An Improved Cellular positioning technique based on Database Correlation

B D S Lakmali¹, S A D Dias²
Department of Electronic & Telecommunication Engineering,
University of Moratuwa.
¹shashika, ²dileeka@ent.mrt.ac.lk

ABSTRACT
The paper proposes an improved network-based positioning technique based on database correlation and presents the application of the proposed technique to GSM. The fingerprint database is created using field measurements and a novel and a more practical method for gathering fingerprints is introduced. The authors propose a Weighted k-Nearest Neighbor (WkNN) method for location estimation and the trial results obtained in urban and suburban environments in Sri Lanka are presented. Comparison with the Geometrical method shows that the proposed method is a competitive alternative for GSM location.

1.0 INTRODUCTION
Positioning of Mobile Stations (MSs) is a key problem in cellular networks. This is the problem of determining the position of a MS using location sensitive parameters. Inexpensive but accurate positioning systems are of great commercial potential due to the emergence of a multitude of location-based services involving a wide range of civilian as well as military applications. These have great potential in areas such as emergency services, personal security, navigation, tourism and entertainment. Providing location-sensitive services such as ATM machines, hospitals etc. are the other areas where the knowledge of user position is important. Due to above mentioned and thousand more reasons there’s no doubt in the importance of achieving more accuracy in location estimation.

Currently the Global Positioning System (GPS) is a widely used solution for positioning. Even though GPS is fairly accurate [1] there are some negative issues like poor indoor coverage, poor coverage in bad weather conditions or shaded areas and the necessity of being equipped with a GPS receiver. On the other hand, cellular networks have provided excellent coverage in both outdoor and indoor environments and the cellular phone is becoming an everyday commodity among general public. These issues encourage the development of accurate cellular positioning techniques.

Cellular positioning techniques can be categorized into Network based methods and Terminal based methods. In network based methods, tracking and evaluation of the user location is carried out at the network side and in terminal based methods the user device performs the location estimation itself. Some of these methods may or may not require modifications to be done at the network side or the terminal side.

The paper presents the research work carried out to come up with a network based Improved Cellular Positioning Technique, which can give the coordinates of a mobile station location to a satisfying accuracy. After comprehensively studying the currently used positioning methods, the authors have focused on developing a positioning technique based on database correlation. The performance of the developed technique has been verified, in urban and sub urban environments through field test trials. The rest of the paper is organized as follows. In section 2, the proposed DCM based positioning technique is described and section 3 presents the encouraging trial results obtained. Finally, section 4 concludes and suggests possible directions for future research.

2.0 CELLULAR POSITIONING WITH DATABASE CORRELATION

The database correlation method (DCM) has been found in previous studies to have promising performance [2,3] in cellular positioning. In Database Correlation, location-dependent network parameters are stored in the form of fingerprints in a database constructed using field measurements or radio wave propagation prediction tools. To estimate the location, a measured fingerprint from a MS is correlated with those in the database and the best matching database fingerprints are used for location estimation. In general, this involves three steps:
1. Obtaining fingerprints
2. Database preparation
3. Location estimation

2.1 Obtaining Fingerprints and Database Preparation

The stored signal information samples in the database are called ‘fingerprints’. Depending on the particular cellular system, the signal fingerprints could include signal strength, signal time delay, or even channel impulse response. Also, it is possible to use measurements performed by the network as well as by the MS [2].
For the proposed method, the Received Signal Strength (RSS) at the MS was selected as the fingerprint variable and the fingerprints were constructed using an extensive set of measurements. Thus, a single fingerprint in the database consists of:

- GPS coordinates of a location
- Received signal strength (from serving base station and other neighboring base stations) in that location

Although a fingerprint collecting process is illustrated in [2], the authors propose a novel and a more practical method. Several drive tests were carried out and measurements were taken continuously along roads using an integrated GSM/GPS mobile measurement unit developed by the authors.

A single measurement constitutes the GPS coordinate of the point at which the measurement is taken and the RSS from all the hearable cells in that location.

It was expected to obtain signal strengths from six cells including the serving cell and five neighboring cells. Thus, the authors first used a single measurement to form a fingerprint as proposed in [2]. However, it was observed that the number of measurable neighboring cells is less than five in most cases in the network environment tested. Also the RSS varies significantly over time at any location. Fig 1 illustrates the RSS variation over time at a particular location for all the cells appeared in that location.

