

Research Article

Nonparametric Bayesian Filtering for Location Estimation, Position Tracking, and Global Localization of Mobile Terminals in Outdoor Wireless Environments

Mohamed Khalaf-Allah

*Institute of Communications Engineering, Faculty of Electrical Engineering and Information Technology,
Leibniz University of Hannover, Appelstrasse 9A, 30167 Hannover, Germany*

Correspondence should be addressed to Mohamed Khalaf-Allah, mohamed.khalaf-allah@ikt.uni-hannover.de

Received 28 February 2007; Revised 16 August 2007; Accepted 10 November 2007

Recommended by Richard J. Barton

The mobile terminal positioning problem is categorized into three different types according to the availability of (1) initial accurate location information and (2) motion measurement data. *Location estimation* refers to the mobile positioning problem when both the initial location and motion measurement data are not available. If both are available, the positioning problem is referred to as *position tracking*. When only motion measurements are available, the problem is known as *global localization*. These positioning problems were solved within the Bayesian filtering framework. Filter derivation and implementation algorithms are provided with emphasis on the mapping approach. The radio maps of the experimental area have been created by a 3D deterministic radio propagation tool with a grid resolution of 5 m. Real-world experimentation was conducted in a GSM network deployed in a semiurban environment in order to investigate the performance of the different positioning algorithms.

Copyright © 2008 Mohamed Khalaf-Allah. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Mobile terminal (MT) positioning is a key problem in wireless environments. It is the most fundamental problem to provide customers with tailored and location-aware services. MT positioning is defined as the determination of the MT geolocation using location-dependent parameters in a specific coordinate system. The key driver for developing MT location technologies in the USA was E-911. In the EU, it was commercial services in the first place, and later E-112 that utilizes the same techniques. Emergency call location has become a requirement in fixed and cellular networks in the USA in 1996 [1] and in the EU in 2003 [2]. Positioning of an MT is considered more critical because MT users and hence MT originated emergency calls are continually increasing. It is estimated that about 50% of all emergency calls in the EU are MT originated, and the expected tendency is rising [2].

The first application of MT location dates back to World War II, when it was critical to locate military personnel rapidly and precisely in emergency situations [3]. Furthermore, nonmilitary interest in this field dates back to about 40

years ago [4, 5]. While emergency call location could be considered the most important of location-based services (LBSs) due to its urgency for life and property safety, commercial LBSs are believed to make increasing revenues for network operators who could provide customers with attractive and tailored services [6]. Therefore, a lot of research is being carried out in this area.

Positioning systems are usually categorized according to the place where location calculations are performed into *network-based* or *mobile-based*, or according to the application environment into *outdoor* or *indoor*. The main approaches of positioning are *global or satellite-based techniques*, and *local or terrestrial-based methods*. Terrestrial-based methods have two variants: *geometric techniques*, and *mapping approaches*. These methods differ in terms of accuracy, coverage, cost, mobile terminal power consumption, and wireless system impact.

Satellite-based technologies are mainly employed for outdoor applications and come in two flavours: stand-alone GPS or assisted-GPS (A-GPS). The first is mobile-based, while A-GPS needs extra signals from reference GPS receivers and

thus increasing the system impact. The main drawbacks are high-power consumption, need of clear view to at least four satellites (for stand-alone GPS), and the costs of integrating GPS receivers into the MTs. Furthermore, A-GPS solutions increase overhead costs due to the requirement to install reference GPS receivers. The satellite-based approach is the most accurate MT positioning technique, and it was only made accessible for commercial applications in the nineties. Also the EU is most likely to follow the US and Japan in requiring high-positioning accuracy of mobile emergency calls from 2010 when the Galileo system will be fully operational [7]. However, the benefits of satellite-based positioning could be limited where location information is still needed due to signal blocking. In such cases, other positioning methods should be triggered in order to backup the failed or degraded satellite signals.

Geometric methods estimate the MT location by triangulation of, for example, time-of-arrival (TOA), time-difference-of-arrival (TDOA), enhanced-observed time-difference (EOTD), angle-of-arrival (AOA) measurements, or relationship between received signal strength attenuation and distance to base stations (RSSAD). The main drawback of TOA measurements is the need of mutual synchronization of the involved base stations (BSs) in order to avoid degraded location accuracy, which is difficult to achieve. Exploiting AOA measurements increases overhead costs due to the need for installation of special antennas at the BSs. At least three BSs are required for TDOA measurements, which cannot always be fulfilled in many situations. RSSAD equations are not really accurate even when using at least three BSs. Although geometric techniques are generally more accurate than mapping methods, their position estimation accuracy degrades severely in multipath environments, which is the dominant condition in built-up areas, and in nonline-of-site (NLOS) situations without accurate environmental information.

Mapping-based mobile location is one way to achieve accuracy improvement of cell-ID positioning. They also appear in the literature under the names *database comparison* or *correlation*, *location fingerprinting*, and *pattern recognition* or *matching*. In these techniques, a database, or map of location-dependent parameters, is constructed using radio wave propagation prediction tools [8–10], field measurements [11, 12], or a combination of both [13]. Later a moving MT collects measurements to be compared with the entries of the database in order to yield location estimates. Propagation prediction tools are advantageous in terms of cost and map construction time. These tools vary in terms of accuracy according to the degree of geographical information precision integrated in the calculations, thus are divided into deterministic (3D), semi-deterministic (2–2.5D), or simple empirical formulas. Field measurements provide more realistic databases but at higher costs and longer construction time that render wide deployment impractical. Nevertheless field measurements in some parts of the deployment environment do help to show the performance upper limit of location estimation algorithms using the mapping approach. Location-dependent parameters usually used for mapping include received signal strength

TABLE 1: Phase II of the FCC's E911 program requirement on location accuracy.

	Network-based	Mobile-based
67%	100 m	50 m
95%	300 m	150 m

levels (RxLevs) from surrounding BSs [8–11, 13] and the channel impulse response (CIR) [12, 14, 15] which is the multipath propagation delay profile of the environment. In GSM systems, the bandwidth is too small, unlike the UMTS system, for accurate positioning based on correlation of CIR only [12]. Also the geometric time-based (TOA, TDOA, EOTD) and angle-based (AOA) methods could be used as location signatures either stand-alone (less accurate) or combined with other location parameters. To the best of the author's knowledge they are not widely used. However, in [16] a network-based fingerprint method composed of TOA and AOA has been proposed for wireless location finding in urban environments, and was found that AOA is more significant than TOA for location discrimination.

Mapping methods often utilize prediction data of RxLev and/or CIR produced during network planning. In the online positioning phase they use only the network available measurements and thus they do not require any expensive hardware installations at BSs or in MTs. Also they have short deployment time and cover current and legacy handsets. This is advantageous in terms of cost, coverage, and system impact compared to the other approaches. Therefore, they seem to be the first alternative to take into consideration, especially for European network operators, since EU mobile location requirement still does not specify any accuracy levels unlike the US mandate, see Table 1. However, mapping-based solutions require continuous update in order to adapt to changes in the environment structure and in the wireless network infrastructure, and to consider the time-varying nature of wireless channels.

The location accuracy of mapping approaches ranges between about 100 m and several kilometers depending on cell size, accuracy of reference maps, mapping resolution, propagation conditions, accuracy of observed measurements, and significance degree of the mapped location-dependent parameter. While CIR maps generally achieve more accurate estimates than RxLev mapping in urban and dense urban environments, they tend to have comparable performance in suburban and rural areas. Therefore, mapping techniques do not fulfil the FCC accuracy requirements in all situations. However, mapping methods are advantageous, because no LOS conditions are needed, it can work even with one BS, and its implementation costs are pretty low. Moreover, mapping techniques will still be needed also when more accurate technologies are fully available. They will achieve positioning for applications with low accuracy requirements; they will be deployed in areas of the network where more accurate methods are not supported; and finally, they will work as backup in case the accurate techniques fail for any reason. Therefore, improving positioning accuracy of mapping approaches is an

TABLE 2: Basic aspects of the different positioning techniques.

	Accuracy	Coverage		Cost	Terminal power consumption	Wireless system impact
		Outdoor	Indoor			
Global or satellite-based methods	High (~15 m)	Yes	None or very poor	Medium	High	Low or medium
Terrestrial geometric techniques	Medium (~100 m)	Yes	Yes	Medium	Low	Medium
Terrestrial mapping approaches	Low (100 m-several km's)	Yes	Yes	Low	Low	Low

active research topic. A comparison of the basic aspects of the discussed positioning approaches is given in Table 2.

In this paper, a mapping-based method for outdoor wireless mobile positioning using the Bayesian filtering formulation is proposed. Prediction of the average received signal strength at reference locations in a working GSM network is calculated using a 3D radio propagation tool. The motion model of the Bayesian filter utilizes simulated inertial measurements. Real-world experiments in a semiurban area have been carried out to study the performance of the proposed techniques.

