Localization in 3GPP LTE Based on One RTT and One TDOA Observation

Kamior Radnosrati, Carsten Fritsche, Fredrik Gunnarsson, Fredrik Gustafsson and Gustaf Hendeby

The self-archived postprint version of this journal article is available at Linköping University Institutional Repository (DiVA):
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-165188

N.B.: When citing this work, cite the original publication.

Original publication available at:
https://doi.org/10.1109/TVT.2020.2968118
Copyright: Institute of Electrical and Electronics Engineers (IEEE)
http://www.ieee.org/index.html
©2020 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
Localization in 3GPP LTE based on one RTT and one TDOA observation

Kamiar Radnosrati, Carsten Fritsche, Fredrik Gunnarsson†, Fredrik Gustafsson, Gustaf Hendeby
Department of Electrical Engineering, Linköping University, Linköping, Sweden
Email: {firstname.lastname}@liu.se
† Ericsson Research, Linköping, Sweden, Email: fredrik.gunnarsson@ericsson.com

Abstract—We study the fundamental problem of fusing one round trip time (RTT) observation associated with a serving base station with one time-difference of arrival (TDOA) observation associated to the serving base station and a neighbor base station to localize a 2-D mobile station (MS). This situation can arise in 3GPP Long Term Evolution (LTE) when the number of reported cells of the mobile station is reduced to a minimum in order to minimize the signaling costs and to support a large number of devices. The studied problem corresponds geometrically to computing the intersection of a circle with a hyperbola, both with measurement uncertainty, which generally has two equally likely solutions. We derive an analytical representation of these two solutions that fits a filter bank framework that can keep track of different hypothesis until potential ambiguities have been resolved. Further, a performance bound for the filter bank is derived. The proposed filter bank is first evaluated in a simulated scenario, where the set of serving and neighbor base stations is changing in a challenging way. The filter bank is then evaluated on real data from a field test, where the result shows a precision better than 40 m 95 % of the time.

Index Terms—CRLB, filter bank, jump Markov models, round trip time, time-difference of arrival, LTE, localization

I. INTRODUCTION

Although Global Navigation Satellite System (GNSS) systems, e.g. global positioning system (GPS), are capable of determining the position of an object with a few meters accuracy in outdoor environments, the robustness of GPS-based methods is always restricted by the availability of GPS signals. Additionally, GPS may be too complex, costly and/or battery consuming for low-complexity devices, while positioning based on cellular measurements can be seen as an asset that comes with the communication needs.

Localization in cellular networks is based on opportunistic timing measurements that vary between different generations. Known reference signals at the receiver are used by the mobile station (MS) to compute the time of arrival (TOA) in relation to an MS specific time reference by correlating the known signal with the received one. In a network with several synchronized base stations (BS) (or at least with known/estimated time offsets), the mobile MS can compute the TOA in relation to an MS specific time reference from each one of them and then form a time difference of arrival (TDOA) measurement for each pair of BS. Each TDOA corresponds to one hyperbolic function, and a unique position can be computed from at least two TDOA pairs collected from three BSs in 2-D space using multilateration techniques. This principle has been extensively studied, where for instance [1]–[5] derive closed-form solutions. The hyperbolas are quadratic functions in position, so to avoid nonlinearities [6] transforms the problem into a set of linear equations and propose a method that is suitable for real-time implementations. To increase the reliability of TDOA positioning algorithms, [7] proposes a model that is applicable either if the MS is located in the near-field or far-field of the BSs. A closed-form two-step non-line-of-sight (NLOS) localization technique for cellular networks by means of TOA measurements is developed in [8]. To deal with NLOS conditions, the authors propose a separate ranging step where an unbiased distance estimate is found. The estimated distances are subsequently used for MS localization using trilateration. For a comparative performance analysis of four circular and hyperbolic geolocation methods see [9]. The authors in [10], use round trip time (RTT) measurements and report the positioning field performance in terms of availability, response time, and accuracy. In [11], the MS measurement data, collected during each call/session, are shown to have valuable information on mobile’s performance metrics and also on signal strength and signal to interference plus noise ratio. The obtained information is used in a binary classifier to infer whether the measurement was generated from an indoor mobile or from an outdoor mobile. If the classifier identifies an outdoor MS, the latitude-longitude of the mobile, when the measurement record was generated, is estimated.

A hybrid approach using one RTT and one TDOA measurement for 2-D localization of MS is investigated in this paper. There is also a rich literature on hybrid approaches. For instance, [12] proposes a hybrid kernel based machine learning using received signal strength while [13]–[15] take advantage of additional information obtained from angle of arrival (AOA) measurements. Accurate AOAs require antenna arrays [16]–[18]. Using coarse sector information or AOA from a few antenna elements provides limited localization improvement in practice, as for instance shown in [19]. Fusion of TDOA and RTT (here denoted two-way time of arrival - TW-TOA) has been investigated in [20] using reference signals transmitted from macro and femto BSs in LTE advanced (LTE-A) heterogeneous networks. The authors in [21] use TDOA and frequency difference of arrival (FDOA) measurements and formulate the localization problem as a weighted least squares (WLS) problem. Then, by performing semidefinite relaxation, they obtain a convex semidefinite programming problem to estimate the location. The results of an extensive
study concerning hybrid network/satellite-based localization systems for wireless networks are reported in [22].

Different wireless network standards enable different kinds of ranging mechanisms. See [23] for WiFi fine timing measurement and [24] for Bluetooth. A typical mechanism in cellular network is the time alignment procedure, designed to align the uplinks (device to base station) of different devices at the base station. The mechanism produces a range estimate that is only available to the serving base station. Mechanisms to configure TDOA measurements, which includes providing the MS with assistance data to inform about what reference signals to monitor and what radio resources that are used to transmit these reference signals, also vary with different network standards. 3GPP OTDOA (Observed TDOA) for LTE and narrowband Internet of Things (NB-IoT) devices is based on downlink TOA estimates, reported as the reference signal time difference (RSTD) between pairs of cells [25]. RSTD is the relative time difference between the Evolved Node B (eNB) \( j \) and the reference eNB \( i \) and is calculated as the smallest time difference between two subframe boundaries received from two different eNBs. It is designed with scalability in mind, and the only signaling cost that scales with the number of users is the assistance data provisioning from location server to device and the location measurements reporting from device to location server. The assistance data provisioning corresponds to a list of cells to monitor, where each cell is associated to a reference signal and radio resource configuration. It can also include a limit in the number of cells to report. The measurement reporting corresponds to the RSTD measurements for pairs of cells. In order to limit the signaling cost, it is possible to limit the number of neighbor cells included in the assistance data, and/or to limit the number of neighbors included in the report. In this work, we restrict the number of reported neighbor cells to only one, and assume that this is the cell with strongest received signal, excluding the serving cell.

