AMERICAN UNIVERSITY OF BEIRUT

HYBRID AND INTELLIGENT POSITIONING TECHNIQUES IN HETEROGENEOUS AND COGNITIVE NETWORKS

by
ALI HUSSEIN YASSIN

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Engineering
to the Department of Electrical and Computer Engineering
of the Faculty of Engineering and Architecture
at the American University of Beirut

Beirut, Lebanon
May 2014

AMERICAN UNIVERSITY OF BEIRUT

HYBRID AND INTELLIGENT POSITIONING TECHNIQUES IN HETEROGENEOUS AND COGNITIVE NETWORKS

by

ALI HUSSEIN YASSIN

Approved by:

Dr. Youssef Nasser, Senior Lecturer
Electrical and Computer Engineering

Advisor

Dr. Mariette Awad, Assistant Professor
Electrical and Computer Engineering

Co-Advisor

Dr. Hassan Artail, Professor
Electrical and Computer Engineering

Member of Committee

Dr. Zaher Dawy, Associate Professor
Electrical and Computer Engineering

Member of Committee

Date of thesis/dissertation defense: May 9, 2014

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: Yassin (Last) Ali (First) Hussein (Middle)

⬭ Master's Thesis ⬭ Master's Project ⬭ Doctoral Dissertation

☐ I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

☐ I authorize the American University of Beirut, **three years after the date of submitting my thesis, dissertation, or project,** to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

_____

Signature Date

This form is signed when submitting the thesis, dissertation, or project to the University Libraries

ACKNOWLEDGMENTS

Special thanks to my parents who always encourage me and give me hope in defeating obstacles, no matter how difficult they are, and accomplishing extraordinary achievements. They are always my source of strength and motivation for accomplishing my Bachelor's degree and Master's Degree at AUB. They are the owners of this achievement, and they will be the owners of higher achievements

Special thanks to my advisor Prof. Youssef Nasser for his precious support and guidance throughout my academic career during my undergraduate and graduate studies at AUB. Prof. Youssef Nasser was not only the academic advisor, but also a brother who always have confidence in me for achieving the PhD degree.

Also, I would thank my co-advisor Prof. Mariette Awad for her precious aid during my academic career at AUB.

I would also thank Prof. Hassan Artail and Prof. Zaher Dawy for serving as members in my thesis committee.

AN ABSTRACT OF THE THESIS OF

Ali Hussein Yassin for Master of Engineering
Major: Electrical and Computer Engineering

Title: Hybrid and Intelligent Positioning Techniques in Heterogeneous and Cognitive Networks

Nowadays, the availability of the location information becomes a key factor in today's communications systems for allowing new location based services. In outdoor scenarios, the Mobile Terminal (MT) position is obtained with high accuracy thanks to the Global Positioning System (GPS) or to the standalone cellular systems. However, the main problem of GPS or cellular systems resides in the indoor environment and in scenarios with deep shadowing effect where the satellite or cellular signals are broken.

This thesis is divided into two main parts. In the first part, we present a potentially good candidate for critical positioning scenarios with the lack of hearability between the Unlocated Mobile Terminal (UMT) and the Anchor Nodes (AN). Indeed, in many cases, only one or two ANs are communicating with the UMT. The proposed positioning algorithm is based on hybrid data fusion and its extension to the tracking step by adopting the Minimum Entropy Criterion to diminish the shortcoming of the Unscented Kalman Filter and Particle Filter, usually used in this field. The proposed solution is divided into two phases: the learning phase and the processing phase. Using Radial Basis Functions, the learning phase allows an accurate model of the probability density function of the positioning error while the processing phase aims at reducing the estimation error. We show that the proposed algorithm reaches an accuracy of 1m squared in terms of Mean Square Error (MSE).

Not only localizing MTs, but also localizing multiple transmitters in a region is also another objective tackled in this thesis. Even though this problem is applicable in different applications, the most prominent one is the cognitive radio context. We are interested in the uncoordinated system where the cognitive node operates in an opportunistic manner. In order to avoid interference, the cognitive system is responsible to recognize the area where there are active primary users. Assuming the location of primary users and their activity are not known, we propose in this thesis the Space-Alternating Generalized Expectation Maximization (SAGE) technique for localizing multiple transmitters based on "smart" initial estimations. We show that the proposed SAGE technique outperforms the conventional techniques such as the Expectation Maximization and the Random Guesses techniques. This improvement is much significant with higher shadowing variance.

LIST OF ILLUSTRATIONS

TABLE OF CONTENTS

CHAPTER I

REVIEW OF THE STATE OF THE ART POSITIONING TECHNIQUES

A. Introduction

Today and future communications systems aim at providing high data rates with ubiquitous service coverage. Nowadays, the availability of the Mobile Terminal (MT) location information at the base stations, i.e. its knowledge by the operators, becomes a key factor in today's communications systems for allowing new location based services [1][2].

In practice, localization techniques are based on Time of Arrival (ToA) [3], Time Difference of Arrival (TDoA) [4], Received Signal Strength (RSS) [5] and Angle of Arrival (AoA) [6]. In outdoor scenarios, the MT position is obtained with high accuracy thanks to the Global Positioning System (GPS) or to the standalone cellular systems. However, the main problem of these positioning systems resides in the indoor environment where the satellite or cellular signals are broken but also in scenarios with deep shadowing effect [7]. Moreover, in homogeneous networks such as cellular networks, the estimation of the Positioning Information (PI) of any device becomes harder as the physical communications resources are more and more valuable.

A potentially good candidate for critical scenarios resides in the class of heterogeneous approaches that combines different radio access technologies (such as cellular systems, WiFi, WiMAX). Indeed, cellular and WiFi networks based localization techniques have recently received increasing interests in both localization and

communication community, e.g. [8][9][10]. This is not only because of the request made by Federal Communication Commission (FCC) about the accurate localization of the MTs, but also because of many other applications that are location sensitive such as billing, fleet management, and mobile yellow pages [11].

Even though all positioning techniques could be exploited in indoor scenarios and homogeneous networks however, in practice, there are limits obtained on the combination of these techniques as well as on the minimal number of anchor nodes (AN) used in such scenarios [12][13]. The main challenge resides in the lack of hearability between the Unlocated Mobile Terminal (UMT) and the ANs [12]. Indeed, in many cases, only one or two ANs are communicating with the UMT. Hence, new techniques based on hybrid data fusion should be proposed and analyzed in this extreme case. In literature, many techniques have been proposed [12][13]. In [12], the authors have proposed a positioning technique based on the combination of ToA and fingerprinting RSS by using one AN and one Located MT (LMT). However, the accuracy of the technique proposed in [12] is not high enough. In [13], the authors have presented a cooperative positioning technique based on the combination of long-range measurements obtained by three ANs and short range measurements obtained by Wi-Fi. The main problem of this technique is that it requires measurements obtained by three ANs and a WiFi hotspot.

B. Literature Review

In this section, we briefly describe the basic standalone positioning techniques used in the context of homogeneous networks. In addition, we explore different research activities on cooperative mobile positioning. Then for tracking purpose, we present a short description of the Unscented Kalman Filter (UKF) estimation.

1. *Location Estimation Techniques*

Training sequences sent by BSs or MTs are used for the location estimation. There are 4 main techniques used for localization, which are ToA, TDoA, AoA, RSS and RSS based fingerprinting. The ToA and TDoA techniques need at least three BSs for localization. The AoA technique requires a minimum of two BSs which means that the estimation error may be large, and the ambiguities of location estimation will exist if the number of the available BSs is less than the minimum requirement.

a. Time of Arrival (ToA)

The ToA approach includes the calculation of the time needed by the signal to travel from the UMT to the ANs. Accordingly, the UMT will be moving on a circle of center given by the AN and with a radius d estimated through the ToA. Hence, to detect the

exact location of the MT, at least three ANs are required. In this case, the estimated position of the UMT is simply within the region of intersection (if it exists) of the drawn circles. It could be easily obtained through any filtering technique such as Least Square (LS) or Weighted Least Square (WLS).

b. Received Signal Strength based Fingerprinting

The RSS approach includes two main methods: the pathloss lognormal shadowing model to deduce a trilateration, and the RSS fingerprinting [12]. Basically, the RSS based fingerprinting firstly collects RSS fingerprints of a scene and then estimates the location of the MT by matching on-line measurements with the closest possible location collected by measurements in a database [8]. Therefore for each possible location, ambiguities points could exist leading then to high estimation errors in standalone positioning scenario.

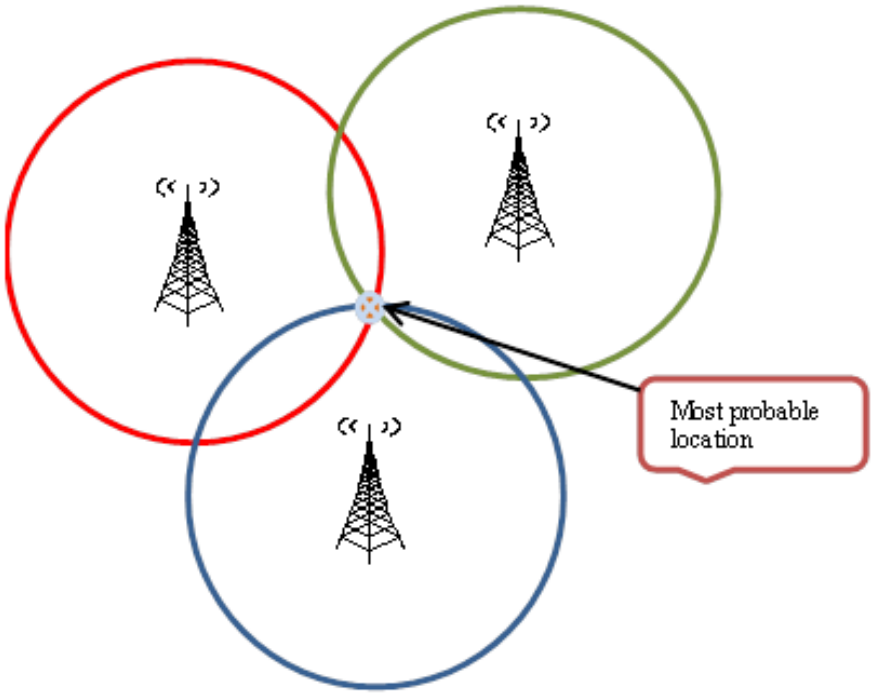

Figure 1 TOA Ranging

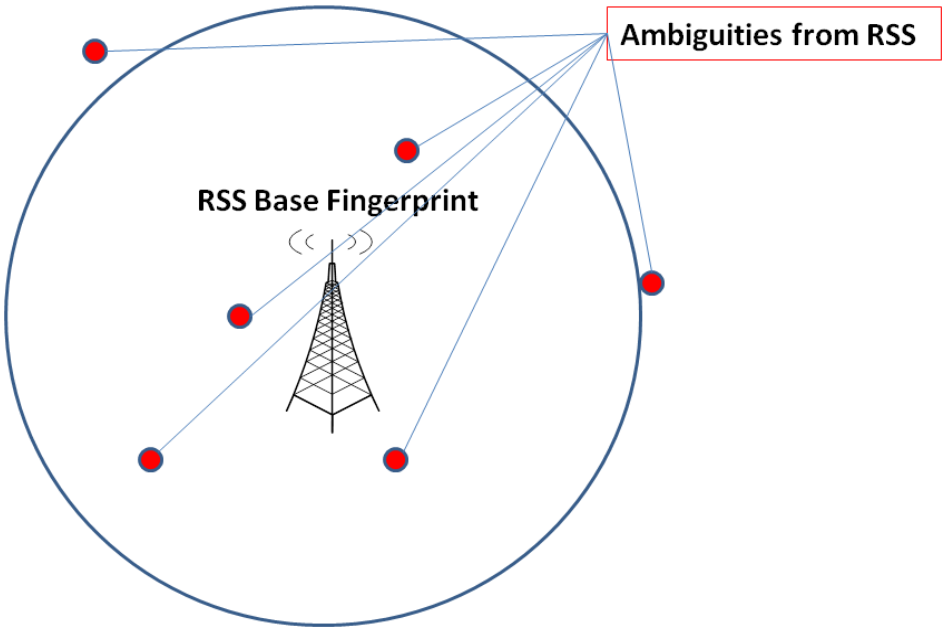

Figure 2 Ambiguities from RSS

c. Angle of Arrival (AoA)

This technique includes the calculation of the angle at which the signal arrives from the UMT to the ANs [6]. Then, the region where the MT could exist can be drawn. Basically, this region is a line having a certain angle with the ANs. Although at least two ANs are needed to estimate the location of the MT, the position estimation error could be large if a small error occurs in the AoA estimation. Therefore, the AoA based technique is with a limited interest for positioning purposes.

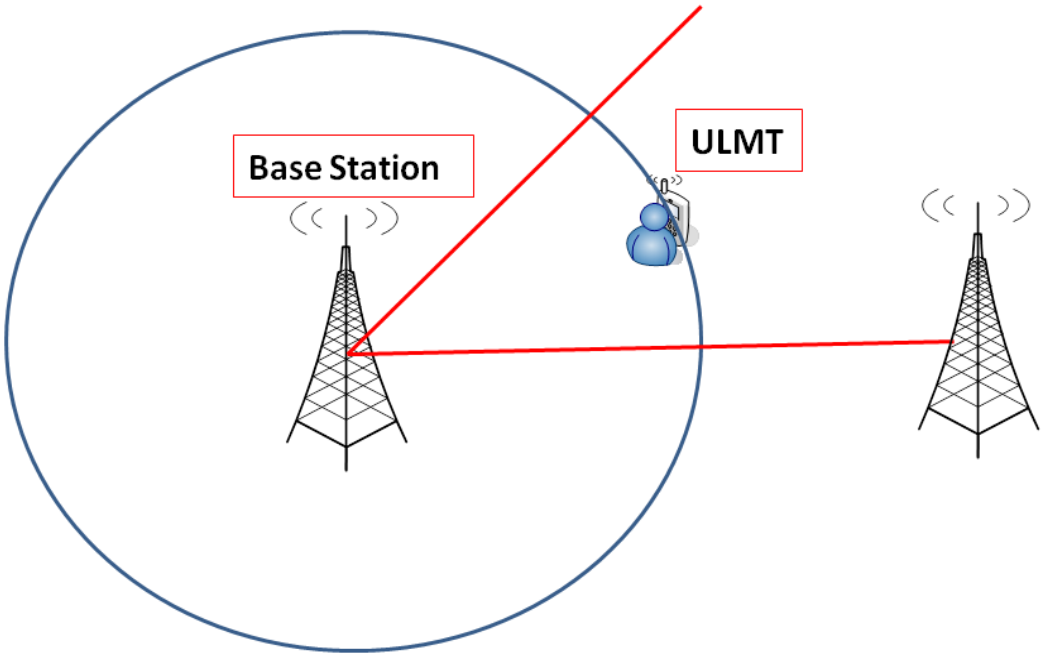

Figure 3 AOA Model

2. *Cooperative Mobile Positioning*

Recently, cooperative mobile positioning is evolving as a new division of wireless location; many research activities are conducted in this domain, such as positioning, tracking, and clustering. The core idea of cooperative positioning relies on the utilization of trustworthy short-range mobile-to-mobile measurements to enhance the accuracy of the location estimation of a wireless system. For instance, the work presented in [14] introduces an example of cooperative mobile positioning system sustained by a hybrid WiMAX/Wi-Fi network.

Eventually, the basic standalone positioning techniques utilized in homogeneous networks includes RSS, TOA, TDOA, AOA, etc. Different combinations of these location estimation techniques have been implemented to enhance the accuracy of location estimation. For instance, [14] proposes a data fusion between TDOA and RSS measurements. Additionally, [15] shows that the hybrid TOA/TDOA and RSS achieves further enhancement in terms of location estimation accuracy when compared to the use TOA or TDOA alone; the RSS approach includes two main methods: the pathloss lognormal shadowing model to deduce a trilateration, and the fingerprinting [16]. Basically, the RSS based fingerprinting firstly collects RSS fingerprints of a scene and then estimates the location of the MT by matching on-line measurements with the closest possible location collected by measurements in a database [8]. Therefore for each possible location, ambiguities points could exist leading then to high estimation errors in standalone positioning scenario.

Moreover, the hybrid TOA-AOA technique presented in [6] proves to achieve higher accuracy in position estimation when compared to TOA or AOA alone. Concerning limitations, it is well known that three ANs are usually required to obtain suitable location accuracy. In GPS based systems, this condition is almost satisfied since there is direct line-of-sight connection between the UMT and the different positioning satellites. However, when the UMT is in deep shadowing or indoor scenarios, the positioning information obtained through satellites will be lost and a terrestrial connection through the cellular system should be established in order to retrieve the location information. The main problem of cellular systems however resides in the hearability restricted conditions of the UMT. The work presented in [17] proposes a hybrid positioning technique, a combination between RSS fingerprints, TOA, and AOA; this work make use of a coupling and decoupling algorithm via UKF to merge these measurements as a try for solving the lack of hearability problem.

The coupling and decoupling algorithm used by [17] and [13] decouples first the relative localization, recognized through the short-range measurements obtained via the secondary anchor node (SAN), and the absolute localization, recognized through the long-range cellular obtained via the estimation of the coordinates and orientation of the group in the cellular network. After the decoupling of the short and long range measurements, an iterative Kalman filter is executed on each measurement to improve the positioning and tracking estimates. Then, the absolute positions are obtained by coupling of the long and short range measurements.

3. *2-Level Kalman Filter (2LKF)*

The paper presented by [13] introduces the 2-Level Kalman Filter (2LKF) as a solution for decoupling the relative localization of the users utilizing peer-to-peer ad-hoc links from the absolute localization of the same users. Basically, the proposed algorithm in [13] analyses the framework of positioning for cooperative schemes. This is mainly useful for scenarios with heterogeneous technologies. Thus, different timing behaviour of the channel measurement procedure can be handled for the long-range and short-range technologies without any additional complexity. Hence, this algorithm is based on decoupling absolute localization recognized via the long-range cellular links and relative localization recognized via the short-range ad-hoc links [13]. Due to the strong non-linear behaviour of the AoA measurements, the UKF is highly recommended for such positioning and tracking scenarios.

Basically, EKF uses the first term of the Taylor series estimation to deal with nonlinear systems. When the system is highly nonlinear, EKF may easily diverge. Thus, UKF was founded on the basis that estimating a probability distribution is easier than estimating an arbitrary nonlinear function. UKF follows a sampling method known as unscented transformation (UT) to define a set of sample points, called sigma points, whose weighted "sample mean and covariance" are close enough to the real mean and covariance. In addition, UKF provides am advantage over the EKF by giving third order accuracy for Gaussian inputs and at least the second order for non-Gaussian inputs.

CHAPTER II

POSITIONING AND CLUSTERING IN HETEROGENEOUS NETWORKS WITH LACK OF HEARABILITY

A. Introduction

In practice, most of the positioning techniques could be exploited in lack of hearability scenarios and homogeneous networks. For instance, in indoor scenarios, the main challenge resides in the lack of hearability between the Unlocated Mobile Terminal (UMT) and the ANs [8]. Hence, the usual minimal number of anchor nodes (AN) used in such scenarios [12][13] is not always available. Indeed, in many cases, only one or two ANs are communicating with the UMT. This imposes new research directives to find and propose adaptive solutions for such critical environment.

Hence, we first propose a hybrid positioning data fusion based on ToA, AoA and RSS fingerprinting. Contrarily to [12], we propose to use the AoA as an additional input for resolving ambiguities. Then, based on an efficient combination of the contributions presented in [12] and [13], we extend the positioning scenario to its extreme case. Contrarily to [13], we assume in our work that only two ANs are available for localization. So, we investigate the combination of short range measurements, obtained via the LMT or via the Wi-Fi hotspot, with the long range measurements to improve the accuracy of the positioning information using the UKF.

Finally, we investigate different clustering approaches for the hybrid data fusion in the heterogeneous context. The main objective of clustering is to improve the PI estimation by exchanging information with the necessary mobile terminals within a cluster while keeping a reduced overhead cost.

B. Problem Statement

It is well known in literature that three ANs are usually required to obtain suitable location accuracy. In GPS based systems, this condition is always satisfied since there is direct line-of-sight connection between the UMT and the different positioning satellites. However, when the UMT is in deep shadowing or indoor scenarios, the positioning information obtained through satellites will be lost and a connection through the cellular system should be established in order to retrieve the location information.

However, the main problem of cellular systems resides in the hearability restricted conditions of the UMT. With the necessary condition of three ANs, one could assume that the UMT is closer to one Base Station (BS- seen as AN) than others. In this case, the signal received from the other BSs will be very weak and might be interpreted as interference. Therefore, the UMT should search for some local solutions such as local Wi-Fi hotspot or another Femto-cell or cooperative communications through the interaction with a LMT to resolve the hearability conditions [18][19]. In this work, we adopt the positioning estimation scenario with restricted hearability conditions such as in indoor scenario, deep shadowing conditions, and insufficient number of ANs.

As depicted in Figure 4, we assume in this paper that the UMT is connected to one BS and a Secondary Anchor Node (SAN such as one Wi-Fi hotspot or a Femto-cell or one LMT). We assume that all the estimated ToA and RSS fingerprinting [20] are collected at the home BS for centralized processing. Denote $a(a_x, a_y)$ as the true 2-D location of the LMT or of SAN, u(ux, uy) as the true location of the UMT. Without loss of generality, the location of the home BS is set as O= (0, 0). We also assume that the location of the SAN is obtained with imperfections and denote $\hat{a} \triangleq (\hat{a}_x, \hat{a}_y)$, as its estimated location modelled as random with Gaussian distribution with variance σ^2_{SAN}.

In our work, we assume that RSS fingerprinting are collected beforehand at the BS. Hence, ambiguities on the location estimation exist. The vector of ambiguities is represented as $s = [s_0, s_1, \dots, s_M]$ while one point is the exact position and M points represent the ambiguities. As given in [12], the (x,y) coordinates of the ambiguities are modelled as random processes with Gaussian distribution and with variance σ^2_{rss}.

