14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

BAYESIAN MOBILE LOCATION IN CELLULAR NETWORKS

Mohamed Khalaf-Allah

Institute of Communications Engineering, University of Hannover
Appelstrasse 9A, D-30167, Hannover, Germany
phone: + (49) 511 762 2845, fax: + (49) 511 762 3030, email: khalaf@ikt.uni-hannover.de
web: www.ikt.uni-hannover.de

Kyandoghere Kyamakya

Department of Informatics Systems, University of Klagenfurt
Universitätsstraße 65-67, A-9020, Klagenfurt, Austria
phone: + (43) 463 2700 3540, fax: + (43) 463 2700 3598, email: kyamakya@isys.uni-klu.ac.at
web: www.isys.uni-klu.ac.at

ABSTRACT

Determining the location of mobile stations could be achieved by collecting signal strength measurements and correlating them to pre-calculated signal strength values at reference locations. This method is advantageous, because no LOS conditions are needed, it can work even with one base station (BS), and its implementation costs are pretty low. However, the correlation process needs an appropriate likelihood function such as that provided by Bayesian statistical estimation approaches. They use all available information surrounding candidate hypotheses to determine their likelihoods. In this paper, we present a Bayesian mobile location algorithm, and show its performance by field measurements in a working GSM network.

1. INTRODUCTION

The key driver for developing mobile station (MS) 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 a MS is considered more critical because MS users and hence MS originated emergency calls are continually increasing. It is estimated that about 50% of all emergency calls in the EU are MS originated, and the expected tendency is rising [2].

While emergency call location could be considered the most important of location-based services (LBS) due to its urgency for life and property safety, commercial LBS are believed to make increasing revenues for network operators who could provide customers with attractive and tailored services [3].

The MS location is usually achieved using satellite-based or cellular system based methods [4], [5]. These methods differ in terms of accuracy, coverage, cost, power consumption and system impact. Satellite-based technologies come in two flavours: stand-alone GPS or Assisted-GPS (A-GPS). Main drawbacks are power consumption, need of clear view to at least four satellites (for stand-alone GPS) and the costs of integrating GPS receivers into the mobile terminals. Furthermore, A-GPS solutions require the additional installation of reference GPS receivers.

Cellular system based techniques include: cell-id (CI), time of arrival (TOA) / uplink time difference of arrival (U-TDOA), enhanced observed time difference (E-OTD) and angle of arrival (AOA). There are many varieties of the cell-id method [6], [7], namely CI, CI+TA (timing advance) of serving cell, CI+TA of several adjacent cells (this is usually not the case in GSM networks), and CI+TA+RxLev (received signal level). Here, the RxLev measurement can be input to an empirical formula or compared with the entries of a look-up table in order to estimate distances to base stations or yield a location estimate of the MS respectively. The later handling is known as *database correlation method (DCM)* [8] – [11].

The location service is divided into three levels [7] according to the accuracy requirements of the different applications. The less accurate is the *basic service level*, which utilizes CI methods. Techniques used for the *enhanced service level* are E-OTD, TOA/U-TDOA, or AOA. A-GPS is usually employed for the *extended service level,* which is the most accurate location service.

Cell-id methods are the simplest to implement because they utilize only network available information. Thus, they are advantageous in terms of cost, coverage, and system impact. TOA/U-TDOA and E-OTD based techniques need mutual synchronization of at least three base stations (BSs) which is difficult to achieve. Installation of special antennas at BSs is a must for the implementation of AOA methods. Location accuracy of TOA/U-TDOA, E-OTD and AOA approaches are severely influenced by multipath propagation which is the dominant propagation condition in built-up environments.

Basic service level location methods 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.

Unlike the US mandate, the EU location requirements do not specify accuracy or standards. This was another reason that pushed toward the implementation of CI methods by EU network operators. However, location accuracy is in the range of hundred meters up to several kilometers depending largely on the environment characteristics, network layout and propagation conditions. Therefore, improving positioning accuracy of CI techniques is an active research topic.

