

The management of disruptions as a research field already exists since the 1970s [9]. Disruption management refers to the structural and procedural organisation of all successive measures, including the elimination of disruptions, the minimisation of the consequences of disruptions and the prevention of disruptions. These three temporal organisational units are also referred to as short-, medium- and long-term disruption management [10]. In disruption management, a distinction is made between different strategies. Prevention strategies deal either with the elimination of the cause (causal strategy) or with the defence against the occurrence (defence strategy), i.e. with the avoidance of the occurrence of disruptions. Reaction strategies have two categories, namely system-oriented strategies that deal with the consequences of disruption, and reactive strategies, that adapts to the new situation resulting from the disruption. Thus, they react to disruptions that have already occurred [11].

While the first approaches in disruption management mainly dealt with the disruptions themselves and their classification [12,10], approaches in recent years focus more on data aspects and simulation [13–16]. However, especially the aspects of possible reaction strategies and the use of ML in disruption management have not yet shown satisfying results.

2.2 Machine Learning in production

Machine Learning (ML) is a subset of Artificial Intelligence (AI) and is capable of discovering underlying patterns and dependencies through examination of data [17]. ML is used across many domains of production, but is mainly used within scheduling, process planning and control [18]. In this chapter, we will cover only applications potentially related to disruption management.

One of the topics gaining attention in the research community is the use of Deep Reinforcement Learning (RL) for solving combinatorial optimisation tasks, such as scheduling [19–21]. This research direction in ML is a potentially good solution for conducting production resequencing in a flexible manner, as a response to production disruptions. Compared to classical Operations Research (OR) approaches, deep RL methods can potentially adapt to changing environments and boundary conditions by retraining without having to redesign the whole solution approach. However, deep RL methods are new. Several questions, such as explainability, validation methods, generalisation capabilities and robustness still need to be addressed before it can be deployed in real production environments [22].

Besides deep RL, unsupervised learning is another subset of ML that does not require collected data to be labelled. There are many scientific works based on unsupervised learning that elaborate on the use of clustering algorithms for automatic detection of similar products, classification of products in product families and anticipating product failures [23–25].

Another promising field is the use of Visual Analytics (VA) in production planning and scheduling. Different works propose the use of various data transformation and visualising techniques to automatically provide experts with decision-relevant information and interactively evaluate possible production scenarios [26–28].
In our opinion, the use of ML and VA approaches on the order level has not yet been discussed sufficiently in the context of disruption management.

3. Approach

The overall goal is to combine a number of ML and VA approaches to enable a data-driven prediction of production disruptions. Therefore, we introduce a practical use case with real production data used to develop and validate our approach.

3.1 Use Case Introduction

We choose a mixed model continuous assembly line producing self-checkout machines for retail to develop and test our disruption management methodology. This is motivated by the complexity of the final product. Many modules of the system, from hardware to software, are customisable to the customer needs. It leads to a vast amount of different products being produced on one line. Often, specific product configurations are produced only a few times, making erroneous planning estimations and disruptions along the assembly process more likely. Products go through a number of stations on a production line with a pre-defined cycle time. Production disruptions of any kind have to be addressed within the short period of time the product stays on a given working station. Otherwise, the assembly sequence cannot be completed and the product will have to be finished separately in the rework area. This requires additional personnel capacities, causes longer production times, higher costs and is limited to a number of products at any given time because of the rework area space and human labour constraints. Possible disruptions, if accounted for at the planning and order release stage, can be mitigated. Therefore, it is important to timely recognise and foresee disruptions.

3.2 Data

One of the most frequently used IT-systems by production companies are Enterprise Resource Planning (ERP) systems [29]. An ERP system facilitates the order processing from supplier to customer, e.g. with production planning and materials management. It leads to a certain degree of homogenisation of available data across companies. Planning data, such as the Bill of Materials (BOM) and production steps, is available for every produced order. The planning data is augmented with historical data, including disruptions in production steps, conducted rework and schedule adherence. This data is not available in every ERP system but is a crucial part in anticipating possible process disruptions [30]. The more information on circumstances, causes and impact of previous process disruptions we have, the more precise we can estimate future disruptions.

3.3 Explorative Data Analysis and General Concept

Similar orders consist of similar materials and require similar steps in the assembly process. Potentially these orders can inherit the same design flaws that make the assembly process prone to errors. Once clusters of similar orders have been identified, it is useful to know how likely each of these clusters are to be finished on time without disruptions or to be moved to the rework area. Originally, for the given use case, the product similarity is defined by four main product families that share common functional and design features. To investigate underlying structures in these product groups we use the unsupervised ML method of hierarchical clustering [31]. The features used to create
such a clustering is a combination of the materials the orders consist of and the production process steps to assemble the order. An agglomerative hierarchical clustering of the data is performed with complete linkage. Figure 1 shows the dendrogram, which is a tree-like structure, representing the results of the hierarchical clustering. This type of representation is integral to understanding how the clustering is performed. The bottom of the dendrogram represents each of the individual orders that have been produced in the past. Moving further up the tree, similar products and product clusters are linked together in larger clusters. The closer to the top the linkage occurs, the lower the degree of similarity within the cluster. At the top of the dendrogram we can recognise four horizontal lines depicting four main product families. The fact that each product family has a complex tree structure demonstrates the high level of variation in the production process. This variation is the result of high customisation possibilities for each product. Therefore, in order to approximate possible behaviours of an order based on similar orders from the past, we need to split given product families in smaller groups with higher similarity first.

Clustering helps with determining which production orders are similar, but we are interested in more than similarity. We would like to abstract this to the behaviour of the orders and would like to determine which orders will be problematic and will most likely be moved to the rework area. This extra layer of information is created by estimating the probability of any order in a cluster to be moved to the rework area, based on historical production data. With this information, the worker can undertake the necessary measures to completely avoid or minimise the impact of potential disruptions.

4. Results

The production data for the given use case contains 6153 orders. We create a one-hot encoded matrix describing all possible configurations, used parts and production steps. Every possible product in this case can be completely described by a vector of length 2462. A Principal Component Analysis (PCA) [32] is performed to reduce dimensionality. As a result, the vector describing each order is reduced to 30 dimensions that capture 93.3% of the variation in the data. It is important to generate an intuitive visualisation combining both similarity information as well as historical production data. t-Distributed Stochastic Neighbour Embedding (t-SNE) [33] is used to map the 30 dimensional production orders onto a 2D plane for the ease of visual inspection. It groups similar observations together in tight clusters while trying to pull dissimilar observations farther apart. Singular points in the 2D space represent product configurations produced only once. To ensure t-SNE visualisations depict true clusters of similar orders we use Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [34] as a validation method. Clustering methods are successfully used for detecting product families based on product similarities [23–25]. Figure 2 shows a t-SNE generated 2D map of orders. On the same visualisation, DBSCAN clusters are colour-coded. From this figure it is clear that all points assigned to the same tight cluster by t-SNE have the same colour, defined by the affiliation to the one or another DBSCAN cluster. Therefore, DBSCAN successfully clusters similar products and t-SNE represents these clusters in a 2D space.
We add another layer of information to the 2D representation of all produced orders by color-coding (in this case encoded as with color gradient from black to white) according to the estimated rework ratio per cluster of similar products, which is represented in Figure 3. This allows the user to quickly see what planned orders can potentially lead to disruptions during production, because it belongs to a cluster with an estimated high rework ratio (clusters with brighter colouring) or because of a new configuration (single point not surrounded by other observations). It allows the user to prepare for possible process disruptions.

5. Conclusion and Outlook

Even in a highly automated and digitized production environment, disruptions will occur due to unforeseen failures of machines, humans or other resources. Therefore, disruption management is an important managerial task that needs to be handled with domain expertise and supported by data-driven approaches, such as ML. The presented paper introduces the concept of using ML to facilitate disruption management by identifying possible problematic orders based on historical data and through the discussion of an application use case. We used unsupervised ML to form clusters of similar product configurations, evaluate the likelihood of rework based on historical data and generate a 2D visualisation allowing to approximate how much rework planned orders are likely to require.

Building on the presented findings, the next step will be the validation of the concept. To do so, we will use data from the presented use case and implement a dashboard for the workers in the assembly line. With the dashboard, they should be able to anticipate possible disruptions and derive possible strategies on how to manage the assembly in the case of a disruption. This will serve as a starting point for the creation of an assistance system that will go a step further than solely helping to detect possible disruptions, but helping to derive possible reaction strategies. Thus, the workers need only to decide between the best alternatives, which should lead to a relief in the amount of work created by managing disruptions.

Acknowledgements

The European Regional Development Fund (ERDF) funded the presented research (grant number EFRE/ERDF-0800924). The authors would like to thank the European Regional Development Fund for the kind support and making this research possible.

References


Biography

Vladimir Samsonov (*1990) is a scientific researcher at the Institute for Information Management in Mechanical Engineering (Cybernetics Lab IMA & IfU) at RWTH Aachen University since 2017. As a part of the Digital Production Technology Research Group, he specialises on questions related to digitalisation and the use of machine learning within the context of manufacturing engineering.

Chrismarie Enslin (*1991) joined the Institute for Information Management in Mechanical Engineering (Cybernetics Lab IMA & IfU) at RWTH Aachen University in 2019. She is part of the Digital Production Technology Research Group and mainly focusses on Machine Learning and Reinforcement Learning within the field of industrial manufacturing.

Ben Lütkehoff (*1990) has been part of the Research Institute of Industrial Management at the RWTH Aachen University since 2015. He works in the department of Production Management in the group Closed-Loop Production Control, where he focusses on combining data from IT-systems with implicit knowledge of production workers.
Felix Steinlein (*1989) has been part of the Research Institute of Industrial Management at the RWTH Aachen University since 2017. His research focus lies within production planning and control, process analytics as well as enhancing operations with digitization and the use of IT-systems (MES / ERP).

Daniel Lütticke (*1984) has been part of the Institute for Information Management in Mechanical Engineering (Cybernetics Lab IMA & IfU) at RWTH Aachen University since 2016. He is leader of the Digital Production Technology group researching on advanced applications and effects of digitalisation and artificial intelligence in the field of industrial manufacturing.

Prof. Dr.-Ing. Volker Stich (*1954) has been head of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 1997. Prof. Dr.-Ing. Volker Stich worked for 10 years for the St. Gobain-Automotive Group and lead the management of European plant logistics. In addition, he was responsible for the worldwide coordination of future vehicle development projects.
Do-it-Together Concept for Production Ecosystems

Andreas Külschbach¹, Tobias Leiting¹, Volker Stich¹
¹Institute for Industrial Engineering (FIR) at RWTH Aachen University, Aachen Germany

Abstract

The do-it-yourself mentality is particularly widespread in the furniture sector. Homemade furniture is very popular. The individualisation of furniture can be observed in internet forums, such as the online platform Pinterest. These creative ideas of potential customers show a need for individualized sustainable pieces of furniture. The current production structures, however, do not allow individual production according to the end customer's specifications. In addition, information logistics faces a major challenge: making the creative ideas of end consumers available to producers in parametric form. Topics such as customer requirements in relation to sustainable production, material specifications, industrial property rights, fair production conditions and traceability are the focus of this data interchange. An open and innovative European furniture ecosystem must be created to connect all stakeholders in the production process. This is made possible by a platform that channels the creativity of consumers and makes it designable and producible through the professional skills of designers. This requires the involvement of manufacturing specialists who can produce personalised products through sustainable intelligent production technologies. An exchange of information must also take place securely and quickly in order to protect the personal rights of the sources of ideas. This is being developed in the EU research project INEDIT - Open Innovation Ecosystem for do-it-together process. By connecting many different stakeholders along the entire value creation process, a change towards efficient collaborative collaboration is achieved. This paper presents a project insight for the development of an international co-creation platform by presenting the problem and linking it to a potential solution.

Keywords

Sustainability; Social Manufacturing; Supply-Chain-Networks; Lean Manufacturing; Information Logistics

1. Introduction

1.1 Initial situation

For decades, industrial change has led to an increase in prosperity. Starting with the first industrial revolution, key technologies such as the steam engine, the use of electricity or the use of computers have increased the industrial performance of individual economies. The increase in economic output is directly related to increased prosperity [1]. As a result of growing prosperity and increasing digitalisation, people are more and more changing their demands on the value of consumer goods as well as material and immaterial services. The current fourth industrial revolution relates to the use of data in the context of production. The driving force behind this evolution is the increased urge for personalized products [2]. In many areas, consumers are looking for products that they can personalise to meet their individual needs [3]. This is reflected in the challenge to adapt the production of products to individual customer requirements. Manufacturing companies try to implement production according to batch size one. However, high costs and the integration
of new technologies are necessary. Furthermore, different customer requirements mean that used resources or the necessary expertise are no longer available centrally but decentrally. This leads to international logistics networks for producing individual goods. In order for the production of individualized goods to be produced according to specific customer wishes, a continuous distribution of information must be ensured. This applies to both internal and external business processes.

1.2 Problem definition and motivation

For years, industrial companies have pursued the goal of designing products in such a way that they can be manufactured as cost-effectively as possible (e.g. automation, lean production), neglecting the needs of customers [4]. Classical product development follows two essential patterns. Either products are developed and manufactured in a customer-neutral way, or the products are based on a market need [5]. Customer-neutral products are usually function-bound and are not subject to any need for individualization. Usually, these are functional products such as pipes, concrete, etc. Products influenced by customers are mainly found in the consumer goods industry. Here, the development process is geared to the needs of the customer. At present, these customer needs are determined by surveys or studies. The knowledge gained is incorporated into the product development process. The manufactured products can then be advertised and brought to market. This entails the danger that the manufactured products may not or only partially address customer needs. This type of product design has shaped society and boosted the growth of individual large companies. Small and medium-sized enterprises (SMEs) in particular have difficulties with their current incremental innovation strategy, organisational structure and business models to withstand market pressure [6]. The trend towards ever more individual products is one of the greatest challenges and at the same time a great opportunity to strengthen the position of SMEs.

With regard to sustainability, issues such as energy management or resource-conserving production are currently not in the focus of manufacturing companies. Companies do not collect data that enables production to be evaluated according to ecological aspects [7,8]. However, the changing market pressure caused by sustainability initiatives is causing companies to rethink their approach in this area as well. Companies are faced with the challenge of incorporating the sustainability aspect into their product development. This means a general rethinking and the creation of new structures to fulfil changing customer requirements and to sell products.

The goal therefore must be, to build an ecosystem of sustainable and interconnected production. This should match and withstand the requirements of digitalisation as well as the requirements of the customers. The involvement of different parties in the life cycle of a product must be guaranteed in order to enable small companies to access new technologies and networks.

2. Objective of research work

Nowadays, companies operate in the social era and try to strengthen their brand by integrating a sales platform. In addition to the product, companies try to sell an experience and enable interaction between customer and product. Many companies rely on the inclusion of platforms on which they carry out a new type of market research. The aim is to create a product in an affluent society that fulfils undiscovered customer needs and thus achieves a market advantage for the company. It illustrates the paradigm shift towards a platform economy. CUSUMANO provided the following definition [9]: “A platform strategy - as opposed to a product strategy - requires an external ecosystem to generate complementary product or service innovations”. In particular, this involves adapting the products to the customers’ requirements and thereby increasing the added value for the customers through an increased emotional bond. The Internet is used to include end consumers on platforms. In Europe, 76% of people in western Europe use the Internet [10]. This is equivalent to around 300 million people. It illustrates the enormous reach that can be achieved through
interaction on the Internet. The furniture industry in particular is lagging far behind other industries when it comes to using the Internet. In the furniture industry, only 16% of business is done online. For other products such as books or clothing, the proportion is over 60% [11]. This means that there is great potential for improvement in this sector.

The platform concept is used in the European research project INEDIT to achieve closer cooperation between a wide range of partners. The platform not only covers a specific area in the life cycle of a piece of furniture, but also covers all life phases from the generation of ideas to production and sales. The overall objective is to design and implement a framework for a new ecosystem that enables stakeholders to collaborate from conceptualisation to materialisation. The core elements of the project are shown in Figure 1.

The platform concept is used in the European research project INEDIT to achieve closer cooperation between a wide range of partners. The platform not only covers a specific area in the life cycle of a piece of furniture, but also covers all life phases from the generation of ideas to production and sales. The overall objective is to design and implement a framework for a new ecosystem that enables stakeholders to collaborate from conceptualisation to materialisation. The core elements of the project are shown in Figure 1.

INEDIT’s customer-focused process on the one hand aims to channel the skills of consumers to enable them to make their own furniture. On the other hand, it enables industrial companies to integrate innovative technologies for machines (e.g. 3D-printing of recyclable plastic or wood), processes and sustainable use of resources into their portfolio. To achieve this goal, INEDIT integrates new methods and technologies to support creativity and design processes within the ecosystem. It supports access to production resources through the development of new sustainable production processes. These are integrated into agile production networks and thus simplify the personalization of furniture through a directed data flow up to the production level. Despite focusing on consumers, INEDIT will develop sustainable business models that are profitable for all stakeholders in an environmentally friendly way. In the following chapter, the core aspects of the platform concept will be discussed.

2.1 Do-it-together

Started with the usage of experimental and construction kits, the urge of the consumer for individualization became more and more entrenched over time. People began to transform their own ideas into objects. From the 1940s onwards, the idea of producing furniture for self assembly spread. By buying individual elements in a compact cardboard box, transport was made easier and the "unfinished" furniture was more affordable [12]. The necessary contribution of the buyer by assembling the individual parts led to an increased appreciation of the end product, as working time and energy were invested [13]. The consumer was involved in the production process. The aspect of an individualised end product which is particularly important today, however, is not yet present in the furniture for assembly [14]. If everything was assembled according to instructions, the similarity of the furniture was more an indication of correct and flawless work. This is where
the so-called "do-it-yourself" concept begins. "Do-it-yourself" refers to all forms of self-made work in all areas [13]. However, the person who did it himself did not receive any professional training, but rather learned the skills himself [12].

The guiding principle of the newly conceived platform is "Design Global - Produce local". This implies the cooperation of different partners and describes the further development of the "do-it-yourself" concept. According to HIRSCHER ET AL., social manufacturing is a way of producing physical products by enabling individuals or groups to contribute to all activities of the development process, such as Idea, design and production [15]. Social manufacturing can be seen as crowdsourcing for manufacturing, which promotes a paradigm shift towards a decentralized and socialized approach of mass individualization of products [16]. It was first tested in the fashion industry, based on collaboration in the production of garments, which allows consumers to participate in different phases of the production process. At the same time, it enables consumers to create new innovations during the design process. This is just an added value created through collaboration. Furthermore, social, economic, environmental, knowledge, experience and emotional benefits can be identified [15]. The social concept approaches of cooperation in value creation networks go hand in hand with the strengthening of small and medium-sized enterprises. The social bond to local companies is also based on good and open communication. Smaller companies therefore have an advantage for open innovation processes [17]. The smaller production series of highly individualised products also fit very well with the smaller capacities of smaller companies. Finally, this type of collaboration in networks offers great potential for new business opportunities leading to new business ethics. This is based on the importance of joint learning, novel value creation and collaborative production in an extensive network [15].

All this is summarized within the concept "do-it-together". The independent individualization of furniture according to the "do-it-together" concept, which is typical for the furniture industry, is to be replaced and the advantages mentioned are to be exploited. As can be seen in Figure 2, the "do-it-yourself" concept is intended to generate added value.

---

**Figure 2: INEDIT Objectives and Targeted Value Elements**
2.2 Sustainability

In today's society, a sustainable way of life is widespread and can be realized especially in wealthy regions through a diverse range of sustainable goods. In contrast to this, for industrial companies the environment is perceived as irrelevant [7]. Companies act according to imposed standards, which, differ worldwide and are predominantly perceived as obstructive. This is one of the reasons why production is relocated globally. In such a way, environmental requirements can be avoided. Furthermore, the prevalence of environmental management systems in companies is very low [8]. Many companies do not collect any data to evaluate the energy consumption of individual machines and systems. Furthermore, no data and protocols are collected on the equipment used that could be used for evaluation purposes. All these aspects do not generate any usable added value for the product for companies as things stand today. For most companies, pure data collection and evaluation as well as evaluation according to sustainability aspects is an activity that does not contribute to sales. Even savings in energy consumption are not very attractive for companies at today's prices for raw materials.

Only end customers can change this. The trend towards sustainable products is obvious [14]. The movement towards a sustainable lifestyle, driven mainly by the younger generations, could also represent a potential opportunity for the manufacturing industry. A sustainably produced product will increase the chance of selling it in the future [18]. To do this, however, companies must first use sustainable raw materials and use them in energy-efficient production processes. The co-creation platform of the INEDIT research project is intended to strengthen this development. Sustainable cooperation is one of the core aspects of the platform. The platform is intended to promote sustainable and resource-saving production. For this purpose, the evaluation of the environmental impacts of newly developed designs is made possible via the platform. The evaluation refers to the complete life cycle of the product. The DIT environmental assessment model calculates a set of Environmental Key Performance Indicators (E-KPIs) on the basis of the production instructions created. These can then be evaluated using an assessment model in order to derive action measures that can reduce part of the calculated environmental impact. Instructions for action should then be translated into real production instructions and transferred to the machining centers in use. However, instructions for action can also be included in the design process (e.g. proposal of an alternative, more sustainable material). In this way, the platform supports the development and production of sustainable furniture.

2.3 Information network and business model

Open innovation platforms explicitly include the business model as a source of value creation for all participants [19,6]. However, many open innovation platforms create products that are unsustainable and have neither a high utility value nor a positive long-term impact on the quality of life of consumers. They also have difficulties establishing themselves in the market. The deficit of the existing platforms is that the business models are not sustainable, focus exclusively on customer benefits and often cover only a part and not the entire value chain of product creation.

The business models of existing platforms address only a part of the stakeholders in the value chain of product design. While business models of some platforms reach a large community of consumers to understand their needs, the business model of other open innovation platforms is primarily production-oriented. RAYNA distinguishes between two types of platforms where product design is supported either by the consumer community or by cooperating manufacturers [20]. Examples of consumer-oriented communities are Quirky, FanVoice or even deinschrank.de. On these platforms, manufacturers are not integrated into the design process. This creates a need for low manufacturing complexity of the products. With the low complexity, mass production can be achieved, which is inflexible to change requests. Representatives for production-oriented platforms are Opendesk and Ideapoke. Complex products can be manufactured on these platforms. However, customers do not have the opportunity to participate intensively
in the creation process, as the ability of the producers is in focus. In order to exploit the full potential, all partners along the value chain must be included in the business model of the co-creation platform [21]. The business model of the INEDIT DIT Open Innovation Platform will include both consumer-oriented and product-oriented open innovation. Designers, developers and manufacturers will be involved to ensure customer-oriented production.

Currently, business models of open innovation platforms are primarily designed for the mass production of products. Individual preferences and sustainability cannot be taken into account. Therefore, mass production must be transformed into mass customisation by addressing customers and creating an interface between digital and physical space. By establishing a co-creation platform where both consumers and professionals work closely together to develop products, INEDIT is contributing to this change by going beyond digital value creation. This link between consumer and production enables the INEDIT platform to develop products with a higher manufacturing complexity than existing platforms. As design and local collaboration spaces are considered necessary to achieve appropriate open knowledge generation across the entire value chain [22], INEDIT will innovate the existing landscape of business models.

With regard to the functions necessary for successful collaboration, many current platforms still lack sufficient components for successful open innovation product development [23]. The lack of tools and the associated restriction of community cooperation ultimately leads to a higher risk for an unsuccessful platform [24]. Therefore, a business model is needed that takes into account a value creation network that encompasses both the scope of online and offline collaboration. A major challenge with existing open innovation platforms is the participation of users in the commercial profit of the platform. Innovative users will only be won over in the long term if they receive sufficient incentives for cooperation [25]. A revenue mechanism that takes into account both monetary and non-monetary rewards has not yet been developed and there are only a few platforms that currently benefit from online collaboration.

3. Outlook and vision

The mentioned problems, which are to be found particularly in the furniture industry, are to be addressed by the co-creation platform. The establishment of the platform should, on the one hand, strengthen the cooperation between partners from different countries and on the other hand strengthen small and medium-sized enterprises in the global market. The sustainability goals are to be anchored in a business model. A reward system is to ensure sustainable cooperation. INEDIT wants more people to be able to implement their own ideas and identify more strongly with their furniture. Due to the higher emotional bond and the more appropriate functions of a piece of furniture, a trend reversal is to be initiated that will put an end to the throw-away society. The newly created production and information networks should result in increased innovative strength. The production technologies will undergo a major change due to the processing of recycled raw materials and the urge for individual production. Compared to today, the corresponding information network for the international exchange of information and knowledge must become much more efficient. In addition, the company’s IT must also enable customer-oriented production. This includes the production of prototypes as well as the flexible reaction to short-term change requests.

Acknowledgment

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement No. 869952
References


Biography

Andreas Külschbach M.Sc. (*1991) has been working as a project engineer at FIR at RWTH Aachen University since 2017. In his current position as part of the Production Management Division, he supports companies in various industries in the design and implementation of efficient production and logistics systems. He also participates in different research projects.

Tobias Leiting M.Sc. M.Sc. (*1992) has been working as a project engineer at FIR at RWTH Aachen University since 2017. In his current position as part of the Service Management Division, he supports companies in various industries in the design and implementation of digital services. He also participates in different research projects.

Prof. Dr.-Ing. Volker Stich (*1954) has been head of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 1997. Prof. Dr.-Ing. Volker Stich worked for 10 years for the St. Gobain-Automotive Group and lead the management of European plant logistics. In addition, he was responsible for the worldwide coordination of future vehicle development projects.
Evaluation of (De-)Centralized IT technologies in the fields of Cyber-Physical Production Systems

Nick Große¹, David Leisen², Tan Gürpinar¹, Robert Schulze Forsthövel², Michael Henke¹, Michael ten Hompel²

¹Chair of Enterprise Logistics, Dortmund, Germany
²Chair of Materials Handling and Warehousing, Dortmund, Germany

Abstract

In the course of the digital transformation, organizations are not only facing increasing volatility of the markets, but also increasing customer requirements and thus an increasing complexity in production and logistics systems. Therefore, production plants need to become more flexible by transforming conventional production systems to Cyber-physical Production Systems (CPPS). CPPS allow organizations to dynamically react to fluctuations in demand and markets and to introduce new product lines quickly and effectively.

