Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 6, June 2014, pg.502 – 509

RESEARCH ARTICLE

Dynamic Clustering Approach to Compress Radio Database for Efficient Localization in Fingerprinting Systems

S.Y. Patil[1], S.M. Chaware[2]
[1]ME Computer, BSCOER, Pune, India
[2]Professor, BSCOER, Pune, India
[1] sheetal.patil77@gmail.com; [2] smchaware@gmail.com

Abstract*—Radio fingerprinting is the process which identifies a cellular phone / other radio transmitter by the unique "fingerprint" which characterizes the signal transmission. Location fingerprinting determines receiver position with the help of database of RSS measurements from different sources .Location fingerprinting a positioning method that explores existing infrastructures e.g. cellular networks or WLANs. Dynamic Clustering technique to compress the radio database in the context of cellular fingerprinting systems has been proposed. The aim of the dynamic clustering technique is to reduce transmission load and computation cost in the mobile-based implementations. The presented method may be called Dynamic Block-based Weighted Clustering (DBWC) technique, which overcomes the issue of BWC that is it requires fix number of cluster which is not feasible in dynamic environment. DBWC is applied in a location-radio signal space, and assigns different weight factors to the location and radio components. The main idea of this work is to propose a dynamic clustering technique well-tailored to the structure of the radio database in a fingerprinting system. Here, we tend to outline the idea of feature types in association with database records. The obtained result confirms the efficiency of DBWC.

***Keywords* —Wireless systems, RSS, location fingerprinting, training phase, BWC**

I. INTRODUCTION

As the need to use wireless networks for monitoring many real time conditions is continuously increasing, such networks presents many challenges to researchers [1]. The concept of localization fingerprinting is resulted into many research problems because localization helps in many wireless network applications for estimating the current position of node in wireless networks. This estimated position of node may be relative or absolute.

In the literature there are many methods presented by various researchers with aim of energy efficiency for the location fingerprinting in wireless networks. Some methods for the node localization are based on global positioning system means GPS. But later many researchers addressed limitations over this approach for estimating the node location. While outdoor localization is almost exclusively performed using the Global Positioning System (GPS), indoor location systems have successfully employed a variety of technologies. However this method can also result in inefficient energy consumption.

Location fingerprinting exploits the existing infrastructures like cellular networks or WLANs. The demand for energy efficient networks and the issues like green networking, clustering technique has been proposed to compress the radio database The aim of the proposed system is to reduce the transmission load and computation cost in the mobile-based implementations. The presented method is called as Dynamic Block -based Weighted Clustering (DBWC) technique, applied in a

Location-radio signal space, which assigns different weight factors to the location and radio components. The obtained results confirm the efficiency of the DBWC technique. [1] An ultimate aim of mobile positioning research is to find a method providing high estimation accuracy to the user within minimum delay and at minimum cost.

To estimate the location of mobile stations (MS) in cellular networks, signal information obtained by an MS is stored in the form of fingerprints in training phase. Difference between the measurement taken at location phase, and the database fingerprints is calculated and Database fingerprints with minimum difference are selected as the nearest fingerprints for the location to be estimated.

A clustering technique is proposed to compress the radio database in the context of mobile based fingerprinting systems. The goal is to minimize the computation cost and to improve the terminal autonomy, with an acceptable positioning accuracy w.r.t. a non -compressed database. To do this, a Block-based Weighted Clustering method is proposed. The BWC technique is applied in a concatenated radio-signal and location space, and assigns different weight factors to the location and radio components. According to the results, we see that the BWC technique improves the performance of standard clustering methods introduced in [12]. The achieved positioning accuracy is shown to be competitive to that of a non-compressed database. Clustering technique outperforms other existing compression methods, while it requires a notably lower computational cost and transmission load. Issue which remains as an open area here is the choice of an optimum number of clusters M. All the clustering algorithms in this work need to fix the number of clusters in advance.

We propose Dynamic Block-based Weighted Clustering (DBWC) to overcome above stated issue. Our work is unique in introducing dynamic clustering of radio map locations as an approach to reduce the computational requirements of the location determination techniques and increase the scalability of the system.

