

# Cell Oscillation Resolution in Mobility Profile Building

Shafqat Ali Shad<sup>1</sup>, Enhong Chen<sup>2</sup>, Tengfei Bao<sup>2</sup>

<sup>1</sup> Laboratory of Semantic Computing and Data Mining,  
Department of Computer Science and Technology  
University of Science and Technology of China  
Huangshan Road, Hefei, 230027 Anhui, China

<sup>2</sup> Laboratory of Semantic Computing and Data Mining,  
Department of Computer Science and Technology  
University of Science and Technology of China  
Huangshan Road, Hefei, 230027 Anhui, China

## Abstract

Mobility profile building became extensively examined area in Location based services (LBS) through extraction of significant locations. Mobility traces are recorded under three reference positioning systems that are Satellite based i.e. GPS, Network based i.e. GSM and Local positioning i.e. WLAN, RFID, IrDA. Satellite based and local positioning due to of high power consumption, additional resource installation, low accuracy and space limitation are less encouraging. So network based positioning i.e. GSM is only viable solution for mobility tracing through Cell global identity (CGI). CGI presents the Cell-ids to extract the significant locations from mobility history. However CGI faces cell oscillation problem, where user is assigned multiple Cell-Ids even at a stationary state for load balancing and GSM cells overlapping. In this paper we proposed two semi-supervised methodology for cell oscillation resolution i.e. semantic tagging and overlapped area clustering, the proposed methodologies are equally useful for the identification of significant places too.

**Keywords:** *GSM Cell Oscillation, Mobility profile mining, Spatial extraction, Trajectory mining.*

## 1. Introduction

Location based services require position mapping for the extraction of trends and significant locations from the user mobility history at mobile terminals, where position is determined through some reference system. This reference system can be coordinate based, address based, areal division or route based. So this reference system in term of LBS is divided into three major types [1] i.e. Satellite based positioning, Network based positioning and local positioning.

Broadly speaking satellite and network based positioning systems are meant to be outdoor systems, while local position system is indoor [2, 3, 4]. Satellite positioning system uses the earth-orbiting satellite system and receivers. Most commonly used satellite positioning is GPS which provide accuracy at meter and sub-level

depending on its configuration and methods used. As for mobile users GPS is discouraging in term of high power consumption and extra equipment installation in network. This makes its use limited for positioning or location extraction. While in case of Local positioning systems i.e. Bluetooth, RFID, Infrared, Wifi, IrDA they operate in limited space area due to of its short range signal capabilities. So only viable solution for position extraction widely spread and examined is Network position system i.e. GSM. While in case of GSM, location exaction can be done through Assisted-GPS (A-GPS), Time difference of arrival (TDOA) and Enhanced observed time difference (EOTD) but all of these require additional resource installation in network [5] which makes Cell Global Identity (CGI) most inexpensive, readily available and suitable location extraction method. Cell Global Identity (CGI) is four header information set i.e. Mobile country code (MCC) assigned uniquely to each country, Mobile network code (MNC) assigned uniquely to each operator in specific country, Location area code (LAC) created by operator for identification & management of specific geographical area and Cell id (Cell ID) assigned uniquely to every user connected over network. So MCC, MNC, LAC and Cell ID together uniquely identify the location of every user over a certain period of time in network.

Over the years many work is carried out regarding the mobility building and its potential applications which cater fall from route prediction to early warning systems because of its spatially rich nature e.g. advertisement [6, 7], early warning systems [8], city wide sensing [9], pollution exposure, route marking and tracking [10], traffic management, social networking and community finding [11], where the mobility date is interpreted to meaningful information by extracting the significant places. For the identification of significant places the distance and time constraint is used along with the discrete location modeling. As mentioned before CGI can be used quite evidently for location extraction, where CGI header are converted into latitude and longitude pair using publically

available Cell Id database e.g. Google API for location representation. However mining significant places is a trivial task as significant locations are meant to be the locations where user stays for a considerable period of time during mobility building. But GSM network face a critical problem of Cell oscillation where user is assigned different Cell Ids even its stationary due of load balancing effect or cell hand over, so apparently in such situation user seemed to be moving which leads to wrong mobility building and significant locations extraction.

So resolution of cell oscillation effect is necessary to extract the correct mobility and significant places. Our work is mainly focused on the same; we introduced two step by step approaches to resolve it a) Location clustering through semantic tag information by users b) Clustering of the overlapped area using the basic GSM network information. However besides resolving cell oscillation effect our proposed methodology is equally useful for the extraction of significant places too later for mobility profile building.