The challenge in implementing CPPS is to handle and store relevant data streams between Cyber-physical objects in a secure but transparent way. As CPPS involve a high level of decentralization, the data storage can either be combined with centralized IT-solutions like a Cloud or utilize decentralized IT-technologies like Edge Computing or Distributed Ledger Technologies (DLT) like Blockchains.

The paper addresses the suitability of centralized and decentralized technologies in terms of dealing with data streams in the fields of CPPS. For this purpose, based on a paper exploration, appropriate evaluation criteria are derived, followed by a comparison of exemplary centralized and decentralized technologies. The outcome is a qualitative evaluation of the supplement of each technology regarding its suitability of dealing with data streams.

Keywords

Cyber-physical Production System; Blockchain; Cloud Computing; Edge Computing; IT Infrastructure

1. Introduction

Cyber-physical Production Systems (CPPS) are envisioned as production system components with information processing and communication capabilities able to execute physical processes within a production system in cooperation with other entities [1]. The expectations towards CPPS are enormous. They are considered as a central factor in the future development of manufacturing [2] or even as the key pillars of the 4th Industrial Revolution [3]. In the field of CPPS, the capture of all data is necessary to monitor the communication between different Cyber-physical Systems (CPS) and for detecting failure potentials [4]. However, the design of CPPS poses a challenge in terms of integration, especially regarding the appropriate IT infrastructure [5]. Various technological approaches are conceivable to create the IT infrastructure. These vary in the degree of decentralisation.
The present paper examines three different possible approaches of an IT infrastructure for a CPPS, each of which represents an extension of the previous:

2. A Cloud extended by the decentral approach of Edge Computing [7]
3. The combination of Cloud and Edge Computing extended by Blockchain (distributed approach) [8]

The three approaches are evaluated in terms of their suitability for a CPPS. The aim of this paper can be subsumed to the following research questions:

RQ 1: Which requirements are determined on an IT infrastructure through the properties of a CPPS?
RQ 2: To what extent do the IT infrastructures fulfill those criteria?
RQ 3: To what extent is the Blockchain suitable in the handling of various data streams in CPPS?

Section 2 deals with the connection of the three mentioned technologies (Cloud Computing, Edge Computing and Blockchain). Section 3 identifies relevant evaluation criteria for IT infrastructures of CPPS. Section 4 consists of an evaluation of the previously mentioned approaches with regard to the identified comparison criteria, which most important insights are gathered in a consolidation table. Section 5 draws a conclusion of the results followed by an outlook for further research.

2. Background

In this paper Cloud Computing is utilized as an example for a central IT infrastructure, while Edge and Fog Computing will be analysed as a decentral-, and Blockchain as a distributed technology. For a better and common understanding, the fundamentals of each technology are explained in more detail below. In order to address the highlighted challenges in CPPS, data streams have to be handled effectively and efficiently. The degree of complexity increases with the number of CPS in terms of material, information and financial flows [9]. With regard to these streams, IT infrastructures have to offer the appropriate level of decentralization. In this paper, central, decentral and distributed technologies are compared in their interaction to realize the managing of localization data of products or machinery, measure certain characteristics, or enable machine-to-machine communication and payments (see Figure 1).

![Figure 1: Categories of System Structures and Data Streams relevant for CPPS](image-url)
According to [7], Edge Computing can enhance the functionalities of Cloud Computing. Combining those two technologies enables a flexible, scalable and reliable production configuration as well as distributed data analytics [10].

3. Review Setting and identified criteria

The central contribution of this paper is the evaluation of the IT infrastructures regarding their suitability for dealing with data streams in CPPS. For this purpose, suitable criteria have to be identified in a first instance, which is done by analysing scientific papers that explicitly deal with IT infrastructures in terms of handling data streams in the topic of CPPS.

The search is composed of scientific papers found by queries done in reputable academic search platforms such as Scopus. Suitable topic-relevant keywords such as ‘Blockchain’ or ‘Edge Computing’ or ‘Cloud Computing’ were chosen and combined with terms referring to the fields of interest such as ‘CPPS’ or ‘CPS’, respectively their full forms. Exclusions are made based on the fact whether the content was referred to CPPS or not. Highlighted terms, which deal with the requirements of data streams referring to CPPS have been chosen as appropriate criteria in this paper. The exploration reveals that, in this context, ‘Scalability’, ‘Latency’, ‘Security’, ‘Processing’ and ‘Flexibility’ were often emphasized requirements and therefore imply a correspondent relevance regarding the problem addressed in this paper [4,11,12]. Further use-case-specific criteria regarding a CPPS are conceivable. However, in addition to their relevance, the criteria already mentioned are primarily influenced by the technologies under consideration, as described in section 4. A more detailed description of each criterion is explained below.

3.2 Scalability

Scalability in this context means the ability of the IT infrastructure to adjust to handle the required data streams of a CPPS, especially concerning the number of connected devices in the network. It is one of the challenges to overcome in manufacturing regarding Industry 4.0 in general [13]. Appropriate structures and methods are necessary to reach a robust CPPS in a changing, uncertain environments [2]. To meet these requirements, the system must be sufficiently scalable. This is also one of the foremost issues in the design of wireless sensor networks [14], which are often part of a CPPS.

3.3 Latency

The implementation of applications in the context of Industry 4.0 in general demands real-time response and reduced latency of the IT infrastructure [15]. It poses a great challenge for industrial data networks to deliver the data to the consumer nodes within the required timeframes [16]. A CPPS, as a distributed embedded system, has additional communication latency in comparison to traditional embedded control systems [17]. However, fulfilling the time constraints is essential for many applications of a CPPS. The operation can become incorrect when exceeding them in a single instance [17] (e.g. an Autonomous Guided Vehicle gets the order to break after a crash). Therefore, the appropriate handling of time in operation systems and computer networks is a main research and development challenge regarding CPPS [2].

3.4 Security

Cyberattacks on CPPS are considered inevitable, which is why Cybersecurity penetration within the manufacturing domain is a need that goes uncontested [18]. The issue of Cybersecurity represents one of the major hurdles in implementing Cyber manufacturing [19] and therefore, is a central issue for future developments regarding CPPS [2]. With an increasing number of CPS with different weak spots and their interconnections, the vulnerability of the whole system also increases [11]. To avoid unintentional disposals or even loss of data, suitable security mechanisms have to be made available [11].
3.5 Flexibility

A certain degree of flexibility of the IT infrastructure is required in a CPPS because of its highly dynamic nature regarding the Computing resources and the physical processes [12]. For instance, the availability of participating devices can change dramatically during deployment [20]. Furthermore, challenges for CPPS regarding the flexibility result from the size of data generated by the devices in modern manufacturing, which can range from terabyte to petabyte for a single data set [21] and the required structures which have to be robust in changing, uncertain environments [2].

3.6 Processing

To capture, manage, and store the extensive data amounts generated in Industry 4.0 applications represents challenges for the industry [16]. To integrate CPS in manufacturing, the systems must be able to analyze big data-information. The generated data sizes for single industrial deployments can reach petabytes for single industrial deployments [13]. Due to the requirement of processing big data in real-time, the use of multimodal interfaces is beneficial [12].

4. Evaluation

4.1 Cloud Computing

Cloud Computing is defined as ‘a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable Computing […] that can be rapidly provisioned and released with minimal management effort or service provider interaction. […]’ [22]. Cloud Computing comes along with the application of the Internet of Things such as IoT-Clouds and fulfills the requirements of flexibility, uptime, cost and redundancy [23]. One of the biggest advantages of Cloud Computing is the ability to handle different scales of data volume [4,24]. IoT applications like Cloud-based solutions may benefit from a strong scalability [25]. Nevertheless, Cloud-based solutions are exposed to plenty challenges. Especially, Cloud-solutions are prone to third party cyberattacks [26]. As well as the integrity, the accuracy of the injected data is a vulnerable aspect regarding the validity of the content inside the Cloud [25]. Moreover, it is difficult to track third party attacks [27]. Therefore, security is a deficit and thus one important challenge of Cloud Computing.

Despite its potential of a high scalability, the quality of Cloud Computing infrastructure is affected or rather restricted by its latency [29,28]. Thus, with regard to Cloud Computing, latency indicates the throughput speed and thus it represents a so-called bottleneck [30]. For use-cases that require low latency, the single use of Cloud Computing alone is not recommended [31]. The disadvantageous latency properties leads to a loss of data processing [25]. Nevertheless, in combination with machine learning methods, Cloud Computing gains the necessary strength to compete with large data-sizes and facilitates the handling of such volumes through suitable separation and allocation of data sets or rather workflows to external units such as dedicated servers [25,32]. Furthermore, in combination with Big Data, Cloud Computing enables a mighty support system for companies to achieve their ideal production scenario [4]. Despite of its big lack regarding security and latency it possesses a high flexibility due to the possibility of target-oriented allocation of data resourced [33]. Besides the scalability, the resource flexibility represents one mighty advantage in the use of Cloud Computing for CPS [24].

4.2 Edge and Cloud Computing

Since the Edge Computing principle aligns with the concept of Fog Computing and the two technologies are often referred to interchangeably, we will only use the term ‘Edge Computing” in the following [29]. Edge Computing brings a distribution of Cloud Computing capabilities to the Edge of the network, which enables
the execution of delay-sensitive and context-aware applications close to the field. The technology also alleviates backhaul utilization and computation at the core of the network [34]. Edge Computing infrastructures move some parts of the system’s Computing power from the Cloud to its Edge nodes, which improves latency and mobility of the overall system and reduces the system load of the Cloud. [29]

Edge Computing improves the scalability through the decentralization of the storage and processing [29]. Due to the improved scalability, Edge Computing facilitates large-scale distributed applications involving multiple plants or factories, which handle data streams from a large number of CPS [29]. When Edge Computing is combined with Cloud Computing, tasks can be delegated depending on their scale. Tasks with a large-scale can be delegated to the Cloud while delegating tasks with a small-scale to Edge-Computing nodes [21]. Besides the possibility of providing low-latency, there is a lack in the security, which leads to a special need in trustworthiness. Due to the mutual change of information between Cloud, Fog and end-user, the whole infrastructure has plenty sources of data retrievals. To maintain the sovereignty of data in the system, corresponding control instances are required. [35] However, Edge Computing devices can be used as a first control instance for encryption and verification [36].

Edge Computing infrastructures move some parts of the system’s Computing power from the Cloud to its Edge nodes, which improves latency and mobility of the overall system and reduces the system load of the Cloud. Furthermore, Edge Computing allows near-real-time applications for analyzing process data at shopfloor level or controlling CPS like machines or industrial robots. It provides low-latency in a distributed network [35]. The wide spread geographical nature of the edge computing technology allows proximity processing close to the shopfloor and consequently a lower latency than in purely Cloud-centric IT infrastructures [29]. The bottleneck problem regarding the data transmission and storage already mentioned in the single Cloud Computing can be bypassed by the principle of Edge Computing [35]. It also supports the processing of tasks immediately or rather in real-time [7]. Edge Computing provides limited processing power [37], whereas the Cloud takes over large or powerful processing tasks [31]. In terms of flexibility, Edge Computing solutions offers the ability to flexibly (re)configure real-time automation flows [29,37].

4.3 Edge-Cloud-Computing with Blockchain-Technology

The third stage observed is an IT infrastructure supplemented by the use of private Blockchain solutions. A Blockchain is a DLT that stores data in time stamped blocks. The blocks are irreversibly chained to their respective predecessors by hash functions [38]. In contrast to public Blockchains, private solutions offer a permissioned access and adjustable level of transparency. In the scenario of handling data streams, the distributed nodes of a private Blockchain can be used as a general purpose database. All data stored, therefore benefits from Blockchain advantages, such as timestamps, immutability and possible verifiability. Especially in an environment with numerous CPS interacting with each other, these characteristics can be of use to store evidence about interrelated communications and payments. [39] The infrastructure of the Blockchain provides the involved parties with an insight into all transactions made, as they are stored in the distributed system in a traceable manner [8]. Furthermore, it is possible to have Smart Contracts run on a Blockchain, which are thus able to trigger and run contract arrangements automatically or in future autonomously and thus opens up an enormous potential for the automation and autonomy of business processes [40].

Blockchain technology has the potential to truly decentralize the way data is stored and managed without the need of a middlemen or third-party involvement. Due to the distribution of data across the network nodes, data is inherently secured by not having a single point of failure. [41] In the case of CPPS, Blockchain-based systems could empower organizational units to act decentralized and autonomously on the basis of shared Smart Contracts [40]. Apart from the system itself, that benefits in terms of security through distributed nodes, there also is the data storage technique itself as a unique feature of Blockchain technology. Already by its nature, entries to the Blockchain are stored in blocks and protected against manipulation by
cryptography and hashing. In case of an interruption of the hash sequence, this would be immediately identified as a manipulation, visible throughout the whole Blockchain. [40]

The supplementary use of Blockchain solutions can enhance flexibility, as involved network partners have a more transparent information situation. Especially when it comes to cross supply chain problems, partners benefit from the data accessibility of a shared ledger and the use of Smart Contracts. Apart from only processing data, they can be directly linked to the exchange of any sort of asset. In this context the potential of the Blockchain goes even further, since financial transactions can be managed additionally to data that is related to the flow of materials or information [40]. Based on these advantages, the use of enterprise Blockchain frameworks can lead to more efficient business processes as well as increased transparency and flexibility [42]. On the other hand, even though Blockchains build system bridges between their network partners, they still have to become interoperable among each other. Today, there are many initiatives working on solutions [42,43], but in fact current Blockchain pilot projects still have lack standards when it comes to consensus and hashing algorithms [44]. Hence, most solutions as of now are designed as standalone systems and interoperability between different frameworks still needs to be established. [45].

Another challenge for Blockchain solutions definitely lies within its scalability and latency [46,47]. In particular, the number of possible transactions within a fixed timeframe, the block size and number of involved network nodes constitute determining factors for these categories. As the factors differ between different frameworks, but are delimiting most of the current private Blockchain solutions, scalability and latency can be seen as main factors for the slow pace of industry adoption. Furthermore, they constitute a reason why most of the current Blockchain projects still remain in proof-of-concept stage. [48,49]

### 4.4 Summary

Based on an exploration of scientific paper dealing with data streams in CPPS, suitable criteria for the evaluation of the selected IT infrastructures were identified. The findings of the evaluation were subsumed in a table, which contains the suitability of the IT infrastructure for CPPS in one column and an explanation in a second column, which underpins the respective assignment (see Table 1).

Table 1: Overview of different supplements on IT infrastructures and its affections referring to the identified criteria

<table>
<thead>
<tr>
<th>Criterion/IT infrastructure</th>
<th>Cloud Computing</th>
<th>Supplemented by Edge Computing</th>
<th>Supplemented by Blockchain Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalability</td>
<td>● - able to handle different sizes of data volume</td>
<td>● - load-specific distribution of data processing tasks</td>
<td>○ - differing transactions per second and number of network nodes - determined storage capacity</td>
</tr>
<tr>
<td>Security</td>
<td>○ - prone to cyberattacks - tracking of third parties very difficult - lack of security</td>
<td>○ - more sources of third-party attacks - Edge devices can perform as a control instances</td>
<td>● - security of the system through redundant data storage - temper-proof data through time-stamped and hashed blocks</td>
</tr>
<tr>
<td>Latency</td>
<td>○ - prone to performance loss through “bottleneck” processes</td>
<td>● - lower latency than purely Cloud-centric IT infrastructures</td>
<td>○ - limited transactions per second due to the limited time for the creation of new blocks</td>
</tr>
<tr>
<td>Processing</td>
<td>● - able to process large data-sizes - strong in combination with machine learning</td>
<td>● - allows proximity processing close to the shopfloor</td>
<td>● - Blockchain has no aim to process large data-sizes - able to process Smart Contracts within a network and involve financial transactions</td>
</tr>
<tr>
<td>Flexibility</td>
<td>● - resource flexibility - allocation of data streams</td>
<td>● - offers the ability to flexibly reconfigure real-time automation flows</td>
<td>○ - Blockchain frameworks still have to become interoperable - flexibility enhancement due to transparent information status</td>
</tr>
</tbody>
</table>

Legend ● extensive supplement ○ partial/conditional supplement ○ no significant supplement
5. Conclusion

The study of scientific work reveals that the requirements for handling data streams in CPPS can be subsumed to the five major criteria ‘Scalability’, ‘Latency’, ‘Security’, ‘Processing’ and ‘Flexibility’. It can be stated, that there are high potentials in the combination of different IT infrastructures, which overwhelms the single use of one technology. The expansion of Cloud Computing through Edge Computing offers advantages, especially in terms of scalability and latency. The adaption of the Blockchain provides potentials in terms of the security. Referring of the scalability, the processing of a high number of transactions, e.g. in a multi-company application, presents a challenge and leads to an interest for further research. Therefore, the supplementary use of Blockchain technology for a CPPS is useful if it requires external entities and a secure and traceable way of data handling. Additionally, the Blockchain-based use of Smart Contracts offers new ways and possibilities to automatize processes, incl. payments, between different CPS.

The criteria were derived from a more technological point view. Thus, financial aspects are not considered in this scope. Furthermore, the evaluation was done qualitatively. Additionally, no differentiation of specific forms of CPPS was done. Apart from Blockchain technology, there is a variety of other DLT such as directed acyclic graphs (DAG) [50], that did not get to reach that much attention in literature yet and aren’t covered in this paper. In order to address our scalability concerns, the handling of CPPS related data streams with different DLT such as DAG should be analysed in future research. Moreover, future research should be performed by testing concrete Blockchain solutions along with Cloud and Edge Computing concerning our derived criteria as well as adding an economic evaluation. As most of the current enterprise projects still remain in proof-of-concept stage, it is necessary to develop more mature pilots for adequate testing.

Acknowledgements: The work presented in this paper was funded by the ‘Deutsche Forschungsgemeinschaft (DFG) - 276879186/GRK2193’.

References


Biography

**M.Sc. Nick Große** (*1992) has been research assistant of the Chair of Enterprise Logistics (LFO) at the University Dortmund since 2019 under the leadership of Prof. Dr. Michael Henke. He is a member of the research training group ‘Adaption Intelligence of Factories in a Dynamic and Complex Environment’(GRK 2193).

**M.Sc. David Leisen** (*1993) had been research assistant of the Chair of Material Handling and Warehousing (FLW) at the University Dortmund since 2019 under the leadership of Prof. Dr. Michael ten Hompel till September 2019. He is working at ‘Deutsche Telekom’-company in Bonn since October 2019.

**M.Sc. Tan Gürpinar** (*1991) has been research assistant of the Chair of Enterprise Logistics (LFO) at the University Dortmund since 2018 under the leadership of Prof. Dr. Michael Henke. His research focuses on blockchain technology in the field of supply chain management.

**M.Sc. Robert Schulze Forsthövel** (*1992) has been research assistant of the Chair of Material Handling and Warehousing (FLW) at the University Dortmund since 2018 under the leadership of Prof. Dr. Michael ten Hompel. He is also member of the research training group GRK 2193.

**Prof. Dr. Michael Henke** (*1971) has been the chairholder of the Chair of Enterprise Logistics (LFO) at the University Dortmund and director of the Fraunhofer Institute of Material Flow and Logistics (IML) since 2013. He is also an adjunct Professor for Supply Chain Management at the School of Business and Management at the Lappeenranta University of Technology in Finland since 2015.

**Prof. Dr. Dr. hc. Michael ten Hompel** (*1958) has been the chairholder of the Chair of Materials Handling and Warehousing (FLW) at the University Dortmund since 2000 and managing director of the Fraunhofer Institute of Material Flow and Logistics (IML) since 2004.
A Framework for Online Detection and Reaction to Disturbances on the Shop Floor Using Process Mining and Machine Learning

Markus Fischer¹, Mahsa Pourbafrani², Marco Kemmerling³, Volker Stich¹

¹Institute for Industrial Management (FIR) at RWTH Aachen, Aachen, Germany
²Chair of Process and Data Science (PADS) RWTH Aachen, Aachen, Germany
³Institute of Information Management in Mechanical Engineering RWTH Aachen, Aachen, Germany

Abstract
The shop floor is a dynamic environment, where deviations to the production plan frequently occur. While there are many tools to support production planning, production control is left unsupported in handling disruptions. The production controller evaluates the deviations and selects the most suitable countermeasures based on his experience. The transparency should be increased in order to improve the decision quality of the production controller by providing meaningful information during his decision process. In this paper, we propose a framework in which an interactive production control system supports the controller in the identification of and reaction to disturbances on the shop floor. At the same time, the system is being improved and updated by the domain knowledge of the controller. The reference architecture consists of three main parts. The first part is the process mining platform, the second part is the machine learning subsystem that consists of a part for the classification of the disturbances and one part for recommending countermeasures to identified disturbances. The third part is the interactive user interface. Integrating the user’s feedback will enable an adaptation to the constantly changing constraints of production control. As an outlook for a technical realization, the design of the user interface and the way of interaction is presented. For the evaluation of our framework, we will use simulated event data of a sample production line. The implementation and test should result in higher production performance by reducing the downtime of the production and increase in its productivity.

Keywords
Disturbance Management; Deviation Detection; Production Control; Process Mining; Machine-Learning; Internet of Production; Decision Support

1. Introduction
The high number of customer demands for individualized products with short delivery times leads to increased perceived complexity for companies. Not only the requirements of end-users but also those within a supply chain are rising and increasing the requirements for companies’ order processing [1]. In order to meet these requirements, companies must design efficient production systems that are optimally aligned with the conflicting goals of logistics. An efficient configuration of the production system also includes robust handling of internally and externally induced disturbances. Currently, the task of disturbance management is the responsibility of the production controller. The production controller faces the challenge of making high-quality decisions for selecting a suitable countermeasure as quickly as possible. The different goals of the production and those of the customers must be weighed against each other, which is very difficult due to
the complex interdependencies on the shop floor. Often the production controller is insufficiently supported by IT systems and therefore relies on experience. In the area of production planning and control, it is expected that decision support systems will improve your decision-making processes and reduce the probability of making the wrong decisions [2]. Classical methods of production control are not able to cope adequately with current developments and ongoing changes, therefore the aim of this research is to develop a decision support system to improve decision quality [3]. To address these issues, the paper presents a framework for online detection and reaction to the disturbances on the shop floor using process mining and machine learning.

2. State of the Art

This section explains the terminology used in production control and deviation management in order to place the work in this context. It continues with a short definition of decision support systems and focuses on the current state of research in the area of detection and response to deviations. For this purpose, the approaches are structured in various groups, which differentiate by the used approaches for supporting the production control in handling the deviations.

2.1 Terminology

**Production Planning and Control**

Production planning and control (PPC) aims to deliver the customers’ products in the right quantity and quality on time with minimal inventory and high utilization of the company’s resources. To meet the requirements of an economic production line, production planning plans the necessary operations for the production of the product [4]. The material, resource and employee capacities are taken into account in the planning process. A distinction is made here between rough and detailed production planning. After the rough planning and the termination of the production orders, production control starts with the detailed planning of the operations. Production control also has the tasks of order and resource control. Within these latter tasks, the production controller tries to ensure adherence to the production plan [5].

**Disturbance Management**

The production is a dynamic environment with disturbances and deviations that must be dealt with as part of production control. For this reason, the task of deviation management plays a decisive role in production control [6]. To understand disturbance management, first, we define the deviation and disturbances in the next paragraph. Thereafter, disturbance management is defined as a part of production management.

Disturbances are regarded as unplanned and unpredictable deviations from a planned state, which have a direct influence on the process chain. Without intervention, a malfunction is accompanied by a reduction in the performance of the production system [6]. Deviations, on the other hand, are identified by comparing planned and actual values. Deviations differ from disturbances in a way that a deviation does not necessarily have negative consequences for a production system. Only when a certain tolerance corridor is violated, it will be referred to as a disturbance. Regular violations of the tolerance corridor are referred to as systematic disturbances. Irregular deviations occur in the case of accidental disturbances. Since disturbances are defined as differences with a negative impact on the plan, it is the task of the production controller - following the definition of production planning and control - to realize plan compliance. Therefore, disturbance management is to be regarded as part of production control [7].

2.2 Decision Support Systems

Decision support systems (DSS) are IT-based systems, which enable the users to access data across the company to analyze it and evaluate different alternatives [8]. The goal of DSS is to improve the quality as
well as the responsiveness within the decision-making process of the user. Due to the tasks of a decision support system, Sauter [9] structures DSS into a data component, a model component, and a user interface. To achieve the goal of a DSS, the data component compiles data from various sources in the company and puts it into logical relation. The data is then processed in the model component to create statements about possible future alternatives. The models are tailored to certain questions and can range from simple to complex mathematical models. The evaluation of the models is done via a user interface, which enables easy operation of the DSS. In this way, the user can not only see the results of the model but also make adjustments to a model. The adaptation of the model as well as the possibility of independent data analysis distinguishes DSS from other decision-support options, e.g., a neural network, which is only adapted to one specific decision situation [9]. For designing a DSS, four criteria are relevant: “robustness, ease of control, simplicity, and completeness of relevant data” [8].

### 2.3 State of the Research

This section focuses on the state of the current research in the fields of disturbance management and compares different approaches to support the production controller in automatic disturbance handling. Various approaches exist to support knowledge-based tasks and decision support in the company. The approaches in production control and disturbance management can be divided into the approaches of methodical support, simulation-based support, and machine learning-based support.

Meissner [7] develops adaptive deviation management which enables the production controller to classify deviations. The classification helps in differentiating the level of criticality and the kind of countermeasure, which is required to handle this deviation situation. Meissner develops a reaction-strategy-matrix for integration in production control; however, automatic support for the production controller is not developed.