The rest of the paper is organized as follows. Section 2 defines three different positioning problems within the context of the mapping approach. Section 3 discusses the basics of Bayesian filtering, introduces world model utilized, and gives implementable algorithms for the different positioning problems. Experiments and numerical results are presented in Section 4. Finally, the paper is concluded in Section 5.

2. TYPES OF MOBILE TERMINAL POSITIONING PROBLEMS USING THE MAPPING APPROACH

Estimation of the MT position in its environment involves using a map of a location-dependent parameter of the environment, network measurement data, and motion information. The estimation accuracy could even be enhanced by utilizing any prior knowledge of the MT location when available.

Motion information is generally the most difficult piece of information to extract. Without dedicated motion sensors, for example, an inertial measurement unit (IMU), motion estimation is either impossible or very inaccurate due to the noisy signal behavior used to derive the MT motion pattern. Accordingly, the MT positioning problem could be divided into *location estimation* and *tracking* based on the availability of motion measurements. Location estimation (LE) algorithms calculate the MT location without incorporating any motion information. Moreover, tracking algorithms could be further categorized according to the availability of prior knowledge into *position tracking* and *global localization*. In position tracking (PT), the initial position of the MT is known, and the problem is to find adequate procedures in order to compensate incremental errors in the motion sensor measurements. In the more challenging global lo-

TABLE 3: Comparison of the three positioning problems.

	Prior knowledge available?	Motion information available?
Location estimation	No	No
Position tracking	Yes	Yes
Global localization	No	Yes

calization (GL) problem, the initial location of the MT is unknown, and consequently the MT position has to be determined from scratch. This positioning problem is more difficult because multiple and distinct hypotheses have to be handled. The three defined positioning problems are summarized in Table 3.

3. BAYESIAN FILTERING FOR MOBILE TERMINAL POSITIONING

3.1. Foundations of the Bayesian filter and basic algorithm

The recursive Bayesian filter (RBF) [17] is a probabilistic framework for state estimation that utilizes the *Markov assumption* (i.e., past and future measurements are conditionally independent if the current state is known). The RBF estimates the posterior belief of the MT position given its prior belief, motion and network measurements, and the model of the world (or environment).

The prior belief is a probability distribution over all possible locations before taking the MT actions and network measurements into account. The posterior belief is the conditional distribution of these locations after incorporating the MT actions and network measurements. The world model is a radio profile map containing predicted received signal strength (RxLev) values at reference locations. The posterior belief distribution is expressed as

$$\text{Bel}(s_t) = p(s_t | o_{0:t}, a_{0:t}, m), \quad (1)$$

where $\text{Bel}(s_t)$ is the posterior belief over the state (or position) of the MT at time t , s_t is the state at time t , $o_{0:t}$ and

$a_{0:t}$ are the network measurement data (or network observations) and the actions performed by the MT from time 0 up to time t , respectively, and m is the world model.

Applying *Bayes rule* to (1) we get

$$\text{Bel}(s_t) = \frac{p(o_t | s_t, o_{0:t-1}, a_{0:t}, m) p(s_t | o_{0:t-1}, a_{0:t}, m)}{p(o_t | o_{0:t-1}, a_{0:t}, m)}. \quad (2)$$

Here, actions and network measurements are assumed to occur in an alternative sequence (every action is followed by a network measurement) although in reality they take place concurrently. They are separated only for convenience and clarity of the mathematical treatment.

Employing *Markov assumption* to the first term in the nominator, and noting that the denominator is a constant probability (denoted η) relative to s_t , (2) is rewritten as

$$\text{Bel}(s_t) = \eta p(o_t | s_t, m) p(s_t | o_{0:t-1}, a_{0:t}, m). \quad (3)$$

With the help of η , which is also called *normalization factor*, the resulting product will always sum up to 1. Thus $\text{Bel}(s_t)$ represents a valid probability distribution.

Expanding the right most term in (3) using the *theorem of total probability* will result in

$$\begin{aligned} \text{Bel}(s_t) &= \eta p(o_t | s_t, m) \\ &\times \int p(s_t | s_{t-1}, o_{0:t-1}, a_{0:t}, m) p(s_{t-1} | o_{0:t-1}, a_{0:t}, m) ds_{t-1}. \end{aligned} \quad (4)$$

Applying *Markov assumption* to the first term in the integration and noting that the second term is simply $\text{Bel}(s_{t-1})$, we obtain

$$\text{Bel}(s_t) = \eta p(o_t | s_t, m) \int p(s_t | s_{t-1}, a_t, m) \text{Bel}(s_{t-1}) ds_{t-1}. \quad (5)$$

Equation (5) is called the *recursive Bayesian filter (RBF)* and is usually computed in two steps called *prediction* and *update*.

Prediction step:

$$\text{Bel}^-(s_t) = \int p(s_t | s_{t-1}, a_t, m) \text{Bel}(s_{t-1}) ds_{t-1}, \quad (6)$$

where $\text{Bel}^-(s_t)$ is the posterior belief just after executing the action a_t and before incorporating the network measurement o_t , and $p(s_t | s_{t-1}, a_t, m)$ is the MT motion model, that is, the transition probability that tells us how the state evolves over time as a function of the MT movements.

Update step:

$$\text{Bel}(s_t) = \eta p(o_t | s_t, m) \text{Bel}^-(s_t), \quad (7)$$

where $p(o_t | s_t, m)$ is the network measurement model that specifies the probabilistic law according to which these measurements are generated from the state, that is, measurements are simply noisy projections of the state [17].

```
(1) Algorithm Basic_RBF ( $\text{Bel}(s_{t-1}), a_{t-1}, o_t, m$ )
(2) for all  $s_t$  do
(3)    $\text{Bel}^-(s_t) = \int p(s_t | s_{t-1}, a_{t-1}, m) \text{Bel}(s_{t-1}) ds_{t-1}$  // Prediction
(4)    $\text{Bel}(s_t) = \eta p(o_t | s_t, m) \text{Bel}^-(s_t)$  // Update
(5) endfor
(6) return( $\text{Bel}(s_t)$ )
```

ALGORITHM 1: The basic recursive Bayesian filter algorithm.

Both motion and network measurement models describe the dynamical stochastic system of the MT and its environment. The state at time t is stochastically dependent on the state at time $t - 1$ and the action a_t . The network measurement o_t depends stochastically on the state at time t . Such a temporal model is also known as *hidden Markov model (HMM)* or *dynamic Bayes network (DBN)* [17]. Algorithm 1 shows a single iteration of the RBF algorithm.

Nonparametric filters (NPFs) [17] provide implementable algorithms for the RBF. They approximate posteriors by a finite number of parameters, each corresponding to a region in the state space, that is, they do not rely on a fixed functional form of the posterior. Moreover, the number of the parameters used to approximate the posterior can be varied. The quality of approximation depends on the number of these parameters. As the number of parameters approaches infinity, NPF tends to converge uniformly to the correct posterior. The NPF approach discussed here approximates posteriors over finite spaces by decomposing the state space into finitely many regions and represents the cumulative posterior for each region by a single probability value. Such an approach is known as *discrete Bayesian filter (DBF)*. The DBF is also referred to as the *forward pass* of a *hidden Markov model*.

The DBF approximates the belief $\text{Bel}(s)$ at any time by a set of n weighted location candidates as

$$\text{Bel}(s) \approx \{s^{(i)}, w^{(i)}\}_{i=1:n}, \quad (8)$$

where $s^{(i)} = \{x^{(i)}, y^{(i)}\}$ is the i th MT location candidate (or state) and $w^{(i)}$ is a probability value (also called *weight*) that determines the importance of $s^{(i)}$. The sum of all weights equals 1 so that $\text{Bel}(s)$ represents a valid probability distribution. However, normalization is not a crucial issue for practical algorithm implementation.

3.2. World model

The utilized world model has been constructed by using two input sources. The first are maps of the predicted average received signal strength in a test semiurban area of 9 km² in Hannover, Germany, created by a 3D deterministic radio propagation tool [18]. These maps are represented by 2D raster arrays with a uniform grid spacing of 5 m. Each array corresponds to a GSM cell antenna working at 1800 MHz. The experimental area contains 6 BSs, each with 3 sectors, and 4 indoor antennas, so that the total number of considered cells equals 22. Figure 1 illustrates the geometry of the

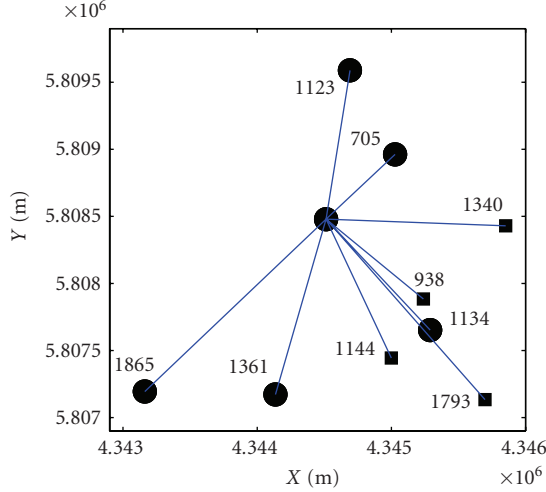


FIGURE 1: Geometry of the base stations in the experimental area. Base stations and indoor antennas are represented by solid circles and squares, respectively.