## 2. Related work

Mobility data has become hot source of information regarding human habit mining and location sensing, where mobility profile building is the main task in all potential applications which mainly focused on extraction of significant places. Location awareness is highlighted in the work of Lee, et. al. [11] where a request is made to the LBS server to find out the context through location agent. While sequential mining approach has been used for location acquisition [12] for the resource allocation in the network. And a guide system is proposed which combines the positioning technique and location-awareness service to provide the surrounding information for users. While for the resolution of positioning Zhuang et.al [13] proposed the signal power approach for the prediction of next cell during the mobility. While Chao et al. [14] carried out the mobility estimation through the time factor over a cell. Shen et al [15] introduced the fuzzy logic approach using the pilot signals and random user movements. While in their work [16] localization information is inferred through the Markov model using the GPS data. Soh et al [17] presented the topographic information usage technique to infer the mobility and extraction of locations. The mentioned work is limited due to of probabilistic model usage where mobile user behavior is totally ignored. While in [18], authors proposed the neural network based approach where vicinity and edge of area is considered for the user movement, where cell area is defined using the coordinate system. Connection traces are periodically studied in their work [19] for the creation of historical record for location extraction and next visit

prediction. This work is not viable in real time large dataset because it need long run for learning and training phase. While in [20] authors presented the mobility finding approach by introducing the clues like position over a particular geographical location and time of stay, later this information is processed over current cell for next visit prediction. This work has main disadvantage that it needs lot of prior knowledge about geographical information and so called clues to be incorporated over mobile device for prediction.

Above mentioned work involve full scale information of network topology and shape which is hidden in most of cases from the public access by the network operator because of security reasons. And most of the work is only emphasized on the mobility analysis without any significant work done for the resolution of cell oscillation which is vital for mobility building and most of these work involve GPS sensors rather GSM CGI information. In our work we presented the approach which is applicable for publically available mobility dataset where network topology information is hidden from the user and lacked movement factors like speed, direction, proximity and geocoding.

In our work we considered GSM CGI information of publically available dataset i.e. MIT Mobility dataset [21] where semantic tag information can play important role for resolution of cell oscillation because each significant location is overlapped with multiples cells. Further according to our work a location which is not semantically tagged by user can be defined significant location if multiple cells are assigned over a period of time within a close area and clustered together using the basic GSM information i.e. Location Area Code and frequency of cell appearance throughout mobility . So in our work we first find out semantically tagged locations and its associated cells to resolve oscillation, and later consider the untagged location through clustering of different cells over a location area defined in GSM network where overlapping area define significant location by introducing the stay point distance and time threshold factor.

## 3. Methodology

### 3.1. Usage of Semantically tagged locations

To resolve the cell oscillation we must consider the basic architecture of any GSM network. Base Transceiver Station (BTS) is the basic units of any cellular network, where each BTS creates its own cell or area where multiple Cell IDs fall in [22].As shown in Fig 1. Mobile Station (MS) is connected with the Base Station Controller (BSC) and Mobile Services Switching Center (MSC).Different

MSCs and BSCs are connected with MSC to provide communication channel between two MS registered with different cells.

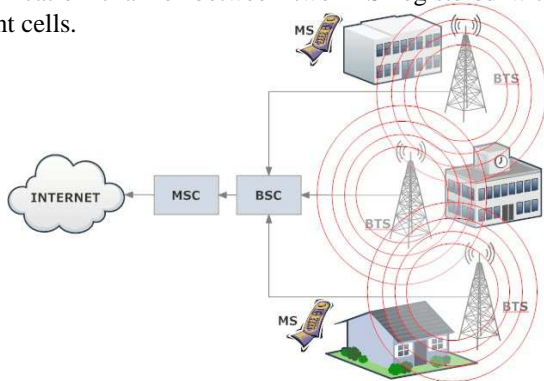


Fig. 1 GSM network basic architecture.

As BTS determines the number of cells in GSM network and Location area code (LAC) is defined by one or more MSCs. So multiple BTS connected with particular MSC share same area code. This LAC is defined by the network operator depending on user density, network architecture and services. This means in city area LAC contains less number of cells to cater fall the requirement of high density users while in country side these LAC contains much high number of cells as user density is significantly low. The mobility data collected carries the semantic tagging e.g. Home, Office, School etc, this information can be used for the resolution of cell oscillation where one semantically tagged location is identified by multiple overlapped cells or one cell can highlight more than one semantically tagged locations as shown in Fig 2.