Galaske and Anderl [10] present a simulation-aided decision support tool for the disruption management of Cyber-Physical Production Systems. Interdependencies between the different disruptions events are analyzed, resulting in a consistency matrix. Based on the consistency matrix, the different events are clustered, allowing to create disruption scenarios for the different categories of events. For every disruption scenario, several actions are defined. A simulation tool is used to evaluate the effect of each action on the corresponding scenario. Based on key performance indicators for each disruption scenario, the best action is chosen. Genc [11] also uses a simulation-based technique with a stronger focus on the early detection of deviations. Therefore, he defines several critical events that should alert the decision-maker to take action. Based on a classification of events, a number of possible actions are given, which is further reduced by using several restrictions. The remaining actions are evaluated using simulation for being able to choose the best possible action for every disruption event.

Priore et al. [12] and Khosravani et al. [13] use the technique of case-based reasoning as an approach to knowledge representation in the field of machine learning. Priore et al. [12] use it for real-time scheduling of flexible manufacturing systems by varying the release rules. Besides CBR they also use Support Vector Machines. Khosravani et al. [13] focus on error identification and correction in an injection molding machine. Both approaches show the potentials of knowledge representation by means of case-based reasoning, but there are no approaches for production control of the shop floor. Krumreich et al. [14] present a reference architecture for complex event processing to analyze continuous production processes in the process industry. The aim of the complex event processing is to predict the future events which may trigger further events. With this approach, the architecture of the system provides many future scenarios based on sensor data provided form the production processes. He aims to provide insights on possible process results and to simulate countermeasures to the different scenarios. The paper pursues a similar research goal but focuses only on a production process in the process industry and not on a production process in the shop floor production of discrete manufacturing as in this paper.
The current state of research shows that a lot of approaches were developed to support the production controller in disturbance management. No approach exists that is using process mining and machine learning to set-up a sustainable solution that also adapts to changes in the production system.

3. Framework for the online detection and reaction to disturbances

This section deals with the developed framework for online detection and reaction to the deviations on the shop floor. As proposed, the aim of the paper is to develop a decision support application for the production controller. This section first introduces the structure of the decision support application and then describes the important subsystems and technologies in detail.

3.1 Structure of the decision support application

The structure of a decision support application was described in section 3. In this section, we focus on the interplay of the subsystem: data component, model component, and user interface. The framework is shown in Figure 1. The decision support application will interact with the databases of the company’s IT and the production controller as the designated user.

![Figure 1: Structure of the framework for a decision support application for online detection and reaction to disturbances (based on [9])]

The decision support application will get the feedback data from the shop floor. Depending on the IT architecture of the company this will be either stored in the Enterprise Resource Planning (ERP) system or in the Manufacturing Execution System (MES). The feedback data includes start- and end-timestamps of operations, as well as the respective resources the order used. Also, the routing will be transferred to the data component, as it will be used as the planned process. This data will be transferred to the first part of the model component, the process mining subsystem. The aim of the process mining subsystem is to analyze the actual process flow in real-time [15]. This will be used to detect and predict deviations and disturbances and to provide information on the performance of the production lines. The process mining can detect deviations regarding a violation of the process flow or a violation of operation times [15]. However, in the next phase, the process mining subsystem can be improved by adding simulation techniques, e.g., system dynamics simulation model [16] to increase the ability of the application in providing more accurate recommendations.
The performance information will be provided to the user interface while the information on detected deviations will be sent to the machine learning subsystem for labeling the disturbances. This subsystem decides whether the detected deviation is a deviation within the acceptable tolerance limits or if it is a disturbance to which a reaction should be derived. The labeling subsystem provides information about disturbed orders to the user interface. There, the production controller can give feedback if the labeling worked and can also train the system with this information after the initialization. Also, the subsystem forwards the information to the second machine learning system, the recommender system. The recommender system comprises two subparts. The first subpart is the pre-trained system, which will be set up initially to suggest reactions to the identified disturbances. The subsystem should suggest up to three countermeasures. The second part involves continuous machine learning based on the controller’s feedback on the proposed measures and their effectiveness. This should enable an adaptation to the constantly changing constraints of production control. The two main advantages are the support of the production controller in the decision-making process to reduce the time to evaluate the situation and the decision making itself. The second advantage is knowledge transfer into the database. This enables intra-company knowledge storage in a field that is often driven by the experience of the employees.

### 3.2 Process Mining

Process mining is a relatively new research discipline which “sits between data science and data mining” [15]. Its idea is to model and analyze processes with the goal of discovering, monitoring or improving the real process based on data. The data for process mining is stored in the forms of event logs [15]. An event log is a set of traces and each trace is a sequence of events and each event includes at least case ID, activity, timestamp, and resource data. Moreover, other attributes can be added for further investigations. The case ID describes a process flow of one real case, e.g. one customer or one order in production. The activities describe the steps which the case performed from the start of the process to its end. The other fields allow further information to be included in the analysis of the process cases. For instance, performing welding (activity) for item 12 (case ID) in the production line by one of the workers Jack (resource) at 12:00:10 10.10.2019 (timestamp) is an event and if it takes ten events for item 12 to be ready in the production line then these 12 event forms a trace in the event log.

As described before, process mining can be used for discovering different process variants based on historical data. If no process model exists (neither target nor actual process), the process model can be used to derive the possible process variants and their characteristics from the data. The advantage of process mining is that all process variants and their frequencies can be displayed. This way, the most frequent process variant and the most frequent deviations can be analyzed. With conformance checking, deviations can be detected and examined using an existing process model. For this purpose, however, a process model must already exist against which the various process cases can be checked. Conformance checking enables process monitoring in the way of disturbance management. Both techniques can be used to improve processes [17].

In this framework, we perform bottleneck analysis within process mining to detect and predict deviations based on the provided data, i.e., event logs. We do so by using conformance checking to detect deviations in the flow of the activities and performance analysis [17] to detect performance deviations. The case ID will be a unique order in the Shopfloor the activities will be work steps provided in the routings.

### 3.3 Machine Learning

The field of machine learning is composed of methods that perform specific tasks by inferring relationships from given examples but without being explicitly programmed for the task at hand. The practical applications of machine learning systems are usually divided into a training phase in which a model is learned and a subsequent phase where the model is applied to previously unseen data or circumstances. A successfully trained model is able to generalize from the data distribution encountered in the training phase in such a way
that it is also able to perform the task on the unseen data [18]. In the field of machine learning, there are different forms of learning, which are relevant to the framework of online detection and reaction to disturbances. For the two tasks of the machine learning system, the approaches are explained in the following.

3.3.1 Labeling of Disturbances

For the recommendation system labeled disturbances are needed as input data. From the process mining part of the framework, deviations will be forwarded to the machine learning system. The first task is to label the data. Therefore a case-based approach was carried out to enhance the current state of the literature on known disturbance classifications. The disturbances found by literature and interviews are classified by the 5M method [19]. This proposed classification of the disturbances builds the knowledge base for the machine learning system. Here, the task of classification will be carried out using supervised learning techniques. The user can label new data by selecting a proposed disturbance and give feedback about whether the system is correct.

3.3.2 Recommender System

The recommender system has the task to propose suitable reaction measures for the identified disturbances. In general, recommender systems suggest items to a user [20]. In this case, an item is a countermeasure to a found disturbance in the Shopfloor. The recommender system gets the classified disturbances as an input. Here, two possible approaches have been identified for building the recommender system: collaborative filtering and reinforcement learning.

Collaborative Filtering is a very promising technology for filtering data based on similarities. Collaborative filtering is often used in e-commerce. There collaborative filtering recommends items that other users bought who also bought the currently viewed item. The collaborative filtering algorithms (CFA) produce a predicted likeliness or a list of top X suggestions for the user based on the input. The input is often user-based, e.g. previous likings or other user’s likings [21,22]. For the production of either a predicted likeliness or a list of top X items, collaborative filtering algorithms can be divided into two classes: memory-based CFA or model-based CFA. The memory-based approach uses the entire user database to produce the wanted output. The system searches for the nearest neighbors of the “active” user based on their history. Then the system uses different algorithms to combine users’ preferences and decisions to predict the suggestion. The model-based CFA first builds a model of user ratings. One approach is the clustering model. The models structure the problem for collaborative filtering as a classification problem and therefore cluster users in classes. The active user is classified and then the particular class is analyzed and a prediction is built upon this class and not the whole user database [21,22]. In the case of the proposed framework, the user would be a production controller and the items would be disturbances and the output would be a list of possible countermeasures to handle the disturbance.

A second approach is to model the problem as a Markov decision process and subsequently solve it using reinforcement learning. A Markov decision process consists of a set of states, a set of actions, a transition function, and a reward function. Each state is a representation of the system at a particular point in time. A reinforcement learning agent interacting with the system perceives this state and performs some action to influence it. The outcome of performing a particular action in a particular state is a new state, which is determined by the transition function. After performing an action, the agent receives this new state as well as a reward to indicate how good the selected action was. While training, the agent learns a policy which determines its behavior in order to maximize its long-term cumulative reward [23]. In the use case described here, the state describes the shop floor with its disturbances, while the actions are countermeasures to the identified disturbances. Finally, the reward function implicitly defines the goals of the agent and thus needs to be carefully designed. A possible approach might be to compare the state of the system after applying the
appropriate countermeasure to some disturbance with the state the system would be in had no disturbance occurred in the first place. In essence, some distance measures between states could be developed in this way such that the agent receives a positive or negative reward proportional to how much closer it moved the state of the system to the ideal state.

3.4 User Interface

The user interface is very important for a decision support application since it is the gateway to provide the relevant information to the user. Its user-friendliness and comprehensibility of the presented information improve the use and trust in the system. The user interface of the decision support application is provided as a web application. This enables the use of mobile devices so that the production controller can retrieve information even during a tour of the plant. The web interface is shown in Figure 2.

![Production Control Cockpit](image)

**Figure 2:** Interactive User-Interface for the decision support application developed to support the decision making in disturbance management

At the top, the user sees a list of all orders for which the decision support system has identified a malfunction. With the aim of enhancement techniques in process mining, the online track of the product would be available through the process. This online tracking would provide this list of orders and their possible deviations. If the user now selects a malfunction, the station is displayed in a plant layout in addition to the performance information from the process mining. The user can then also indicate whether the displayed deviation is actually a disturbance. If the production controller wants to see all open orders, he can have them displayed via a filter function and, if necessary, display the open orders as malfunctions. This information is then passed on to process mining and machine learning subsystems.
When a disturbance is selected, the user is also shown three countermeasures. The implementation of the measures lies with the production controller. Here, however, the user can specify whether he has implemented the measure and later also evaluate the effectiveness of the measure. This allows machine learning subsystems to learn continuously and improve their performance. This makes it possible to adapt to changed constraints such as new products or new routings. In addition, the user can enter new measures and thus expand the knowledge database of reaction strategies. The user can specify whether the measure is applicable to this or all fault classes.

4. Conclusion

This paper proposed a framework for decision support for a production controller in order to react to identified and predicted disturbances. Besides the framework, the general idea for the use of process mining in disturbance management has been described and the current state of the research was discussed. Moreover, process mining techniques provide enhanced analyses which makes the online track and detection of disturbance possible. In addition, the article showed how to use machine learning next to process mining to enable a recommender system to support the production controller. The framework and subsystems of the framework were described to test the system with real data. The paper at hand should form a working and discussion basis for further research in disturbance management based on the proposed system framework.

In the future, this framework will be tested with simulation and real data within the research program Cluster of Excellence “Internet of Production”.

Acknowledgments

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References


Biography

Markus Fischer (*1992) is a research associate at the Institute for Industrial Management (FIR) at the RWTH Aachen in the research group Production Control. He studied mechanical engineering at the RWTH Aachen University, specializing in production engineering and mechanical engineering. Currently, he is working in the research project “Internet of Production”, which aims to develop data-driven decision support systems.
Mahsa Pourbafrani (*1989) is a research associate in the Data and Process Science group of RWTH Aachen University. Her research includes simulation and what-if analysis in process mining techniques regarding the performance metrics of a process. She is also working as a scientist on “Internet of Production” project with the focus of supporting decisions in production lines using process mining.

Marco Kemmerling (*1993) is a research associate at the Institute for Information Management in Mechanical Engineering at RWTH Aachen. He studied Data Science at Maastricht University and is currently doing research in applications of data science and machine learning. Among other projects, he is working on applications of reinforcement learning within the project “Internet of Production”.

Prof. Dr.-Ing. Volker Stich (*1954) has been CEO of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 1997. Prof. Dr.-Ing. Volker Stich worked for 10 years for the St. Gobain-Automotive Group and led the management of European plant logistics. In addition, he was responsible for the worldwide coordination of future vehicle development projects.
Efficient Task Realizations in Networked Production Infrastructures

Marcus Grum¹, Benedict Bender¹, Norbert Gronau¹, Attahiru S. Alfa²,³

¹University of Potsdam, Potsdam, Germany
²University of Pretoria, Pretoria, South Africa
³University of Manitoba, Winnipeg, Canada

Abstract

As Industry 4.0 infrastructures are seen as highly evolutionary environment with volatile, and time-dependent workloads for analytical tasks, particularly the optimal dimensioning of IT hardware is a challenge for decision makers because the digital processing of these tasks can be decoupled from their physical place of origin. Flexible architecture models to allocate tasks efficiently with regard to multi-facet aspects and a predefined set of local systems and external cloud services have been proven in small example scenarios. This paper provides a benchmark of existing task realization strategies, composed of (1) task distribution and (2) task prioritization in a real-world scenario simulation. It identifies heuristics as superior strategies.

Keywords

Industry 4.0; CPS; Decentral Decision Making; Industrial Analytics; Case Study

1. Introduction

Faced with an increase in the complexity of company IT infrastructures, such as an increasing number of networked machines and their heterogeneity in hardware and software [1], companies are often challenged with dimensioning capacities for processing of time-critical analytical tasks. Since companies aspire to realize tasks using the least resources as possible, new infrastructure trends emerged to cost-efficiently use existing resources: examples refer to Cloud-Computing, Edge-Computing, hardware outsourcing and adaptable enterprise architectures. Beside costs, as a trade-off, further criteria need to be considered, so that e.g. jobs are realized in time. Assuming to have internal and external processing systems with different sets of rights per layer and capabilities, generic multilayer hardware infrastructure models seem to be promising ways to achieve scalable and flexible infrastructures without maintaining oversized IT-infrastructures [2].

Given the many heterogeneous computing devices in modern production infrastructures, questions about the efficient use of Industry 4.0 resources further complicate decisions about the organization of analytical infrastructures. For this, a variety of task distribution approaches within networked infrastructures have evolved, each exhibiting characteristic advantages and disadvantages [3]. Under these, one can find a tendency for decentralized, heuristic approaches [2], since these simplify the strategy determination.

Former studies approached the evaluation based on hypothetical data [2], which seems reasonable as a start. To confirm validity of research results established, this current contribution is based on a real-world practical case study and thereby expands the collection of small and theoretical benchmarks considered so far. By the use of a realistic example containing numerous systems and public cloud services, working on real system data satisfies complexity of Industry 4.0 production. So, we aim to answer the following research question:
How can analytical tasks be efficiently distributed and processed within networked production infrastructures?

2. Theoretical Foundation

For the examination close to reality, this contribution considers different forms of analytical tasks, that are realized by various levels of computing infrastructures within the context cyber-physical production systems. The first subsection therefore describes theoretical terms and principles corresponding to the related literature. The second subsection then presents underlaying concepts, on which the contribution builds: the benchmark carried out considers a comparison of different task realization strategies as well as the new heuristic approach called NDM, and further the performance is measured by an adequate framework.

2.1 Related Literature

Cyber-Physical Production System. As a result of integrating cyber-physical systems (CPS) into production systems, more and more manufacturing components become connected. The combination of multiple cyber-physical systems within one common production setting can be referred to as Cyber-Physical Production System (CPPS) [4, 5]. Because of intelligent components being distributed within production environments, a cooperative planning can replace the traditional linear, hierarchical planning and control [6]. To realize the production process, a complex interplay of CPPS components is necessary similar to autonomous logistic systems (e.g. [7]). Therefore, each CPS interacts autonomously based on its decision strategies to realize a cooperative planning solution. These are considered as analytical tasks from hereon.

Computing Infrastructures. Different concepts exist to classify computing infrastructures of companies. While each company might have a unique infrastructure, common elements, structures and hierarchies can be identified. Issuing the distribution of tasks, a generic concept is desirable that firstly provides typical infrastructure levels, e.g. for differentiating the processing power, and secondly is able to specify levels dependent on the system’s individual situation. In the following, a separation in three typical infrastructure levels is used [2, 8], by which computing systems are grouped into three levels (Table 1).

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPS level</td>
<td>Computing resources being situated at the shop floor level. Subsumed as CPSs, they are able to sense their environment, process their perception in order to derive decisions, communicate with other CPS, and carry out decisions with the aid of actuators. Hence, they are attractive to carry out analytical tasks by themselves.</td>
<td>[4]</td>
</tr>
<tr>
<td>Local Cloud</td>
<td>Computing resources being interconnected by local or virtual networks within and across company sites [9]. These are typically more powerful than a single CPS [8] and therefore attractive to support CPS level systems for the realization of heavy analytical tasks.</td>
<td>[8], [9]</td>
</tr>
<tr>
<td>Public Cloud</td>
<td>Computing resources of external parties being rented for the heaviest computing tasks. Being integrated as service, they subsume highly scalable resources supplementing a company’s infrastructure. Hence, they are attractive to quickly disburden the company in the realization of analytical tasks.</td>
<td>[10]</td>
</tr>
</tbody>
</table>

Industrial Analytics. Oriented to [11], analytical tasks of Industry 4.0 production systems can be categorized by eight task types being able to be processed in parallel [12]. These task types refer to Business Intelligence (BI) categories, such as reports, spontaneously instructed reports, drilldowns and alarms, as well as Business Analytics (BA) categories, such as statistical analyses, intra- and extrapolation, prediction models and scenario-based optimizations (increasing order). According to the assumption of Davenport and Harris, the effort of higher task types rises while resulting in higher competitive advantages.
2.2 Underlying Concepts

**Efficient Task Realizations.** Being faced with an architectural design, that provides various computing infrastructure layers for the realization of an analytical task, these differ in processing capability (supply), and task types requiring for individual processing capabilities for an analytical task, the optimal fit of demand and supply needs to be identified according to economic objectives to be identified as efficient.

The realization of an analytical task is operationalized by the following three: (1) the analysis of available tasks at any system being considered as origin, (2) the transferring of generated tasks based on the system’s transfer strategy, this can be referred to as task distribution, and (3) the processing of arriving tasks based on a processing strategy, this can be referred to as task prioritization and assigns an order to arriving tasks [2]. While the transferring strategy e.g. can refer to heuristic strategies, processing strategies can refer to consumption sequence procedures, such as Last-In-First-Out (LIFO), First-In-First-Out (FIFO), Slowest-In-First-Out (SIFO) according to the collection of [2].

**The New Decision Maxim (NDM).** The NDM proposed by [2] serves as heuristic decision guideline for CPSs to decide at which system to process analytical tasks. It is so suitable to be a transferring strategy of tasks. Depending on the level where the decision is taken, different options exist. According to the current design of the generic architecture presented, three layers and four options can be identified. These are visualized in Table 2 and described thereafter.

<table>
<thead>
<tr>
<th>Option / Level</th>
<th>Vertical-up</th>
<th>Vertical-down</th>
<th>Horizontal</th>
<th>Local</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPS level</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Local cloud</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Public Cloud</td>
<td>-</td>
<td>X</td>
<td>(X)</td>
<td>X</td>
</tr>
</tbody>
</table>

First, tasks can be processed locally, whereby the entity computes the results itself. Task can be distributed vertically. This includes the distribution to higher levels (upward distribution) or lower levels (downward distribution). Tasks can also be distributed to systems on the same level (horizontal distribution). Dependent on collaboration strategies of public cloud service providers, a horizontal distribution is not allowed, which is the reason for the (X) in Table 2.

**NDM Configuration.** While the NDM provides ability to conduct different criteria, we employ the same escalation prioritization as [2] in order to ensure comparability of benchmark results. By this, our contribution refers to the validation of the NDM approach in a realistic case study and performing on real data. So, Figure 1 presents the NDM configuration and clarifies the heuristically escalation strategy selection of an example system of the local cloud level.
The example task is characterized by a task effort of 0.9, a divisibility of 0.2, a priority of 0.5, and a current load of 0.8. Being masked by allocation options of Table 2 and weighted with regard to the evaluation prioritization of Figure 1, the best option is identified by the maxima. For the example task presented, the vertical upward escalation shows the highest value and therefore refers to the best solution in this case.

**Performance Evaluation Framework.** In order to compare different approaches to distribute tasks efficiently, a common evaluation framework is required. This ensures the quantitative compatibility of the results on the one hand and allows for the quantitative comparison of approaches on the other hand.

A framework for the measurement of the observed decision strategy performances was developed by [2]. In there, key performance indicators (KPIs) were defined and optimization functions proposed. Besides single system-specific KPIs, global measures across systems were defined. Further, a global objective function, that considers task realization strategies in regard with a collection of scenarios was presented. Since results are intended to be reproduced and to transfer benchmarks to a greater case study, the following aspects are considered for the evaluation of the global systems. Further details can be found in [2].

- The **total processing costs** as accumulated costs of the individual systems. Processing costs combine cost for the task realization of a certain task, such as processing costs, transportation costs, etc.
- The **total traffic intensity** as the systems current job capacity.
- The **total number** and **total time of waiting jobs** as the number or time of jobs waiting to be processed and being accumulated across individual system levels.
- The **total number of jobs** realized in time, as the number of jobs which were conducted in time, i.e. before a delay occurs in the global system as a result of unfinished jobs.
- The **common objective function** as the joint consideration of relevant criteria over scenarios following [13]. This includes e.g. a trade-off of the waiting time, the remaining time, processing costs, minimal transfer costs and job importance.

### 3. Benchmarks in Real-World Scenarios

Following the procedure model by [2], this section demonstrates a real-world case. The numbering corresponds to the phase numbering as presented.

**1) Project Initiation.** The printing case [14] was chosen with the intention to check the performance of the NDM in a practical simulation setting. Following the Industry 4.0 definition by [4], a typical production system consists of production machines (e.g. printer) and workpieces (e.g. raw materials that are to be processed). Within production, machines are transforming one or more input materials into one or more...
output products [4]. To connect the machinery, conveying belts are used, which ensure that the printing output of a former production step is transported to the machinery of the next production steps.

Each printer can handle different qualities, sizes and variations of paper, such that machinery and workpieces can find different production cycles through comparisons. Further, it is necessary to include parameters like speed, or variables like electricity costs, or even adapting production parameters like a changing degree of quality. It is feasible to include parameters, distinguishing between the machine agents, such as the individual computing power, the CPS specific memory, its proper sensor, actuator and communication capabilities.

2) Objectives. The general objective is to compare three different approaches, that is (1) local processing (no distribution), (2) a human-made manual task distribution approach and (3) the NDM strategy for analytical tasks. While being faced with various material flows through the CPPS, the examination of analytical task realization starts at the generation of analytical tasks at each CPS because of material-consuming production tasks. We so consider analytical tasks only and disregard the flow of materials.

3) Systems and Connections. The arrangement of the machines within the CPPS can be seen in Figure 2(a). Here, one can see machines visualized by dark rectangles and conveyors visualized by gray rectangles. While the process starting point and ending point are visualized with help of white rectangles and serve as interface for the production process entry and outlet, human workers are visualized with help of yellow rectangles. CPS are surrounded by green dashed lines, which shall represent separate rooms or buildings.

In order to consider the missing elements of analytical infrastructures (see section 2.1), Figure 2(b) visualizes the CPS within its analytical context: Above the CPS level having the lowest computing power, one can find one level for local cloud machines and one level for public cloud machines. These contain local or public cloud machines having the same computing power. The limitation to have only three levels is only for this graphical purpose. Since machines do seldomly obtain the same computing power in reality, practically, the model provides several levels of CPS, local clouds and public clouds.

In analogy to room separations visualized in green in in Figure 2(a), one can find dashed lines at the local cloud and public cloud levels as well. These rather represent collaboration spaces allowing horizontal distributions. The valid kind of vertical distribution is visualized by blue associations in Figure 2(b).

4) Scenario Collection. During the fourth phase of the procedure model, typical scenarios that impact task realizations were identified. The scenarios include the reference scenario, price increase of public clouds, power enhancement of CPSs, adding computing resources, and breakdown of system elements. During a first simulation, scenarios beside the reference have been disregarded. However, given the practical importance, they are to be considered in the future.
5) Initialization Parameter. The initialization parameters were collected in workshops with the customer. Here, system administrators, production and process owners as well as analytics experts were surveyed. Considering the scenario design of Figure 2(a) and (b), the production setting can be transferred to a computational model as follows. The scenario requires $N=15$ CPS, $M=5$ local cloud systems and $K=3$ public cloud systems. Corresponding to the underlying processes and production runs, each generates independently analytical tasks from eight task types as they have been drawn in section 2.1.

6) Transfer and Processing Strategies. Suppose we have three transfer strategies: A first transfer matrix provides a strategy called 'Do-Not-Transfer-Anything'. Using matrix structure, this looks like an identity matrix since 100% of each task type stays at its origin system.

A second transfer matrix stands for a strategy called 'Workshop-Based'. This has been conceptualized by a group of experts as the procedure model requires in phase six. Having held sessions over the period of two days with 5 experts being responsible for the analytical processing, 4 business process management experts and 4 production managers, various transfer strategies were established individually. These were discussed in groups in regard with the reference scenario selected. Finally, a joint transfer strategy was established, whose visualization can be found in Figure 3. This characterizes a good and reasonable task realization.

A third transfer matrix, coming from the 'New-Decision-Maxim' presented in section 2.2, can be found in Figure 4. Compared with Figure 3, here we cannot see a well-planned, diverse transfer strategy, which results because of the limited set of the four escalation strategies of the heuristic. It characterizes a third party focus.

7) Simulation, Analysis and Comparisons. The initial configuration results in a workload, as can be seen in Figure 5(a). Each task type has been coloured by separate colour. Here, one can see that the traffic intensity of Human 1, Robot 1, Printer 5 and the Waiting Queue exceeds the limit of 1.0, which means that those systems are overloaded and will break down.