Fusion of one RTT and one TDOA observation is an interesting problem, that geometrically corresponds to finding the intersection of a circle with a hyperbola. Both these are quadratic functions in the position, so there are algebraically two solutions. In the normal case, both solutions are real-valued and corresponds to two ambiguous positions. There is a special case with a double root, when the circle and hyperbola meet in one point only. Observation noise can imply that the two functions never intersect and in this degenerated case the algebraic solution becomes complex-valued, which can be interpreted as no geometrical solution.

The proposed solution in [20] is based on fusion of multiple RTT and TDOA measurements and uses the extended pedestrian A (EPA) and the extended typical urban (ETU) channel models to simulate the multipath environment and is tailored for indoor to outdoor scenarios. The proposed method in this work, however, is more suitable to an outdoor-only scenario and we do not include counter-measures for multipath in our algorithm. The fusion of RTT and TDOA, collected from two BSs, appears to be rather unexplored problem, with the exception of [26], which investigates the position estimation of a vessel in an automatic identification system using TOA and TDOA measurements. The authors propose a Taylor-linearization of the nonlinear TOA and TDOA equations and derive a least squares solution to the linearized problem. The performance of the estimator is evaluated in terms of geometric dilution of precision (GDOP). In this work, we evaluate the performance of the proposed algorithm in terms of the Cramér-Rao lower bound (CRLB) which takes the effect of different weighting of the variances into account. See [27] for more details.

We first derive the two analytical solutions to the geometrical problem, and then formulate the problem in a statistical signal processing setting with realistic model of observation noise. In this way, we can derive a covariance matrix of the position estimate for both solutions that should reflect the geometry of the BS and MS configuration. However, the ambiguity cannot be resolved from such a snapshot solution.

To resolve the ambiguity, we complement the observation process with a simple motion model for the MS. The two solutions are represented by a discrete mode parameter in an otherwise linear Gaussian state space model. The optimal solution to this problem formulation is given by a bank of an exponentially increasing number of Kalman filters. Each filter corresponds to a sequence of modes for the solution. We limit the complexity using a pruning mechanism to avoid exponential increase of hypotheses.

The remainder of this paper is organized as follows. Sec. II briefly introduces RTT and TDOA measurements in 3GPP LTE. Sec. III illustrates the geometry of the problem and derives a closed-form snapshot solution, while Sec. IV introduces the filter bank framework. Fundamental performance bounds of both the snapshot positioning and the filter bank approach are derived and presented in Sec. V. Sec. VI evaluates the performance of the proposed method for both simulation and real experimental scenarios followed by concluding remarks in Sec. VII.

II. TDOA AND RTT IN 3GPP LTE

This section provides more background on how the TOA and TDOA measurements are computed in the network, and what the typical precision is. Timing-based positioning in cellular systems is based on TOA, RTT and TDOA measurements. TOA is estimated by cross-correlating the received signal with a replica of the transmitted signal waveform, and it is the basis for both TDOA and RTT. Downlink TDOA or OTDOA is estimated as the difference between TOA of downlink signals from two different base stations, typically a neighbor BS and a reference base station. Since the network can post-process the reported measurements to change the reference base station, we assume without loss of generality that the reference base station is the serving base station. RTT can be determined based on the time alignment procedure, where the serving BS provides the MS with an timing advance offset, indicating the start of an uplink frame in relation to a received downlink frame. The serving BS can determine the RTT from the timing advance offset and the difference between uplink reception and downlink transmission in the base station.
while supporting a reasonably short measurement time. The sufficient time diversity and avoiding too much overhead, over the PRS occasion. The periodicity is configured to ensure 1 PRS occasions, where each occasion corresponds to one or time-slot muting pattern, all to suppress intersite interference.

The performance analysis in [6], based on different PRS configurations subject to additive white Gaussian noise, provides results on TOA accuracies. The TOA error standard deviation when the PRS is received at a sufficient SNR, a SNR of −13 dB or better in the considered scenario, is determined as

\[ \sigma_{\text{TOA, 20MHz}} = 2.4 \text{ ns, (corresponding to 0.7 m)} \]  
\[ \sigma_{\text{TOA, 1.4MHz}} = 66 \text{ ns, (corresponding to 20 m)} \]

A similar analysis in [30] evaluated a 10 MHz PRS considering a three-dimensional evaluation scenario used in 3GPP and a realistic MS receiver model. For sufficient SNR (at least −10 dB), the determined TOA error standard deviation is

\[ \sigma_{\text{TOA, 10MHz}} = 27 \text{ ns, (corresponding to 8 m)} \]

These numbers will be used later on in the simulation studies, and only downlink signals received at an SNR of at least −10 dB will be considered useful. Multipath is assumed to be resolved reasonably well so that it is reasonable to assume the same variance along the trajectory. The uplink signal is subject to power control to ensure that it is received at sufficient SNR in the serving BS. Two signal bandwidths will be considered, 10 MHz and 1.4 MHz, corresponding to PRS/CRS and PSS/SSS respectively in the downlink and DMRS or SRS in the uplink.