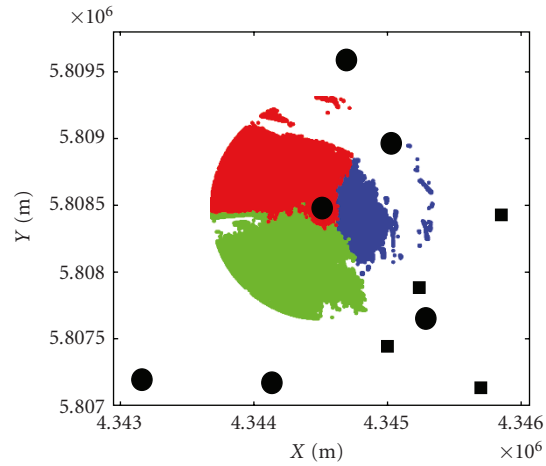


FIGURE 2: Locations served by sector cell antennas up to distances corresponding to TA = 0.

involved BSs and distances from the area centric BS to the rest.

After several preprocessing steps as in [19, 20], the maps are rearranged so that each raster array contains only the reference locations served by a certain cell antenna. Moreover, each raster array is further divided into smaller arrays according to timing advance (TA) values; see Figure 2. This is very useful for the reduction of computational costs. Each array element contains x - y coordinates and average predicted RxLev of all involved BSs.

The second input was geographical information system (GIS) data to assist in discriminating between the different environmental features, for example, indoor, outdoor, water, green, and so forth, with very high resolution of 30 cm. Before the arrays that resulted from the preprocessing steps were further divided according to the land feature, which is also very helpful for the computational efficiency of the pro-

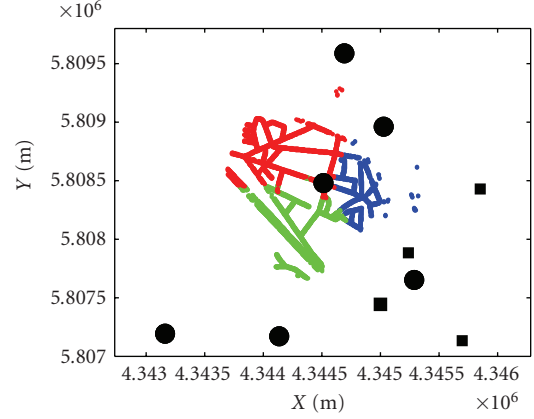


FIGURE 3: Outdoor pedestrian locations served by sector cell antennas up to distances corresponding to TA = 0.

```

(1) Algorithm LocationEstimation( $o_t, m_t$ )
(2)  $\text{Bel}(s_t) = 0, \hat{s}_t = 0, m_t = \text{DB}_{\text{cell-ID}}$ 
(3)  $o_t = \{\text{cell-ID}_t, \text{TA}_t, \text{RxLev}_t^{(j)}\}$ 
(4) for  $i = 1 : n$  do
(5) Compute the weight  $w_t^{(i)}$ 
(6)  $\text{Bel}(s_t) = \text{Bel}(s_t) \cup \{s_t^{(i)}, w_t^{(i)}\}$ 
(7) endfor
(8)  $\text{Bel}(s_t) = \text{sort}(\text{Bel}(s_t))$  // Descending sort
(9) Calculate  $\hat{s}_t$ 
(10) return( $\hat{s}_t$ )
    
```

ALGORITHM 2: The location estimation algorithm.

posed algorithms, the GIS data resolution was adapted to the 5 m resolution of the radio propagation prediction maps. Figure 3 shows outdoor pedestrian locations served by their main sector cell antennas for TA = 0. Arrays as depicted in Figure 3 were the ones used in the three positioning algorithms.

Furthermore, the raster arrays have been re-sampled to 10 m, 15 m, ..., 50 m resolutions for use only with the location estimation algorithm.

3.3. Location estimation

As mentioned in Section 2 the location estimation algorithm calculates the MT position without any prior information about the accurate initial location of the MT or any motion measurements from dedicated sensors. Thus line 3 in Algorithm 1 could not be executed. Consequently, the algorithm computes only the output probability of the network measurements, which is merely a table-lookup procedure.

Algorithm 2 depicts a single iteration of the location estimation algorithm to estimate the MT state at time t . It is initialized (in line 2) by allocating memory space for the location belief $\text{Bel}(s_t)$ and the final MT location estimate \hat{s}_t . The inputs (lines 2 and 3) are the network measurements o_t and the world model m_t , where $\text{DB}_{\text{cell-ID}}$ is the database that

contains location information and expected RxLev values (of the main and neighboring cell antennas) of the areas covered by the main (or serving) cell antenna (or BS) at time t , and $\text{RxLev}_t^{(j)}$ is the measured received signal strength from the j th observed BS. The weight of the location candidate i is calculated (in line 5) as

$$w^{(i)} = w_{\text{MM}}^{(i)} + w_{\text{ND}}^{(i)} + w_{\text{SN}}^{(i)}, \quad (9)$$

where $w_{\text{MM}}^{(i)}$, $w_{\text{ND}}^{(i)}$, and $w_{\text{SN}}^{(i)}$ are the weights according to the *measurement model*, *neighborhood degree*, and *strongest neighbor*, respectively. They are calculated as

$$\begin{aligned} w_{\text{MM}}^{(i)} &= p(o_t | s_t^{(i)}, m) \\ &= \prod_{j=1}^M \frac{1}{\sigma_{\text{RxLev}} \sqrt{2\pi}} e^{-\frac{(\text{RxLev}_t^{(j)} - \text{RxLev}_{\text{DB}_j})^2}{2\sigma_{\text{RxLev}}^2}}, \end{aligned} \quad (10)$$

where M is the number of observed BSs (main and neighboring), that is, $M_{\text{max}} = 7$ in typical GSM network measurements, σ_{RxLev} is the standard deviation of the measured RxLev, and $\text{RxLev}_{\text{DB}_j}$ is the database RxLev prediction value of the j th observed BS at $s_t^{(i)}$:

$$w_{\text{ND}}^{(i)} = l, \quad (11)$$

where l is the number of observed neighbor BSs that coincide with the list of the predicted six strongest neighbor BSs at $s_t^{(i)}$, that is, $l_{\text{max}} = 6$:

$$w_{\text{SN}}^{(i)} = \alpha_{\text{SN}}, \quad (12)$$

where α_{SN} is a constant bonus value and equals 1. It is assigned if the strongest observed neighbor BS coincides with the predicted *first* or *second* strongest neighbor BS at $s_t^{(i)}$. Otherwise, $w_{\text{SN}}^{(i)} = 0$.

Intuitively, the summation in (9) should be multiplication. However, summation has two advantages over multiplication. First, summation will prevent the assignment of zero to the total weight of any location candidate in case a weighting criterion, for example, $w_{\text{SN}}^{(i)}$, equals zero. Second, multiplication cause many candidates to have very low weights, which will be considered as zero weights if the computer that runs the algorithm has limited numerical precision. Zero weights can cause many problems especially when sorting location candidates according to their weights. The correct order of candidates cannot be determined.

After weight calculation, the location candidate is added to the belief (line 6) together with the assigned weight. This is done for all location candidates before sorting them (line 8) in a descending order with respect to their weights. The aim is not just to find the belief distribution of the MT state, but an estimate of the state called *point estimate*. This point estimate is simply the final MT location estimate that is output by the algorithm (line 10). There are several ways to calculate point estimates (line 9), for example, *maximum a posteriori (MAP)*, *weighted average estimate (WAE)*, and *trimmed average estimate (TAE)*.

```

(1) Algorithm PositionTracking( $s_{t-1}, a_{t-1}, o_t, m_t$ )
(2)  $s_{t-1} = (x_{t-1}, y_{t-1})$  // Input
(3)  $a_{t-1} = (\text{trans}_{t-1}, \theta_{t-1})$  // ...
(4)  $o_t = \{\text{cell-ID}_t, \text{TA}_t\}$  // ...
(5)  $m_t = \text{DB}_{\text{cell-ID}_t} = \langle x_j, y_j, w_j \rangle$  // ...
(6)  $x_t^- = x_{t-1} + \text{trans}_{t-1} \cdot \cos \theta_{t-1}$  // Prediction
(7)  $y_t^- = y_{t-1} + \text{trans}_{t-1} \cdot \sin \theta_{t-1}$  // ...
(8) for  $j = 1 : n$  do // Update
(9)  $w_j = 1/\sqrt{(x_t^- - x_j)^2 + (y_t^- - y_j)^2}$ 
(10) endfor
(11)  $m_t = \text{sort}(m_t)$  // Descending sort
(12)  $s_t = (x_t, y_t) = (x_1, y_1)$ 
(13) return( $s_t$ )

```

ALGORITHM 3: The position tracking algorithm.

