A FAST DATABASE CORRELATION ALGORITHM FOR LOCALIZATION OF WIRELESS NETWORK MOBILE NODES USING COVERAGE PREDICTION AND ROUND TRIP DELAY

Rafael Saraiva Campos and Lisandro Lovisolo

Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro
rafael_saraiva@ig.com.br, lisandro@uerj.br

ABSTRACT
In this work a fast database correlation algorithm is proposed for the localization of mobile stations in wireless metropolitan area networks. The algorithm is entirely based on built-in functionalities of mobile nodes. The set of measured parameters used for localization is called radio-frequency fingerprint. To provide a position estimate, it is correlated with a database built from field measurements or propagation modeling. The latter alternative was selected, as it allows a less expensive and faster database update. A novel use of the round trip delay to reduce the correlation space and improve positioning accuracy is presented. Vehicular field tests and Monte Carlo simulations were conducted to evaluate the algorithm’s accuracy in GSM and WCDMA cellular networks, respectively.

1. INTRODUCTION
Different levels of location awareness are inherent to any metropolitan area network (MAN) with wireless radio access, in order to allow packet routing to any mobile station (MS) and session continuity when the MS moves from one base transmission station (BTS) coverage area to another. These built-in location capabilities can be extended to provide added value location services (LCS) to the subscribers, which include, among others, location based billing and emergency call location. This latter application has been receiving much attention from government authorities and is already mandatory in cellular networks in some countries [1][2].

Recent wireless network standards have specifically addressed MS location issues. The European Telecommunication Standards Institute (ETSI) has defined MS location methods for Global System for Mobile Telecommunications (GSM) radio access networks [3]. Similarly, the Third Generation Partnership Project (3GPP) has defined User Equipment (UE) location methods for Universal Mobile Telecommunications System (UMTS) Wideband Code Division Multiple Access (WCDMA) networks [4]. Both specifications propose location methods based on Round Trip Delay (RTD), Angle of Arrival (AOA), Time of Arrival (TOA), Observed Time Difference of Arrival (OTDOA), Global Positioning System (GPS) or Assisted GPS (AGPS). Some of those location methods require deployment of new network elements, like antenna arrays in the AOA method [5] or Location Measurement Units (LMU) in the OTDOA method using measurements of asynchronous Node Bs in WCDMA networks [4]. Other solutions require changes in the MS, from software updates in the OTDOA [6] to MS replacement, as in the GPS or AGPS based methods. Those solutions are either quite expensive or do not allow positioning of legacy MS. Thereby, location methods that require no MS modifications and are entirely network based have been the focus of several studies [7][8][9]. Even in a scenario where all MS were GPS compliant, network based methods would still play an important role as fallback during GPS unavailability, as in indoor environments or in street canyons in dense urban areas. By doing so, network based methods would extend LCS availability.

The time alignment based location methods include the CI+TA (Cell Identity and Timing Advance) and CI+RTT (Cell Identity and Round Trip Time) algorithms. Both make use of forced handovers and require at least three BTS to allow a position fix [6][9]. Even though they might yield fairly accurate position estimates, especially in wideband systems, they impose an additional processing load at the Base Station Controller (BSC), due to the forced handovers. Besides that, the MS is not always within reach of at least three BTS, especially in suburban and rural areas.

Another class of location methods comprise the DCM (Database Correlation Method) algorithms, which correlate measured RF (Radio Frequency) fingerprints with a RF fingerprint database to provide position estimates [10][11]. Those fingerprints contain RF parameters measured by the MS and reported through Network Measurement Reports (NMRs), which are periodically sent by the MS to the serving BTS. The correlation databases (CDBs) might be built from field measurements or from propagation modeling. The fast DCM algorithm proposed in this work is extensive to any wireless network using TDM (Time Division Multiplex) or CDM (Code Division Multiplex) in the physical layer. In this group, which includes WCDMA, CDMA, GSM and WiMax (Worldwide Interoperability for Microwave Access) networks, synchronization between the MS and at least the serving BTS is required. This allows the insertion of the RTD in the RF fingerprint, which reduces the computational cost and improves accuracy, especially in areas with low BTS density. The proposed algorithm is capable of yielding a position fix with only one server, which increases the LCS availability, and does not increase the processing load at the BSC, as it does not make use of forced handovers.

