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

Title: Genetic Algorithm Optimized Grid-based RF Fingerprint Positioning in Heterogeneous Small Cell Networks

Year: 2015

Version:

Genetic Algorithm Optimized Grid-based RF Fingerprint Positioning in Heterogeneous Small Cell Networks

Riaz Uddin Mondal
Department of Mathematical Information Technology
University of Jyvaskyla
Jyvaskyla, Finland
riaz.u.mondal@student.jyu.fi

Tapani Ristaniemi
Department of Mathematical Information Technology
University of Jyvaskyla
Jyvaskyla, Finland
tapani.ristaniemi@jyu.fi

Jussi Turkka
Magister Solutions Ltd.
Tampere, Finland
jussi.turkka@magister.fi

Abstract— **In this paper we propose a novel optimization algorithm for grid-based RF fingerprinting to improve user equipment (UE) positioning accuracy. For this purpose we have used Multi-objective Genetic Algorithm (MOGA) which enables autonomous calibration of grid-cell layout (GCL) for better UE positioning as compared to that of the conventional fingerprinting approach. Performance evaluations were carried out using two different training data-sets consisting of Minimization of Drive Testing measurements obtained from a dynamic system simulation in a heterogeneous LTE small cell environment. The robustness of the proposed method has been tested analyzing positioning results from two different areas of interest. Optimization of GCL is performed in two ways: (1) array-wise calibration of the grid-cell units using non-overlapping GCL and (2) creating an overlapping GCL to cover of whole simulation area with different rectangular grid-cell units. Simulation results show that if sufficient amount of training data is available then the proposed method can improve positioning accuracy of 56.74% over the conventional grid-based RF fingerprinting.**

Keywords- Grid-based RF fingerprinting; Minimization of Drive Tests; Multi-objective Genetic Algorithm; Kullback-Leibler Divergence.

I. INTRODUCTION

Positioning in wireless networks is dependent on the mobility of users and the dynamic nature of both the environment and radio signals. Users expect the same level of performance whether they are indoors or outdoors in a rural or urban environment. So far no single positioning method, including GPS, works well in all environments [1]. Receivers in Global Navigation Satellite Systems (GNSS) such as GPS or GLONASS tend to output inaccurate location estimations while operating in urban regions, mostly due to the density of tall buildings, which often block a receiver's line of sight to the navigation satellites [2]. Among the non-standard positioning methods included in LTE Release 9, RF fingerprinting is the most cost-efficient solution for indoor WLAN positioning [3], [4], [5] as well as for outdoor mobile cellular positioning in densely built urban environments [6], [7]. RF fingerprinting, also known as database correlation method (DCM) finds a user's position by mapping RF measurements obtained from the UE onto an RF map, where the map is typically based on detailed RF predictions or site surveying results.

An ideal positioning system should be self-learning and environmentally adaptive, capable of building up information databases that store actual observations, and employ smart data analysis mechanisms [1]. In order to achieve such a goal, a functionality known as Minimization of Drive Tests (MDT) has been proposed in LTE Release 10 which reduces the huge cost and efforts associated with the conventional drive test measurement procedure. MDT provides a framework for gathering user reported location-aware radio measurements from commercial mobile phones that can be used for creating and maintaining RF fingerprint training databases [8], [9]. It allows operators to autonomously build and update large training database for RF fingerprinting from various locations of user experience along with available location information from UEs without extra hardware installation.

In [10] and [11] grid-based RF fingerprinting has shown good positioning performance in rural, urban and heterogeneous small cell networks scenario using MDT simulated data. An overlapping grid-cell layout (GCL) based RF fingerprinting approach was proposed in [12] which further improves the positioning accuracy. The performance results of [11] indicate that using a common GCL in different network scenarios does not provide the best positioning accuracy in both 68%-ile and 95%-ile of positioning error (PE) estimation. It was also found

from [12] that positioning performances varies of a GCL that is composed of same square grid-cell units when amount of training data is varied. From literature review we did not find an explicit method for optimizing GCL in grid-based RF Fingerprint positioning. In [13] GA was used to reduce the correlation space in RF fingerprinting location method to improve location accuracy and authors have claimed their method to be suitable for UE positioning in urban environments. Authors in [14] have proposed a location detection algorithm which employs cell-id positioning enhanced by triangulation. The process was accelerated through the application of a GA.