### III. Snapshot Estimate

Denote the 2-D position of the MS by \( \mathbf{p} = (p_x, p_y)^T \) and the position of BS \( i \) by \( \mathbf{l}_i = (l_{x,i}, l_{y,i})^T \). The distance between BS \( i \) and the MS is then

\[ r_i = \|\mathbf{p} - \mathbf{l}_i\| = \sqrt{(p_x - l_{x,i})^2 + (p_y - l_{y,i})^2}. \]

Assuming that BS 1 is the serving cell and BS 2 is the neighbor cell, the available RTT and TDOA measurements are noisy measurements of \( 2r_1 \) and \( r_{21} = r_2 - r_1 \), respectively. Let \( r = (r_1, r_{21}) \), the first task is to compute the inverse mapping relates \( r \) to the MS position \( \mathbf{p} \).

### A. Noise Free Measurements

To facilitate the derivation, the first step is to temporarily use a local coordinate system, where the two BS are symmetrically located along the \( x \)-axis as illustrated in Fig. 2. This is achieved with a rotation \( \mathbf{\Gamma} \) and a translation \( \mathbf{t} \), \( \mathbf{\tilde{p}} = \mathbf{\Gamma p} + \mathbf{t} \) and \( \mathbf{\tilde{l}_i} = \mathbf{\Gamma l_i} + \mathbf{t} \), such that \( \mathbf{\tilde{l}_1} = (D/2, 0)^T \) and \( \mathbf{\tilde{l}_2} = (-D/2, 0)^T \). Let the solution to the positioning problem in the local coordinates be \( \mathbf{\tilde{p}} \). The inverse transformation is given by \( \mathbf{p} = \mathbf{\Gamma}^{-1}(\mathbf{\tilde{p}} - \mathbf{t}) \). Note that the rotation matrix \( \mathbf{\Gamma} \) is scale preserving, so the distances \( r_i \) are the same in local coordinates, where the bit 1 in the pattern indicates that the base station is transmitting the PRS in the corresponding PRS occasion, and not transmitting anything but mandatory signals (system information, synchronization signals, cells specific reference signal, etc) with bit 0. The TDOA measurements are referred to as reference signal time difference (RSTD) which is the relative time difference between the BS \( j \) and the serving BS \( i \).
coordinates, and the BS separation $D = ||l_1 - l_2||$ is the same in global and local coordinates.

Geometrically, the solution in the noise-free case is given by the intersection of a circle and a hyperbola, defined by

$$\begin{align*}
\left(\bar{p}_x - \frac{D}{2}\right)^2 + \bar{p}_y^2 &= r_1^2, \\
\frac{\bar{p}_x^2}{a^2} - \frac{\bar{p}_y^2}{b^2} &= 1,
\end{align*}$$

where $a^2 = \frac{1}{4}r_{21}^2$ and $b^2 = \frac{1}{4}(D^2 - r_{21}^2)$. Algebraically, the intersections are given by

$$\bar{p} = \bar{g}(r, \delta) = \begin{bmatrix} \bar{g}_x(r_1, r_{21}) \\ \delta \bar{g}_p(r_1, r_{21}) \end{bmatrix},$$

with

$$\bar{g}_x(r_1, r_{21}) = \frac{r_{21}(2r_1 + r_{21})}{2D},$$

$$\bar{g}_p(r_1, r_{21}) = \frac{\sqrt{(2r_1 + r_{21})^2 - D^2}}{2D},$$

where $\delta$ is a discrete parameter representing the possible intersection points. Note that the circle and hyperbola can intersect in zero, one or two positions, where two solutions is the normal case where $r_1 + r_2 > D$ and $\delta \in \{\pm 1\}$. Having one solution is a degenerated case where $r_1 + r_2 = D$ and $\delta = 0$, hence the solution lies on the $x$ axis. The case of no solution, corresponding to $r_1 + r_2 < D$ and $\delta \in \emptyset$, can happen for noisy measurements and may need special care.

### B. Noisy Measurements

Let $\epsilon_{dl,1}$ and $\epsilon_{ul,1}$ denote the TOA estimation error of the serving BS in the downlink and uplink directions, respectively. The error of the neighbor BS has only contributions from the downlink and is denoted by $\epsilon_{ul,2}$. The noisy RTT and TDOA measurements, as defined in the previous section, are thus defined by

$$\begin{align*}
z_1 &= 2r_1 + \epsilon_{dl,1} + \epsilon_{ul,1}, \\
z_2 &= r_2 - r_1 + \epsilon_{dl,2} - \epsilon_{ul,1},
\end{align*}$$

where $\epsilon_{dl,i} \sim \mathcal{N}(0, \sigma_{dl}^2)$ and $\epsilon_{ul,1} \sim \mathcal{N}(0, \sigma_{ul}^2)$, and the considered TOA estimation error standard deviation (1c). It is worth noting that due to reciprocity of the downlink and uplink channel, $\epsilon_{dl,i}$ and $\epsilon_{ul,i}$ might be correlated. However the correlation is milder by the fact that these can be representing different time instants and can be at different frequency bands. The interference and noise is also different, so it can still be seen as a reasonable first order approximation that these quantities are uncorrelated. We define the measurement vector $z = (z_1, z_2)^T$ where $z \sim \mathcal{N}((2r_1, r_{21})^T, R)$ with $R$ given by

$$R = \text{cov}(z) = \begin{bmatrix} \sigma_{dl}^2 & \sigma_{dl}^2 \\
\sigma_{dl}^2 & -\sigma_{ul}^2 & 2\sigma_{dl}^2 \end{bmatrix}.$$
Algorithm 1 Unscented Transformation for \( n \) variate Gaussian Distribution.

**Input:** \( R, g(z - e, \delta) \).

**Output:** \( \tilde{p}(\delta), \Lambda(\delta) \)

1. Define \( \sigma_i \) and \( u_i \) using a singular value decomposition of \( R = U \Sigma U^T \), where \( U^T U = I \) and \( \Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2) \), as

\[
R = U \Sigma U^T = \sum_{i=1}^{n} \sigma_i^2 u_i u_i^T,
\]

where \( u_i \) is the \( i \)-th column of \( U \).

2. Form a set of \( 2n + 1 \) sigma points

\[
\tau^{(0)} = 0, \quad \tau^{(\pm i)} = \pm \sqrt{n + \lambda} \sigma_i u_i, \quad i = 1, \ldots, n.
\]

where \( \lambda \) is a design parameter.

