

2nd International Conference on System-Integrated Intelligence: Challenges for Product and Production Engineering

Evolutionary approach for an optimized analysis of product life cycle data

Roland Lachmayer^a, Iryna Mozgova^a, Bastian Sauthoff^a, Philipp Gottwald^{a*}

^aLeibniz Universität Hannover, Institut für Produktentwicklung und Gerätebau, Welfengarten 1A, 30167 Hannover, Germany

Abstract

The application of life cycle data of smart products offers new opportunities for the product development process. Nowadays, products often consist of adaptive design variants of an existing product. Taking this into account, the new product development approach called technical inheritance is developed analogous to biological inheritance. This approach considers the intergenerational evolution of design characteristics.

Enhanced smart products are developed within the Collaborate Research Center (CRC) 653 called "Gentelligent Components in Their Lifecycle". These features the capabilities to sense, collect and transfer life cycle data inherently by using genetic product properties and artificial intelligence. By using technical inheritance optimization strategies are currently investigated and the design of gentelligent components is researched. During the technical inheritance various monitoring concepts are applied to realize a targeted algorithmic feedback of lifecycle information from smart products. For a targeted algorithmic feedback of product life cycle information methods of data mining are applied. These include the objectives of data beneficiation as well as information detection. The boundaries of the investigations are determined through the gentelligent components. Therefore highly mechanical loaded systems are in focus. It follows that the physical aspects and specific life cycle incidents are major objects for the monitoring concept of the product life cycle. The approach aims at the integration of an evolutionary algorithm to identify the component specific critical loads as well as the optimal allocation of loads cases. The results of this concept are exemplified by a wheel suspension which is part of the demonstrator of collaborate research center.

© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Peer-review under responsibility of the Organizing Committee of SysInt 2014.

Keywords: Data Mining; Life Cycle Information; Evolutionary Algorithm

* Corresponding author. Tel.: +49 511 762 3333; fax: +49 511 762 4506.
E-mail address: gottwald@ipeg.uni-hannover.de.

1. Introduction

The application of smart products in the product development processes holds a few challenges for the research [1]. For example these products could be used for methods of X-in-the-Loop [2] or the allocation of life cycle information [3]. In context of the CRC 653 enhanced smart products, called gentelligent components, are developed. Thereby new manufacturing technologies and materials are investigated [4].

In the course of this research center the approach of the technical inheritance is developed. The focused project in the subproject N4 of “CRC 653” is the transmission of principles from the biology into the technical processes and applications [5]. For the technical inheritance the complete life cycle of smart products has considerable meanings. In contrast of the nature-inspired process model [6] and the autogenetic design theory [7] this approach creates an intergenerational view of development processes.

Often new products represent only an adaptive design of previous product generations. The approach of the technical inheritance occupies this fact [8]. One resultant challenge in the research demonstrates the data analysis of life cycle information from gentelligent components. Therefor principles of biology were analyzed and evaluated. At the example of a demonstrator from the CRC 653 a data mining algorithm which is adapted from biological systematics is applied to get useful information for the development process. This includes the objective to adapt the next product generation to their environmental influences.

2. Methodology

Present research investigates the algorithmic feedback of product life cycle information for a design evolution. This process is developed taking account of the technical inheritance. Therefore an intergenerational development process is generated and depicted in Fig. 1.