In this paper, we propose a novel method based on MOGA to develop an autonomous grid-based RF fingerprinting by optimizing the GCL in order to achieve the best possible positioning accuracy despite changes in cellular network scenarios and amounts of training data. This would render the positioning system an adaptive one which can make necessary adjustments to the grid-cell size along with the structural and environmental changes of the surrounding for an optimal RF fingerprint positioning performance.

The following section contains a brief description of the conventional RF fingerprinting using MDT measurements. In section III first a brief description of multi-objective GA is given, and then the MOGA optimized RF fingerprint positioning (GAFP) method is explained. Finally section 4 discusses the performance evaluation of GAFP in the light of extensive system simulations.

II.   RF FINGERPRINTING USING MDT MEASUREMENTS

A.   Minimization of drive tests

Conventional drive tests consume significant time and human efforts to obtain reliable data [8]. MDT is a feature introduced in 3GPP Release 10 that enables operators to utilize UE to collect radio measurements and associated location information, in order to assess network performance while reducing the large operation expenditure associated with traditional drive tests. Location information for MDT can be categorized into two different types: detailed location information and RF fingerprint. Detailed location information is typically obtained by GPS or Global Navigation Satellite System (GNSS) positioning method, but can also be obtained by other positioning methods supported by the UE and the network, e.g., Observed Time Difference of Arrival (OTDOA), Assisted -GNSS (Assisted-GPS) or Enhanced Cell ID (E-CID). Whereas in RF fingerprint type of MDT a profile of measured signal strength from neighboring cells is created. A major difference between conventional drive tests and MDT is that conventional drive tests use a controlled application with known traffic characteristics, whereas MDT uses ordinary user traffic with largely unknown characteristics [9].

Thus operators are able to build a large RF fingerprint training database from UE's in different network environments and up-date it as needed. This urges for an autonomous RF fingerprinting method which makes necessary calibration to its GCL with the updated MDT training database.

B.   Grid-cell Based RF Fingerprinting

In a conventional grid-cell layout (CGCL) based RF fingerprinting method using Kullback-Leibler Divergence (KLD) has two main phases [3], [10], [11]:

Training Phase: First an offline processing and manipulation of MDT correlation database takes place. A layout of adjoining rectangular or square grid-cell units are formed over the whole geographical area of interest. Each of these grid units g_i is associated with a center point having coordinates $c_i=\{x_i, y_i\}$ as shown in Fig.1. Then unit-wise creation of training signatures is performed by gathering MDT measurements having signal strength values, e.g., Reference Signal Received Power (RSRP) from same set of base-stations (BSs). Hence depending upon the size of a grid-cell unit I and the availability of MDT measurement samples within that grid, it may contain multiple training signature $S_{i,j}=\{s_{i,1}, s_{i,2}, ..., s_{i,j}\}$, where j is the total number of signatures. Each signature consists of multiple MDT measurements collected from that particular grid-cell unit having RSRP values from similar BS Ids along with detail location information of corresponding MDT measurement positions. For example the signature $s_{1,1}$ as shown in Fig. 1 consists of four MDT samples and can be expressed by:

$$s_{1,1} = \begin{pmatrix} c_{1,1} & ID_{1,1} & rsrp_{1,1} \\ \vdots & \vdots & \vdots \\ c_{1,4} & ID_{1,4} & rsrp_{1,4} \end{pmatrix} \qquad (1)$$

where, $c_{i,j}$ is a vector of x and y coordinates for an MDT measurement, $ID_{i,j}$ stands for the cell identities which is same for all the samples that form a signature and the vector $rsrp_{i,j}$ contains RSRP values of corresponding MDT samples.
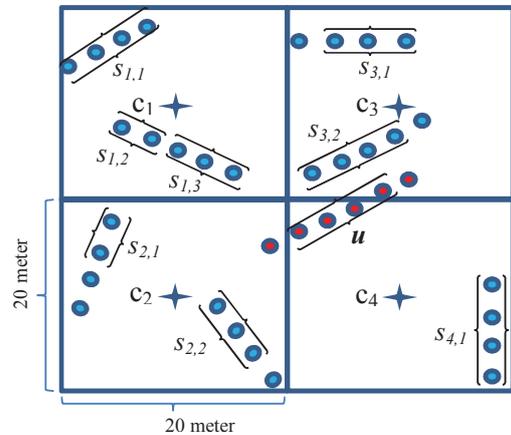

