

**This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.**

Author(s): Mondal, Riaz; Turkka, Jussi; Ristaniemi, Tapani

Title: An efficient grid-based RF fingerprint positioning algorithm for user location estimation in heterogeneous small cell networks

Year: 2014

Version:

Please cite the original version:

Mondal, R., Turkka, J., & Ristaniemi, T. (2014). An efficient grid-based RF fingerprint positioning algorithm for user location estimation in heterogeneous small cell networks. In Proceedings of 2014 International conference on localization and GNSS (ICL-GNSS) (pp. 1-5). IEEE. International Conference on Localization and GNSS. <https://doi.org/10.1109/ICL-GNSS.2014.6934169>

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

An Efficient Grid-based RF Fingerprint Positioning Algorithm for User Location Estimation in Heterogeneous Small Cell Networks

Riaz Mondal^{1,2}, Jussi Turkka³, Tapani Ristaniemi¹

¹Department of Mathematical Information Technology, University of Jyväskylä, Jyväskylä, Finland

²Department of Information and Communication Engineering, University of Rajshahi, Rajshahi, Bangladesh

³Magister Solutions Ltd, Tampere, Finland

riaz.u.mondal@student.jyu.fi, jussi.turkka@magister.fi, tapani.ristaniemi@jyu.fi

Abstract—This paper proposes a novel technique to enhance the performance of grid-based Radio Frequency (RF) fingerprint position estimation framework. First enhancement is an introduction of two overlapping grids of training signatures. As the second enhancement, the location of the testing signature is estimated to be a weighted geometric center of a set of nearest grid units whereas in a traditional grid-based RF fingerprinting only the center point of the nearest grid unit is used for determining the user location. By using the weighting-based location estimation, the accuracy of the location estimation can be improved. The performance evaluation of the enhanced RF fingerprinting algorithm was conducted by analyzing the positioning accuracy of the RF fingerprint signatures obtained from a dynamic system simulation in a heterogeneous LTE small cell environment. The performance evaluation indicates that if the interpolation is based on two nearest grid units, then a maximum of 18.8% improvement in positioning accuracy can be achieved over the conventional approach.

Keywords— *Grid-based RF fingerprint; Kullback-Leibler Divergence; Minimization of Drive Tests*

I. INTRODUCTION

Various location-based services in wireless communication networks depend on mobile positioning. Commercial examples range from low-accuracy methods based on cell identification to high-accuracy methods combining wireless network information and satellite positioning. These methods are typically network centric, where the position is determined in the network and presented to the user via a specific service. Due to logistical reasons, the position is estimated from static snapshot measurements, possibly provided by the mobile station (MS) [1]. A major driving force to estimation mobile user location is the requirement for E-911 emergency positioning in the North American market. This requirement calls for accuracies of 50 meters (68%) and 150 meters (95%) in the terminal-dependent case and 100 meters (68%) and 300 meters (95%) in the network-dependent case [2]. Today, the Global Navigation Satellite System (GNSS) is the most effective positioning technology in the outdoor open environments [3]. However it has limitations such as poor performance in built-up areas and high power consumption. These limitations led to the development of positioning techniques based on the wireless networks. These technologies

include a variety of time-of-arrival techniques (ToA and TDoA), angle-of-arrival techniques (AoA), and location fingerprinting techniques [4]. Typically, RF fingerprinting refers to a database correlation method where the position is estimated by comparing the radio measurements e.g., the RF fingerprint of the user equipment (UE) with the training fingerprints in the correlation database. The training fingerprints consist of received signal strength (RSS) radio measurements from several base stations (BS) that are used to provide a fingerprint of the radio conditions at a specific geographical location. Typically, this location is determined with an accurate positioning method, for example GNSS. Hence, fingerprinting is a positioning method which exploits the already existing infrastructures such as cellular networks [5] and WLANs [6].

One of the biggest challenges of RF fingerprinting is the burden of creating and maintaining the correlation database of the training fingerprints. Operators can maintain the correlation database by conducting extensive and expensive periodical drive test campaigns to collect the required measurements. However, the concept of Minimization of Drive Tests (MDT) provides a framework for gathering user reported location-aware radio measurements from commercial mobile phones that can be used for creating and maintaining such training databases [7], [8]. In fact, one of the benefits of MDT is that it provides an efficient way to automate the collection of training fingerprints in Universal Terrestrial Radio Access Network (UTRAN) and Evolved-UTRAN (EUTRAN) cellular systems. The MDT procedure allows operators to collect radio measurements, i.e. received signal strength and quality, with UE location information and a time stamp. It is worth noting that for immediate MDT, the network can make a request to UE to attempt to make GNSS location information available for MDT [8].