In this section, we propose to use the estimation algorithm studied in [12] as our basic coarse positioning algorithm. In other words, we combine between the RSS fingerprints and the ToA measurements. However, in contradiction with [12], we propose to use the AoA information to reduce the number of ambiguities analyzed through the RSS fingerprinting. This will reduce the domain search of the possible location of the UMT even though is not necessary for the flow and execution of the PDF based algorithm proposed in the paper. The ToA based distances estimation between the ANs and the UMT could be modelled as:

$$\hat{d}_{BS} = d_{BS} + \omega_{TOA} * \mathcal{N}\ (0,1) + \varepsilon_{NLOS} \quad (1)$$

$$\hat{d}_{SAN} = d_{SAN} + \omega_{TOA} * \mathcal{N}\ (0,1) + \varepsilon_{NLOS} \quad (2)$$

where $\hat{d}_{BS}$ is the estimated distance between the base station and the UMT, d_{BS} is the true distance between the base station and the UMT. Similarly, $\hat{d}_{SAN}$ and d_{SAN} are respectively the estimated and true distances between the SAN and the UMT. Also, ε_{NLOS} is the error due to the non-line of sight (NLOS) following an exponential distribution with a probability distribution function (pdf) [21] of p(b) given below [22] .

$$p(b) = \begin{cases} \lambda e^{-\lambda b}, & b \geq 0 \\ 0, & b < 0 \end{cases} \quad (3)$$

where b denotes the NLOS error and $E(b)=1/\lambda$

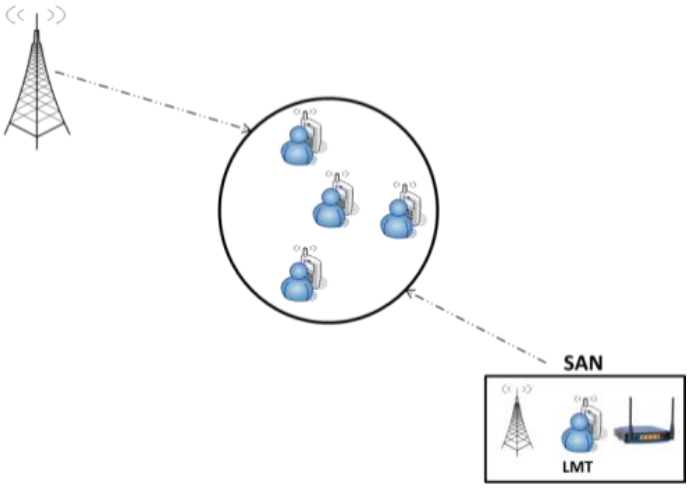

Figure 4 Positioning Scenario with lack of hearability

Using ToA technique, we should be able to get the distances from different ANs to the UMT. Denote A and B the two possible points provided by the ToA approach, one of which is the true position and the other one is an ambiguity point. These solutions could be

obtained by finding the intersection points of the two circles of respective centers O(0,0) and $\hat{a}(\hat{a}_x, \hat{a}_y)$ and of respective radius $\hat{d}_{BS}$ and $\hat{d}_{SAN}$. These intersection points could be easily written as [12]:

$$A = \begin{cases} x = \dfrac{y_{SAN}\left(\gamma - \sqrt{(\ \gamma^2 - 4\beta\theta)}\right)}{2\beta x_{SAN}} + \alpha \\ y = \dfrac{-\gamma + \sqrt{(\ \gamma^2 - 4\beta\theta)}}{2\beta} \end{cases} \tag{4}$$

$$B = \begin{cases} x' = \dfrac{y_{SAN}(\gamma + \sqrt{(\ \gamma^2 - 4\beta\theta)})}{2\beta x_{SAN}} + \alpha \\ y' = \dfrac{-\gamma - \sqrt{(\ \gamma^2 - 4\beta\theta)}}{2\beta} \end{cases} \tag{5}$$

$$\beta = 1 + \frac{y_{SAN}^2}{x_{SAN}^2}, \quad \gamma = -2\frac{y_{SAN}}{x_{SAN}}\alpha, \ \theta = \alpha^2 - d_{BS}^2$$

where x_{SAN} represents the abscissa of SAN, y_{SAN} is the ordinate of SAN, d_{BS} is the distance between the UMT and the base station and d_{SAN} is the distance between the UMT and the SAN.

In order to solve for the most suitable location, the RSS fingerprinting ambiguities points could be used in combination with the ToA intersection points. The proposed solution could be obtained by taking the midpoint between the closest RSS ambiguity point to one of the solutions A and B obtained in, say S. If the ToA error variances are available then a weighted combination could be applied.

C. Coupling and Decoupling Algorithm

In [13], a 2-level Kalman filter was proposed to deal with short range and long range measurements and to come-up with a suitable positioning solution. The main problem of [13] is that the authors were based on the long range measurements using three BSs (in addition to the short range measurements) from the cellular network to estimate the UMT position. Unfortunately, this scenario is quite ideal for investigation and the application of the 2-level Kalman filter would be with restricted impact. The main question would then be how much the Kalman filter and the combination of heterogeneous networks would improve the positioning estimates in critical scenarios. To deal with this scenario, we propose to firstly apply the coarse positioning estimation as shown in Figure 6 and then the coupling-decoupling algorithm proposed in [13] as shown in Figure 7. Contrarily to [13], we assume here that all UMT are moving with a velocity *v*. In addition, we assume the non-linear RSS model is described by:

$$\hat{p}_{k|k-1}^{(i,j)} = \alpha - 10\beta \, log\left(d^{(i,j)}\right) + \varepsilon_{NL,RSS} \tag{6}$$

where $\hat{p}_{k|k-1}^{(i,j)}$ is the received power at the UMT j from the AN i, $d^{(i,j)}$ is the distance between the two nodes, α is a variable taking into account the shadowing effect and β is the path loss exponent. $\varepsilon_{NL,RSS}$ is a measurement error taking into account the mismatch between the path loss model and the real measurement. Contrarily to [13], we propose in this work to firstly decouple the relative localization, realized by using the short-range measurements obtained via the SAN, and the absolute localization, realized by using the long-range cellular by estimating the coordinates and orientation of the group in the cellular

network. After decoupling the short and long range measurements, UKF is applied on each measurement to improve the positioning and tracking estimates. UKF is governed by a transition function and an observation function as shown in (7) and (8). The transition function f is based on the state x_t at time *t,* and the state x_{t+1} at time *t+1*. This relation between the two states includes is a transition noise q_t based on a predefined transition model. On the other hand, the observation function h relates between the observation y_t at time *t,* the state x_t, and the observation noise r_t at time *t* based on a predefined observation model. In UKF, q_t and $r_t$are normally distributed with zero mean and covariance Q_t and R_t respectively.

$$x_{t+1} = f(x_t) + q_t \tag{7}$$

$$y_t = h(x_t) + r_t \tag{8}$$

Then, a coupling of the long and short range measurements is applied again to obtain the absolute position of the UMTs. Additionally, the coupling and decoupling are applied on the center of mass of UMTs in order to reduce overhead cost. The proposed framework runs in a cyclic way where iterations are performed whenever observations are available. The framework considers the following steps:

- Assume that UMT 1 is the reference and MT 2 is on the x-axis of the relative coordinate system as shown in Figure 5.
- Find the absolute coordinates of different MTs through the cellular network.
- Decoupling into relative coordinates and center of mass coordinates through a transformation of coordinates.

- Find the Current Transformation Matrix (CTM) that corresponds to a translation followed by a rotation of the axis.

- Assuming that the absolute and relative coordinates of a MT *i* are defined by $x^{(i)}$ and $x^{(i)rel}$ respectively and that T_{ctm} is the CTM of the transformation of coordinates, we can write the following:

$$\begin{bmatrix} x^{(i)rel} \\ 1 \end{bmatrix} = T_{ctm} \begin{bmatrix} x^{(i)} \\ 1 \end{bmatrix} \tag{9}$$

  Knowing that we added the last component of the vector on the right-hand-side in order to perform transformations independent of $x^{(i)}$.

- T_{ctm} is obtained by defining a translation equivalent to absolute position $x^{(1)}$ of MT 1 followed by a rotation equivalent to the angle of the segment between MT 2 and MT 1 with respect to the absolute coordinate system ($\theta = \arctan(y^{(1,2)}/x^{(1,2)})$)

$$T_{ctm} = \begin{bmatrix} \frac{x^{(1,2)}}{d^{(1,2)}} & -\frac{y^{(1,2)}}{d^{(1,2)}} & 0 \\ \frac{y^{(1,2)}}{d^{(1,2)}} & \frac{x^{(1,2)}}{d^{(1,2)}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -x^{(1)} \\ 0 & 1 & -y^{(1)} \\ 0 & 0 & 1 \end{bmatrix} \tag{10}$$

  knowing that $x^{(1,2)} = x^{(2)} - x^{(1)}$, $y^{(1,2)} = y^{(2)} - y^{(1)}$ and $d^{(1,2)}$ is the Euclidean distance between MT 1 and MT 2.

- Depending on whether available measurements are from short- or long-range technology, we apply

- A single iteration of a stochastic filter for estimating relative localization.
- A single iteration of a stochastic filter for estimating the coordinates of the center of mass.
- Little number of iterations in the case of mobile scenarios.

- Coupling estimations of relative and center of mass coordinates in order to have absolute estimators by doing the inverse of (9).
- It is worth reminding that in our scenario we assume that all UMT are moving and then a recursive process is applied on the coupling and decoupling algorithm.

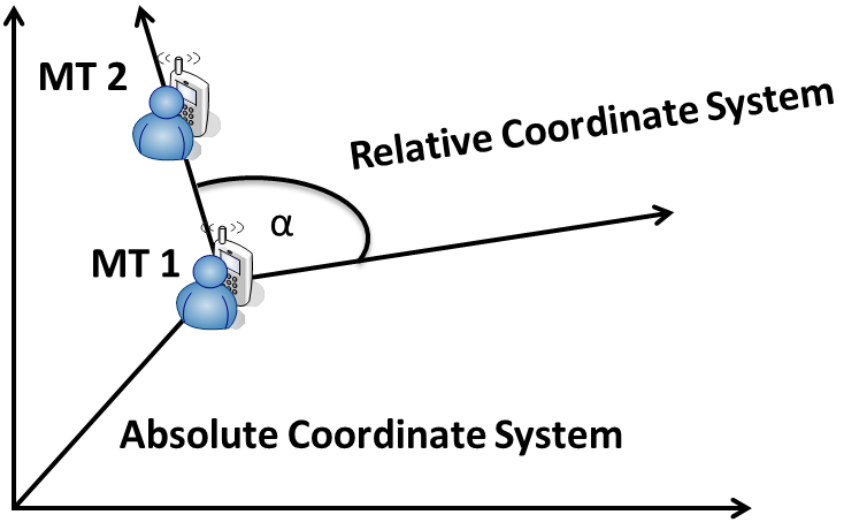

Figure 5 Coordinate System Transformation

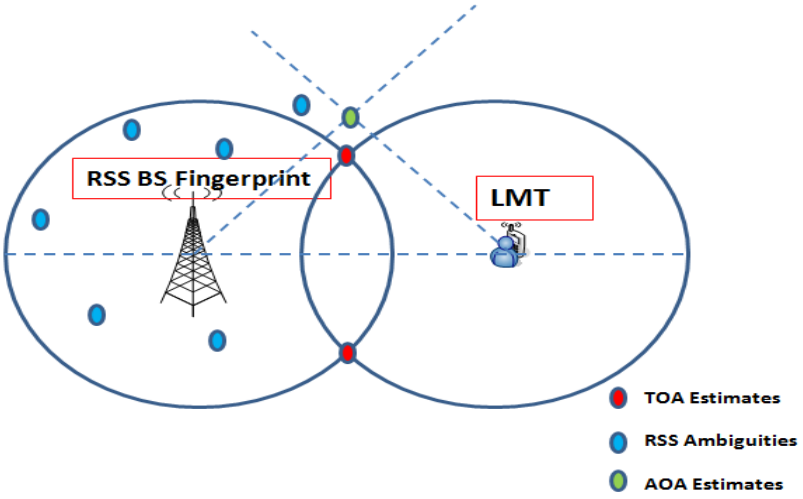

Figure 6 Combination of TOA, AOA, and RSS Fingerprints

We remind that the coupling and decoupling are applied on the center of mass of UMTs. However, this requires defining this center for all MTs in a proper way. Moreover, the coupling and decoupling will be with a limited gain due to the possible large distance among UMTs. Hence, we propose in this work efficient clustering algorithms where the center of mass and the clusters could be efficiently defined so that the coupling/decoupling becomes more interesting. Therefore, the coupling and decoupling will be applied on few centers while keeping the overhead cost reduced. In this work, we propose different clustering algorithms and we compare them with the advanced genetic based clustering algorithms.

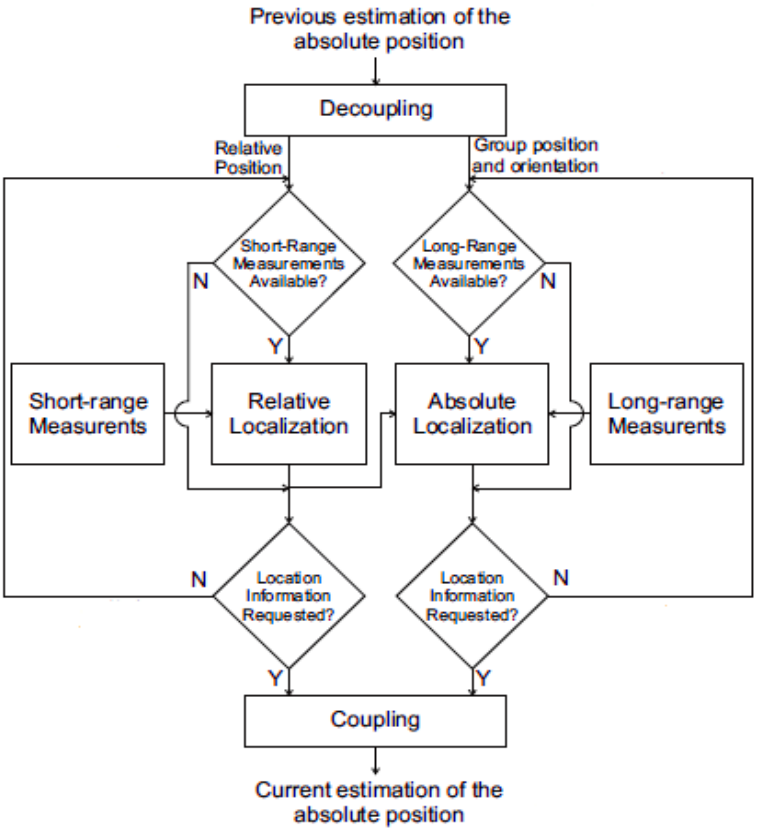

Figure 7 Coupling and Decoupling Algorithm [13]

D. Proposed Clustering Algorithms

The main objective of this section is to describe and propose efficient clustering algorithms suitable for coupling and decoupling the short and long range measurements. Basically, the proposed clustering procedure is used to segregate MTs in such a way to improve the positioning estimation while reducing the overhead cost.

In our work, we introduce three clustering techniques: the RSS based clustering, the hierarchical clustering, and the genetic algorithm based clustering.

1. RSS based Clustering

The RSS clustering algorithm is based on dividing the MTs into 2 clusters. The first cluster center is the closest MT (highest RSS) to the serving BS (or AN) and the second cluster center is the farthest MT (lowest RSS) from the serving BS.

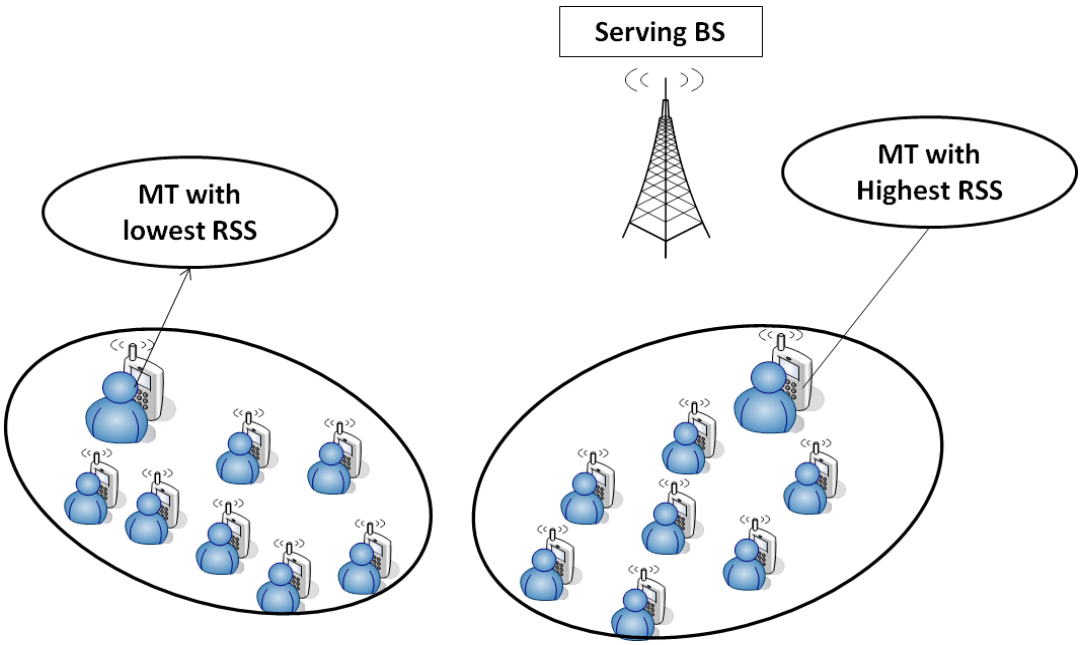

Figure 8 RSS Clustering Model

2. Hierarchical Binary Clustering

The aim of hierarchical clustering is to build a hierarchy of clusters from a binary tree. Based on the long range measurements, a matrix $D = \left(\{d_{i,j}\}; i = 1, \dots, N; j = 1, \dots, N\right)$ of the distances between all MTs is formed. The first step in hierarchical clustering is to search for the pair of MTs that are the closest in terms of Euclidean Distance. Then, a single linkage method based on the Euclidean distance between all the mobile terminals is proposed and implemented. It is again based on the selection of MTs pairs with the smallest distance. The point at which the pair of MTs is joined is called a node. Then, we repeat these steps over all MTs until we form a Hierarchical Binary Tree (HBT) as shown in Figure 9. Basically, using this method, the distance between the merged pair and the other MTs will be the minimum distance of the pair in each case. For instance, if the distance between MT2 and MT1 is 5, while the distance between MT3 and MT1 is 6.5; thus, we choose the minimum of the two, 5, to quantify the distance between (MT2, MT3) and MT1. As a result, we will obtain the binary tree where at each step two MTs were merged. Thus, for *N* MTs, we will obtain *N*-1 nodes. Finally, we are one step ahead from creating the clusters based on this hierarchical clustering. Mainly, we should specify a minimum and a maximum number of clusters. Then, we segregate the MTs into clusters ranging from the minimum till the maximum defined number. However, since in this work we impose a condition of at least 2 mobile terminals per cluster (to perform the combination between the short range and the long range measurements), the condition on the maximum allowed number of clusters is removed.

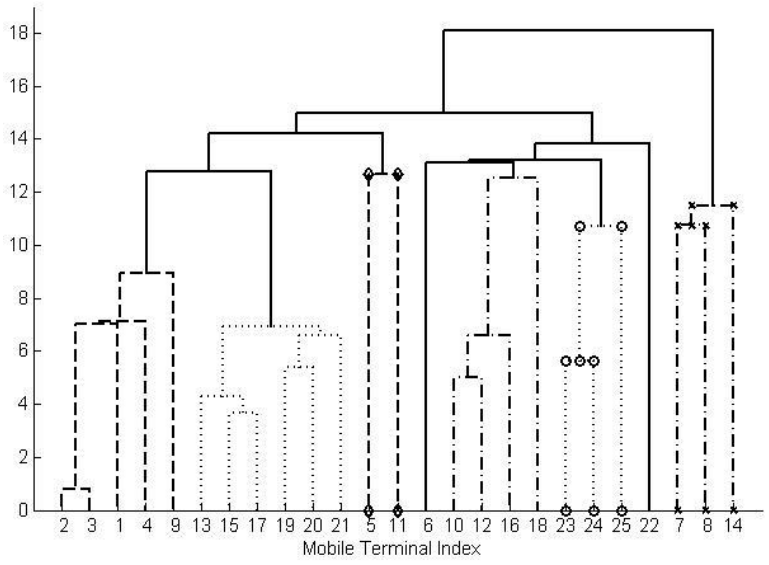

Figure 9 Hierarchical Binary Tree

3. Genetic Clustering

In genetic based algorithms, six main components have to be necessarily defined for establishing the similarity with the wireless cellular networks. This includes the definition of the following:

- Genotype
- Population Initialization
- Fitness Function
- Selection Operator
- Crossover operator

- Mutation Operator

A Genotype is the genetic makeup of a cell, an organism, or an individual [23]. By making an analogy from genetics to cellular networks, the chromosome of N genes represents a cell in cellular network with N MTs. Hence, a gene in a chromosome is mapped to a position of a mobile user in a cell. Secondly, the initialization of a population in genetics can also be mapped on cellular networks in such a way that the position of each UMT is selected randomly in the studied area. The fitness function is a third factor in genetic algorithms. It describes the objective function designed to satisfy some conditions such as minimizing the mean square error. In [23], Azimi et al. introduced an algorithm with two stages of fitness functions: the intra-cluster fitness and the extra-cluster fitness. Thus, the final fitness value is calculated by subtract *Intra-Clstr-Fit* from *Extra-Clstr-Fit*. The selection operator in genetics process selects individuals from the mating pool directed by the survival of the fitness concepts of natural genetic systems. Finally, the crossover operator in genetics represents a probabilistic process that generates at least two child individuals by exchanging information between two parent individuals. Therefore, the pathway of the analogy from genetics to cellular network consists of the following main points:

- The Genotype is a kind of representation for the mobile users in the cellular network.
- The population initialization for cellular network could be the initialization of the mobile users in the network after setting their positions.