DCM solutions in the literature [10][12] usually employ CDBs built from field tests. To keep those CDBs up-to-date, drive tests must be carried out after any change in radio access network elements, turning this solution into an impractical one. Therefore, this paper proposes the use of a RF fingerprint CDB built from propagation modeling, which allows quick and inexpensive CDB upgrade and prevents accuracy degradation due to the use of out-of-date network parameters in the correlation process. However, to maximize the accuracy of the fast DCM algorithm, additional efforts might be required in initial stages to tune the propagation models used to predict the best server areas.

The propagation modeling and the coverage map construction are introduced in Section II. The proposed RF fingerprint correlation algorithm is presented in Section III. The results of the GSM 850
MHz field tests are analyzed in Section IV. The results of a Monte Carlo simulation developed to assess the algorithm’s performance in a WCDMA 850 MHz network are discussed in Section V. Conclusions and future developments are presented in Section VI.

II. RF COVERAGE PREDICTION

II-A. Propagation Model

The propagation model used to build the predicted Received Signal Strength (RSS) map was Okumura-Hata [13], which is largely applied to RF planning in cellular networks. The average propagation loss is given by:

\[ L = k_1 + k_2 \log(d) + k_3 \log(h_b) + k_4 L_f + k_5 \log(h_b) \log(d) \]  

(1)

where \( d \) is the distance in meters from the BTS to MS, \( h_b \) is the BTS antenna effective height and \( L_f \) is the diffraction loss attenuation, calculated by the Epstein-Peterson method [14]. The MS height is assumed to be 1.5 m. The model coefficients \( k_i \) depend on the area morphology and transmission frequency. As both the GSM test and the WCDMA simulation in Sections IV and V were carried out in the 869-881 MHz band, the model was applied at the central frequency of 875 MHz. The coefficients values are \( k_1=-12.1 \) (urban) and \( k_2=-2.2 \) (suburban), \( k_3=-44.9 \), \( k_4=-5.83 \), \( k_5=0.5 \) (urban) and \( k_4=0.4 \) (suburban), \( k_5=6.55 \). All those values are the standard Okumura-Hata values for urban and suburban morphologies, except \( k_5 \), which was empirically defined by the authors for use in the two test areas listed in Section IV.

II-B. The Predicted RSS Map

The topography of the region where the LCS is to be offered is represented by a matrix \( T = (a_{i,j})_{I \times J} \), where \( a_{i,j} \in \mathbb{R}^+ \) for any \( i \in \{1,2,\ldots,I\} \) and \( j \in \{1,2,\ldots,J\} \). Each matrix element \( a_{i,j} \) stores the terrain height averaged over a \( r^2 \) square. The \( T \) matrix might also contain, added to the terrain height, the buildings heights. If the region covers a total surface of \( L \times W \) square, then \( T \) will have \( \lceil \frac{L}{r} \rceil \times \lceil \frac{W}{r} \rceil \) elements, where each one of them referred to as a pixel. The \( r \) parameter is the \( T \) matrix planar resolution.

The RSS prediction for an isolated antenna is obtained by the application of Eq. (1) to the terrain profile between the BTS antenna position and each pixel in the test area. The transmitting antenna characteristics - geographical coordinates, azimuth, radiation pattern, effective isotropic radiated power (EIRP), etc - are considered in the propagation prediction process. The RSS map is obtained by the superposition of each antenna predicted coverage, resulting in a matrix with \( \lceil \frac{L}{r} \rceil \times \lceil \frac{W}{r} \rceil \times V \) elements, where \( V \) is the number of strongest servers considered at each pixel. The RSS map has a planar resolution \( r' \), where \( r' \geq r \). Each element of the RSS map stores the predicted RSS and the correspondent cell identity.