In our previous work [9], [10], we have evaluated the RF fingerprint positioning accuracy using Kullback-Leibler Divergence (KLD) and Mahalanobis distance with a single grid layout (SGL) in rural, urban and heterogeneous small cell networks. Although the positioning accuracy in dense urban scenario was good, the positioning error (PE) did not fulfill the E-911 requirements in the rural case.

In this work, we propose a novel RF fingerprint algorithm to enhance the positioning accuracies further for rural case and

also in the dense urban scenario. The proposed enhancements are based on an overlapping grid layout technique to produce the training signatures and to determine the estimated position using the two smallest KLD weighted values among the matched training signatures. The effect of using the overlapping grid layout (OGL) is that it increases the training signatures as well as reduces the distances among the surrounding grid centers and thereby improving the positioning accuracy. However it increases the computational time needed to estimate the testing UE positions as compared to that of the traditional SGL method by a small amount.

This paper is organized as follows. Section 1 gives an introduction to the research problem. In Section 2, first conventional SGL RF fingerprinting and then the OGL approach using Kullback-Leibler Divergence method is described. Finally in Section 3, the performance evaluation of the enhancements is discussed in the light of extensive system simulations.

II. GRID-BASED RF FINGERPRINTING FRAMEWORK

A. Single Grid Layout

In SGL approach, RF fingerprinting correlation database is compressed to present the geographical space $G=\{g_1, g_2, \dots, g_N\}$ e.g., “area of interests” as a regular tessellation of N square grid units g_i as depicted in Fig. 1. Each grid unit g_i is associated with a center point having coordinates $c_i=\{x_i, y_i\}$ and a set of RF fingerprint training signatures $S_i=\{s_{i,1}, s_{i,2}, \dots, s_{i,k}\}$ which will depend upon different sets of same BS id MDT samples. Hence, a training signature $s_{i,j}$ is the j th signature associated with i th grid unit and it is constructed from n different MDT measurement samples. Each sample in $\mathbf{M}_{i,j}$ contains detailed location information e.g., x and y coordinate obtained from GNSS receiver, and m RSS measurement e.g., Reference Signal Received Power (RSRP) in EUTRAN, from different BSs.

$$\mathbf{M}_{i,j} = \begin{bmatrix} x_1 & y_1 & rsrp_{1,1} & \cdots & rsrp_{1,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n & y_n & rsrp_{n,1} & \cdots & rsrp_{n,m} \end{bmatrix}, \quad (1)$$

where, $rsrp_{l,m}$ is the 1st RSRP measurement with a BS identifier m . Moreover, Euclidean distance between the detailed locations of the n measurement samples and the center point of c_i suggests that the nearest grid unit is g_i . For ensuring an efficient and fast processing of the measurement data, the MDT training measurements are compressed e.g., for each grid g_i , the signatures in S_i stores only the means and the covariance of the received signal strengths between the detected BSs in $\mathbf{M}_{i,j}$ as discussed in [9], [10].

B. Overlapping Grid Layout

In Overlapping Grid Layout (OGL) approach, the whole area of interest is divided into square grids similar to the SGL approach but there are now multiple layouts which overlap each other. In the present work, OGL consisted of two grid