Figure 1. 20-by-20 meter grid-cell layout of a conventional RF fingerprint positioning

Fig.1 illustrates a conventional grid-based RF fingerprinting method. Here the area of interest has been divided into four square grid-cells, all of which have the same arm length of 20 meters and corresponding grid centers as depicted by c_1, c_2, c_3 and c_4. The training MDT samples are shown here by blue dots while the testing samples by red dots. It can be seen from this figure that no all the training samples which belong to a grid-

cell form training signatures, but only those which matches in BS ids.

Testing Phase: During this phase testing signatures are formed from the MDT samples obtained from a particular UE. As we can see from the Fig.1 we have multiple training signatures and one testing signature depicted by u in this example. At first training signatures are collected which match with the testing signature BS IDs. Here if we assume that we have three such matched training signatures- $s_{1,3}$, $s_{2,2}$, and $s_{3,2}$, then KLD is calculated between u and matched training signature RSRP values. Now if the KLD between u and $s_{3,2}$ gives the minimum value then the center location of grid-cell 3, which is $\mathbf{c_3}$ is chosen to be the estimated position for the MDT samples of testing signature u. The closed form KLD equation is given by:

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

where, $\mathbf{u}_{u,}$ and $\hat{\mathbf{u}}_{i,t}$ corresponds to the mean received signal strength values, while Σ_u and $\Sigma_{i,t}$ represents the covariance matrices of the received signal strength values of the testing and training signatures respectively. Here, $tr()$ 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 $p_{i,t}$.

III. GRID-CELL LAYOUT OPTIMIZATION USING GENETIC ALGORITHM

A. Motivation

From our previous work in this field it is clear that GCL optimization is necessary to make RF fingerprinting an adaptive positioning technique which would deliver its optimal performance in both 68%-ile and 95%-ile of PE despite changes in the BS positions, surrounding structures or amount of MDT training samples used [10],[11],[12]. Here our goal is to autonomously select a GCL which would deliver optimal positioning performance in both the 68%-ile and 95%-ile value of PE. From simulation results we found that using square GCs between 20m-by-20m to 40m-by-40m, for different square Grid-cell layouts the 68 and 95 percentiles of positioning error tends to be conflicting - when 68%-ile value of PE decreases then 95%-ile value increases and vice-versa. Therefore we were motivated to use a popular Multi-Objective Evolutionary Algorithm (MOEA) know as Multi-Objective GA (MOGA) to select the proper GCL for RF fingerprinting, where the objective functions are 68th and 95th percentiles of positioning error. The reason for employing MOGA is that other stochastic search strategies (e.g., simulated annealing, ant colony optimization or particle swarm optimization) do not guarantee to find the true Pareto optimal set but, instead, aim to generate a good approximation of such set in a reasonable computational time. On the other hand, MOEAs are particularly well-suited to solve multi-objective problems because they operate over a set of potential solutions. This feature allows them to generate several elements of the Pareto optimal set in a single run. Furthermore, MOEAs are less

susceptible to the shape or continuity of the Pareto front than traditional mathematical programming techniques, require little domain information and are relatively easy to implement and use.

