Experimental Assessment of a Novel Geolocation Algorithm Based on OTDOA Technique in Real UMTS Networks

Juan Antonio García-Fernández, Antonio Jurado-Navas, and Mariano Fernández-Navarro
Department of Communications Engineering, University of Málaga, 29071 Málaga, Spain
Email: {jagf, navas, mariano}@ic.uma.es

Carlos Úbeda
Ericsson E-28033, Madrid, Spain
Email: carlos.ubeda@ericsson.com

Abstract—In this paper, we propose a new tool based on Observed Time Difference of Arrival (OTDOA) technique in order to estimate both the position of the different User Equipments (UEs) and the Relative Time Differences (RTDs) between Node Bs in real UMTS networks from measurements reports providing the TM parameter and the locations of the sites. This technique results in a non-linear least squares estimation problem which is solved by employing an iterative method based on the Levenberg-Marquardt (LM) algorithm and taking advantage of a particular and frequently found spatial geometry of sites: the star configuration. Modifications proposed in this paper increase its inherent accuracy while maintaining a reduced computational complexity, a fast convergence and a high robustness. Obtained results employing real data from various drive tests measured in different locations are compared with the performance of a standard LM algorithm run on these same measurements. We show that remarkably high geolocation and RTD accuracy is achieved considering only filtered events beforehand.

Index Terms—Levenberg-Marquardt, measurements reports, geolocation, TDOA, UMTS, nonlinear system of equations, RTD

I. INTRODUCTION

In the Third Generation (3G) mobile communication systems, geolocation of the User Equipments (UEs) plays an essential role for Location Based Services (LBS) [1] and Location Assisted Network Management (LANM) such as Location-Aided Handover (LAH) [2]. In the last few years, many efforts have been performed to adopt an implementation of the Time Difference of Arrival (TDOA) technique in downlink. Hence, the 3rd Generation Partnership Project (3GPP) standardized network-based UMTS location scheme known as Observed Time Difference of Arrival (OTDOA) [3]. It is specifically designed for UMTS, and based on the information reported in Measurements Reports (MRs) from UEs to the UMTS Terrestrial Radio Access Network (UTRAN). This technique utilizes multilateration [4] defined as the positioning process given by the points representing the intersection of hyperbolas whose foci are the sites. Through this paper, the terms “node B”, “base station” and “site” are considered equivalent.

However, OTDOA technique requires the Relative Time Difference (RTD) between base stations to be known. This is an important drawback because the UTRAN transmitters are unsynchronized. Furthermore, the RTD varies with time as a direct consequence of clock drift on each Node B [3]. As a result, fixed Location Measurement Units (LMUs) are required to perform timing measurements of all the local transmitters, obtaining the RTDs which are stored in databases and used later for the geo-location of the UEs. Nevertheless, it would be advantageous not to depend on the deployment of LMUs since, among other reasons, this deployment derives in a high and unessential economic cost from the point of view of the telecommunication operators. A solution supported by the Round Trip Times (RTTs) was proposed in [5], notwithstanding the RTT is not always available in MRs.

3GPP defines a type of event called MR to be performed by UEs, considering that events are an indication that something of interest has occurred; so for instance, a change in the pilot signal level or a soft handover procedure. Thus, here we concentrate on determining the RTDs and events positions by profiling only the $T_M$ measurements provided by the traces of the MRs [6], i.e. without the need of LMUs. In this sense, the $T_M$ is a parameter that represents the subframe offset remaining when the integral frame offset is subtracted from the total offset between a frame of a Dedicated Physical Channel (DPCH) from the current base station and its associated frame of the Primary Physical Common Control Channel (P-CCPCH) from a neighboring base station. The $T_M$ parameter has a resolution of one chip and a range of from 0 to 38399 chips [7].

The resulting system of equations can be solved with the aid of the Levenberg-Marquardt (LM) algorithm. This iterative procedure has been proven to give the best trade-
off between accuracy and computational complexity [8], [9] and it will be supported by the selection of a smart spatial geometry of the sites [10] as well as a filtering stage prior to processing the events [6]. The scope of this paper is to describe a tool based on the formulation of the nonlinear system of equations.