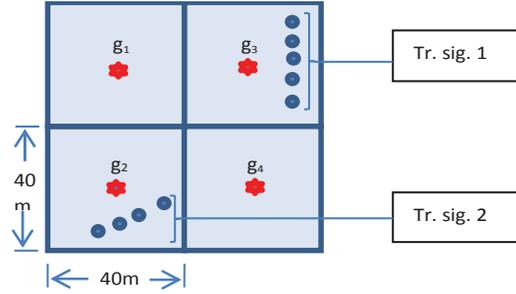


Fig. 1: Single grid layout based conventional RF fingerprint positioning

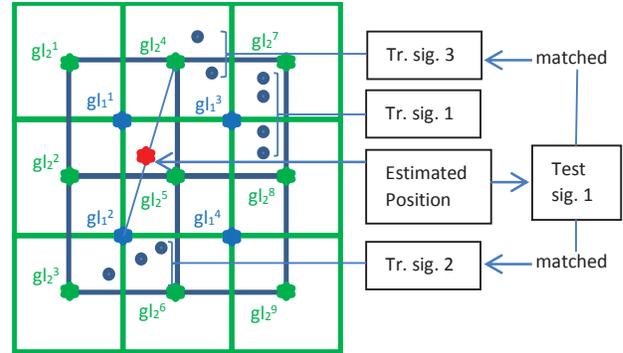


Fig. 2: Two overlapping grid-cell based RF fingerprint positioning

layouts denoted by OGL_1 and OGL_2 having same sized square grids. OGL_2 was placed in a fashion that an OGL_1 grid unit is overlapped by $1/4$ th area of an OGL_2 grid unit. In this way OGL_1 grid unit is fully overlapped by four OGL_2 grid units as shown in Fig. 2, here the squares with blue border lines belongs to OGL_1 layout and the blue stars represents the corresponding grid centers, whereas the squares with green border lines and green stars depicts the OGL_2 layout. Thus the center points of four OGL_2 grid units fall on the four corners of the overlapped OGL_1 grid unit.

Training signatures are formed in grid-wise manner. A training signature comprises of a set MDT samples that are located within a grid unit and these MDTs were reported to have the same BS ids. The same set of training samples is used to form the training signatures for OGL_1 and then for OGL_2 . Because of the offset between the layouts mostly different training signatures will be associated with the layouts. Inherently some common signatures will exist between these layouts. All the training signatures belonging to OGL_1 and OGL_2 layouts are put together to create the whole training signature database.

During the testing phase, all training signatures that match with the testing signature BS ids are searched, and then the KLD distances are calculated between the testing signature and those training signatures. In KLD method, training and testing phase signatures e.g., mean vectors \mathbf{u} and covariance matrices Σ parameterize multivariate Gaussian distributions $p(\mathbf{x}|\mathbf{u}, \Sigma)$, and therefore, the method aims to exploit interdependencies among the received signal strengths between BSs by using covariance matrices for training and

testing phase signatures as given by the following closed form KLD equation:

$$d(p_u \| p_{i,t}) = \frac{1}{2} \left((\mathbf{u}_u - \hat{\mathbf{u}}_{i,t})^T \hat{\Sigma}_{i,t}^{-1} (\mathbf{u}_u - \hat{\mathbf{u}}_{i,t}) + \text{tr}(\hat{\Sigma}_i \hat{\Sigma}_{i,t}^{-1} - \mathbf{I}) - \ln |\hat{\Sigma}_i \hat{\Sigma}_{i,t}^{-1}| \right) \quad (2)$$

where, \mathbf{u}_u , and $\hat{\mathbf{u}}_{i,t}$ corresponds to the mean received signal strength values, while $\hat{\Sigma}_u$ and $\hat{\Sigma}_{i,t}$ represents the covariance matrices of the received signal strength values of the testing and training signatures respectively. Here, $\text{tr}()$ denotes the trace of matrix, Σ^{-1} denotes the inverse of covariance matrix Σ and \mathbf{I} is the identity matrix. It is a non-symmetric measure of the difference between testing and training signature probability distributions p_u and $p_{i,t}$. Two grid units are selected which correspond to the two smallest KLD distance values among all the matched training signatures. Now the estimation point for the testing signature is calculated to be in between the center points of these two grids according to the KLD weight factor of the corresponding grids. For example, in Fig. 2, the testing signature 1 i.e., ‘test sig.1’, has two matched training signatures: ‘tr.sig. 2’ which belongs to OGL_1 and ‘tr.sig 3’ of OGL_2 layout. In order to determine the estimated position for the testing signature, the KLD values of these two training signatures are used as weights and depending upon the weighted value the estimation point moves along the line joining the center points of the corresponding grid units as described in next section.

C. Weighted Overlapping Grid Layout

The enhanced algorithm for a grid-based RF fingerprinting uses the weighted geometric center of a set of nearest grid units to determine the location of a testing signature in a two dimensional grid. Formally, the geometric center (or centroid) is an arithmetic mean of a set of points weighted by special weight i.e., point density or mass. In the context of this paper, the KLD metric is used to find the n nearest grid units and to weight the arithmetic mean.

To determine the location of a testing signature s_i , a subset $G_i = \{g_1, g_2, \dots, g_k\}$ of nearest grid units based on KLD metric is chosen depending upon the matched BS ids of the testing signature. In addition, each grid unit in G_i is assigned with a weight w_j and therefore a set of weights $W_i = \{w_1, w_2, \dots, w_k\}$ is obtained. Since, the KLD metric is smallest for the nearest grid unit, the weighting factors are inversely proportional to KLD metric for ensuring that more weight is given to the nearest grid unit. The weighting factor w_j for j th grid unit g_j in subset G_i is given as,

$$w_j = \frac{1}{\zeta_j} = \frac{1}{\sum_{l=1}^k \frac{1}{\zeta_l}} \quad (3)$$

where, ζ_j is the KLD metric between testing signature s_i and j th grid unit in the subset G_i of the k nearest grid units. Note that the weights in W_i are normalized so that they sum up to 1.