- The fitness function proposed by Azimi et al. [23] could be applied on cellular networks, or another metric could be used such as RSS indicator or Euclidean distance metric.

The selection and crossover operators are the main operators required for aggregating MTs into an optimal number of clusters and finding the most probable solution.

In the following, we describe the genetics based algorithm applied in the heterogeneous context for positioning purposes.

a. Genotype

A chromosome is a set of genes in genetics. In our representation, the gene is a cluster center; hence, the chromosome is a set of cluster centers. Therefore, the total number of clusters determines the size of a chromosome. So, an initial clustering method has to be applied on the mobile terminals in order to proceed with the genetic algorithm. We can use one of the two clustering methods discussed before i.e. the RSS clustering and the hierarchical clustering. The flow chart in Figure 10 presents the proposed genetic algorithm for clustering once an initial clustering technique has been applied.

b. Population Initialization

The chromosome is a set of cluster centers, and the population is a set of chromosomes. Hence, the population in our case is a set of all different combinations of the cluster centers. For instance, in the case of 3 clusters having 5 MTs, 3 MTs, and 4 MTs respectively as cluster members, we will have 5*3*4=60 different combinations of cluster centers. Hence, the population contains 60 chromosomes and the maximum population size will be 60.

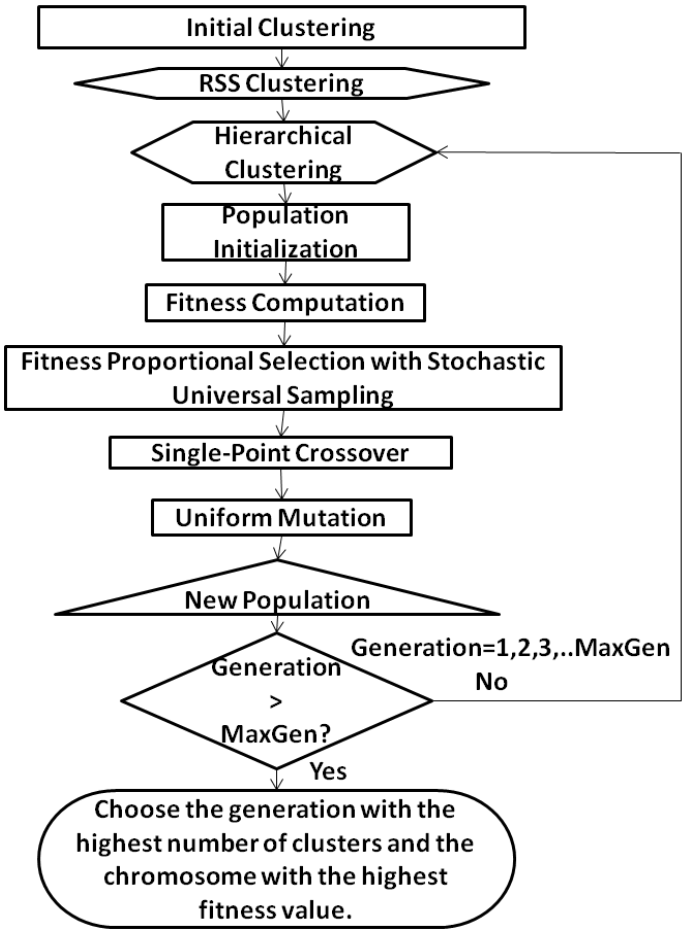

Figure 10 Genetic Clustering Algorithm

c. Fitness Function

Once an initial clustering is applied, the genetics based algorithm requires the definition of the fitness function. In our work, we propose a clustering metric μ_i for each cluster 'i' defined by:

$$\mu_i = \sum_{x_j \in C_i} \|x_j - z_i\| \tag{9}$$

where x_j is cluster member, z_i is the cluster center and $\|.\|$ is the Euclidian distance. Then, we define μ as the sum of the Euclidian distances of all clusters by:

$$\mu = \sum_{i=1}^{K} \mu_i \tag{10}$$

where K is the total number of clusters. The fitness function is defined by:

$$Fit = \frac{1}{\mu} \tag{11}$$

In our clustering approach, we aim at maximizing the fitness function or equivalently maximizing the Euclidian distance among clusters. The fitness function plays a central role in the clustering process.

d. Mutation

We pre-generate two "repositories" of random binary digits from which the masks used in mutation and crossover will be picked up. Then, a 'xor' operation is implemented

between the population and the new mask. This mutation will lead to the initiation of a new population with the same size but different binary representations, and that is the aim behind mutation. After the mutation, a new population is evolved. The population size is originally determined by the multiplication of the number of MTs in each cluster. Thus, we apply a demapping from binary to decimal. Then, the maximum value obtained in the demapping will be the new population size. Then, a factorization will be applied as it leads to the new maximum number of the clusters. For instance, the new population size is 90; the factorization will result into 2*3*3*5=90. Hence, the new maximum number of the clusters will be 4.

e. Genetic Algorithm over multiple Generations

The latter will give us the maximum possible number of clusters with at least two mobile terminals per cluster. Consequently, the entire genetic algorithm is run over many generations. At the end of each generation, we select the chromosome that has the highest fitness value. Finally, at the end of all the generations we select the generation with the highest number of clusters, and the chromosome with the highest fitness.

E. Simulation Results

The aim of this section is to evaluate the hybrid positioning technique in parallel with the proposed clustering approaches. In our simulations, we compare between the long range positioning estimation obtained through the BSs and the SAN (i.e. coarse positioning

estimation) and that obtained after combination with the UKF. Simulations are applied on N = 50 UMTs. They were implemented on the micro cells with 1 Km range placing the BS (equivalently the AN) at position (0,0), and the SAN at $(\sqrt{3}, 1)$ km or $(\sqrt{3}, -1)$ km. Along this work, the BS with position O(0,0) will be used as a reference. Without loss of generality, we assume that all MTs are uniformly distributed around a cluster head, assumed to be the first Mobile Terminal (MT 1), within a range radius of 50 m. The simulation was run for 3000 realizations. The number of iterations of the UKF is equal to 2 while the number of observations is equal to 5. It is worth mentioning that the positioning estimation could be improved if the number of iterations is increased. All results are given in terms of Mean Square Error (MSE).

Figures 11 to 15 show that, independently of the parameter, the proposed hybrid data fusion using UKF coupling/decoupling significantly outperforms the stand-alone long range measurements. In all these figures, no clustering has been applied. Figure 11 shows the improvement made by UKF in combining long range and short range measurements just by considering 5 observations per iteration. Figure 12 presents the effect of the number of ambiguities obtained from RSS fingerprints on the performance of the positioning algorithm. As expected, it is clear that a higher accuracy is obtained when the number of the RSS ambiguities increases. This is true for both long range measurements and the combination through UKF. In both figures, a noticeable positioning improvement is observed. Figure 13 explores the effect of the NLOS parameter λ on the accuracy of our positioning algorithm. λ is inversely proportional to accuracy. As shown in this figure, the MSE strictly increases from 150 till 280 for long range measurements and from 130 till 250

for the combination, when λ increases from 0 till 0.9. Thus, the improvement made by the combination using UKF is still preserved.

Figure 14 explores the effect of the error obtained by GPS measurements. As the GPS error variance increases from 0 till 10, the MSE for the long range measurements and for the combination slightly increases. This means that the utilization of a LMT as a SAN does not really affect the positioning estimation. Similar conclusions could be drawn from Figure 15.

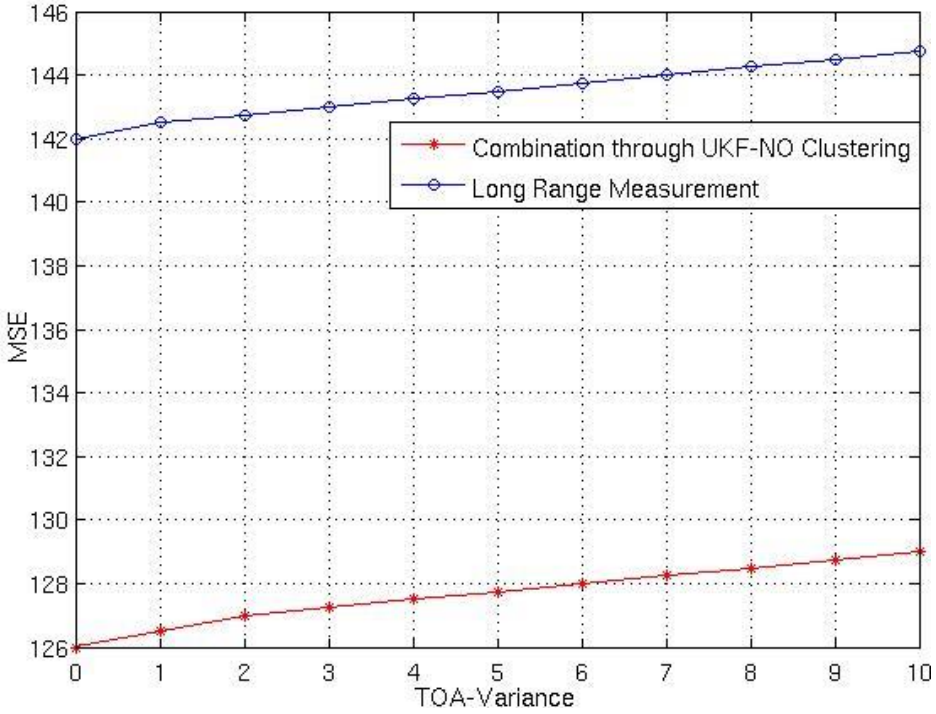

Figure 11 No-Clustering: MSE vs. TOA-Variance

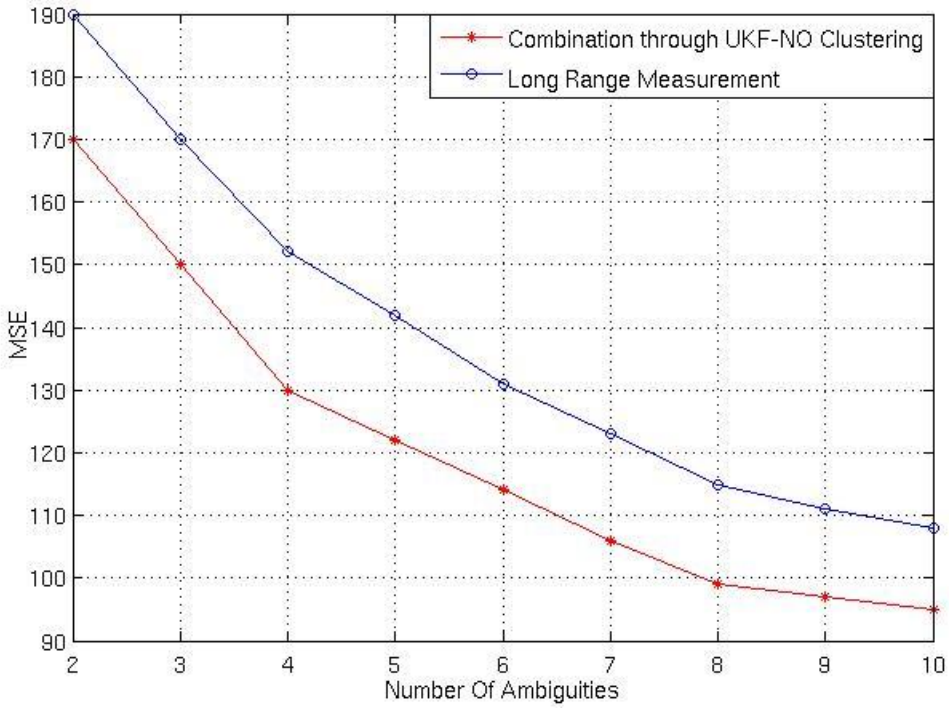

Figure 12 No-Clustering: MSE vs. Number of RSS Ambiguities

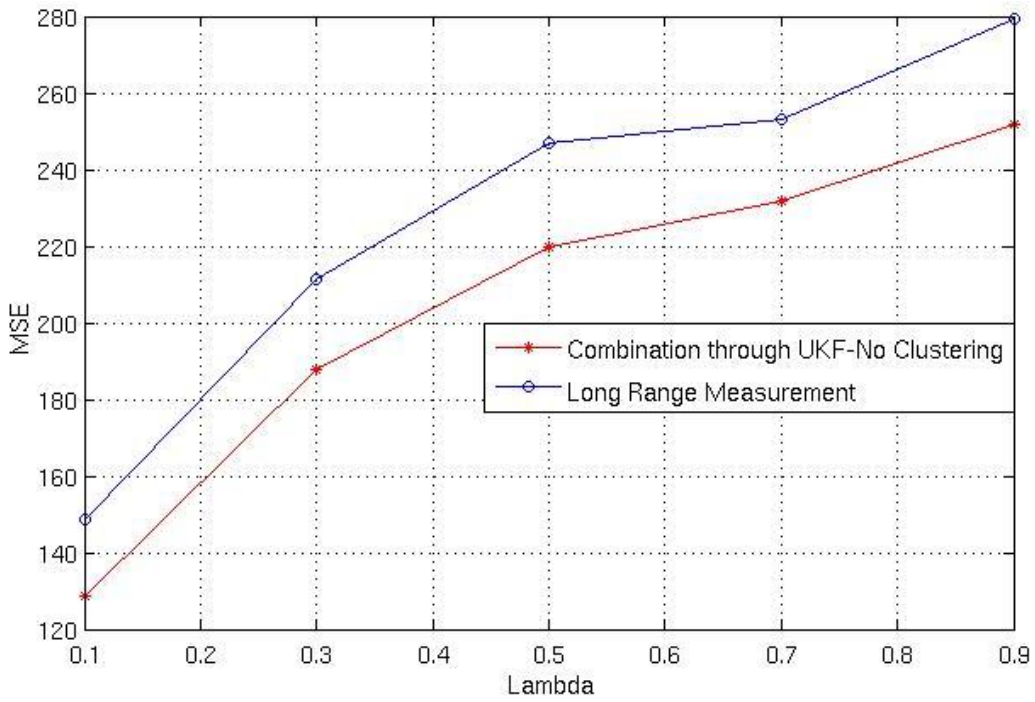

Figure 13 No-Clustering: MSE vs. Lambda

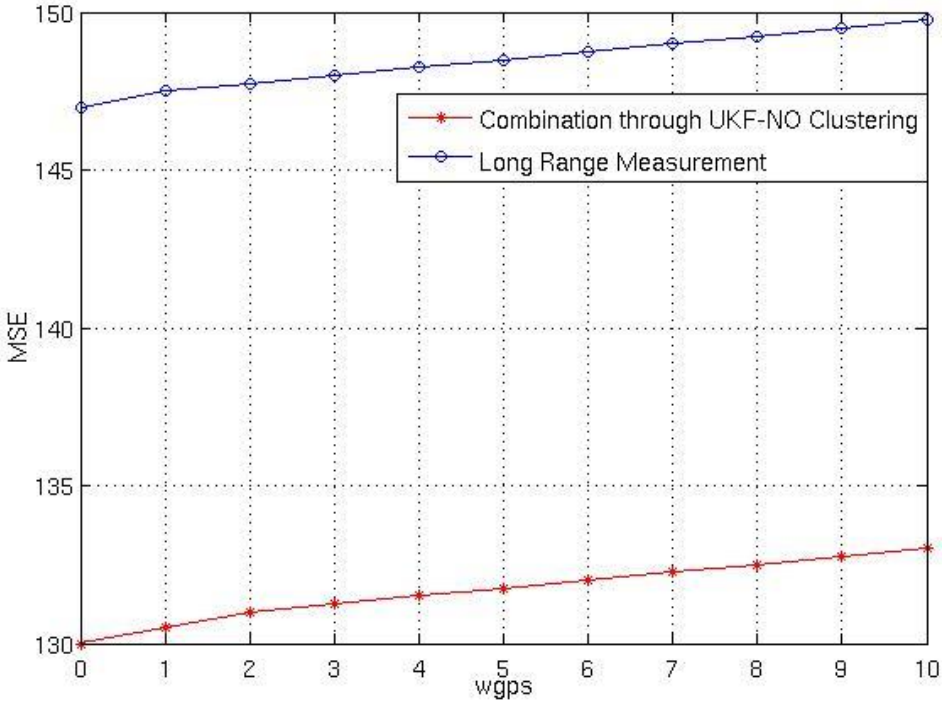

Figure 14 No-Clustering: MSE vs. GPS Error

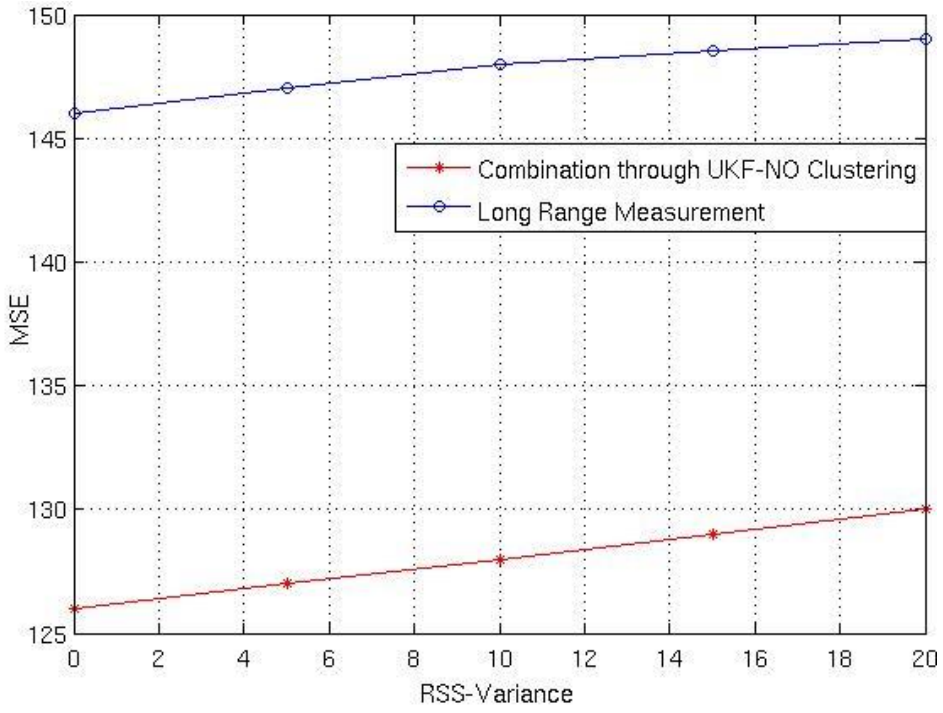

Figure 15 No-Clustering: MSE vs. RSS Variance

Figures 16 to 18 present the effect of clustering on the obtained results. It is clear that, independently of the clustering algorithm, some gain is obtained in the simulated scenario. Figure 16 shows that the RSS clustering provides a little improvement over the No-Clustering case. This is compatible to our expectations due to the fact that RSS clustering leads to the formation of two clusters. That explains our motivation behind implementing a clustering technique based on more advanced approach.

Figure 17 shows that the hierarchical clustering provides a significant improvement over the RSS clustering and over the No-Clustering case. This improvement is due to the fact that the hierarchical clustering leads to the formation of higher number of clusters. Compared with Figure 11, the MSE reduces from around $125m^2$ till 82 m^2 while the RSS clustering approach presents a MSE around $115m^2$.

Figure 18 presents the performance of the genetic based clustering. It is clear that, as for the hierarchical clustering, it presents a large improvement. Nevertheless, this gain is comparable to that obtained through the hierarchical clustering. This is due to the fact that the genetic based clustering implicitly utilizes hierarchical clustering. However, the added value provided is that genetic based clustering generates the highest possible number of clusters whose cluster centers attained the highest fitness value. However, for larger number of MTs, we expect a significant improvement made by the genetic based clustering as it leads to the formation of larger number of clusters than that generated by the hierarchical clustering. Moreover, clusters centers will be chosen in a better way since they are based on the maximization of the clusters centers.