Maximum a posteriori is simply the location candidate with the highest assigned weight and is expressed as

$$\hat{s}_t = \arg \max \text{Bel}(s_t). \quad (13)$$

If many candidates have the same weight, the returned location estimate will depend on the stability of the sorting scheme. Stable sorting algorithms maintain the relative order of the location candidates, that is, a location candidate with the highest weight that appeared first in the unsorted belief will also appear first in the sorted belief. This is very disadvantageous as an arbitrary candidate could be returned as the location estimate though other candidates also assigned with the same highest weight would be more accurate. However, this negative aspect could be reduced by computing the *weighted average* of all candidates representing the posterior belief distribution. Thus the location estimate would be

$$\hat{s}_t = \frac{1}{\sum_{i=1}^n w^{(i)}} \sum_{i=1}^n s_t^{(i)} \times w^{(i)}. \quad (14)$$

The WAE is the *mean value* of the updated belief distribution and it will coincide with the MAP estimate only for *unimodal* and *symmetric* distributions, which is not often the case. The *trimmed average estimate* calculates the MT location as the average of the k best weighted candidates as follows:

$$\hat{s}_t = \frac{1}{k} \sum_{i=1}^k s_t^{(i)}, \quad (15)$$

where $k < n$ and n is the total number of location candidates.

3.4. Position tracking

A single iteration of the position tracking algorithm is given in Algorithm 3. The inputs are the initial position (line 2) $s_{t-1} = (x_{t-1}, y_{t-1})$, the IMU data (line 3) $a_{t-1} = (\text{trans}_{t-1}, \theta_{t-1})$, where trans_{t-1} and θ_{t-1} are the translation (after twice integration of the IMU acceleration measurement) and orientation (IMU compass) in a 2D Cartesian coordinate system at time $t - 1$, respectively, the network measurement o_t (line 4), and the corresponding world map m_t

(line 5), where w_j is the weight of the j th location candidate and initially set to zero. Note that the proposed algorithm updates only one position hypothesis, that is, n in expression (8) equals 1.

The position tracking algorithm propagates the known initial MT location s_{t-1} using IMU data in the prediction step (lines 6 and 7). The propagated location is then updated by matching it to the set of candidate locations (lines 8–10) that are covered by the current serving cell antenna, after descending sort of the candidates with respect to weight (line 11), the new MT position (line 12) is simply the candidate of the minimum Euclidean distance to the location computed in the prediction step.

3.5. Global localization

The global localization algorithm has no information about the accurate MT position at the beginning. Thus, it has to resolve the location ambiguity and converge to the true position of the MT by tracking all probable location candidates and determine their weights every time the algorithm is run. When this task is successfully fulfilled, the algorithm is allowed to run in the position tracking mode (line 30 in Algorithm 4).

As depicted in Algorithm 4, the global localization algorithm is initialized by setting the travelled distance as measured by the IMU ($trvld_dist$) to 0, and $Mode$ also to 0, that is, global localization mode (line 3). The inputs (lines 4–7) are the same as in Algorithm 3 except (line 5) that the global localization algorithm tracks a number of hypothetical candidates, unlike the position tracking algorithm. The global localization mode will run as long as the number of location candidates n in the belief distribution $Bel(s_{t-1})$ is greater than a certain threshold α (line 9). During this mode, the prediction and update steps will only run if the MT's travelled distance is greater than or equal to the database (or map) resolution DB_{res} (line 11), in order to allow position state transition using the world model. The updated candidate will only be added to the new belief, if the location it is matched to is not greater than DB_{res} away (lines 19–21). Therefore, the number of location candidates will decrease after every run of the algorithm until their total number is equal to or less than the threshold α . In this very event, the updated MT position is simply estimated as the average of the remaining candidates, and the algorithm is switched to the position tracking mode (lines 25–28). Note that the algorithm returns no position estimates in the global localization mode. First after switching to the position tracking mode, location estimates are returned at the end of every update run, see Algorithm 3. For both global localization and position tracking algorithms only the cell-ID and TA but no RxLev values of the network measurement report have been utilized, see line 4 in Algorithm 3 and line 7 in Algorithm 4, respectively.

The update step of the position tracking and global localization algorithms has different roles. In the position tracking algorithm, the position estimate is decided upon the result of the update step, where in the global localization algorithm, the update step works to reduce the size of the position belief

```

1: Algorithm GlobalLocalization( $Bel(s_{t-1}), a_{t-1}, o_t, m_t$ )
2: // Initialization, only at the first run of the algorithm
3:  $trvld\_dist = 0, Mode = 0$ 
4: // Inputs
5:  $Bel(s_{t-1}) = DB_{cell-ID_t} = \langle x_i, y_i \rangle, i = 1, \dots, n$ 
6:  $m_t = DB_{cell-ID_t} = \langle x_j, y_j, w_j \rangle, j = 1, \dots, q, \langle w_j \rangle = 0$ 
7:  $o_t = \{cell-ID_t, TA_t\}, a_{t-1} = (trans_{t-1}, \theta_{t-1})$ 
8: if  $Mode = 0$  // Global localization mode
9:   if  $n > \alpha$ 
10:     $trvld\_dist = trvld\_dist$ 
11:     $+ \sqrt{(trans_{t-1} \cdot \cos \theta_{t-1})^2 + (trans_{t-1} \cdot \sin \theta_{t-1})^2}$ 
12:    if  $trvld\_dist \geq DB_{res}$ 
13:      for  $i = 1 : n$  do
14:         $x_i^- = x_i + trvld\_dist \cdot \cos \theta_{t-1}$  // Prediction
15:         $y_i^- = y_i + trvld\_dist \cdot \sin \theta_{t-1}$  // ...
16:        for  $j = 1 : q$  do
17:           $w_j = 1/\sqrt{(x_i^- - x_j)^2 + (y_i^- - y_j)^2}$  // Update
18:        endfor
19:         $\langle w_j \rangle = \text{sort}(\langle w_j \rangle)$  // Descending sort
20:        if  $(1/w_1 \leq DB_{res})$ 
21:          add  $(x_1, y_1)$  to  $Bel(s_t)$ 
22:        endif
23:      endfor
24:     $trvld\_dist = 0$ 
25:    endif
26:  else if  $n \leq \alpha$ 
27:     $Mode = 1$ 
28:     $s_t = (\sum_i x_i/n, \sum_i y_i/n)$ 
29:  endif
30: else if  $Mode = 1$  // Position tracking mode
31:   PositionTracking ( $s_{t-1}, a_{t-1}, o_t, m_t$ ) // Algorithm 3
32: endif

```

ALGORITHM 4: The global localization algorithm.

and makes it converge to a single estimate before allowing the position tracking algorithm to run.

3.6. How global localization works

Solving the global localization problem for an MT in a GSM network is described and illustrated in Figure 4. Location state space, MT location belief, ground truth, and position estimation (when available) are depicted in green, red, solid blue diamond, and black, respectively. At start, the MT location is not known and the algorithm has to handle all probable locations. Therefore, the location belief covers the whole state space, see Figure 4(a). After approximately 27 m of motion, many location candidates have been found improbable and thus have fallen out of consideration, as in Figure 4(b). After another 38 m of movement, the location belief has concentrated on two parallel streets, see Figure 4(c). As the MT moved further, the location belief has almost converged to the true position as in Figure 4(d). Figure 4(e) shows how the MT location ambiguity has been resolved after a total movement of about 145 m with a position estimation error of approximately 16 m.