Therefore, the weighted geometric center of the testing signature is obtained from,

$$[x_i, y_i] = (\mathbf{w}_i)^T [\mathbf{x}_i, \mathbf{y}_i], \quad (4)$$

where \mathbf{w}_i , \mathbf{x}_i , and \mathbf{y}_i are column vectors containing the weights and x and y coordinates of the k nearest grid units in G_i . Moreover, $(\cdot)^T$ is a vector transpose operation. Hence, (4) gives the scalars x_i and y_i which is the location of the weighted arithmetic mean of the points in G_i and the estimated location of s_i . In this study, k is equal to 2, so only two nearest neighbors where considered for analyzing the location estimates for testing signatures.

III. PERFORMANCE EVALUATION

A. Simulation Scenario

In this study we have conducted a performance evaluation of our enhanced RF fingerprinting method having two overlapping grid layouts with the traditional single grid layout. For this purpose, we have performed two simulations in two different network scenarios, (i) sparse regular macro (RM) and (ii) heterogeneous small cell (HSC) network as illustrated in Fig. 3 and Fig. 4 respectively. Both scenarios are the same as in [10] and those were simulated by using a state-of-the-art LTE Rel’10 dynamic system simulator to model both the downlink and the uplink in an OFDM symbol resolution with several radio resource management, scheduling, mobility, handover and traffic modeling functionalities. Simulation parameters and mathematical models are based on the 3GPP specifications, (especially in the simulation assumptions in 3GPP TR 36.839) defining parameterization for used bandwidth, center frequency, path-loss, slow fading, and fast fading [11]. Moreover, UE measurements e.g., RSRP, were implemented in the simulator taking into account the technical requirements for the absolute and relative measurement errors and -6 dB \hat{E}_s/lot cell detection criterion as in [12].

The RM network consisted of 57 BSs operating on primary frequency band (CC0) and 36 sectored high power Pico BSs operating on non-overlapping secondary carrier (CC1). Inter-

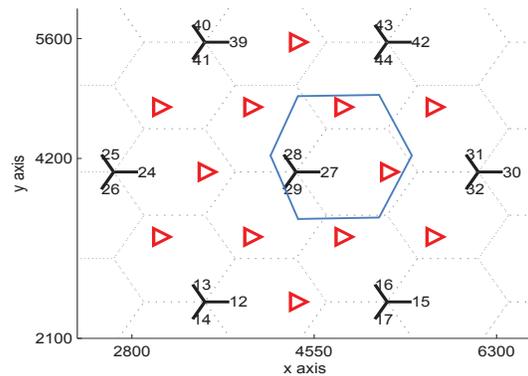


Fig. 3: Sparse regular macro scenario

site distance between the BSs on CC0 and CC1 are 1750m and 875m, respectively. Locations of three-sectored Pico sites on CC1 were shifted to be in between the macro BSs on CC0 as depicted in Fig. 3 with red triangles. In RM scenario a total of 82320 MDT samples were generated for the analysis. For this case we have picked up the MDT samples from an area enclosed by the blue hexagon as shown in Fig. 3 in order to maintain size similarity between the RM and HSC scenarios.

The HSC network consisted of 57 macro BSs having inter-site distance of 500m and operating on CC0 band. In addition,

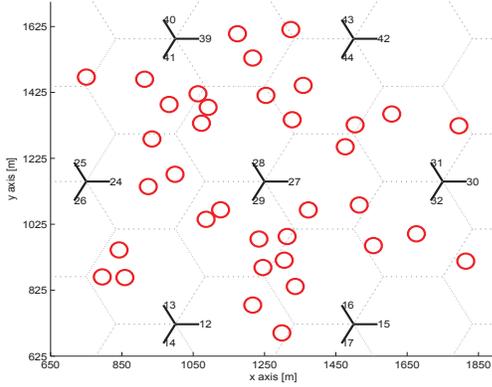


Fig. 4: Heterogeneous small cell scenario

36 small cells with omni-directional antennas were randomly deployed in the coverage area of 12 centermost macro BSs as depicted in Fig. 4 with red circles. Distance to the nearest inter-site small cell varied from 50m to 170m with average distance being 95m. In HSC scenario, UEs were moving only in the area of 12 centermost macro BSs but were able to monitor all detected cells. In RM scenario, UEs were moving in one specific area with area size similar to the simulation area in HSC scenario. In total, 1200 randomly moving outdoor vehicular UEs (30 km/h) were distributed uniformly to the simulation area. Traffic profile consisted of data generated by MDT reports which were sent once per second, however, 100% resource block loading was used for creating interference limited simulation environment which is more challenging from cell detection point of view. More details about the used simulation parameters can be found in the Table I in [10].