The paper is organized as follows. Section II reviews the formulation of the nonlinear system of equations. Section III presents the manner to process and filter the events from MRs. Section IV describes how the total area where Node Bs are situated is divided into small scenarios formed by 4 sites. The modified Levenberg-Marquardt algorithm intended to solve the system of equations is introduced in Section V. Results with real data are shown in Section VI. The paper ends with concluding remarks in Section VII.

II. NONLINEAR SYSTEM OF EQUATIONS

In the multilateration technique, commonly only three sites are required since the intersection of any two hyperbolas defines the point corresponding to an unknown position (xy-coordinates) i.e., each pair of sites determines a different hyperbola taking one of them as a reference. Nevertheless, if the RTDs are also considered as unknowns, then the minimum number of sites required is four as can be seen in Fig. 1. The equations describing this behaviour are deduced in [6], [8], being the resulting system for 4 sites as follows:

\[
\begin{align*}
T_M[2,1] &= \rho \cdot d_{[1,2]} + \text{RTD}[2,1] \\
T_M[3,1] &= \rho \cdot d_{[1,3]} + \text{RTD}[3,1] \\
T_M[4,1] &= \rho \cdot d_{[1,4]} + \text{RTD}[4,1] \\
T_M[2,1] &= \rho \cdot d_{[1,2]} + \text{RTD}[2,1] \\
T_M[3,1] &= \rho \cdot d_{[1,3]} + \text{RTD}[3,1] \\
T_M[4,1] &= \rho \cdot d_{[1,4]} + \text{RTD}[4,1] \\
T_M[2,1] &= \rho \cdot d_{[1,2]} + \text{RTD}[2,1] \\
T_M[3,1] &= \rho \cdot d_{[1,3]} + \text{RTD}[3,1] \\
T_M[4,1] &= \rho \cdot d_{[1,4]} + \text{RTD}[4,1]
\end{align*}
\]

where the operator X [a, b] = X [a] - X [b]. Moreover, \( T_M[a,b] \) is the estimate of the difference of \( T_M \) parameters related to the hyperbola formed by the sites \( a \) and \( b \) for the \( i \)-th event and \( \rho \) is the meter-to-chips conversion parameter (1 m corresponds to 0.0128 chips). Finally, \( d_{[a,b]} \) is the distance difference and RTD \( [a,b] \) is the relative time difference between sites \( a \) and \( b \). It is also worth noting that, without loss of generality, we use site 1 as the reference Node B.

Hence, from (1), at least three events reported by the same four sites are necessary to generate a system of equations with as many unknowns as equations. Additionally it is possible to add redundancy to the system by including more events with the aim of increasing the accuracy, the reliability and the robustness. Thus, from the fourth event, each additional one introduces two new variables (its position in cartesian coordinates) and three new equations, since the RTDs involving the process are the same for all events contained in a particular scenario.

III. READING AND PROCESSING OF MRS

In the first step, the essential input data for the tool are extracted from MRs or added by our tool: thus, identifiers of the events (EventIDs), \( T_M \) parameters measured and sectors where such measurements were made. The next step is grouping reports from sectors within a site as follows [6]:

\[
\hat{T}_M[i] = T_M[i] + 256 \cdot \text{Tcell}[i]
\]

where \( T_M[i] \) is the normalized \( T_M \) parameter, \( i \) is the site to which sector \( i \) belongs to, and \( \text{Tcell} \) is a known network parameter that relates the synchronization references of the different sectors within a site.

Now we can relate events to sites. However, some events may have been reported by two sectors of the same site but with different \( T_M \) parameters. This contradiction is owing to measurement errors produced by, for instance, multipath. Thus, in order to avoid corrupt events, we filter out those whose \( T_M \) parameters for a same site differ in more than 3 chips. Otherwise, the \( T_M \) values are averaged.