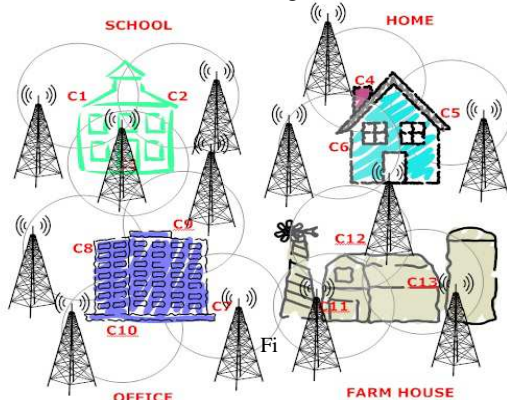


Fig. 2 Semantically tagged locations.

As discussed previously in the GSM network architecture there exist many overlapping cells for handover and handoff tasks so it is obvious that one semantic location can be tagged by the user with different cell ids over different visits and in same way one cell can appear at multiple nearby significant locations. By assuming this situation we have proposed the clustering technique where

cells appearing over one semantically tagged location clustered together for resolution of oscillation phenomena. We have developed a semantic based clustering algorithm to resolve this problem. Let us define the set of all locations to be  $L$  and set of all cells to be  $C$ . We define two mappings  $\emptyset$  and  $\Psi$ .  $\emptyset: L \rightarrow C$  such that  $\emptyset(l_k) = \{c_i \in C \mid c_i \text{ contains } l_k\}$  where  $l_k \in L$  whereas  $\Psi: C \rightarrow L$  such that  $\Psi(c_j) = \{l_p \in L \mid l_p \text{ is in } c_j\}$ . Now we choose a location say  $l_q$  and we calculate  $\emptyset(l_q) = \{c_{iq}\}$ . Next for every  $c_{iq}$ , we calculate  $\Psi(c_{iq}) = \{l_{iq}\}$ . we sort  $c_{iq}$ 's according to  $|\Psi(c_j)|$  and then store this sorted data against  $l_q$ . we also store  $|\emptyset(l_q)|$  against  $l_q$ . The pseudo code of the algorithms is as:

**Algorithm 1:** Semantic based location clustering

Select the complete list of the semantically tagged locations from the subject data

Treat each semantic location as a single cluster and repeat step 1-5 for each semantic location

1. Select all the Cell IDs associated with the selected semantic location
2. Compute the frequency of each Cell ID as per its appearance in the semantically tagged data
3. Sort the Cell IDs in descending order on the basis of their frequency
4. Compute the number of Cell IDs associated with the semantic location
5. Mark the semantic location as a single cluster which uniquely identifies all associated Cell IDs.

The presented technique can resolve the problem of cell oscillation as the cells appearing at one semantically tagged location can be clustered together, so during the mobility profile building by introducing the time constraint over the sequence of oscillation pairs we can determine whether the oscillation pairs fall in same cluster or not, if they fall in same cluster than it is quite obvious that the user is stationary and cell switching is only because of load balancing or handover effect.

Further the proposed technique is quite evident of calculating the significance of a particular location during mobility building alongside significance calculation of every cell visited by user. As shown in step 2 we are calculating the frequency of the each cell appeared in all semantically tagged locations and at step 4 we are assigning the frequency to a semantic location on the basis of number of cells observed on that particular location over user mobility history. This is very important to compute two frequencies for example if user is a student and stays most of time in school where most of the semantically tagged locations are geographically nearby each other e.g. Lab, Dormitory, Dining hall, Library, Play ground etc, later these frequencies can determine the significance

parameter for location or cell level extraction during user mobility building.

### 3.2. GSM network and Cell Oscillation resolution

As described in the previous section, apparently the GSM network is considered to be of hexagonal shape for the simplification of topology [23, 24, 25] as shown in Fig 3 (a) where radio cells are concatenated. But in reality this is not the case these cells are overlapped and of polygon spread over the network as shown in Fig 2 (b), where distribution is entirely dependent on the network operator strategy defined in term of mobile user density, load balancing and services.

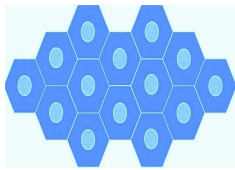


Fig. 3(a) Hexagonal shaped.

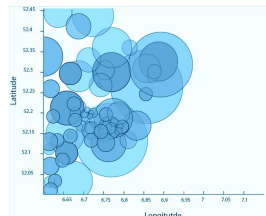


Fig. 3(b) Polygonal shaped cells.

Because of the overlapping and polygonal shape cell switching is obvious phenomena in GSM network. There is four kind of cell switching usually known as Handoff or Handover i.e. channel switching within same cell, switching under same BSC, Switching under same MSC and Switching under different MSCs [26]. Where first two handovers are internal because they involve BSC only while other two are known as external due to of MSC involvement. While first one effect in no change in Cell ID but other three results in change of Cell ID. So these three are most important for proximity detection and user actual movement.