B. Performance Results

In this study the positioning accuracies were analyzed using two different amounts of training data and testing data. First 90% of the MDT data samples were used in training and the estimation was done for the remaining 10% MDT samples. Then 10% of the MDT data samples were taken for training and 90% for the testing purpose. Such an approach was selected for studying how the amount of training data affects the performance of OGL method. The training and the testing database samples are selected in call-wise manner e.g., all MDT reports from single UE belongs always either on the training or the testing database. This is done to avoid too

optimistic positioning results in cases where consecutive measurements from one UE are found in training and testing databases. The number of training signatures used and analyzed test signature percentages in SGL and OGL methods are given in Table I. From this table we can find that number of training signatures used in OGL is about twice the number used in SGL for all the different simulations.

TABLE I. ANALYZED RF FINGERPRINT SIGNATURES OF SGL AND OGL LAYOUTS

Grid-cell size	Scenario	Training Data (%)	Total no. of Training Signatures (Absolute)		Analyzed Test Signatures (%)	
			SGL	OGL	SGL	OGL
10-by-10 M	RM (ISD 1750M)	90%	16443	32758	83.19	84.86
		10%	2044	4092	62.19	69.07
	HSC (ISD 500M)	90%	48808	97687	71.66	73.39
		10%	6319	12707	47.60	55.14
40-by-40 M	RM (ISD 1750M)	90%	7090	14222	82.64	85.72
		10%	1709	3401	64.02	73.80
	HSC (ISD 500M)	90%	22079	44219	70.50	74.62
		10%	5321	10677	47.74	61.43

TABLE II. POSITIONING ACCURACY PERFORMANCE EVALUATION

Scenario	Training Data (%)	RF Fingerprint Algorithm	For 10-by-10 m Grid		For 40-by-40 m Grid	
			68% PE (m)	95% PE (m)	68% PE (m)	95% PE (m)
RM (ISD 1750 M)	90%	SGL Based	29.73	165.29	43.53	196.30
		OGL Based	31.41 (+5.6%)	147.49 (-10.7%)	40.86 (-6.1%)	161.75 (-17.6%)
	10%	SGL Based	72.00	228.80	72.48	225.45
		OGL Based	63.96 (-11.1%)	206.05 (-9.9%)	65.03 (-10.2%)	203.70 (-9.6%)
HSC (ISD 500 M)	90%	SGL Based	21.12	58.08	33.73	76.43
		OGL Based	19.45 (-7.9%)	50.94 (-12.2%)	27.57 (-18.2%)	64.87 (-15.1%)
	10%	SGL Based	27.23	73.61	34.83	80.86
		OGL Based	25.14 (-7.6%)	66.47 (-9.6%)	28.28 (-18.8%)	68.71 (-15.0%)

It is noticeable from Table I that for 90% of training data the increased percentage of analyzed test signatures for OGL as compared to SGL method is about 1.7% and 3% for 10 by 10 m and 40 by 40m grid-cells respectively. Whereas for 10% of training data the improvement shown by OGL method over SGL is about 7% and 10% for 10 by 10 m and 40 by 40m grid-cells respectively in analyzing the test signatures.

The simulation results for the conventional SGL and the proposed OGL RF fingerprint positioning in RM and HSC scenarios are given in Table II. Here 68 and 95 percentiles give the positioning error (PE) in meters for SGL and OGL methods using different training data sets during simulations. The percentage within the first bracket under the PE value of each of the OGL based method indicates the degradation (with a + sign) or improvement (with a - sign) in positioning as compared to the SGL method. For example, in the HSC 10% training case with 40m by 40m grid the SGL fingerprint positioning algorithm has given 34.83 meters and 80.86 meters of PE in the 68 and 95 percentiles respectively whereas that of OGL is 28.28 meters and 68.71 meters respectively. Hence OGL method has reduced the PE by 18.8% and 15.0% in 68 and 95 percentiles respectively as compared to that of the SGL method. Also for HSC 90% training case with the same grid size OGL has shown 18.2% and 15.1% of improvement in 68 and 95 percentiles respectively over that of the SGL. Table I depicts that almost in every simulated study OGL RF Fingerprinting performs better than SGL method. Moreover, it is worth highlighting that if only 10% samples are available for constructing the training signatures, 40-by-40 meters OGL can outperform 10-by-10 meters SGL performance.

