Bayesian Filtering for Localization of Mobile Terminals

Mohamed Khalaf-Allah and Kyandoghere Kyamakya

***Abstract*— A discrete Bayesian filter (DBF) was developed to work as a pattern matching technique with database correlation for mobile location. The database is constructed using a 3D deterministic radio wave propagation prediction model. Three methods to yield final location estimates are presented and compared according to their accuracy. We carried out field measurements in a GSM network deployed in a suburban environment, which is very common and widespread in Europe. We show that more work has to be done in order to make positioning accuracy as good as in urban and dense urban areas.**

***Index Terms*— Discrete Bayesian filter (DBF), database correlation, mobile location.**

I. Introduction

A key problem in wireless networks is the positioning of mobile stations (MSs). In this problem, the position of a MS is determined by the utilization of location sensitive parameters. The FCC adopted standards for location accuracy and reliability of emergency calls [1] that further motivated the interest in the field of mobile location, which can be traced back to the 1970s [2]. It is believed that inexpensive but still accurate MS positioning systems are of great commercial importance. Therefore, a lot of research efforts are carried out in the area. Throughout the literature, the terms mobile location, positioning and localization are used interchangeably. An overview of localization methods are provided in [3].

MS positioning is usually performed using one of these methods: Time-of-arrival (TOA), Angle-of-arrival (AOA), Network-assisted GPS (A-GPS), Time-difference-of-arrival (TDOA) and enhanced observed time-difference-of-arrival (E-OTD), and Enhanced cell-id. The evaluation criteria for the different positioning methods include accuracy, cost, coverage, system impact, and power consumption.

TOA techniques need mutual synchronization of the base stations (BSs), which is difficult to achieve leading to poor location accuracy. AOA methods suffer from large positioning errors caused by multipath propagation. Moreover, special antennas have to be installed at BSs. The main shortcomings of A-GPS solutions are power consumption, the need to a clear view to at least four satellites, and the installation of additional hardware. In addition, multipath propagation degrades TOA estimation. TDOA based techniques need at least three BSs, which could not be fulfilled in rural areas.

Enhanced cell-id methods provide an attractive alternative as they utilize only network available information and do not require any additional hardware installations at BSs or in MSs. This is advantageous in terms of cost, coverage and system impact compared to other methods. However, the accuracy is ranging from about 100 m up to a couple of kilometers depending on type and characteristics of the area covered by the network.

One solution to improve positioning accuracy of enhanced cell-id techniques is the database correlation method (DCM). This method is also referred to as database comparison, location fingerprinting, pattern recognition and pattern matching. In such techniques, a database of location dependent parameters is constructed using field measurements [4], [7] or radio wave propagation prediction tools [5], [6]. Later a moving MS collects measurements to be compared with the entries in the database in order to yield position estimates.

Different location dependent parameters could be used with DCM. In [4]-[6], the received signal levels (RxLev) from surrounding BSs are used, whereas in [7], the channel impulse response (CIR) is utilized for this purpose. However, the bandwidth of GSM is too small for accurate positioning based on database comparison of the CIR only [7].

We present a DCM based on Bayesian filtering that works as a pattern matching technique, and three methods to yield location estimates. We applied our technique to measurements of the received signal strength and timing advance (TA) in a GSM network covering a suburban area. The DCM is introduced in section II. Section III presents the basics of Bayesian filtering. Experiments and results are given in section IV. Finally, the paper is concluded in section V.

II. Database Correlation Method

A. *Database Construction and Utilization*

A 3D deterministic radio propagation prediction model described in [8] was used to construct our database. Robust location estimation with DCM depends mainly on three factors. The first is to assure that searching in the database is

Mohamed Khalaf-Allah is with the Institute of Communications Engineering, University of Hannover, Appelstrasse 9A, D-302167 Hannover, Germany (phone: +49 511 762-2845; fax: +49 511 762-3030; e-mail: khalaf@ikt.uni-hannover.de).

Kyandoghere Kyamakya is with the Department of Informatics Systems, University of Klagenfurt, Austria (e-mail: kyamakya@isys.uni-klu.ac.at).

restricted to an area where the MS is really located. This could be achieved by utilizing TA (timing advance) and sector information of the serving cell. The second factor relies on the pattern matching technique used to evaluate location candidates. For this purpose we use Bayesian filtering, which is a robust technique in multi-hypotheses situations usually occur in the context of RxLev-based mobile location. The third is to decide how to yield a final location estimate from the available hypotheses. We present three methods in section III.B.