There are two kinds of algorithms operating for these handovers i.e. a) Minimum acceptable performance b) Power budget, these algorithms enable MSC and BSC to let MS get connected to different Cells. But these algorithms result into cell oscillation where user is assigned different Cell ids over a time even the user is in stationary mode because of change in signal quality like atmosphere changes, people interference in the signals; or other mobile phones enter into the same cell or other input parameters of the handover algorithms change [27].

In our work we are interested in clustering the Cells through usage of basic GSM network properties where overlapping area determines the significant location. As overlapping means stationary state of the user at significant location due to of high switching occurrence among particular cells residing in same cluster and distance constraint is not violated for transition shown in Fig 4.

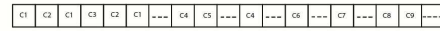
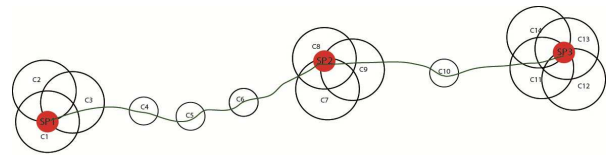


Fig. 4 Significant places with overlapped area.

On the basis of GSM network topology and user mobility trend we proposed the overlapping cell area determination algorithm to resolve the cell oscillation where user is stationary even assigned different cell id over a period of time. Let C be the sequence of Cell ID trajectories where  $C=c_1, c_2, c_3, \dots, c_n$  along with their location area codes. Define a set of untagged locations cluster L as an empty set, where  $L=l_1, l_2, l_3, \dots, l_n$  and each l is resultant sequence of overlapping area of multiple cells in vicinity under same location area code as  $l=c_1, c_2, c_3, \dots, c_n$ . We initialize the first sequence with first cell observed in user mobility history, we firstly check if the cell belongs to the cluster generated through the semantic tagging algorithm or not if it belongs to any of the cluster then we assign it there and move to the next cell. We repeat these steps until the entire C are consumed up. For every unassigned cell we calculate the distance with its adjacent cell alongside calculating the sum of radius of both to determine the actual movement of user. For every  $c \in l$  we calculate the distance between  $c_i$  and  $c_{i+1}$  using their latitude and longitude properties and compare this distance with sum of radius of both cells. If the distance is lower than the sum of radius of two cells this shows user is under overlapped area and there is no actual movement. Where distance is calculated through Eq. (1).

$$\text{Distance}(c_i, c_{i+1}) = R \times \arcsin$$

$$\sqrt{\sin^2(\Delta \text{Lat}/2) + \cos(c_i \text{Lat}) \cos(c_{i+1} \text{Lat}) \sin^2(\Delta \text{Long}/2)} \quad (1)$$

Where R donates the radius of equator and  $\Delta \text{Lat} = (c_i \text{Latitude} - c_{i+1} \text{Latitude})$ ,  $\Delta \text{Long} = (c_i \text{Long} - c_{i+1} \text{Long})$ . After determining the stationary position of user we find out if the user stay time i.e.  $T_{\min} = c_{i+1} \text{Time of arrival} - c_i \text{Time of leaving}$  is less than the minimum stay time threshold set for stay point, if it is true then we find out  $c_i \text{Location area code} = c_{i+1} \text{Location area code}$ , if it is true this ensures that user is under overlapping area and these cells can be placed under one cluster of sub sequence, we append the cells together under same sub sequence. We repeat it until distance criteria become false. After extracting the complete list of overlapped cells for the sub sequence we assign it to the set of cluster L as an individual cluster  $l_n$ , where each  $l_n$  represents the overlapped area or significant place for mobility building with complete resolution of cell oscillation effect.

**Algorithm 2:** Overlapped area based cell oscillation resolution algorithm

Select the complete list of the user movement sequences in form cell IDs and location area codes

Initialize the untagged location cluster set as an empty set, which will later replace the cell ids in user history for mobility profile building.