B. Genetic Algorithm Optimized RF Fingerprinting

The term genetic algorithms refer to a subset of evolutionary algorithms that model biological processes to optimize highly complex cost functions. GA is capable of yielding a robust search by implicitly sampling hyper-plane partitions of a search space. A single hyper-plane, commonly referred to as schema, is the theoretical foundation on which the algorithm was developed as first introduced by John Holland in 1975 [15]. Here a population represents a group of potential solution points and a generation represents an algorithmic iteration. A chromosome is comparable to a design point and a gene is comparable to a component of the design vector. Given a population of designs, three basic operations are applied: reproduction, crossover, and mutation. Reproduction involves selecting design vectors from the current generation to be used in the next generation and whether or not a design is selected depends on its fitness value. Fitness, which is determined by a fitness function, is an indication of how desirable a design is in terms of surviving into the next generation. The selection probability represents the chance for survival and is proportional to a design's fitness value. Once a new generation of designs is determined, crossover is conducted as a means to introduce variations into the population of designs. Crossover is the process of combining or mixing two different designs. The next operation, which also is used to introduce variations into the population, is mutation. It is a random process that entails altering part of a design's genetic string. In our simulations we have used the multi-objective GA function 'gamultiobj' of Matlab R2014a which uses a controlled elitist genetic algorithm, a variant of NSGA-II [16]. This controlled elitist GA favors individuals that can help increase the diversity of the population even if they have a lower fitness value. It is very important to maintain the diversity of population for convergence to an optimal Pareto front. This is done by controlling the elite members of the population as the algorithm progresses. Two options 'ParetoFraction' and 'DistanceFcn' are used to control the elitism. The Pareto fraction option limits the number of individuals on the Pareto front and the distance function helps to maintain diversity on a front by favoring individuals that are relatively far away on the front. Here a crowding distance for each member is calculated and it is used in the selection process in order to spread the solutions along the Pareto front.

In the present work two approaches were followed to optimize the GCL using MOGA:

(1)   Non-overlapping GCL (NoGCL) approach: Here the length of a chromosome is twice the number obtained dividing the length of the area of interest by the lower bound. The first half genes of a chromosome are allocated as lengths of the grid-cell units (GCUs) sequentially in a row one after another from left to right. While the second half genes are for the heights of the corresponding grid-cell units. In this approach MOGA tries to select GCUs design which is replicated

column-wise to cover the whole area of interest. For the example shown in Fig. 2, the NoGCL approach will use a chromosome length of 60 genes for 10m-by-10m square GCL. Using MOGA operators the GAFP searches for the optimal GCL which consists of different rectangular GCUs along a row while he same GCU is used along a column.

Figure 2. Allocation of a chromosome for overlapping GCL approach
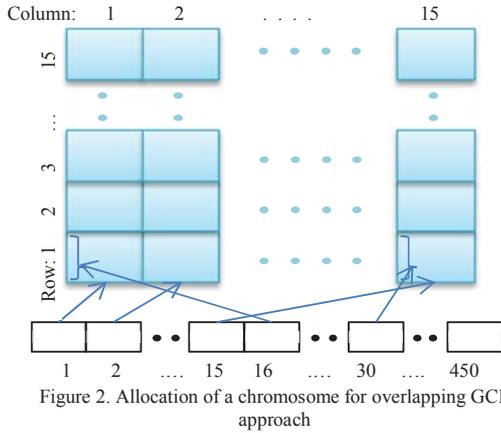
(2) Overlapping GCL (OGCL) approach: In the second approach the total gene number of a chromosome equals the multiplications of the numbers obtained dividing the length and height of the area of interest by the lower bound. The chromosome allocation used in the present research is depicted in Fig. 2 where the area of interest spans over 300m-by-300m and only 20m-by-20m square GCUs constructs the GCL. So in this case there are 15 GCUs along the rows as well as 15 GCUs along the columns. As we can see from the Fig.2 a chromosome comprises of 450 genes. In each of the chromosomes 30 consecutive genes starting from the left of the chromosome string is allocated to the lowest row of GCUs as indicated by number 1 in the figure, the next 30 consecutive genes belongs to the 2nd row and so on. In each of these sets of genes the first 15 genes are allocated to the lengths of GCUs starting from left side and rest of the 15 genes are allocated to the heights of the corresponding GCUs. In this approach, we have fixed the base-position of each of the GCUs according to the lower bound of the GCU, which is 20m along the columns heights. So, only for the 20m-by-20m square GCU there will be no overlapping, otherwise for all GCL overlapping GCUs will cover the area of interest. Hence for any chromosome GCU layout begins from row 1 according to the first 30 consecutive genes and then the next 30 genes are allocated to row 2 where for each of the GCU the lower base is 20m above the base of row1. Hence there will be fixed number of overlapping GCUs along the columns, except the 20m-by20m GCL. The OGCL approach is computationally more expensive as compared to that of the NoGCL approach.