3. Map the sigma points to \( \tilde{p}^{(i)}(\delta) = g(z - \tau^{(i)}, \delta) \) by propagating them through nonlinear function \( g(\cdot) \) for \( i = -n, \ldots, n \).

4. Compute the constant weights \( w^{(n)} \)

\[
w^{(0)} = \frac{\lambda}{n + \lambda}, \quad w^{(\pm i)} = \frac{1}{2(n + \lambda)}, \quad i = 1, \ldots, n.
\]

5. Estimate the mean and covariance of the transformed variable from the mapped sigma points

\[
\mathbb{E}[g(z - e, \delta)] \approx \tilde{p}(\delta) = \sum_{i=-n}^{n} w^{(i)} \tilde{p}^{(i)}(\delta),
\]

\[
\text{cov}[g(z - e, \delta)] \approx \Lambda(\delta) = \sum_{i=-n}^{n} w^{(i)} (\tilde{p}^{(i)}(\delta) - \tilde{p}(\delta)) \times (\tilde{p}^{(i)}(\delta) - \tilde{p}(\delta))^T.
\]

which measurement to use, until enough information has been collected.

A. System models

Models that change behavior based on a mode indicator are called jump Markov models (JMM) when the mode follows a Markov chain. Here we consider the MS to move according to a constant velocity model with the state \( x_t = [x_{xt}, y_{xt}, v_{xt}, v_{yt}]^T \), where \( v_{xt} \) and \( v_{yt} \) are the velocity along the \( x \)- and \( y \)-axis, respectively. The different solutions to the measurements are modeled as different modes, \( \delta_t \), in the measurement equation, yielding the following linear JMM

\[
x_{t+1} = F x_t + \omega_{t+1},
\]

\[
0 = H x_t - \tilde{p}(z_t | \delta_t) + e_t(\delta_t),
\]

\[
\omega_{t+1} \sim \mathcal{N}(0, Q),
\]

\[
e_t(\delta_t) \sim \mathcal{N}(0, \Lambda_t(\delta_t)),
\]

\[
p(\delta_{t+1} | \delta_t) = \Pi_{\delta_{t+1} | \delta_t},
\]

with

\[
F = \begin{bmatrix} I_2 & T I_2 \\ 0_2 & I_2 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},
\]

\[
Q = \sigma^2 \begin{bmatrix} T^2 I_2 & \frac{T^2 I_2}{T I_2} \\ \frac{T^2 I_2}{T I_2} & \frac{T^2 I_2}{T I_2} \end{bmatrix},
\]

where \( \Pi_{\delta_{t+1} | \delta_t} \) is the transition probability in the mode Markov chain at time \( t \), \( \sigma_\omega \) is the process noise standard deviation, and \( T \) is the sampling interval. It is assumed that the

IV. Filter Bank

When using the converted measurements introduced above as measurement in a filter, the different solutions give rise to different modes corresponding to which of the solutions are used as measurement. Hence, the KF does not apply. However, a filter bank can be used, where the idea is to enumerate all possible mode sequences, conditional on each of the different modes apply the KF, and then combine all the filters based on how well they fit to the measurements. The filter bank is an optimal solution to the filtering problem, if all modes are considered and the individual filters are optimal for the conditional problem. Maintaining several possible modes in the filter bank allows for delaying hard decisions about the conditional problem.
mode at time \( t \) is time independent, thus \( \Pi^{S,\delta} = \frac{1}{|\Delta_1|} \), where \( |\delta_1| \) represent the number of modes (measurements) at time \( t \). This corresponds to not being able to beforehand decide or give prior to any of the solutions.

**B. Kalman Filter Bank**

The solution to the filtering problem given by model (10) is a mixture of all solutions to the considered mode sequences, \( \delta_{1:t} = (\delta_1, \ldots, \delta_t) \). Let \( \Delta_t \) denote the set of considered mode sequences, \( \delta_{1:t} \). Then the posterior distribution becomes

\[
p(x_t | z_{1:t}) = \sum_{\delta_{1:t} \in \Delta_t} w_t^{(\delta_{1:t})} p(x_t | \delta_{1:t}, z_{1:t}),
\]

where the mixing probability is \( w_t^{(\delta_{1:t})} = p(\delta_{1:t} | z_{1:t}) \).

In order to correctly represent the posterior distribution, all possible mode sequences \( \delta_{1:t} \) must be present in \( \Delta_t \). The number of possible mode sequences grows exponentially in the number of time steps. Any tractable on-line algorithm hence has to reduce the number of components to be tractable, i.e., limit the number of components in \( \Delta_t \). The number of modes can be reduced either by pruning unlikely mode sequences, and/or by merging similar sequences. See [31], [33] for a discussion about approximations to avoid the exponential growth of \( \Delta_t \).

In this work, a filter bank which prunes low-probability branches is applied to compute the posterior distribution. After each measurement update, all but the \( M \) most likely mode sequences are eliminated. The used filter bank algorithm is given in Algorithm 2.

The result of running parallel filters introduced in Algorithm 2 can be used to construct a set of quadruples \( (\delta_{1:t}, \hat{x}_{1:t}^{(\delta_{1:t})}, P_{1:t}^{(\delta_{1:t})}, w_{1:t}^{(\delta_{1:t})}) \). The surviving modes typically include both the desired solution, the true MS position, and the secondary solution. See Fig. 4, which illustrates the indistinguishable symmetrical solutions arising in the problem.