Since the traffic intensity shown here is not changed by transfers of the first transfer strategy ('No-Transfers-At-All'), its workload after transfers is the same and an urgent transfer is required. Focusing on the traffic intensity of all local and public clouds, one can identify a potential in free capacities. Ideally, those capacities are used for the processing of tasks exceeding local capacity.

As jobs are transferred, best task realization runs for each transfer strategy can be found in Figure 5(b-c). These have been identified by the objective function, as it was presented by [13]. Figure 5(b) shows system specific traffic intensity following a 'Workshop-Based' transfer strategy. Beginning with the processing of tasks that came in last (Last-In-First-Out-based processing strategy), one can see a significant transfer of tasks to systems having a greater computing power. So, a system overload can be avoided successfully.
Figure 5(c) shows system-specific traffic intensities after NDM-based task transfers. Beginning with the processing of tasks having the lowest remaining time (First-Remaining-In-First-Out-based processing strategy), one can identify even more tasks to be transferred to systems having a greater computing power than the 'Workshop-Based' transfer strategies demands for. Here, a system overload can be avoided successfully, too.

8) Target Concepts. Comparing the results of the previous phase (“Simulation, Analysis and Comparison”), a change to the NDM is suggested. Here, the processing strategy shows minor importance, since all combinations of processing strategies and the new decision strategy improve the situation of the customer. A detailed evaluation can be found in the following chapter.

4. Evaluation

Having rebuilt the simulation setting and results of [2] with the aid of a Python implementation, which is deterministic, as results are reproducible by the same initial conditions, discrete, as simulation time is based on fix time points and process-oriented, as the process steps of analysing, transferring and processing have been considered, the simulation setting has been verified. Having then applied the production scenario described in section 3, the following evaluates given artefacts and creates a trade-off between various approaches. A first subsection deepens the evaluation in focusing the great printing example. Further, the common objective function [13] is verified.
4.1 Heuristic Validation

The simulation of the printing example, as it was shown in section 3, resulted in different performance levels for each transfer strategy. An overview can be seen in Table 3.

Table 3: Transfer Strategy Performance of Printing Center

<table>
<thead>
<tr>
<th>Transfer Strategy</th>
<th>NDM-Based Transfers</th>
<th>Workshop-Based Transfers</th>
<th>NTAA-Based Transfers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Processing Costs</td>
<td>127 734.8158</td>
<td>110 501.3583</td>
<td>94 368.7225</td>
</tr>
<tr>
<td>Total Traffic Intensity</td>
<td>0.4793</td>
<td>0.6828</td>
<td>0.8770</td>
</tr>
<tr>
<td>Total Number of Waiting Jobs</td>
<td>12.1521</td>
<td>17.7485</td>
<td>23.4977</td>
</tr>
<tr>
<td>Total Time of Waiting Jobs</td>
<td>1.5947</td>
<td>1.6604</td>
<td>1.7260</td>
</tr>
<tr>
<td>Total Job Realization in Time</td>
<td>{183 - 184}</td>
<td>{182 - 184}</td>
<td>{164 - 169}</td>
</tr>
</tbody>
</table>

Here, one can see that the most expensive, system-wide task realization is caused by transfers of the New-Decision-Maxim. Workshop-Based transfers come second and No-Transfers-At-All third.

Faced with the total number of jobs, that were realized in time, extra costs can be justified. New-Decision-Maxim-Based transfers and Workshop-Based transfers realized the most tasks in time, although Workshop-Based transfers show a higher variance. Showing 23 systems with 8 job types per system, all 184 jobs types could only be realized in time following the New-Decision-Maxim or Workshop-Based transfers. While New-Decision-Maxim-Based transfers contain 183 out of 184 jobs as well, Workshop-Based transfers show 183 and 182 out of 184 jobs as well. Only 164-169 out of 184 jobs have been realized in time following the No-Transfers-At-All transfer strategy. Mostly, better results are achieved because of their transfer focus on more powerful computers. This is why further KPIs improve with better approaches, as can be seen at the decreasing number of waiting jobs and the decreasing total time of waiting jobs.

Table 4: Process Strategy Performance of Printing Center (Sorted by Total Job Realization in Time)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N-D-M</td>
<td>First-Remaining-In-First-Out</td>
<td>0.0659</td>
<td>31.5115</td>
<td>184</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Low-Importance-In-First-Out</td>
<td>0.0661</td>
<td>31.5404</td>
<td>184</td>
</tr>
<tr>
<td>N-D-M</td>
<td>First-In-First-Out</td>
<td>0.0664</td>
<td>31.5227</td>
<td>184</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Slowest-In-First-Out</td>
<td>0.0671</td>
<td>31.7141</td>
<td>184</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Last-In-First-Out</td>
<td>0.0908</td>
<td>27.8890</td>
<td>184</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Most-Expensive-In-First-Out</td>
<td>0.0914</td>
<td>27.7934</td>
<td>184</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Slowest-In-First-Out</td>
<td>0.0914</td>
<td>27.9510</td>
<td>184</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Task-Type-Ascending</td>
<td>0.0917</td>
<td>27.8663</td>
<td>184</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Most-Expensive-In-First-Out</td>
<td>0.0673</td>
<td>31.6426</td>
<td>183</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Task-Type-Ascending</td>
<td>0.0676</td>
<td>31.7473</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Low-Importance-In-First-Out</td>
<td>0.0920</td>
<td>27.8207</td>
<td>182</td>
</tr>
<tr>
<td>W-B-T</td>
<td>First-Remaining-In-First-Out</td>
<td>0.0915</td>
<td>27.6229</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>High-Importance-In-First-Out</td>
<td>0.0916</td>
<td>27.7376</td>
<td>183</td>
</tr>
<tr>
<td>N-D-M</td>
<td>High-Importance-In-First-Out</td>
<td>0.0670</td>
<td>31.6085</td>
<td>183</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Last-Remaining-In-First-Out</td>
<td>0.0688</td>
<td>31.5933</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Last-Remaining-In-First-Out</td>
<td>0.0915</td>
<td>27.7487</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>First-In-First-Out</td>
<td>0.0913</td>
<td>27.6859</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Task-Type-Ascending</td>
<td>0.0898</td>
<td>27.4073</td>
<td>183</td>
</tr>
</tbody>
</table>

Considering further processing strategies, performance levels could be found as Table 4 shows. Here, task realization strategies (this includes transfer and processing strategies) are sorted by the number of task types, that were realized in time (first criteria), and the common objective function (second criteria).

As best strategies, New-Decision-Based realization strategies can be identified. Those are followed by Workshop-Based realization strategies. The middle field shows a mixture of New-Decision-Based and Workshop-Based realization strategies. Lastly, realization strategies on base of No-Transfers-At-All-Based realization strategies can be identified without exceptions.
While a concrete ranking of all task realization strategies can be found within the table, best strategies focus the remaining time, the importance and the processing costs. In Section 3, only best candidates per category have been presented in detail.

4.2 Objective Function

Focusing the printing case, Table 5 shows a ranking of task realization strategies sorted by the common objective function of [13]. Since this results in a similar order of task realization strategies as a list ordered by the total job realization in time, this is an indicator for a working objective function. A less complex scenario with less conflicting results would underline this perfectly.

As in this complex scenario, a ranking may not be trivially built on base of the number of task types, that were realized in time (first criteria) and the total job remaining time (second criteria), rather a complex trade-off of the remaining time, the importance of a task, etc. was required. One example for this kind of trade-off can be found at the objective function. Considering this specific trade-off, without exceptions, best task realization strategies build on the New-Decision-Maxim. Those are followed by the Workshop-Based task realization strategies without exception. Lastly, No-Transfer-At-All-Based strategies can be found. Since the New-Decision-Maxim considers parts of the objective function, this is not a surprising result. But this serves as indicator for a functioning of the objective function.

Table 5: Process Strategy Performance of Printing Center (Sorted by Objective Function)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N-D-M</td>
<td>First-Remaining-In-First-Out</td>
<td>0.0659</td>
<td>31.3116</td>
<td>184</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Task-Type-Descending</td>
<td>0.0660</td>
<td>31.3825</td>
<td>183</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Low-Importance-In-First-Out</td>
<td>0.0661</td>
<td>31.5404</td>
<td>184</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Fastest-In-First-Out</td>
<td>0.0661</td>
<td>31.4111</td>
<td>183</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Cheapest-In-First-Out</td>
<td>0.0663</td>
<td>31.4920</td>
<td>183</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Fastest-In-First-Out</td>
<td>0.0664</td>
<td>31.5327</td>
<td>184</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Last-In-First-Out</td>
<td>0.0666</td>
<td>31.5545</td>
<td>183</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Last-Remaining-In-First-Out</td>
<td>0.0668</td>
<td>31.5833</td>
<td>183</td>
</tr>
<tr>
<td>N-D-M</td>
<td>High-Importance-In-First-Out</td>
<td>0.0670</td>
<td>31.6085</td>
<td>183</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Slowest-In-First-Out</td>
<td>0.0671</td>
<td>31.7414</td>
<td>184</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Most-Expensive-In-First-Out</td>
<td>0.0673</td>
<td>31.6420</td>
<td>183</td>
</tr>
<tr>
<td>N-D-M</td>
<td>Task-Type-Accending</td>
<td>0.0676</td>
<td>31.7473</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Task-Type-Descending</td>
<td>0.0898</td>
<td>27.4073</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Cheapest-In-First-Out</td>
<td>0.0900</td>
<td>27.5762</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Fastest-In-First-Out</td>
<td>0.0904</td>
<td>27.5857</td>
<td>182</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Last-In-First-Out</td>
<td>0.0908</td>
<td>27.6086</td>
<td>184</td>
</tr>
<tr>
<td>W-B-T</td>
<td>First-In-First-Out</td>
<td>0.0913</td>
<td>27.6850</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Most-Expensive-In-First-Out</td>
<td>0.0914</td>
<td>27.7934</td>
<td>184</td>
</tr>
</tbody>
</table>

Table 5: Process Strategy Performance of Printing Center (Sorted by Objective Function)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>W-B-T</td>
<td>Slowest-In-First-Out</td>
<td>0.0914</td>
<td>27.9510</td>
<td>184</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Last-Remaining-In-First-Out</td>
<td>0.0915</td>
<td>27.7487</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>First-Remaining-In-First-Out</td>
<td>0.0915</td>
<td>27.6229</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>High-Importance-In-First-Out</td>
<td>0.0916</td>
<td>27.7376</td>
<td>183</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Task-Type-Ascending</td>
<td>0.0917</td>
<td>27.9663</td>
<td>184</td>
</tr>
<tr>
<td>W-B-T</td>
<td>Low-Importance-In-First-Out</td>
<td>0.0926</td>
<td>27.6829</td>
<td>182</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>Fastest-In-First-Out</td>
<td>0.1224</td>
<td>33.5232</td>
<td>169</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>Task-Type-Task</td>
<td>0.1230</td>
<td>33.3024</td>
<td>168</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>Slowest-In-First-Out</td>
<td>0.1244</td>
<td>23.5164</td>
<td>168</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>First-Remaining-In-First-Out</td>
<td>0.1248</td>
<td>23.5807</td>
<td>168</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>First-In-First-Out</td>
<td>0.1257</td>
<td>23.7219</td>
<td>166</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>Last-In-First-Out</td>
<td>0.1261</td>
<td>23.6695</td>
<td>167</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>Slowest-In-First-Out</td>
<td>0.1267</td>
<td>23.6994</td>
<td>164</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>Most-Expensive-In-First-Out</td>
<td>0.1271</td>
<td>23.8776</td>
<td>167</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>High-Importance-In-First-Out</td>
<td>0.1275</td>
<td>23.8132</td>
<td>165</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>Last-Remaining-In-First-Out</td>
<td>0.1277</td>
<td>23.7813</td>
<td>164</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>Task-Type-Ascending</td>
<td>0.1286</td>
<td>24.0758</td>
<td>166</td>
</tr>
<tr>
<td>N-T-A-A</td>
<td>Slowest-In-First-Out</td>
<td>0.1289</td>
<td>24.0453</td>
<td>165</td>
</tr>
</tbody>
</table>

5. Conclusion

5.1 Summary

The paper starts with an overview of different important aspects of Industry 4.0 and research that is relevant analytical infrastructures. The combination of the different aspects has resulted in a mathematical model by [2], which builds the basis for the experimentation of a real-world case study presented in this contribution. We provide a first real-world case study, that benchmarks common task realization strategies. These consist of (1) different transfer (distribution) and (2) processing strategies. As such, we contribute to validate the common objective function of [13], that has been applied to a small, theoretical example so far. The results highlight the importance of distribution and task prioritization in networked production infrastructures.

The research question (How can analytical tasks be efficiently be distributed and processed within networked production infrastructures?) can be answered by the selection of the most promising task realization strategy and infrastructure design. The real-world case study demonstrates the strengths of heuristic approaches and is in accordance with results of [2]. This includes the perspective of common KPIs and the common objective...
function of [15]. Considering the objective function as optimization criteria, with this, the AI-based task transfers within the networking infrastructure is prepared.

5.2 Limitations and Outlook

Even though modern analytical infrastructures as well as production settings have been evaluated carefully, individual setups, especially cross-company combinations with multiple sites [15], might not match the structure of this model. Even though the model allows for the modification of arbitrary many and complex levels, insights were derived with three hierarchy levels (Figure 2). For the sake of demonstration simplification was needed. Future research should incorporate more complex and multi-site setups.

Regarding the proposed NDM, further case studies, incl. sensitivity analyses, should be employed to allow for generalization of the results and further increase the current level of validity. The extension beyond the Analytics context allows to apply the results in related domains. Finally, dynamics in distributed infrastructures should be incorporated in the simulation, as these are proposed by disregarded scenarios.

References

Biography

**Marcus Grum** (*1988*) M.Sc.mult. Marcus Grum is working on his Ph.D. at the University of Potsdam. His main research interest is the integration of artificial intelligence in economic contexts.

**Benedict Bender** (*1989*) M. Sc. Benedict Bender is working on his Ph.D. at the University of Potsdam. His main research is the integration of digital platforms and business ecosystems.

**Norbert Gronau** (*1964*) Univ.-Prof. Dr.-Ing. Norbert Gronau holds the chair of Business Information Systems, especially Processes and Systems at the University of Potsdam. His main research activities concentrate on the areas of Enterprise Resource Planning and Knowledge Management.

**Attahiru Sule Alfa** (*1948*) Attahiru S. Alfa is Professor Emeritus at the University of Manitoba, Department of Electrical and Computer Engineering and also a UP/CSIR co-hosted SARCHI Chair Professor at the University of Pretoria, Department of Electrical, Electronic and Computer Engineering. Dr. Alfa’s research covers queueing theory and applications, modelling of communication networks.
Abstract
In the course of the fourth industrial revolution, a rapid technological change proceeds in the manufacturing industry. Numerous new technologies enable multiple opportunities for industrial applications. In order to keep pace with this development, companies are forced to cope with a high amount of new technologies and arising application trends. For a successful positioning, knowledge of the industrial relevance of possible applications and the technologies associated with their implementation is required in particular. In this context, the Fraunhofer Institute for Production Technology IPT and the Centre of Excellence in Production Informatics and Control (EPIC CoE) conducted a two-stage case study to identify and evaluate promising industry 4.0 based application fields, such as self-optimizing production scheduling. The case study is proceeded as part of the European Union's Horizon 2020 research project under grant No. 739592. Within the first stage of this project a systematic screening for industrial application fields was conducted. Several potential application fields were identified and their advantages and disadvantages outlined. Furthermore, the application fields were evaluated according to their potential industrial impact and maturity level. In the second stage, technologies for the implementation of the most promising application fields were identified. At this, technologies were investigated and evaluated according to their readiness level for the identified application fields. In this paper, the methodology as well as the results of the first stage of the study are presented.

Keywords
Industry 4.0; Trend Analysis; Potential Evaluation; Industrial Impact
applications of upcoming technological trends has become a critical competitive factor for success. However, this identification makes companies often facing major challenges [6].

First, the technology foresight in particular constitutes various difficulties for companies. Technology foresight is necessary in order to recognize technologies at an early stage to enable radical innovations [7]. Especially the implicit, unpublished knowledge gives an edge over the competition. Generating this knowledge, however, is particularly difficult for companies in an early stage, due to a lack of access to the required knowledge. The situation is further exacerbated by the fact that companies today operate in a complex environment with many different players and influences [8]. The challenge is to find relevant trends and technology fields in the flood of data and the multitude of data sources [9]. In addition, companies are forced to determine the benefits for their customers and to select appropriate technologies [10]. As a result, the interpretation of complex technological trends is increasingly difficult for companies in practice.

In order to face these challenges and to successfully position the company, knowledge about the industrial relevance of potential applications and the technologies required for their implementation is of major importance. Therefore, the department of technology management of the Fraunhofer IPT conducted a two-stage case study on technology trend scouting for EPIC CoE, within the context of Horizon 2020 research and innovation program. The main objective of this technology trend scouting approach is to identify potential industry 4.0 based application fields for production planning and control (PPC) and to support the associated cyber-physical system of EPIC CoE.

2. Definitions and related work

2.1 Definitions

The following section serves to achieve a consistent understanding on relevant fundamental terms. Therefore, the meaning of industry 4.0 and trend will be detailed below.

2.1.1 Industry 4.0

The given understanding of industry 4.0 exceeds the mere networking of machines and products via the internet. Although modern technologies make it possible to build up an ever broader database, the use of the underlying potential depends as much on the organisational structure and culture of the company. Accordingly, organisational and cultural areas of a company must also be transformed in the course of digitisation. The overall goal is to build a learning, agile company that can continuously adapt to a changing environment [2]. Sub goals of the implementation of industry 4.0 are the reduction of costs, the sustainable production and the optimization of the collaboration productivity, which lead to an improvement in infrastructure, to a cultural proximity and to an advancement in expertise. According to GAUSEMEIER and KLOCKE, these factors are essential for a competitive advantage [11].

2.1.2 Trend

According to SCHUH ET AL., a trend describes the general direction of a development in a certain area or subject area and can thus include the coherent development of different variables in this area. Trends often arise from the interaction of different factors following a continuous process. The content focus includes social, political, economic, scientific or technological effects and the impact focus is based on global or regional effects. [10].
2.2 Related work

Technology foresight is a component of company-wide strategic early detection (Business Intelligence) [12], [13]. The aim of this early detection is to provide relevant information about changes in the entire environment of the company in order to identify potential opportunities and risks at an early stage. While early detection is focused on any future development and event in the corporate environment, technology foresight as part of these activities focuses on the analysis and prediction of the technological potential of new technologies and the determination of technological performance limits of existing technologies [12]. The objective is to identify developments in relevant fields of technology as a basis for technology decisions within the company [14].

During the technology foresight, the evaluation task is of huge importance. The main objective of the evaluation is to assess the significance of technologies or technological developments, by comparing them to each other or in relation to the company. Depending on the given evaluation background, different criteria can be derived [15]. To be able to focus on the most important criteria, first all conceivable criteria need to be taken into account. After considering these criteria it is possible to specify on the essential criteria [16].

In the context of technology foresight, the aim of a trend analysis is to derive the corresponding technological effects from the predicted development. For this purpose, an observation object is selected and the underlying trend is examined. The application of the trend analysis is especially useful in the phases of determining information needs and evaluating information [15]. When evaluating the trend, the effect of the development on the company and the probability of occurrence of the forecast are taken into account [17].

To be able to establish a successful trend analysis and to support the evaluation process, different methods can be used. The most common methods in the context of this paper are the scenario technique, the lead user analysis, the TRIZ-like methodology and the trend extrapolation. Based on the scenario technique potential future scenarios are examined and evaluated with regard to the business environment [16]. To evaluate future technologies from the key customer perspective, the lead user analysis can be used [18]. The TRIZ-like methodology helps focusing on a general problem and enhances the solution space in order to identify as many technological solutions as possible. Finally, these solutions are transferred to a specific problem [19]. The trend extrapolation helps to transform a predicted trend into a technological application [15].

3. Methodology and results of the case study

Based on the existing scientific methods in the context of technology foresight a systematic process for generically identifying potential application fields of technologies was derived. Within a case study with EPIC CoE this process was developed and implemented. The methodology as well as the achieved results are addressed in this section.

3.1 Methodology of the case study

In the first stage of the case study, a systematic screening of technological trends was conducted in order to identify and evaluate promising technological application fields for production planning and control in the context of industry 4.0. The general procedure of the trend analysis and the results of the implementation in the case study are shown in Figure 1. The main process of identifying potential application fields of technologies is preceded by a trend analysis in which mega trends are broken down into sub trends in order to analyze their influences on different observation areas, e.g. industrial, economical, technological or political observation areas. For identifying industry 4.0 based technological application fields the mega trend of connectivity serves as the starting point in the conducted trend analysis. Connectivity can be further specified in its various sub trends e.g. Internet of Things (IoT), Big Data or Smart Devices.
Analyzing the influences of a specific sub trend on the different observation areas thus enables the derivation of relevant fields of application, which can be further detailed into specific applications. In the context of the case study, the influences of the sub trend IoT were considered in an industrial context focusing on industry 4.0. In the course of this, different application clusters such as production planning and control, digital factory, production networks or machine learning could be identified and detailed in specific application fields. This paper focuses on the application cluster production planning and control in the context of industry 4.0 and is detailed in six specific application fields: production performance monitoring, advanced planning and scheduling, self-optimizing production planning systems, self-optimizing production and plant operation, decentralized production control and remote control. These six application fields were identified within a joint workshop of the members of the EPIC CoE.

In order to identify the relevance of industrial application fields, their impacts on industry and their probability of occurrence must be evaluated. In the case study, such an evaluation is carried out using a trend portfolio, which is visualized in Figure 2. While the standard trend portfolio consists of the two axis impact on company and probability of occurrence, an adapted version was derived according to the requirements of the project’s scope. Instead of the specific classification (impact on company), a more general classification (industrial impact) is used. The classification according to the industrial impact is based on the target industries of EPIC CoE. The probability of occurrence is determined on the basis of several factors, such as the technology readiness level of the underlying technologies. A period of 3 - 5 years is taken into account for the evaluation. Using the trend portfolio, the application fields can be classified into different clusters to derive actions according to future strategic activities. For example, an application field classified in cluster I indicates a particularly high potential due to its significant industrial impact and its high probability of occurrence. Thus, the application field should be immediately integrated in the strategic decisions.

Figure 1: General procedure of trend analysis according to SEITER and OCHS [20]1

1 The contents were created by experts in several workshops by transferring the general approach to the case study.
3.2 Results of the case study

This chapter examines the partial results of the implementation of the first stage of the case study. First, the application cluster production planning and control is described. Secondly, the specific application fields are detailed by describing them, addressing their advantages and disadvantages and evaluating their effects with regard to their industrial impact and their probability of occurrence. For assessing industrial impact and probability of occurrence, only the criteria with the greatest influence on the respective application field are briefly explained below. The following results were developed in workshops and are based on the expert knowledge of all institutes involved in EPIC CoE (Institute for Computer Science and Control, Hungarian Academy of Sciences (SZTAKI), Fraunhofer Institute for Manufacturing Engineering and Automation (IPA), Fraunhofer Institute for Production Technology (IPT), Fraunhofer Institute for Production Systems and Design Technology (IPK), Fraunhofer Austria Research GmbH (FhA) and Budapest University of Technology and Economics).

3.2.1 Industry 4.0 based applications for production planning and control

Production planning and control deals with the operative, temporal and quantitative planning, monitoring and control of production systems as well as the administration of all processes necessary for the production of goods and merchandise. The overall objective is the economic design and the smooth running of the production processes, which today takes place not only in particular companies but also in networks of several companies and organizations. The field of production planning and control therefore focuses on the development and implementation of process-oriented management systems and paperless production monitoring systems, in order to maximize the productivity of manufacturing processes. This can be done by providing appropriate models capable of solving complex planning and scheduling problems and suitable IT solutions for automated or semi-automated production planning and control.

3.2.1.1 Production performance monitoring

In the digital factory, large amount of data is collected to monitor production performance. With the increasing amount of data it becomes essential to prepare and visualize the data so that employees can quickly access and understand the relevant information. Statistical algorithms and machine learning can be utilized to derive additional information and to conduct root cause analyses. There are numerous software
tools available, which provide information, such as the overall equipment efficiency and profitability of assets in real-time as well as the assembly station performance and the identification of bottlenecks. The most advantageous aspect of this sub trend is the increase in transparency in production and the associated support in decision-making. However, the effects on production performance are implicit, and thus, its industrial impact is relatively low, even if transparency in production is relevant for most companies. The probability of occurrence can be seen as high, because several software solutions for advanced monitoring of production performance are commercially available.

3.2.1.2 Advanced planning and scheduling
Advanced planning and scheduling (APS) includes software tools to optimize supply chain and production planning and scheduling in real-time. Those software tools require predictive algorithms to forecast the demand, as well as optimization algorithms for production schedules, which consider product and capacity availability. APS tools are typically linked to Enterprise-Resource-Planning (ERP) or Supply-Chain-Management (SCM) systems to retrieve the required information locally and to return the optimized schedules. APS tools usually encompass supply chain planning as well as factory planning and scheduling. Supply chain planning includes demand and sales forecasting as well as inventory and transportation planning in contrast to factory planning and scheduling, which considers e.g. lead times and delivery times. The advantages of APS range from optimized production schedules over improved delivery reliability and optimized inventory level to real-time reaction to unexpected events. In contrast, the associated complex and costly implementation is a major disadvantage. The probability of occurrence of APS is high, due to its creation in the 1990's and the continuous improvement being made to the prediction and optimization algorithms. In terms of industrial impact, APS software tools are relevant for all manufacturers with complex, make-to-order production and products, which are composed of a large number of parts. Thus, the industrial impact is also high.

