Mobile Location in GSM Networks using Database Correlation with Bayesian Estimation

Mohamed Khalaf-Allah
Institute of Communications Engineering
University of Hannover, Germany
khalaf@ikt.uni-hannover.de

Kyandoghere Kyamakya
Department of Informatics Systems
University of Klagenfurt, Austria
kyamakya@isys.uni-klu.ac.at

Abstract

We present a database correlation method combined with Bayesian estimation for mobile location in GSM networks. The final location request can be estimated using different ways. Three methods are introduced and investigated in terms of location accuracy. Field measurements with off-line simulations have been performed in a suburban area, which is the dominant environment type in many parts of European cities. We think that location accuracy requirements in such environments should be relaxed, otherwise more work have to be conducted to achieve accuracies comparable to those in urban and dense urban areas.

1. 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: (1) Time-of-arrival (TOA), (2) Angle-of-arrival (AOA), (3) Time-difference-of-arrival (TDOA) and enhanced observed time-difference-of-arrival (E-OTD), (4) Network-assisted GPS (A-GPS), and (5) 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 the BSs. TDOA based techniques need at least three BSs, which could not be fulfilled in rural areas. The main shortcomings of A-GPS solutions are power consumption, the need of clear view to at least four satellites, and the installation of additional hardware (reference GPS receivers and GPS receivers in MSs).

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 as the location dependent parameter. 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 estimation that works as a pattern matching technique, and three

methods used to conclude final location estimates. The proposed technique was applied to measurements of the received signal strength (RxLev) and timing advance (TA) in a GSM network covering a suburban area. The location accuracy has been investigated by off-line simulations.

The rest of the paper is organized as follows. The database setup is described in section 2. Section 3 presents the basics of Bayesian estimation and its implementation together with three methods for final location calculation. Experimental results and conclusion are given in sections 4 and 5 respectively.

2. Database Preparation

A 3D deterministic radio propagation prediction model, described in [8], was used to construct our database of RxLev's at known locations. Robust location estimation with DCM depends mainly on three factors: (1) assuring that searching in the database is 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. (2) The pattern matching technique used to evaluate location candidates. For this purpose we use Bayes estimation (section 3), which is a robust technique in multi-hypotheses situations usually occurring in the context of RxLev-based mobile location. (3) Deciding how to yield a final location estimate from the available hypotheses. We present three methods for this factor (section 3.2) and compare their performances (section 4).

The localization algorithm can take advantage if the 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 possible location candidates. Thus the prediction database 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 data, predicted RxLev at that location, and location distance to the serving BS. Figure 1 shows an example of the preprocessing step results. Here, the BS has three sector antennas. The locations served by each antenna are depicted in different colors. The black dot represents the location of the BS.
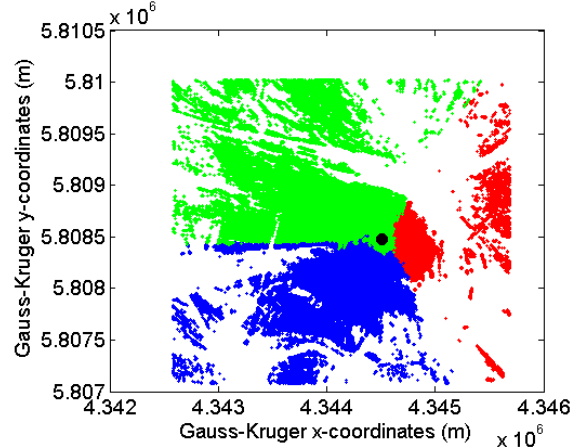

Figure 1. Results of the preprocessing step for a sectorized cell

3. Bayesian Estimation

The Bayesian estimator (BE) [9] is a concept that only provides a probabilistic framework for state estimation. There are two different implementations of BE 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.

In the context of mobile location using DCM, the state space is divided into grids (pixels) with a specified resolution. Therefore, we will be concerned with the discrete version of the BE. BE 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.

3.1. Mathematical Principles of Bayesian Estimation

The BE 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 the *belief* and is denoted

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

Where $Bel(s_t)$ is the MS belief state at time t, s_t is the state at time t, $o_{t...0}$ denote the measurement data delivered from time 0 up to time t, $a_{t-1...0}$ 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 in order *Bayes rule*, *Markov assumption*, *theorem of total probability*, and once again the *Markov assumption* to expression (1). Hence, we obtain the following recursive equation

$$Bel(s_t) = \eta \cdot p(o_t \mid s_t, m)$$
$$\int p(s_t \mid s_{t-1}, a_{t-1}, m) \cdot Bel(s_{t-1}) \, ds_{t-1} \qquad (2)$$

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 [10].

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

3.2. Implementation of the Discrete Bayesian Estimator

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

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

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

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

Where M is the number of the main and neighboring observed BSs ($M : 7$ 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:

- 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)*.

$$\hat{s} = \arg\max \, Bel(s) \qquad (5)$$

- 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.

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

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

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

4. Experimental Results

4.1. 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 the proposed algorithm. The prediction data available with us covers an area of 9 km^2 (with a resolution of 5 m) and contains 6 BSs (each with three sector antennas) and four indoor antennas.

4.2. Simulation Results

We investigated the accuracy of the localization algorithm by off-line simulations using the three proposed methods for location estimation. The percentile localization error curves for the different methods are depicted in Figure 2 and summarized in Table 1. 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 WA provides better estimations as it considers all samples with respect to their weights, thus reducing the effects that degraded the previous estimation method. However, the trimmed average method still yields better location estimates than WA. This method has shown less sensitivity to erroneous RxLev measurements and helped neglecting outlier candidates. Simulation results have also shown that the real MS location is almost always in the region of the 10% best weighted location candidates.

5. Conclusion

We presented simulation results of mobile terminal localization in a GSM network operating in a suburban area. The proposed localization technique is a database correlation method that utilizes Bayesian estimation. The Bayesian estimator is an efficient probabilistic framework for parameter estimation in multi-hypotheses contexts such as mobile location based on the received signal strength. 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.

It is planned to integrate an inertial measurement unit (IMU) in our localization system in order to fully exploit equation (2) for the pedestrian case.
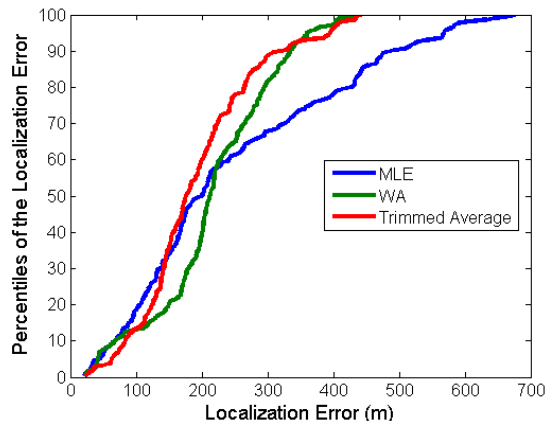

Figure 2. Cumulative distribution functions (CDF) of the localization error using the three methods for final location estimation

Table 1. Summary of the mean, 67%, and 95% percentiles localization errors for the location estimation methods

Localization Error	MLE	WA	Trimmed Average
67%	295 m	254 m	216 m
95%	567 m	361 m	395 m
Mean	248 m	217 m	194 m

6. Acknowledgment

7. 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. 26, Feb 1977, pp. 43-46.

[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, Oct 1996, pp. 33-41.

[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, April 2002, pp. 496-506.

[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] 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.