In this paper, we present a database correlation method (DCM) within a Bayesian statistical framework for mobile location in GSM networks. The proposed location algorithm is a generic one that could also be applied to other cellular systems and wireless networks. The mathematical derivation and the practical implementation of the proposed Bayesian filtering algorithm are provided in the next section. In section 3, we discuss the model of the wireless environment. Experimental results are given in section 4. Section 5 summarizes the paper.

2. THE BAYESIAN LOCATION ALGORITHM

2.1 Mathematical Derivation of the Bayes Filter

Bayes Filter (BF) [13], [14] 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). In the case of mobile location, BF estimates the posterior belief distribution of the MS position given its prior belief, a series of RxLev measurements, and a model of its world (environment).

The prior belief is a probability distribution over all locations of the given cell combined with the TA measurement before taking the MS actions and RxLev measurements into account. The posterior belief is the conditional distribution of these locations given the MS actions and RxLev measurements. The world model is a database that contains predicted RxLev at the candidate locations.

The posterior belief distribution is expressed as

$$Bel(s_t) = p(s_t \mid o_{0:t}, a_{0:t}, m) \tag{1}$$

Where $Bel(s_t)$ is the posterior belief over the state (position) of MS at time t, s_t is the state at time t, $o_{0:t}$ are the measurement data from time 0 up to time t, $a_{0:t}$ are the actions performed by the MS from time 0 up to time t, and m is the world model.

Applying *Bayes rule* to equation (1) we get

$$Bel(s_t) = \frac{p(o_t \mid s_t, o_{0:t-1}, a_{0:t}, m) \;\; P(s_t \mid o_{0:t-1}, a_{0:t}, m)}{p(o_t \mid o_{0:t-1}, a_{0:t}, m)} \tag{2}$$

Here, actions and measurements are assumed to occur in an alternative sequence, 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, equation (2) is rewritten as

$$Bel(s_t) = \eta \;\; p(o_t \mid s_t, m) \;\; p(s_t \mid o_{0:t-1}, a_{0:t}, m) \tag{3}$$

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

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

$$Bel(s_t) = \eta \;\; p(o_t \mid s_t, m)$$
$$\int p(s_t \mid s_{t-1}, o_{0:t-1}, a_{0:t}, m) \;\; p(s_{t-1} \mid o_{0:t-1}, a_{0:t}, m) \;\; ds_{t-1} \tag{4}$$

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

$$\begin{aligned} Bel(s_t) &= \eta \;\; p(o_t \mid s_t, m) \\ &\int p(s_t \mid s_{t-1}, a_t, m) \;\; Bel(s_{t-1}) \;\; ds_{t-1} \end{aligned} \tag{5}$$

Expression (5) is a recursive equation that is usually computed in two steps called *prediction* and *update* [13], [14].

Prediction step:

$$Bel^{-}(s_t) = \int p(s_t \mid s_{t-1}, a_t, m) \;\; Bel(s_{t-1}) \;\; ds_{t-1} \tag{6}$$

Where $Bel^{-}(s_t)$ is the posterior belief just after executing the action a_t and before incorporating the measurement o_t, and $\int p(s_t \mid s_{t-1}, a_t, m)$ is the MS motion model, i.e. the transition probability that tells us how the state evolves over time as a function of the MS movements. These movements are undeterminable without an extra measurement source, i.e. inertial measurements.

Update step

$$Bel(s_t) = \eta \;\; p(o_t \mid s_t, m) \;\; Bel^{-}(s_t) \tag{7}$$

Where $p(o_t \mid s_t, m)$ is the measurement model that specifies the probabilistic law according to which measurements are generated from the state, i.e. measurements are simply noisy projections of the state [14].