3.2.1.3 Self-optimizing production planning systems
In large and dynamic production environments, production scheduling problems become highly complex and finding global optimums becomes a time consuming challenge. Machine learning algorithms can be used to create virtual production models capable of autonomously optimizing the production plan. Besides the interface to the production planning system, some sort of feedback is required to monitor the actual production and to detect deviations from the planned production schedule. Potentials for self-optimization production planning systems can exist at various levels of the value chain, e.g. in supply chain design, production management or assembly management. Example applications include ramp-up decision support, whereby potential errors, bottlenecks etc. are predicted and avoided. Self-optimizing production planning systems have the potential to lead to optimized production schedules, greater flexibility and faster response to unexpected events. Apart from these advantages, there are some disadvantages, such as long learning or training period in case of complex systems. In addition, machine learning models are black box models and, therefore, difficult to predict in unknown situations. The probability of occurrence can be assessed based on the degree of the incorporation of the required machine learning algorithms into production planning systems. Machine learning algorithms are only recently incorporated into production planning systems, while holistic approaches are still limited to research projects. Thus, the value is still to be considered low. Additionally, self-optimizing production planning systems can have a potentially high impact for manufacturers with a complex and highly dynamic production. Dedicated from this, the industrial impact can be seen as relatively high.

3.2.1.4 Self-optimizing production and plant operation
Production lines are often not running at optimal conditions. Self-optimizing control systems constantly monitor process parameters, intermediate and product quality and important Key Performance Indicators
(KPIs). They are able to detect, if the process is not running optimal and readjust the process autonomously. Those control systems require accurate models of the whole process to be able to predict the process behavior. For processes, which are too difficult to describe with analytical models, machine learning algorithms provide a promising alternative. For instance, waste water plants monitoring the incoming waste water contents can adjust the process parameters accordingly in real time.

The greatest advantages of such self-optimizing production and plant operations are automated process control, increased process efficiency and improved process reaction to raw material changes. On the contrary, high costs and high expenditure of time due to a complex implementation needs to be considered. The core technologies for self-optimizing production line are available. However, implementing such systems for large production processes is very complex. Thus, the probability of occurrence is evaluated as medium. In terms of the industrial impact, the fact that even small increases in production performance can provide significant return on investment for large process plants is encountered by high costs for implementation of advanced process control systems, often impeding the use in process plants. However, the industrial impact of self-optimizing production and plant operation is evaluated as relatively high.

3.2.1.5 Decentralized production control

In industry 4.0, machines, assets and work pieces become cyber-physical systems (CPS). Those CPS are equipped with sensors and logical controllers and are capable to communicate directly with each other (machine-to-machine communication) in real-time via standardized communication protocols. This enables decentral process control, whereby the production planning system is merely required to control the overall production goals. For instance, work pieces can receive their production order from the ERP and autonomously find a path through the production.

Beneficial about decentralized production control are the rapid reaction to unexpected events as well as the reduced complexity of the production control, as only a local optimum remains to be striven for. However, in contrast to centralized controls, the decentralized control is currently less efficient. Although the described examples have already been implemented at production sites in pilot scale, completely decentralized production control systems still remain a concept. Therefore the probability of occurrence of decentralized production control is currently rated slightly more than medium. In terms of industrial impact, partially decentralized control can help to manage production environments of high complexity, though it is unlikely to completely replace centralized production control.

3.2.1.6 Remote control

With increasing connectivity throughout factories, machines, assets and systems become remotely controllable. Production planning systems can be accessed and controlled anywhere via mobile devices. Management can rapidly react to changes or adapt production schedules without the need of being on site. Machine operators do not need to be at the machine to monitor the current status and can receive notifications in case of unexpected events. In addition, entire plants, such as offshore oil platforms, can be controlled from central operation control centers, which reduces stress and risk for the human workers.

Remote control has the potential to realize faster responses to changes or unexpected events as well as increased transparency. Nevertheless, with increasing connectivity, network security is also becoming more important, since an internet connection is required. This online access can result in a vulnerability. In addition, physical task still require technicians on site.

Remote control for IT systems and large production sites are common in the production industry. Furthermore, machines, which can be controlled via a mobile device, are a more recent development. Thus, the probability of occurrence is evaluated as high. The medium industrial impact of remote control is based to its limited relevance for large multi-site companies, especially in the process industry. But, remote control for individual machines via mobile devices can be useful for machine operators in larger production sites.
Summarizing, the six identified application fields for production planning and control in the context of industry 4.0 can be classified in the trend portfolio as shown in Figure 3, in order to identify relevant and promising application fields for future development and strategic focus of EPIC CoE.

As a result, advanced planning and scheduling as well as self-optimizing production planning systems have a particularly high potential for EPIC CoE, since both application fields fall into the core competencies of the centre and have a high industrial impact for the targeted industries.

4. Summary and Outlook

In this paper, a case study for identifying potential application fields for production planning and control in the context of industry 4.0 was presented. Therefore, the authors derived a systematic process from existing scientific approaches on trend scouting. Within the case study, the process was implemented and validated by identifying and evaluating promising industry 4.0 based application fields for production planning and control. In the first stage, a total of six application fields in the cluster of production planning and control were identified: production performance monitoring, advanced planning and scheduling, self-optimizing production planning systems, self-optimizing production and plant operation, decentralized production control and remote control. The application fields were described in detail according to their application focus as well as advantages and disadvantages. Furthermore the potential industrial impact and the probability of occurrence of the application fields were outlined. Finally the identified application fields were classified using the trend portfolio to be considered within future strategic decisions of the EPIC CoE.

Implementing the trend portfolio for EPIC CoE has proven to fundamentally helpful in deriving relevant application fields for the future strategic alignment of the centre. Thus, the classification can serve as a guideline for aligning further activities of EPIC CoE as well as future development focuses. However, the results represent only a general orientation which requires subsequent detailing in the form of an implementation strategy for the centre's activities. The added value is achieved in particular in the expert
discussions, the methodology’s systematic approach for the identification of the different application fields as well as in the estimation of impact and probability.

To further validate the results of the first stage, a critical reflection on the positioning of the application fields in the trend portfolio will be accomplished through expert interviews. Subsequently, in the second stage, technologies for the implementation of the most promising application fields are to be identified. Hereby, the technologies will be investigated and evaluated according to their readiness level for the identified application field.

Acknowledgements
The work for this paper was supported by the European Union’s Horizon 2020 research and innovation programme under grant No. 739592.

References


Biography

Prof. Dr.-Ing. Dipl.-Wirt. Günther Schuh is the head of the chair of production systems at RWTH Aachen University and a member of the directorate of the Laboratory for Machine Tools and Production Engineering (WZL) at RWTH Aachen University and of the Fraunhofer Institute for Production Technology IPT in Aachen. In addition, he is Director of the Institute for Industrial Management (FIR) at RWTH Aachen University. Professor Schuh is also a member of several supervisory boards and directorates and was prorector for industry and commerce at RWTH Aachen University from 2008 to 2012. His most important research findings include relevant methods and instruments for complexity management, resource-oriented process cost accounting and participative change management, as well as the concept of the virtual factory.

Patrick Scholz, M.Sc. RWTH studied mechanical engineering at the RWTH Aachen University with a specialization in production engineering. Since 2017, Mr. Scholz has been a research assistant in the Technology Management Department and since 2019 also Manager of the Business Unit Lightweight Production Technology at the Fraunhofer Institute for Production Technology IPT in Aachen. As part of his work at the Fraunhofer IPT, he has already carried out various consulting projects in technology and innovation management. His research is concerned with the selection and evaluation of technologies in the early phase of the innovation process. The focus is on the efficient design of decision processes on the basis of individual potentials and risks.
Development of a Model for describing, evaluating and designing Communication Concepts in Factories in the Context of Industry 4.0

Yeong-Bae Park¹, Lennart Hingst¹, Peter Nyhuis¹, Michael Kelber², Jörg Rainer Noennig²

¹Institute of Production Systems and Logistics/ Leibniz University Hannover
²Wissensarchitektur Laboratory of Knowledge Architecture/ TU Dresden

Abstract

Far-reaching changes in the technical and organisational structure have influenced the communication between the elements of a factory. In particular, Industry 4.0 as a contemporary trend in production technology and many related areas of science places new requirements for communication. Thus, expectations comprise a wide variety of effects and potentials of Industry 4.0 on communication in factory systems. Exemplarily, efficiency and productivity increases are most frequently named as effects of industry 4.0. They seem advantageous by industrial enterprises, whereby the raising of these potentials is also a necessary competitive factor. However, such influences can trigger other effects with interdependent impacts on the components of factory systems. These effects can have both positive and negative impacts in factory systems. Still a comprehensive and generally valid description of these aspects is not yet available. Similarly, there is no understanding of the cause-effect relationships between the requirements for communication between the system components of a factory arising from Industry 4.0. The lack of specific understanding of Industry 4.0 caused effects and the cause-effect relationships between the requirements for communication between the factory system components lead to the inability to design effective communication concepts in factory systems. Therefore, existing communication systems remain exposed to undesired effects and leave desired effects underutilized. In order to close this research gap, a holistic model for the description, evaluation and the design of effective communication concepts in factories in the context of Industry 4.0 is in development within the frame of "Komm 4.0", a research project of the Institute for Production Systems and Logistics at Leibniz University Hannover and the WISSENSARCHITEKTUR Laboratory of Knowledge Architecture at TU Dresden. The pursuit of this research project is to describe and evaluate existing communication concepts, as well as design more effective concepts and to adapt previous recommendations in terms of communication concept design where necessary.

Keywords

Factory planning; Communication concepts; Industry 4.0; Description model; Effect model

1. Introduction

The factory as a socio-technical system and its various system components are linked by communication [1]. A factory defines a specific location where value is created by the manufacture of industrial goods through the utilization of division labour and production factors [2]. Profound developments in the field of technology and organisational structure influence the communicative and informational connections of the factory components (technology, organisation, employees, location and buildings) in particular [3] [4]. With
the arrival of such developments such as industry 4.0, communication in factory systems has become more import to the operative tasks of the industry [5]. This contemporary trend particularly places new demands on communication in production technology and in related areas of science [6] [4] [7]. The establishment of these new demands depicts an influence of Industry 4.0 on the communication and therefore on the communication and information processes in production technology. Despite the importance of communication, it often receives insufficient attention [8]. These Industry 4.0 induced communication and information processes are here defined as "Communication 4.0" and are the result of the respective communication concept. These resulting processes aim to achieve individual corporate goals. Based on numerous definitions of the communication concept in general in the operational environment, it describes a process that transmits information from a sender to one or more receivers [9] [10]. In addition to this fundamental understanding of communication, further properties are of relevance [11] [12]. These properties include the communication type [13], communication means [14], communication structure [15], communication quality [16], communication environment [14], communication direction [17], communication intensity [14] and communication distance [14], which are brought together under the idea of the communication concept.

As factory systems are evolving over time communication concepts that have grown organically and involuntarily in the rapid course of Industry 4.0 can lead to unwanted inefficiencies and loss of competitive advantages. In order to use the effect of communication concepts on the efficiency of an enterprise, more specifically on a factory, a comprehensive understanding of Industry 4.0 caused influencing factors on the configuration of the communication concept is necessary. Well-founded knowledge of the cause-and-effect relationships can enable a structured design of communication concepts. The application of this knowledge at an early stage can prevent potential loss of performance due to communication inefficiencies and support to realize performance potentials comprehensively.

2. Need for Research

The preliminary work on the topic of communication concept design in factories resulted in a uniform recommendation to increase the degree of organisational, procedural and spatial networking. The aim is to increase the frequency and quality of communication and consequently to increase the efficiency of the factory [18] [19]. Nevertheless, the increase in communication frequency and communication quality as the maxim of an increase in performance must be viewed in a more differentiated way. The application of findings of the attention economy to the context of communication concept design in factories, leads to deviating conclusions. The core idea of the attention economy understands human attention as the scarcest
commodity [20] [21]. Consequently, this attention should only be used for those actions that represent maximum marginal utility [22] [23]. With reference to previous recommendations, the increase in the degree of networking does not show a constant marginal utility. If, for example, the communication time or the number of communication interfaces extent, this can lead to decreasing efficiency and effectiveness. Based on these findings, increasing the degree of networking in factories in the course of Industry 4.0 does not appear to be a direct maxim. Because of the contradictory findings from different scientific fields, this research project will, among other things, review and, if necessary, adapt the previous recommendations for increasing the degree of networking in the context of Industry 4.0.

Furthermore, previous research projects at the Institute for Factory Systems and Logistics have examined the importance of communication in factories and have highlighted it as an essential competitive factor [19] [24]. These projects emphasize the targeted design of the communication concepts by factory planners and architects as an essential interdisciplinary component. Moreover, previously carried out work only forms descriptive models under constant conditions or evaluation models for the special situation of the change of the type of order processing. Furthermore, various studies predict versatile effects and potentials that emanate originate from Industry 4.0 on communication in the factory system [6]. These effects and potentials can drive changes. A comprehensive and generally valid description of these change drivers is not yet available. This also applies to a detailed understanding of the cause-effect relationships between the requirements for communication between the systems components of the factory arising from Industry 4.0. Therefore, no investigation of the possibilities of the targeted reinforcement of desired changes and the mitigation of negative effects of Industry 4.0 by generic measures is present.

Since neither a holistic model nor a method of procedure exists to describe, evaluate and design communication concepts in factories under the influence of Industry 4.0 as of now, the aim of this research project "Komm 4.0" is to close this research gap.

3. Approach to the Development of a Model and a Method for describing, evaluating and designing Communication Concepts in Factories in the Context of Industry 4.0

3.1 Approach

In order to close the research gap, this research project aims to develop a model and a method for describing, evaluating and designing communication concepts in factories in the context of Industry 4.0. The model should make it possible to design effective communication concepts and thus to support the maintenance and increase of the innovation ability in factories as a competitive advantage. The following approach aims to reach the formulated objectives within the frame of this research project.

The first step is the development of a description model of communication in the context of Industry 4.0. This model should cover all relevant changes as well as the corresponding drivers that Industry 4.0 has on the communication and information processes between the systems components of a factory (technology, organization, employees, location and building). Based on the knowledge gained, the next step is the development of future requirements for communication in factories. Finally, the objective is to operationalize these requirements in order to integrate them into the existing communication concept as new components. Subsequently, the next step is to identify and describe communication barriers at the interfaces of the systems components. The phase ends with the definition of measures for the removal of communication barriers. The consolidation of the findings above will form a description model.

In a second phase of the approach, it is to verify and supplement the findings from the previous phase, which are consolidated within the description model. A further aspect of this phase is the identification of correlations between the components of the description model. A case study will verify the findings in a two-step approach. The first step comprises the comparison of the contents of the description model with
information from industrial companies. The comparison includes the gathering of information on the elements of the description model from the respective companies. For example, Industry 4.0 technologies, which are already in use by enterprises both long-term and short-term, are examined regarding their impact as change drivers. Future requirements for the communication and information processes as well as existing communication barriers and the converted measures will be part of the documentation in the same way, as described for the verification of the elements of the description model. The comparison serves as a basis for formulating initial hypotheses on the relationships between the elements. In the second step of the case study, the formulated hypotheses will undergo a check with selected experts.

In the next phase after examining the formulated hypotheses, the development of a model for the evaluation and design of the communication concept of factories in the context of Industry 4.0 is done. This overall model is composed of the three components. An effect model, a process method based on it and an application tool. The effect model enables the visualisation of connections between change drivers, operationalized components of the communication concept and generic measures. The method describes the necessary steps for the evaluation and organization of the communication concept in detail. Based on this, the application tool enables a systematic evaluation of the existing communication concept as well as the selection of generic measures for the design of the communication concept.

The fourth phase comprises a final evaluation of the consolidated model for the evaluation and design of the communication concept. The practical tests in partner companies as well as adjustments based on these studies will be part of this final phase if necessary. The documentation and preparation during and after the processing of the preceding phases is an essential part of this project.

3.2 Current Findings

With regard to the first work package and thus the development of the description model, first results have already been achieved. Based on an extensive research of scientific literature as well as studies and surveys, relevant trends and technologies in the context of Industry 4.0 have been identified [25] [26] [27]. The positioning of the identified aspects in the factory system is based on the dimensions "hierarchy levels of a factory" and "main business processes of a factory" (Figure 2). Accordingly, it is analysed in which main business processes of a factory and at which hierarchy level the technologies are relevant.

Figure 2: Positioning of the change drivers in the factory system
However, since the focus of the research project presented here is on communication in factories, the identified and located change drivers (Figure 2) are compared with the communication concept (Figure 1) described above. The influence of the change drivers of Industry 4.0 on the components of a communication concept have been analysed. The result of these analyses are hypotheses about the influence of the change drivers on the components of the communication concept. In the course of processing this step, it became clear that the components communication means, communication distance and communication environment are not to be taken into account for further action in the context of Industry 4.0. In the case of the communication distance and the communication environment, the reason for this is that Industry 4.0 has negligible or no influence. In the case of the means of communication, the influence is determined by the dependency on the component Communication Type. Further influences have not been identified.

At the same time, the formulated hypotheses represent the communication requirements resulting from the consideration of the already known communication concept within the framework of Industry 4.0. In a further step, the influences of Industry 4.0 have been further investigated by identifying new communication requirements associated with Industry 4.0. For this purpose, communication processes have been analysed and their future characterization under the influence of Industry 4.0 was examined. By considering the future communication processes between the system components of a factory, new communication requirements have been identified. In order to integrate these new aspects into the already known communication concept, an operationalization of the newly identified communication requirements has been carried out. The operationalization has been realized by the definition of different degrees of fulfilment of the new requirements. The integration of the new components into the existing communication concept leads to the communication concept in the context of Industry 4.0 (Figure 3).

The communication processes between the system components of a factory have been again used to identify action measures to reduce or even remove communication barriers. Particular attention has been paid to the
communication interfaces between the system components. In a first step, communication barriers have been worked out by means of a comprehensive research. Since each interface is related to a specific future communication requirement, the obstacles at an interface can provide information about which measures are necessary to be able to meet the respective communication requirement. The identification of the potential measures to remove the identified communication barriers at the respective interface has been carried out by means of a second literature search. All measures then have been assigned to the Industry 4.0-side communication requirements according to their interface and sorted according to their department-specific or interdepartmental character. The configuration thus put together describes the catalogue of measures. Described in the following for the example of the technical-technical interface. Previous research work has been able to identify the communication barrier "media breaks due to missing standard interfaces and heterogeneous IT landscapes" at this interface. [28] [29] [30] [31]. Within the framework of Industry 4.0, this obstacle stands in complete contrast to the required interface compatibility, which describes a continuous and media-break-free exchange of data and information between the IT systems used. In order to meet this future communication requirement, it is therefore necessary to eliminate the media breaks between the technical systems within a factory. A potential department-specific action at this point is the introduction of new, compatible software in the department concerned. If media breaks occur between several departments, the definition of standards with regard to uniform software must take place as a cross-departmental measure. Further measures can be taken, which can be found in the catalogue of measures.

4. Conclusion

Overall, the catalogue of measures developed forms the basis for the subsequent design of the communication concept within the framework of Industry 4.0. For the design of the communication concept in the context of Industry 4.0, however, it is necessary that the current findings are validated and adapted depending on the results of this validation. This step will be included in the next work package. To this end, case studies will be developed and discussed with industry partners to validate and supplement the description model and to identify correlations between the elements of the description model.

Acknowledgements

The Deutsche Fördergemeinschaft (DFG) - project number 405932399 - supports the research project described above.

References


Biography

Yeong-Bae Park, M.Sc. (*1994) has been a research associate in the specialist factory planning group at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover since 2019. He previously studied mechanical engineering at the TU Ilmenau (B.Sc.) and at the Karlsruhe Institute of Technology (M.Sc.).

Lennart Hingst, M.Sc. (*1989) has been a research associate in the specialist factory planning group at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover since 2019. He previously studied technical logistics at the University of Duisburg-Essen (M.Sc.) and holds years of industry expertise.

Prof. Dr.-Ing. habil. Peter Nyhuis (*1957) has been head of the Institute of Production Systems and Logistics (IFA) at Leibniz University Hannover since 2003. Prof. Dr.-Ing. habil. Peter Nyhuis was a member of the German Science Council (2014 – 2018) and has been Chairman of the Science Council of the AiF (GAG3) since 2015. He is also a member of the German Academy of Technical Sciences acatech and is a member of the Scientific Society for Production Technology (WGP).

Dipl.-Ing. Michael Kelber (*1992) has been a research associate in the special field of working place of the future at Wissensarchitektur - Laboratory of Knowledge Architecture at Technische Universität Dresden since 2017, where he previously studied architecture (Dipl.-Ing) at the Technische Universität Dresden.

Prof. Dr.-Ing. Jörg Rainer Noennig (*1973) is the director of the Wissensarchitektur – Laboratory of Knowledge Architecture at TU Dresden, and Professor for Digital City Science at the CityScienceLab at HafenCity University of Hamburg.
Bio-inspired Factories of the Future

Arvid Hellmich¹, Brandon Sai², Marian Süße³, Martin Schreiber⁴, Torben Wiese¹, Steffen Ihlenfeldt¹, Thomas Bauernhansl², Matthias Putz³, Gunther Reinhart⁴

¹Fraunhofer Institute for Machine Tools and Forming Technology IWU, Dresden Germany
²Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Stuttgart Germany
³Fraunhofer Institute for Machine Tools and Forming Technology IWU, Chemnitz Germany
⁴Fraunhofer Research Institution for Casting, Composite and Processing Technology IGCV, Augsburg Germany

Abstract

The biological transformation of added value is seen as one of the key aspects in applied research. Bio-inspired methods and technologies will affect factories of the future and enable them to cope with changing boundary conditions and the rising necessity of sustainability. This results in a higher demand for flexibility and transformation ability of the comprised production systems. To elaborate topics like these, Fraunhofer initiated strategic collaborative research projects. The current project aims at developing aspects of the biological transformation, whereof organic bio-inspired factories is one. Different research focal points were identified as enabling technologies on different levels of the well-established automation pyramid. The paper highlights the aspects “facility layout planning”, “behavioral modeling of production systems” and “skill-based controller programming” as enabling technologies. Solution approaches for the addressed aspects are discussed and future steps towards a flexible and sustainable production are shown.

Keywords

Biological Transformation; Production; Factory Planning; Behavioral Models; Controller Programming

1. Introduction

Over the last years, research activities in manufacturing have been mainly focused on the digital transformation, an evolution, driven by the motivation to achieve higher efficiency and interconnectivity – not an agenda for more sustainability. This is where biological transformation comes in – a process that describes the increasing technological utilization of materials and principles from nature to establish a sustainable economy. [1] With the goals of the digital transformation the requirements on how a system has to operate are defined. But only with the biological transformation the activities’ purpose is clearly named. Hence, digital transformation and biological transformation are complementary strategies. Each is necessary, but neither is sufficient in itself. They must, therefore, be reasoned out and pursued in combination. [2,3]

In fact, biological transformation can help to achieve the sustainable development goals set out by the United Nations (SDGs), which serve as a reliable benchmark against which to measure sustainability. In terms of manufacturing the biological transformation addresses the SDG No. 9: “Industry, Innovation and Infrastructure” and SDG No. 12: “Responsible Consumption and Production”. [3]
For SDG No. 9 the adaptation of biological principles can play a key role in making technological systems and infrastructure more resilient. In addition, the use of natural resources is a prerequisite for releasing global industry from its dependence on materials based on fossil fuels. SDG No. 12 addresses a responsible consumption and production. The replacement of fossil fuels by renewable biological materials, and the establishment of closed-loop material cycles, will make a significant contribution towards reducing industry’s carbon footprint. Using biological processes and biomimetics can help make manufacturing processes more efficient and, thus, reduce the consumption of raw materials. Similarly, the application of biological principles can reduce logistical requirements and a setup of industry at particularly impacted locations. [4]

In this paper, versatility and adaptability of nature is focused. Abilities which are highly required either by changing boundary conditions or by the rising necessity of sustainability.

2. Adaptability of nature in different areas of production

In order to achieve a change towards biological inspired and sustainable production, all aspects, starting from the factory layout down to the education of workers have to be transformed. Therefore, the vision of a biological bio-inspired factory is introduced as a role model (Figure 1). Here, stationary production resources such as forming machines, casting machines or rolling mills form the centered element of a factory in analogy to a cell nucleus. All other means of production (resource input, energy supply, cutting and assembly cells and shipping stations) are flexible in quantity and situated radially around the nucleus. Here, the analogy of a biological cell becomes visible, since they are also flexible e.g. in mitochondria based on the purpose of the biological cell (muscle cell, skin cell, …).

Biological bio-inspired factories comprise different kinds of manufacturing units. Besides stationary machines, a large number of highly flexible units such as robot cells or adaptable storing devices enhance the process chains. The stationary units can remain more or less unchanged, when a factory layout is transformed e.g. to a more bio-inspired variant. However, the position and also the manufacturing or storage task of flexible cells can vary in wide ranges. This requires

- completely new design solutions,
- new ideas of power, air and information supply,
- new IT solutions in interlinkage and information distribution,
- new optimization paradigms in Facility layout planning as well as production optimization,
- new approaches of supervision and condition monitoring as well as
- a new paradigm of PLC programming.

In the following paragraphs, three aspects are selected and specific approaches towards the introduced biological bio-inspired factory layout are introduced: Facility factory planning, Behavioral modeling and PLC (programmable logic controller) programming. While in the segments of factory planning and
production optimization adaptability approaches are proposed, the skill-based PLC programming represents one of the necessary enabling technologies.

2.1 Facility layout planning

The biological transformation incorporates several characteristics or tools of nature which are applied in technology. Not only is the contribution of each of these tools measured economically, but also with regard to its ecological and social impact. Hence, the development, clearance, use and demolition of a production plant is characterized by growing complexity, since a contribution to the above mentioned SDGs becomes a necessity. During this process, layout planning must provide the basis for sustainable factories and existing processes. Algorithms and software solutions require adjustments.

Planning facility structures or layouts is part of distinct phases in the standardized and well-defined factory planning procedure, described in the guideline VDI 5200 [5] of the German Association of Engineers. Therefore, developing factory layouts can be seen as a part of the concept planning phase and detailed planning phase of factories. Several authors like Müller [6], Pawellek [7] or Wiendahl [8] describe more specific procedures and detailed sequential steps for layout elaboration.