B. Database Preprocessing

The localization algorithm can take advantage if the pixels (locations) that are served by every BS antenna are determined. In this case, it is guaranteed that no position outside the coverage area of the BS antenna would be returned by the algorithm when the deviation between predicted and measured power levels are large or when the situation is highly ambiguous due to an increased number of probable location candidates. Thus, the prediction data is preprocessed so that for every BS antenna an array of the locations belonging to its coverage area is constructed. Every array entry consists of location coordinates, predicted received power level at that location, and distance to the serving BS. Note that the midpoint of every pixel is considered as its coordinates. Fig. 1 shows an example of the preprocessing step results. Here, the serving BS has three sector antennas. The locations covered by each antenna are depicted in different colors. The black dot represents the location of the BS.

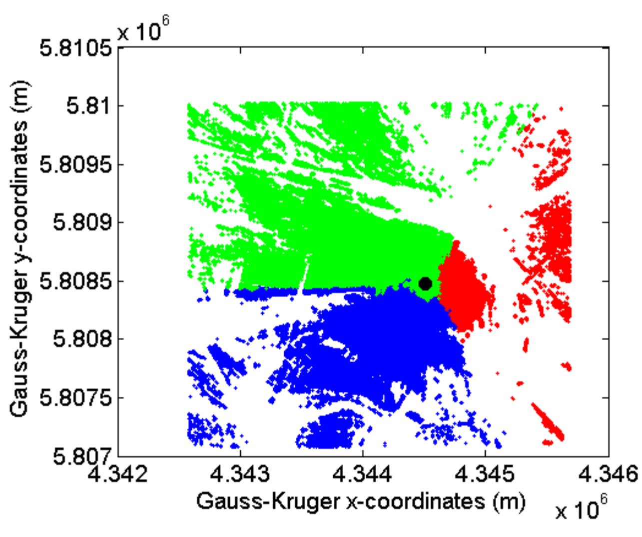

Fig. 1. Results of the preprocessing step for a sectorized cell where the pixels (locations) served by each sector antenna are depicted in different color. The black dot represents the BS location.

III. BAYESIAN FILTERING

Bayesian filtering (BF) [9] is a concept that only provides a probabilistic framework for state estimation. There are two different implementations of Bayesian filter that differ mainly in the way they represent *belief distributions* over the state space. The first implementation represents *continuous* belief distributions, where the second represents *discrete* distributions.

As the state space is divided into pixels with a specified resolution in the context of mobile terminal localization, we will be concerned with the discrete version of the Bayesian filter. BF estimates the posterior belief distribution of a MS position given a map (database) of predicted signal strengths and a series of signal strength measurements. In other words, it estimates the state of a dynamical system, i.e. *partially observable Markov chain*, using measurement data. In this context, the dynamical system is the mobile terminal and its environment, and the state is the MS position relative to that environment.

A. Mathematical Foundations of Bayesian Filtering

BF assumes that the environment is *Markovian*, i.e. if the current state is known; past and future data are conditionally independent. The key idea is to estimate a *posterior probability density* over the state space conditioned on the measurement data. This posterior is called *belief* and is denoted

$$Bel(s_t) = p(s_t \mid o_t, a_{t-1}, o_{t-1}, a_{t-2}, \ldots, o_0, m) \tag{1}$$

Where $Bel(s_t)$ is the MS belief state at time *t*, s_t is the state at time *t*, $o_{t\ldots0}$ denote the measurement data delivered from time *0* up to time *t*, $a_{t-1\ldots0}$ denote the actions (movements) performed by the terminal user from time *0* up to time *t-1*, and *m* is the model of the environment, i.e. a map (or database) of predicted RxLev. It is assumed that measurements and terminal user actions occur in an alternative sequence.