Both motion and measurement models describe the dynamical stochastic system of the MS and its environment. The state at time t is stochastically dependent on the state at time $t-1$ and the action a_t. The 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)* [14].

2.2 Practical Implementation

The Bayes Filter (BF) algorithm derived in the previous section cannot be directly implemented on a digital computer. However, nonparametric filters [14] provide implementable algorithms for the BF. Nonparametric filters (NPF) approximate posteriors by a finite number of parameters, each corresponding to a region in the state space, i.e. 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 under specific smoothness assumptions [14]. 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 Bayes Filter (DBF)* [14]. The DBF is also referred to as the forward pass of a hidden Markov model.

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

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

Where $s^{(i)}$ is the *i*-th MS location candidate 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 $Bel(s)$ represents a valid probability distribution. At any time, the weight of a location candidate is calculated as

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

Where $w_{MM}^{(i)}$, $w_{ND}^{(i)}$ and $w_{SN}^{(i)}$ are the weights according to the *measurement model*, *neighbourhood degree*, and *strongest neighbour* respectively. They are calculated at time t as

$$w_{MM}^{(i)} = p(o_t \mid s_t^{(i)}, m) = \prod_{j=1}^{M} \frac{1}{\sigma_{RxLev}\sqrt{2\pi}} e^{-\frac{(RxLev_j - RxLev_{DB_j})^2}{2\sigma_{RxLev}^2}} \tag{10}$$

Where M is the number of observed BSs (main and neighbouring), i.e. $M_{\max} = 7$, 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 the *j*-th observed BS at $s^{(i)}$.

$$w_{ND}^{(i)} = l \cdot \alpha_{ND} \tag{11}$$

Where l is the number of observed neighbour BSs that coincide with the list of the predicted six strongest neighbour BSs at $s^{(i)}$, and α_{ND} is a constant bonus value, i.e. $l_{\max} = 6$.

$$w_{SN}^{(i)} = \alpha_{SN} \tag{12}$$

Where α_{SN} is a constant bonus value, and is assigned if the strongest observed neighbour BS coincides with the predicted first or second strongest neighbour BS at $s^{(i)}$. Otherwise, $w_{SN}^{(i)} = 0$.

The final location estimate $\hat{s}$ is calculated from the belief $Bel(s)$ as

$$\hat{s} = \frac{1}{k}\sum_{i=1}^{k} s^{(i)} \tag{13}$$

Where $k < n$, and $Bel(s)$ is sorted according to $w^{(i)}$. Thus, $\hat{s}$ is the average of a certain number (k) of the best weighted location candidate. Expression (13) is also known as *trimmed average estimate (TAE).*

TABLE I depicts the implementation of the proposed Bayesian mobile location algorithm when run at time t. Note that no motion model is integrated, because network measurements are the only source of information. The prior Belief at time t (denoted $Bel(s_t)_{prior}$) is initialized over the whole state space of the MS candidate locations using the CI and TA at time t with initial weights $w_t^{(i)} = \frac{1}{n}$, i.e. the prior belief is a uniform distribution over the determined state space.

3. ENVIRONMENT MODEL

The utilized database has been constructed using a 3D deterministic radio propagation prediction model, described in [12], with a resolution of 5 m. This database is a by-product of the network planning stage and contains location dependent parameter values (e.g. signal strength in GSM networks) at reference locations. The provided cell information in the interest area include antenna geographical location, antenna height, azimuth and tilt, effective isotropic radiated power, channel numbers, cell identifiers, etc.

The MS acquires information about its environment (or world) through the network measurements. However, the MS environment is a stochastic system. Therefore, the network RxLev measurements are often noisy and deviate from the prediction RxLev values, which are in turn not precise.

In order to enhance the prior belief of the discrete Bayes filter, as much information as possible could be extracted from the prediction database. This would enhance the correlation process of measurements with knowledge about the MS world. Achieving this needs reorganization, partitioning, and clustering of the initial prediction database.