**Algorithm 2 Kalman filter bank with pruning.**

**Input:**

\[
\Delta_{t-1} \equiv \{ w_{t-1}^{(\delta_{1:t-1})} x_{1:t-1}^{(\delta_{1:t-1})}, P_{1:t-1}^{(\delta_{1:t-1})}; \delta_{1:t-1} \in \Delta_{t-1} \}, z_t
\]

1. **State time update:** \( \forall \delta_{1:t-1} \in \Delta_{t-1} \)

\[
\begin{align*}
\hat{x}_{1:t}^{(\delta_{1:t-1})} & = F_{t-1}^{(\delta_{1:t-1})} x_{1:t-1}^{(\delta_{1:t-1})}, \quad (12a) \\
\tilde{P}_{1:t}^{(\delta_{1:t-1})} & = F_{t-1}^{(\delta_{1:t-1})} T F_{t-1}^{(\delta_{1:t-1})} + Q. \quad (12b)
\end{align*}
\]

2. **Branching:** Where \( \hat{\delta}_1, \ldots, \hat{\delta}_N \) are the possible modes at time \( t \)

\[
\Delta_t = \bigcup_{\delta_{1:t-1} \in \Delta_{t-1}} \{ (\delta_{1:t-1}, \hat{\delta}_1), \ldots, (\delta_{1:t-1}, \hat{\delta}_N) \}. \quad (13)
\]

3. **Weight update:** \( \forall \delta_{1:t-1} \in \Delta_t \)

\[
\begin{align*}
\bar{w}_{t}^{(\delta_{1:t})} & = w_{t-1}^{(\delta_{1:t-1})} \Pi_{t}^{\hat{\delta}_{1:t-1}} N(0 | \hat{y}_{t}^{(\hat{\delta}_{1:t})}, S_{t}^{(\hat{\delta}_{1:t})}), \quad (14a) \\
\tilde{w}_{t}^{(\delta_{1:t})} & = \sum_{\delta \in \Delta_t} \bar{w}_{t}^{(\delta_{1:t})}. \quad (14b)
\end{align*}
\]

where \( \hat{y}_{t}^{(\hat{\delta}_{1:t})} \) and \( S_{t}^{(\hat{\delta}_{1:t})} \) are the predicted measurement and innovation covariance, given the mode sequence \( \delta_{1:t} \), as given by the Kalman filter for the mode.

4. **Pruning:** Reduce \( \Delta_t \) to only the \( M \) most likely components, as given by \( w_{t}^{(\delta_{1:t})} \)

5. **State measurement update:** \( \forall \delta_{1:t} \in \Delta_t \)

\[
\begin{align*}
S_{t}^{(\hat{\delta}_{1:t})} & = H P_{t-1}^{(\hat{\delta}_{1:t})} H^T + A_t(\hat{\delta}_{t}), \quad (15a) \\
K_{t}^{(\delta_{1:t})} & = \tilde{P}_{t-1}^{(\delta_{1:t})} H^T (S_{t}^{(\hat{\delta}_{1:t})})^{-1}, \quad (15b) \\
\hat{x}_{1:t}^{(\delta_{1:t})} & = \hat{x}_{1:t-1}^{(\delta_{1:t-1})} + K_{t}^{(\delta_{1:t})} (\hat{y}_{t}^{(\delta_{1:t})} - H \hat{x}_{1:t-1}^{(\delta_{1:t})}), \quad (15c) \\
P_{t}^{(\delta_{1:t})} & = \tilde{P}_{t-1}^{(\delta_{1:t})} - K_{t}^{(\delta_{1:t})} H \tilde{P}_{t-1}^{(\delta_{1:t})}. \quad (15d)
\end{align*}
\]

6. return \( \Delta_t \), \( \{ w_{t}^{(\delta_{1:t})}, \hat{x}_{1:t}^{(\delta_{1:t})}, P_{t}^{(\delta_{1:t})}; \delta_{1:t} \in \Delta_t \} \)

The MS starts at \( x = 600 \) and \( y = 0 \) and goes through the network counterclockwise. The information needed to identify the actual and the mirrored solution becomes available when a handover changes the symmetry of the problem. Delaying the hard decision about which of the two principal solutions is the correct one until a handover happens, ensures the right solution is kept.

Fig. 5 illustrates the first 19s of the algorithm applied to the simulated network with \( M = 2 \) and ideal noise free measurements. At time \( t = 0 \), the MS starts its path at point \( (600, 0)^T \). The serving BS is BS5 and the neighboring BS is BS1. As the figure suggests, the mode sequence \( \Delta_{18} \) contains the MS’s estimated positions \( \hat{\delta}_{1:18} \), and the shadow solutions, \( \hat{\delta}_{2:18} \), both equally likely given the symmetry of the problem. The measurement arriving at \( t = 19 \) results in four possible mode sequences, \( \Delta_{19} = \{ (\hat{\delta}_{1:18}, \pm 1), (\hat{\delta}_{2:18}, \pm 1) \} \). However, due to the change of symmetry axis, only one of these are likely to be true \( \hat{\delta}_{1:19} = (\hat{\delta}_{1:18}, \pm 1) \) as illustrated in Fig. 5. The correct solution can now be backtracked, as the mode sequence represented by \( \hat{\delta}_{1:19} \).
The minimum variance estimate of the states, from a filter bank is obtained as the weighed average of the contained filter solutions,

\[
\hat{x}_{t|t}^{\text{MSE}} = \frac{1}{\ell} \sum_{\delta_1 \in \Delta_t} w_{t|t}^{(\delta_1,1)} \hat{x}_{t|t}^{(\delta_1,1)},
\]

\[
P_{t|t}^{\text{MSE}} = \frac{1}{\ell} \sum_{\delta_1 \in \Delta_t} w_{t|t}^{(\delta_1,1)} \left( \mathbf{P}_{t|t}^{(\delta_1,1)} + (\hat{x}_{t|t}^{(\delta_1,1)} - \hat{x}_{t|t}^{\text{MSE}})(\cdot)^\top \right),
\]

where notation \((\cdot)\) denotes that this parenthesis has the same elements of its adjacent parenthesis.

V. PERFORMANCE BOUNDS

In this section, two performance bounds for the position estimates are derived. The bounds studied both answer the question: Given a certain trajectory, what is the lower bound of the covariance of the estimate that can be obtained? The first bound studied is the parametric CRLB, which is a local, and hence overly optimistic bound for Jump Markov linear systems (JMLS). Secondly, a MMSE bound is derived, taking inherent ambiguities into account, but assuming the mode distribution known (as can be obtained in the studied application).