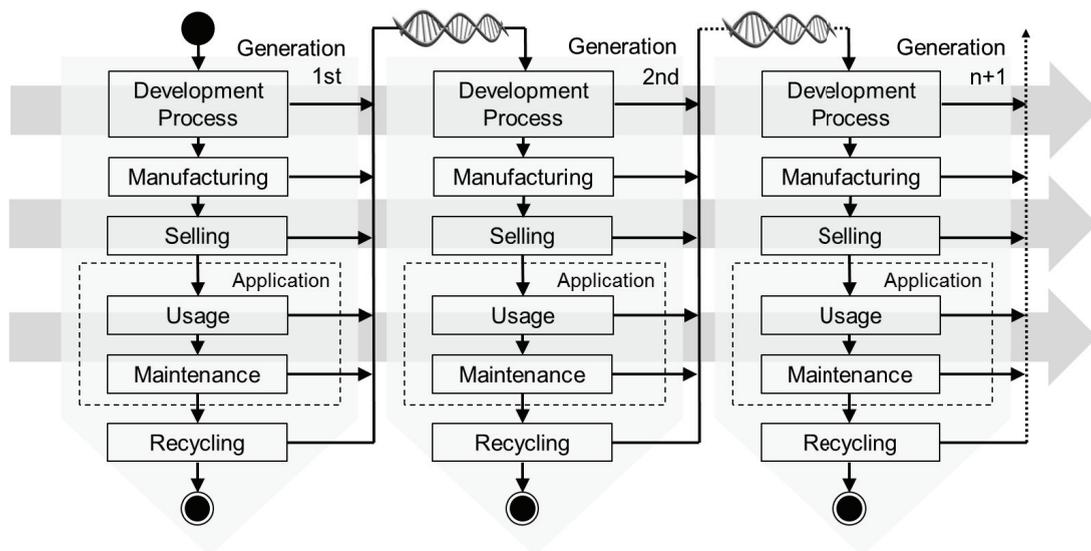


Fig. 1. Intergenerational development process of the technical inheritance.

The fundamental investigation regarding the realization of the technical inheritance is divided into three working packages [9]. The first package investigates the development with the gentelligent technology. Gentelligent components are featured to collect, save and transmit product life cycle information inherently. Methodologies to integrate single gentelligent components into assemblies are investigated to get the usually possible product life cycle information for example usage data. Also approaches for transforming the inherent product data for the application in development process, like elongation at defined points on the component to acting forces at junction points of the assembly, are analyzed. In addition methods of data mining are integrated in the development process

to achieve a targeted feedback of life cycle information [10]. The implementation and the application of a data mining algorithm is focused in the next sections. The data analysis itself is part of the second working package, defined as “statistic operator“.

The objective of the investigations in the statistic operator represents the identification of develop-relevant information, which for example influences the specifications, in the product life cycle data sets. The results of such analysis form the basis for the third package the so called “design optimization”. The generative parameter models, which are developed in the design optimization, complete the investigations of the design evolution and the subsequent process of the technical inheritance [11].

3. Data Analysis

The Formula Student race car of the Leibniz Universität Hannover represents one of the demonstrators in the CRC 653 [12]. The wheel suspension is in focus, where some intelligent technologies are integrated to collect life cycle data, especially the applied loads. With different techniques it is possible to measure the forces during a race. Such information builds the basis for the development of the next product generation. The wheel carrier has three develop-relevant junction points, at which the forces are of particular interest. Additionally equivalent to races multi-body simulations of the race car are done.

The data set of such simulations represents the analyzed data. Therefore the forces $F_{x,y,z}$ for every dimension were obtained at each junction point A, B, C of the simulations. In Fig.2 the forces are depicted.

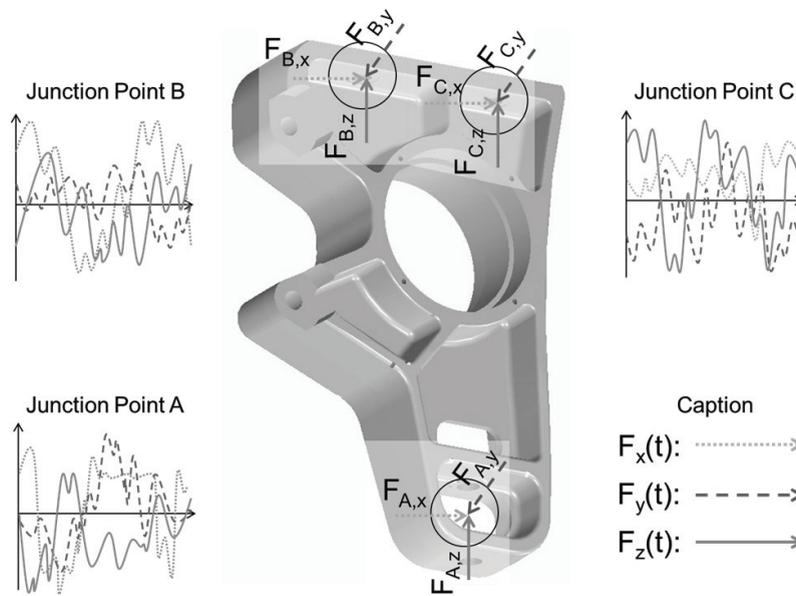


Fig. 2. Forces at the three junction points of the wheel carrier.