Finally, both events that are not reported by enough sites and sites that are not reported by enough events are filtered. The cut-off thresholds could vary according to the area to be analyzed provided that, at least, one event is reported by four sites and one site is reported by a minimum of three events, as indicated in Section II. Hence, our tool filters all sites that are not reported, at least, by seven events. Thus, redundancy in scenarios is achieved.

IV. DIVISION OF SCENARIOS

The division of the total area into different scenarios is made in an efficient and smart way since there exists a huge number of possible candidates. First, the sites are sorted in a decreasing order according to the reported
events, i.e. the sites most reported have priority. Next, such sorted sites are taken in turn as the reference site to accomplish a search for the best scenarios. The procedure consists of the following steps:

- **Step 1:** Define the radius that delimits the area where neighbour sites can be found (an example is shown in Fig. 2). The value of the radius must consider what kind of environment is being examined; for instance, urban areas should have radii smaller than sub-urban areas. Thus, a parameter indicating the maximum number of neighbour sites is required with the purpose of providing flexibility. By proceeding so, if the value of this parameter is exceeded, then the more distant neighbour sites from the reference one are ignored.

- **Step 2:** Generate all possible combinations of sites in sets of 4 (the reference site and three neighbour sites).

- **Step 3:** Apply filtering criteria to discard scenarios. These criteria assume that the values of $T_M[a; b]$ are constrained within a limited range: two times the intersite distance, i.e., the distance between two sites. In a general case, the hyperbolas generated by the foci placed at the sites have a range from $-|d[a; b]|+RTD[a; b]$ to $|d[a; b]|+RTD[a; b]$. This range is which we call a “band”. Fig. 3 shows three hyperbolas corresponding to three UEs located in different positions of the band between the sites. The UE called ‘A’ is on the left side of the band, and thus the difference distance has a negative value, with a $T_M$ value below the RTD. However, the UE called ‘C’ on the right side gives a difference distance positive, and a $T_M$ value above the RTD. An intermediate situation would be the position of the UE called ‘B’, placed just in middle of both sites and yielding a $T_M$ value equal to the RTD. Then, the extreme cases are determined by UEs placed at the foci from which the range for the $T_M$ values can be deduced, as indicated above. Thus, a first criterion is established by calculating the size of the band defined by each pair of sites. This is performed by subtracting the maximum and the minimum $T_M$ values, which are taken from among events taking part in that scenario. Then, considering that the ideal size of a band is two times the intersite distance, we can easily calculate the percentage of occupation of each band. If the result for any band exceeds 10% of two times the intersite distance, then the scenario is discarded since it is considered that the measurements add too much error. A second criterion is associated with the distribution of the events on the band. A complete and homogeneous distribution offers more information to the system of equations, increasing considerably the accuracy of the final results. Hence, if none of the bands is more than 50% occupied, the scenario is discarded.

- **Step 4:** Sort the scenarios in accordance with its reliability. The space distribution of the sites determines a particular geometry that may affect the accuracy of the UE geo-location estimate. A star topology (displayed in Fig. 4) has been proved as an efficient and robust geometry to avoid local minima [10]. Hence, scenarios with this geometry are weighted by a factor to assign them a priority factor. Apart from this feature, if the bands between the sites of a scenario have a high percentage of occupation, for instance, over 75%, then this scenario is also weighted in a positive manner.

**Figure 2.** Selection of neighbor sites.

**Figure 3.** Behavior of the hyperbolas generated by three UEs in function of its position between the sites. $T_M, i \in A, B, C$ represents the $T_M$ parameter measured by each UE, where $d[i; 1, 2], i \in A, B, C$ is the distance difference between the UE $i$ and the sites #1 and #2.

**Figure 4.** Sites placed in a star-geometry formed by four sites named ‘A’, ‘B’, ‘C’, ‘D’, with ‘A’ being the site of reference. Angles between two of these sites ($\alpha, \beta, \gamma$) must be between 100º and 140º i.e., 1.75 and 2.44 rads. On another note, the distance between the most distant site and the nearest site with regard to the site of reference, ‘A’, should not exceed a 3:1 proportion [10].
• Step 5: Limit the number of scenarios to be simulated. The aim is to reduce the computational load selecting, among the most reliable scenarios, the minimum number of them that allow all sites participate in, at least, one scenario. For example, if the scenarios [Site-A Site-B Site-C Site-D] and [Site-B Site-C Site-E] were simulated, then the scenario [Site-A Site-B Site-D Site-E] would be discarded and the RTD[D, E] could be obtained by means of linear combination of the two previous ones.