The general scheme may be separated into the following phases; first, structuring and dimensioning is conducted based on the required machines and production plants in order to achieve the previously defined system performance. Hence, the necessary area for all production elements is calculated. Second, several versions of rough layouts are developed, evaluated and discussed. Third, with the stepwise implementation of restrictions more detailed designs are created and finally one fine layout for the implementation is generated. The following Figure 2 illustrates the phases of the planning process and specifies the referring steps for layout planning.

![Figure 2: Delineation of factory planning and layout planning processes (based in VDI 5200)](image)

During ideal planning and real planning, optimization of the arrangement plays a major role to find an optimal or at least good solution regarding distinct target criteria (e.g. material flow or transport costs). This accounts for both greenfield as well as brownfield planning cases. In mathematical optimization and operations research different formulations of the facility layout planning problem occurred during the last decades and diverse scientific research on algorithms and solution methods has been conducted. Depending on the circumstances and boundaries, further specification into different subcategories (e.g. dynamic facility layout problems or multi-floor facility layout problems) exists. A structured overview with according solution methods is given by Anjos and Vieira [9], Drira et al. [10] and Ahmadi [11]. In complexity theory the non-deterministic polynomial-time hardness (NP-hardness) characterizes a class of problems as possibly intractable, i.e. their solution requires an exponentially growing amount of time with increasing problem size and it is still a part of research if these problems can be solved in polynomial time [12]. Because of their complexity, facility layout problems are often characterized as NP-hard [9]. Therefore sophisticated heuristics are applied to generate improved solutions in an appropriate amount of time. As an example, Tuzkaya et al. [13] compare meta-heuristic approaches for generation of facility layouts.
In industrial planning projects heuristic schemes like the triangle method of Schmigalla or circle method of Schwerdtfeger serve as well-established approaches for the creation of rough layouts or block layouts [7,6]. On the one hand, they generate quick arrangements for an ideal layout based on simple rules which provide transparent results. On the other hand, the development of innovative layout structures is limited and the applicability gets reduced when restrictions are increased. In addition to that, current and ongoing improvements in information technology are not exhausted with these heuristics.

However, a large gap between scientifically investigated approaches and applied layout optimization heuristics can be determined. As shown by Lohmer et al. [14] software tools for a factory layout design only comprise greedy heuristics, the actual potential of optimization approaches in realistic environments is still unexploited. Furthermore, layout development and design is mostly tied to rectangular concepts and building structures which leads to efficiency potentials of innovative and nature-inspired concepts.

Different concrete approaches for the adoption of bio-inspired designs for factory layouts were developed by Tinello and Winkler [15,16]. They investigated their applicability for diverse use cases and production environments. Some of their findings are compared in the following table:

<table>
<thead>
<tr>
<th>Honeycomb</th>
<th>Nautilus</th>
<th>Spiderweb</th>
</tr>
</thead>
<tbody>
<tr>
<td>– Minimal circumference for same space compared to triangles or rectangles</td>
<td>– Reduces required space for transport</td>
<td>– Reflows besides main production flow can be handled easily</td>
</tr>
<tr>
<td>– Basis for efficient material flow with centrally focused section</td>
<td>– Material flows are reduced compared to established heuristics in factory planning</td>
<td>– Basic structure can be easily expanded</td>
</tr>
</tbody>
</table>

Besides these advantages it should be pointed out that these concepts highly contribute to an improved changeability of factories since nature-based structures are intrinsically built for an adoption towards changing circumstances and comprise efficient ways for necessary growth processes. Referring to this, Tinello and Winkler show that some of the procedures for biomimetic layout development are more complex compared to the already mentioned and well-known heuristics. [15]. It becomes obvious that the presented, nature-oriented shapes are still limited to the stage of an ideal layout which means that investigation regarding their efficiency and applicability has only been conducted for an earlier planning stage with less restrictions. Thus, several further steps for examination become apparent.

With reference to the aforementioned bio-inspired factory, several analogies of the structure and mechanisms in eukaryotic cells could be derived in order to generate creative as well as efficient planning results. As an example the cytoskeleton of organic cells serves as a solid and robust structure and likewise provides enough flexibility for cells to adapt their internal partitioning. Hence, specific ways to model, assess, compare and finally implement biomimetic factory layouts on such principles are necessary. As already indicated with the biologically inspired layouts in Table 1, this necessitates further development of heuristics and analytic optimization methods that are based on additional shapes besides rectangular grid structures.

Resuming investigation of bio-inspired layouts is conceivable in advanced planning phases (e.g. real planning) taking further restrictions into account. This implies the adoption of different legal requirements and norms or prevailing supply infrastructure. A third point of further examination is the transfer of scientifically investigated and tested approaches into appropriate software tools. Therefore, the evolution of existing software tools as well as the development of new software solutions for planning engineers is imaginable.
The fourth, but not less relevant aspect is the investigation of ecologic criteria and their consideration in layout planning. With reference to material flows, the energy demand (e.g. of automated guided vehicles) may already be a relevant aspect. As soon as the ecologic reflection is extended towards further aspects of lifecycle assessment-based impact categories, their quantification and optimization in factory planning stages has to be considered. Hence, aspects like biodiversity, land use or transformed area could become relevant target criteria in factory layout design.

2.2 Behavioral modeling of production systems

The complexity of modern production systems is increasing. In addition, achievements in development are more frequent while the time span between them decreases. As a consequence, the role of the human in manufacturing is changing. Formerly, the focus was on adding value, mostly by handcraft. Today, in modern plants the role is to ensure a smooth running of the automated production system, i.e. the role of human has been changed towards production optimization. [17] In terms of optimization, the overall equipment effectiveness is becoming the most important benchmark with targeting the elimination of losses in the process. However, due to the continuously increasing complexity, this task is also slowly becoming more difficult. Therefore, intelligent systems which help to eliminate losses are required. Not just for identifying losses, but for identifying the root cause and its possible impacts. In analysis there are different approaches with different goals. For explanation, a promising approach is the diagnosis. In diagnosis, the main concept incorporates a reference, for systems with lower complexity, a single value can be sufficient to represent a goal. The difference between the target and actual value delivers the optimization action and the explanation inherently. For systems with higher complexity, a single value is hardly sufficient, therefore, a behavior model can be used as a reference, with the only change that the systems normal behavior is the new reference.

The manual creation of behavior models is a tedious process and represents the bottleneck in the development of approaches for intelligent diagnosis. For the modeling itself, much expertise is required and all interactions in the plant have to be known. Additionally, not all physical effects can be captured and modeled in detail. [18,19]

To achieve adaptability, i.e. to be flexible, in production optimization the modeling of the systems behavior has to be automated. In addition, the system itself has to be autonomous [1]. Thus, whenever circumstances are changing the system has to adjust. Even if the circumstances and boundaries are set, the system has to learn continuously to provide the optimization target. Hence, the algorithm for diagnosis has to have the following characteristics:

- Online algorithm: Because the modeling is based on observation, its characteristics defines the main requirements of the algorithm. As an observing or reading system, the input has to be processed piece by piece in a serial fashion, i.e. in the order that the input is fed to the algorithm, without having the entire input available from the start.
- Passive algorithm: While passive learning algorithms have to cope with a given set of observations to learn the model, active learning algorithms can ask for additional observations, if needed.
- Fault data: Because the systems has to learn live during the production, fault-free data cannot be expected.

The main task in the automated modeling of a systems behavior is the extraction of a noised core sequence. Different approaches, either model-based or data-driven have been designed to fulfill this assignment.

One promising approach comes of the field of theoretical computer science and discrete mathematics. An automaton is a construct made of states designed to determine, if the input sequence should be accepted or rejected. An automaton has a finite set of states which are used to assess if the current state or transition from one state to another is receivable. With this construct a sequence, i.e. a behavior, can be defined. [18] Another model-based approach comes along with the Hidden Markov Model (HMM). A statistical Markov model in which the system being modeled is assumed to be a Markov process with unobservable (i.e. hidden) states.
In simpler Markov models, the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters, while in the hidden Markov model, the state is not directly visible, but the output, dependent on the state, is visible. [20]

Possible data-driven approaches can be either modeled with a recurrent neural network (RNN) or an autoencoder. An RNN is a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence. This allows for temporal dynamic behavior. An autoencoder is an artificial neural network that learns to copy its input to its output. It has an internal (hidden) layer that describes a code used to represent the input, and it is constituted by two main parts: an encoder that maps the input into the code, and a decoder that maps the code to a reconstruction of the original input. [21]

Performing the copying task per se would be meaningless, and this is why usually autoencoders are restricted in ways that force them to reconstruct the input only approximately, prioritizing the most relevant aspects of the data to be copied.

![Figure 3: Structure of a hidden Markov model (left); Structure of an autoencoder (right)](image)

Regarding nature, a self-optimizing system is the goal. To achieve this goal, a first step towards adaptability has to be made, which requires that a system is able to observe the actual and target value, i.e. the actual and normal behavior. Therefore, it is necessary to continue research in automated behavior modeling. The mentioned approaches deliver possible and promising solutions.

### 2.3 Skill-based controller programming

Automation of production means is currently mainly a technical development to increase production volume. As mentioned, for the introduced bio-inspired factory, state of the art construction principles, IT-solutions and automation paradigms have to be changed. Only if flexibility and transformation ability in hardware is regarded in combination with flexibility in designing the automation solution, IT and optimization, a true flexibility in manufacturing can be reached. Different solutions in this manner are already under development, currently mainly with a focus on organization. In the following paragraph, the focus is laid on automation. Here, state of the art in programmable logic controllers (PLC), which are the common solution for e.g. robot cells, is a cyclic processing comprising the steps:

- input scan (reading all inputs and memory states)
- execution of a problem-oriented automation program (PLC-program) to generate output and memory values
- output update (writing values to outputs and memory)

The PLC program incorporates all necessary information to generate output signals based on the input and memory state. Therefore, the complete automation task is specified and transformed into the PLC program, which is usually developed, implemented, tested and maintained by an automation technician of the machine/unit manufacturer [22]. This approach is focused on the specific automation problem (problem-
oriented) and can only handle changed conditions or tasks to the extent that they were already known to the programmer at the design stage and taken into account in designing the control program. Modification, adaptation or addition of command sequences, positions, process sequences is usually not possible. Conventional control programs are, therefore, structured individually and project-specific based on the automation problem (Figure 4, left). The effort required for programming, testing and commissioning control software is growing disproportionately with the increase in the scope and complexity of control functionality [23]. This mismatch is aggravated due to the fact, that machine users cannot initiate even slight adaptions of process chronology, such as reordering process steps, position changes, changes of loading aids etc.. Therefore, even minor changes to the program are costly and time-consuming because specialized commissioners are required.

Along with Industrie 4.0, initial approaches to modularize the control landscape, architecture and programming arise. One course of action is to simplify the task of programming PLC by transforming it to a parametrization action. Hence, the controller is not only equipped by a program to solve one specific complex automation task but with a large set of modular skills, covering all possible abilities of the manufacturing unit. Programming and teaching this skillset still requires an automation technician but is extremely flexible in further application and individual utilization. The transformation from the well-known but inflexible task-based programming to a skill-based parametrization requires a new point of view and a new definition of the automation procedure. At first, all possible skills of the flexible manufacturing unit such as handling, measuring, orienting, loading/unloading are regarded as jobs. All jobs consist of a sequence of skills, such as movement, opening/closing a gripper, call to a camera etc.. Hence, the skill “movement” is furthermore a combination of basic skills such a move linear or move circle (Figure 4, right). A rather complex robot path is composed from basic movement elements. An analogy to biology can be found when the DNA is regarded, where also very basic elements are combined to a highly complex information storage system.

The manufacturing cell is equipped with all hardware modules and a complete skillset. This skillset is taught and programmed by the automation technician but remains open for a variety of applications. After a basic commissioning and software test, the machine operator can combine the skills and jobs to an automated process sequence or adapt the given sequence to a new setting with support of a graphical user interface (GUI). Figure 5 shows such a GUI for a robot cell to manipulate, measure and stack up automotive parts. The cell consist of a KUKA robot and four stations, realizing different process steps. The left part shows the parametrized job list followed by a specification for each job. A visualization in the right part provides an interaction with the operator and visualizes the abstract robot actions, associated with the manufacturing layout and task. Here, the programming paradigm interacts with factory facility planning, when the state-
actual model is incorporated in the GUI. In the middle part, control buttons allow for connecting with the PLC, downloading the program to the plc and running specific jobs.

![Figure 5: GUI of the skill-based automation for a flexible robot cell](image)

The new parametrization paradigm uses flexible robot jobs, parametrized by a GUI and subsequently downloaded into the PLC. The program, job/skill parameters and the chronology of skills can be controlled, adapted and reorganized through a GUI. After basic commissioning of the complete skillset, no automation specialist is needed anymore. Job and process flow adaption and testing can be realized within a few minutes. Users without PLC or robot programming knowledge can implement changes and the risk of errors in programming is reduced.

Beside the mentioned robot cell for manipulating automotive parts, also a different robot cell for highly flexible utilization to load and unload machine tools was automated, based on the new parametrization paradigm.

3. Conclusion

Bio-inspired autonomous systems in manufacturing incorporate flexibility in hardware and software and adaptability in concepts. Both, flexibility and adaptability, were discussed in this paper. Promising approaches to achieve adaptability on different levels, e.g. facility planning and production optimization, have been pointed out and discussed. For facility planning specific ways to model, assess, compare and finally implement biomimetic factory layouts are necessary. In addition, ecologic criteria and their consideration in layout planning need further examination. Moreover, scientifically investigated and tested approaches must be transferred into appropriate software tools. For production optimization, intelligent diagnosis systems are needed. Their development is limited by the manual process of behavior modeling. Promising solutions have been discussed for both, model-based and data-driven approaches. On the lowest level of the automation pyramid a flexibility of controller planning is needed to enable adaptability in manufacturing. Therefore, a rethink of programming PLC is shown, simplifying the task of programming...
PLC by transforming it to a parametrization action. Here, the controller is not only equipped by a program to solve one specific complex automation task but with a large set of modular skills, covering all abilities of the manufacturing unit.

Acknowledgement

The authors wish to acknowledge and thank Professor Reimund Neugebauer, President of the Fraunhofer-Gesellschaft, for his initiation and all involved Fraunhofer-institutes for the possibility of carrying out this research project. They also thank Dr. Sophie Hippmann, Head of Think-Tank, Dr. Miriam Leis and Ms. Petra van Heek for their continuous support. In addition, the authors thank Dr. rer. nat. Wolfgang Karl Hofbauer of Fraunhofer IBP, Dr. Malte von Ramin of Fraunhofer EMI and Dr. Bernhard Ellinger of Fraunhofer IME for their input.

References

[14] Lohmer, J., Klausnitzer, A., Lasch, R. Advanced scientific algorithms in digital factory design applications.


Biography

Thomas Bauernhansl (*1969) has been Head of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart and the Institute of Industrial Manufacturing and Management at the University of Stuttgart since 2011 Prof. Dr.-Ing. Bauernhansl is author of many books and member of many scientific societies. Both the digital and the biological transformation of value creation are core research fields in his institutes.

Arvid Hellmich (*1982) is Head of the Department Technical Cybernetics of the Fraunhofer Institute for Machine Tools and Forming Technology IWU Dresden. After his PhD thesis in 2014 he is currently working in different projects on data acquisition and –analysis, function integration in industrial controller systems, intuitive controller programming and simulation of electromechanical systems.
Steffen Ihlenfeldt (*1971), studied at the TU Braunschweig and has been a member of the Fraunhofer Institute for Machine Tools and Forming Technology IWU in various positions since 1997 at. He received his Doctorate in 2012. Since 2015 Prof. Dr.-Ing. Steffen Ihlenfeldt has been called to the Position of Professorship for Machine Tool Development and Adaptive Control systems at the TU Dresden. In addition to his work at the TU Dresden he is currently Head of the Department for Cyber-Physical Production Systems at Fraunhofer IWU.

Matthias Putz (*1957) studied Mechanical Engineering at the St. Petersburg State Polytechnic University. He gained his Doctor’s degree in 1986 in the Field of Machine Tool Construction. Professor Putz joined Fraunhofer IWU in 1994. Since 2000 he has had leading positions at the Fraunhofer IWU. In the year 2007 he obtained the Honorary Professorship at the Dresden University of Applied Sciences. Since April 2014 Professor Putz is Institute Director of the Fraunhofer IWU, responsible for the Scientific Field of Machine Tools, Production Systems and Machining and he has additionally been holding the Professorship for Machine Tools and Forming Technology at the Chemnitz University of Technology.

Gunther Reinhart (*1956) studied Mechanical Engineering at the Technical University of Munich and finished his PhD Thesis in 1987. He worked as Executive Employee for the BMW AG and IWKA. He is the Professor of Management Sciences and Assembly Technology at the iwb of the TUM together with Prof. Dr.-Ing. Michael Zäh since 2007. Since 2016 he is a Member of the Management of the Fraunhofer Research Institution for Casting, Composite and Processing Technology IGCV in Augsburg.

Brandon Sai (*1990) has been a Research Associate of the department of autonomous production optimization of the Fraunhofer Institut of Manufacturing Engineering and Automation IPA in Stuttgart and a doctoral student at the Institute of Industrial Manufacturing and Management at the University of Stuttgart since 2018. He is currently working on behavioral models of production systems.
Marian Süße (*1992) started working at the Fraunhofer-Institute for Machine Tools and Forming Technology IWU as a Student Assistant in 2014. Since 2017 he is working as Research Associate in the Department Factory of the Future where he is in charge of the Group for Factory Design and Simulation since 2018. In his research and PhD-Thesis he focusses on the development of methods for factory layout planning and layout optimization approaches as well as ecologic impacts of production systems.

Martin Schreiber (*1989) is a Research Assistant at the Fraunhofer Research Institution for Casting, Composite and Processing Technology IGCV since 2016. His research fields are production management, production planning and operations research. He is also specializing in simulation and optimization of production systems as well as knowledge generation using Machine Learning.

Torben Wiese (*1991) studied mechanical engineering at the TU Dresden. Afterwards he started working at the Fraunhofer-Institute for Machine Tools and Forming Technology IWU in 2016. He works as a Research Associate in the Department Technical Cybernetics where he specializes in programming and engineering of industrial controllers. His focus lies in the real time part, such as path planning, axis transformation, process model integration as well as data acquisition and data analysis.
Abstract
In manufacturing, adherence to delivery dates is one of the main logistic goals. The production control department has to cope with short-term deviations from the planned route sheets. Because of unforeseen disruptions, e.g. machine breakdowns or shortage of material or personnel, in some situations, the promised delivery date to the customer is at stake. In practice, a fast and reasonable decision on how to deal with the delayed order is required. This decision process is often based on a qualitative analysis relying on the planner’s subjective assessment of a complex situation. To improve the quality of possible countermeasures this paper presents an application, which supports the decision process through a quantified analysis using real-time data from business application systems in combination with a simulation of the value stream. The developed app is part of the decision process and estimates the effect of selected countermeasures to accelerate a delayed order. Performance indicators illustrate the effect of the countermeasures on the specific order as well as the whole system. This approach empowers the planner to assess unforeseen situations and aims to improve the quality of the decision-making process. This paper describes the architecture of the application, its simulation ecosystem, the relevant data and the decision process to select the most effective countermeasures.

Keywords
Production Control; Decision Support; Discrete Event Simulation; Adherence To Delivery Dates; Disturbance Management

1. Introduction
The order fulfillment process in manufacturing companies is developing a higher complexity due to the increasing number of product variants and rising demand for individualized products [1]. In addition to this internal complexity, the market side demand fluctuations are of high and growing relevance for companies pushing them to increase the flexibility of their production system [2]. To cope with internal and external complexity, companies have to establish efficient and effective production control systems. These systems are characterized through a fast and robust adaption process in the context of unexpected disturbances. Therefore, production controllers face the challenge to deliver high-quality decisions in a complex environment to maintain an optimum between logistic performance and logistic costs. The main logistic target of manufacturing companies is adherence to delivery dates [3] and its importance is increasing within the supply chain [2]. It is expected that knowledge-based systems such as decision support systems can play a vital role to support the decision process in production planning and control [4]. In this paper the decision support system, which is developed within the research project “EkuPro – Decision support for short-term
production control”, is described. The focused research question of this paper is as followed: How must a decision support system be designed to support production control in improving adherence to delivery dates?

2. Motivation

Successful companies strive towards low inventories and high adherence to delivery dates. Such companies are leaders in punctually delivering to their customers by which they gain trust and avoid high costs and penalties from delayed orders [5]. To realize their logistic performance they have implemented high performing production control systems achieving a closed-loop production control [6]. Production control monitors the execution of the planned production program and therefore constantly compares the planned and actual output of the production system and engages if deviations from the plan occur. The tasks of production control include the releasing and sequencing of orders as well as the capacity control of the resources [5]. Unforeseen incidents, such as machine breakdowns, shortage of material or personnel challenge companies to meet promised delivery dates. In this situation, the production planner has to deliver a fast and profound analysis of the situation in order to decide how to proceed with the affected orders. If an order falls late due to the disturbance with respect to its production schedule, various actions to accelerate exist. The range of actions varies from shortly increasing production capacity (additional shifts) through prioritizing the delayed order to shifting the release of competing orders to a later timeframe.

Operational application systems such as ERP (Enterprise Resource Planning), MES (Manufacturing Execution System) or APS (Advanced Planning and Scheduling) support in the production planning [7] but don’t assist in the selection process of the right countermeasure. Their functionality supports the optimal scheduling of all manufacturing orders under the restriction of the available resources and the promised delivery dates. Due to this focus, these IT-systems do not support daily operational decisions. These choices concern for example how to handle a single delayed manufacturing order or the question, which measures are necessary to accelerate that delayed order to meet a promised delivery date. Therefore, production planners rely on their subjective assessment of the situation, which results in an unstable decision quality as well as a low adherence to delivery dates. Today’s application systems lack to support with a quantified analysis as well as a prognosis of the effects of possible countermeasures.

In the proposed methodology, an application is developed which focuses on the decision support for the acceleration of one specific delayed order. The developed decision support application considers monetary and logistic effects of possible countermeasures. The impact of the measures is estimated via a discrete event simulation of the considered production system and delivers a quantified basis within the decision process.

3. State of the art

The state of the art shows the research focusing on system support for production planning and control, simulation and decision support systems for production control in the context of adherence to delivery dates.

3.1 System support for production planning and control

In practice, companies rely on ERP, MES or even APS systems to support their tasks in production planning and control. Born out of MRP (material requirements planning) systems, which handled the material planning of a company, ERP systems are nowadays the information backbone for most companies across all industries [8]. They support the complete order fulfilment process from a customer order over the product manufacturing to the delivery and the invoicing of the order. ME systems gather information from the shop-floor level on personnel, work orders and capacity of resources and link them to the planning process in the ERP systems [9]. Therefore ERP and ME systems operate in different timeframes whereas ME systems have a shorter timespan since they mainly focus only on production-related processes. APS systems are a
functional extension for ME and ERP systems. They provide more powerful methods based on mathematical optimization for order sequencing, scheduling and resource allocation in comparison to the basic concept of material requirements planning (MRP). In the context of decision processes within production, ERP and ME systems provide information on the order status, the availability of resources and material as well as aggregated key performance indicators such as OEE or adherence to delivery dates. Therefore, the information basis for the decision support application, which is described in this paper, also relies on data from the ERP or ME system.

3.2 Simulation in production and logistics

Discrete event simulation (DES) is a method for analysing dynamical independencies in production and logistics with the goal to improve the design, control and operation of the material flow, the resources and the information flow [10,11]. Latest research has been conducted within the fields of simulation integration in enterprise applications, automatic model generation and collaborative modelling and simulation [11]. In automatic model generation approaches a simulation model is not created manually but is generated from external data sources [12]. Automated model generation reduces the effort to build up models and helps to standardize simulation models using three ways. Parametric approaches use predefined building blocks stored in a simulation library, which are then combined and parameterized for the modelling of a specific system. Structural approaches use a structural description e.g. a factory layout and hybrid approaches combine both methods. Most of today’s approaches are hybrid since they are using both parametric and structural data [13]. In the development of a simulation for a production system, the distinction is made between planning related and operations related simulation. In contrast to simulation in the production planning phase, operation related simulation supports the analysis of the actual behaviour of the production system under various aspects [14]. The operation related approach, discrete simulation helps to quantify the production system behaviour in a certain state through simulation scenarios. By modelling different strategies for production control tasks, e.g. adapting the order release policy or resource capacities in the simulation, the production control department is able to use simulation scenarios as decision support for the evaluation of their control options.

3.3 Decision support systems

“A decision support system (DSS) is a computer-based system that supports the choice by assisting the decision-maker in the organization of information and modelling of outcomes” [15]. Every decision a human makes is the attempt to change an unsatisfactory situation into a future satisfactory situation [16]. The decision process starts with an estimation of the current situation following the anticipation about the effect of certain or multiple actions, which create the desired future situation. Therefore, every decision process inherits a guess about the future. In the process of decision making several alternatives have to be studied and evaluated in terms of their implications on the system and thus their purpose in reaching the desired goal [17]. In a complex environment, the number of alternatives might be too large for a quick decision or the interdependencies of the regarded system are too multi-layered for a human to comprehend. In this case, a decision support system helps to better understand the interdependencies of the system [18] and supports in filtering for relevant data in the assessment of the situation. The focus on relevant parameters of the system helps the decision-maker to build up actionable knowledge [19]. This leads to a reduced effort within the decision process and a reduced count of suboptimal decisions. Therefore, decision support systems improve the process of decision making as well as the quality of the decision [20].
A decision support system consists of three main components (see Figure 1). The data component contains the necessary functions to gather and transform data to draw conclusions from it. The model component simplifies the considered system for understanding its behaviour. The modelling process summarizes and accumulates the data in order to compare different alternatives and their effect on the system. Different types of models from mathematical, optimization or financial models are possible within a decision support system [21]. Finally, the user interface receives data input and presents analytics, which is carried out by the DSS.