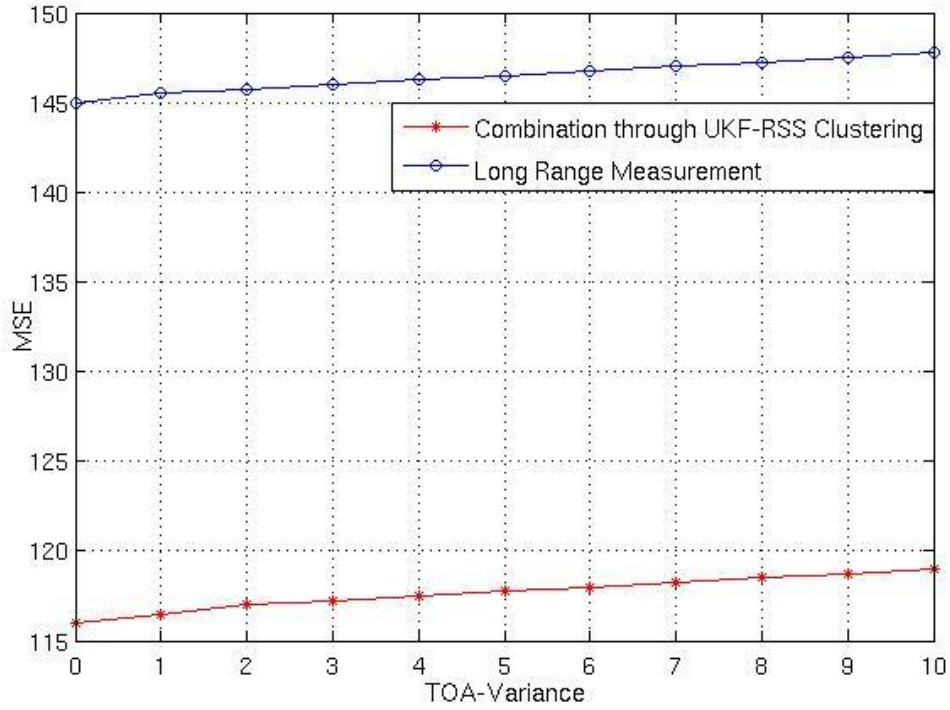

Figure 16 RSS Clustering: MSE vs. TOA-Variance

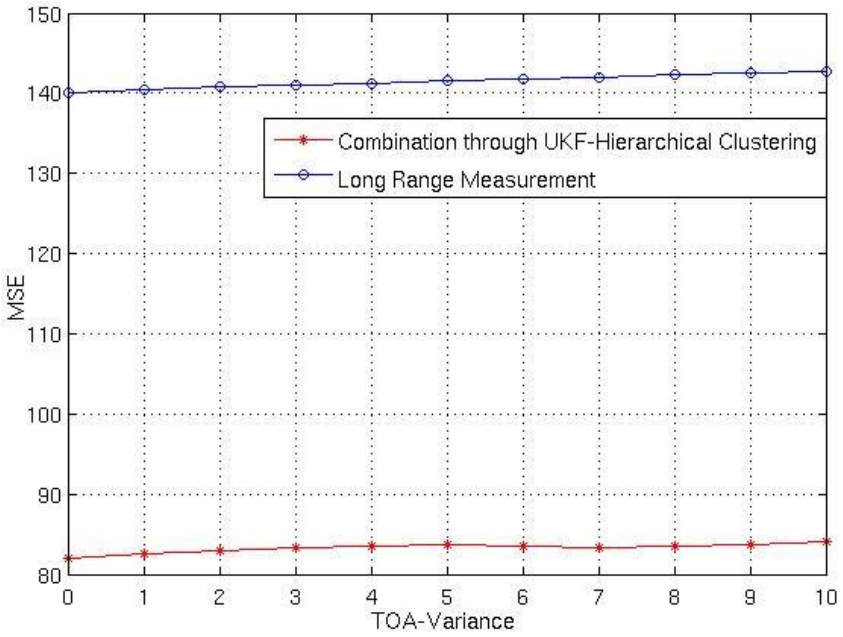

Figure 17 Hierarchical Clustering: MSE vs. TOA-Variance

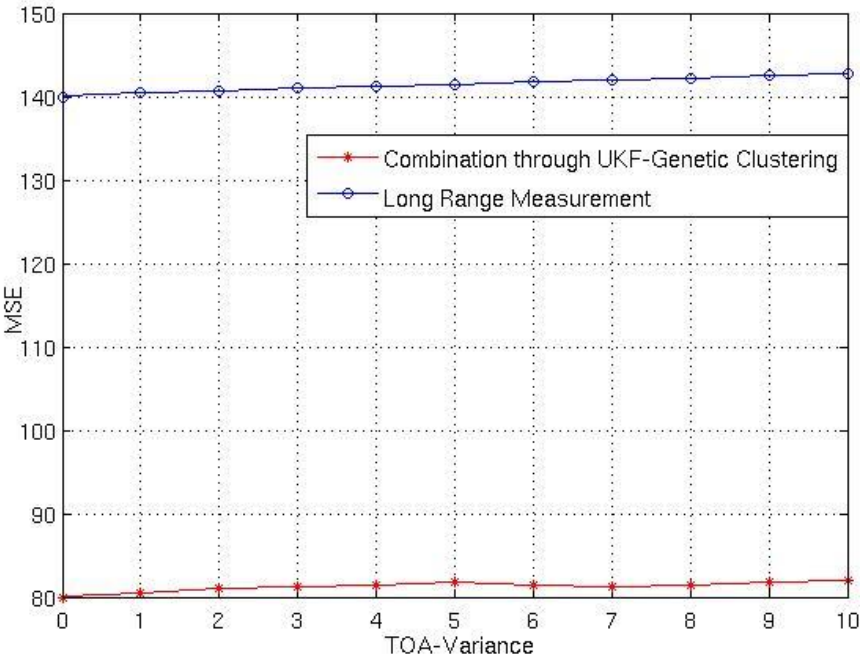

Figure 18 Genetic Clustering: MSE vs. TOA-Variance

F. Conclusion

We presented in this chapter the coupling/decoupling algorithm that merges long range measurements with short range measurements. We applied UKF on each measurement to improve the positioning and tracking estimates. It is shown through simulations that the proposed hybrid data fusion algorithm provided significant improvement in terms of position estimation accuracy. Then, we proposed RSS based clustering, Hierarchical Binary clustering, and genetic clustering for the sake of obtaining higher position estimation accuracy. Simulations proved that genetic and hierarchical binary clustering techniques showed higher accuracy enhancement when compared to the proposed RSS clustering technique.

CHAPTER III

HYBRID POSITIONING DATA FUSION USING UNSCENTED KALMAN FILTER WITH LEARNING APPROACH

A. Introduction

Different tracking filters have been proposed and applied to handle the mobility of the MTs. Kalman Filter (KF) [24] is an adaptive recursive solution for the estimation of linear Gaussian process based on Least Square Error (LSE). Nevertheless, non-linear processes are not handled with the standard KF solutions. Extended Kalman Filter (EKF), Second-order extended Kalman Filter (SEKF) [25], and Unscented Kalman Filter (UKF) are examples from the family of Kalman filters. EKF is the first extension of the KF; it performs a first order linearization of the nonlinear system and then applies the conventional KF estimation. SEKF preserves the second-order terms via the Taylor series development of the transition and measurement equations. UKF is based on selecting a set of sigma points and performing unscented transform. The conditional mean and variance are computed by UKF with third order of accuracy for Gaussian noise. Recently, cubature Kalman filter (CKF) [26] was introduced based on the utilization of a spherical radial cubature rule for estimation of the Gaussian filter. The Monte Carlo based filtering, called also particle filters, handles complex nonlinear systems [27]. Particle filters utilize a large number of independent random variables defined as particles in order to update the

posterior probability. The location, weights, and propagation of the particles are adjusted according to the Bayesian rule.

Recently, nonlinear filters based on probability density function (PDF) have gained some attention since PDF captures all the statistical characteristics of a random variable. Mainly, two kinds of criteria are generally used in PDF filtering: PDF shaping and entropy minimization. PDF shaping is based on picking filtering parameters so that the PDF of the residual follows a narrow distributed zero mean Gaussian PDF [28]. On the other hand, entropy minimization techniques aim at minimizing the entropy of the filtering residual [29]-[31].

The aim of this chapter is multifold. Firstly, we consider a real scenario with lack of hearability and/or insufficient number of ANs. Secondly, we handle the problem of positioning accuracy in such scenarios by adopting the PDF based UKF with Minimum Entropy Criterion (MEC). Thirdly, we tackle the learning phase effect in such scenarios to enhance the positioning accuracy.

The contributions of this chapter are summarized as follows:

- ✓ Proposition of a learning phase for tracking: The contribution here consists in the addition of a learning phase to minimize the error between the measured and predicted position.
- ✓ Adoption of the MEC to reduce the innovation (error) in the UKF tracking filter: The MEC is proposed as it changes the shape of the error PDF to a narrow and peaky shape.

- ✓ Proof-of-concept in realistic scenarios: this includes some mobility models, in contrast with the literature review on this topic.

B. Proposed UKF with learning approach

The UKF filter achieves a good performance under Gaussian noise. However, its performance declines under non-Gaussian noise. On the other hand, PF requires highly complex computations to deal with non-Gaussian nonlinear systems. Thus, the proposed work in this paper is to obtain an optimal estimation based on the MEC in order to diminish the shortcoming of the UKF and PF [32]. Moreover, in most of the UKF based algorithms, the mean square error (MSE) is mainly used to describe the deviation of the estimation from the true value. However, it only accommodates for the second moment due to the second order logarithmic property.

In non-Gaussian environment, the MSE based method is not suitable since higher order moments should be considered in the filter updates. To do so, we propose in this paper to consider the minimum error entropy (MEE) as it allows a good measure of the dispersion of a random variable.

To introduce the MEE or equivalently the MEC in the tracking process, we propose to divide the filter into two phases, the learning and processing phase. During the learning phase, we observe the innovation (error) between the current measurement and the predicted positioning information. We propose to use the Radial Basis Function (RBF) in the learning phase to recompense for the system non-Gaussianity. The RBF is used since it

allows a good description of the error in the UKF filter. Based on the RBF network conducted in the learning phase, we will process the estimated and measured data in the processing phase. This will allow an enhancement of the state estimation and measurement prediction in the UKF filter as shown in Figure 19 and detailed below.

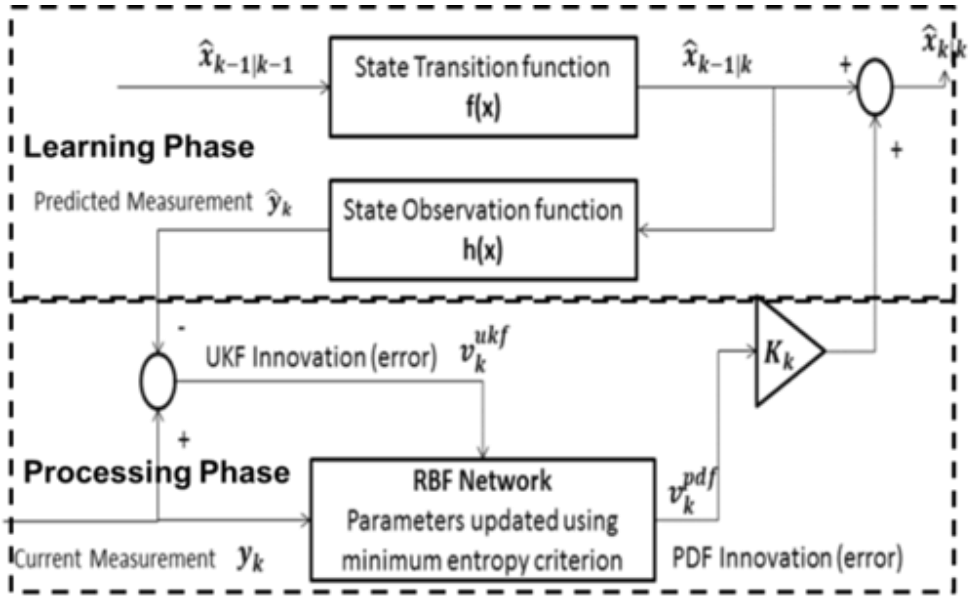

Figure 19 Schematic of UKF based PDF filter

1. *Learning Phase*

This phase is related to the execution of the UKF. Consider the nonlinear system described by equations (7) and (8), which defines the state and the measurement of the position transition and observation functions.

We also define the following terms to describe the error in the UKF as follows:

- $e_k^- = \hat{x}_k - \hat{x}_{k|k-1}$ defined as prior error

- $e_k = \hat{x}_k - \hat{x}_{k|k}$ defined as posterior error
- $v_k = y_k - \hat{y}_k$ (12)

The term v_k known as "innovation" describes the discrepancy between the measurement and the predicted state as shown in Figure 19.

In general, the Kalman gain *K* is obtained from the current state estimate and the relative uncertainty of the measurements. Tuning *K* can be done to attain certain performance. The filter tracks measurements more closely with high gain and tracks predictions more closely with low gain. In our proposed UKF based PDF filter, the UKF gain *K* is obtained so that the posterior error covariance is minimized.

The learning about the innovation is introduced in this phase through a neural network. In our approach, the neural network is based on set of RBF functions referred to as RBF network. The latter is obtained via the multiplication of the radial basis functions with the corresponding weights. Hence, the problem resides here in tuning the RBF weights.

Accordingly, we update the innovation term defined in (12) as follows:

$$v_k = v_k^{ukf} + v_k^{pdf} = y_k - \hat{y}_k + v_k^{pdf} \tag{13}$$

where the input to the RBF network is the UKF innovation given by:

$$I_k = v_k^{ukf} \tag{14}$$

and the neural network update is given by:

$$v_k^{pdf} = W_k^T \Phi(I_k) \; where \; W_k \in \mathbb{R}^{pxm} \tag{15}$$

In (15), W_k^T represents the weight matrix of the RBF-network at time k and $\Phi(.)$ represents the radial basis function which depends on the PDF specifications of the innovation term given in (14). The specifications of the RBF network, i.e. $\Phi(.)$, are described in section A of Appendix I. In our work, the weight matrix is chosen to minimize the entropy criterion. In order to complete the second phase of the UKF based PDF, i.e. innovation estimation update and processing, an estimation of the PDF parameters of the innovation term has to be firstly done so that the RBF functions are constructed. The additional term v_k^{pdf} should have a zero mean in order to achieve unbiased estimation. It remains to define the weight matrix according to the MEC. This will be described in the next section.

2. *Processing Phase*

Based on the PDF estimation of the innovation term, i.e. I_k, and the RBF network, the processing phase consists in updating this innovation according to the MEC. Hence, the randomness and the entropy of innovation are minimized. This will be done according to the following methodologies.

a. Probability Density Estimation

In this work, we use the kernel density estimation (KDE), known also as Parzen window method, to estimate the probability density due to its relationship with the Renyi's Entropy [32]. The estimation of the PDF $f(x)$from the samples given by:

$$\hat{f}(x) = \frac{1}{N}\sum_{i=1}^{N} G_{\sigma}(x - x_i) \tag{16}$$

where $G_{\Sigma}(.)$ is the kernel function with bandwidth σ and $\{x_1, \dots, x_N\}$ are the measurements. The KDE is explicitly described in Section D of Appendix I.

b. Updating the UKF innovation

Mainly, the innovation term could be described by combining (13), (14) and (15):

$$v_k = v_k^{ukf} + W_k^T \Phi\left(v_k^{ukf}\right) = v_k^{ukf} + W_k^T \Phi(I_k) \tag{17}$$

However, a series of measurements has to be taken to obtain the PDF estimation of the innovation term. Thus, a pseudo innovation (PI) is going to be utilized instead of real innovation in solving this problem [22]. At time step *k*, the pseudo innovation is expressed, based on (17), as follows:

$$PI_i^k = v_i + [W_k^T - W_i^T]\Phi(I_i) = v^{ukf} + W_k^T \Phi(I_i) \quad for\ k - N + 1 \le i < k \tag{18}$$

where N is the number of samples used in the estimation of the PDF. So, we obtain PI_i^k, the pseudo innovation term that represents the former error at time step k. PI_i^k shares the same linearly dependent term W_k^T in addition to its simple use in the estimation of PDF and the adaptive tuning process.

c. Updating Weights via Entropy Minimization

The minimum entropy criterion is used in the processing phase to reduce the uncertainty in the innovation term. Ideally, the PDF of the innovation has to be Gaussian with unbounded innovation term or uniform with bounded term [32]. Practically, these results are not achievable because of the contaminated measurement. Hence, this problem will be solved by solving the MEE problem as follows:

$$\min_W H(PI)\ s.t.\ PI = v^{ukf} + W^T \Phi(I) \tag{19}$$

given a sequence of pseudo innovation data PI_i^k and input I_k. $H(.)$ is the Renyi's quadratic entropy and W is the weight of the neural network. The minimization of the entropy criterion could be replaced with the maximization of the information potential given in equation (48) of section B in Appendix I, i.e.

$$\min_W H(PI) = \max_W V(PI) \tag{20}$$

Hence, the locally optimal weight which minimizes the entropy is obtained by:

$$\frac{\partial V(PI)}{\partial W} = 0 \tag{21}$$

Basically, an optimal solution is obtained iteratively for online processing. Hence, the adaptive law is written by utilizing the gradient descent method.

$$W_{k+1} = W_k + \mu \nabla w V(PI^k) \tag{22}$$

Using (51) of section 0 in Appendix I, the potential information of the pseudo innovation at time *k* is calculated as follows:

$$V(PI^k) = \frac{1}{N^2} \sum_{i=k-N+1}^{k} \sum_{j=k-N+1}^{k} G_{\Sigma}(PI_i^k - PI_j^k) \tag{23}$$

Formulating $\Phi(I_k)$ as $\Phi(k)$ for simplicity, the gradient of the information potential is expressed as:

$$\begin{aligned} \nabla_W V(PI^k) &= \frac{\partial}{\partial W}\left(\frac{1}{N^2} \sum_{i=k-N+1}^{k} \sum_{i=k-N+1}^{k} G_{\Sigma}\left(PI_i^k - PI_j^k\right)\right) \\ &= \frac{1}{N^2} \sum_{i,j=k-N+1}^{k} \frac{\partial G_{\Sigma}\left(PI_i^k - PI_j^k\right)}{\partial\left(PI_i^k - PI_j^k\right)} \frac{\partial\left(PI_i^k - PI_j^k\right)}{\partial W} \\ &= \left[\frac{1}{N^2}\Sigma^{-1} \sum_{i,j=k-N+1}^{k} G_{\Sigma}\left(PI_i^k - PI_j^k\right)\right][\Phi(i) - \Phi(j)]\left(PI_i^k - PI_j^k\right)^T \end{aligned} \tag{24}$$

d. <u>The overall Algorithm</u>

The proposed UKF based PDF algorithm with learning approach is summarized as follows:

1. Selection of the UKF parameters.

2. Selection of the kernel bandwidth $\sum$, normally a diagonal matrix, and the selection of RBF network parameters, such as initialization of the weight vector to be equal to zero ($W_1 = 0$).
3. Record the state x_k^{ukf} using UKF method and the innovation term v_k using (17).
4. Update the state estimation during the processing phase by $x_k = x_k^{ukf} + K_k W_k^T \Phi(k)$.
5. Construct the pseudo innovation $PI_i^k = v_i^{ukf} + W_k^T \Phi(i)$ during the processing phase
6. Update W using (22) and (24) by the properly chosen μ.
7. $t + 1 \rightarrow t$, go back to 3)

C. Simulation Results

After evaluating the hybrid positioning technique in parallel with the proposed clustering approaches, we will evaluate the enhancement added on the UKF via introducing UKF+PDF filter. In previous simulations, we compare between the long range positioning estimation obtained through the BSs and the SAN (i.e. coarse positioning estimation) and that obtained after combination with the UKF.

Simulations are applied now on N = 3 UMTs. Without loss of generality, we assume that all MTs are uniformly distributed in a cellular area of radius 1 Km range placing the BS (equivalently the AN) at position (0,0), and the SAN at ($\sqrt{3}$, 1) km or ($\sqrt{3}$,

−1) km. The simulations are run on a very large number of realizations. Along this work, the BS with position O(0,0) will be used as a reference. We ran the simulations for three different mobility models of the MTs:

- Linear variation of only the abscissa of the MTs,
- Linear variation of both, the abscissa and the ordinate of the MTs,
- Sinusoidal variation.

To have a real simulation scenario, we have also considered the case when the MTs are moving according to the different mobility models above but with some noise on the trajectory.

Concerning the RBF parameters, we consider that $\sigma_{rbf}^2 = 0$, $\mu = 10$, and the number of samples for PDF estimation is equal to 50. All results are given in terms of Mean Square Error (MSE).

Figure 20, Figure 23 and Figure 26 (respectively Figure 21, Figure 24 and Figure 27), show the MSE (respectively the CDF of the MSE) of the position estimation using the hybrid algorithm only, data fusion with UKF, and data fusion with UKF based PDF filter with/without a noise imposed on the trajectory of the MTs corresponding to each of the three mobility models. It is clearly shown that the data fusion with UKF highly enhances the MSE measurements. Additionally, the use of the UKF based PDF filter, i.e. with learning approach, enhances the UKF position estimation accuracy. These figures show that the MSE could be reduced to less than 1 m, on both average and CDF. Finally, it is obvious

that imposing noise over the trajectory of the MTs deteriorates the accuracy of position estimation especially with the case hybrid data fusion. This slightly decreases the accuracy of the UKF based PDF approach.

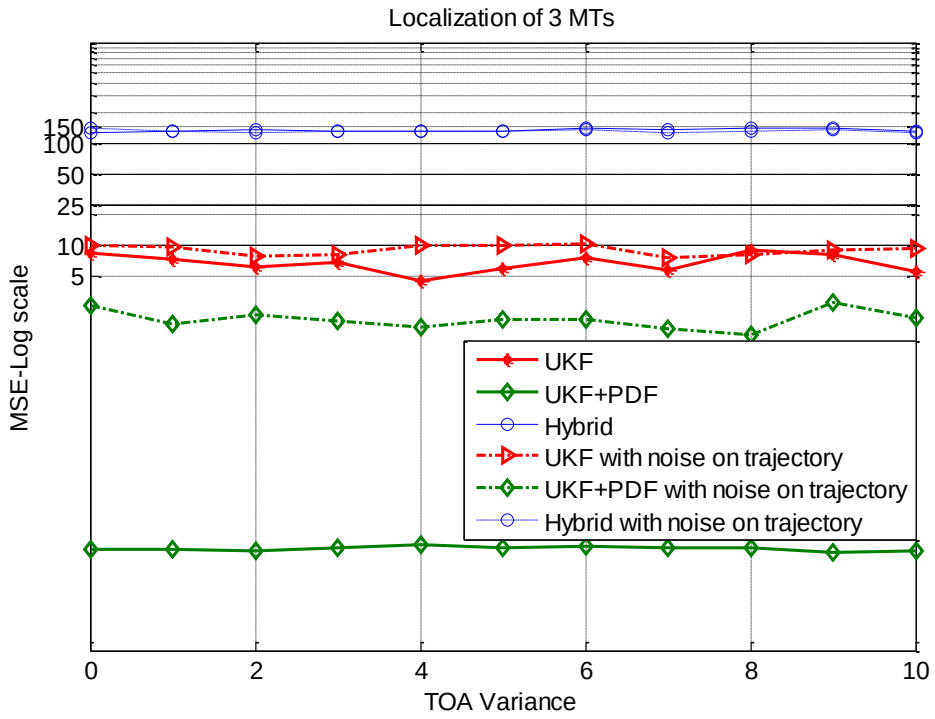

Figure 20 MSE for Model 1

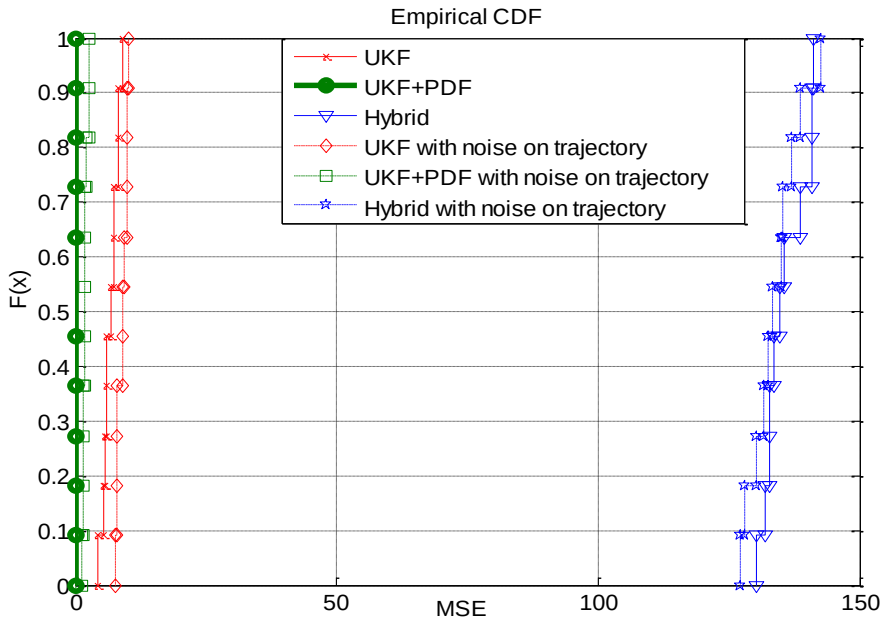

Figure 21 CDF of the MSE for Model 1

Figure 22, Figure 25 and Figure 28 illustrate the PDF of the innovation term before and after applying the learning step (i.e. UKF filter only, and UKF+PDF) corresponding to the three mobility models. It can be directly observed how the probability density of the innovation term is transformed from a Gaussian like one in the case of UKF to a narrow and peaky shape in the case of UKF based PDF, which stands for the minimization of the entropy.