Initial population was generated using real valued chromosome genes. For the NoGCL approach the lower and upper bounds of the GCU were 10m and 30m respectively. For the OGCL approach we have selected the bounds to be

20m and 40m respectively in order to reduce the chromosome length. In both the approaches initial population contains GCL having square GCUs ranging between the lower and upper bounds covering the whole area of interest while the rest of the chromosomes comprise of random real valued genes within the lower and upper bounds of the respective approaches. It is worth mentioning that for creating the training signatures only those GUCs are selected which fall inside the area of interest, since for various chromosome structures some GCL will cover far bigger area as compared to our area of interest.

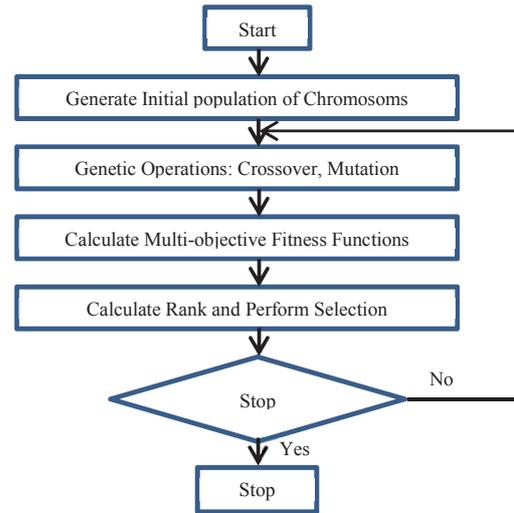The MOGA operation is given by the following flowchart:

Figure 3. A flowchart of MOGA

The fitness function of the proposed GAFP method is given below:
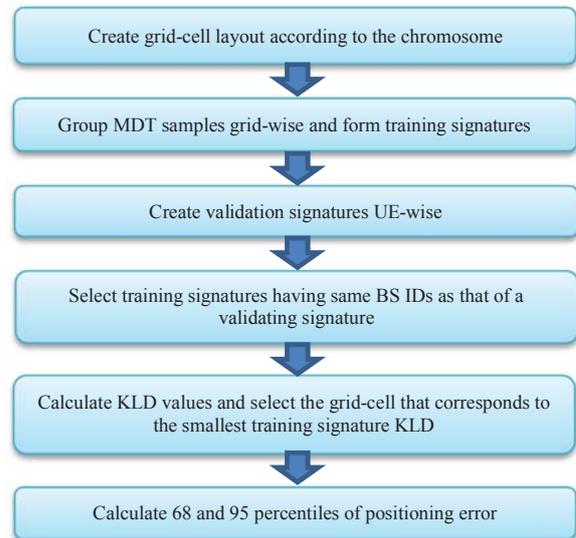

Figure 4. A block diagram representation of GAFP fitness function

We have divided the total number of MDT samples obtained from the area of interest into three sets: (i) training data, (ii) validation data and (iii) test data. As we can see from Fig.4, GAFP uses the training data to create training signatures and the validation data to calculate the fitness values of the chromosomes in different generations. Finally the test data is used to evaluate the performance of the GAFP optimized GCL.

GAFP delivers no single optimum solution but a series of equally optimal ones. In order to select a single solution we have adopted the approach of the evolution line as a posteriori preference articulation technique that uses weights to prioritize objectives after the Pareto front has been obtained [19]. The evaluation line is drawn in the objective space using a weighted function that reflects the preferences towards the objectives. The advantage of the function is that it leans the evaluation line towards the preferred objectives. Here Pareto-optimal solution is evaluated based on its point-line distance from the evaluation line. This criterion overcomes the issue of non-convex areas of the Pareto front, as each solution is evaluated based on its closeness to the evaluation line. In our simulations equal weight was given to the objective functions to select the optimal solution while evaluating the GAFP positioning accuracy.