V. ITERATIVE METHOD

Once the scenario is selected, the nonlinear system of equations must be solved. We proposed a novel modified LM algorithm for the RTDs and position estimates.

A. Adding Redundancy

Before running the numerical method, it is possible to include redundancy to the system by employing events reported by only three of the four sites. To this end, the RTDs variables are fixed to the values obtained in the previous iteration of the LM algorithm. In such a context, each event reported by three sites introduces two unknowns (Cartesian coordinates of its position) and two equations, which can be written as

\[
\begin{align*}
\hat{T}_{\text{d}1 \text{e}1} & = \rho \cdot d_{\text{a}1} + \text{RTD}[2,1] \\
\hat{T}_{\text{d}1 \text{e}1} & = \rho \cdot d_{\text{a}1} + \text{RTD}[3,1] \\
\hat{T}_{\text{d}1 \text{e}1} & = \rho \cdot d_{\text{a}1} + \text{RTD}[2,1] \\
\hat{T}_{\text{d}1 \text{e}1} & = \rho \cdot d_{\text{a}1} + \text{RTD}[4,1] \\
\hat{T}_{\text{d}1 \text{e}1} & = \rho \cdot d_{\text{a}1} + \text{RTD}[3,1] \\
\hat{T}_{\text{d}1 \text{e}1} & = \rho \cdot d_{\text{a}1} + \text{RTD}[4,1]
\end{align*}
\]

where n1, n2, n3 represent events that are reported by sites #1–#2–#3; #1–#2–#4; and #1–#3–#4, respectively.

B. Modified Levenberg-Marquardt Algorithm

Levenberg-Marquardt is a gradient descent hybrid method between Steepest Descent (SD) and Gauss-Newton (GN) methods. GN algorithm provides quadratic convergence, but it may diverge if initial values are very distant from the global minimum. On a different matter, the SD method searches for a minimum based only on the first derivatives of the function, ensuring the convergence in a linear manner.

LM suggested an improvement with respect to the GN method with the purpose of yielding a more robust implementation [8], [9]. It is based on a damping parameter, \( \lambda \), which is added to the diagonal elements of the pseudo-Hessian matrix, \( AA^T \), in order to make it always invertible. The parameter \( \lambda \) may be adaptive, changing in each iteration in such a manner to fluctuate between the SD and GN techniques. GN method is approximated by a small \( \lambda \) value, whilst a larger value approaches to the SD algorithm, i.e., LM scales the gradient to the curvature so that there is larger movement in the directions where the gradient is minor. This circumstance avoids slow convergence along the direction of small gradients [11].

Therefore, from [8], [9], the procedure followed by the classical LM algorithm is given by

\[
\begin{align*}
\hat{z}_{i}^{k+1} & = \hat{z}_{i}^{k} + \left( A^T A + \lambda^k I \right)^{-1} A^T \hat{T} \\
& = \left( \hat{T}_M - \hat{T}_m (x, y, \text{RTD}) \right)
\end{align*}
\]

where \( I \) is the identity matrix, \( A \) denotes the Jacobian matrix, and \( \hat{z}_i \) represents the vector of unknowns.