III. RF FINGERPRINT CORRELATION ALGORITHM

The algorithm proposed in this work (RF-FING+RTD-PRED - Predicted RF Fingerprint with RTD) obtains the MS position estimate through correlation of a measured RF fingerprint (RF-FINGM) with predicted RF fingerprints (RF-FINGP) stored in a CDB. The RF-FINGM is obtained from a subset of the parameters available in the NMR or at the serving node. The parameters measured by the MS and reported in the NMR include the serving and neighbor nodes identities and control channel RSS. The RSS measurement is made on the control channel because it is transmitted with constant power. The RTD parameter is typically measured by the serving node in networks using CDMA or TDM in the physical layer. The RTD can be included in the RF fingerprint to further reduce the confidence region, i.e., the area where the MS can be located, thereby enhancing positioning accuracy. Assuming a neglectful processing time at the MS, the RTD can be directly related to the signal propagation distance between the serving node antenna and the MS [15][16] by:

\[ \text{RTD} = \frac{2d/c}{T_s} \]  

(2)

where \( c \) is the speed of light, \( d \) is the propagation distance and \( T_s \) is the symbol period.

Each pixel in the RSS map will have an associated RF-FINGP. The received signal levels in the RF-FINGP are directly obtained from the RSS matrix values. The predicted RTD value (\( \text{RTD}_{\text{P}} \)) in each pixel is calculated by Eq. (2), assuming \( d \) as the LOS distance between the pixel and the antenna of the predicted serving node, which is the node whose predicted RSS is the highest at that pixel. The \( \text{RTD}_{\text{P}} \) may diverge from the measured value, especially due to NLOS (Non Line Of Sight) propagation.

Let \( S \) be the set of all pixels in the test area, i.e., the elements of the planar projection of the RSS matrix. The RF-FINGP of the \( i \)th pixel is the pair \( (\text{RTD}_{\text{P},i}, \hat{\nu}_{\nu,i}) \), where \( \text{RTD}_{\text{P},i} \) is the predicted RTD value calculated by Eq. (2) and \( \hat{\nu}_{\nu,i} \) is the matrix

\[ \hat{\nu}_{\nu,i} = \begin{bmatrix} \text{ID}_{\nu,1} & \text{RSS}_{\nu,1} \\ \vdots & \vdots \\ \text{ID}_{\nu,N} & \text{RSS}_{\nu,N} \end{bmatrix} \]  

(3)

The \( \hat{\nu}_{\nu,i} \) matrix has \( N \times 2 \) elements, where \( N \) ranges from 1 (when there is only one serving sector) to \( N_{\text{max}} \), which is the maximum number of access nodes that the MS can monitor and report in the NMR. Each line of \( \hat{\nu}_{\nu,i} \) matrix contains the identification (ID) and RSS of each serving node. The lines are organized in descending order of RSS level. The RSS range and quantization step vary depending on the wireless access technology. The RF-FINGP CDB is completed after the pair \( (\text{RTD}_{\text{P},i}, \hat{\nu}_{\nu,i}) \) has been calculated for all pixels in the RSS map.

To estimate the MS position one must carry out the correlation between the measured and the predicted RF fingerprints stored in the CDB. The MS position can be assumed as that of the pixel whose predicted RF fingerprint is most similar to the measured RF fingerprint. However, this approach can be misleading due to propagation prediction errors, MS inherent RSS measurement inaccuracy and channel impairments. Thus, instead of considering just one pixel as the MS position estimate, a set of at least \( K \) pixels - with the RF fingerprints which are most similar to the one reported by the MS - is averaged to estimate the MS location. One way to evaluate the similarity between the RF-FINGM and each RF-FINGP is by means of the squared difference. However, it would not be feasible to calculate the squared difference for all pixels in the service area. So a pixel filtering process to reduce computational load without impairing the location accuracy is required. The RF-FING+RTD-PRED algorithm uses a four step filtering process:

1. Selection of the pixels in \( S \) that are within the best server area of the cell identified by \( \text{ID}_{M,1} \), obtaining the set \( A = \{ i \in S \mid \text{ID}_{P,1} = \text{ID}_{M,1} \} \), where \( \text{ID}_{P,1} \) is the predicted best server ID at the \( i \)th pixel and \( \text{ID}_{M,1} \) is the measured best server ID. Note that \( A \subseteq S \).