II. BACKGROUND

a) Location Fingerprinting:-

Location Fingerprinting works in two phases, First training phase, in which radio map is constructed over the area of interest. During the localization phase, by matching mobile terminals received signal to the radio map entries a mobile terminal is localized. The stored signal information in the radio map may be the signal time delay, Received Signal Strength (RSS), channel impulse response or any other location-dependent information.

b) Database Compression :-

The size of the radio database is an influential factor in regards to issues such as transmission loads and computation. Suggested approach for database compression is to reduce the dimension of the radio feature space [8], [9], and [10].

c) Block –Based Weighted Clustering (BWC) :-

Block-based Weighted Clustering (BWC) scheme imposes equal weights to blocks of Components belonging to the same feature type, in the clustering cost function; these assigned weight factors are optimized during the clustering process. The BWC technique improves the performance of standard clustering methods mentioned in [12].The BWC outperforms other existing compression methods, As it requires a notably lower transmission load and computational cost.

D) Drawbacks of Previous Work

Issue which remains as an open area here is the choice of an optimum number of clusters M. All the clustering algorithms in this work need to fix the number of clusters in advance .which also may not be feasible for dynamic environment.

III. PROBLEM DEFINATION

There are different methods for location fingerprinting based on different techniques with different aims and goals to achieve, however on the other side they having their limitations. In recent time new method was presented in [1], efficient method in which clustering technique is proposed to compress the radio database in the context of mobile based fingerprinting systems. The main goal of this method was to reduce the computation cost and hence to improve the terminal autonomy, with acceptable positioning accuracy with respect to a non-compressed database. But this method is having some limitations such as clustering algorithm used here not guarantee to find the global minimum of the Objective function. Another limitation is the choice of number of clusters. Currently all clustering algorithms requires fixed number of clusters in advance which may not be efficient technique for dynamic wireless networks.

IV. PROPOSED ARCHIETECTURE AND DESIGN

In this project firstly we are presenting clustering technique for radio database compression, which takes into account location and the radio components of the recorded measurements for investigation. The main idea of this work is to propose a clustering algorithm well-tailored to the structure of the radio database in a fingerprinting system. We define the concept of feature types association with database records. All the stored parameters in a record that belong to the same nature is defined as feature types. Based on this, a Block-based Weighted Clustering (BWC) technique is proposed, which imposes equal weights to blocks of components belonging to the same feature type, in the clustering cost function; In this project we are presenting new clustering method which is dynamically works for radio database compression and location fingerprinting. Dynamic clustering approach is presented method to improve the performance of investigating method. Fig. 1 shows the proposed architecture.

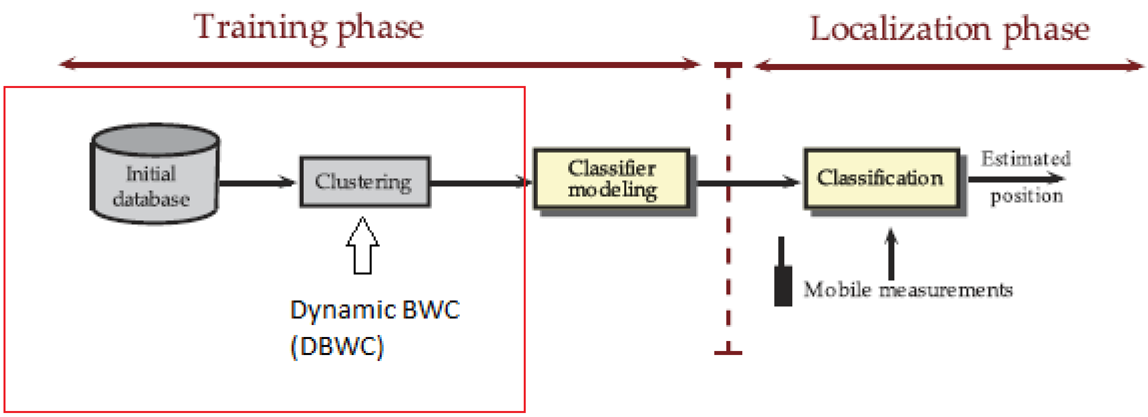

Fig 1**.** Architecture of the fingerprinting system, with database clustering (DBWC).