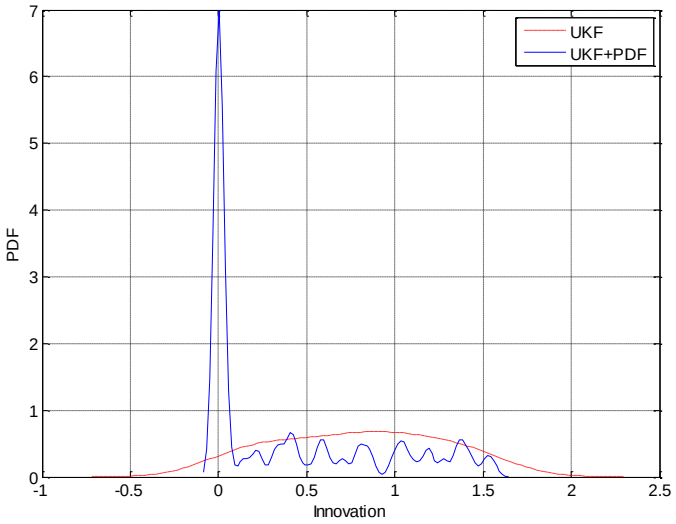

Figure 22 PDF for the innovation for Model 1

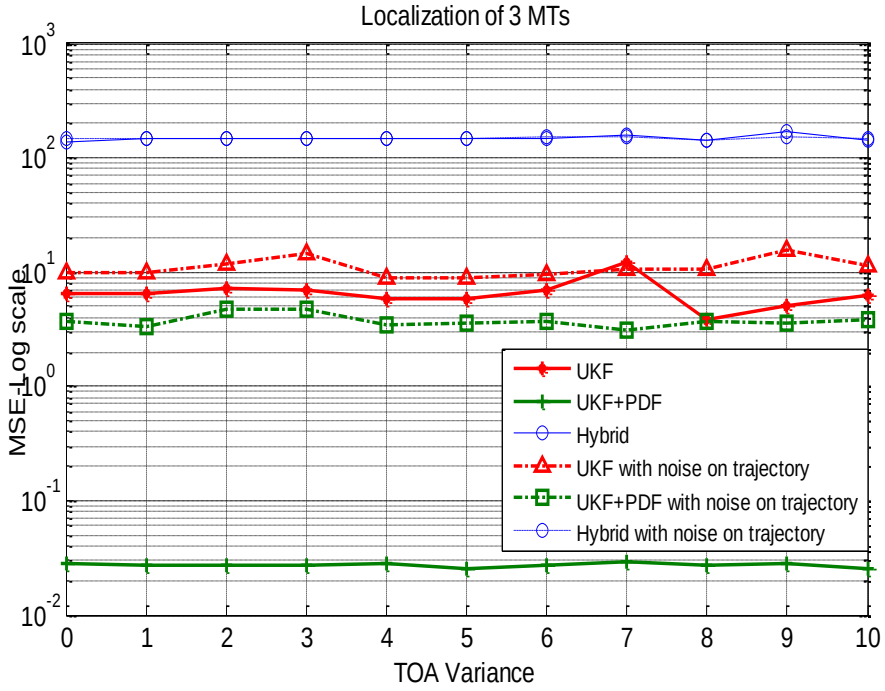

Figure 23 MSE for Model 2

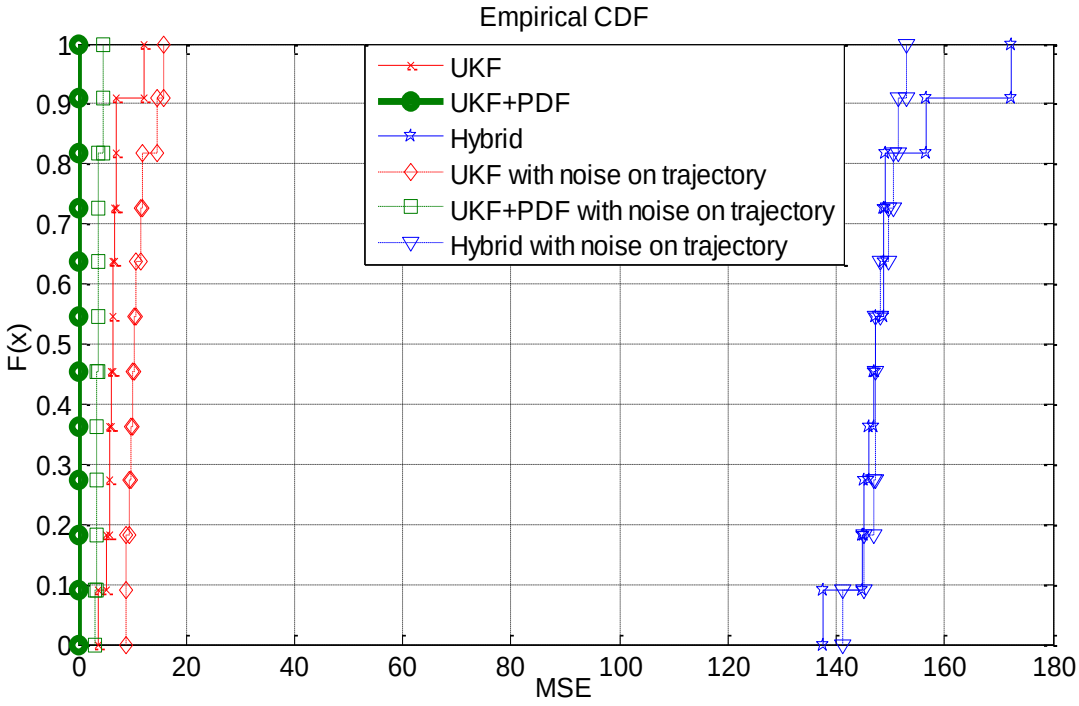

Figure 24 CDF of the MSE for Model 2

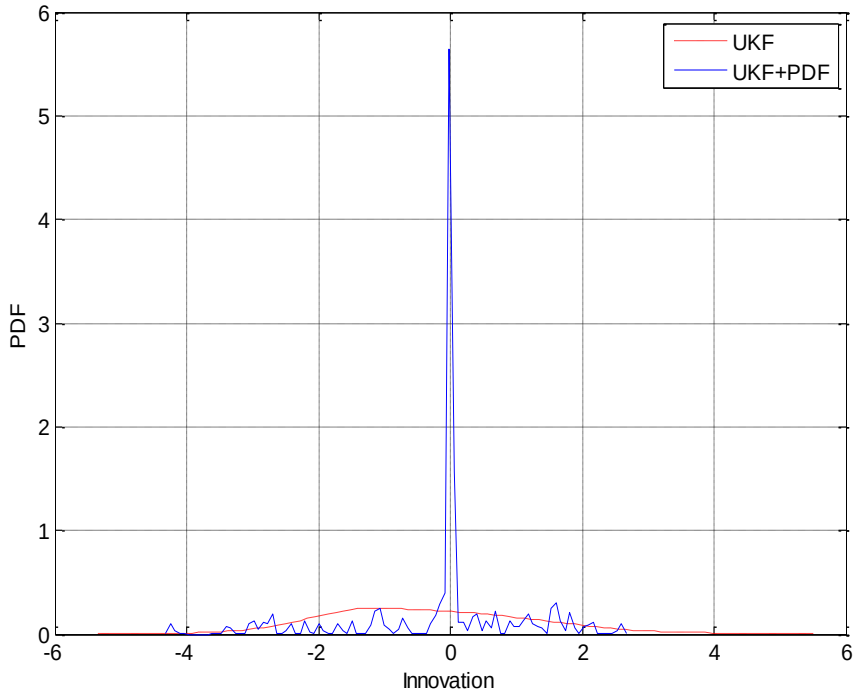

Figure 25 PDF of the innovation for Model 2

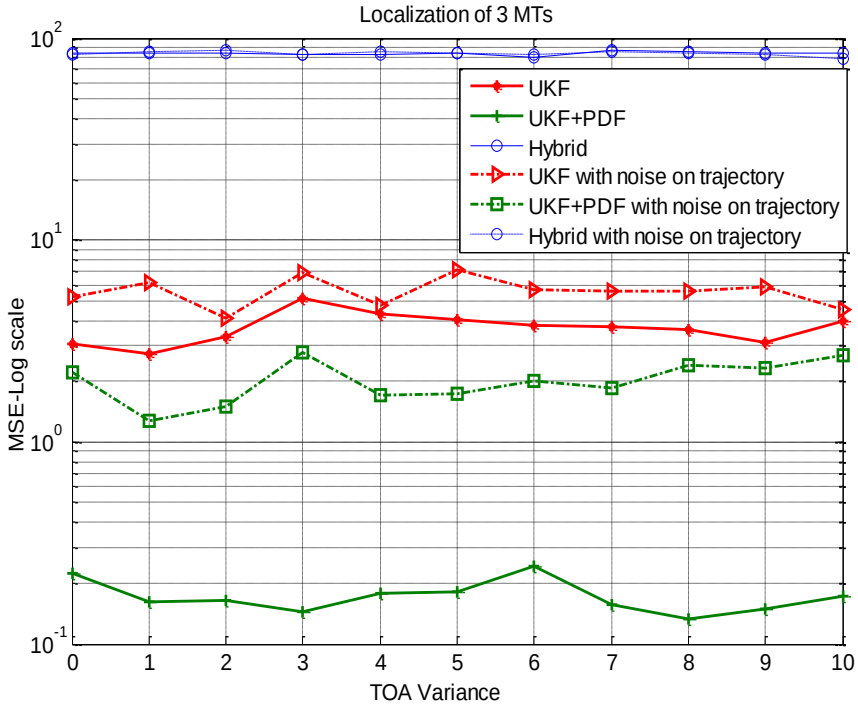

Figure 26 MSE for Model 3

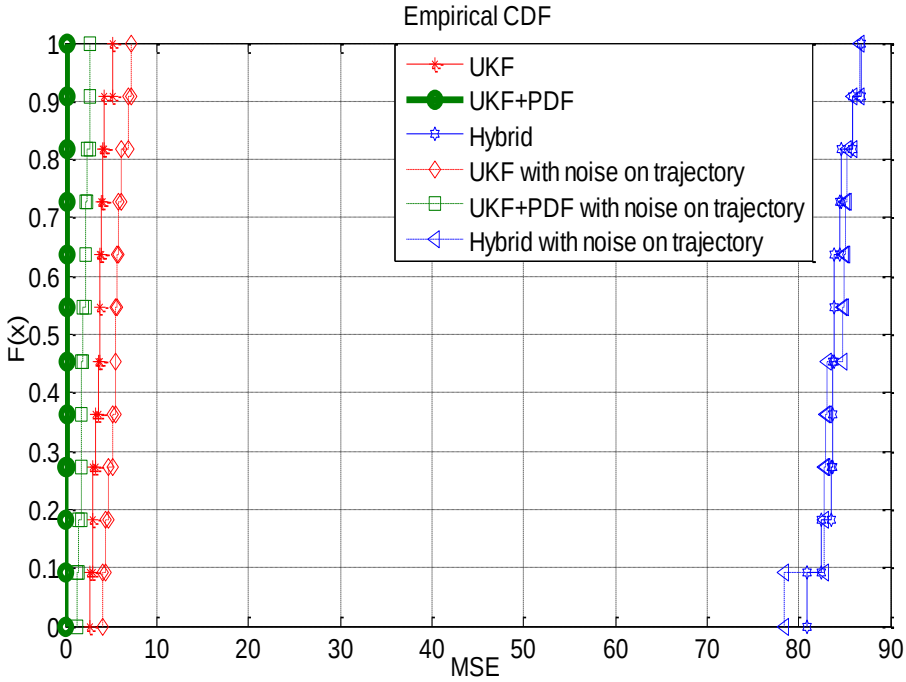

Figure 27 CDF of the MSE for Model 3

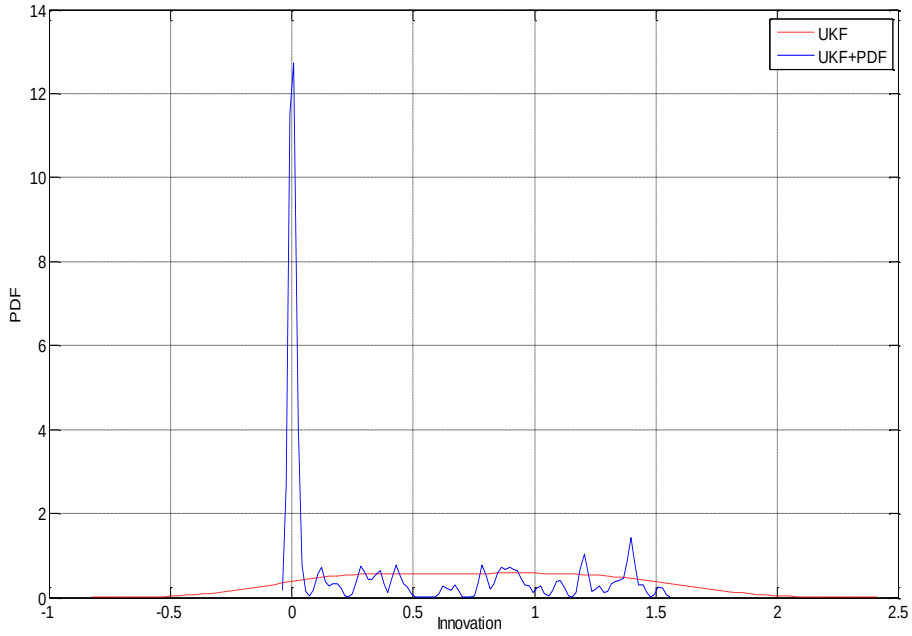

Figure 28 PDF of the innovation on the long range measurements for Model 3

D. Conclusion

In this chapter, we have compared the results of the hybrid localization approach based on the combination of different positioning measurements presented in the previous chapter with the proposed UKF based PDF filter including a learning phase. Simulation results have shown that the proposed positioning technique based on a learning phase has highly increased the positioning accuracy in all considered scenarios. Moreover, we have shown that the addition of noise on the mobility trajectories of the MTs (very) slightly affects the proposed UKF filter with learning phase.

CHAPTER IV

QUASI-SAGE BASED POSITIONING OF COGNITIVE TRANSMITTERS

A. Introduction

Localizing multiple transmitters in a region by making use of the received signal strength is one of the main problems that are assessed by researches nowadays. Even though this problem is applicable in different applications, the most prominent one is the cognitive radio environment [34][35]. Basically, a cognitive radio system can be classified as coordinated or uncoordinated systems [36]. In the coordinated environment, the cooperation with primary systems is necessary to establish a reliable transmission. In the uncoordinated environment, the system should operate in an opportunistic manner by identifying spectral holes, usually under-utilized by the primary systems [37]. The identification of the spectral holes is a corner brick in the cognitive environment. It is based on the measurements in time or frequency or space of the signal (and of its specifications) received from legacy transmitters.

Basically, the opportunistic transmission should be done in such a way it does not interfere with the legacy receivers. Hence, the cognitive system is responsible to recognize the area where there are active primary users. Assuming the location of primary users and their activity are not known, the nodes identify the areas where primary users are active based on the received power measurements. In this domain, we can make use of erroneous

detection for proclaiming spectral hole as the detection might be affected by deep shadowing [38][39]. However, the argument nowadays is not being restricted to simple techniques for detection [40] due to the fact that such techniques leads to conservative transmission. Fundamentally, if prior information about transmitter locations is available, we can achieve more efficient exploitation of the spectrum.

The benefit behind localizing transmitters is mainly significant when nodes in a cognitive network are mobile. Knowing that the detection techniques require power measurements repeatedly at the nodes, the availability of transmitter locations allows a reduction of the amount of power measurements by the cognitive nodes. Hence, the recognition of spectral holes by making use of transmitter localization requires simple tracking of the transmitters' locations.

Transmitter localization can be obtained using trilateration technique by making use of the received power measurements at three locations. However, this technique doesn't generate an accurate solution due to the error generated by the signal propagation model and the imperfections in the sensors. The work proposed in [41] tackled these challenges by proposing the global optimization based on "smart" initial conditions generated by clustering. The authors of [37] handled these challenges by introducing an expectation-maximization (EM) method for localizing multiple transmitters based on the power measurements monitored by a set of arbitrarily-located receivers. Basically, previous solutions have tackled multiple transmitters in an additive white Gaussian noise scenario [41]. The authors of [42] tackled these challenges by proposing a quasi-expectation-maximization (Quasi-EM) technique for localizing multiple transmitters based

on the received power measurements under lognormal shadowing. However, quasi-EM has slow convergence, and it doesn't employ initial conditions.

In this chapter, we will tackle these challenges by proposing the Quasi-Space Alternating Generalized Expectation Maximization (Quasi-SAGE) technique for localizing multiple transmitters. Quasi-SAGE method is based on the observation of the power levels by a set of arbitrarily-placed receivers. The proposed technique is based on a judicious selection of the initial conditions of the transmitters. The main advantages of quasi-SAGE are to diminish the dimensionality of the maximum-likelihood estimation problem and to obtain faster convergence.

B. Literature Review

A better exploitation of the spectrum can be done via using cooperative sensing to generate location estimations of transmitters given the power levels observed by radio nodes. The radius around each transmitter, in which opportunistic communication must be avoided, can be determined after estimating transmitter locations and identifying the properties of legacy systems, government regulations, and the maximum probability of interference.

In this section, we will briefly describe some cooperative communication systems used in cognitive network. Then, we will explore Bayesian tracking approach of primary

users to be used by mobile cognitive radio nodes for identifying spectral hole at a location without employing measurement at that location.

1. *Location Assisted Wireless Systems*

The cognitive radio base station (CRB) creates a radio map of the environment in a CR network for cooperative communications as shown in Figure 29. This map is obtained after defining the location of PUs [43]. Radio environment map permits obtaining assisting CR nodes and dynamic spectrum management within the network in condition that there is no interference between PU and CR transmissions. Mainly, a Radio Environment Map (REM) provides information related to frequency, power, and space. Additionally, making use of REM has two advantages in reducing interference and in terms of radio transmission. First, CR nodes can transmit in directional way in condition that the locations of the PUs in the surrounding are known so that to avoid any interference. Second, the transmission power can be controlled by CR nodes to avoid again any interference with PUs in the surroundings.

2. *Cooperative Localization Model*

The work presented in [43] presents the utilization of cooperative localization in CR network where each CR node generates its own estimates of the time of arrival (ToA) and

direction of arrival (DoA) on the nearest target; then, the obtained estimations are passed to the CRB. The link between CRB and CR nodes is assumed to be an error free channel. Also, the network is assumed to have awareness of the locations of its CR nodes, and CRB is assumed to be the point of origin of the 2D-plane. The system model of this work is shown in section A of Appendix II.