The desired posterior is estimated by applying *Bayes rule*[1], *Markov assumption*, *theorem of total probability*[2], and once again the *Markov assumption* to expression (1) in this stated order. Hence, we obtain the following recursive equation

$$\begin{aligned} Bel(s_t) &= \eta \cdot p(o_t \mid s_t, m) \\ &\int p(s_t \mid s_{t-1}, a_{t-1}, m) \cdot Bel(s_{t-1}) \; ds_{t-1} \end{aligned} \tag{2}$$

Where η is a normalization factor, $p(o_t \mid s_t, m)$ is the measurement model, and $p(s_t \mid s_{t-1}, a_{t-1}, m)$ is the MS motion model. Detailed derivation of equation (2) is provided in [11].

As the actions performed by the terminal user (pedestrian) cannot be directly obtained (without additional inertial sensors) unlike vehicles that have complete motion models,

[1] $p(a \mid b, c) = \dfrac{p(c \mid a, b) \cdot p(a \mid b)}{p(c \mid b)}$

[2] $P(A) = \int p(A \mid B_i) \cdot p(B_i) \; dB_i$

we have decided to implement equation (2) non-recursively only in the pedestrian positioning case. Here the prior or initial belief $Bel(s_{t-1})$ is initialized by a uniform distribution over the state space for every run of the filter.

B. Implementation of the Discrete Bayesian Filter

The key idea is to represent the *belief* $Bel(s)$ at any time by a set of *n* weighted location candidates distributed according to $Bel(s)$ as follows

$$Bel(s) \approx \{s^{(i)}, w^{(i)}\}_{i=1,\dots,n} \tag{3}$$

Where $s^{(i)}$ is the location candidate *i*, and $w^{(i)}$ is a non-negative numeric value called *weight* that determines the importance of the location candidate *i*.

All weights sum up to 1, thus, the *continuous belief* $Bel(s)$, is approximated by a *discrete probability function* defined by the location candidates. $w^{(i)}$ is calculated for every location candidate according to the measurement model as

$$w^{(i)} = p(o_t \mid s_t, m) = \prod_{j=1}^{M} \frac{1}{\sigma_{RxLev}\sqrt{2\pi}} \; e^{-\frac{(RxLev_j - RxLev_{DB_j})^2}{2\sigma^2_{RxLev}}} \tag{4}$$

Where M is the number of the main and neighboring observed BSs ($M \le 7$ in GSM networks), σ_{RxLev} is the standard deviation of the measured RxLev, $RxLev_j$ is the measured RxLev from the *j*-th observed BS, and $RxLev_{DB_j}$ is the database RxLev prediction value of *j*-th observed BS at $s^{(i)}$.

The final location estimate $\hat{s}$ is calculated using one of the following three methods:

1. Taking the location candidate with the highest weight as the location estimate. This is the estimate at which the posterior is maximum and known as the *maximum likelihood estimate (MLE).*

$$\hat{s} = \arg\max \; Bel(s) \tag{5}$$

2. Taking the weighted average of all candidates representing the *belief* as the location estimate. This is the *mean value* of the *posterior distribution* and known as the *weighted average estimate (WAE)*. It will coincide with MLE only in case of *unimodal* and *symmetric* posterior distributions.

$$\hat{s} = \frac{1}{\sum_i w^{(i)}} \sum_i s^{(i)} \cdot w^{(i)} \tag{6}$$

3. Taking the average of the k ($k < n$) best weighted candidates as the location estimate. This is called the *trimmed average estimate (TAE).*

$$\hat{s} = \frac{1}{k} \sum_k s^{(k)} \tag{7}$$

IV. EXPERIMENTAL RESULTS

A. Experimental Setup

Field tests have been carried out in an E-Plus GSM 1800 MHz network operating in a suburban area in Hannover, Germany. The measurements were collected every 4 seconds by a pedestrian along a route of about 2.4 km with a total number of 250 measurement reports. Every report was stamped by a GPS position, which is considered as the true location reference for evaluation purposes of our algorithm. The prediction data available with us covers an area of 9 km^2 and contains 6 BSs (each with three sector antennas) and four indoor antennas. The database resolution is 5 m.

B. Simulation Results

We investigated the accuracy of our localization algorithm by off-line simulations using the three proposed methods for location estimation (see III.B). The experiments have been performed using the prediction database with and without the preprocessing step. This has allowed us to see how much accuracy improvement is achieved using the preprocessed database. When experimenting without preprocessing, the location candidates were selected according to the azimuth information of the BS sector antennas. Each antenna covers a sector with an angle of 120°. The sector radius is determined using TA information ±275 m [10] (assuming a maximum error of ±0.5 bits for TA measurement). In this case, the area of the sector is taken as the coverage area of the sector antenna. Fig. 2 shows for TA=0 the coverage area of a sector antenna as determined by the preprocessing step and the borders of the area that would be considered using antenna azimuth without preprocessing. Here, the radius of the sector equals 550+275=825 m. It is clear that the preprocessed database considers only locations that are actually served by the sector antenna.