The challenge is to identify the develop-relevant load cases. Develop-relevant load cases are divided into three aspects: The first one involves the significant load cases. This means that the most frequently occurring couple of forces has to be identified. The second aspect represents the maximum load cases which accord to the highest expired load case. The third one constitutes critical load cases. These aspects always refer to the shape of the component. The third aspect can be neglected, because the actual geometry is not critical claimed by the present loads. To analyze the simulated life cycle data methods of data mining have to be applied.

Among the methods of patterns recognition on the data set it should be noted different techniques like classification, clustering or association [13]. Usually data mining approaches propose the following tasks [14]:

- Classification *(Classifying a data set into several predefined categories)*
- Regression *(Mapping a data set to real-valued prediction)*
- Clustering *(Assignment of a dataset to clusters; Cluster are grouped by similarity matrix or probability density models)*
- Rule generation *(Deriving classification rules)*
- Discovering association rules *(Describing the association between different attributes)*
- Summarization *(Initialization a compact description for a subset of data)*
- Dependency modeling *(Describing significant dependencies between variables)*
- Sequence analysis *(Developing a sequence pattern; Objective is the extraction and displaying variations and trends)*

The objective of the applied algorithm is to localize the significant and maximum load cases for the wheel carrier expired during a simulated race. This information is important for an optimal adaption of the design of the next wheel carrier generation.

3.1. Selection of Natural Based Algorithm

The technical inheritance prefers the adaption of natural principles in the development process. Based on this fact four possible algorithms were identified:

- Genetic algorithm
- Particle swarm optimization
- Ant colony optimization
- Simulated annealing

The benefit of the genetic algorithms is the robust and efficient calculating with complex and large data sets. The type of the genetic algorithms is inspired from the biological approaches of genetics and evolution [15]. The algorithms of the particle swarm optimization apply the behavior of biological swarm intelligence. At the beginning of the analysis the swarm is distributed over the entire search space. By a rapid concentration of the swarm the global optimum could be identified [16]. The algorithms of the ant colony optimization are deduced from the foraging of ants. There agents were used to identify the shortest way to the aliment. The advantages of these algorithms are the easily application and the possibility of parallel calculation [17]. The simulated annealing is similar to the physical behavior of cooling materials. In contrast of the genetic algorithm the simulated annealing just use physical processes instead of evolutionary mechanism, like mutation [18].

The four natural based algorithms were compared in different characteristics regarding the required identification of significant load cases and reduction of data volume. An excerpt of the reference analysis is depicted in Table 1.

Table 1. Analysis of natural based algorithms.

Characteristic	GA	PSO	ACO	SA
Optimization of multiple properties	+	+	-	-
Runtime	-	-	-	-
Solutions easily changed	+	-	-	0
Center-based Clustering	+	+	-	-
Broad-based Clustering	+	-	-	-

Legend: + = Suitable; 0 = Neutral; - = Inapplicable

GA = Genetic Algorithm; PSO = Particle Swarm Optimization; ACO = Ant Colony Optimization; SA = Simulated Annealing

The analysis of this characteristics revealed that genetic algorithms are the most suitable method for the planned data mining. The advantages of the genetic algorithms are in addition to center-based clustering that a broad-based clustering can be applied. So the global optimum of the most frequently load cases can be identified as well as the maximum expired load cases.

3.2. Integration of the Genetic Algorithm

For computational analysis of the applied loads the force per time of every dimension x,y,z for every application point is measured by a intelligent component or determined from a multi-body simulation. Instancing of three application points at the wheel carrier, there are 9 arrays of force values per time. For detection of significant load cases all 9 dimensions have to be analyzed in combination. The 9 dimensional force spaces are divided in a number of 9 dimensional force ranges. For evaluation the force frequency of every range is calculated by counting the force points of the data files according to the range. The higher counted frequency of load cases in the different force ranges, the more significant it is. The force frequency is dependent of the width of the 9 dimensional force ranges, therefore the edges of the range have to be varied to get an optimized configuration of width values. Optimally there exist a lot of ranges, which exhibit a less force frequency and a small number of such ranges with a high force frequency.

The calculation of the optimal range width configuration is realized by a genetic algorithm. The positions of range edges are the design parameters of this optimization problem. The operation mode of the implemented genetic algorithm is schematically illustrated by an activity diagram of the Unified Modeling Language (UML), depicted in Fig. 3.