A. Cramér-Rao Lower Bound

The covariance of an unbiased estimator, \(\hat{\theta}\), of the parameter \(\theta\) given the measurements \(y = h(\theta) + e\), is under mild regularity conditions (see [34]) bounded by the Cramér-Rao lower bound

\[
\text{cov}(\hat{\theta}) \succeq J^{-1}(\theta^0) = \mathbf{P}^{\text{CRLB}}(\theta^0),
\]

where \(J(\theta^0)\) denotes the Fisher information in the true parameter value \(\theta^0\). The Fisher information is given by

\[
J(\theta^0) = -\mathbb{E} \nabla^2 \log p(y|\theta) = H^\top(\theta^0) \mathcal{I}_e H(\theta^0),
\]

where \(H(\theta^0) = \nabla^2 h(\theta^0)\), and \(\mathcal{I}_e = -\mathbb{E} \nabla^2 p_e(e)\) is the intrinsic accuracy of the noise [33]. Assuming Gaussian measurement noise with covariance \(\mathbf{R}, \mathcal{I}_e = \mathbf{R}^{-1}\).

The CRLB concept can be extended to cover the estimation of the state of dynamic systems. One of the available results bounds the filtering performance along a nominal state trajectory, \(x_{0:t}\). This is denoted the parametric CRLB [35], which states

\[
P_{t|t} \succeq J^{-1}_t(x_{0:t}|t) = P_{t|t}^{\text{CRLB}}(x_{0:t}),
\]

where \(J^{-1}_t(x_{0:t})\) is obtained from the covariance recursion in the EKF

\[
J^{-1}_{t+1}(x_{0:t+1}) = F_t(x_{0:t}) J^{-1}_t(x_{0:t}) F_t^\top(x_{0:t}) + \mathcal{I}_w^{-1},
\]

\[
J_{t|t}(x_{0:t}) = J_{t-1|t}(x_{1:t}) + H_t^\top(x_{0:t}) \mathcal{I}_e H_t(x_{0:t}),
\]

linearized about the nominal trajectory, and substituting noise covariances with their inverse intrinsic accuracy. Note that the CRLB for the parameter estimate enters (19b) as the last term, which represents the new information provided by the measurement. The information form (19b) is used to highlight this fact. This has a nice implication here, the original nonlinear measurement equation can be substituted for a trivial linear one based on the snapshot estimates of the position using the Fisher information as intrinsic accuracy.

It should be noted that the CRLB is a local property, it does not take the ambiguities of the estimate, i.e., multi-modality of the posterior, into account. This is illustrated by the fact that the CRLB for the JMLS in this case can be computed assuming the mode to be known. Hence, the CRLB is overly optimistic when ambiguities arise. Next a bound taking this into consideration is considered.

B. Minimum Mean Square Estimate Bound

An optimal minimum mean square estimator (MMSE) for a JMLS with Gaussian noise can be obtained by constructing a filter bank considering all possible mode sequences \(\delta_{1:t}\), which then perfectly represents the true posterior distribution as a sum of weighted KF. Based on the filter bank representation, of the posterior distribution, the optimal estimate \(\hat{x}_{t|t}\) and its covariance \(P_{t|t}\) can be found as

\[
\hat{x}_{t|t} = \sum_{\delta_1 \in \Delta_t} \omega_{t|t}^{(\delta_1,1)} \hat{x}_{t|t}^{(\delta_1,1)},
\]

\[
P_{t|t} = \sum_{\delta_1 \in \Delta_t} \omega_{t|t}^{(\delta_1,1)} \left( \mathbf{P}_{t|t}^{(\delta_1,1)} + (\hat{x}_{t|t}^{(\delta_1,1)} - \hat{x}_{t|t})(\cdot)^\top \right),
\]
the ambiguities into account, providing a tighter bound than the local CRLB. This then becomes a suitable bound for the achievable performance under ambiguities when these can be expressed as JMLS. In most cases the exponential growth of possible components in the filter bank renders the computation of (20) infeasible in practice. However, in some situations the structure of the problem or some additional information allows for limiting the number of components in the filter bank, to make the computation feasible. If additional information is introduced, e.g., information about the mode probabilities, the bound loses tightness but remains considerably tighter than the pure CRLB.

The case studied in this paper, the position ambiguity arising from the raw measurements possesses a certain symmetry around the line between the two BSs used at any time, as visible in Fig. 6. With this knowledge of the problem setup it is possible to, in each time instance, identify the feasible modes that cannot be distinguished between at each time instance. For instance, when the used BSs change, only new modes that match an existing mode remains and no old modes without support in the new measurement are kept. The result is at most two modes at any time, and that when more than one mode exists they have equal weight. Hence, the proposed MMSE bound can easily and efficiently be computed given the probability of each mode sequence.

VI. PERFORMANCE EVALUATION

A simulation study is used to evaluate positioning capability using two BSs from single measurements, and based on the filter bank solution proposed in Sec. IV. The filter bank solution is further evaluated using experimental data.

A. Simulation Results

The simulated cellular network consists of 7 macro sites each having 3 cells, located in a hexagonal grid with an intersite distance of 1500 m. The MS starts at a certain point in the cellular network and passes through the network counterclockwise along a pre-defined flower-shaped path with the speed 1 m/s. The flower shape of the trajectory is selected to excite the key aspects of timing-based localization [36]. Fig. 4 illustrates the cellular network deployment in which the macro sites, each with three cells, are marked with black circles and the green line represents the true MS trajectory.

The scatter plot of two solutions obtained when solving the snapshot problem are also indicated in Fig. 4. The mode of the system changes along the trajectory depending on the set of measuring BSs and their geometrical positions. For instance, in the starting point marked with cyan, \( \theta = [600, 0]^\top \), BS5 and BS1 are the set of measured BSs. The true position of the system thus corresponds to the mode \( \delta = 1 \). As the MS moves along the trajectory at the position \( \theta \approx [680, 400]^\top \) the measured base stations are BS5 and BS2 and the position of the system then corresponds to mode \( \delta = -1 \).