Also it was observed that the set of cells that contribute to consecutive measurements, along a road, were not always the same especially in urban areas. Fig 2 illustrates the RSS variation of 10 consecutive measurements taken along a road for all the cells appeared in those 10 measurements.
According to Fig 1 it is evident that a single measurement cannot be used to characterize the RSS at a location. But if fingerprints are created using number of stationary measurements at several locations along a road, it will not characterize the RSS in between the fingerprint locations. Because according to Fig 2 it is evident that the RSS characteristics vary significantly along a road. Therefore it was decided to take several consecutive measurements (along a road) together to form a single fingerprint. In order to have a sufficient contribution from most of the hearable cells, ten consecutive measurements were taken to create a single fingerprint by averaging the received signal strengths (for each hearable cell). The location of the fingerprint was considered as the median location of the corresponding ten points.

This reduces the fingerprint resolution from a factor of 10 when compared with that of [2]. Thus a ‘Sliding window’ approach is proposed to reduce the distance between two consecutive fingerprints. For the implementation of the sliding window approach, only five overlapping measurements were used in consecutive fingerprints. Fig 3 illustrates this sliding window approach of creating fingerprints. It was evident from the results that this approach resulted in better accuracies than using separate sets of measurements for each fingerprint. The fingerprint resolution can be further increased by increasing the number of overlapping measurements in the fingerprints. Also the drive test was performed at a very low speed (less than 20km/h) to get a higher resolution and it was possible to obtain a fingerprint resolution less than 100m.

2.2 Location Estimation

For location estimation, the RSS of the location is to be measured and given to the DCM algorithm as the input fingerprint. Due to the fact that only one measurement cannot be used to characterize the RSS at a particular location, several measurements should be obtained at the test location and the averaged RSS of each hearable cell should be calculated and input to the DCM algorithm.

The effect of the number of measurements taken at the required location was studied and observed that better performance can be obtained with higher number of measurements. Accordingly it was proposed to use 10 measurements at the location to be estimated.

The authors examined the performance of both the Nearest Neighbor (NN) and the Weighted k Nearest Neighbor (WkNN) methods for location estimation and proposed a location estimation technique based on WkNN approach.

The paper introduces two approaches for identifying the nearest neighbors. The first approach (approach-I) is based on the distance in signal space. The implemented DCM algorithm use the original concept of Least Mean Square (LMS) approach published in [2]. But as in (1), the authors introduce a new cost function based on Manhattan distance along with a new approach for the penalty term.

\[
d(k) = \sum_i (f_i - g_i(k)) + \sum_j (f_j - g_j(k)) \times w_j + \sum_{i,j} (l_{\text{max}} - g_i(k)) \times w_{ij}
\]

Where,
- \( d(k) \) = The difference between the input fingerprint and the \( k^{\text{th}} \) database fingerprint.
- \( f_i \) = the averaged RSS of the \( i^{\text{th}} \) hearable cell in the input fingerprint
- \( g_i(k) \) = the averaged RSS of the same cell in the \( k^{\text{th}} \) database fingerprint
- \( f_j \) = the averaged RSS of the \( j^{\text{th}} \) hearable cell in the input fingerprint which is not hearable in the \( k^{\text{th}} \) database fingerprint
- \( g_j(k) \) = the averaged RSS of the \( k^{\text{th}} \) hearable cell in the database fingerprint which is not hearable in the input fingerprint
- \( l_{\text{max}} \) = replace the missing signal strength value
- \( w_j, w_k \) = corresponding weights of the penalty terms

The weights of the penalty terms were calculated based on the contribution of the corresponding cell for the 10 measurements.

The summation is taken over the cells that are found in both fingerprints. The database fingerprints with lower \( d(k) \) values are considered as the Nearest neighbors of the location to be estimated.

The second approach (approach-II) is based on defining a valid RSS range for each cell in each fingerprint using the averaged RSS and the standard deviation of the RSS of the corresponding cell. Here the difference is calculated using (2).

\[
c(k) = \sum_i \delta(f_i, g_i(k), \text{std}_i(k))
\]
Where,
\( c(k) \) = the total number of cells in the input fingerprint which falls within the valid RSS range of corresponding cells in the \( k^{th} \) database fingerprint.
\( f_i \) = the averaged RSS of the \( i^{th} \) hearable cell in the input fingerprint
\( g_i(k) \) = the averaged RSS of the same cell in the \( k^{th} \) database fingerprint
\( \text{std}_i(k) \) = the standard deviation of the same cell in the \( k^{th} \) database fingerprint
\( \delta (a,b,c) = \) a function which returns 1, if \( a \) falls within the region \( b-c \) and \( b+c \). Otherwise it returns 0.