There are some activities of algorithm initialization including the calculation of the force range for a given data set as well as the number of load ranges. For the genetic algorithm the size of the population and also some evolutionary parameters instancing the rate of mutation are defined. After starting the iteration loop the range edges of every individual of population are randomly varied in a small range by the mutation operator. The absolute width of every force range causes the normalization and requires the recalculation of the frequency. To evaluate the results of the variation parameters the objective function is implemented.

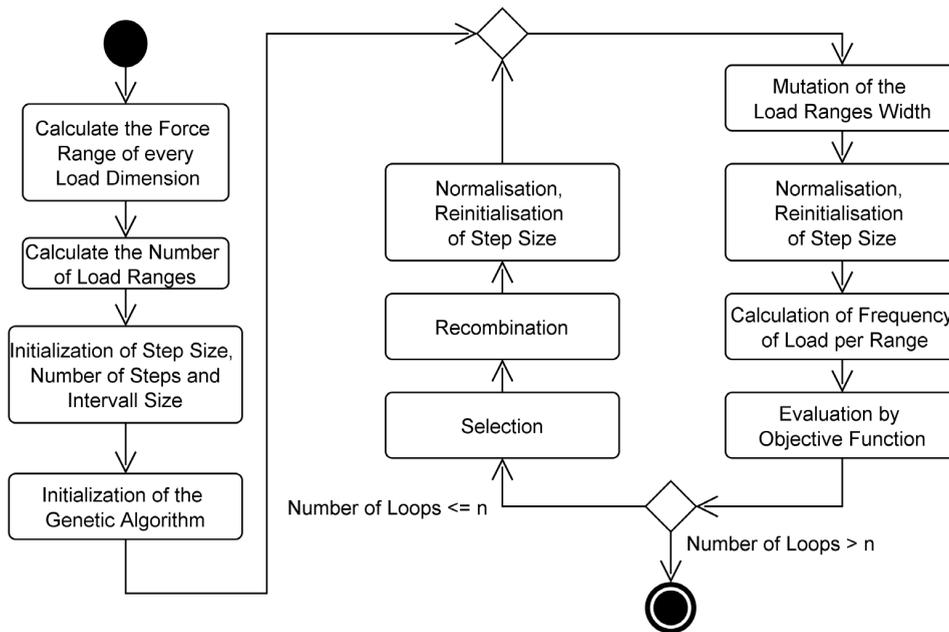


Fig. 3. Activity diagram of applied genetic algorithm.

This function includes the three following criteria, calculated for every range of every individual:

- If the frequency of a range n in the evolution step $i \leq$ the frequency of the according range n of the evolution step $i-1$ and the product of the width of all 9 dimensions of a range n in the evolution step $i \geq$ the product of the width of the according range n of the evolution step $i-1$, then the range n is evaluated with +1, else 0.
- If the frequency of a range n in the evolution step $i \geq$ the frequency of the according range n of the evolution step $i-1$ and the product of the width of all 9 dimensions of a range n in the evolution step $i \leq$ the product of the width of the according range n of the evolution step $i-1$, then the range n is evaluated with +1, else 0.
- If the number of ranges in the individual of population in the evolution step $i \geq$ the number of ranges in the individual in the evolution step $i-1$, then the individual is evaluated with +5 additionally.

By summation the values for every range the evaluation of every individual is completed. If the number of defined loops is reached or a criteria of convergence is satisfied, the algorithm finishes. Otherwise the calculation goes on with the selection operation. According to the size of population the best 25% of individual as well as the weak 25% of individual are selected for further calculations. To enlarge the population to the original size a recombination operation is implemented. The rejected individuals are replaced by a randomly recombination of range widths of the well evaluated individuals. After the normalization and the re-initialization of the step size the loop repeats itself. As a result of optimization the best individual is selected to get the configuration of optimal range widths.

4. Results

The implemented genetic algorithm is applied for the data analysis and identified the different load cases. Therefore the 9 arrays of forces were analysed which the wheel carrier expired during a simulated race. The load cases, structured according to their frequency, are depicted in Fig. 4.