In our present study the MDT samples comprising the testing signatures tend to be distributed over several training grid units, as a result positioning accuracies were improved when an overlapping grid layout is used for the formation of training signatures and then the centroid point is determined from the two smallest weighted KLD grids. The OGL method has constructed more training signatures with different combinations of BS IDs than the SGL method. Therefore, it is more probable to find matching combination of BS IDs in OGL method. Hence the proposed algorithm not only improves the positioning accuracy but it can also analyze more testing samples.

IV. CONCLUSION

This paper proposed two enhancements for grid-based RF fingerprint position estimation framework. First enhancement is an introduction of overlapping grid layout to form training signatures. In the second enhancement, the location of the testing signature is estimated to be a weighted geometric center of a set of nearest grid units. This is different from the traditional grid-based RF fingerprinting where only the center point of the nearest grid unit is used for determining the user location. The performance evaluation of the enhanced RF

fingerprinting algorithm was conducted by analyzing the positioning accuracy of the RF fingerprint signatures obtained from a dynamic LTE system simulation.

The proposed enhancements can increase the number of training signatures that needs to be analyzed for finding the nearest grid but in addition the positioning accuracy is increased, specifically in cases where only a limited amount of data is used for constructing the training signatures. The performance evaluation indicates that the proposed overlapping grid layout method using KLD can provide a maximum of 18.8% improvement in positioning accuracy as compared to that of the conventional single grid layout approach.

ACKNOWLEDGEMENT

The present work was carried out within the framework of European Celtic-Plus project SHARING (Self-organized Heterogeneous Advanced Radio Networks Generation).

REFERENCES

- [1] F. Gustafsson and F. Gunnarsson, "Mobile Positioning Using Wireless Networks," *IEEE Signal Processing Magazine*, July 2005, pp. 41-53.
- [2] The FCC, *Fact Sheet—FCC Wireless 911 Requirements*, Jan. 2001, FCC.
- [3] E.C.L. Chan, G. Bacieu, "Wireless tracking analysis in location fingerprinting," in *Proc. IEEE International conference on Wireless and Mobile Computing*, October 2008, pp. 214–220.
- [4] A. Kupper, *Location based services: Fundamentals and Operation*. John Wiley & Sons Ltd, England, August 2005.
- [5] C. Takenga, K. Kyamakya, "A low-cost fingerprint positioning system in cellular networks," in *Proc. International Conference on Communications and Networking*, August 2007, pp. 915 – 920.
- [6] P. Bahl, V.N Padmanabhan, "Radar: an in-building rf-based user location and tracking system," in *Proc. Joint Conference of the IEEE Computer and Communications Societies*, vol. 2, March 2000, pp. 775 – 784.
- [7] Wuri A. et al., "Minimization of Drive Tests Solution in 3GPP", in *IEEE Communications Magazine*, Vol. 50, No. 6, June 2012.
- [8] Johansson J., Hapsari W.A, Kelley S. and Bodog G., "Minimization of drive tests in 3GPP release 11", in *IEEE Communications Magazine*, Vol. 50, No. 11, pp. 36-43, November 2012.
- [9] Mondal R., Turkka J., Ristaniemi T. and Henttonen T., "Performance Evaluation of MDT RF Fingerprinting Framework", in *Proc. of 7th International Conference on Mobile Computing and Ubiquitous Networking (ICMU2014)*, January 2014, Singapore.
- [10] Mondal R., Turkka J., Ristaniemi T. and Henttonen T., "Positioning in Heterogeneous Small Cell Networks using MDT RF Fingerprints", in *Proc. of the First IEEE International Black Sea Conference on Communications and Networking*, July 2013, Batumi, Georgia.
- [11] 3GPP TR 36.814, "Further advancements for E-UTRA physical layer aspects", version 9.0.0, March 2009, available at <http://www.3gpp.org>.
- [12] 3GPP TS 36.133, "Requirements for support of radio resource management", version 11.4.0, March 2013, available at <http://www.3gpp.org>.