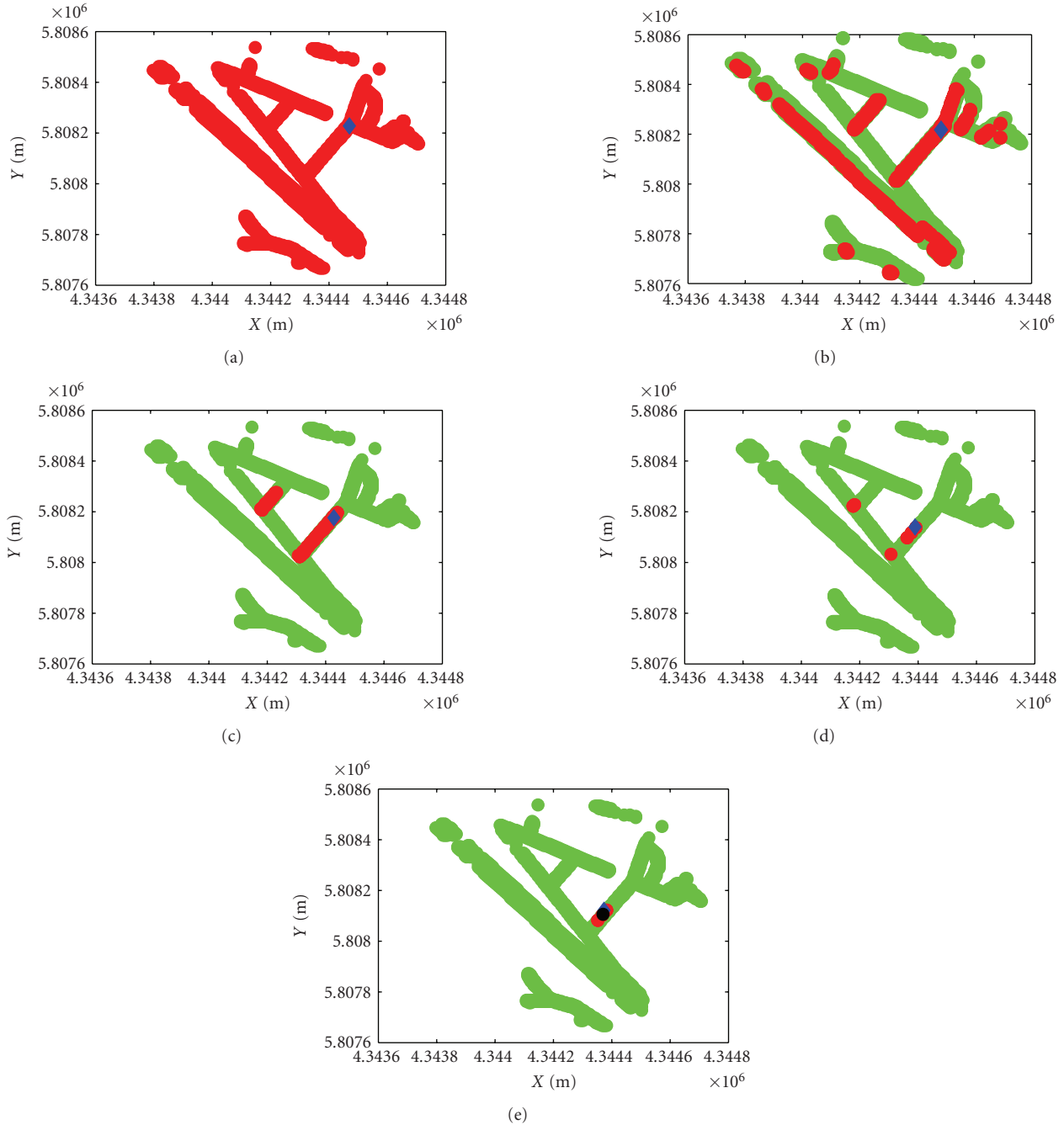


FIGURE 4: Global localization of a mobile terminal in a GSM environment.

4. EXPERIMENTS AND NUMERICAL RESULTS

4.1. Experimental setup

Measurements have been carried out in an E-Plus GSM 1800 MHz network by a pedestrian along a route of about 1940 m long in a 9 km² semiurban environment in Hannover, Germany. There are six BSs, each with three sectors, and four indoor antennas in the test area. RxLev measurements of the serving BSs and up to six neighboring stations along with GPS position fixes for ground truth have been

logged into a file for later offline evaluation. Furthermore, the GPS positions have been used to generate IMU pseudomeasurements to simulate real ones in order to investigate the feasibility of real IMU employment. Experimental results are based on a single network measurement report (NMR) at 172 data points made during active calls. Each NMR contains cell-IDs and signal strength levels of the serving BS antenna and up to 6 neighbor BS antennas, and TA of the serving BS. Signal strength levels from the serving BS recorded during active calls are those of the traffic channel which undergoes power management. However, the position tracking and

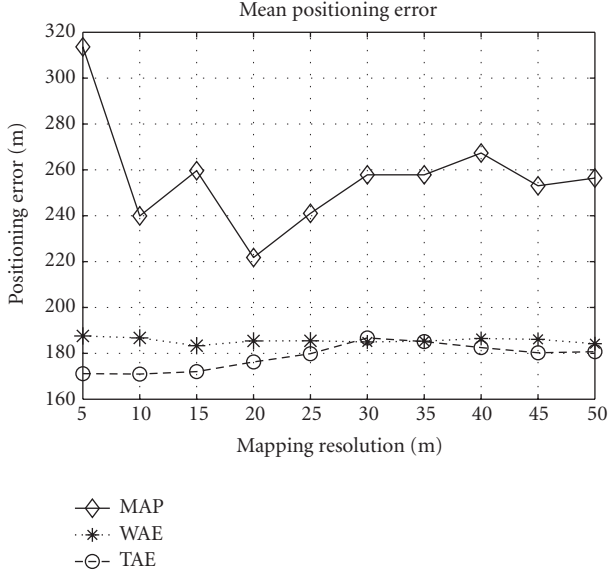


FIGURE 5: Mean positioning error of the location estimation algorithm.

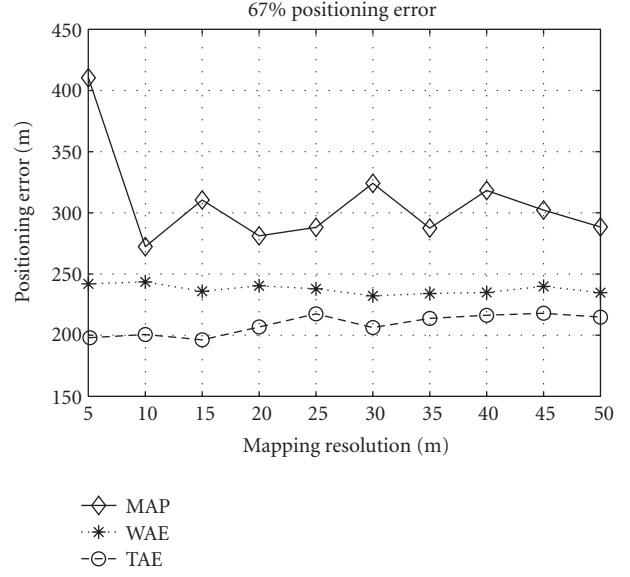


FIGURE 6: Sixty-seven percentile positioning error of the location estimation algorithm.

global localization algorithms depend only on the TA measurements that correspond to the serving BS wireless coverage, which can be sufficiently determined offline, taking account of power management effects. Thus, both algorithms are not affected by power management operations. For the location estimation algorithm, the network operator would need to keep prediction information for all possible range of the power management scheme in order to avoid the decrease in accuracy performance.

4.2. Location estimation results

The positioning accuracy of the location estimation algorithm has been investigated for the three presented point estimators and using different mapping resolutions. Figures 5–7 show the mean, 67 percentile and 95 percentile positioning error, respectively, of the different point estimators with varying world model resolution.

It can be seen that WAE and TAE always outperform the MAP estimator. This is logical as both WAE and TAE consider more location candidates of the posterior belief and not only one candidate as the MAP estimator. Because in the context of mobile terminal positioning using RxLev mapping, multimodal posterior belief distributions are generated; MAP estimation will choose only one peak of the posteriors which is not a suitable estimation decision. On the contrary, WAE and TAE consider more than the one peak and thus can better represent the multimodal property of the posterior distributions.

Figure 6 also shows that TAE outperforms WAE at the 67 percentile positioning error for all mapping resolution. This might be due to the fact that WAE represents the whole posterior belief distribution, while TAE considers only the upper areas of the posteriors, that is, location candi-

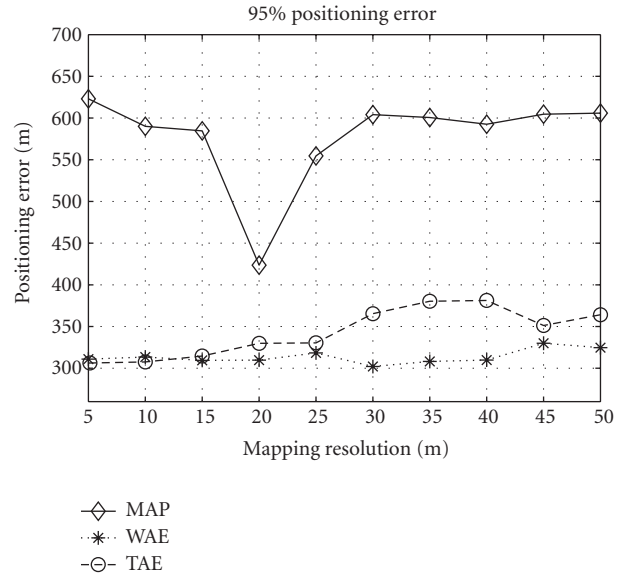


FIGURE 7: Ninety-five percentile positioning error of the location estimation algorithm.

dates of higher weight. In Figure 5 we can see that the TAE mean positioning error outperforms that of WAE only up to the resolution of 25 m. For the 30 m and 35 m resolutions both TAE and WAE perform almost the same. Starting from the 40 m resolution, the TAE further slightly outperforms the WAE. However, this does not indicate the superiority of TAE for all cases. In Figure 7, at the 95 percentile positioning error, the TAE is slightly better than the WAE up to the 10 m resolution. From the 15 m resolution, the WAE starts to perform obviously better than the TAE.