1) Snapshot Evaluation: To better understand the properties of the simulated setup, expression (17) is used to obtain a lower bound of the covariance of the position estimate, as a function of the position and the two BSs used. The results are given in Fig. 7, which illustrates the point-wise positioning accuracy along the evaluated trajectory assuming \( \sigma = 8 \) m, and the BS selection scheme discussed in Sec. III. While an error of below 80 m, in terms of the square root of trace of CRLB, is expected in most of the area inside the BS coverage, there are certain points where the expected positioning error is worse. Hence, it is obvious that the geometry of the problem influences the obtainable performance.

2) Filter Bank Evaluation: For the proposed filter bank solution, a constant velocity motion model is used, with the process noise standard deviation \( \sigma_w = 1 \) m/s². The snapshot estimates are used as measurements, and the measurement noise is given by \( \Lambda(\delta) \), which is estimated using the UT transform in Algorithm 1. The filter is initialized randomly inside a circle of radius 50 m centered at the closest BS and the initial velocities along the \( x \) and \( y \) axis are set to zero,
$v_{x_{0}|0} = v_{y_{0}|0} = 0 \text{ m/s}$. The initial uncertainty, $P_{0|0}$, is given by $\sigma_{x_0} = \sigma_{y_0} = 10 \text{ m}$ and $\sigma_{e_{x_0}} = \sigma_{e_{y_0}} = 1 \text{ m/s}$.

Fig. 8 illustrates the normalized weights after the pruning stage of each mode along the trajectory. The filter bank is allowed to have $M = 2$ branches. The handover times are marked with the dashed lines. As soon as a handover occurs, the information obtained from the change in the set of involved BSs results in a jump in the probability of modes. For instance, as Fig. 8 suggests, initializing the filter bank with equal probabilities for two modes, results in equal weights until $t = 19 \text{ s}$. At $t = 19 \text{ s}$, when the set of involved BSs change, the other mode, marked with black, contains the estimated MS position.

The performance of the proposed filter is evaluated using 1000 MC simulations. At first, the effect of the number of competing modes, $M$, and the sampling period, $T$, is investigated in terms of the cumulative distribution function (CDF) of the positioning error of the MSE estimate (16). To stress the impact of bandwidth, evaluations are performed for both 10 MHz and 1.4 MHz signal bandwidths. Fig. 9 and Fig. 10 present the empirical CDF plots of the positioning accuracy for $M \in \{2, 4, 16, 32\}$ and $T \in \{1, 10, 20, 50, 100\}$ s with 10 MHz and 1.4 MHz signal bandwidths, respectively. As shown in the figures, in all scenarios, except for $M = 32$, the positioning error is very large when the sampling period is below $T = 50 \text{ s}$. The reason for such a behavior is that when the two solutions get very close to each other, each of the existing hypotheses get equal weights. One remedy to this problem is to allow a large number of modes to compete resulting in a deeper mode tree in the filter bank.

The 50th and 90th percentile of horizontal positioning error, extracted from the empirical CDFs, for 10 MHz and 1.4 MHz bandwidths are reported in Table II and I, respectively. The result for the 10 MHz bandwidth scenario indicates that in order to guarantee 60 m error 90% of the time, the sampling time must be more than 20 seconds. For the 1.4 MHz bandwidth, such a sampling time guarantees 50 m error 50% of the time.

Although increasing the number of hypotheses allows us to have smaller sampling periods, the complexity of the problem increases exponentially. In the rest of this section, we let $M = 4$ and $T = 50 \text{ s}$ and further investigate the performance of the proposed filter.

Fig. 11(a) presents the comparison between the numerical RMSE of the state estimate and the square root of $\text{tr}(P_{\text{MMSE}})$ introduced in (20b) corresponding to 20 MHz signal
TABLE I: 50%/90% horizontal positioning error statistics for 1.4 MHz bandwidth.

<table>
<thead>
<tr>
<th>T [s]</th>
<th>M 2</th>
<th>M 4</th>
<th>M 16</th>
<th>M 32</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>96/473</td>
<td>96/472</td>
<td>95/472</td>
<td>95/472</td>
</tr>
<tr>
<td>10</td>
<td>95/464</td>
<td>93/462</td>
<td>87/453</td>
<td>79/440</td>
</tr>
<tr>
<td>20</td>
<td>86/447</td>
<td>70/402</td>
<td>56/280</td>
<td>54/263</td>
</tr>
<tr>
<td>50</td>
<td>48/176</td>
<td>46/139</td>
<td>45/145</td>
<td>44/140</td>
</tr>
<tr>
<td>100</td>
<td>45/154</td>
<td>45/148</td>
<td>45/140</td>
<td>45/139</td>
</tr>
</tbody>
</table>

TABLE II: 50%/90% horizontal positioning error statistics for 10 MHz bandwidth.

<table>
<thead>
<tr>
<th>T [s]</th>
<th>M 2</th>
<th>M 4</th>
<th>M 16</th>
<th>M 32</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48/466</td>
<td>48/465</td>
<td>44/461</td>
<td>43/460</td>
</tr>
<tr>
<td>10</td>
<td>46/460</td>
<td>41/451</td>
<td>33/406</td>
<td>29/355</td>
</tr>
<tr>
<td>20</td>
<td>32/392</td>
<td>26/267</td>
<td>21/105</td>
<td>20/81</td>
</tr>
<tr>
<td>50</td>
<td>19/62</td>
<td>19/59</td>
<td>19/60</td>
<td>19/57</td>
</tr>
<tr>
<td>100</td>
<td>19/61</td>
<td>18/61</td>
<td>18/60</td>
<td>18/60</td>
</tr>
</tbody>
</table>

bandwidth. As the figure suggests, both the bound and the numerical RMSE are large when the two possible positions are equally likely and far apart. However, as soon as a new BS pair is assigned, the positioning error decreases quickly. Fig. 11(b) compares the square root of \( P_{t|t} \) (b) Cramér-Rao lower bound (19) and positioning error of the state estimate using the estimated mode, obtained in the simulation in logarithmic scale, for 10 MHz signal bandwidth. The handover times are indicated with red bars in (a).