Every cell antenna of the test area has acquired a separate database, called the *cell database (CDB)*, which contains only the locations served by it. Each database entry consists of location ID, location coordinates, prediction RxLev from serving cell, prediction RxLev values and IDs of the strongest neighbour cells, and distance to the serving cell antenna.

Furthermore, every CDB has been divided into sub-databases according to all possible TA values (with an assumed error of ± 0.5 bits); each called *cell TA database (CTADB)* and labelled with a stamp indicating its TA value. The location algorithm will process only the CTADB matching the TA measurement, thus, reducing the online computational burdens to a minimum.

Another interesting aspect can be explained by the help of Figure 1, which illustrates the location of a sector cell antenna (black dot), locations served by the cell antenna for $TA = 0$ (red spots), and the sector boundary using the azimuth and coordinates of the cell antenna (depicted in black), also at $TA = 0$. The white areas inside the boundary are locations served by other cell antennas. Such locations could be determined along with their serving CI and the other information as above at all possible TA values for every cell, and then stored in separate databases, each called *outsider locations database (OLDB).*

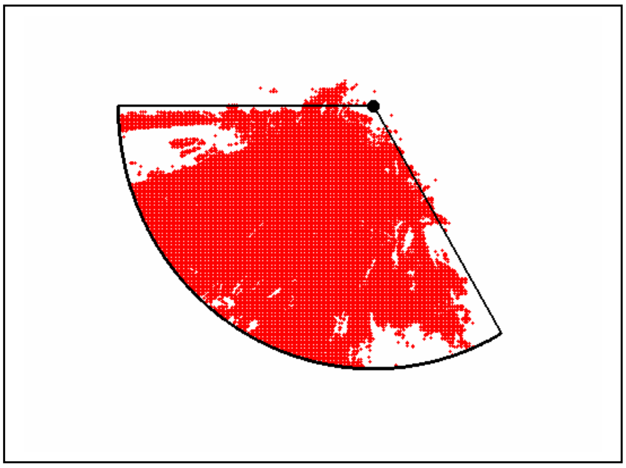

Figure 1 – Definition of outsider locations

When the actual network measurement reports a switching to a new serving cell, it is most probably that the true location of the MS is somewhere in the white areas (as explained above) at least in the first period of time after switching. This is very advantageous for the discrete Bayes filter, in which the state space is more specified by the concentration of the prior belief on locations of more likelihood.

Inputs: $Bel(s_t)_{prior} = \{s_t^{(i)}, w_t^{(i)}\}_{i=1:n}$

$RxLev_{j=1:M}$

$Bel(s_t) = 0$ // Initialize posterior belief

$\eta = 0$ // Initialize normalization factor

$\hat{s} = 0$ // Initialize location estimate

for $i = 1:n$ ***do***

// Compute weights

$w_t^{(i)} = w_t^{(i)}{}_{MM} + w_t^{(i)}{}_{ND} + w_t^{(i)}{}_{SN}$

// Insert into posterior belief

$Bel(s_t) = Bel(s_t) \cup \{s_t^{(i)}, w_t^{(i)}\}$

// Update normalization factor value

$\eta = \eta + w_t^{(i)}$

endfor

// Normalize weights

for $i = 1:n$ ***do***

$w_t^{(i)} = w_t^{(i)} / \eta$

endfor

// Sort posterior belief according to weights in descending order

$Bel(s_t) = sort(Bel(s_t))$

// Estimate the location using TAE

for $i = 1:k$ ***do***

$\hat{s} = \hat{s} + s^{(i)}$

endfor

$\hat{s} = \hat{s} / k$

return $\hat{s}$

TABLE I – Implementation of the discrete Bayes filter

4. EXPERIMENTS

Field measurements have been collected in a working GSM-1800 network by a pedestrian using a notebook connected to a GSM modem and a GPS receiver that provided true position references. The test field is a 9 km^2 suburban area in

Hannover, Germany, with 18 and 4 sector and indoor cells respectively.