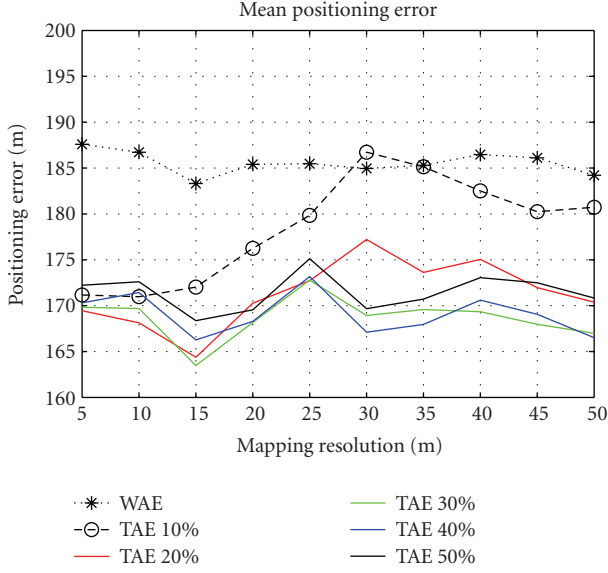


FIGURE 8: Mean positioning error of the location estimation algorithm using WAE and TAE ($k = 0.1*n - 0.5*n$).

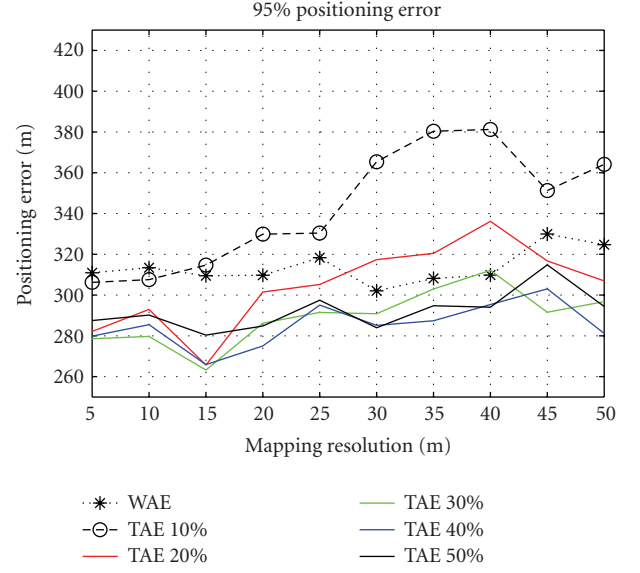


FIGURE 10: Ninety-five percentile positioning error of the location estimation algorithm using WAE and TAE ($k = 0.1*n - 0.5*n$).

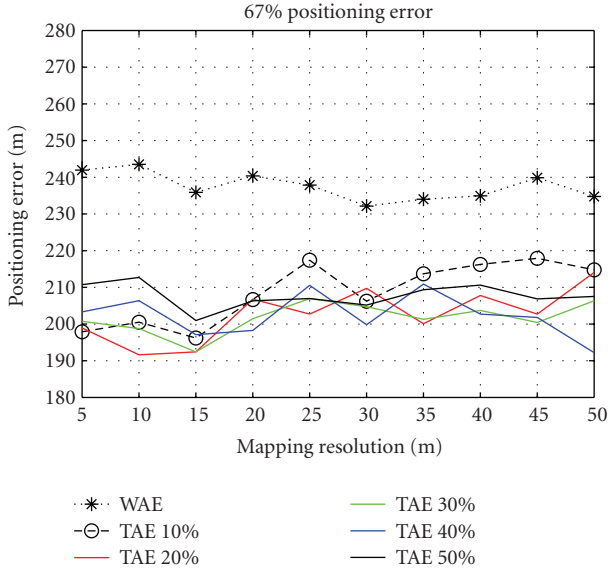


FIGURE 9: Sixty-seven percentile positioning error of the location estimation algorithm using WAE and TAE ($k = 0.1*n - 0.5*n$).

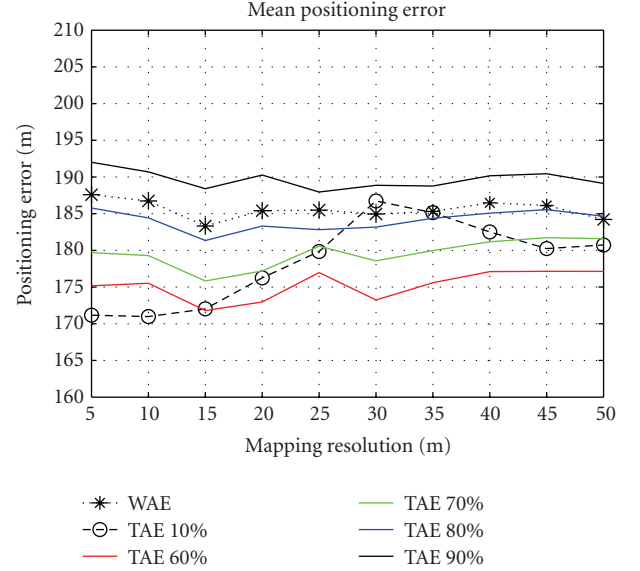


FIGURE 11: Mean positioning error of the location estimation algorithm using WAE and TAE ($k = 0.6*n - 0.9*n$).

The explanation is that for lower mapping resolution, considering only upper areas of the posterior belief distributions to calculate a point estimate, as the TAE, will not correctly keep the information represented by the posterior distributions, and thus considering the whole distribution area, as the WAE, is more representative.

In Figures 5, 6, and 7, TAE was calculated by averaging the best 10% weighted location candidates, that is, $k = 0.1*n$ in (15). The explanation in the previous paragraph can be confirmed if we look at the results obtained when k is increased to $0.9*n$.

Figures 8 and 9 show that increasing the number of location candidates to average ($k = 0.2*n - 0.5*n$) for TAE with decreasing mapping resolution enhances the performance of TAE at the mean and 67 percentile errors and always outperforms the WAE. We can notice the same tendency in Figure 10. However, k had to be over $0.2*n$ in order to outperform the WAE at the 95 percentile positioning error with decreasing mapping resolution.

In Figure 11 we can see that for lower resolutions, increasing k over $0.7*n$ does not enhance the TAE mean positioning error anymore. TAE will even perform worse than

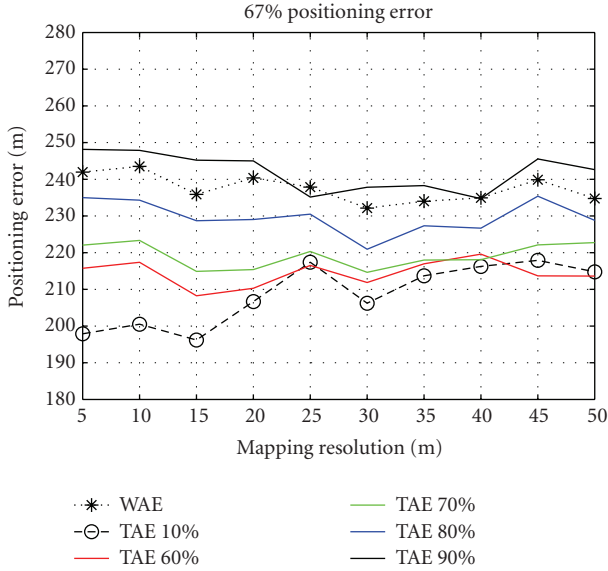


FIGURE 12: Sixty-seven percentile positioning error of the location estimation algorithm using WAE and TAE ($k = 0.6 * n - 0.9 * n$).