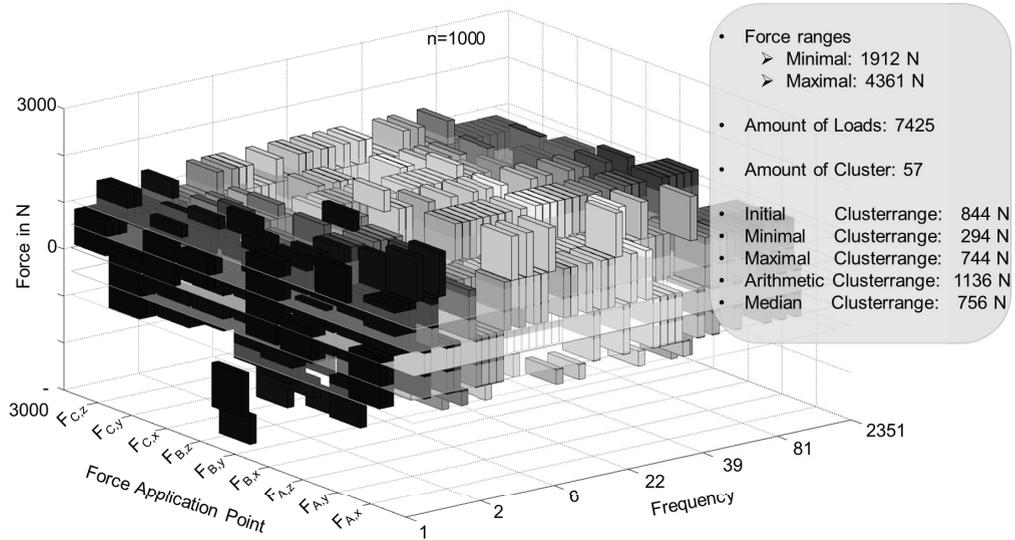


Fig. 4. Frequency of different load cases after 1000 loops.

After calculating 1000 loops with the genetic algorithm 57 different cluster are determined, in which all 7425 detected load cases can be classified. The ranges of the cluster width are located between 744 N and 294 N. The reduction of the clusterrange width indicates that the starting value for the width was selected too high. The maximum absolute force of the 9 force arrays constituted 4361 N. The minimum force of a cluster is amount of 1912 N. Also the analysis demonstrates that the frequency distribution increases progressively and not linear.

The objective of the simulated data analysis is divided into two aspects. The first examination investigates the maximal load cases, the second shows the significant ones. To demonstrate the difference between the variations of the range width, the results are presented for the configuration of 100 and 1000 loops of iterations. For each examination the five meaningful clusters are illustrated.

The both diagrams with the five largest absolute force clusters are depicted in Fig. 5. First of all, the difference between the location as well as the range width are demonstrated. Moreover the absolute values of forces are increased from the 100 to the 1000 loops of iteration by the genetic algorithm. Another point of view represents the change of the maximum and minimum width of the clusterrange. The minimal value of the clusterrange gets smaller while the maximum value gets higher. A further effect which is observed represents the approximation of the arithmetic and median mean. This fact demonstrates that the width of the clusterrange gained a better distribution after 1000 loops of iteration.

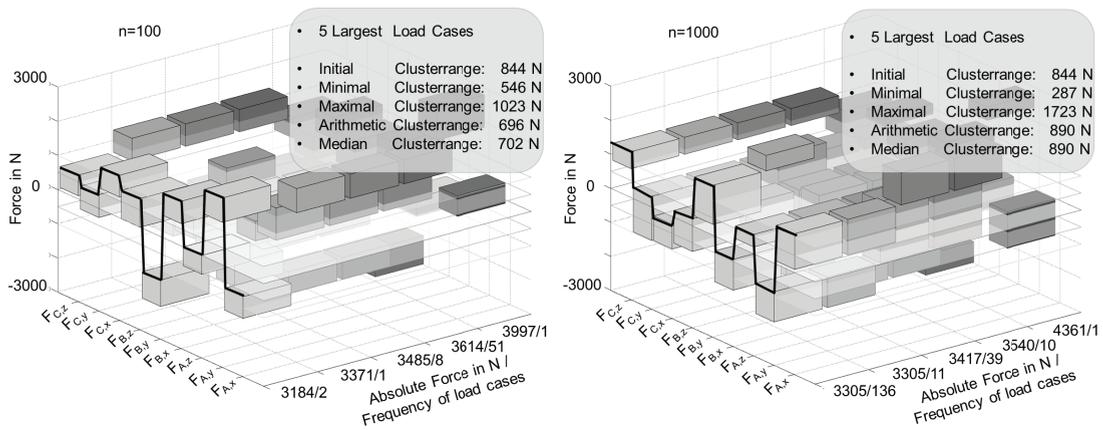


Fig. 5. Maximum load cases after (a) 100 loops and (b) after 1000 loops.