Set the minimum stay time threshold

Set the initial cluster with the first cell observed in the user mobility history and repeat step 1-7 until all the cell id in user mobility are consumed

1. Find if the cell belongs to any of the cluster generated through Algorithm 1 defined over semantic tagging, if it is true assign the cell to a semantically tagged location using the weight property defined over it where cell belongs to multiple locations.
2. If cell does not belong to any semantically tagged cluster defined in Algorithm 1 then repeat step 3-7.
3. Determine the user movement by calculating the distance between the current cell and adjacent cell in the mobility sequence along with the sum of radius of both cells.
4. If the distance between the two cells is less than sum of their radius this means the subject is under overlapped area and there is no actual movement. This result shows that multiple cell assignment is only due to of ping pong effect of cell oscillation in GSM network.
5. Determine if stay time is less than threshold time, find out whether the two cells belong to same location area code or not and if they belong to same location area code then append the cells in the sub sequence for clustering together.
6. Define sub sequence as a single cluster with overlapped area as an effect of cell oscillation
7. Assign the cluster to the untagged location cluster set, which can be used as determination of significant places later during mobility profile building (As our current task is only related to cell oscillation resolution so significant places role in mobility profile building is beyond the scope of this paper)

The proposed methodology can resolve the cell oscillation problem precisely through determining the overlapped area and presenting the sub sequence detail. This can be used for precise mobility profile building later; however mobility profile building is beyond scope of this paper so we discussed it briefly here.

#### 4. Dataset

As mentioned earlier we are using the reality mining dataset collected by MIT Media labs. The data is collected by 100 subjects using the Nokia 6600 mobile phone sets with installed application that can record context and

potential information between cell tower transitions. The data is collected over 9 months and contain an activity span of 350K Hours with the size of about 1GB. As the data contains only the partial CGI information i.e. LAC, Cell ID headers with partial semantic tag information, extraction of user mobility profile is trivial task for it.

We used Google location API for the extraction of Longitude and Latitude information of corresponding Cell towers for location extraction and mapping using CGI information. Google API provides the largest cell Id database and it is equally effective for partial CGI information of MIT data.

#### 5. Experiments and results

In first step we retrieved the location information using the CGI headers available against the subject mobility data using Google API using a batch tool. The unique locations visited by subject is located in **S(n).all\_locs** tables in Reality mining dataset, for the selected subject there are 1744 unique locations logged. However the data collected is old and due to of rapid change in GSM network over time only 768 locations are recovered. However after removal of spatial outliers from data we finally retrieved 698 locations for mobility building process.

In next step we implemented our semantic tagging based algorithm on the subject data, in reality mining dataset the semantic information is located in **S(n).cellnames** table. There are total of 75 unique locations tagged by the user **X**. For user **X** after implementing the semantic based algorithm we found out that most of the locations are identified by more than one cell, this clearly satisfies our assumption related to cell oscillation phenomena. For example “Home” and “Office” are identified by 4 cells each, which is evident that even if the user is at stationary mode at semantic location, it can be assigned multiple cells over time. So these cells can be clustered together and later during mobility building can be used for cell oscillation problem. As shown in Fig 5.

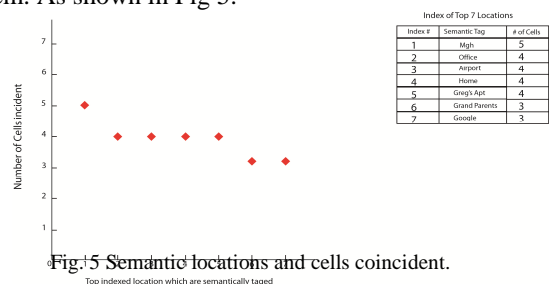


Fig.5 Semantic locations and cells coincident.

While as mentioned earlier one cell can be incident for multiple semantic locations where the mobility is across nearby places, for example consider a scenario of mobile

user life who is a student and live most of time within the University boundaries where semantically tagged locations are in vicinity of each other i.e. Lab, Dormitory, Dining hall, Play ground etc. Fig 6 shows the same where one cell is incident for multiple semantic locations in user mobility history, an average of 2 locations identified by one cell so we can also assign it a weight for voting in the mobility building.

Fig. 6 Frequency of each unique cell.

For our further analysis we numbered the semantic locations in ascending order then we draw a graph for the distribution of user mobility over these semantic locations with respect to time period spent over them. We took the time span of two months and plotted a scattered graph against average time spent by user at each location as shown in Fig 7. As shown in the graph the sites 2 and 4 are the places where the user X spent most of its time, the sites 2 and 4 are infact "Home" and "Office". While average spends time at majority of location is between 1 to 2 Hours per day.

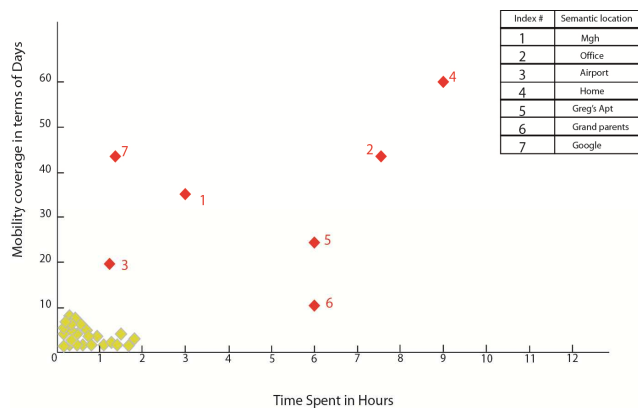


Fig. 7 Significant places appearance.