4. The concept and design of the decision support system

4.1 Countermeasures to accelerate a delayed order and their cause-effects

The development of the proposed decision support application aims to support the comparison of different countermeasures for a delayed manufacturing order. The main goal is to empower a production planner in the decision making on how to deal with a delayed order and to assess if it is possible to catch up and deliver in time. In the research project, six practical and applicable countermeasures were developed with industry employees to form a manageable set of alternatives. This straightforward approach helps the planner to quickly model different scenarios, which are then simulated. The six possible countermeasures are a “batch split”, “delaying a production order”, “changing order priorities”, “substituting resources”, “changing resource capacities” and “subcontracting” (see Table 1). In the following, the different countermeasures are explained by their chain of action in relation to the manufacturing control model [5]. Each countermeasure develops its effect by influencing the actuating variables of the control model.

Through a batch split the batch parts are processed in parallel, thus the first part is started earlier at the succeeding work system. Therefore, the actual load is decreased in that timeframe which results in earlier completion of the production order. Delaying a competing order in favour of the delayed one frees up resources (reduction in actual or planned input) and accelerating the remaining orders through an earlier start. A change in the priority of the delayed order lifts its status to a rush order. Therefore, it is preferentially processed at the workstations (change in planned or actual sequence), which reduces the waiting times, and thus the throughput time of the order. Through the substitution of a resource, e.g. the worker or the machine the process times or the planned start of a manufacturing order can be adjusted resulting in earlier or faster production. The same logic applies to the increase in resource capacities. By adjusting personnel or machine shift plans the actual output of the production system is increased, which accelerates the delayed as well as the rest of the production orders. Subcontracting production orders decreases the load (reduction in planned or actual input) on the production system by freeing up capacity for the delayed order.
4.2 The architecture of the decision support app

The decision support systems consist of two main components. The first one is the web application, which presents an interface to the user, and the second one is the simulation model, that runs in the background (see Figure 2). The user interface is designed as a web application to offer access from a variety of devices. It is used to model different decision scenarios and to analyse the simulation results. A scenario forms a set of multiple countermeasures, which are possible to accelerate the delayed order. The simulation is realised in “Tecnomatix Plant Simulation” which is a commercial DES tool developed by Siemens PLM Software. The simulation model is based on the toolbox and platform WOPS (“Wertstromorientierte Produktionssteuerung”), which is a verified toolset to simulate production systems developed at the Laboratory for Machine Tools and Production Engineering (WZL) [22]. For the decision support app which is realized in “EkuPro” WOPS was extended to correctly model the six relevant countermeasures.

In order for the tool to be used in a real production control environment, some requirements in the design have to be considered. The user is guided through the decision process, which is structured in three sections (Figure 2). Most importantly, the DSS has to produce realistic estimations about the production system for selected alternatives in order to be credible and helpful. Therefore, the simulation relies on live data from the business application systems (ERP) to have a realistic starting point (data component). The data which is used from the application systems is described in Table 2.

Table 2: Relevant input data from application systems to parametrize the simulation

<table>
<thead>
<tr>
<th>Input data</th>
<th>Source</th>
<th>Changeable control properties in a decision situation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Master production schedule</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scheduled production orders</td>
<td>ERP</td>
<td>resource order allocation, planned start time, order priority</td>
</tr>
<tr>
<td>Partially confirmed production orders</td>
<td>ERP</td>
<td>resource order allocation, order priority</td>
</tr>
<tr>
<td>Production order properties</td>
<td>ERP</td>
<td></td>
</tr>
<tr>
<td><strong>Resource availability</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worker shift plan</td>
<td>ERP</td>
<td>worker attendance timeframe, worker qualification, workforce quantities</td>
</tr>
<tr>
<td>Machine shift plan</td>
<td>ERP</td>
<td>machine availability time frames</td>
</tr>
<tr>
<td><strong>Product master data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Route sheet</td>
<td>ERP</td>
<td></td>
</tr>
<tr>
<td>Bill of material</td>
<td>ERP</td>
<td></td>
</tr>
</tbody>
</table>
The live data is gathered via an export from the ERP system and then uploaded from the user in the first phase of the decision process. The data is used to parametrize the simulation model in order to reflect the current state of the production system.

5. Decision process with app support

The decision process starts with a relevant disruption (Step 1 in Figure 2) from the production schedule for a certain production order. This disruption can be the sum of previous delayed process steps or can occur instantly if a necessary resource becomes unavailable. In this situation, clearance is required on how to proceed with the delayed order. First, an assessment is carried out if the order is going to be finished on time considering the current production schedule and utilization of the production system. Therefore, the current state of the production system is modelled. It is based on a snapshot of the production system using data from the ERP or ME system, which is imported via the user interface (Step 2). The user enhances the dataset with the latest state of information about resource availability and chooses a simulation timeframe. By finishing this step, the decision support system incorporates all the necessary information to simulate the status quo. Within the tool, we refer to this scenario as a base scenario, which does not incorporate accelerating measures.

The planner can simulate the base scenario on its own or he can directly add scenarios to the simulation run if countermeasures seem to be necessary by all means (Step 3). The additional scenarios are modelled within the web application. While activating certain countermeasures, the planner has to check their general feasibility in the context of working time acts, agreements with the working council or availability of extra needed staff. After modelling the scenarios, the simulation is triggered and the results are presented in the user interface (Step 5). In this step, the planner evaluates which scenario is the best choice to accelerate the delayed order on the basis of time related as well as monetary key performance indicators. The simulation gives an estimation of a realistic internal completion date for each possible scenario. With this information, it becomes transparent if the promised delivery date is at stake or if a certain countermeasure scenario can accelerate the order. The evaluation uses dashboards and lists to compare the scenarios. The dashboards visualize the order network via Gantt chart, a histogram for the adherence to delivery dates, key performance indicators (WIP, throughput time, utilization, operational costs) as well as a diagram to identify capacity bottlenecks (Figure 3). Relying on this set of transparent information the planner either refines the decision.

![Figure 2: Steps of the decision process within “EkuPro”](image-url)
scenarios and starts a new simulation run or decides to execute the set of countermeasures. This marks the end of the decision support process.

Histogram for adherence to delivery dates in daily time buckets

Key performance indicators for scenario comparison

Figure 3: Example analytics dashboards of the DSS

5.1 Expected improvements using the DSS

By estimating the effects of alternative countermeasures or their combination, the decision support app aims to improve the decision process through an increase in decision quality as well as analysis and decision latency (see Figure 4). By standardizing the analysis step through a standard procedure, which uses a simulation delivering an estimation of the production system behaviour, we aim to reduce the time span of this phase. The user-focused visualization of the simulation results delivers a reduction potential for the decision latency. In comparison to the decision without the support system, the decision-maker does not have to carry out manual analytic steps or search for information in the ERP or ME system. He is guided through the evaluation through familiar dashboards, which also help to decrease the time span of the decision latency. A major improvement is expected within the decision quality, in this case, the probability that a certain set of countermeasures accelerates the delayed order. By estimating the system’s behaviour, countermeasure alternatives, which have an insufficient effect on the internal delivery date, are ruled out within the decision process.

Figure 4: Improvements in the decision process based on [23]
6. Summary and outlook

The proposed decision support system is currently finalized. After that the real life evaluation is going to be started within medium sized companies. In the evaluation phase, the app is going to be used to support the production control of a manufacturer for clamping technology. Out of this phase, we are going to gather feedback through a structured questionnaire on the usability as well as the usefulness in various decision situations.

In summary, we introduced a decision support system, which empowers production planners to choose the most effective set of countermeasures in order to accelerate a delayed order. The proposed decision support system is based on discrete event simulation to estimate the behaviour of the production system by applying different countermeasure scenarios. The described DSS is developed within the research project “EkuPro – Decision support for short-term production control”. The next step within the research project is the validation of the tool in the operative daily routine by the industry project partners. Based on the research results from the practical evaluation an assessment is possible to what extent a decision support system which uses discrete event simulation can help within production control to improve internal adherence to delivery dates.

References


Biography

**Felix Steinlein** (*1989) works as a project manager and scientific assistant at the Institute for Industrial Management (FIR) at the RWTH Aachen since 2017. Before that, he gained experience in operations at Bosch and developed a strong background in lean management. His research focus lies within production planning and control, process analytics as well as enhancing operations with digitization.

**Yuan Liu** (*1987) works as a project manager and scientific assistant at the Laboratory for Machine Tools and Production Engineering (WZL) of the RWTH Aachen University since 2017. Before that, he worked for several years as a strategy consultant. His research focus lies within process optimisation and production systems as well as enhancing Business models in mechanical engineering.
Prof. Dr.-Ing. Volker Stich (*1954) has been head of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 1997. Prof. Dr.-Ing. Volker Stich worked for 10 years for the St. Gobain-Automotive Group and lead the management of European plant logistics. His research focuses on operations management, logistics and business applications systems.
1st Conference on Production Systems and Logistics

Tactical Production Planning for Customer Individual Products in Changeable Production Networks

Johannes be Isa¹, Thomas Zimmermann², Philipp Scherwitz¹, Gunther Reinhart³

¹ Fraunhofer IGCV, Am Technologiezentrum 2, 86159 Augsburg, Germany
² BSH Home Appliances, Savinjska cesta 30, 3331 Nazarje, Slovenia
³ Technical University of Munich, Boltzmannstraße 15, 85748 Garching bei München, Germany

Abstract

The requirements of future production are characterized by increasing demand volatility as well as very short delivery times and high timeliness in the order-based production environment. Furthermore, the trend to customer individual products leads to additional production planning challenges. Therefore, to react fast to these market trends, changeable production networks is a key to success. This technical contribution describes a method for tactical production planning for customer individual products in changeable production networks. The design of the method contains three main process modules (1) order-capability-comparison, (2) capacity planning, and (3) order-specific network structure. Underlying the former described modules, a data model is necessary and introduced. Furthermore, the simulation of the applied system on a prototypical implementation at BSH Hausgeräte GmbH, the largest home appliances manufacturer in Europe, is shown and explained. Finally, limitations are discussed and an outlook into future work for the research field in production planning of production networks is given.

Keywords

Changeability; Mass Customization; Production networks; Production planning; Production as a Service

1. Introduction

During the last 30 years, a strong trend towards global production in networks has been recognized [1]. Already today, more than 80% of the turnover of goods takes place in production networks [2]. Furthermore, the current business environment confronts the companies with high demand fluctuations and high competitive pressure [3]. In addition, saturated markets lead to an unbroken trend towards customer-specific products [4]. The production of these customer-individual products opposes the classical industrial approach of mass production [5]. In order not to produce these individualized products in a one-off production process, as it was the case in the traditional manufactory with high process costs, the research field of mass customization was created. In the area of mass customization, a distinction is often made between soft customization and hard customization [5]. Soft customization is characterized by comparatively simple adjustments to the sales product. Hard customization, on the other hand, represents a physical product individualization in production and for this reason; it places very high demands on changeability of production systems. Moreover, the market demands ever-shorter delivery times and high adherence to delivery dates [6]. These harsher conditions lead to higher complexity and increased requirements for future production [7]. Production networks, which have grown historically, tend to have rigid and inert structures [8]. Furthermore, the heterogeneous production capabilities are not taken into account and mapped in the...
production network. However, this mapping of the technical capabilities of each location at the production network level is necessary in order to make quick decisions for the acceptance of a customer-specific order. Another challenge is the capability-based production planning across locations [9]. This is necessary in order to avoid resource overbooking due to a lack of knowledge about existing production capabilities. Consequently, production planning based on capabilities in order-related production networks is required to break out of the existing rigid network structures. In order to meet these increased requirements and to master the complexity, it is necessary to apply new digital technologies [4]. The globally available internet as well as modern information and communication technologies (e.g. cloud approaches) make it possible to connect single production sites further to form integrated production networks [10]. According to the joint Fraunhofer IPA and Trovarit (software enterprise) study in 2017, the use of manufacturing execution systems to transform multiple single production sites to an integrated production network and thereby ensure deep networking of these locations at the production planning and control (PPC) level is increasing [11]. The PPS-Report 2017/2018, a regularly study concerning production planning and control, shows that more than 45% of companies in Germany plan to change or implement major changes to their production planning systems within the next one to two years to boost higher network integration [12].

Integrated production networks offer a high potential to balance the demand volatility described above by intelligently distributing customer-specific orders over the whole network in a short and medium term perspective. After this brief introduction, the state-of-the-art of research will be considered. Subsequently, the method for tactical production planning of order-related production networks will be presented. In the fourth part, the prototypical implementation, which has been realized at a world-leading producer of household appliances, is described. The fifth part concludes this work with a summary and an outlook.

2. State of the art

2.1 Production planning aspects

The configuration and coordination of production networks requires a wide variety of planning efforts due to the high degree of interdependence between the locations. In literature there are existing diverse approaches to structure the production planning aspects. According to Wittek et al. [13], the following four dimensions can be formed: (1) Planning tasks according to the company divisions e.g. procurement, production, sales; further along the (2) production system levels e.g. work station, cell, system, site, network; additionally according the hierarchical order of the (3) management level (strategic, tactical, operational) and the (4) planning horizon (long, medium and short-term) are common. The development of the Aachen PPC model contributed greatly to the understanding and classification of production planning especially on the network level [14]. Production planning, in turn, can be divided into three temporal areas. (1) Strategic production planning, which deals with long-term production issues, (2) operational production planning, which deals with short-term issues a few days before order release, and (3) tactical production planning, which is located in the medium term and lies between the strategic and operational time horizon [15]. The research concerning the medium term time horizon is relatively weakly developed compared to the two other temporal areas [16]. Tactical production planning is used to make decisions about the service areas, capacities and production organization in order to design an efficient production process. Based on the current market conditions described above, this tactical time horizon becomes more important [16]. The main focus in this work is therefore the fast and adaptive reconfiguration of networks to customer-specific products.
2.2 Classification of production networks

There are different definitions for production networks; the following one given by Röhrs [17] is technical but universally valid:

“A production network can be seen as a network where nodes adopt subtasks of a production process and maintain service exchange relations based on material and information flow.” [17]

In general, production networks can be classified and typologized at different levels. A common classification according to Rudberg and Olhager is into four classes: (1) cross-company networks, which describe several locations and several organizations, (2) internal networks, which describe several locations of one organization, (3) supply chain, which describes a location and several organizations serving this location, (4) factory, which describes one location of one single organization [18].

A further appropriate division into five categories of so-called phenotypes world factory, sequential or convergent, network structure, hub and spoke, local production plants takes place according to Abele et al. [12]. In addition, there are other classifications, but these should be more understood as extensions of those previously mentioned. The classification into different phenotypes is very common in the scientific literature but has the disadvantage of creating very rigid network structures as described above. Therefore the latest research linked to changeable, reconfigurable and fast scalable production networks can be found in the area of Production-as-a-Service which will be considered in this work [19, 20].

2.3 Production as a Service

The Production-as-a-Service approach has its origin in the information technology sector and is based on the idea of sharing temporarily unused resources. It can be seen as an adaptation of IT cloud approaches, Platform-as-a-Service, Infrastructure-as-a-Service and Software-as-a-Service to the operational technology sector [11]. Another term describing which is often used synonymously to describe the Production-as-a-Service approach is Cloud Manufacturing. An extensive investigation of this research area can be found in Siderska & Jadaan [21]. In general the approach of Production-as-a-Service is the focus on production capabilities rather than specific resources. This means further that the production planning in a Production-as-a-Service environment can be designed differently. A production process step can be interpreted as a micro service. This approach to the planning of production capabilities in an internal company network is further investigated in this paper. This leads to the two questions: (1) How does the mapping of these changeable production networks look like? (2) How can the data model of these networks be designed?

2.4 Customer individual products

In order to meet the rising customer requirements in the industrial sector, an attempt has been made in recent decades to meet these requirements with a wide variety of product variants [5]. In the meantime, however, it has become apparent that the management of these variants is a very complex and time-consuming task. In particular, keeping all variants in stock would go beyond storage locations and contradicts lean approaches. In order to reduce exactly these stock keeping units, many companies changed from the simple make-to-stock production approach to a combined make-to-stock and make-to-order approach [5, 10]. As mentioned above, there is also a trend towards customer-specific products in the private consumer sector due to the saturated markets. The general idea of mass customization was to produce individualized products to the time, process and cost of traditional mass production [22, 23].

The first step towards the customer-specific production described above is the reduction of batch size production and the determination of an appropriate customer decoupling point. This point separates production into a customer-related and customer-neutral control loop. The customer-related control cycle represents a direct link between the customer order and the production order. The customer-neutral control cycle represents production orders in prefabrication without direct reference to the customer [15]. The
customer individual demand requires fast adaption of the production capabilities and capacity. In this work, mass customization is related to the hard customization approach and will be later applied by the developed method. Further this work contributes to the development of solutions at the network level to the challenges outlined.

The comprehensive consideration and consolidation of the areas of mass customization, tactical production planning and changeable production networks will be discussed and deepened in the next sections of this paper.

3. Tactical production planning of changeable order-based production networks

In order to convey the setting of this work, the production network is depicted in Figure 1. An exemplary network with six distributed production sites of a European company is shown.

Furthermore, a simplified aggregated capacity utilization profile can be seen for each site. This profile also shows the flexibility corridor, which illustrates the frequently available possibility of scaling to compensate for minor demand fluctuations in single production areas or locations [24]. The constant aim of the companies is to find an operating point within this corridor in order to produce cost-efficiently. In the exemplary network representation, supply relationships already exist in some locations, represented by the dashed lines (logistics link) in Figure 1. This representation is well suited to visualize locations with over- or underutilization. For example, locations A, C and E show an underutilization. The utilization rate is below the lower bound of the flexibility corridor. This means existing resources are not used efficiently because there are not enough orders available at that time. Locations B and F are overloaded and therefore cannot provide all orders at the desired delivery time. In this simple example, however, a snapshot of a particular planning period is given. Under certain circumstances, it may also be the case that the capacity utilization of a location can look completely different at another planning period.

Both states described above should be avoided. The state of underutilization is by far the more critical problem, since high losses are usually caused in this state due to fix costs of workforce and production systems. In a state of overload, unrealised profits often result. This capacity utilization problem is caused by the difficulty to balance the supply and demand of capacity. Production sites around the world have to face this challenge.
The initial goal is therefore the optimal allocation of orders to the production resources in the production network of a company and the automated derivation of recommendations for action. In order to solve this problem, a methodological approach is required.

The following research questions were derived:

1) How should the data model for production planning for order-related production networks be designed?
2) How can the technical order capability comparison be mapped at the production network level?
3) How does tactical capacity planning take place in order-related production networks?
4) How can a cost-minimal network structure and the corresponding network allocation plan be set up?

The following sections attempt to further specify these research questions and present new approaches to solving them.

3.1 Data model

The aim of the data model is to illustrate the information requirements for tactical production planning of changeable production networks. The graphical modelling language Unified Modelling Language (UML) was used to specify and visualize the requirements see Figure 2. The data model supports the developed method and is divided into four submodels: order model, product model, logistics model and network model. The order model is used to describe the order specifications e.g. quantity and delivery date. To ensure data transferability, standards such as DIN 6789 were taken into account [25]. The product model defines the technical and geometrical information as well as the bill of material and the bill of processes. The logistics model serves to characterize the supply relationships, the means of transport and the costs incurred by the respective use. The network model is the most comprehensive one and is structuring, assembly as well as manufacturing capabilities according to DIN 8580 [26]. Furthermore, the network model depicts the locations, production systems and their capabilities as well as the site-related production costs. The capacity offered and the respective capacity utilization for each capability are also defined in this model. As a result, all four submodels together represent the overall data model and provide the information basis for the following method.

![Figure 2: Overview of the data model and the four submodels for changeable production networks](image-url)
3.2 Order-capability-comparison

The first method component of the order-capability-comparison aims to match the order requirements per process step with the existing technical capabilities of the individual production sites in the network. First a completeness check and an analysis of the order data is fulfilled. In the second step, the bill of materials and the bill of processes, which are available in the product model, are extracted. In the third step, every single process step is compared with the capabilities portfolio of the individual sites. An order capability matrix maps the necessary process steps on the one hand and the technically capable locations on the other. As a result of this method step a solution space is created. For this purpose, binary variables are used to indicate whether a technological capability is available at the production site (1) or not (0). Figure 3 shows a basic example considering three operations. The process starts with the source (S), followed by operations O122 Injection Molding, O517 Labeling and O412 Assembly and finally the sink (D).

![Figure 3: Order-capability matrix for three linked operations](image)

3.3 Capacity planning

The second module of the method is the production capacity planning on network level. The aim here is to determine the capacity plan for the solution space. The results from the previous method step will be enriched with further information regarding the production site capacity. The goal is a solution space that brings together technically capable and capacitively available locations in a comprehensive matrix. The lower section in Figure 4 displays an example of the solution space in a matrix structure with the three required process steps and the respective available locations.

The capacity planning on a tactical period is required to determine the capacity availability of the individual processes. It differs from operational planning, which is focused at detailed sequence planning. Tactical planning is done a few weeks before the start of production, so an initial rough estimation is appropriate. However, in order to deal with these order-specific requests, a new approach to capacity analysis is required. Up to now, planning in practice has usually been based on finite capacity, but often also on infinite capacity. However, the problem described above requires a dynamic capacity supply and, accordingly, dynamic capacity planning [27].

To accomplish this, an analogy was created from the yield management of airline operations. Airlines use an overbooking rate to maximize the utilization of their aircraft fleets. This means, for example, that 88 flight tickets are sold for a flight with 80 physically available seats. The reason for this overbooking rate is that, based on the airline's previous experience, e.g. around 10% of passengers do not board their flights. This means that according to the 10% no show rate 80 passengers can still board their plane. This allows the airline to achieve a very high load factor and thus to operate economically [28]. A knowledge transfer linked
to this overbooking Airline approach can also be created in the production environment. Production orders, however, have more diverse characteristics than standardized seats in an aircraft. Therefore, the airline approach is only transferable under certain restrictions. In addition to the usual fixed production capacity limit, this requires a nonphysical capacity limit. The term virtual capacity limit is well fitted to describe this dynamic approach and create a plannable capacity unit [27]. This virtual capacity limit thus allows an overbooking of the physical capacity. However, it is different from planning against infinite capacity. The virtual capacity limit is designed dynamically and therefore does not represent a fixed amount. This dynamic approach allows a fast response to a change in the order behaviour. Rather, this figure depends on the time or planning period, demand volatility, and the order probability based on historical data [27].

The virtual capacity limit, see Figure 4, is used to compare the capacity requirements and the available capacity of each capability in the planning horizon for each individual process step. Also for the capacity availability binary variables are modelled. The result of this method module are pre-selected suitable locations, which have both the technical capability and the capacity availability for each process step. This pre-selection of suitable locations thus forms the solution space for the next method module.

### 3.4 Order-specific network structure

In the last method module, the order-specific network structure and the corresponding production plan are created. For this purpose, the dimension of costs details the previously defined solution space. Both production and logistics costs are considered. Each technological capability of a location has specific production costs. These are qualitatively represented in Figure 5. A non-linear course and a minimum are recognizable here. However, usually step costs are also represented in such cost curves. In this study, an idealized curve progression of the costs is depicted. In the final solution space, two locations - Prague and Ljubljana - are technically and capacitively suitable for the injection molding process. For each of these locations, a cost curve is shown in Figure 5.

![Figure 4: Virtual capacity limit and solution space](image1)

![Figure 5: Integration of production and logistics costs](image2)

Each of these cost curves in Figure 5 (grey and black, representing a production site) show a local minimum for a certain production quantity. The local optimum of this idealized curve shape for one particular process step is relatively easy to calculate. The search for the global minimum, on the other hand, becomes complex when all the necessary process steps are taken into account.
Next to production costs also logistics costs always play an important role within the overall costs of production networks. Therefore, matrices for each mode of transport were developed in the data model. This information is taken into account in this method module for the integrated consideration of production and logistics costs. The lower part of Figure 5 shows a process variant graph according to Käschel et al. [29]. This was extended by logistics operations, which are essential for the integrated time and overall cost consideration. In this simplified example two processes, injection molding and assembly are considered. Both processes can be manufactured at two locations. The “exclusive or gate” (XOR) allows one process step of an order to be distributed to several locations. This process variant graph approach supports the modelling of an order splitting. The order splitting in turn allows companies to accept large-volume customer-specific orders that a single location could not handle in terms of capacity. The most-economic network structure and the corresponding production plan, including the order splitting, are determined in the last step by applying a solving algorithm. Due to the high degree of complexity of such a problem and the modesty of an approximate solution in the area of medium-term production planning, a heuristic procedure was chosen to solve the problem [30].

4. **Prototypical implementation at a large household appliances manufacturer**

A prototypical implementation was tested on the production network at BSH (former Bosch-Siemens Hausgeräte GmbH) a leading manufacturer of home appliances.

In the initial situation, as with many companies, there was a rigid assignment of products or product families to production sites. The Industrial Engineering Department by means of a forecast, which was designed on historical production figures, created the offered dimension and capacity for future production figures.

![Figure 6: Future production structure & concerning process steps (PlantSimulation)](image)

The major challenge here lies in the strong volatility in demand within a few weeks. To tackle this challenge, the production processes of two different kitchen appliances and a cumulative production number of over 5 million units per year were analyzed. To simulate the individual customer platform over 220 individual products based on the two different kitchen appliances were ordered. For each of the individual product five process steps were examined in detail: coating, assembly, testing, marking and finally packaging of the finished product. The technical capabilities required to successfully complete the production step were analyzed. Different variants with different process times and individual test mechanisms had to be considered. Furthermore, the provision of correct and sufficient material in all production areas must be ensured during planning. Within the industrial project, the prototype implementation could be validated by running a material flow simulation. The result is a uniformly high utilization of the individual production...
areas, which enables more cost-efficient production. The network approach can be interpreted in the prototypical implementation by the eight production systems distributed within the site.

5. Conclusion

Nowadays, a very high proportion of value creation takes place in production networks. However, these networks have mostly grown historically and have rigid structures. This work has scientifically described the relevance and potential of networked and, in particular, changeable production networks. Furthermore, a methodological approach was described for the tactical production planning of customer-specific products in changeable production networks. This consists of three method modules (1) order-capability-comparison, (2) capacity planning, (3) order-specific network structure, all of which were presented in this publication. Finally, the prototypical implementation of a part of the method in an industrial project in the field of household appliances production was briefly discussed.