2. Selection of the pixels in \( A \) whose predicted and measured RTD values are equal, obtaining the set \( B = \{ i \in A \mid \text{RTD}_{P,i} = \text{RTD}_{M} \} \), where \( \text{RTD}_{P,i} \) is the predicted
IV. TRIALS IN A GSM 850 MHZ NETWORK

IV-A. Drive Test Scenario

Field tests were performed in a GSM 850 MHz network in Downtown (Area 1) and Santa Cruz (Area 2) regions, both in Rio de Janeiro city, to evaluate the RF-FING+RTD-PRED algorithm performance in a real propagation environment. The two test areas characteristics are summarized in Tab. I. The test set was composed of a standard legacy GSM phone and a GPS receiver, both connected to a laptop placed inside a moving vehicle. The MS was in active mode and for each transmitted NMR the current location was calculated by the GPS. The RTD values, known as Timing Advance (TA) in GSM systems, were also registered. Every TA, NMR and GPS measurement were recorded for further processing: 4501 records in Area 1 and 7864 in Area 2. The high number of samples ensures a reasonable confidence for statistical analysis. The GPS location estimate is given by the arithmetic mean of the coordinates of the pixels in D. Parameters N and K may have strong impact in the algorithm performance and must be carefully selected [8].

\[
d_i = \sum_{ID \in F_i} \left( \frac{(RSS_{IDM} - RSS_{IDP}) \delta^{-1}}{} \right)^2
\]

The mean location error was 134 meters, with displacements of 119, 155 and 297 meters for the 50th, 67th and 95th percentiles, respectively. Fig. 1 also shows the CDF of the location error yielded by the same algorithm, but considering only those NMRs whose reported best server (BS\(_{PM}\)) was equal to its predicted best server (BS\(_{P}\)). The BS\(_{PM}\) associated with a given NMR is obtained from the pixel of the RSS map located where the NMR was collected. A priori, equivalence between BS\(_{P}\) and BS\(_{PM}\) will result in better accuracy, as the real MS position will not be excluded from the confidence region already in the first filtering step. This subset with BS\(_{PM}\)=BS\(_{P}\) contains 2029 NMRs (45% of the full set). For this subset, the mean location error was 106 meters, with displacements of 96, 122 and 245 meters for the 50th, 67th and 95th percentiles, respectively. The CDF of the Cell ID (CID) location method [8] is also shown for comparison.

A similar analysis was carried out in Area 2, using the RF-FING+RTD-PRED with \(r' = 200\) m, \(N=3\) and \(K=2\). The mean location error was 876 meters, with displacements of 745, 969 and 1982 meters for the 50th, 67th and 95th percentiles, respectively. The subset of NMRs were BS\(_{PM}\)=BS\(_{P}\) had 6123 NMRs (78% of the full set). As in Area 1, the algorithm provided a much better accuracy for this subset, achieving a mean location error of 705 meters, with displacements of 639, 850 and 1567 meters for the 50th, 67th and 95th percentiles, respectively. The higher location errors in Area 2 are due to the much lower GSM cell density in this area in comparison to Area 1, as shown in Tab. I.

Fig. 1 shows the cumulative distribution function (CDF) of the location error achieved by RF-FING+RTD-PRED considering all valid collected NMRs in Area 1, with \(r' = 10\) m, \(N=2\) and \(K=2\).
these results, except the 50th percentile location error in [10] (94 m), which was 22% lower than the one achieve by RF-FING+RTD-PRED in the full set (119 m). However, if only the subset is considered, RF-FING+RTD-PRED yields approximately the same accuracy (96 m).