For the implementation of algorithm described in section 3.2 we need to define a time threshold for minimum stay time for cell oscillation resolution to determine the actual user movement, therefore we plotted the duration of likely time over different locations by choosing the sample data of about 3 months from user X mobility history as graph shown in Fig 8. This data is stored in **S(n).locs** table.

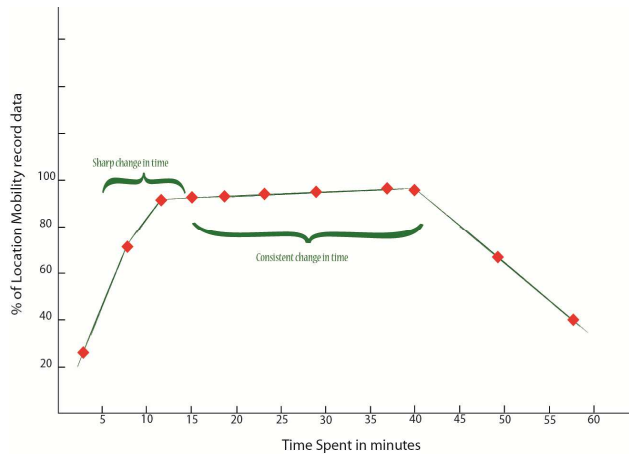


Fig. 8 Time threshold.

We divided the stay time into equal intervals of 5 minute span starting from 1 to 60 minutes. As the graph shows the stay time period with respect to location grows with time but at reaching the time of 6 minute there is sharp until 11 minute, then it becomes almost static or there is slight change until it reaches about 30 minutes, after 30 minutes the coverage starts to decline. So it is obvious to consider that critical time is 11 minute which can be set as Minimum time of stay. And the places fall under this time are about 91% of user mobility histories which make 11 minute an obvious choice for minimum stay time threshold. So it can easily be inferred that if the user is infact stationary and on some significant place he must stay there at least for 11 minutes. In another scenario if he is staying within a overlapped area then he will be assigned multiple cell ids over time where his stay time over these ids will be less then this threshold however there will be no change in distance between transitioned cells. So by determining if the distance is not significantly changed and by determining if time of stay is less then threshold at these cells, it becomes obvious that a cell oscillation phenomenon is only reason in change of cell ids. These overlapped cell ids are result of handovers and handoff effects which can be clustered together for true mobility building of user.

For the illustration of effectiveness of our proposed methodology we have computed the distance matrix for the two sets of latitude; longitude values i.e. a) raw longitude; latitude values from user mobility b) Clustered latitude, longitude values retrieved through proposed methodology to check the similarity of covered area. Distance matrix for matrices X (longitude; latitude values for sample dataset from user X mobility history) and Y (longitude; latitude values for cells clustered together with overlapped area) is list of coords where the (i,j)th entry is the distance between the i'th and j'th row of X and Y respectively. Then we have plotted the distance matrix over its

coordinates in Km scale Fig 9, as a result a straight line is retrieved which makes it clear that the area shape covered by the both sets of coordinates is similar and has no difference. So we can rely on the set of values retrieved through the clustering.

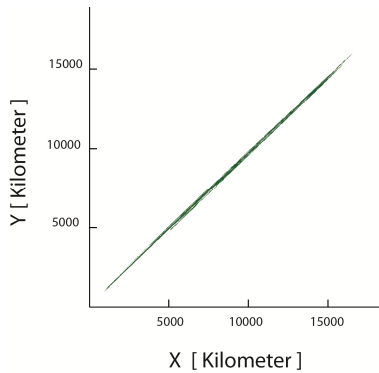


Fig. 9 Distance matrix plot

In Fig 10. We plotted the sample data of user's one day mobility history (as on 18<sup>th</sup> March, 2005) by considering the sequence of cell Ids and location information retrieved from Google API, this figure shows that the proposed methodology successfully resolved the cell oscillation because user defined semantic locations fall on line constructed through clustered cells, where clustered cells are result of cell oscillation and overlapped area. The gray line shows the trajectory of the user and diamonds shows the stay points. As shown clearly the diamond with red color are places user already tagged semantically and yellow diamond are places identified as a result of overlapping of cells.