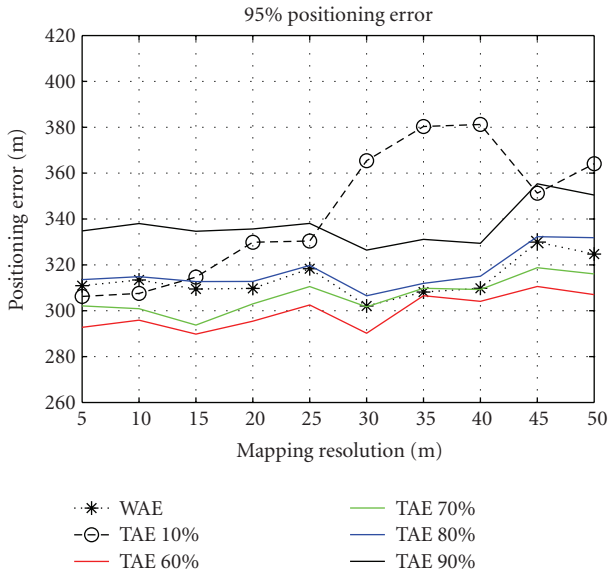


FIGURE 13: Ninety-five percentile positioning error of the location estimation algorithm using WAE and TAE ($k = 0.6 * n - 0.9 * n$).

WAE for k over $0.8 * n$. Also at 67 percentile positioning error in Figure 12 no TAE enhancement was achieved by increasing k . However, at the 95 percentile in Figure 13 TAE performed better till k reached $0.7 * n$.

From the previous discussion we can conclude that TAE performs better with lower resolution mapping, that is, up to 15 m, when k is increased up to $0.5 * n$.

We can also generally notice that for all point estimation methods, there is no significant decrease in the positioning accuracy with decreasing mapping resolution. This shall be further investigated by comparing these results with a lower bound, for example, the Barankin bound [21] in a future

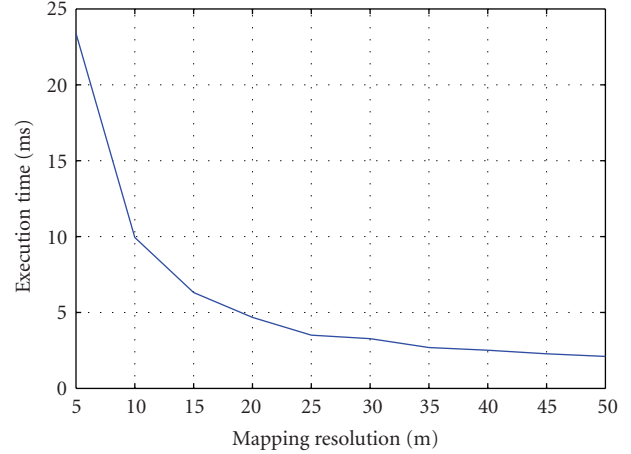


FIGURE 14: The average execution time needed for a single iteration of the location estimation algorithm using different mapping resolutions on a standard PC with 2.2 GHz processor.

work. However, according to this notice it was of interest to calculate how the run time of the location estimation algorithm changes with varying mapping resolution. Figure 14 depicts the average computation time needed for a single iteration on a standard PC with 2.2 GHz processor. At the 5 m resolution the execution time was only 23 milliseconds. Computation time then drops down exponentially to under 3 milliseconds as the mapping resolution decreases. However, execution time is linearly proportional to the number of location candidates.

These results can even suggest providing mobile-based implementation for the location estimation algorithm, which will supply customers with position information for low accuracy applications at very low costs. World models can initially be installed in the mobile terminals and updated as needed.

4.3. Position tracking results

Within position tracking experiments the initial location of the MT is known. We have investigated the performance of the tracking algorithm by varying σ_{trans} from 1% to 10% of the performed translation and σ_{orient} between 1° and 6° . The quality of performance is determined according to *reliability* and *positioning errors* in meters. We consider the MT position is reliably tracked if the final position estimate error over the whole experiment route of 1940 m is not greater than 50 m. All experiments have been repeated 100 times in order to get reasonable results. It can be seen in Figure 15, as expected, that the higher σ_{trans} and/or σ_{orient} are the lower the reliability of the tracking algorithm along the test route. However, for σ_{trans} up to 4% and σ_{orient} up to 2° , reliability is over 90% of all repeats. With σ_{orient} up to 2° and σ_{trans} up to 10%, slightly less than 70% of the cases are successfully tracked. When σ_{orient} is increased up to 5° , reliability is at least 60% of all repeats even with the worst case of σ_{trans} . For σ_{orient} equals 6° , the reliability drops below 60% as σ_{trans} is over 4%.

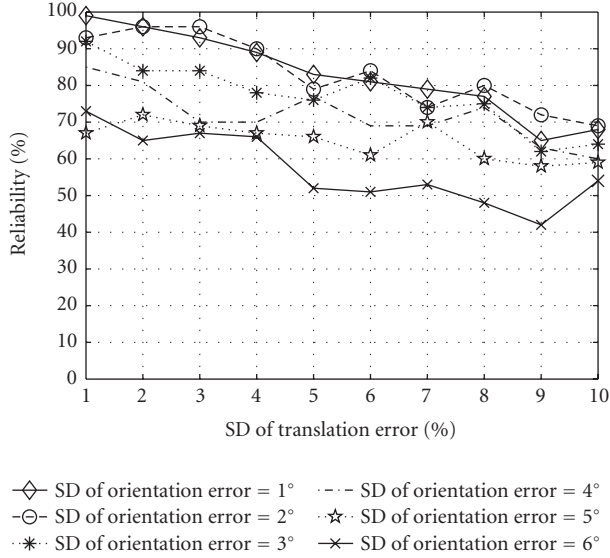


FIGURE 15: Reliability of position tracking with varying standard deviation (SD) of IMU translation and orientation.

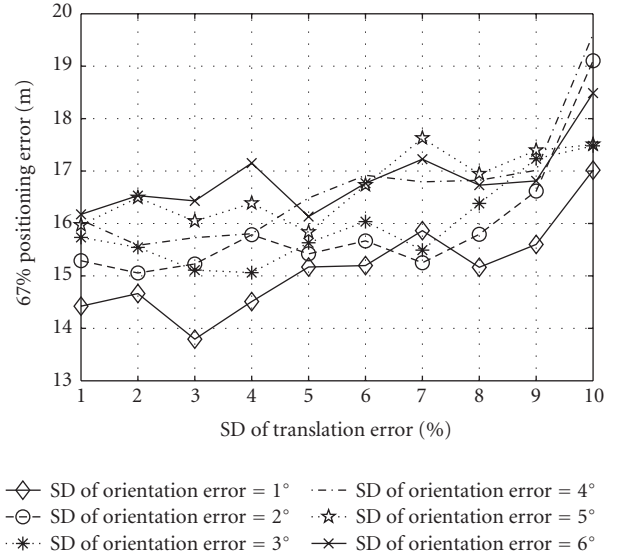


FIGURE 17: Sixty-seven percentile position tracking error.

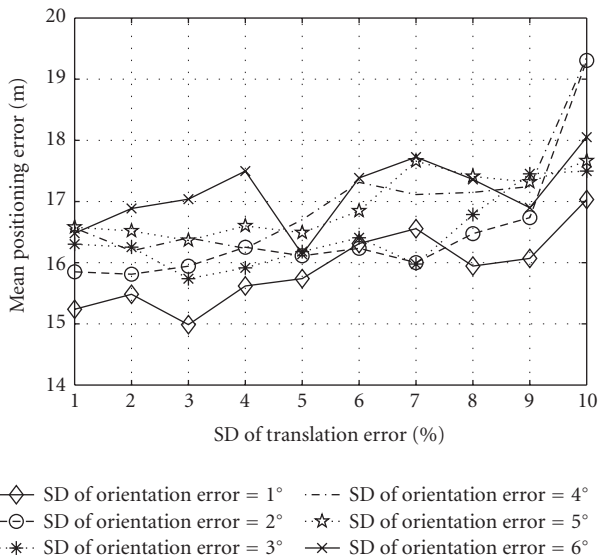


FIGURE 16: Mean position tracking error.

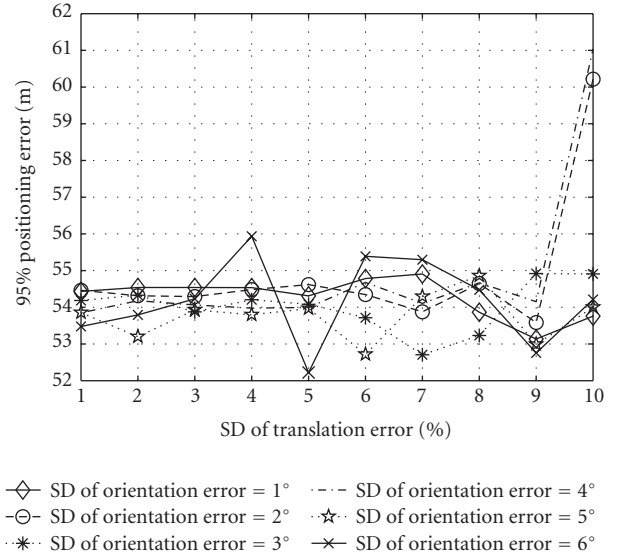


FIGURE 18: Ninety-five percentile position tracking errors.