The accuracy degradation due to the use of an out-of-date CDB has been demonstrated in previous works. In [12], it resulted in an increase of location mean errors from the 100-200 meters range to the 250-350 meters range. In [10], it increased the 50th percentile error from 94 to 245 meters. In DCM methods using propagation modeling such degradation is less likely to occur, as no drive tests are needed to update the CDB.

V. SIMULATIONS IN A WCDMA 850 MHZ NETWORK

V-A. Simulation Scenario

To evaluate the RF-FING+RTD-PRED performance in a wide-band system, a simulation was carried out in a WCDMA FDD (Frequency Division Duplex) 850 MHz network with a 1:1 overlay over the GSM network in Areas 1 and 2. The selected overlay ratio and frequency band allowed the use of the same predicted RSS map built for the GSM test. Therefore, the predicted RF fingerprint (RF-FINGP) CDB and the measured RF fingerprints (RF-FINGM) collected during the GSM test were used, after replacement of the TA values by the proper RTT (Round Trip Time) values. The predicted RTT (RTTpr) of each RF-FINGP was calculated using Eq. (2). If the processing time at the MS is assumed to be negligible, the “measured” RTT (RTTm) of each RF-FINGM can be defined as:

$$\text{RTT}_m = \frac{2\tau_0 + \tau_{md} + \tau_{mu}}{\beta s}$$

where \(\tau_0\) is the propagation delay from the Node B to the MS in LOS conditions; \(\tau_{md}\) and \(\tau_{mu}\) are the mean excess delays due to NLOS propagation in the downlink and uplink, respectively; \(s\) is the WCDMA symbol rate and \(\beta\) is the oversampling factor. Oversampling factors down to 1/16 of the chip period are supported [9]. However, sampling rates higher than 4 times the chip rate are unusual in practical scenarios [17].

Due to the WCDMA FDD duplex separation (45 MHz at the 850 MHz band), multipaths in the uplink and downlink are a priori assumed to be independent at any given MS position. Therefore, for each RTTm, the mean excess delays \(\tau_{md}\) and \(\tau_{mu}\) were treated as independent random variables. They are always positive, as NLOS produces a positive bias in RTT estimation. Both in the uplink and downlink, the mean excess delay \(\tau_m\) is lognormally distributed at any given Node B to MS distance [18], as follows:

$$\tau_m = kT_0D^\epsilon X$$

where \(k\) is the proportionality constant between \(\tau_m\) and \(\tau_{ms}\) (root mean square delay spread), \(T_0\) is the \(\tau_{ms}\) median value at \(D=1\) km, \(D\) is the distance in kilometers between the Node B and the position where the NMR was collected, and \(\epsilon\) is a constant. \(X\) is a lognormal random variable at distance \(D\), such that \(\ln(X)\) is normally distributed with zero mean and standard deviation \(\sigma\). Parameters \(T_0, k, \epsilon\) and \(\sigma\) are environment dependent. Typical values for those parameters in macrocell urban and suburban environments [18][19] are shown in Tab. II, as well as the values used in this work for the simulations in Areas 1 (Urban) and 2 (Suburban).

After replacing the downlink and uplink mean excess delays in Eq. (5) by the expression in Eq. (6), Eq. (5) can be rewritten as:

$$\text{RTT}_m = \frac{2\tau_0 + kT_0D^\epsilon (X_d + X_u)}{\beta s}$$

where \(X_d\) and \(X_u\) are independent lognormal random variables at distance \(D\), both with zero mean and standard deviation \(\sigma\).

To achieve reliable MS positioning results, a Monte Carlo simulation was run. At each realization, lognormally distributed values were randomly selected for \(X_d\) and \(X_u\), and the RTTm for each NMR was calculated as defined in Eq. (7). The final location error for each NMR was obtained by averaging its location error over all runs. Each simulation comprised 50 realizations. This proved to be enough to provide convergence, since two independent simulations of 50 runs each yield approximately the same location error CDF for any given NMR.