Algorithm 1 Modified Levenberg-Marquardt

1: \( k \leftarrow 0 \)
2: \( \lambda^{(0)} \leftarrow 1 \)
3: \( v^{(0)} \leftarrow 2 \)
4: \( \text{uplam}^{(0)} \leftarrow 0.5 \)
5: \( \alpha \leftarrow 0.65 \)
6: \( \text{upJ} \leftarrow 1 \)
7: \( \text{networkRadius} \leftarrow 3000 \)
8: while \( k \leq 100 \) do
9: if \( \text{upJ} = 1 \) then
10: Calculating \( \hat{T}_m \)
11: Calculating the Jacobian matrix
12: if \( k = 0 \) then
13: \( \text{residue}^{(0)} \leftarrow \left\| \hat{T}_m - \hat{T}_u \right\| \)
14: end if
15: end if
16: if \( \text{residue}^{(k)} \leftarrow \left( \hat{T}_m - \hat{T}_u \right) \left( \hat{T}_m - \hat{T}_u \right) \)
17: \( \text{residue}^{(k)} \leftarrow \alpha \cdot \text{residue}^{(k)} \)
18: \( \alpha \leftarrow 0.8\alpha \)
19: if \( \text{upJ} = 1 \) then
20: \( \text{RTD} \) values are not updated
21: end if
22: end if
23: Calculating the new residue
24: \( \rho \leftarrow \frac{\text{residue}^{(k)} - \text{residue}^{(k-1)}}{\text{residue}^{(k-1)}} \)
25: if \( \text{residue}^{(k)} < \text{residue}^{(k-1)} \) then
26: Updating the solution and the residue
27: \( \text{uplam} \leftarrow \frac{1}{(2-\alpha)^2} \)
28: end if
29: if \( \text{uplam} > 1 \) then
30: \( \text{uplam} \leftarrow 0.99 \)
31: end if
32: \( \lambda^{(k+1)} \leftarrow \lambda^{(k)} \cdot \text{max}(0.5, \text{uplam}) \)
33: \( \lambda^{(k+1)} \leftarrow 1.1 \)
34: \( \text{upJ} \leftarrow 1 \)
35: else
36: \( \lambda^{(k+1)} \leftarrow \lambda^{(k)} \cdot \text{max}(0.5, \text{uplam}) \)
37: \( \lambda^{(k+1)} \leftarrow 1.2\lambda^{(k)} \)
38: end if
39: \( \text{upJ} \leftarrow 0 \)
40: if \( \lambda^{(k+1)} > 1000 \) then
41: \( \lambda^{(k+1)} \leftarrow 1.01 \)
42: \( \text{upJ} \leftarrow 1 \)
43: end if
44: \( k \leftarrow k + 1 \)
45: end while

If distance between any estimated position and the mass center is more than \( 2 \text{networkRadius} \), then the scenario is discarded by divergence.
Now, in order to improve the performance of this numerical method, several modifications to the LM algorithm presented in [9] have been proposed, as described in Algorithm 1. In such a context, the \( \lambda \) parameter is especially relevant since it provides high robustness against local minima. This parameter takes into account the estimates of the previous iteration to avoid abrupt changes between two consecutive iterations, which could be caused by local minima. Furthermore, it is also worth noting that the adjustment of some parameters and values has been carried out with help of several drive tests realized in a European sub-urban region. Consider that, slight loss of accuracy may be obtained in non-suburban areas.

Regarding the initial values or seeds, it has been heuristically checked that placing all events at the center of mass of the sites gives better outcomes. On another note, the initial values for the RTDs are obtained by employing the concept of band represented in Fig. 3. The midpoint of the bands is obtained by subtracting the minimum \( M_T \) from the maximum \( M_T \) with the resulting value divided by two. This value should be ideally the exact RTD. Therefore, if the bands are highly occupied, then the initial value will be a good approximation.

C. Applying of Known RTDs

Lastly, when all selected scenarios have been solved, the values of RTDs calculated more than once are averaged, and all possible pairs of RTDs are determined by means of linear combinations. In this manner, we attain approximately a synchronized UMTS network for the analyzed region.

At this point, all RTDs have been obtained. Then, the system of equations presented in Eq. (1) becomes overdetermined, which means that one event reported by three sites is enough to generate a system of equations resolvable, i.e., constituted by two equations and two unknowns given by its geographical position in Cartesian coordinates. By proceeding in this term, we can employ the modified LM algorithm described in previous subsection to locate, for instance, events reported by three sites.