Concept and Notations

Radio database R is a set of records. A record is a vector r =[x; s] where X belongs to geographical position and S belongs to measurement vector in radio space, which describes radio signal at location X. feature type is defined as all the stored parameters in a record that belong to the same nature. Two different feature types in each record: Location feature type and a single Radio Access Technology (RAT) RSS feature. By applying a clustering algorithm during the training phase, the initial database could be divided into M (M <N) subsets. N is set of data points.

Clustered algorithm:
Input: {Records Containing X, Y location and RSS measurement}

Step 1: Estimating no. of clusters
Assume the data have been clustered via any technique, such as k-means, into k clusters.
For each datum i, let $a(i)$ be the average dissimilarity of i with all other data within the same cluster.
Any measure of dissimilarity can be used but distance measures are the most common.
We can interpret $a(i)$ as how well i is assigned to its cluster (the smaller the value, the better the assignment). We then define the average dissimilarity of point i to a cluster c as the average of the distance from i to points in c.
Let $b(i)$ be the lowest average dissimilarity of i to any other cluster which i is not a member. The cluster with this lowest average dissimilarity is said to be the "neighbouring cluster" of i because it is the next best fit cluster for point i. We now define:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

Which can be written as:

$$s(i) = \begin{cases} 1 - a(i)/b(i), & \text{if } a(i) < b(i) \\ 0, & \text{if } a(i) = b(i) \\ b(i)/a(i) - 1, & \text{if } a(i) > b(i) \end{cases}$$

From the above definition it is clear that

$$-1 \le s(i) \le 1$$

For $s(i)$ to be close to 1 we require $a(i) \ll b(i)$. As $a(i)$ is a measure of how dissimilar i is to its own cluster, a small value means it is well matched. Furthermore, a large $b(i)$ implies that i is badly matched to its neighbouring cluster. Thus an $s(i)$ close to one means that the datum is appropriately clustered. If $s(i)$ is close to negative one, then by the same logic we see that i would be more appropriate if it was clustered in its neighbouring cluster. An $s(i)$ near zero means that the datum is on the border of two natural clusters.

The average $s(i)$ over all data of a cluster is a measure of how tightly grouped all the data in the cluster are. Thus the average $s(i)$ over all data of the entire dataset is a measure of how appropriately the data has been clustered. If there are too many or too few clusters, as may occur when a poor choice of k is used in the k-means algorithm, some of the clusters will typically display much narrower silhouettes than the rest. Thus silhouette plots and averages may be used to determine the natural number of clusters within a dataset.

```
Step 2: foreach (document di in document collection dc), d Ɛ  dc
        If(di is in list)
                Break;
         For(document dj=i+1)
                Foreach (word wi in di), wi Ɛ  di
                        If (wi.equal(wj Ɛ  dj))
                        Count++;
                End foreach
                If count > Threshold value
                        Form cluster of (di, di+1)
                        Store clusters in dic{di, di+1}
                        Add to list(di, di+1)
                End if
        End for loop;
End foreach
```

Output: {Dynamic clustered data}

The above algorithm is works same for BWC but the only difference is number of clusters that is the value of M is fix in BWC where as in DBWC is dynamic depend on location of access points and RSS fields mentioned in initial database. DBWC considers the added and removed access points during localization phase. A clustering technique can reduce notably the transmission load and computation cost in location fingerprinting systems.

BWC and DBWC technique: here, the clustered database is transferred to the terminal. Simply, we have

bTot = M(DGbG + DRbR)

Where
bG is the number of bits required to code a single geographical coordinate
bR is the average number of bits required to code a single radio parameter
bTot is the total number of bits required to code the database.
M is number of clusters, For BWC it’s fixed and for DBWC it’s dynamic.