The cumulative distributions of the localization error for the different methods of location estimation are depicted in Fig. 3 and summarized in Table I. The results show that taking the MLE as the location estimate is not recommended as this is highly affected by noisy RxLev measurements and database inaccuracies. The WAE provides better estimations as it considers all candidates with respect to their weights, thus reducing the effects that degraded the previous estimation method. However, taking the average of a specified proportion (k = 10%) of the *best weighted candidates* still yields better location estimates than WAE. TAE has provided less sensitivity to erroneous RxLev measurements and helped neglecting outlier candidates. Simulation results have shown that the real MS location is almost always in the region of the *10% best weighted candidates*. It is also shown that the preprocessed database has improved the standard deviation and the mean of the localization error for the three location estimation methods.

V. CONCLUSIONS

We presented simulation results of mobile terminal localization in a GSM network operating in a suburban area. Our localization technique is a database correlation method that utilizes Bayesian filtering for pattern matching. Bayesian filtering is an efficient probabilistic framework for parameter estimation in multi-hypotheses contexts such as mobile location based on the received signal strength. A 3D deterministic radio wave propagation prediction tool, used for cellular network planning, was employed for the database construction.

Different methods for final location estimation has been presented and compared in terms of localization accuracy. The experiments were performed in a suburban environment, which is very common in European cities. Our goal was to achieve positioning accuracy comparable to that reached in urban areas that have a high BS-to-area ratio. We belief that more work should be done to further enhance the positioning accuracy in suburban areas, which have a considerable number of network subscribers also provided with location-based services. We have achieved results with a mean positioning error of 194 m and a standard deviation of 216 m. However, this accuracy is still acceptable for many services, taking into account the characteristics of the test environment (low BS density).

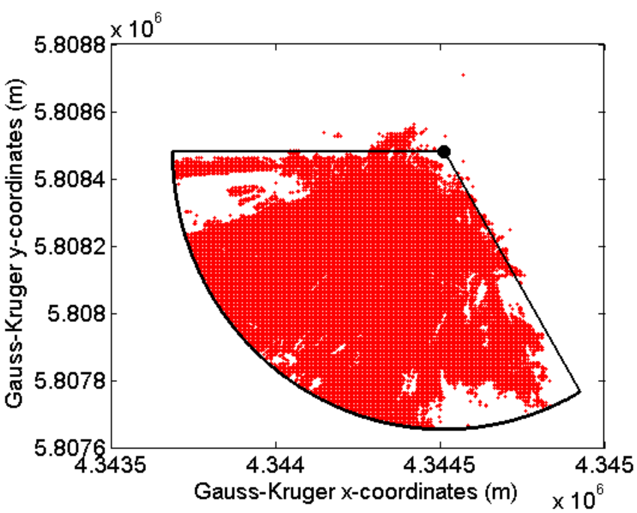

Fig. 2. For TA=0, the coverage area of a sector antenna as determined by the preprocessing step are illustrated (red spots) along with the borders of the area that would be considered when using only the antenna azimuth information (black arc and lines).

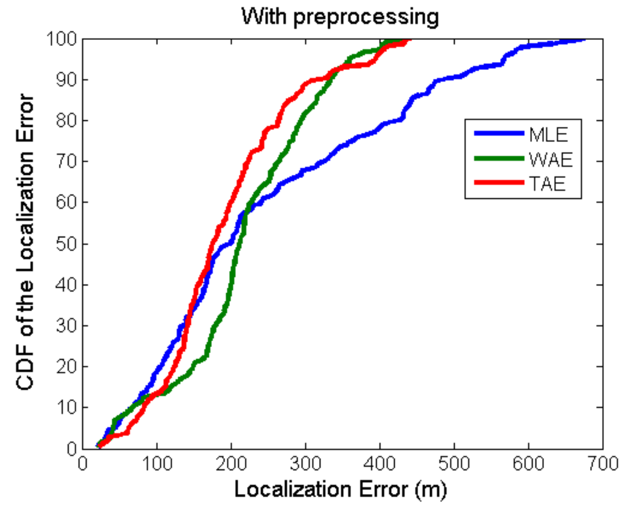

Fig. 3a. Cumulative distribution functions (CDF) of the localization error using the three methods of location estimation with preprocessing.