The analysis demonstrates that five maximum load cases are identified. Moreover 1000 loops generate a distributed value of clusters as well as defining the five largest load cases which the wheel carrier expired during the simulated race.

The second aspect of data analysis by the genetic algorithm embraced the significant load cases. Equal to the maximal load cases investigations after 100 and 1000 loops are executed. The most important result represents the increasing clusterrange for the five significant clusters from 100 to 1000 loops. It follows that the value of the frequency in the cluster is accumulated. The diagrams of the five significant cluster are depicted in Fig. 6.

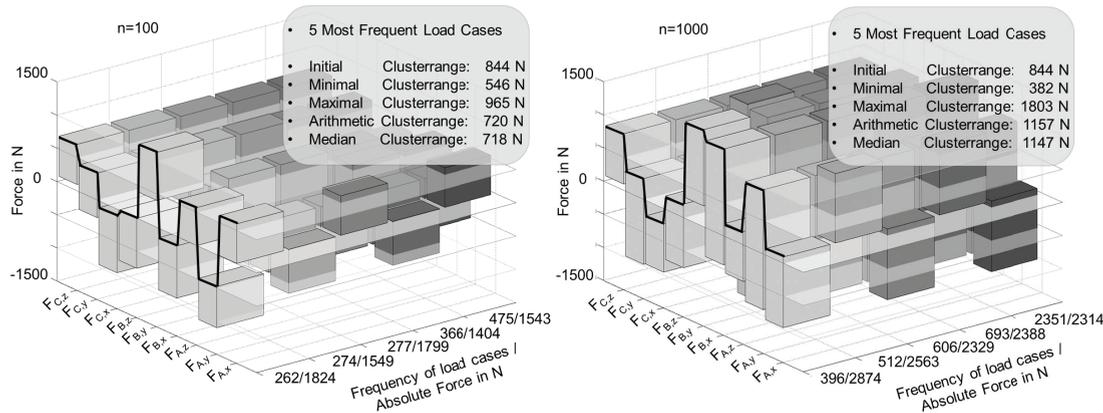


Fig. 6. Significant load cases after (a) 10 loops and (b) after 100 loops.

Another point of view demonstrates the displacement of the clusters width range. Another attitude is represented by the approximation of single forces like $F_{B,x}$ to the same largest value in each five clusters.

Overall the examination of the loads expired during a simulated race are performed by the implemented genetic algorithm. The significant and the maximal loads are identified. This information is retraced back in the steps of product development to adapt the design of the wheel carrier.

5. Conclusion

The analysis demonstrates how to investigate load cases in life cycle data by the transfer of biology principles into technical approaches. This transmission shows the relevance to support the intergenerational development process of the technical inheritance. For this propose product life cycle information is passed to the next generation. Different algorithms, based on natural effects, were analyzed in regard to their suitable application. The requirements of the algorithm analysis consist of an optimal characterization of different load cases as well as the identification of individual develop-relevant values of forces in the data set. The evaluation of four algorithms offered that genetic algorithms represent the pursue solution.

After the implementation the genetic algorithm to analyse the loads which a wheel carrier expired during a simulated race the significant and maximum loads are identified. The investigations have shown that with a bigger count of loops a better distribution of load cases in the cluster can be achieved. The classification of the significant and maximum load cases reduces the huge data set of the simulated race to 10 values. These contain 9 forces for each array at the wheel carrier junction points for every load case. As a result, with an applied genetic algorithm, the significant load cases are identified as well as the reduction of the data volume is accomplished.

The development process of the next generation is attributed with this information. By using the product life cycle information in the design optimization the adaption of the wheel carrier can be realized. By the combination of all three working packages the design evolution can be implemented and the technical inheritance is generated.

The further investigations contain the identification of the critical load cases which always is set in relation to the components design. The critical load cases have to be identified for the actual design. Afterwards their influence to the adaption by the design optimization has to be analysed. Another research aspect is the analysis of bigger data sets. Moreover the evaluation of the objective function and the resulting counts of the loops can be optimized. In Addition to the results the challenges of overfitting should be elaborated.

Acknowledgements

The results presented in this paper were obtained within the Collaborative Research Centre 653 "Gentelligent Components in their Lifecycle", Subproject N4. The authors would like to thank the German Research Foundation (DFG) for its financial and organizational support.