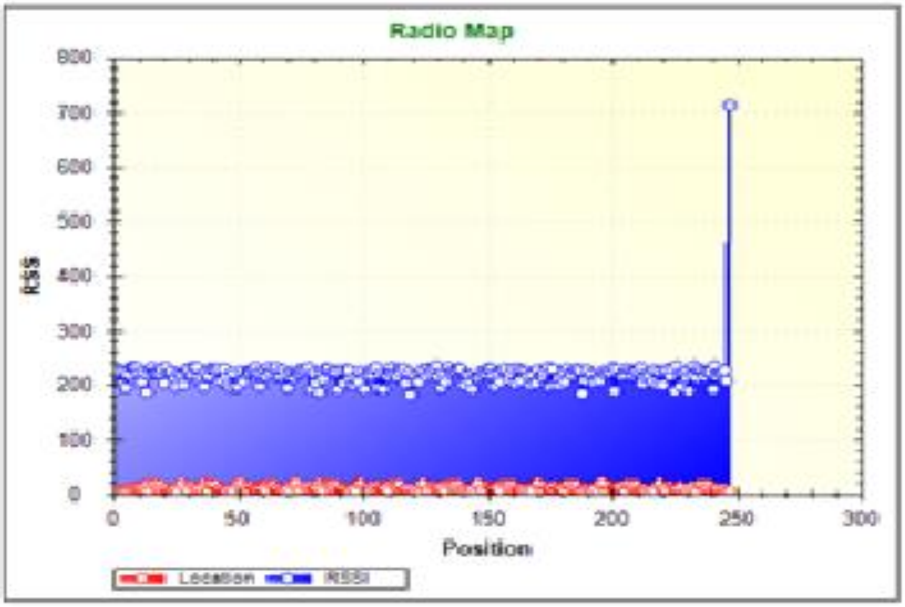

Fig.2 shows Radio Map .

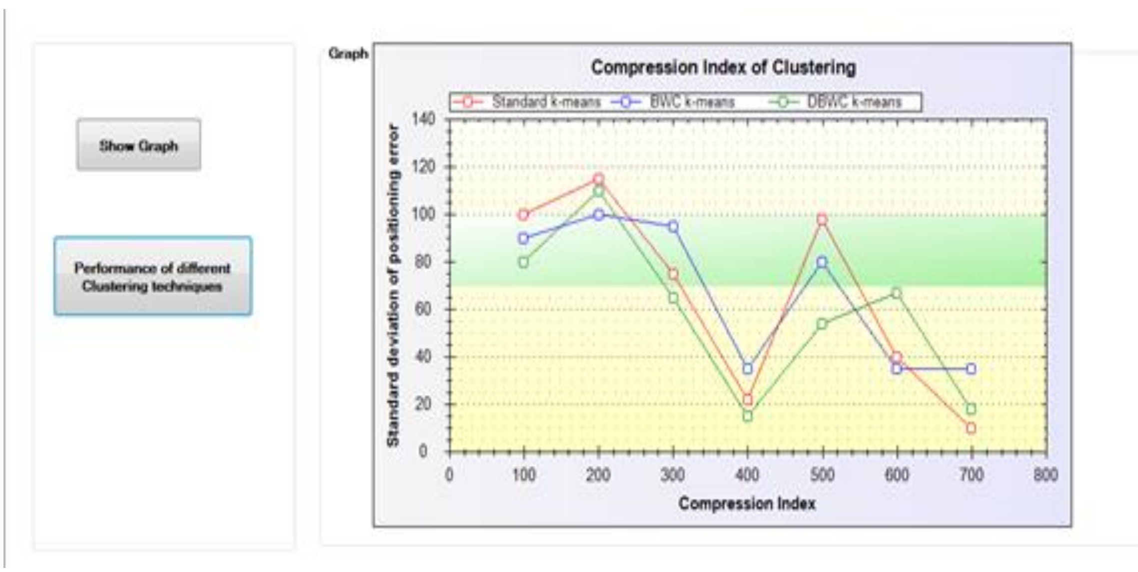

fig 3 Shows compression index for non-compressed data , compressed data with BWC and DBWC

V. CONCLUSION AND FUTURE WORKS

In this paper, a clustering technique is proposed to compress the radio database in the context of mobile based fingerprinting systems. The goal is to minimize the computation cost and improve the terminal autonomy, with an acceptable positioning accuracy w.r.t. a non-compressed database. To do this, a Dynamic Block-based Weighted Clustering (DBWC) method has been proposed. The obtained results confirm the efficiency of the DBWC technique, and it shows the improvement in the performance of standard k-means and hierarchical clustering methods and BWC. The achieved positioning accuracy is shown to be competitive to that of a non-compressed database. Moreover, the proposed clustering technique outperforms other existing compression methods, while it requires a lower transmission load and computational cost. In the future, we will extend the proposed clustering technique to apply it in applications, as radio network planning.