![Fig. 11: (a) Minimum mean square error bound (20) and positioning error of the state estimate \( \hat{x}_{t|t} \). (b) Cramér-Rao lower bound (19) and positioning error of the state estimate using the estimated mode, obtained in the simulation in logarithmic scale, for 10 MHz signal bandwidth. The handover times are indicated with red bars in (a).](image)

Fig. 12: (a) Minimum mean square error bound (20) and positioning error of the state estimate \( \hat{x}_{t|t} \). (b) Cramér-Rao lower bound (19) and positioning error of the state estimate using the estimated mode, obtained in the simulation in logarithmic scale, for 1.4 MHz signal bandwidth. The handover times are indicated with red bars in (a).

![Fig. 12: (a) Minimum mean square error bound (20) and positioning error of the state estimate \( \hat{x}_{t|t} \). (b) Cramér-Rao lower bound (19) and positioning error of the state estimate using the estimated mode, obtained in the simulation in logarithmic scale, for 1.4 MHz signal bandwidth. The handover times are indicated with red bars in (a).](image)

B. Experimental Results

To experimentally evaluate the proposed method, three separate antennas were used to collect measurements in the Kista area in Stockholm, Sweden. The network mimics a macro-cell deployment of LTE in urban areas. The inter-site distance of BSs were between 350 m to 600 m where the height of all the antennas were a few meters above the average height of the buildings in the surrounding environment. The synchronization of the transceivers were highly accurate, with standard deviation of less than \( 10^{-12} \) s. The ground truth trajectory is obtained by logging GPS measurements. The whole trajectory took 18 minutes to traverse. For more detailed description of hardware and the measurement campaign see [37]. Fig. 13 presents the taken path.

The performance of the proposed method is evaluated for the experimental scenario in terms of the scatter plot of the estimated position and the positioning error. The sampling period of the filters for the real data is \( T = 5 \) s. The filters are initialized at the closest BSs with \( v_{x_{0|0}} = v_{y_{0|0}} = 0 \). The quantities for the initial uncertainties are the same as the simulation scenario. The observed standard deviations, used to form the measurement covariance matrix, are constant and equal to 8 m for all base stations.

![Fig. 13: The estimated positions using MMSE, CRLB, and numerical RMSE from the proposed method for the real network scenario. The measurements are indicated with red bars.](image)

Fig. 13 presents the estimated positions, marked with white, on top of the true position, marked with red, for the whole path in the network. In the majority of the trajectory, the estimated positions follow the true trajectory with high accuracy. However, in the area \( x \approx [500, 600] \) and \( y \approx [400, 500] \), the measurements suffer from NLOS conditions [37]. The poor accuracy in some areas of the real network deployment scenario is due to both the geometry of the antenna placements and the MS, and also a result of NLOS effects in the measurements.

The positioning error of the real field experiments is calculated relative to the GPS logged positions. Fig. 14 presents the RMSE of the proposed filter bank and of the proposed filter bank allowing for a lag of \( T = 5 \) s to resolve mode ambiguities. The reported values corresponding to the filter bank are based on the MMSE estimate of the position without any further knowledge obtained from the change in the set of serving BSs. As expected, the error increases when the two possible solutions are both likely. The estimation error shrinks as soon as the set of serving BSs change. Incorporating this additional information, as in the proposed filtering solution,
We have proposed a filter bank framework for positioning based on an RTT and a TDOA measurements obtained from two base stations. We first derived the two possible analytical solutions of the intersection of an RTT circle and a TDOA hyperbola. The nonlinear mapping between the collected noisy measurements to the 2D positions of the MS was then approximated using an estimator based on the unscented transformation. In order to estimate the MS position from the two UT estimates, they were then fed into a filter bank as pseudo-measurements of the position of the MS. The filter bank keeps track of all possible solutions until more information is available. While the MS moves through the network, the serving BS changes through a handover procedure. The additional information obtained from the handover is then automatically utilized by the filter bank. The developed method is evaluated in terms of RMSE for both simulated network and real-field experiments collected from the network in the Kista area.

VII. CONCLUSIONS

We further evaluate the performance of the proposed filter by comparing it to a conventional TDOA approach with no measurement ambiguity. In the considered conventional approach, the positions are estimated using a standard extended Kalman filter (EKF) in which the transition of the four-component state vector is modeled using a constant velocity model. At each time, the measurement vector of the EKF is given by \( \tilde{z} = [\tilde{z}_1, \tilde{z}_2, \tilde{z}_3, \tilde{z}_4] \) where \( \tilde{z}_1 \) denotes TOA and \( \tilde{z}_i, i \neq j \) denotes TDOA. Note, this solution hence uses 50\% more measurements than our solution! Hence, this method solves a much easier problem, as the ambiguities observed with only two measurements, can be resolved using the additional measurement. Alternatively, comparisons could be made with existing methods proposed in the literature. For instance, the authors in [26] propose a method based on Taylor linearization of TOA and TDOA measurement models. However, linearization without an extremely good prior results in poor performances with the considered nonlinear measurement models.

Fig. 15 presents the CDF of the positioning error for the proposed filter with two base stations and the EKF using \( \tilde{z} \) as measurements. Although the 67\% percentile of the EKF reads 19 m error that is 4 m less than the proposed filter, the 95\% percentile of the estimation error of both algorithms is around 40 m. The proposed method is hence marginally worse than the standard method, but does at the same time reduce the need of TDOA measurements to half. The poor performance obtained using a standard nonlinear weighted least squares (NWLS) approach, marked with red in Fig. 15, indicates that Taylor linearization methods do not perform well in this application. The NWLS is initialized at the closest BS position and refined iteratively to estimate the MS’s position.

Fig. 13: The final KF bank estimates, marked with white, on top of the GPS-logged trajectory of the MS, marked with red. The three base stations are marked with green.

Fig. 14: RMSE of the estimated position corresponding to the experimental data, for the state estimates marked with blue, and the estimate that is delayed to obtain more mode information marked with black.

Fig. 15: Empirical error CDF of the horizontal positioning error of the KF bank estimator applied to the data from real experiments, marked with blue. The result of using NWLS estimator with two base stations is marked with red and the result of using EKF with three base stations is marked with black.
Stockholm, Sweden. The results indicate good performance for data from both networks.

REFERENCES