IV. PERFORMANCE EVALUATION

A. Simulation Scenario

In this study we have performed test simulations in heterogeneous small cell (HSC) network scenario as illustrated in Fig. 5. It was 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 TR 36.839 specifications, defining parameterization for used bandwidth, center frequency, path-loss, slow fading, and fast fading 0. 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 Ês/Iot cell detection criterion as in 0.
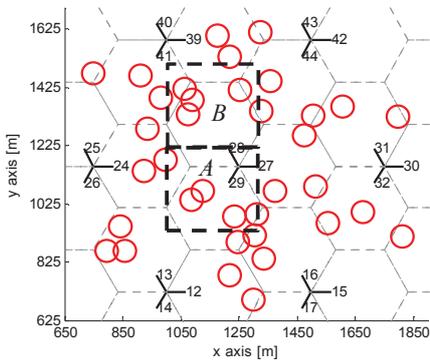

Figure 5. Heterogeneous small cell scenario

The HSC network consisted of 57 macro BSs having inter-site distance of 500m and operating on CC0 band. In addition, 36 small cells with omni-directional antennas were randomly deployed in the coverage area of 12 centermost macro BSs as depicted in Fig. 5 with red circles. Distance to the nearest inter-site small cell varied from 50m to 170m with average distance being 95m. Here, UEs were moving only in the area of 12 centermost macro BSs but were able to monitor all detected cells.

In order to check the robustness of our proposed method the total area of interest which is 300m-by-600m has been equally divided into two squares, each of 300m-by-300m. In Fig 5 these square areas are indicated by alphabetic letters- A and B with dotted black lines showing their boundaries. Within area-A there were 36184 MDT samples and within area-B the number of MDT samples are 36375. In total, 1623 randomly moving outdoor vehicular UEs (30 km/h) were distributed uniformly to the simulation area, where 100% resource block loading was used for creating interference limited simulation environment which is more challenging from cell detection point of view. Traffic profile consisted of data generated by MDT reports which were sent once per second. More details about the used simulation parameters can be found in [11].

B. MOGA Parameters

The parameters used in the MOGA are summerized in Table I.

TABLE I. MOGA PARAMETERS USED IN SIMULATIONS

Parameters	Type/Value
Selection type	Tournament
Crossover type and	Scattered
Crossover fraction	0.8
Mutation function	Constraint dependent
Fitness functions	68 percentile value of PE, 95 percentile value of PE
Chromosome length	60 and 450 for Non-overlapping and Overlapping approaches respectively
Population size	60 and 100
Stopping criteria	200 generations and spread of Pareto solutions less than tolerance: 0.0001
Pareto fraction	0.35

### C. Simulation Results

Robustness Evaluation: For this purpose we have chosen two different areas- A and B with different training, validation and testing datasets in the simulations. In both of the simulations only 10% training data were used to form the training signatures, 30% MDT samples were used for validation purpose of the GAFP fitness evaluation and rest of 60% data for testing the optimized GCL. Here we have used the computationally less expensive NoGCL approach with a GA population size of 50; it has taken 127 and 132 GA generations for area A and B respectively. A comparison study between CGCL and NoGCL methods positioning performance is shown in Table II. Here we have results from three different CGCL grid-cell layouts: 30m-by30m, 20m-by-20m and 10m-by-10m following the conventional grid-based RF fingerprinting. We have chosen these CGCL grid-cell layouts

as the grid-cell size parameter used in NoGCL simulations range between 30m to 10m. The Pareto front and average distance between chromosomes are shown in Fig. 6(a) and 6(b) respectively, obtained from NoGCL simulation in area *B*. We can see from the 68%-ile and 95%-ile results in Table II that the GAFP NoGCL approach gives less PE as compared to CGCL in both the areas *A* and *B*. In area B the NoGCL approach shows an improvement of 10.5% in the 95%-ile PE over that of the 10m-by-10m CGCL method. The analyzed test MDT sample percentage is bit higher in area *A* and bit lower in area *B* for the NoGCL than that of the CGCL method. The third row in Table II shows the performance results for the whole area of interest, i.e. combined area *A* and *B*. Here we have used the optimized GCL solutions from the NoGCL method simulated separately in areas *A* and *B*. Here also the NoGCL has better positioning accuracies in both 68%-ile and 95%-ile as compared to the CGCL. All these results show the effectiveness of the proposed GAFP method despite changes in cellular network structure.