VI. RESULTS

In order to assess the performance of the proposed tool, three graphs were produced from real data collected in drive tests. These drive tests were made in a European sub-urban area and a North-American urban area. It is important to keep in mind that both areas may have a different impact on the frequency and the characteristics (potential multi-path, number of reported sites, etc) of the collected MRs, and hence, on the achievable accuracy.

Fig. 5 shows a geo-location cumulative distribution function (cdf) comparing the accuracy of the proposed tool based on the modified LM algorithm and the classic LM procedure. These results correspond to the events reported by 4 sites that have been solved jointly to the RTDs. We can see that results obtained by the modified LM algorithm are the most accurate for both areas, with a noticeable improvement in the American area. This fact is somewhat expected since the parameters of both methods are optimized for the European region. Nevertheless, the modified LM algorithm is much more flexible and adaptive to solve different types of scenarios. For instance, the 75th percentile error has been reduced 200 m in the European area with the new proposed tool, whereas in the North-American area has been reduced over 1000 m.

For the sake a better observation of the improvements introduced by the modified LM algorithm results from a concrete example of a real scenario are displayed on a map of the North-America area in the Fig. 6 (see next page). It can be seen that the estimated positions by the modified LM are closer to the real positions, keeping the distance error below 140 m in all cases.

On the other hand, Fig. 7 represents the comparison in accuracy for the events reported by three sites, i.e. once the RTDs are set, the position of these events is solved one by one. In this case, obtained results in the European scenario with both strategies are precise and very similar, which means that the obtained RTDs are comparable. However, in the American scenario, the proposed tool provides better performance due to the fact that RTDs are closer to the real values and therefore, the problem of divergence is mitigated.
VII. CONCLUSIONS

In this paper, a tool to jointly estimate the position of the different UEs and the RTDs between sites has been proposed. This tool has been divided into three main parts: first, a processing and a filtering stage of the events read from MRs; second, the division of the whole area into small scenarios of four sites, and third, the application of an iterative method based on a modified LM algorithm. Furthermore, we have employed real networks to assess the performance of our proposed tool, which has demonstrated a remarkably improvement of accuracy, convergence, adaptation and flexibility in both analyzed areas. The principal reasons for this improvement are: the elimination of events with too much error that could introduce negative effects in the system of equations, a smart selection of the scenarios according to its geometry and the distribution of the events, and the use of enhanced numerical method.

ACKNOWLEDGMENT

This work was supported by Ericsson.

REFERENCES


Juan Antonio García-Fernández was born in Málaga, Spain 1988. He received the M.S. degree (2012) in Telecommunication Engineering from the University of Málaga (Spain) with honors. Since 2012, he has been with Ericsson Spain, where he is working on geolocation tools for third generation mobile networks. His current research interest includes optimization of mobile networks and chipless RFID.

Antonio Jurado-Nava received the M.S. degree (2002) and the Ph.D. degree (2009) in Telecommunication Engineering, both from the University of Málaga (Spain). He joined the Department of Communications Engineering at the University of Málaga in 2004. In 2011, he became an Assistant Professor in the same department. Since 2012, he has been with Ericsson Spain, where he is working on geolocation tools for third generation mobile networks. His current research interests include mobile communication systems and channel modeling in addition to atmospheric optical communications, adaptive optics and statistics.
Mariano Fernández-Navarro received his M.S. in Telecommunication Engineering from the Polytechnic University of Madrid in 1988 and the Ph.D. degree from the University of Málaga, in 1999. He is on the staff of the Communications Engineering Department at the University of Málaga since 1992, after 3 years as design engineer at Fujitsu Spain S. A. His research interests include optimization of radio resource management for mobile networks and location-based services and management.

Carlos Úbeda received his MSc. degree in Telecommunication Engineering from Miguel Hernández University of Elche (Spain) in 2006. After finishing his Master’s Thesis at Aalborg University (Denmark), he worked as External Researcher for Nokia Siemens Networks in Aalborg. In 2008 he joined Optimis as R&D Engineer in Málaga. Since 2011 he is a R&D System Developer for Ericsson in Madrid. His main research interests include radio resource management and radio network optimization.