References

- [1] Abramovici M, Stark R (Editors). Smart Product Engineering, Proceedings of the 23rd CIRP Design Conference; Bochum; Germany; 2013.
- [2] Düser T. X-in-the-Loop – ein durchgängiges Validerungsframework für die Fahrzeugentwicklung am Beispiel von Antriebsstrangfunktionen und Fahrerassistenzsystemen. Forschungsberichte IPEK; Dissertation; Band 47; 2010.
- [3] Neubach M. Wissensbasierte Rückführung von Produktnutzungsinformationen in die Produktentwicklung im Rahmen einer Product Lifecycle Management (PLM)-Lösung. Dissertation; Maschinenbauinformatik, Bochum: Shaker Verlag; 2010.
- [4] Denkena B, Henning H, Lorenzen L-E. Genetics and intelligence: new approaches in production engineering, Production Engineering. Volume 4; Number 1; 2010. p. 65-73
- [5] Lachmayer R, Sauthoff B, Gottwald, P. Product Optimization by Analysis of Gentelligent Life Cycle Information. In Proceedings of SysInt; 1st Joint International Symposium on System-Integrated Intelligence 2012: New Challenges for Product and Production Engineering; 2012. p. 28-31
- [6] Parvan M, Miedl F, Lindemann U. Nature-inspired Process Model for Concept Selection and Evaluation in Engineering Design. NordDesign; Aalborg; Denmark; 2012. p. 486-493
- [7] Clement S. Erweiterung und Verifikation der Autogenetischen Konstruktionstheorie mit Hilfe einer evolutionsbasierten und systematisch-opportunistischen Vorgehensweise. Dissertation; Otto-von-Guericke-Universität; Magdeburg; 2005.
- [8] Lachmayer R, Sauthoff B, Gottwald P. Technical Evolution Process - An Approach for Product Development and Optimization. Norddesign 2012 Conference; Aalborg; Denmark; 2012. p. 342-349
- [9] Lachmayer R, Mozgova I, Sauthoff B, Gottwald P. Product Evolution and Optimization based on Gentelligent Components and Product Life Cycle Data. In Proceedings of CIRP Design 2013; Bochum; 2013. p. 687-694

- [10] Lachmayer R, Mozgova I, Deiters A, Sauthoff B, Gottwald P. Design Evolution by Integration of Life Cycle Information in Knowledge Repositories. *Consilience and Innovation in Design*; 5th International Congress of International Association of Societies of Design Research; Tokyo; Japan; Vol. 2; 2013. p. 309-318
- [11] Sauthoff B, Lachmayer R. Optimierung von Bauteilen auf Basis von Produktnutzungsinformationen. Beitrag zum 24.DFX Symposium; Jesteburg; 2013. p. 119-130
- [12] Denkena B, Mörke T, Krüger M, Schmidt J, Boujnah H, Meyer J, Gottwald P, Spitschan B, Winkens M. Development and first Applications of Gentelligent Components over their Life-Cycle. *CIRP Journal of Manufacturing Science and Technology*; 2014.
- [13] Sumathi S, Sivanandam S N. Introduction to data mining and its applications. Berlin, Heidelberg: Springer-Verlag; 2006.
- [14] Ghosh A, Dehuri S, Ghosh S. Multi-objective evolutionary algorithms for knowledge discovery from databases. Berlin, Heidelberg: Springer-Verlag; 2008.
- [15] Maulik U, Bandyopadhyay S, Mukhopadhyay A. Multiobjective Genetic Algorithm for Clustering, Application in Data Mining and Bioinformatics. Berlin, Heidelberg: Springer-Verlag; 2011.
- [16] Vajna S, Weber C, Bley H, Zeman K. *Cax für Ingenieure, eine praxisbezogene Einführung*. 2. völlig neu bearbeitete Auflage; Berlin, Heidelberg: Springer-Verlag; 2009.
- [17] Hamada M. Ants-Like Agents: A Model and Analysis Based on Natural Ants Behavior. Third KES International Symposium; KES-AMSTA 2009; Uppsala; Sweden; Berlin, Heidelberg: Springer-Verlag; 2009. p. 20-29
- [18] Sivanandam S N, Deepa S N. Introduction to Genetic Algorithms. Berlin, Heidelberg: Springer-Verlag; 2008.