Next we have evaluated the GAFP positioning accuracy through increasing amount of training data: 50% data for training signature creation, 25% for validation and 25% for testing. In this case area *A* was chosen for simulation and both NoGCL and OGCL approaches were used. In both methods the GA population size was 100 and number of GA generations was 200. The results are shown in Table III; here we have three different CGCL simulations: 40m-by-40m, 30m-by-30m and 20m-by-20m. For both the NoGCL and OGCL approaches the grid-cell size ranges between 40m to

20m. Here we can find that the with CGCL method the lowest PE values in 68%-ile and 95%-ile are 44.03 and 127.18 meters whereas for the GAFP NoGCL approach we have much reduced PE values of 24.52 and 55.01 meters respectively. Hence the proposed method shows an excellent improvement in positioning accuracy of 44.31% and 56.74% in the 68%-ile and 95%-ile as compared to the conventional grid-based RF fingerprinting.
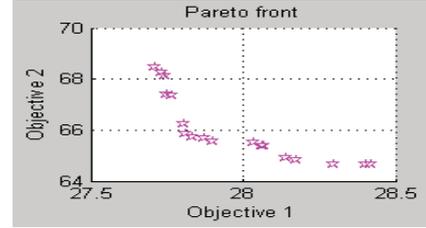
(a)
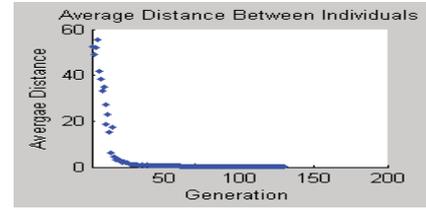
(b)

Figure 6. (a) Pareto front obtained from NoGCL method (b) Average distance between chromosomes

TABLE II. POSITIONING ACCURACY PERFORMANCE EVALUATION OF CGCL AND GAFP NoGCL METHODS WITH 10% TRAINING DATA

Area of Interest	Positioning Error 68 %-ile (m)				Positioning Error 95 %-ile (m)				Analyzed Test MDT Samples (%)			
	CGCL			GAFP	CGCL			GAFP	CGCL			GAFP
	30 m	20 m	10 m	NoGCL	30 m	20 m	10 m	NoGCL	30 m	20 m	10 m	NoGCL
AREA A	51.53	49.21	47.27	44.96	131.64	128.47	130.58	115.75	57.99	57.33	55.95	59.94
AREA B	40.12	34.93	33.02	30.73	93.29	85.21	76.32	65.85	52.84	56.16	52.60	52.82
AREA AB	45.99	40.05	42.13	38.07	115.02	109.32	110.27	98.49	55.83	54.63	57.06	56.84

TABLE III. PERFORMANCE EVALUATION OF CGCL, GAFP NoGCL AND GAFP OGCL METHODS WITH 50% TRAINING DATA

GAFP Type	Positioning Error 68 %-ile (m)				Positioning Error 95 %-ile (m)				Analyzed Test MDT Samples (%)			
	CGCL			GAFP	CGCL			GAFP	CGCL			GAFP
	40 m	30 m	20 m		40 m	30 m	20 m		40 m	30 m	20 m	
NoGCL	48.25	47.58	44.03	24.52	127.18	152.70	135.96	55.01	63.07	62.40	64.53	65.32
OGCL	48.25	47.58	44.03	26.37	127.18	152.70	135.96	58.13	63.07	62.40	64.53	71.77

It can be seen from Table III that positioning accuracies shown by GAFP OGCL is bit less than the NoGCL but it has analyzed 6 more percentage of test samples as compared NoGCL method. The histogram of the chosen chromosome from OGCL method in shown in Fig. 7. As we can see the widths and heights of the grid-cell units are mostly chosen to be in-between 21 to 27 meters.
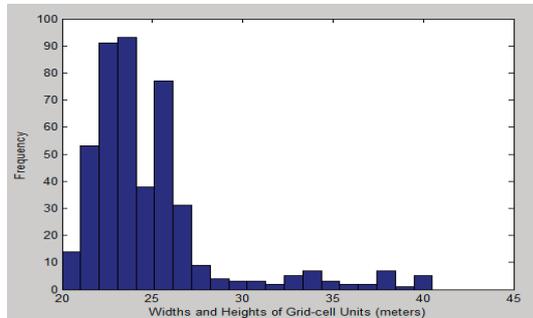

Figure 7. Histogram of the selected chromosome from GAFP OGCL method simulated in area *A*.