The summation is taken over the cells that are common in both fingerprints. The database fingerprints with higher \( c(k) \) values are considered as the Nearest neighbors of the location to be estimated.

After identifying the nearest neighbors using one of the above two approaches, the location is estimated using the Weighted k-Nearest neighbors (WkNN) method as below.

\[
p = \sum_i p_i w_i
\]

\( p \) = estimated location
\( p_i \) = location of the \( i^{th} \) nearest neighbor
\( w_i \) = weight of the \( i^{th} \) nearest neighbor

The results were tested with three different weights. They are:

I. \( w_i = c(i) / \sum_i c(i) \)

II. \( w_i = \left( \frac{1}{d(i)} \right) / \sum_i \frac{1}{d(i)} \)

III. \( w_i = \left( \frac{c(i)}{d(i)} \right) / \sum_i \frac{c(i)}{d(i)} \)

Also the performance of the WkNN estimation was studied with the geographical K-Mean Clustering algorithm with K=2. However, a significant improvement of the performance could not be observed with clustering.

As proposed in [1] the number of searches was limited based on the serving cell of the input fingerprint and each database fingerprint. But this criterion failed for some locations in the urban environment. Thus the authors came up with a new solution to find out the cell or cells which have the highest occurrence from all the hearable cells and to treat this cell(s) also as a serving cell(s) when extracting the database fingerprints.

### 3.0 RESULTS & DISCUSSION

Field trials were performed in two environments: urban environment (around Bambalapitiya, in the heart of Colombo, Sri Lanka,) and a sub urban environment (area around University of Moratuwa). Encouraging results were obtained for the tested environments.

Best results were obtained by identifying the nearest neighbors using approach-I. For urban environment best results were obtained with weight I and for sub urban environment best results were obtained with weight III.

The performance of the WkNN estimation was studied with different number of neighbors (i.e. \( k \)) and best results were obtained with \( k=3 \) and \( k=4 \) for urban and sub urban environments respectively. Performance of the WkNN estimation degrades with higher number of neighbors due to the affect of distant neighbors for location estimation.

Fig. 4 and Fig. 5 show the CDF-wise comparison of the error for the best results obtained for urban and sub urban areas respectively. The results of the trial were compared with the accuracies given by the Geometrical method, in which the location is estimated by calculating the distance between the MS and BS for at least three MSs using Hata model and applying trilateration [5,6].

For urban environments, in 80% of the estimates, the error is less than 150m and for sub urban environment it is less than 350m. The 80% error with the Geometrical method is 400m and 1250m in urban and sub urban environments respectively. Although DCM is less accurate in sub urban environment it shows excellent performance when compared with the Geometrical method.

Table 1 and Table 2 summarize the results of the trials in urban and sub urban areas respectively.

<table>
<thead>
<tr>
<th></th>
<th>DCM</th>
<th>Geometrical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Error (m)</td>
<td>87</td>
<td>220</td>
</tr>
<tr>
<td>80 % Error (m)</td>
<td>150</td>
<td>400</td>
</tr>
<tr>
<td>Average Error (m)</td>
<td>108</td>
<td>284</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>93</td>
<td>263</td>
</tr>
</tbody>
</table>

Table 1: Summary of results for Urban environment

<table>
<thead>
<tr>
<th></th>
<th>DCM</th>
<th>Geometrical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Error (m)</td>
<td>177</td>
<td>787</td>
</tr>
<tr>
<td>80 % Error (m)</td>
<td>350</td>
<td>1250</td>
</tr>
<tr>
<td>Average Error (m)</td>
<td>240</td>
<td>1032</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>207</td>
<td>1458</td>
</tr>
</tbody>
</table>

Table 2: Summary of results for Sub urban environment
4.0 CONCLUSION
The proposed DCM based cellular positioning technique shows improved positioning accuracy for both urban and sub urban environments. This would be a better solution for cellular positioning not only in urban environments but also for less densely built environments where other positioning techniques such as geometrical method would fail to provide acceptable accuracy. Being a network based method which requires no modification at the network side or the terminal side is an added advantage of the proposed method.

5.0 ACKNOWLEDGEMENT
The authors are grateful to Dialog Telekom Ltd. for their sponsorship of this research.

6.0 REFERENCES
[4] Binghao Li, Yufei Wang, Hyung Keun Lee, Andrew Dempster,Chris Rizos, “ A New Method of Yielding a Database of Location Fingerprints in WLAN”, School of