Further topics were to be scientifically examined in the context of fast changeable production networks:

- Is it possible to expand to changeable cross-company production networks?
- How does the qualification and evaluation of production sites take place in order to participate in a changeable production network?
- How can social aspects be introduced in rapidly changeable production networks and how can their compliance be verified?
- How can high resource efficiency and environmental protection in changeable production networks be evaluated, ensured and rewarded?

References

Rapid and Long-term Measures for Prevention and Mitigation of Communication Barriers in Production Networks

Michael Kelber¹, Jörg Rainer Noennig¹, Peter Nyhuis²

¹WISSENSARCHITEKTUR - Laboratory of Knowledge Architecture, Technische Universität Dresden, Germany
²Institute of Production Systems and Logistics, Leibnitz University, Hannover, Germany

Abstract

Due to the constantly increasing complexity and dynamic of modern production networks within value chains, effective communication between the individual corporations is required. Communication barriers, can lead to delays in the value creation process, which can lead to losses of efficiency in the entire production network. Within the framework of "InterKom", a joint research project of the WISSENSARCHITEKTUR Laboratory of Knowledge Architecture at TU Dresden and the Institute for Factory Systems and Logistics at Leibnitz University Hanover, corresponding measures were developed on the basis of previously identified communication barriers in the context of production networks. These insights can be used to prevent communication barriers between suppliers, producers or consumers or to mitigate them as soon as they occur. The aim of the project is to provide a catalogue with precautions and countermeasures for corporations that supports the removal of communication barriers in the future.

Keywords

Production Network; Communication; Communication Barriers; Precautions; Countermeasures

1. Introduction: The Role of Communication Barriers in Production Networks

Communication is a decisive core competencies of networked corporations [1] and needs to be included as a key factor in new production factor systems [2]. According to BRUHN, communication in the context of production networks, serves for internal agreement-finding between departments and employees within a corporation, as well as for external consultation with customers [1] or different corporations at different locations, which according to Rudberg and Olhager can be categorized as an "internal network" [3]. Especially in production networks in which information between collaborating corporations flows in both directions – upstream and downstream – along the value chain, [4], inter-organisational processes must be coordinated by consistent communication. Delays or false transmission of information at some communication interface can be called “communication barriers”. According to Shannon, a communication barrier can be defined as an obstacle between the transmitter and receiver of a communicative exchange [5]. Errors, that occur through a communication barrier, continue upstream or downstream and can lead to a slowdown of the entire value creation process - a phenomenon that is called "bullwhip effect" [6] [7]. According to Lee et al., causes of the bullwhip effect are updates of the demand forecasting, order batching, price fluctuation or rationing and shortage when product demand is too high [8], as well as lead times [9]. This vulnerability of the value creation process of manufacturing corporations by communication barriers requires what Röhner & Schütz call “communication competence” [10] – only thus error transmission can be avoided in advance, and occurring communication barriers be mitigated timely.
2. Framework: A Descriptive Model to Identify Communication Barriers

In the framework of "InterKom", a DFG-funded research project of the WISSENSARCHITEKTUR Laboratory of Knowledge Architecture at TU Dresden and the Institute for Factory Systems and Logistics at Leibniz University Hanover, communication barriers and corresponding generic measures for avoidance and mitigation in the context of production networks were collected. Especially in non-physical spaces, MOESLEIN sees a need for the visible and tangible in an "initially largely invisible world of information" [11]. In this context, "InterKom" presents a descriptive model that maps the communication between employees of different departments or corporations by so-called "communication paths" [12]. The descriptive model distinguishes between different layers, which can localise the communication barriers in the overall structure of the value chain, as shown in Figure 1.

On the first layer (Organisation Layer) the organisational classification of the corporation is carried out in the overall structure of the production network. Here, communication interfaces appear within a corporation (intra-organisational) and between corporations of a production network (inter-organisational) [13]. The second layer (Responsibility Layer) represents the organizational assignment of responsibility areas of the corporations that communicate within the production network. Here, communication interfaces appear within an area of responsibility, such as a department or sub-department, at (intra-departmental) or between different departments (inter-departmental) layer. The third layer (Operative Layer) considers the employees of a department or sub-department of an organization. It thus represents the smallest possible scale in a production network, which is necessary to localise and characterize communication processes. The fourth layer (Communication Layer) describes the medium [5] of the communication object to be transmitted in a communication context [14] between sender and receiver according to SHANNON, which can basically be divided into visual media (e.g. letter, e-mail), auditory media (e.g. telephony, voice messages) and audiovisual media (e.g. video telephony, face-to-face) [11]. In order to analyse inter-organisational production networks, the focus of the communication barriers should be on the communication paths that take place between the same or different departments of different corporations. Examples are the communication between suppliers and producers, or the communication between end customers and producers.

Figure 1: Descriptive Model of Communication Paths within a Production Network
3. Problem Description: Measurement of Communication Barriers in Production Networks

Communication barriers in production networks can occur in corporations at all formal and informal communication interfaces [15] between sender and receiver. Formal communication can be understood as everything that is involved in processing the product to be produced, such as order confirmations, material orders or price negotiations. Formal communication thus follows the formal structure or hierarchy of the corporation [16] and, since it takes place explicitly, can be identified by theoretical and empirical means. In contrast, there are very few theoretical models and empirical studies on informal communication. WAGNER cites as a reason that informal communication usually leads to the transmission of confidential information, the exchange of which is based on trust and secrecy and the transmission of which would be critical [17]. Informal communication channels such as grapevine and gossip can often fill the gaps left by formal communication [16]. Examples include recommendations from partners, discounts in price negotiations or the transmission of confidential information. The importance of communication within corporations, especially via information and communication technologies (ICT) requires, according to LUO & BU, the three interrelated elements of access, search and codification as well as a „significant attention for not only theoretical advancement but also practical implementations“ [18].

The focus of "InterKom" is on inter-organisational communication (formal and informal) between corporations that have joined forces to form a common value chain. The constellation shown in Figure 1 is thus to be examined according to the descriptive model explained above, i.e. by using communication paths: On the Organizational Layer, communication takes place outside the corporation; on the Responsibility Layer, communication takes place inside or outside the department; on the Operational Layer, communication barriers are identified. Foridentifying communication barriers, as a first step, general communication barriers as well as barriers in production environment were compiled by literature research, e.g. WIK-WAH FONG & CHU [19], NIIMINÄKI et al. [20] and GRANHAGEN JUNGER, et al. [21]. In the next step, all collected communication barriers were validated in qualitative interviews and surveys by representatives of industries working in the field of logistics, production or distribution as part of a value chain; they either confirmed or classified certain communication barriers as irrelevant. Also, they added further communication barriers to the list, if such occurred in their environment and were not yet mentioned in the literature research. The communication barriers validated that way could be assigned to three different causes, which also represent the classification of communication barriers in the following sections: 1) organisational, 2) personnel and 3) technical communication barriers. Organizational communication barriers comprise all processes that deal with the overarching coordination of the value creation process, such as the responsibility of employees for certain activities. As shown in Figure 1, the cause of organisational communication barriers can primarily be found in the administration, i.e. on the organisational layer of the descriptive model. Organisational communication barriers are, for example, lack of standard processes, unclear responsibilities or unclear decision-making power. Personnel communication barriers comprise all processes that can be attributed to personal concerns, abilities and deficits of employees, such as different mother tongues, use of technical language, or use of "insider-terms". Causes for personnel communication barriers can primarily be traced back to the human factors [22] of employees at the responsibility level, such as the personal expertise, abilities and weaknesses of individuals. Technical communication barriers include all processes that are dependent on technological aids and exclude human factors. Thus the cause of their occurrence lies neither in the administration nor in the employee, but on the communication level, which includes the medium of communication. Technical communication barriers include, for example, poor local infrastructure, poor transmission quality or large spatial distances.
4. **Approach: Rapid and Long-term Measures for Prevention and Mitigation of Communication Barriers in Production Networks**

The next step in the InterKom project was to collect measures for the communication barriers described above. First, approx. 20 generic measures in the context of production networks and other domains were collected by literature research. Subsequently, representatives from industries active in logistics, production or distribution as part of a value chain were shown the previously collected communication barriers in a qualitative survey. About 25 measures were formulated by the participants for all communication barriers which play a role in their environment and which affect them in their working routine. It could be observed that most of the measures mentioned are theoretical recommendations, but are not practically applied in the corporation. The next step was to list the measures collected from the literature and to let representatives indicate which measures were relevant or applicable to their own working environment. A validated catalogue of generic measures could be derived from this. Below is a selection of generic measures with their corresponding classifications:

<table>
<thead>
<tr>
<th>Precautions / Countermeasures</th>
<th>Rapid Implementation</th>
<th>Long-term Implementation</th>
<th>Prevention</th>
<th>Mitigation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training of employees who communicate with foreign-language customers/partners</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integration of native speakers/interpreters</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Transmitter (who knows both cultures)</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Qualification: for example through team trainings</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Technological aids e.g. language software</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Active listening: Explain facts to others in order to understand them better</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Translation into easy language</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Definition of standardised processes, formats and procedures</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Direct Mailing</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Definition of minimum requirements for documents</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Periodic checking of the up-to-dateness of electronic tools</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Implementation of a Wiki (knowledge store + definitions)</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

As shown in Table 1, measures collected from respondents as well as from literature could be classified in two ways. The first classification distinguishes measures that can be applied rapidly when a communication barrier occurs in the corporational environment from those requiring a lead time for implementation. For example, a change of communication medium can take place immediately, whereas the introduction of...
standard processes requires a lead time within a department or the entire corporation. The second classification distinguishes between measures that either prevent communication barriers as precautionary measures or reduce them as countermeasure in case of already occurred communication barriers or both. For example, periodically checking the up-to-dateness of electronic devices can prevent their susceptibility to errors, whereas language software can reduce a lack of language skills.

5. Findings: Consolidation of Communication Barriers and Precautions / Countermeasures

In a last step, the validated communication barriers were juxtaposed to validated precautions / countermeasures. According to ANDERSSON, these measures should aim at implementing routines [23] that prevent or mitigate communication barriers. As shown in Figure 2, one or more suitable measures were assigned to each communication barrier (arranged according to organisational, personal and technical aspects). For the representation, 15 arbitrary communication barriers were selected from the previously validated barriers.

Figure 2: Exemplary linkage of communication barriers with validated precautions / countermeasures

The selection of the 15 communication barriers shown makes clear that measures can usually be applied to several communication barriers. They can also be applied across all layers, for example by training teams to clarify unclear responsibilities (organisational barriers) or by helping out with lack of technical knowledge (personal barriers). However, it should be noted that it was not possible to find suitable measures for every communication barrier up to this point of the research project. For the practical avoidance and mitigation of communication barriers in the context of production networks, an iterative approach is required, as shown in Figure 3.
The aim of communication is to ensure smooth and secure transmission of information without barriers, for both sender and receiver. If a barrier occurs during communication between sender and receiver, corresponding countermeasure must be initiated. Depending on the classification of countermeasure, it can be initiated rapidly (B1) or with delay (B2). The latter can be used if the communication barrier causes a longer after-effect, such as a poor local infrastructure. If the short-term and long-term countermeasures lead to successful communication, the sender and receiver usually learn and acquire knowledge. These insights can be used to initiate precautions to prevent communication barriers in advance. In this case, they can be used rapidly (A1) or with a planned lead time (A2). It is recommended that measures A1 and B1 be implemented whenever possible (by low efforts). Measures A2 and B2, on the other hand, should be used if there are serious communication barriers that are to be expected or have occurred in the past.

The list of precautions and countermeasures for previously validated communication barriers is constantly being expanded within the framework of InterKom, targeting a comprehensive catalogue of measures that companies can implement in their daily work routines.

Acknowledgements
Deutsche Forschungsgemeinschaft (DFG) supports the research project descriced above [Grant number NY 4/59-1].

References


Biography

**Michael Kelber, Dipl.-Ing.** (*1992) has been a research associate at WISSENSARCHITEKTUR – Laboratory of Knowledge Architecture at Technische Universität Dresden since 2017. He previously studied architecture (Dipl.-Ing) at the Technische Universität Dresden.

**Prof. Dr.-Ing. Jörg Rainer Noennig** (*1973) has been director of the Wissensarchitektur – Laboratory of Knowledge Architecture at TU Dresden since 2015 and Professor for Digital City Science at the CityScienceLab at HafenCity University of Hamburg since 2016.

**Dr.-Ing. habil. Peter Nyhuis** (*1957) has been head of the Institute of Production Systems and Logistics (IFA) at Leibniz University Hanover since 2003. Prof. Dr.-Ing. habil. Peter Nyhuis was a member of the German Science Council (2014 – 2018) and has been Chairman of the Science Council of the AiF (GAG3) since 2015. He is also a member of the German Academy of Technical Sciences acatech and the Scientific Society for Production Technology (WGP).
Assessing product portfolios from a production logistics perspective

Tim Kämpfer¹, Torben Lucht¹, Philipp Schäfers¹, Peter Nyhuis¹

¹ Leibniz University Hannover, Institute of Production Systems and Logistics, Germany

Abstract

The increasing individualization and the growing customer demand for product variety leads to a constant shortening of product life cycles and to the necessity of periodically rationalizing product portfolios. For this reason, approaches to product portfolio assessment offer methods that allow a financial or market-oriented valuation of existing products in portfolios. When assessing products in product portfolios, conventional approaches do not explicitly take the logistical impact of products on the logistics performance or costs of the production into account. The consequence of neglecting the logistical assessment dimension to product portfolios is that products, that have a negative impact on the logistics performance of a company, are not part of a critical examination. This paper therefore presents an approach that aims at developing a methodology to assess product portfolios both from a logistical as well as from financial or market-oriented perspectives. To this end, the approach initially works the influence of individual products and product characteristics on the logistics performance and logistics costs of production out. The consolidation of these findings with further evaluation variables then enables a product portfolio optimization with explicit consideration of a logistic assessment dimension.

Keywords

Product portfolio analysis; portfolio optimization; portfolio assessment; production logistics

1. Introduction

Global competition, the opening and development of new sales markets as well as dynamically changing customer requirements lead to a constantly growing and more individualized product variety [1, 2]. Furthermore, the constant shortening of product life cycles requires frequent adjustments to the product portfolio [3]. The above-mentioned factors lead to an increased inclusion of new products and variants in product portfolios of manufacturing companies. Exotic products with low production volumes can cause high complexity induced production costs, while high-volume standard products cause lower costs. Higher complexity induced production costs result from higher set-up times and non-productive time in production [4]. Especially in the case of product portfolios that have grown over the long term, a lack of transparency regarding the impact of exotic and standard products on the logistics performance or logistics costs of companies can lead to the retention of exotic products with high logistics expenditures product portfolios. The deletion of products from portfolios being an unappealing managing decision reinforces this effect [5].

The retention of exotic products can lead to their cross-subsidisation by standard products, whereby the spread between the costs caused by exotic products and their selling price is being subsidised by the profit of standard products in the portfolio. This leads to a competitive disadvantage regarding standard products caused by a loss of profits through said cross-subsidisation [4]. Figure 1 illustrates this problem.
Figure 1: Need for action in managing complexity in product portfolios [4]

The systematic analysis and control of variant diversity induced complexity is of high importance and at the same time a challenge for industrial companies. A joint study conducted by several production technology institutes shows, that half of the 170 industrial companies within the survey are highlighting complexity through variant diversity as a key challenge for their production planning and control activities [6]. Existing approaches of strategic product portfolio management offer strategies regarding the continuation or rejection of products or support the identification of target areas for successful marketing based on known strengths and weaknesses of a company [7, 8]. However, these approaches mostly focus on economic aspects while neglecting logistical aspects of product portfolios. This circumstance carries the risk that products, that meet financial margins or demands regarding their respective market potential, but at the same time have a negative influence on the logistics performance of a company, are not subject to a critical examination. The deletion of such products has an impact on the entire supply chain [4]. Increasing transparency with regard to logistics-relevant characteristics of product portfolios in the course of product portfolio management is therefore of great importance. For this reason, this paper describes an approach that focuses on the development of a methodology to logistical product portfolio assessment in section 3. A literature review regarding existing approaches for the assessment of product portfolios and products in Section 2 precedes the description of the approach. These approaches form an important basis for multi-criteria portfolio assessment.

2. Approaches to the assessment of product portfolios and products

The origin of portfolio assessment approaches lies in the financial sector. In 1952, MARKOWITZ published an essay on the optimized composition of securities [8]. According to MARKOWITZ, the basic problem of portfolio composition is the uncertainty of future returns. He therefore models financial returns as random variables with an expected value, standard deviation and correlation. The expected values of the returns should be maximised, whereas the standard deviation of the return, i.e. the financial risk, should be kept as low as possible. In the context of contribution margin accounting [9], cost and performance accounting evaluates the profitability of products within a portfolio. The difference between revenues and variable costs, which is the amount available to cover the fixed costs in the company, is calculated. Depending on the contribution margin, strategies can be derived for corresponding products. For example, it is conceivable to withdraw products with a low or even negative contribution margin from the market and to invest more in products with a high contribution margin in order to strengthen their market position.

The Portfolio Market Matrix [10] establishes a link between a company's products and their target market. This approach considers the extent to which the company has already penetrated a market and the extent, a
change in the product portfolio is necessary. The growth–share matrix [11] sorts products of companies in a matrix based on their relative market shares and forecasted market growth. The approach arranges the products into four categories, while connecting the categories with strategies such as market or product development. Another prominent approach to product portfolio assessment is the GE multifactorial analysis [12], developed in a collaboration between McKinsey and General Electric. The approach encourages selecting a list of influencing factors on the market attractiveness and business strength of products at the beginning of the analysis. Instead of four categories, the approach defines nine categories with corresponding standard strategies and arranges analysed products into said categories. Due to the refined breakdown of variables and the multifactorial assessment, more differentiated strategies can be derived with this portfolio analysis than with the approach described in [11]. Due to the subjective selection of the parameters, however, the approach bears the risk of misinterpretation and insufficient comparability.

KLIMKE shows that a change in the portfolio position within the growth-share matrix also influences production-related variables and thus logistical target values [13]. In his analysis, the products are first classified using the market share market growth portfolio. This classification and the defined standard strategies form the basis for deriving possible adjustments to the production. These adjustments include material flow optimization, changes to production control methods or changes to the goods distribution systems.

The variant tree [14] developed by SCHUH represents an approach to control the diversity of product variants in companies. Through its application, variants can be identified which do not offer any significant financial benefit to the company. With the aim of developing a variant-oriented product design method, CAESAR adapted SCHUH'S ideas and developed a methodology based on the Failure Mode and Effect Analysis (FMEA) to design a wide variety of serial products in accordance with cost aspects [15]. The so-called Variant Mode and Effect Analysis (VMEA) pursues the goal of technical and cost-based control of variant diversity. Here, a reference to the disciplines product planning, product development and product design is established [15]. Thus, the VMEA is not really an assessment approach for product portfolios, but rather a methodology for strategic portfolio optimization through a cost-oriented design of product variants. Besides aiming at a reduction of the number of variants, the approach aims at increasing the average profit contribution per variant.

LÖSCH describes an approach for controlling and designing the diversity of variants, taking the cost and benefit effects of the assessed variants into account [16]. The essential idea of this approach is the determination of the optimal diversity of variants. BROSCH has presented an approach for capturing and evaluating product-variation-induced complexity as well as to point out possibilities for its reduction [17]. Both the product and the value chain are included in this process - this is why BROSCH refers to it as Design for Value Chain. The methodology consists of two method blocks: The identification of strategic fields of action and the support in product development. As part of the identification of strategic fields of action, the approach records the external and internal product diversity and identifies the complexity drivers. By prioritizing these drivers and assigning them to generic fields of action in complexity management, BROSCH’s approach enables the derivation of company-specific strategies for reducing complexity. Based on this, the approach records the target product variety and various alternatives for product and order processing (e.g. alternatives for the positioning of the customer order decoupling point and variant development point).

RIESENER et al. analyse correlations among portfolio-relevant corporate key performance indicators [18]. The methodology uses a neural network to model correlations and to predict future trends for analysed key performance indicators. The methodology aims at supporting companies in proactively managing their product portfolio by anticipating the product portfolios future development. While RIESENER et al. do analyse cause-effect correlations of portfolio-relevant KPI’s, they do not take driver variables of products or
portfolios like the number of products in a portfolio into account. In addition, RIESENER et al. do not systematically link said driver variables to logistical target values.

BOHL investigates the complex interdependencies between product and production complexity [19]. By modelling the complexity-related dependencies between product and production, he enables standardization. Initially, BOHL defines a system with the four sub-areas: product range, product architecture, production structure and supply chain. He captures specific parameters that characterize complexity for each subarea. By marking the dependencies between the sub-areas and the parameters, the approach creates a qualitative model of cause-effect relationships within the system. BOHL subsequently models important cause-effect relationships in the form of characteristic curves. The work of BOHL outlines some of the relevant cause-effect relationships that are relevant for the development of the approach that this paper describes.

3. Development of a logistical product portfolio assessment methodology

The approaches to product portfolio assessment presented here show gaps in the consideration of cause-effect relationships between the product portfolio and logistics performance and cost parameters. For this reason, the Institute of Production Systems and Logistics (IFA) develops an approach for the logistical assessment of product portfolios and their products.

The primary goal of the approach described here is to be able to identify problematic products from a logistical point of view and, in addition, to derive optimization strategies. Figure 2 shows the approach to the development of a logistical product portfolio assessment methodology. This chapter subsequently describes necessary steps to create said methodology in detail.

Figure 2: Approach to the development of a logistical product portfolio assessment methodology
A mere assessment of a product portfolio according to logistical criteria does not appear to make sense. If, for example, a product has high sales numbers and a high margin, any company will continue to rely on this product, even if it is to be rated critically from a logistical point of view. Here other measures than a discontinuation of the product are conceivable as for example a product design adapted to the requirements of production. Section 2 shows that a large variety of methods for product portfolio assessment already exists. The majority of these methods focus on financial and market fixated assessment parameters. The aim of the approach presented in this paper is to combine these valuation dimensions with the logistical assessment dimension. Figure 3 shows an exemplary basic model of the multi-criteria product portfolio evaluation.

Figure 3: Possible dimensions of a multi-criteria product portfolio assessment method

The assessment of products in a portfolio from a logistical point of view requires the linkage of logistical target values with driver variables. Logistics performance and logistics costs are the two categories of logistical target values. To be able to link driver variables to target values of the logistics performance and costs such as delivery reliability, step 2 of the presented approach requires the identification of relevant variables. Driver variables can be properties of products, of product portfolios as well as properties of the supply chain. Figure 4 shows examples of possible driver variables (such as set-up-requirements of products, mean set-up-requirements of the portfolio or the position of the variant formation point in the supply chain) and its assignment to the three mentioned levels.

<table>
<thead>
<tr>
<th>Driver variables</th>
<th>Logistical target values</th>
<th>Logistics performance</th>
<th>Logistics costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level product</td>
<td>e.g. set-up-requirements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level portfolio</td>
<td>e.g. mean set-up-requirements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level supply chain</td>
<td>e.g. variant formation point</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: Possible hierarchy of driver variables

The identification of relevant driver variables of products and product portfolios forms a fundamental basis for the development of qualitative cause-effect models. These cause-effect models describe the relationships between driver variables and logistical target values. The relationships can be used in the estimation of effects of changes to products or product portfolios with regard to logistical target values such as delivery
reliability or delivery time. For this reason, the modelling of cause-effect relationships represents a key step in the development of the logistical product portfolio assessment method.

The previously described identification of techniques, that assess products and portfolios from a market or financially oriented perspective in step 1 and the modelling of cause-effect relationships between driver variables and logistical target values in step 3 enables their combination within one multi-criterial portfolio assessment method.

In order to ensure a purposive use of the resulting assessment method, the approach develops a workshop concept. The workshop concept represents the methodical framework of product portfolio assessment and optimization. A software demonstrator supports the workshop concept in the visualization and implementation of results. Systematic tests based on real data records from industry partners verify the workshop concept and the software demonstrator.

4. Conclusion

This paper describes necessary steps to develop a methodology for assessing product portfolios from a production logistics perspective. To this end, the paper discusses relevant approaches and methods for assessing product portfolios and respective products. The discussion enabled the derivation of research gaps in the field of product portfolio management. The gaps identified are apparent in the form of missing links between product portfolios and logistical target values.

The presented approach aims at closing these research gaps by developing a methodology that allows the assessment of products and product portfolios from a production logistics perspective. In order to achieve this goal, the approach links product and product portfolio properties with logistical target values within cause-effect models as a first step. Subsequently, the approach combines the developed cause-effect models with financially and market oriented approaches to portfolio assessment. The combination of traditional and logistics-oriented assessment dimensions enables a multi criterial assessment and optimization of product portfolios. In form of a workshop concept, the applicability of research findings in relevant industrial companies is to be assured. The concept supports manufacturing companies in product portfolio management by deriving measures to optimize portfolios within a step-by-step procedure.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 423857520. The authors kindly thank the DFG for the financial support to accomplish the research project “Logistische Produktportfoliobewertung”.

References


Biography

Tim Kämpfer, M.Sc. (*1992) studied production engineering and logistics at Leibniz University Hannover and works as a research associate in the field of production management at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover since 2019.

Torben Lucht, M.Sc. (*1991) studied industrial engineering with the focus on production technology at RWTH Aachen University. Since 2018, he works as a research associate in the field of production management at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover.
**Dipl. Ing. Philipp Schäfers, MBA** (*1987) studied Mechanical Engineering at Karlsruhe Institute of Technology. Subsequently he completed a MBA programme. From 2014 to 2019 he worked as a research associate at IFA of Leibniz University Hannover.

**Prof. Dr.-Ing. habil. Peter Nyhuis** (*1957) studied mechanical engineering at Leibniz University Hannover and subsequently worked as a research assistant at IFA. After completing his doctorate in engineering, he received his habilitation before working as a manager in the field of supply chain management in the electronics and mechanical engineering industry. He is heading the IFA since 2003. In 2008 he became managing partner of the IPH - Institut für Integrierte Produktion Hannover gGmbH.