The collected measurements have been processed offline using the proposed Bayesian location algorithm. We investigated the performance by running the algorithm once with only the CTADBs and another once with both CTADBs and OLDBs as explained in the previous section.

Using only the CTADBs, the achieved location accuracy was as shown in TABLE II.

Error percentiles	**67%**	**95%**	**mean**
Location error	240 m	419 m	216 m

TABLE II – Location accuracy using only CTADBs

TABLE III depicts the enhancement of the performance accuracy when incorporating OLDBs into the location algorithm. The improvement of the 67 and 95 percentiles, and mean error is 12%, 9%, and 11% respectively. The utilization of OLDBs has enhanced prior beliefs when serving cell changed, accordingly the Bayesian filtering process could perform better with more useful information.

Error percentiles	**67%**	**95%**	**mean**
Location error	211 m	382 m	191 m

TABLE III – Location accuracy using CTADBs and OLDBs

Location accuracy depends strongly on the cell size. The performance of our algorithm is still more accurate than those presented in, e.g. [10], [11], for similar cell sizes also using database correlation methods.

5. CONCLUSION

The Bayes filter (BF) algorithm calculates the posterior over the state conditioned on the measurement data. It is well-suited to represent complex multimodal beliefs as is the case in the problem of MS positioning in wireless networks. BF assumes that the world is *Markovian*. This assumption could be considered somehow severe, because it is already violated in building the world model and during real measurements due to the fact that unmodeled dynamics (e.g. people and cars) are not included in calculation despite their influence on, e.g. multipath, and hence on the resultant RxLev value at different locations. However, the proposed approach is robust in the face of such assumptions, noisy measurements, and other inaccuracies in the environment model. They are handled as close-to-random effects. Another limitation is the approximation of posterior distributions in continuous environments. This is, however, unavoidable in order to make the location algorithm computationally feasible. Field experimental results showed good performance accuracies of the implemented algorithm in a suburban environment with low BS density.

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, July 2003.

[3] T. M Rantalainen, M. A. Spirito, and V. Ruutu, “Evolution of location services in GSM and UMTS networks,” in *Proc. The 3rd Int. Symp. on Wireless Personal Multimedia Communications (WPMC 2000)*, Bangkok, Thailand, Nov. 2000, pp. 1027-1032.

[4] H. Laitinen (editor), “Cellular Location Technology,” *Public deliverable of IST/CELLO project*, 2001.

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

[6] M. Schreiner, M. Tangemann, and D. Nikolai, “A new network-based positioning method for location services in 2G and 3G mobile communications,” in Proc. *5th European Personal Mobile Communications Conference (Conf. Publ. No. 492)*, 22-25 April 2003, pp.162 - 168.

[7] M. Weckström, M. Spirito, and V. Ruutu, “Mobile Station Location,” ch. 4 in *GSM, GPRS, and edge performance: evolution towards 3G/UMTS*, T. Halonen (ed.), 2nd ed., Chichester, Wiley, 2003.

[8] H. Laitinen, J. Lähteenmäki, T. Nordström, “Database Correlation Method for GSM Location,” presented at *IEEE Vehicular Technology Conference (VTC2001-Spring)*, Rhodes, Greece, May 2001.

[9] F. Erbas, K. Kyamakya, K. Jobmann, “On the user profiles and the prediction of user movements in wireless networks,” presented at *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC ’02)*, Lisbon, Portugal, 2002.

[10] 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,” presented at *IEEE Vehicular Technology Conference (VTC2003-Spring)*, Jeju, South Korea, May 2003.

[11] 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,” presented at *IEEE Vehicular Technology Conference (VTC2004-Fall)*, Los Angeles, USA, Sept. 2004.

[12] 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.

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

[14] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. Cambridge, Mass.: MIT Press, 2005.