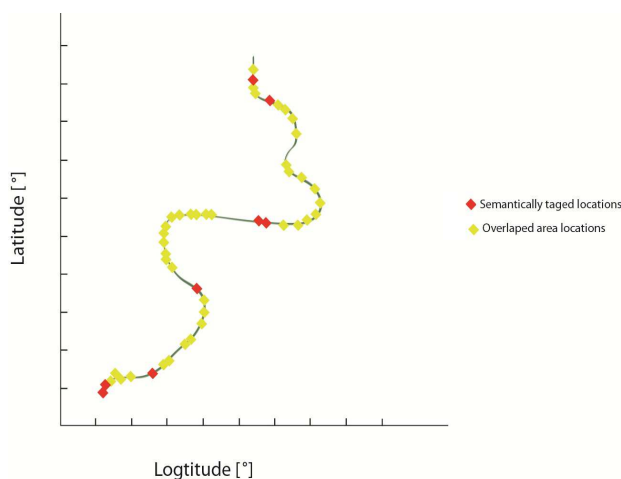


Fig. 10 User movement after clustering on

## 6. Conclusion and future work

In this paper our work mainly focused on the cell oscillation resolution which is essential in mobility building and is not previously addressed significantly. As if the cell oscillation is not resolved the mobility profile building may lead to inappropriate results. Our proposed methodology is not only significant for resolution of cell oscillation but beside this it is evidently useful for significant place identification as well. As true mobility building is basis of Location based services, the proposed methodology can be adopted for its precise delivery.

In future work we will identify the significant places from the user mobility and build up a mobility profile, where profile building will be done by using the Markov chain model for its time slicing over time and day. Later this information can be used for inferring similar users based on their mobility profile similarity measure. We are interested in devising the efficient mobility profile building model with proper resolution of mobility session pruning and similarity measure.

## Acknowledgments

The work described in this paper was supported by grants from Natural Science Foundation of China (Grant No. 60775037), The National Major Special Science & Technology Projects (Grant No. 2011ZX04016-071), The HeGaoJi National Major Special Science & Technology Projects (Grant No. 2012ZX01029001-002) and Research Fund for the Doctoral Program of Higher Education of China (20093402110017, 20113402110024).

## References

- [1] Paikannussanasto, "Vocabulary of Positioning," *Tekniikan sanastokeskus, TSK 30, Helsinki*, 2002.
- [2] Harter and Hopper, "A distributed location system for the active office," *IEEE Network*, vol. 8, pp. 62-70, 1994.
- [3] A. Harter, *et al.*, "The Anatomy of a Context-Aware Application," *Wireless Networks*, vol. 8, pp. 187-197, 2002.
- [4] "IEEE Computer Society LAN/MAN Standards Committee " *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-speed Physical Layer in the 5 GHz Band, The Institute of Electrical and Electronic Engineers, New York*, 1999.
- [5] "Location Technologies for GSM, GPRS and WCDMA Networks," *White Paper, SnapTrack, A QUALCOMM Company*, 2001.
- [6] M. Couceiro, *et al.*, "Data Stream Processing on Real-time Mobile Advertisement: Ericsson Research