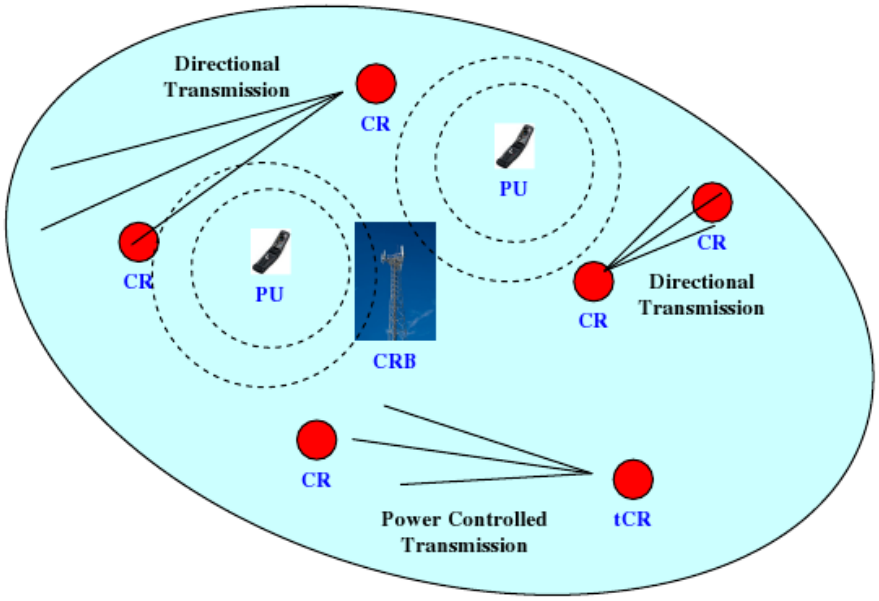

Figure 29 Typical Radio Environment Map and radio scene analysis in CR network [43]

3. *Bayesian Tracking of Primary Users*

When we have knowledge about the *posteriori* distribution of the noise process, Bayesian estimation and tracking is a good candidate to be used in tracking for the sake of improving the accuracy of localization. Particle filtering method is a Bayesian tracking

algorithm [43]; the system model is described in section B of Appendix II. Particle filtering is based on sequential Monte-Carlo process for sequential signal processing. Considering the discrete model of given by $x_j[k]$ and Φ_j[k] with k=0,1,…N, and the observation vector $z_k = \left[x_j[k], \phi_j[k]\right]^T$ for the parameter vector $\mu_k = \left[d_j[k], \Phi_j[k]\right]^T$ at the observation time k, the Bayesian formula for the iterative estimation of the posteriori pdf is given by:

$$p(\mu_k|z_k) = \frac{p(z_k|\mu_k)p(\mu_k|z_{k-1})}{p(z_k|z_{k-1})} \tag{25}$$

Knowing that:

$$p(z_k|z_{k-1}) = \int p(z_k|\mu_k)p(\mu_k|z_{k-1})d\mu_k \tag{26}$$

In the computation of the posteriori density function, the normalizing component of (25) and (26) can be neglected due to the fact that μ_k elements are of first order Markovian processes. As a result, the Bayesian estimate is the expected value of the latest density function $p(\mu_k|z_k)$, given by:

$$\hat{\mu}_k = \int \mu_k p(\mu_k|z_k)d\mu_k \tag{27}$$

C. Problem Statement and the Maximum Likelihood Solution

1. *Problem Statement*

In literature, different spectrum sensing techniques have been proposed to estimate the activity of the primary transmitters hence, the availability of the spectral holes. Alternatively, the exploitation of the spectrum holes can be done by using cooperative

sensing to generate location estimations of transmitters. Hence, based on the transmitter estimated positions, its transmitted power and coverage, a cognitive node can determine if it can transmit or not from a specific location before reaching its destination. The location of the transmitter can be estimated from three power measurements via trilateration when only one transmitter is present. However, when multiple transmitters contribute with unidentified proportions to the received power, the non-cooperative localization technique doesn't provide a direct solution.

To generalize the problem, we consider *M* legacy transmitters and *N* cognitive radio nodes. The locations of the *M* transmitters are assumed to be unknown, and they are denoted by $\theta = [\theta_1\ \theta_2\ ...\ \theta_M\ ]^T$ where θ_i represents the two-dimensional location of the i^{th} transmitter. On the other hand, the *N* sensors are assumed to be with known arbitrary locations. Let $d_j(\theta_i)$ represents the two-dimensional Euclidean distance between the i^{th} transmitter and the j^{th} receiver.

Knowing that $d_j(\theta_i)$ represents the Euclidean distance from the transmitter positioned at θ_i to the j^{th} receiver, the noise-free received power P_{ij} from the i^{th} transmitter to the j^{th} receiver is denoted by:

$$P_{ij} = \rho P_{TX,i}\left(\frac{d_0}{d_j(\theta_i)}\right)^n \quad \text{with } \rho = \rho_o \times \rho_{NLOS} \tag{28}$$

where $P_{TX,i}$ is the power transmitted by the i^{th} transmitter, ρ_o is a constant representing the antenna and carrier frequency properties, ρ_{NLOS} is a constant representing the error due to

the effect of Non-Line-Of-Sight (NLOS), n is constant representing the pathloss exponent, and d_0 is representing the reference distance [44], usually taken equal to 1m.

Assuming independent lognormal shadowing corresponding to each transmitter-receiver pair, the unknown power h_{ij} measured by the j^{th} receiver from the i^{th} transmitter could be modeled as:

$$h_{ij} = P_{ij} 10^{\frac{x_{ij}}{10}} \tag{29}$$

where $x_{ij} \sim N(0, \sigma^2)$ in dB, represents the shadowing variable between transmitter i and receiver j. Hence, the total received power r_j at the j^{th} receiver is given by:

$$r_j = w_j + \sum_{i=1}^{M} h_{ij} \tag{30}$$

where w_j is the Additive White Gaussian Noise (AWGN) with zero-mean and variance σ^2.

2. *The Maximum Likelihood (ML) Solution*

The aim of this algorithm is to determine the ML estimate $\hat{\theta}$ of the locations θ of the M transmitters based on the power measurements obtained at each cognitive node. This can be represented by the following equation:

$$\hat{\theta} = arg \max_{\theta} P(r|\theta) \tag{31}$$

where $r = [r_1\ r_2\ \ldots\ r_N\ ]^T$.

Given the transmitter location θ, the measured power at the sensor denoted by r_j is a Gaussian random variable with mean $\sum_i P_{ij}^R$ and variance σ^2. Let $P_i^R = \left[P_{i1}^R\ P_{i2}^R\ \ldots\ P_{iN}^R\right]^T$ such that r is a Gaussian random vector with mean $\mu_r = \sum_i P_i^R$ and covariance matrix $\sigma^2 I_N$, where I_N is *N*x*N* identity matrix. Thus, the likelihood is obtained by:

$$P(r|\theta) = (2\pi\sigma^2)^{-N/2}\exp\{-\frac{1}{2\sigma^2}\left(r - \sum_{i=1}^{M} P_i^R\right)^T\left(r - \sum_{i=1}^{M} P_i^R\right)\} \tag{32}$$

Thus, the log-likelihood function could be obtained from (32) as follows:

$$L(r|\theta) = -(r - \sum_{i=1}^{M} P_i^R)^T(r - \sum_{i=1}^{M} P_i^R) \tag{33}$$

Using (30) and (33), the likelihood function of unknown received power conditioned on the transmitters' locations θ is computed as the product of MN lognormal densities given by:

$$p(H = h|\theta) = \prod_{i=1}^{M}\prod_{j=1}^{N} \frac{10\,log_{10}\,e}{h_{ij}\sigma\sqrt{2\pi}} e^{-\frac{(10\,log_{10}(h_{ij}) - 10\,log_{10}(P_{ij}))^2}{2\sigma^2}} \tag{34}$$

Basically, the log-likelihood doesn't provide a direct solution due to its complexity as it often has multiple local maxima. Thus, alternative solutions should be proposed to approach the ML technique. In this paper, we tackle this problem through the Quasi-SAGE algorithm as described in the next section.

D. Proposed Quasi-SAGE

1. *From EM based techniques to Quasi-SAGE*

The EM based techniques [45] are efficient for iterative parameters estimation. The classical computation of an EM method consists of supplementing the observed measurements (incomplete data) with a single complete-data space whose association to the parameter space simplifies the estimation. Basically, the EM algorithm alternates iteratively between an E-step and an M-step.

- *The E-step*: It is based on computing the conditional expectation of the complete-data in the log-likelihood.

- *The M-step*: It is based on maximizing simultaneously that expectation relative to all of the unknown parameters, i.e. $d_j(\theta_i)$, ρ and $P_{TX,i}$.

The implementation of the EM based techniques is most beneficial in applications where performing the M-step is simpler than maximizing the original likelihood. Furthermore, updating the parameters simultaneously makes the classical EM technique to converge slowly due to the need for overly informative complete-data spaces. The convergence rate of an EM method has an inverse association with the Fisher information of its complete-data space [45]. Moreover, it is shown that enhanced asymptotic convergence rates are obtained with less-informative complete-data spaces [46]. Additionally, larger step sizes and greater likelihood increases are obtained with less informative complete-data

spaces in the early iterations [47]. Hence, there is a tradeoff between complexity and convergence rate.

In our work, we extend the concepts in [46] to avoid this tradeoff by introducing the SAGE based technique. The latter is applicable in scenarios where parameters update is done sequentially in small groups of elements. In addition, it associates with each group of parameters a hidden-data space instead of utilizing just one large complete-data space. Moreover, the flexible admissibility criterion is introduced to ensure the monotonic increase of the penalized-likelihood objective in the algorithm.

Basically, one of the two motivations for the SAGE technique is the convergence rate. For instance, applications where the parameter space is very large, such as imaging restoration applications, regularizing via smoothness penalties is often necessary. On the other hand, a SAGE method decouples the parameter updates by utilizing a separate hidden-data space. Astonishingly, this technique not only simplifies the maximization, but also enhances the converge rate. Also, SAGE technique guarantees the monotonicity, and it is based on statistical considerations. For the sake of simplicity of the paper, more details regarding SAGE formulation are stated in Section 0 of the Appendix II.

2. *The proposed Quasi-SAGE Algorithm*

We address the problem of localizing multiple transmitters based on power measurements at multiple receivers under lognormal shadowing. As stated previously,

accurate localization of legacy-system transmitters increases the degree of identification and exploitation of unused spectrum by cognitive radio nodes without triggering interference.

Basically, the quasi-SAGE algorithm is applied on a wise choice of the index set, on which the expectation and maximization steps will be applied. Similar to the quasi-EM algorithm, the proposed quasi-SAGE technique alternates between the two steps:

- *The E-step*: It concerns the estimation of each transmitter's locations independently via the designated percentage of the received power at each receiver

- *The M-step*: It concerns the allocation of a percentage of the received power at each receiver to each transmitter proportional to the expected received power given the estimates of the last transmitter location.

Nevertheless, quasi-SAGE differs from quasi-EM in the selection of the set over which we apply the previous two steps that are analogous to the expectation and maximization steps in the classical EM algorithm.

Unlike the quasi-EM technique, the proposed quasi-SAGE algorithm does an efficient selection of some receivers (not all as in quasi-EM) that are mostly useful for localizing the existing transmitters. In our approach, once an initial estimate of the transmitters is randomly generated, a set of receivers that are closest in terms of the Euclidean distance to the estimated transmitters' positions is selected. The number of these receivers is to be determined empirically.

The quasi-SAGE solution can be summarized by the following steps:

1. Generate randomly the initial estimate $\hat{\theta}$ of the locations for the *M* transmitters.
2. Select a set of N' receivers, from the initial set of *N* receivers, that is closest in terms of the Euclidean distance to the transmitters. The ratio N'/N is a parameter tested through simulations.
3. Compute the expected power in dB from the i^{th} transmitter at the j^{th} receiver for j =1 to $N'$and i=1 to M:

$$\begin{aligned} e_{ij} &= E\left[10 \times \log_{10}\left(h_{ij}\right)\right] = E\left[10\log_{10}\left(P_{ij}10^{\frac{x_{ij}}{10}}\right)\right] \\ &= 10\log_{10}(\rho P_0) + 10n\log_{10}(\frac{d_0}{d_j(\hat{\theta}_i)}) \end{aligned} \tag{35}$$

4. Normalize the expected values e_{ij} so that the expected total power at each receiver is equal to the observed power at that receiver:

$$\tilde{e}_{ij} = 10\log_{10}\left(\frac{r_j 10^{\frac{e_{ij}}{10}}}{\sum_i 10^{\frac{e_{ij}}{10}}}\right) \tag{36}$$

5. Re-estimate the locations of transmitter via the minimization of the sum of squared dB error by utilizing the expected values $\tilde{e}_{ij}$.

$$\hat{\theta}_i = \arg\min_{\tilde{\theta}_i} \sum_{j=1}^{N'} (\tilde{e}_{ij} - 10\log_{10}\left(\frac{\rho P_0 d_0^n}{d_j(\tilde{\theta}_i)^n}\right))^2 \tag{37}$$

6. Return to Step 2 after a given number of iterations

Due to the random initial conditions, the transmitter locations are finally estimated by searching over the different iterations for the lowest sum-of-squared log-power errors given by the following global minimum cost function:

$$C(\hat{\theta}) = \sum_{j=1}^{N'} (\log_{10} r_j - \log_{10} \sum_{i=1}^{M} \left(\frac{\rho P_0 d_0^n}{d_j (\hat{\theta}_i)^n} \right))^2 \quad (38)$$

It can be easily proven that this cost function increases the likelihood criteria defined in (34).

E. Simulation Results

In this section, we provide simulation results on the proposed quasi-SAGE algorithm and we compare its performance to quasi EM technique and to random guessing, proposed in literature. The area of interest is considered one-kilometer square. We firstly assume that the distance of separation between transmitters to be equal to 200 meters since in reality primary transmitters utilizing same frequency band can't be too close in terms of distance to avoid interference. In addition, we assume that the minimum distance between two receivers is equal at least twice the reference distance, i.e. d_0=1m.

To run the quasi-SAGE and quasi-EM technique, M^2 different uniformly random initial estimates $\hat{\theta}$ of the transmitters' positions are generated. For each initial estimate, and for the sake of fairness, both techniques are executed 10 times (10 iterations). This also applies on the random guesses algorithm.

The simulation results have been normalized to represent a square of unit area. Figures 1 through 6 represent the mean square error over 5000 different random draws for M=2 and 3 transmitters, N=2M to 40 receivers, and shadowing variance $\sigma^2 = 4$ dB and $\sigma^2 = 16$ dB.

Figure 30 and Figure 31 represent the mean normalized squared distance error respectively for 2 and 3 transmitters with shadowing variance equals to 4 dB. It is shown that the Quasi-EM and Quasi-SAGE techniques perform better performance than the random guess technique. Additionally, Quasi-SAGE algorithm shows an additional enhancement over the quasi-EM technique.

Figure 32 and Figure 33 designate the mean normalized squared distance error respectively for 2 and 3 transmitters with shadowing variance equals to 16 dB. Similarly, the quasi-EM and quasi-SAGE algorithms offer higher accuracy in terms of transmitters' position estimation. The quasi-SAGE technique provides an additional improvement over the quasi-EM technique. It is very clear that the improvement provided by the quasi-SAGE over the quasi-EM technique is more significant with shadowing variance equals to 16 dB when compared to the results with shadowing variance equals to 4 dB.

As a result, the quasi-SAGE technique generates the smallest MSE for M=2 and 3 transmitters and $\sigma^2 = 4$ dB and $\sigma^2 = 16$ dB. As we increase the number of receivers from N=4 to N=40, the performance improvement becomes larger for the proposed quasi-SAGE over the quasi-EM, and random guess localization approaches. As expected, the MSE decreases for all the three localization methods with the increase in the number of receivers;

yet, the performance error ultimately flattens. Basically, the performance enhancements are not significant as the number of receivers increases due to the fact that it less likely to obtain independent information for the power measurement offered through each additional receiver.

Figure 34 explores the effect of the number of the receivers used by the quasi-SAGE approach for localizing multiple transmitters. We define the parameter α as the ratio between the number of receivers used by the quasi-SAGE approach with respect to the quasi-EM technique. As α increases, the number of receivers used by quasi-SAGE algorithm becomes closer to the number of receivers used by quasi-EM algorithm. It is shown that for $\alpha = 1/6$ the improvement made by quasi-SAGE technique is more significant than that obtained with $\alpha = 2$.

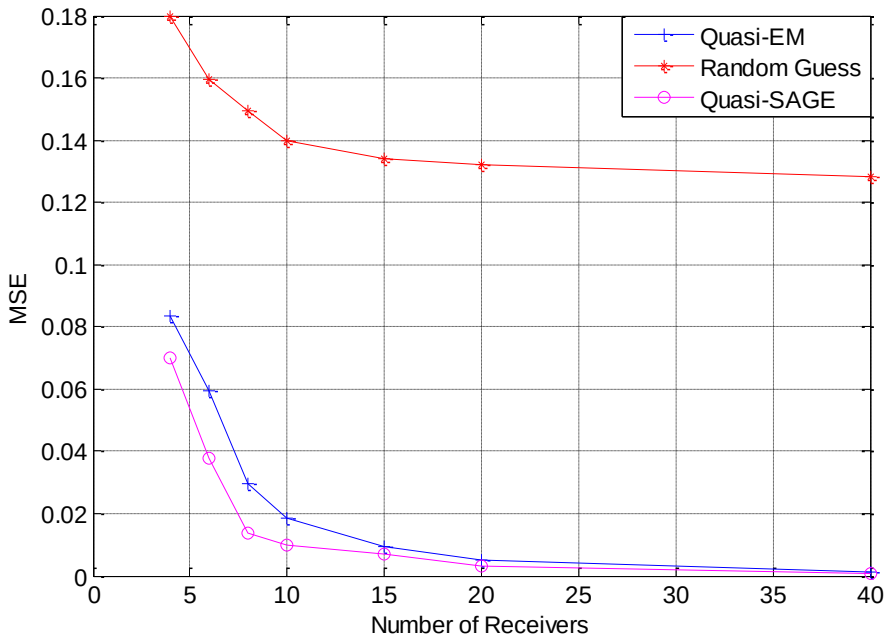

Figure 30 MSE-Normalized: M=2 and sigma=4dB

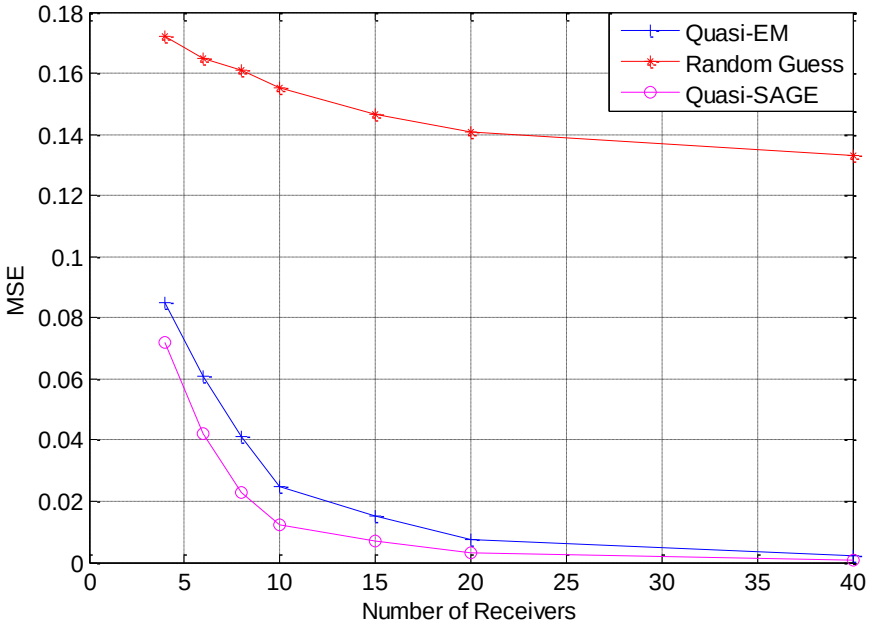

Figure 31 MSE-Normalized: M=3 and sigma=4dB

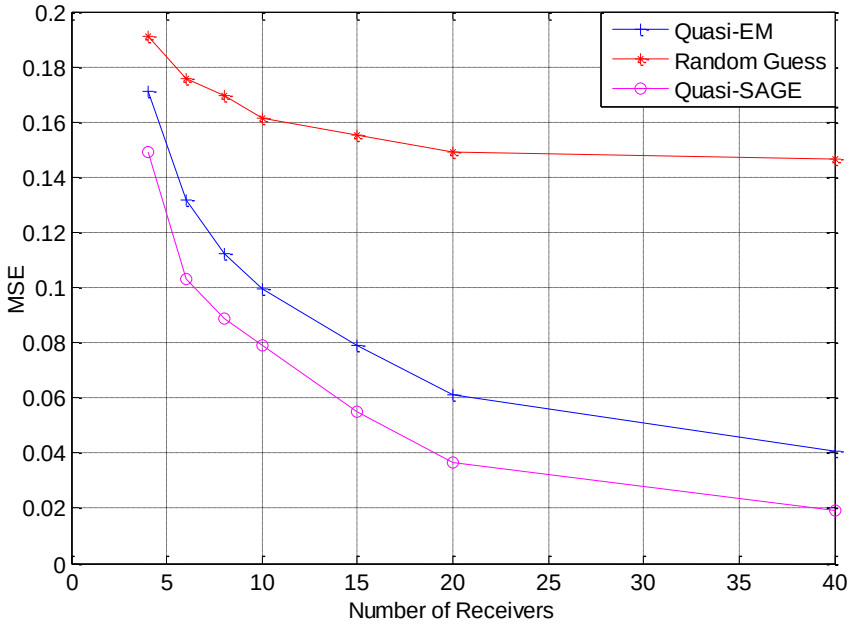

Figure 32 MSE-Normalized: M=2 and sigma=16dB

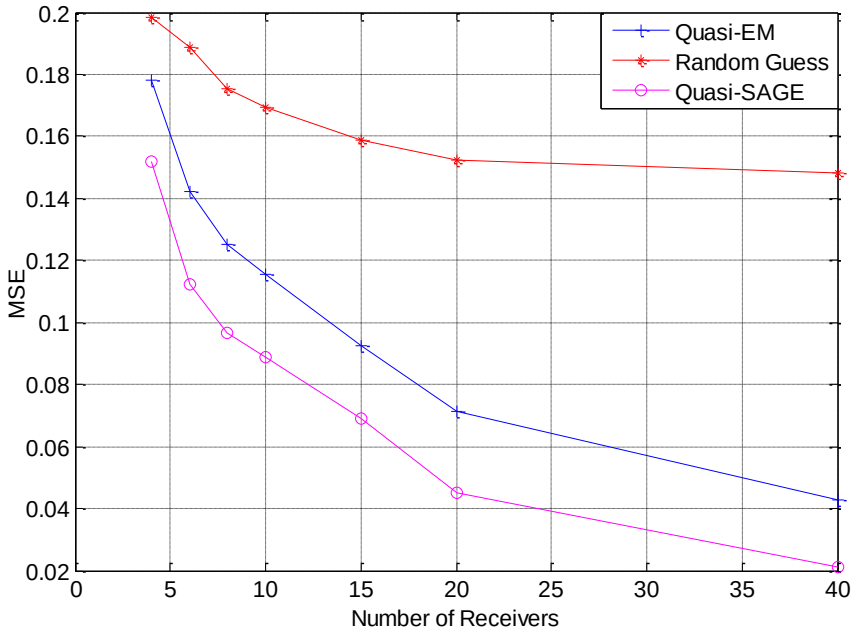

Figure 33 MSE-Normalized: M=3 and sigma=16dB

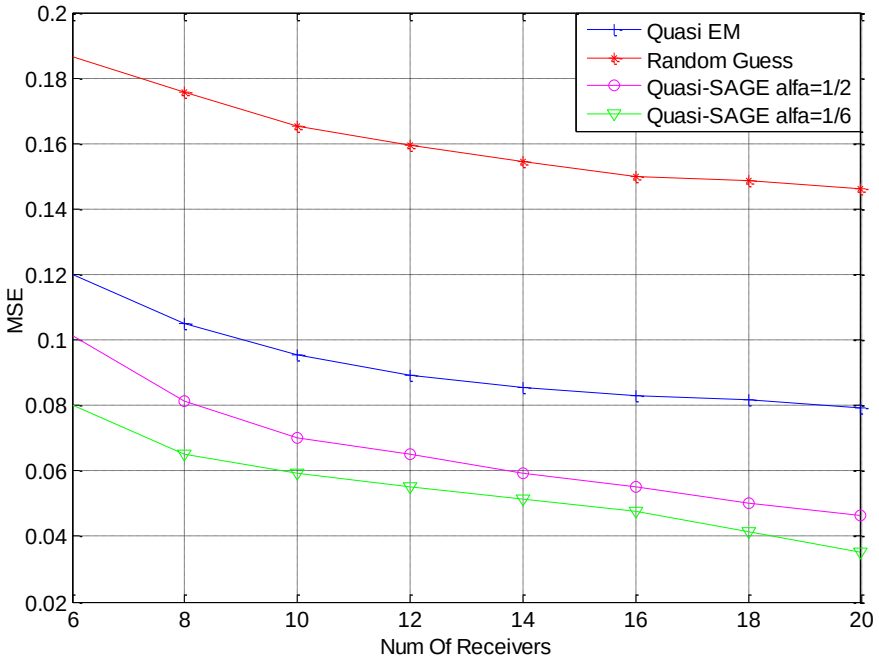

Figure 34 MSE-Normalized with M=2, $\boldsymbol{\sigma^2}$=20dB, two values of α

Furthermore, the effect of the NLOS factor added over the proposed quasi-SAGE algorithm is shown in Figure 35. It is shown how the error of transmitters' localization increases with the addition of the NLOS effect. Finally, the true position of 2 transmitters and the corresponding estimated positions obtained via the quasi-SAGE and quasi-EM, and random guessing techniques are shown within the square region of interest in Figure 36. Again, this figure shows the outperformance of the quasi-SAGE algorithm versus the other techniques.