V. CONCLUSION

In this paper we propose a novel method to improve the positioning accuracy of grid-based RF fingerprinting through autonomous calibration of grid-cell layout. For this purpose multi-objective GA is used for selecting the best possible grid-cell layout over the area of interest so that RF fingerprinting gives the optimal output. Comparison study with conventional grid-based RF fingerprinting shows that GA optimized RF fingerprinting has the ability to calibrate the grid-cell layout in an optimal way for improved RF fingerprinting despite the change in the cellular network. It was found that with sufficient amount of training data the proposed method is able provide very good positioning accuracy as compared to that of the conventional RF fingerprinting.

REFERENCES

[1] Catherine Mulligan, "EPC and 4G Packet Networks: Driving the Mobile Broadband Revolution," ISBN: 978-0-12-394595-2, Academic Press, 2013.
[2] Ben-Moshe B., E. Elkin, H. Levi, A. Weissman, 'Improving Accuracy of GNSS Devices in Urban Canyons', *In Proc. of 23rd Canadian Conference on Computational Geometry (CCCG'11)*, August 10-12, 2011.
[3] D. Milioris et al., "Low-dimensional signal-strength fingerprint-based positioning in wireless LANs," Ad Hoc Networks, 2012, doi:10.1016/j.adhoc.2011.12.006.
[4] Chan E. C. L., G. Baciu, S.C. Mak, "Wireless Tracking Analysis in Location Fingerprinting," in *Proc. of IEEE International Conference on Wireless & Mobile Computing, Networking & Communication,* October 12-14, 2008.
[5] P. Bahl, V.N Padmanabhan, "Radar: an in-building rf-based user location and tracking system," in *Proc. of Joint Conference of the IEEE Computer and Communications Societies*, vol. 2, March 2000, pp. 775 –784.
[6] Zhu J., G. D. Durgin, "Indoor/Outdoor Location of Cellular Handsets Based on Received Signal Strength," in *Proc. of Vehicular Technology Conference*, May 30 - 1 June, 2005.
[7] Laitinen H., J. Lahteenmaki, T. Nordstrom, "Database Correlation Method for GSM Location," in *Proc. of Vehicular Technology Conference*, May 6-9, 2001.
[8] Wuri A. et al., "Minimization of Drive Tests Solution in 3GPP," *in IEEE Communications Magazine*, Vol. 50, No. 6, June 2012.
[9] 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.
[10] 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.
[11] 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.
[12] R. Mondal, J. Turkka and T. Ristaniemi, "An Efficient Grid-based RF Fingerprint Positioning Algorithm for User Location Estimation in Heterogeneous Small Cell Networks," in *Proc. of International Conference on Localization and GNSS (ICL-GNSS)*, Helsinki, Finland, June 24-26, 2014.
[13] Campos R. S. and L. Lovisolo, "Mobile Station Location using Genetic Algorithm Optimized Radio Frequency Fngerprinting," in *Proc. of International Telecommunications Symposium*, Manaus, Brazil, September 2010.
[14] M. J. Magro and C. J. Debono, "A Genetic Algorithm Approach to User Location Estimation in UMTS Networks," in *EUROCON - The International Conference on Computer as a Tool*, Warsaw, Poland, pp. 1136 – 1139, September 2007 .
[15] J. Holland, "Adaptation in Natural and Artificial Systems," University of Michigan Press, 1975.
[16] Deb, Kalyanmoy, "Multi-Objective Optimization using Evolutionary Algorithms," John Wiley & Sons, Ltd, Chichester, England, 2001.
[17] 3GPP TR 36.814, "Further advancements for E-UTRA physical layer aspects," version 9.0.0, March 2009, available at http://www.3gpp.org.
[18] 3GPP TS 36.133, "Requirements for support of radio resource management," version 11.4.0, March 2013, available at http://www.3gpp.org.
[19] Vergidis. T., Vergidis. K., Tiwari. A., "The evaluation line: A posteriori preference articulation approach," *IEEE Congress on Evolutionary Computation*, pp. 2694 - 2700, 2008.