Figure 16 shows that the mean positioning error for the different cases is between 15 m and 20 m. This is accurate enough for most positioning applications and confirms the suitability of IMU employment for reliable position tracking. The 67 percentile positioning error is always less than 20 m for all cases as illustrated in Figure 17. Figure 18 depicts the 95 percentile position tracking error which is almost always between 52 m and 56 m and less than 62 m in the worst cases.

4.4. Global localization results

In the global localization experiments, the reliability for the different values of σ_{trans} and σ_{orient} has been investigated. Global localization is considered reliable, that is, success-

ful if the MT position estimation error just before switching to position tracking mode (line 30 in Algorithm 4) is not greater than 50 m in order to also allow reliable position tracking. As shown in Figure 19, the global localization reliability is over 80% and 65% for σ_{orient} up to 3° and 6°, respectively.

The effect of σ_{trans} on the results is almost not significant, because of the 5 m map resolution that makes the update step insensitive to the range of translation errors assumed. Moreover, there is a slight tendency to increase the reliability of global localization with increasing σ_{trans} especially when σ_{orient} also increases, which seems counter intuitive. However, the fact is that large errors caused by high σ_{orient} values are compensated by increasing σ_{trans} and the low map resolution that prevents quick deviation from the true path.

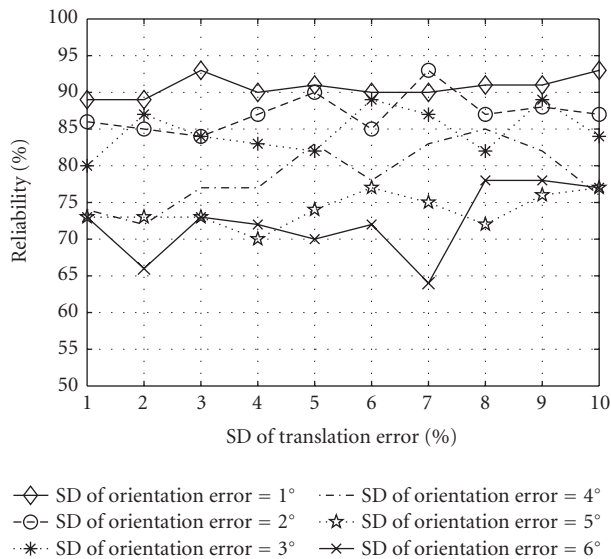


FIGURE 19: Reliability of global localization with varying standard deviation (SD) of IMU translation and orientation.

5. CONCLUSIONS AND FUTURE WORK

In this paper, the mobile terminal positioning problem was first classified into three types according to the availability of (1) prior knowledge about the accurate initial position of the MT and (2) motion measurement data. Solutions for the three positioning problems have been suggested within the Bayesian filtering framework. Also implementation algorithms have been provided and the world model has been described. Finally, experimental results in a live GSM network have been presented and discussed. The paper showed that reliable accurate position information can be obtained and maintained for mobile terminal users by combining environment radio maps with IMU data.

As mentioned in Section 4.2, calculation of a theoretical lower bound, for example, Barankin bound, on the location estimation performance will be a topic for future work. Also increasing the reliability of the position tracking and global localization algorithms is a possible extension of the work. This shall be achieved by further analysis of the behavior of the algorithms when they incorrectly estimate the MT position. Results should help develop mechanisms to recognize and handle such situations. It is apparent that the proposed algorithms could be applied in indoor environments as well as for vehicle navigation applications.

ACKNOWLEDGMENTS

This work has been conducted under the supervision of Professor Kyamakya. The help of many parties was essential. The author would like to thank the Institute for Communications Technology at Braunschweig Technical University and E-Plus Mobilfunk GmbH & CO KG in Düsseldorf, Germany, for providing the network and radio prediction data. The author is also grateful to the Institute of Cartography and Geoinfor-

matics at the Leibniz University of Hannover for supplying the GIS data utilized in this research.

REFERENCES

- [1] Federal Communications Commission (FCC) Fact Sheet, "FCC Wireless 911 Requirements," 2001.
- [2] EU Institutions Press Release, "Commission Pushes for Rapid Deployment of Location Enhanced 112 Emergency Services," DN: IP/03/1122, Brussels, Belgium, July 2003.
- [3] K. Pahlavan and A. H. Levesque, *Wireless Information Networks*, John Wiley & Sons, Hoboken, NJ, USA, 2nd edition, 2005.
- [4] W. G. Figel, N. H. Shepherd, and W. F. Trammell, "Vehicle location by a signal attenuation method," *IEEE Transactions on Vehicular Technology*, vol. 18, no. 3, pp. 105–109, 1969.
- [5] G. D. Ott, "Vehicle location in cellular mobile radio systems," *IEEE Transactions on Vehicular Technology*, vol. 26, no. 1, pp. 43–46, 1977.
- [6] T. M. Rantalainen, M. A. Spirito, and V. Ruutu, "Evolution of location services in GSM and UMTS networks," in *Proceedings of the 3rd International Symposium on Wireless Personal Multimedia Communications (WPMC '00)*, pp. 1027–1032, Bangkok, Thailand, November 2000.
- [7] Berg Insight, "GPS and Galileo in Mobile Handsets," Research Report, Berg Insight, Gothenburg, Stockholm, Sweden, November 2006.
- [8] H. Schmitz, M. Kuipers, K. Majewski, and P. Stadelmeier, "A new method for positioning of mobile users by comparing a time series of measured reception power levels with predictions," in *Proceedings of the 57th IEEE Semiannual Vehicular Technology Conference (VTC '03)*, vol. 3, pp. 1993–1997, Jeju, Korea, April 2003.
- [9] D. Zimmermann, J. Baumann, M. Layh, F. Landstorfer, R. Hoppe, and G. Wolffe, "Database correlation for positioning of mobile terminals in cellular networks using wave propagation models," in *Proceedings of the 60th IEEE Vehicular Technology Conference (VTC '04)*, vol. 7, pp. 4682–4686, Los Angeles, Calif, USA, September 2004.
- [10] M. Khalaf-Allah and K. Kyamakya, "Database correlation using bayes filter for mobile terminal localization in GSM suburban environments," in *Proceedings of the 63rd IEEE Vehicular Technology Conference (VTC '06)*, vol. 2, pp. 798–802, Melbourne, Australia, May 2006.
- [11] H. Laitinen, J. Lahtenmaki, and T. Nordstrom, "Database correlation method for GSM location," in *Proceedings of the 53rd IEEE Vehicular Technology Conference (VTC '01)*, vol. 4, pp. 2504–2508, Rhodes, Greece, May 2001.
- [12] T. Nypan, "Mobile terminal positioning based on database comparison and filtering," Dissertation, Norwegian University of Science and Technology, Trondheim, Norway, 2004.
- [13] J. Zhu, "Indoor/outdoor location of cellular handsets based on received signal strength," Technical Specification PG-TR-060515-JZ, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Ga, USA, August 2006.
- [14] M. Layh, U. Reiser, D. Zimmermann, and F. Landstorfer, "Positioning of mobile terminals based on feature extraction from channel impulse responses," in *Proceedings of the 63rd IEEE Vehicular Technology Conference (VTC '06)*, vol. 3, pp. 1078–1081, Melbourne, Australia, September 2006.

- [15] S. Ahonen and H. Laitinen, "Database correlation method for UMTS location," in *Proceedings of the 57th IEEE Vehicular Technology Conference (VTC '03)*, vol. 4, pp. 2696–2700, Jeju, Korea, April 2003.
- [16] I. Kelly, D. Hai, and L. Hao, "On the feasibility of the multipath fingerprint method for location finding in urban environments," *Applied Computational Electromagnetics Society Journal*, vol. 15, no. 3, p. 232, 2000.
- [17] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, MIT Press, Cambridge, Mass, USA, 2005.
- [18] T. Kurner and A. Meier, "Prediction of outdoor and outdoor-to-indoor coverage in urban areas at 1.8 GHz," *IEEE Journal on Selected Areas on Communications*, vol. 20, no. 3, pp. 496–506, 2002.
- [19] M. Khalaf-Allah and K. Kyamakya, "Mobile location in GSM networks using database correlation with bayesian estimation," in *Proceedings of the 11th IEEE International Symposium on Computers and Communications (ISCC '06)*, pp. 289–293, Cagliari, Italy, June 2006.
- [20] M. Khalaf-Allah and K. Kyamakya, "Bayesian mobile location in cellular networks," in *Proceedings of the 14th European Signal Processing Conference (EUSIPCO '06)*, Florence, Italy, September 2006.
- [21] H. Koorapaty, "Barankin bounds for position estimation using received signal strength measurements," in *Proceedings of the 59th IEEE Vehicular Technology Conference (VTC '04)*, vol. 5, pp. 2686–2690, Milan, Italy, May 2004.