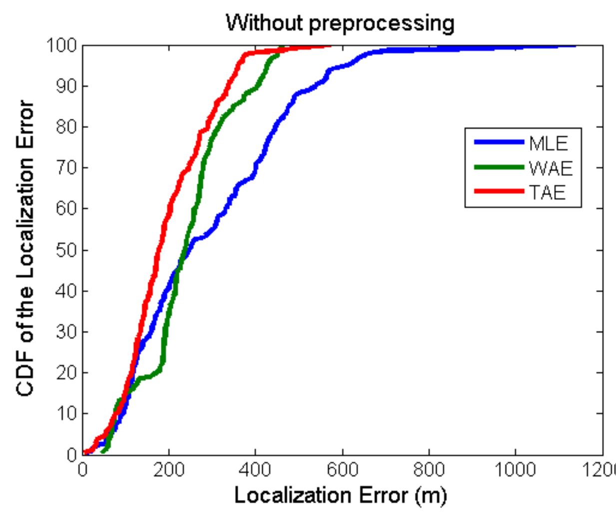

Fig. 3b. Cumulative distribution functions (CDF) of the localization error using the three methods of location estimation without preprocessing.

TABLE I
ACCURACY OF THE PROPOSED METHODS FOR FINAL LOCATION ESTIMATION WITH AND WITHOUT PREPROCESSING

Localization Error		MLE	WAE	TAE
With Preprocessing	67%	295	254	216
	95%	567	361	395
	mean	248	217	194
Without Preprocessing	67%	378	275	229
	95%	610	428	361
	mean	294	240	197

ACKNOWLEDGMENT

We would like to thank Thomas Kürner, Peter Unger, and E-Plus Mobilfunk GmbH & Co KG for providing network and prediction data utilized in our research.

REFERENCES

[1] Federal Communications Commission (FCC) Fact Sheet, "FCC Wireless 911 Requirements," 2001.

[2] G. D. Ott, "Vehicle location in cellular mobile radio system," *IEEE Trans. Vehicle Technology*, vol. VT-26, pp. 43-46, Feb 1977.

[3] T. S. Rappaport, J. H. Reed, and D. Woerner, "Position location using wireless communications on highways of the future," *IEEE Communication Mag.*, vol. 34, pp. 33-41, Oct 1996.

[4] H. Laitinen, J. Lähteenmäki, and T. Nordström, "Database Correlation Method for GSM Location," *VTC 2001 Spring*, Rhodes, Greece, May 2001.

[5] H. Schmitz, M. Kuipers, K. Majewski, P. Stadelmeyer, "A new method for positioning of mobile users by comparing a time series of measured reception power levels with predictions," *VTC 2003 Spring*, Jeju, South Korea, May 2003.

[6] D. Zimmermann, J. Baumann, M. Layh, F.M. Landstorfer, R. Hoppe, G. Wölfle, "Database Correlation for Positioning of Mobile Terminals in Cellular Networks using Wave Propagation Models," *VTC 2004 Fall*, Los Angeles, USA, Sept. 2004.

[7] T. Nypan, "Mobile terminal positioning based on database comparison and filtering," Dissertation at the Norwegian University of Science and Technology, 2004-65, July 2004.

[8] T. Kürner, 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, April 2002.

[9] D. Fox, J. Hightower, L. Liao, D. Schulz, G. Borriello, "Bayesian Filters for Location Estimation," *IEEE Pervasive Computing*, vol. 2, no. 3, 2003.

[10] B. Walke, "*Mobile radio networks: networking and protocols*," Chichester , Wiley, 2000, ch. 3.

[11] M. Khalaf-Allah, and K. Kyamakya, "Database Correlation using Bayes Filter for Mobile Terminal Localization in GSM Suburban Environments," in *Proc. IEEE 63rd Semi-Annual Vehicular Technology Conference (VTC2006-Spring)*, May 7-10, 2006, Melbourne, Australia.