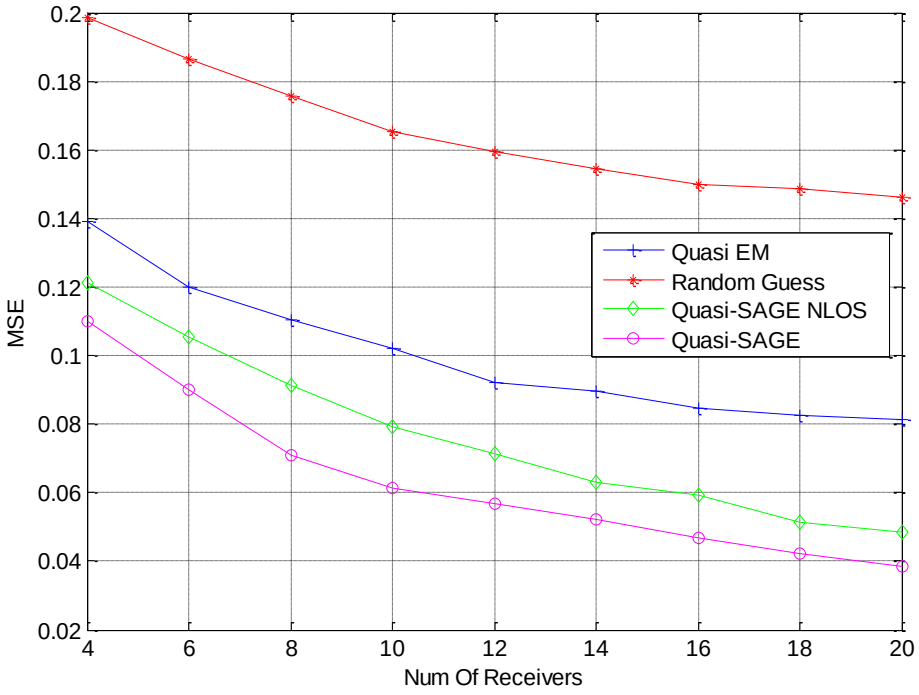

Figure 35 MSE-Normalized with M=2 and sigma=4dB in addition to the NLOS factor on quasi-SAGE

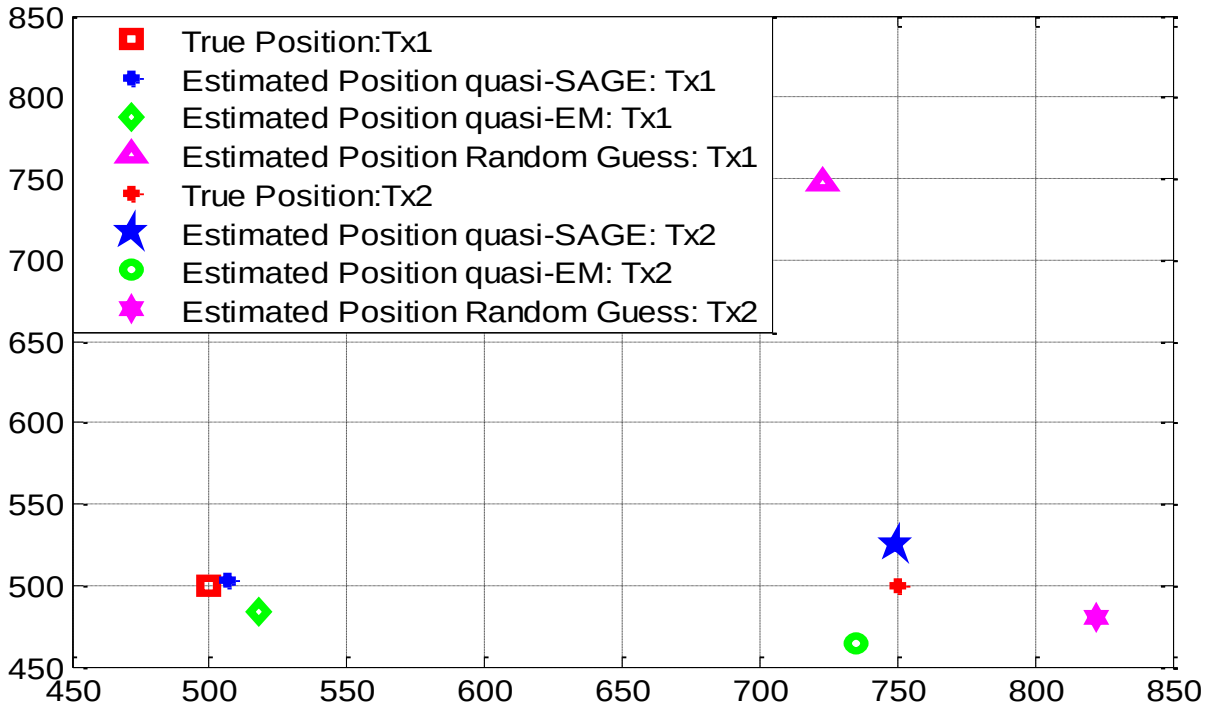

Figure 36 True and Estimated Positions of 2 Transmitters

F. Conclusion

In this paper, we have considered the localization of multiple transmitters based on power measurements for a cognitive radio environment. To solve this problem, quasi-SAGE algorithm has been proposed due to its advantages over the solutions existing in literature. It has been shown that the proposed quasi-SAGE technique achieves major error reduction when compared to quasi-EM and random guessing techniques. This improvement is much significant with higher shadowing variance. In the future, some real measurements will be considered for the proof-of-concept.

CHAPTER V

CONCLUSION

The existence of the location information at the base stations turns out to be a vital factor in today and future communications system for permitting innovative location based services. Practically, Time of Arrival (ToA), Time Difference of Arrival (TDOA), Received Signal Strength (RSS) and Angle of Arrival (AoA) are the main techniques used for positioning. Localization of a Mobile Terminal (MT) is estimated with high accuracy in outdoor scenarios due to the Global Positioning System (GPS) or to the standalone cellular systems. Nevertheless, this is not the case with indoor scenarios where satellite or cellular signals are broken and with scenarios where deep shadowing effect exists. Heterogeneous networks, where a combination of radio access technologies exists (such as cellular systems, WiFi, WiMAX), presents good candidate for critical scenarios. Definitely, significant interests are present for cellular and WiFi networks based localization techniques in both communication and localization community.

In this thesis, we have considered a particular positioning problem with lack of hearability or with a limited number of anchor nodes. We have presented a hybrid localization approach based on the combination of the ToA, AoA and RSS based fingerprinting techniques. Simulations have shown that the proposed hybrid approach outperforms the stand-alone ToA and RSS fingerprinting techniques in this critical transmission scenario. Based on a 2-level UKF, we have shown that significant positioning estimations could be obtained. Moreover, hierarchical and genetic based clustering

approaches have been proposed and combined with the UKF to improve the accuracy of the estimation algorithm. The simulation results have shown the outperformance of our proposed algorithm with respect to the standalone techniques proposed in the literature.

Moreover, we proposed the use of UKF based PDF filter in this data fusion results in further enhancement in the accuracy of the position estimation. The RBF is tuned using minimum entropy criterion to achieve an optimal estimation. Additionally, the proposed adjustment of the PDF filter parameters is done every time step instead of using a batch mode, where parameters are fixed during each batch and are updated between two adjacent batches. Thus, using batch mode makes the estimation error to be reduced batch by batch causing slow convergence; however, the weight vector of the RBF network in the proposed technique is updated at each time step enabling a higher convergence rate. Regarding the convergence analysis, the proposed UKF based on PDF filter is shown to be asymptotically convergent as long as μ is small enough. Thus, an appropriate step-size μ has to be chosen to trade-off between convergence speed and misadjustment.

Finally, not only localizing MTs, but also localizing multiple transmitters in a region is one of the main problems that are assessed by researches nowadays; hence, it is also assessed in our thesis. Even though this problem is applicable in different applications, the most prominent one of common interest is cognitive radio systems. Localizing active primary users is required by the cognitive radio system in order to avoid interference and obtain a successful transmission for secondary users. Hence, we considered localization multiple transmitters based on power measurements at multiple receivers under lognormal shadowing. Different solutions were proposed in the literature to solve the problem of

localizing multiple transmitters. Quasi-EM technique is one of the proposed methods used to localize multiple transmitters based on the received power measurements under lognormal shadowing model. The quasi-EM technique was proposed due to the unfeasibility of presenting an analytic distribution for the sum of lognormal random variables representing the observed power. We have proposed the quasi-SAGE algorithm through wise selection of the index set and application of the analogous expectation and maximization steps corresponding to the received power measurements under the lognormal shadowing model over this set. It is shown that the proposed quasi-SAGE technique achieves major error reduction when compared to random guessing technique. Also, the proposed technique shows an improvement over the quasi-EM technique. This improvement is much significant with higher shadowing variance.

APPENDIX I

UKF BASED PDF PRELIMINARIES

A. RBF Network Specifications

Let's consider the RBF network $y = W^T\Phi(x): \Re^n \mapsto \Re^m$ to be defined as follows:

$$W = \left[w_1, w_2, \dots, w_p\right]^T \tag{39}$$

$$\Phi(x) = \left[\phi_1, \phi_2, \dots, \phi_p\right]^T \tag{40}$$

$$\phi_i(x) = \exp\left(-\frac{\|x - c_i\|^2}{\sigma_{rbf}^2}\right) - \exp\left(-\frac{\|x + c_i\|^2}{\sigma_{rbf}^2}\right), i = 1,2, \dots, p \tag{41}$$

If c_i and w_i are symmetrically distributed, that is:

$$c_i = -c_{2p-i} + 1 \tag{42}$$

$$w_i = -w_{2p-i} + 1 \tag{43}$$

Then, the RBF network will have a zero-mean output for any stochastic inputs with symmetric PDF distribution, as:

$$E[y] = E[W^T\Phi(x)] = E\left[\sum_{i=1}^{2p} w_i\phi_i(x)\right] \tag{44}$$

Substituting (40) and (41) in (44), we will obtain the following mean:

$$E[y] = \sum_{i=1}^{2p} w_i E\left[\phi_i(x) - \phi_{2p-i+1}(x)\right] \tag{45}$$

Assuming that $p(x)$ is the PDF of x, the mean of the PDF term will be equal to zero: (knowing that $\phi_{2p-i+1}(x)$ and $p(x)$ are odd and even functions of *x* respectively)

$$\begin{gathered} E\left[\phi_i(x) - \phi_{2p-i+1}(x)\right] = \int p(x)\left[\phi_i(x) - \phi_{2p-i+1}(x)\right] = 0 \\ E[y] = \sum_{i=1}^{2p} w_i E\left[\phi_i(x) - \phi_{2p-i+1}(x)\right] = 0 \end{gathered} \tag{46}$$

As a result, the term v_k^{pdf} achieves a zero mean if we have symmetrically distributed condition for the RBF-network since the input to the network is sigma-point innovation with a symmetric PDF.

B. Inclusion of Renyi's Entropy

According to the information theory, Shannon presented in 1948 the entropy as a measure of the uncertainty relative to random variables. Renyi's entropy, named after Alfred Renyi [32], quantifies the randomness of a random variable. It is described as:

$$H_\alpha = \frac{1}{1-\alpha} log \int p^\alpha(x)dx, \;\; \alpha > 0 \; and \; \alpha \neq 1 \tag{47}$$

The most commonly used type of Renyi's entropy of a variable *X* is the quadratic one where $\alpha = 2$, given by:

$$H_2 = -log\left[\int_X p^2(x)dx\right] = -log\big(V(X)\big) \tag{48}$$

C. Renyi's Quadratic Entropy Calculation using KDE

The probability density of a variable ***X*** is estimated using KDE as follows:

$$p(x) = \frac{1}{N}\sum_{i=1}^{N} G_{\Sigma}(x - x_i) \tag{49}$$

where $G_{\Sigma}(.)$ is the Gaussian function defined by the following equation:

$$G_{\Sigma}(x - x_N) = (2\pi)^{-\frac{n}{2}}(\det\Sigma)^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(x - x_i)^T\Sigma^{-1}(x - x_i)\right\} \tag{50}$$

Consequently, the Renyi's quadratic entropy is computed using the KDE as follows,

$$\begin{aligned} H_2 = -log\big(V(X)\big) &= -\log\int\left(\frac{1}{N}\sum_{i=1}^{N} G_{\Sigma}(x - x_i)\right)^2 \\ &= -\log\frac{1}{N^2}\int\left(\sum_{i=1}^{N}\sum_{j=1}^{N} G_{\Sigma}(x - x_i)G_{\Sigma}\big(x - x_j\big)\right)dx \\ = -\log\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\int G_{\Sigma}(x - x_i)G_{\Sigma}\big(x - x_j\big)\,dx &= -\log\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{\sqrt{2}\Sigma}\big(x_i - x_j\big) \end{aligned} \tag{51}$$

Basically, the fourth equality is based on the fact that the convolution of two Gaussian functions is also Gaussian.

D. Probability Density Estimation

The probability $p(x)$ represents the PDF of the random vector **x** in a Lebesque measure. x_k describes the true state in time step k; let $\hat{x}_k(\hat{x}_{k|k})$ and $\hat{x}_{k|k-1}$ describes the filtered and the predicted state of $\hat{x}_k$.

A kernel has to be non-negative real valued integrable function sustaining the following conditions:

- $\int_{-\infty}^{+\infty} G_\sigma(x)dx = 1$
- $G_\sigma(-x) = G_\sigma(x), \forall x$

The first condition guarantees that the output of the KDE is a probability density function and the second condition ensures stability of the expectation of the process. Despite the existence of various kernel functions, the most common one is the Gaussian kernel that is expressed in the following equation:

$$G_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{x^2}{2\sigma^2}) \tag{52}$$

Concerning the KDE, the mean and the variance of the estimated PDF are expressed by:

$$Bias[\hat{p}(x)] = E[\hat{p}(x)] - p(x) \approx \sigma^2/2p''(x)C_1 \tag{53}$$

$$Var[\hat{p}(x)] = E[(\hat{p}(x) - E[\hat{p}(x)])^2] \approx \frac{1}{N\sigma} C_2 p(x), N\sigma \to \infty$$

where $C_1$and C_2 are two constants related to specific kernel and p'' is the second order derivative of the PDF. Based on the above equations, the best choice of the bandwidth (σ) compromises between low bias and low variance since it affects them in opposite manners.

For a Gaussian kernel, we can select the bandwidth as follows:

$\sigma = 0.2\hat{\sigma}$ or $\sigma = \hat{\sigma}(4N^{-1}(2d+1)^{-1})^{\frac{1}{d+4}}$ where $\hat{\sigma}$ is the data standard deviation, *N* is the number of samples, and *d* is the dimensionality of the data [24].

APPENDIX II

LOCALIZATION OF COGNITIVE TRANSMITTERS

A. Cooperative Localization Model in Cognitive Network

First, the parameters estimated at the CR node are defined as follows:

$$\Lambda = [R_{ij}(t), \theta_{ij}(t)]^T \tag{54}$$

where $R_{ij}(t)$ is the range obtained from the ToA estimates and $\theta_{ij}(t)$ is the angle of arrival that represents the directional information of the target PU-j at the *i*th CR node, given by:

$$R_{ij} = \begin{bmatrix} R_{11} & R_{12} & .. & R_{1L} \\ R_{21} & R_{22} & .. & R_{2L} \\ . & . & .. & . \\ R_{K1} & R_{K2} & .. & R_{KL} \end{bmatrix} \tag{55}$$

$$\theta_{ij} = \begin{bmatrix} \theta_{11} & \theta_{12} & .. & \theta_{1L} \\ \theta_{21} & \theta_{22} & .. & \theta_{2L} \\ . & . & .. & . \\ \theta_{K1} & \theta_{K2} & .. & \theta_{KL} \end{bmatrix} \tag{56}$$

The use of maximum likelihood (ML) estimation technique at the CRB along with the *a*-priori knowledge of the CR node locations, we can obtain the following estimates for the targets PU-j as shown in the following equations [48].

$$x_j(t) = d_j(\Lambda) + w_d^j(t) \tag{57}$$

$$\Phi_{\mathrm{j}}(\mathrm{t}) = \phi_{\mathrm{j}}(\Lambda) + w_\phi^j(t) \tag{58}$$

Where $d_j(\Lambda)$ and $\phi_j(\Lambda)$ are the true distances and the phase angles of the PU-j's w.r.t the CRB and $w_d^j(t)$ and $w_\phi^j(t)$ are the additive noise components relative to the ML approximations. It is assumed that the signal has no fading component. ρ_j and ζ_j are the range and phase respectively describing the signal-to-noise ratio of the received signals prior to ML estimation.

$w_\phi^j(t)$ is modeled as a zero mean random process and $w_d^j(t) \alpha N(0, \sigma_{d_i}^2)$.

Additionally, the mobility of PUs is modeled by 2-first order Markov processes for the phase angle and distance described by the density functions $p(\theta_j^t | \theta_j^{t-1})$ and $p\left(d_j^t \middle| d_j^{t-1}\right)$ knowing that θ_j^t and d_j^t are the phase and the distance at time t. Then, tracking the target PU-j is performed in the network by making use of the positional estimates x_j and Φ_j.

B. Particle Filtering: A Bayesian Tracking Algorithm

Particle filtering track the desired parameters d_j and w_j based on the model defined in [43]. The respective probability distributions are estimated by assuming the discrete model of the distributions, and they are given by the following:

$$p(\mu|z) \approx \sum_{n=1}^{M} W_n \delta(\mu - u_n) \tag{59}$$

where $u_n = \left[u_{d_i}^n, u_{\phi_i}^n\right]^T$ are known as the particles that are drawn from each posteriori distribution $p(\mu|z)$, $W_n = \left[w_{d_i}^n, w_{\phi_i}^n\right]^T$ are the weights assigned to the particles, and M is the total number of particles. The initial value of the weights is basically $w_{d_i}^n = 1/M$, and $w_{\phi_i}^n = 1/M$ at k=0; then, they are iteratively calculated using (25). Then, the weights are updated using the analytical solutions to the *posteriori* distributions of the desired parameters. So, updating the weights is done according to the following equations.

$$w_{d_j}^{-n}(k) \propto w_{d_j}^{-n}(k-1) F_{d_j}(x_j(k)|d_j(k-1)) \tag{60}$$

$$w_{d_j}^{n}(k) = w_{d_j}^{-n}(k) \Big/ \sum_{n=1}^{M} w_{d_j}^{-n}(k) \tag{61}$$

$$w_{\emptyset_j}^{-n}(k) \propto w_{\emptyset_j}^{-n}(k-1) F_{\emptyset_j}(\Phi_j(k)|\phi_j(k-1)) \tag{62}$$

$$f_{w_\phi^j}\left(w_\phi^j\right) = \begin{Bmatrix} \exp(-\zeta j/2)\left[\frac{1}{2\pi} - \exp\left(\frac{T}{2}\right)\left(\frac{T}{2\pi}\right)^{\frac{1}{2}} Q\left(\sqrt{T}\right)\right] & for\ \pi/2 < w_\phi^j \le \pi \\ \exp(-\zeta j/2)\left[\frac{1}{2\pi} + \exp\left(\frac{T}{2}\right)\left(\frac{T}{2\pi}\right)^{\frac{1}{2}} \left(1 - Q\left(\sqrt{T}\right)\right)\right] & for\ -\pi/2 < w_\phi^j \le \pi/2 \\ \exp(-\zeta j/2)\left[\frac{1}{2\pi} - \exp\left(\frac{T}{2}\right)\left(\frac{T}{2\pi}\right)^{\frac{1}{2}} Q\left(\sqrt{T}\right)\right] & for\ -\pi < w_\phi^j \le -\pi/2 \end{Bmatrix} \tag{63}$$

Knowing that:

$$T = \zeta_j cos^2\left(w_\phi^j\right) \; and \; Q(z) = \frac{1}{2\pi}\int_z^\infty \exp(-u^2/2)\, du \tag{64}$$

$$w_{\emptyset_j}^n(k) = w_{\emptyset_j}^{-n}(k) \Big/ \sum_{n=1}^{M} w_{\emptyset_j}^{-n}(k) \tag{65}$$

$F_{d_j}(x_j(k)|d_j(k-1))$ and $F_{\emptyset_j}(\Phi_j(k)|\phi_j(k-1))$ are known as importance functions. Thus, based on the *posteriori* distributions, the two importance functions are expressed as follows:

$$F_{d_j}\left(x_j(k)\middle|d_j(k-1)\right) = \exp(-\left(u_{d_j}^n - x_j(k)\right)^2) \tag{66}$$

$$F_{\emptyset_j}\left(\Phi_j(k)\middle|\phi_j(k-1)\right) = \gamma \exp(\gamma)\,(1 - \mathrm{Q}(\gamma)) \tag{67}$$

where $\gamma = \cos\left(u_{\phi_j}^n - \Phi_j(k)\right)$. Then, we make use of the newly calculated particle weights to estimate the parameter via calculating the expectation of the discrete sample set as shown below:

$$\hat{d}_j(k) = \sum_{n=1}^{M} u_{d_j}^n w_{d_j}^n \;, \hat{\phi}_j(k) = \sum_{n=1}^{M} u_{\phi_j}^n w_{\phi_j}^n \tag{68}$$

After performing little estimation, the weights of some particles will be almost zero; those particles are no more needed in the process of estimation; importance sampling is used in such cases to diminish the complexity of computation in the estimation [49][50].