V-B. Simulation Results

Fig 2 shows the CDF of the location error achieved by RF-FING+RTD-PRED in Areas 1 and 2. The parameters \(N\) and \(K\) in each area have the same values used in the GSM test. The same applies to \(r^4\) in Area 1 (10 m). However, in Area 2, a smaller value was selected to \(r^4\) (25 m). The \(r^4\) value used in the GSM test (200 m) in Area 2, even though adequate for the TA spatial resolution (554 m), was an order of magnitude above the RTT spatial resolution (78 m). The use of a pixel corresponding to a smaller area in the WCDMA simulation in Area 2 resulted in accuracy improvement, even in the presence of NLOS propagation.

In Area 1, RTT oversampling in presence of NLOS has not improved the algorithm’s accuracy. In Area 2, due to the larger pixel size (25 x 25 m²), oversampling has not been tested. Therefore, the reported results in both areas were achieved with a sampling rate equal to the WCDMA chip rate.

In Area 1, the algorithm achieved a mean error of 144 meters in the full set. This result indicates no improvement in relation to the GSM results. However, if only the NMR subset where BS_P=BS_M is considered, the mean error reduces to 80 meters, with displacements of 68, 89 and 181 meters for the 50th, 67th and 95th percentiles, respectively. The same applies to Area 2, where the mean error for the full set was 817 meters, while for the subset it was 568 meters, with location errors of 459, 680 and 1377 meters for the abovementioned percentiles.

Due to the narrower RTT spatial resolution in comparison to TA, the relative accuracy gain between the NMR full set and subset is much higher in WCDMA. In the GSM test, the reduction of the mean location error in the subset in comparison to the full set was 10% in Area 1 and 19% in Area 2. In the WCDMA simulation, this reduction was 44% in Area 1 and 30% in Area 2. However, by the same token, a mismatch between BS_M and BS_P degrades accuracy more in WCDMA than in GSM. This mismatch might be caused by poor propagation modeling, small scale variations of the received signal or MS measurement inaccuracy. The first impairment can be mitigated by means of propagation model tuning. The second is reduced with time averaging of the received signals (in GSM, the averaging period is approximately 480 ms). The last impairment is partially reduced with the insertion of the \(\delta\) parameter in Eq. (4).
V-C. Comparison with Previous Results

A DCM method using Pilot Correlation has been proposed in [20]. It uses a CDB built from pre-measured WCDMA pilot strength samples and has achieved an accuracy of 90 and 190 meters in a dense urban area for the 67th and 90th percentiles, respectively. In the WCDMA simulation, RF-FING+RTD-PRED has achieved similar results in the NMR subset where BS<sub>P</sub>=BS<sub>M</sub>, with errors of 89 and 140 for those same percentiles. Both methods have met the FCC E911 (Enhanced 911) accuracy requirements for network based location methods [1].

VI. CONCLUSIONS

This paper proposes a fast DCM localization algorithm using CDBs built from propagation modeling. The algorithm can also be employed with CDBs built from field measurements, but the use of predicted RSS maps justify itself by making CDB updating faster and less expensive. The correlation space, which is the set of pixels where the correlation process will be carried out to produce a location estimate, is initially restricted to the predicted best server area of the reported serving sector. The RTD information is then used in a novel way to further restrict this confidence region, thereby enhancing accuracy and reducing computational load. The algorithm performance was evaluated by means of drive tests in a GSM network and a semi-empirical simulation in a WCDMA network. It reached reasonable levels of accuracy, especially when the best server was correctly predicted. In such conditions in a WCDMA network in an urban area, the proposed algorithm met the E911 accuracy requirements for network based location methods [1].

VII. REFERENCES


[3] European Telecommunications Standard Institute, “ETSI TS 101 724 v8.9.0 (2004-06) - Digital telecommunications system (Phase 2+); Location Services (LCS); Functional description; stage 2 (3GPP TS 03.71 version 8.9.0 Release 1999),” 2004.