REFERENCES

[1] Azin Arya, Philippe Godlewski, Marine Campedel, and Ghislain du Chene, "Radio Database Compression for Accurate Energy-Efficient Localization in Fingerprinting Systems", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 6, JUNE 2013.

[2] K. Pahlavan, Wireless Information Networks. Wiley, 2005.

[3] A. Kupper, Location Based Services. Wiley, 2005.

[4] C. Takenga and K. Kyamakya, "A Low-Cost Fingerprint Positioning System in Cellular Networks," Proc. Int'l Conf. Comm. And Networking, pp. 915-920, Aug. 2007.

[5] C. Dessiniotis, "Motive Project Deliverable 2.1," technical report,European Project FP6-IST 27659, Sept. 2006.

[6] P. Bahl and V.N. Padmanabhan, "Radar: An In-Building RF-Based User Location and Tracking System," Proc. IEEE INFOCOM '00,vol. 2, pp. 775-784, Mar. 2000.

[7] B.D.S. Lakmali, W.H.M.P. Wijesinghe, and K.U.M. De Silva,"Design, Implementation and Testing of Positioning Techniques in Mobile Networks," Proc. Int'l Conf. Information and Automation for Sustainability, pp. 94-99, Dec. 2007.

[8] S. Fang, T. Lin, and P. Lin, "Location Fingerprinting in a Decorrelated Space," IEEE Trans. Knowledge and Data Eng., vol.20, no. 5, pp. 685-691, May 2008.

[9] Y. Chen, J. Yin, X. Chai, and Q. Yang, "Power-Efficient Access-Point Selection for Indoor Location Estimation," IEEE Trans.Knowledge and Data Eng., vol. 18, no. 7, pp. 877-888, July 2006.

[10] M.A. Youssef, A. Agrawala, and A. Udaya Shankar, "WLAN Location Determination via Clustering and Probability Distributions," Proc. Conf. Pervasive Computing and Comm., pp.
143-150, Mar. 2003.

[11] A. Arya, P. Godlewski, and P. Melle´, "Performance Analysis of Outdoor Localization Systems Based on RSS Fingerprinting," Proc.Int'l Symp. Wireless Comm. Systems, pp. 378-382, Sept. 2009.

[12] A. Arya, P. Godlewski, and P. Melle´, "A Hierarchical Clustering Technique for Radio Map Compression in Location Fingerprinting Systems," Proc. Int'l Conf. Vehicular Technology, pp. 1-5, May 2010.

[13] D. Zimmermann, J. Baumann, M. Layh, F.M. Landstorfer, and R.Hoppe, "Database Correlation for Positioning of Mobile Terminals in Cellular Networks Using Wave Propagation Models," Proc.Conf. Vehicular Technology, vol. 7, pp. 4682-4686, Sept. 2004.

[14] M. Khalaf-Allah and K. Kyamakya, "Database Correlation Using Bayes Filter for Mobile Terminal Localization in GSM Suburban Environments," Proc. Conf. Vehicular Technology, vol. 2, pp. 798-802, May 2006.

[15] M. Brunato and R. Battiti, "Statistical Learning Theory for Location Fingerprinting in Wireless LANs," J. Computer Networks, vol. 47, no. 6, pp. 825-845, 2005.

[16] C. Takenga and K. Kyamakya, "A Hybrid Neural Network-Data Base Correlation Positioning in GSM Network," Proc. IEEE Int'l Conf. Comm. Systems, pp. 1-5, Oct. 2006.

[17] Z. Wu, C. Li, J. Kee-Yin Ng, and K.R.P.H. Leung, "Location Estimation via Support Vector Regression," IEEE Trans. Mobile Computing, vol. 6, no. 3, pp. 311-321, Mar. 2007.

[18] J. Junfeng Pan, J.T. Kwok, Q. Yang, and Y. Chen,"Multidimensional Vector Regression for Accurate and Low-Cost Location Estimation in Pervasive Computing," IEEE Trans.
Knowledge and Data Eng., vol. 18, no. 9, pp. 1181-1193, Sept.2006.