- Approach," *In IEEE International Conference on Mobile Data Management*, vol. 1, pp. 313 - 320 2011.
- [7] Ad Orchestrator: <http://prodcat.ericsson.se/frontend/category.action?code=FGD%20101%>.
- [8] C. Yang, *et al.*, "Use of Mobile Phones in an Emergency Reporting System for Infectious Disease Surveillance after the Sichuan Earthquake in China," *Bulletin of the World Health Organization*, vol. 87, pp. 619-623, 2009.
- [9] S. S. Kanhere, "Participatory Sensing: Crowdsourcing Data from Mobile Smartphones in Urban Spaces" *in IEEE International Conference on Mobile Data Management on Mobile Data Management* vol. 2, pp. 3-6, 2011.
- [10] J. W. Lee, *et al.*, "Location-aware Agent using Datamining for the Distributed Location-based Services," *In LNCS*, vol. 3894, pp. 867-876 2006.
- [11] G. Yavas, *et al.*, "A Data Mining Approach for Location Prediction in Mobile Environments," *In IEEE conference on Data & Knowledge Engineering*, vol. 54, pp. 121-146, 2005.
- [12] Y. Huang and Chuang. WP, "Improving the Museum's Service by Data Mining and Location-aware Approach," *In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, vol. 13, pp. 2646-2651 2004.
- [13] W. Zhuang, *et al.*, "Measurement-based dynamic bandwidth reservation scheme for handoff in mobile multimedia networks," *In IEEE International Conference on Universal Personal Communications*, vol. 1, pp. 311 - 315, 1998.
- [14] C. Choi, H., *et al.*, "Adaptive bandwidth reservation mechanism using mobility probability in mobile multimedia computing environment," *IEEE2000*, 2000a.
- [15] X. Shen, *et al.*, "User mobility profile prediction: An adaptive fuzzy interference approach," *Wireless Network*, vol. 6, pp. 362-374, 2000.
- [16] D. Ashrook and T.Staruer, "Learning significant locations and predicting user movement with gps," *In Proceedings of the 6th International Symposium on Wearable Computers TSWC'02*, 2002.
- [17] W. Soh, Hyong, and S. Kim, "Qos provisioning in cellular networks based on mobility prediction techniques," *IEEE Communication Magazine*, pp. 86-92, 2003.
- [18] S. Liou and H. Lu, "Applied neural network for location prediction and resource reservation scheme," *In Wireless Network, Proceedings of ICCT'03*, 2003.
- [19] J. Capka and R. Boutaba, "Mobility prediction in wireless networks using neural networks," *In International Federation for Information Proceeding IFIP'04*, 2004.
- [20] Jean-Marc François and G. Leduc, "Entropy-based knowledge spreading and application to mobility prediction," *In Proceedings of conference on Emerging network experiment and technology* 2005.
- [21] "<http://reality.media.mit.edu/download.php> as on 15-04-2012."
- [22] "Digital Cellular Telecommunications System (Phase 2+), Location Services (LBS), Functional Description, Stage 2, ETSI TS 101 724 V8.8.0; <http://webapp.etsi.org/key/queryform.asp>," 2004.
- [23] S. Nandi and M. K. Raushan, "An efficient implementation of distance-based update scheme using directional cell identification codes," *In Distributed computing: IWDC*, p. 535, 2004.
- [24] M. Hassan and R. Jain, "High performance TCP/IP networking: Concepts, Issues and Solutions," *Pearson Prentice Hall, Upper Saddle River, NJ 07458*, 2004.
- [25] Ling Chen, *et al.*, "A personal route prediction system based on trajectory data mining," *Information Sciences*, vol. 181, pp. 1264-1284, 2011.
- [26] J. Schiller, "Mobile communications," *Addison-Wesley Pearson Education, Edinburgh Gate, Harlow, England*, vol. 2nd, 2003.
- [27] H.-P. Lin., *et al.*, "Validation of an improved location-based handover algorithm using gsm measurement data," *IEEE Transactions on Mobile Computing*, vol. 4, pp. 530-536, 2005.

**Shafqat Ali Shad:** Mr. Shafqat Ali Shad is second year doctoral student at USTC. He obtained his Master degree in Computer Science from COMSATS Institute of Information Technology, Pakistan with the distinction of Chancellor Gold Medal in 2004. He had worked as National ICT Consultant for Asian Development Bank for Health sector reforms program in Pakistan mainly focused on Pakistan millennium development goals in 2010. He also worked with Planning Commission, Government of Pakistan as Deputy Director for ICT policy making and implementation of Five year plans, Annual plans and Vision 2030 from 2004 to 2009. Beside this He worked as freelance ERP consultant for SAP and Oracle E-business suite implementations in large enterprises for almost 5 years. His professional certifications include Oracle 10g DBA, Juniper Certified Internet Specialist-ER and PMP. His current research interests include Mobile mining, Recommender systems, Machine learning and ERP systems. He is the author of 3 international journal publications.

**Prof. Enhong Chen:** Dr. Chen Enhong, born in July 1968, currently works as a professor and doctoral supervisor at the Laboratory of Semantic Computing and Data Mining, University of Science and Technology of China (USTC). Prof. Chen is also a Senior Member of the Institute of Electrical and Electronics Engineers (IEEE). Chen obtained his Ph.D. in Computer Software from USTC. Prof. Chen currently serves as Vice Dean of School of Computer Science and Technology of China, Deputy Director of the MOE - Microsoft Key Laboratory of Multimedia Computing and Communication, University of Science and Technology of China, Member of the Knowledge Engineering Committee and the Machine Learning Committee of the Chinese Association for Artificial Intelligence, Member of the Artificial Intelligence and Pattern Recognition Committee of the China Computer Federation. Besides, Prof. Chen also serves on



the program committees for over 20 international academic conferences. Prof. Chen has authored more than 90 research papers and invited papers published in international and domestic academic journals or submitted to international academic conferences. His paper presented to the top international conference on data mining, namely, the KDD2008, has won the Best Application Paper Award.

**Tengfei Bao:** Tengfei Bao is a research student at USTC he is author of many international journal and conference publications. He worked for Nokia Research Center Asia and made major contributions in Mobile user's habit mining research area.