C. Proposed Quasi-SAGE based positioning algorithm Preliminaries

1. *SAGE Formulation*

Let the observation Y have probability density function $f(y, \theta_{true})$ where θ_{true} is a parameter representing the true location of the M transmitter based on the power measurement at each sensor represented in subset Θ of the *p*-dimensional space $\mathbb{R}^p$. Given a measurement realization Y=*y*, the maximum penalized-likelihood estimate $\widehat{\theta}$ of θ_{true} is computed as follows:

$$\begin{aligned} \widehat{\theta} &\triangleq \arg\max_{\theta \in \Theta} \Phi(\theta) \text{ where} \\ \Phi(\theta) &\triangleq \log f(y;\theta) - P(\theta) \end{aligned} \tag{69}$$

However, the complexity of f or the coupling in P makes the direct maximization intractable. Hence, we have to make use of iterative techniques by considering subset of the elements of the parameter vector θ . The following definition formalizes the idea.

Definition 1: *A set S is defined to be an index set if it i) is nonempty, ii) is a subset of the set {1,…,p}, and ii) has no repeated entries. The set* $\tilde{S}$ *denotes the complement of S intersected with {1,…,p}.*

Let *m* denotes the cardinality of S; thus, θ_S represents the *m* dimensional vector consisting of *m* elements of θ indexed by the members of S. Similarly, $\theta_{\tilde{S}}$ represents the *p-m* dimensional vector consisting of the remaining elements of θ. For instance, if p=5 and S={1,3,4}, then $\tilde{S} = \{2,5\}$; thus, $\theta_S = [\theta_1 \theta_3 \theta_4]'$ and $\theta_{\tilde{S}} = [\theta_2 \theta_5]'$. Also, we define $\Phi(\theta_S, \theta_{\tilde{S}}) = \Phi(\theta)$.

The "grouped coordinate-ascent" technique is based on sequencing through different index sets $S = S^i$ and updates only the elements θ_S of θ while holding the other parameters $\theta_{\bar{S}}$ fixed [51]. θ_S^{i+1} is assigned to the argument that maximizes $\Phi\left(\theta_S, \theta_{\bar{S}}^i\right)$ over θ_S ; however, there is no analytical form for the maximum of $\Phi(\theta_S, \theta_{\bar{S}})$ over θ_S in some applications such as the imaging application even if the index set S consists of a single element. The evaluation of $\Phi\left(\theta_S, \theta_{\bar{S}}^i\right) - \Phi\left(\theta^i\right)$ for many values of θ_S is complex in terms of computations.

Therefore, the SAGE technique, originated from the EM method, utilizes the underlying statistical structure of the problem to use simple maximizations instead of the expensive numerical maximizations. Basically, SAGE uses the maximization of functional $\phi^S(\theta_S; \theta^i)$ instead of the maximization of $\Phi\left(\theta_S, \theta_{\bar{S}}^i\right)$ over θ_S by presenting the "hidden-data" space for θ_S. Maximizing $\phi^S(.; \theta^i)$ can be done analytically when wise choice of hidden-data space is done; hence, avoiding the use of line searches. Eventually, line searches for maximizing $\phi^S(.; \theta^i)$ is less expensive than line searches for maximizing $\Phi\left(., \theta_{\bar{S}}^i\right)$. Additionally, ϕ^S are initiated to guarantee that the increase in ϕ^S leads to an increase in Φ.

2. *Hidden-Data Space*

Generating the functions ϕ^S for each index set S of interest requires the identification of an admissible hidden-data space defined in the following sense:

Definition 2: *A random vector X^S with probability density function $f(x;\theta)$ is an admissible hidden-data space with respect to θ_S for $f(y;\theta)$ if the joint density of X^S and Y satisfies*

$$f(y,x;\theta) = f(y|x;\theta_{\tilde{S}})f(x;\theta) \tag{70}$$

i.e., the conditional distribution $f(y|x;\theta_{\tilde{S}})$ must be independent of θ_S . Hence, X^S must be a complete-data space for θ_S given $\theta_{\tilde{S}}$ is known.

The following explains the association between this definition and the related methods.

- The complete-data space for the classical EM algorithm [45] is taken as a special case of Definition 2 by selecting S={1,…,p} and Y as a deterministic function of X^S.
- When decomposing (34), we can represent Y as the output of a noisy channel that may depend on $\theta_{\tilde{S}}$ but not on θ_S
- We describe X^S using the term "hidden" is used instead of "complete" because X^S will not be complete for θ in the original sense of Dempster *et al.*[45]. Also, the aggregate of X^S will not be an admissible complete-data space for θ over all S.
- The conditional distribution of Y on X^S is permitted to depend on all of the parameters $\theta_{\tilde{S}}$.
- The cascade EM algorithm [52] presents an alternative generalization based on a hierarchy of nested complete-data spaces. Actually, SAGE could be generalized by permitting hierarchies for eachX^S.

3. *SAGE Algorithm*

SAGE algorithm is based on the conditional expectation of the log-likelihood of X^S:

$$Q^S(\theta_S;\bar{\theta}) = Q^S(\theta_S;\bar{\theta}_S,\bar{\theta}_{\tilde{S}}) \triangleq E\{\log f(X^S;\theta_S,\bar{\theta}_{\tilde{S}})\big|Y=y;\bar{\theta}\} \tag{71}$$

$$= \int f(x|Y)=y;\bar{\theta})\log f(x;\theta_S,\bar{\theta}_{\tilde{S}})dx$$

This expectation is combined with the penalty function to obtain:

$$\phi^S(\theta_S;\bar{\theta}) \triangleq Q^S(\theta_S;\bar{\theta}) - P(\theta_S;\bar{\theta}_{\tilde{S}}) \tag{72}$$

Let $\theta^0 \in \Theta$ be an initial parameter estimate. A generic SAGE algorithm generates a sequence of estimates $\{\theta^i\}_{i=0}^{\infty}$ through the following recursion:

For *i=0,1,...{*

1. Choose an index set $S = S^i$
2. Choose an admissible hidden-data space X^{S^i} for θ_{S^i}
3. E-step: compute $\phi^{S^i}(\theta_{S^i};\theta^i)$ using (9).
4. M-step:

$$\theta_{S^i}^{i+1} = \arg\max_{\theta_{S^i}} \phi^{S^i}(\theta_{S^i};\theta^i) \tag{73}$$

$$\theta_{\tilde{S}^i}^{i+1} = \theta_{S^i} \tag{74}$$

5. Optional: Repeat steps 3 and 4.

where the maximization in (73) is over the set:

$$\theta^{S}(\theta^{i}) = \{\theta_{S^i} : (\theta_{S^i}, \theta^{i}_{\tilde{S}^i}) \in \Theta\} \tag{75}$$

If an appropriate choice of the index sets and hidden data spaces is made properly, we can combine the E-step and the M-step through an analytical maximization into a recursion of the form $\theta_{S^i}^{i+1} = g^{S^i}(\theta^i)$.

4. *Monotonicty*

Let S represent the index set and X^S represents the hidden data space to be used in SAGE algorithm. Applying Bayes' Theorem under regularity conditions [45] we obtain the following:

$$\begin{aligned} Q^S(\theta_S; \bar{\theta}) &= \int f(x|Y = y; \bar{\theta}) \log f(x; \theta_S, \bar{\theta}_{\tilde{S}}) dx \\ &= L(\theta_S, \bar{\theta}_{\tilde{S}}) + H^S(\theta_S; \bar{\theta}) - W^S(\bar{\theta}) \end{aligned} \tag{76}$$

$$where\ L(\theta_S, \bar{\theta}_{\tilde{S}}) \triangleq \log f(y; \theta_S, \bar{\theta}_{\tilde{S}}$$

$$H^S(\theta_S; \bar{\theta}) \triangleq E\{f(X^S|Y = y; \theta_S, \bar{\theta}_{\tilde{S}})|Y = y; \bar{\theta}\} \tag{77}$$

$$\text{Then, } W^S(\bar{\theta}) \triangleq \int f(x|Y = y; \bar{\theta}) \log f(y|x; \bar{\theta}_{\tilde{S}}) dx$$

As shown in the (77), W^S is independent of θ^S; thus, it will not affect the maximization in (73). Additionally, we can show the following using these definitions and Jensen's inequality [45]:

$$H^S(\theta_S; \bar{\theta}) \leq H^S(\bar{\theta}_S; \bar{\theta}), \forall \theta_S, \forall \bar{\theta} \tag{78}$$

Based on that, the following theorem follows directly.

Theorem: *let θ^i represents the sequence of estimates generated by a SAGE algorithm* (73)*. Then, 1) $\Phi(\theta^i)$ is monotonically non-decreasing, 2) if $\hat{\theta}$ maximizes Φ , then $\hat{\theta}$ is a fixed point of the SAGE algorithm, and 3)*

$$\Phi(\theta^{i+1}) - \Phi(\theta^i) \geq \phi^S(\theta_S^{i+1}; \theta^i) - \phi^S(\theta_S^i; \theta^i) \tag{79}$$

On the other hand, standard numerical methods necessitate the evaluation of $\Phi(\theta^{i+1}) - \Phi(\theta^i)$ to guarantee monotonicity; however, this prerequisite is hindered for SAGE algorithm due to the monotonicity theorem above.

5. *Convergence*

The monotonicity property guarantees that the sequence $\{\theta^i\}$ will not diverge; however, there is no guarantee that it will converge to a local maximum of Φ [53]. Basically, we can summarize the following:

- Initializing a SAGE algorithm in a region suitably close to a local maximum in the interior of Θ makes the sequence of estimates to converge monotonically in *norm* to it.
- The region of monotone convergence in norm for strictly concave objectives is ensured to be nonempty.
- Choosing a less informative hidden-data space enhances the asymptotic convergence rate of a SAGE method. Basically, faster convergence is achieved with

less informative hidden-data spaces; however, simpler M-steps are obtained with more informative hidden-data spaces [47][54].

BIBILIOGRAPHY

[1] A. K¨upper. Location-based services. Wiley, 2005

[2] A.H. Sayed, A. Tarighat, and N. Khajehnouri. "Network-based wireless location: challenges faced in developing techniques for accurate wireless location information". IEEE Signal Processing Magazine, 22:24–40, July 2005

[3] J. Caffery and G. L. Stuber, "Subscriber location in CDMA cellular networks", *IEEE Trans. Veh. Technol*, vol. 47, no. 2, pp. 406–416, May. 1998.

[4] Y. Zhao, "Standardization of mobile phone positioning for 3G systems", *IEEE Commun. Mag.*, vol. 40, pp. 108–116, Jul. 2002.

[5] A. J. Weiss, "On the accuracy of a cellular location system based on RSS measurements", *IEEE Trans. Veh. Technol*, vol. 52, no. 6, pp. 1508–1518, Nov. 2003.

[6] L. Cong and W. Zhuang, "Hybrid TDOA/AOA mobile user location for wideband CDMA cellular systems", *IEEE Trans. Wireless Commun.*, vol. 1, no. 3, pp. 439–447, Jul. 2002.

[7] F. L. Piccolo, "A new cooperative localization method for UMTS cellular networks", in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2008.

[8] C. Mensing, S. Sand, A. Dammann, and W. Utschick, "Interferenceaware location estimation in cellular OFDM communications systems", in *Proc. IEEE Int. Conf. Commun.*, Jun. 2009.

[9] C. Mensing, S. Sand, A. Dammann, and W. Utschick, "Data-Aided Location Estimation in Cellular OFDM Communications Systems", in *Proc. IEEE Global Telecommun. Conf.*, December 2009.

[10] 3GPP TS 36.305, "Stage 2 functional specifications of UE positioning in E-UTRAN", *http://www.3gpp.org*, May. 2009.

[11] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location", *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 12–23, Jul. 2005.

[12] Ziming He; Yi Ma; Tafazolli, R.; , "A hybrid data fusion based cooperative localization approach for cellular networks," *Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International*, pp.162-166, 4-8 July 2011, doi: 10.1109/IWCMC.2011.5982409

[13] Figueiras, J.; Frattasi, S.; Schwefel, H.-P.; , "Decoupling Estimators in Mobile Cooperative Positioning for Heterogeneous Networks," *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th* , vol.,no.,pp.1-5,21-24Sept.2008 doi:10.1109/VETECF.2008.433.

[14] Mayorga, Carlos Leonel Flores, et al. "Cooperative positioning techniques for mobile localization in 4G cellular networks." *Pervasive Services, IEEE International Conference on*. IEEE, 2007.

[15] Gezici, Sinan, et al. "Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks." *Signal Processing Magazine*, IEEE 22.4 (2005): 70-84.

[16] Ziming He; Yi Ma; Tafazolli, R.; , "A hybrid data fusion based cooperative localization approach for cellular networks," *Wireless Communications and Mobile Computing Conference (IWCMC),* 2011 7th International, pp.162-166, 4-8 July 2011, doi: 10.1109/IWCMC.2011.5982409

[17] Yassin, Ali; Awad, Mariette; Nasser, Youssef, "On the hybrid localization in heterogeneous networks with lack of hearability," *Telecommunications (ICT), 2013 20th International Conference on* , vol., no., pp.1,5, 6-8 May 2013

[18] C. Mayorga, F. Rosa, S. Wardana, G. Simone, M. Raynal, J. Figueiras, and S. Frattasi. Cooperative Positioning Techniques for Mobile Localization in 4G Cellular

Networks. In *Proceedings of the IEEE International Conference on Pervasive Services (ICPS 2007)*, pages 39–44, July 2007.

[19] S. Frattasi, M. Monti, and R. Prasad. Cooperative Mobile User Location for Next-Generation Wireless Cellular Networks. In *Proceedings of IEEE International Conference on Communications, 2006. ICC '06*, volume 12, pages 5760–5765, 2006.

[20] G. Sun, J. Chen, W. Guo, and K. J. R. Liu, "Signal processing techniques in network-aided positioning", *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 12–23, Jul. 2005.

[21] Q. Cui, J. Liu, X. Tao, and P. Zhang, "A novel location model for 4G mobile communicaton networks", in *Proc. IEEE Veh. Technol. Conf.*, Oct.2007.

[22] L. Cong and W. Zhuang. Non-line-of-sight error mitigation in TDOA mobile location. In *Proceedings of the IEEE Global Telecommunications Conference, 2001. GLOBECOM'01*, volume 1, pages 680–684, 2001.

[23] Ghaemi, Reza, et al. *"A review: accuracy optimization in clustering ensembles using genetic algorithms."* Artificial Intelligence Review 35.4 (2011): 287-318.

[24] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME-J. Basic Eng*., vol. series D, no. 82, pp. 35–45, 1960.

[25] M. Roth and F. Gustafsson, "An efficient implementation of the second order extended Kalman filter," *in Proc. 14th Int. Conf. Inf. Fusion*, Chicago, IL, USA, 2011, pp. 1–6.

[26] I. Arasaratnamand S. Haykin, "Cubature kalman filters," *IEEE Trans. Autom. Control*, vol. 54, no. 6, pp. 1254–1269, 2009.

[27] Z. Chen, “Bayesian filtering: From Kalman filters to particle filters, and beyond,” Adaptive Systems Lab., Univ.McMaster, Hamilton, Ontario, Canada, Tech. Rep., 2003.

[28] J. Zhou, D. Zhou, H. Wang, L. Guo, and T. Chai, “Distribution function tracking filter design using hybrid characteristic functions,” *Automatica*, vol. 46, no. 1, pp. 101–109, 2010.

[29] J. Ding, C. Tianyou, and H. Wang, “Offline modeling for product quality prediction of mineral processing using modeling error pdf shaping and entropy minimization,” IEEE Trans. Neural Netw., vol. 22, no. 3, pp. 408–419, 2011.

[30] B. Chen, Y. Zhu, J. Hu, and Z. Sun, The Institute of Control, Robotics and Systems Engineers and The Korean Institute of Electrical Engineers, “Adaptive filtering under minimum information divergence criterion,” *Int. J. Contr., Autom. Syst.*, vol. 7, no. 2, p. 157, 2009, co-published with Springer-Verlag GmbH.

[31] L. Guo and H.Wang, “Minimum entropy filtering for multivariate stochastic systems with non-Gaussian noises,” *IEEE Trans. Autom. Control*, vol. 51, no. 4, pp. 695–700, 2006.

[32] Yu Liu; Hong Wang; Chaohuan Hou, "UKF Based Nonlinear Filtering Using Minimum Entropy Criterion," Signal Processing, IEEE Transactions on , vol.61, no.20, pp.4988,4999, Oct.15, 2013 doi: 10.1109/TSP.2013.2274956

[33] A. Renyi, Probability Theory. Budapest, Hungary: Dover, 2007.

[34] J. Mitola, "*Cognitive radio: An integrated agent architecture for software defined radio,"* PhD Thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, 2005.

[35] S. Haykin, "*Cognitive radio: Brain-empowered wireless communications*,"IEEE Journal on Selected Areas in Communications, vol. 23,no. 2, pp. 201–220, February 2005.

[36] P. Kolodzy, "*Application of cognitive radio technology across the wireless stack,*" IEICE Transactions on Communications, vol. E88-B,no. 11, pp. 4158–4162, November 2005.

[37] Nelson, J.K.; Gupta, M.R., "An EM Technique for Multiple Transmitter Localization," Information Sciences and Systems, 2007. CISS '07. 41st Annual Conference on , vol., no., pp.610,615, 14-16 March 2007 doi: 10.1109/CISS.2007.4298380

[38] A. Ghasemi and E. Sousa, "*Collaborative spectrum sensing for opportunistic access in fading environments*," in IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, November 2005, pp.131–136.

[39] G. Ganesan and Y. Li, "*Cooperative spectrum sensing in cognitive radio networks,*" in IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, November 2005, pp. 137–143.

[40] D. Cabric, S. Mishra, D. Willkomm, R. Brodersen, and A. Wolisz, "*A cognitive radio approach for usage of virtual unlicensed spectrum*," in Proceedings of the 14th IST Mobile and Wireless Communications Summit, June 2005.

[41] Nelson, J.K.; Hazen, M.U.; Gupta, M.R., "Global Optimization for Multiple Transmitter Localization," *Military Communications Conference, 2006. MILCOM 2006. IEEE* , vol., no., pp.1,7, 23-25 Oct. 2006 doi: 10.1109/MILCOM.2006.302280

[42] Nelson, J.K.; Gupta, M.R.; Almodovar, J.E.; Mortensen, W.H., "A Quasi EM Method for Estimating Multiple Transmitter Locations," *Signal Processing Letters,*

IEEE , vol.16, no.5, pp.354,357, May 2009 doi: 10.1109/LSP.2009.2016003

[43] Kandeepan, S.; Reisenfeld, S.; Aysal, T.C.; Lowe, D.; Piesiewicz, R., "Bayesian Tracking in Cooperative Localization for Cognitive Radio Networks," Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th , vol., no., pp.1,5, 26-29 April 2009

[44] A. Goldsmith, Wireless Communications, Cambridge University Press, New York, NY, 2005.

[45] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. J. Royal Stat. Soc. Ser. B, 39(1):1–38, 1977.

[46] A. O. Hero and J. A. Fessler. Asymptotic convergence properties of EM-type algorithms. Technical Report 282, Comm. and Sign. Proc. Lab., Dept. of EECS, Univ. of Michigan, Ann Arbor, MI, 48109-2122, April 1993

[47] J. A. Fessler, N. H. Clinthorne, and W. L. Rogers. On complete data spaces for PET reconstruction algorithms. IEEE Tr. Nuc. Sci., 40(4):1055–61, August 1993.

[48] Randolph L. M., Dushyanth. K, Robert M. P.A Self-LocalizationMethod forWireless Sensor Networks, EURASIP Journal on Applied Signal Processing 2003:Vol.4,pp 348358

[49] M. S. Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp, A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking, IEEE Trans on Signal Proc, Vol.50, No.2, Feb 2002

[50] P. M. Djuric et. al. ”Particle Filtering” , IEEE Signal Processing Magazine, September 2003.

[51] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. Numerical recipes in C. Cambridge Univ. Press, New York, 1988

[52] M. Segal and E. Weinstein. The cascade EM algorithm. Proc. IEEE, 76(10):1388–1390,October 1988.

[53] Bohanudin, S.; Ismail, M.; Abdullah, M., "Hybrid localization techniques in multilayer heterogeneous network under low hearability," Wireless Technology and Applications (ISWTA), 2012 IEEE Symposium on , vol., no., pp.16,21, 23-26 Sept. 2012
doi: 10.1109/ISWTA.2012.6373835

[54] J. A. Fessler and A. O. Hero. New complete-data spaces and faster algorithmsfor penalized-likelihoodemission to- mography. In Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf., volume 3, pages 1897–1901, 1993.