# An Activity Based Mobility Prediction Strategy Using Markov Modeling for Wireless Networks

**R.V. Mathivaruni and V.Vaidehi**

*Abstract:* **The foremost objective of a wireless network is to facilitate the communication of mobile users regardless of their point of attachment to the network. The system must discern the location of the mobile terminal, to afford flawless service to the mobile terminal. Mobility prediction is widely used to assist handoff management, resource reservation and service pre-configuration. Prediction techniques that are currently used don't consider the motivation behind the movement of mobile nodes and incur huge overheads to manage and manipulate the information required to make predictions. This paper proposes an activity based mobility prediction technique that uses activity prediction and Markov modeling techniques to devise a prediction methodology that could make accurate predictions than existing techniques.**

## 1. INTRODUCTION

The fundamental requirement for the wireless networks to provide flawless service to the mobile terminal is that it must know the point of attachment of the mobile terminal (MT) at any point of time. Mobility prediction is widely employed to aid network management tasks. At network level, mobility prediction can be used to assist handoff management, resource allocation, congestion control and quality of service provisioning [1]. At the application level, it's exploited to provide Mobile Location Services. Mobility prediction is an important maneuver that determines the location of the mobile terminal by carefully manipulating the available information. The prediction accuracy depends on the user movement model and the prediction algorithm used.

Two different techniques are already available for mobility prediction. First technique uses the historical movement patterns of the user to predict the user's whereabouts in future. The second technique uses the contextual knowledge.

Pattern based technique [2] is used based on the fact that often moving objects have some degree of regularity in their motion. In regular pattern based prediction techniques, the user mobility pattern is created using the time sequence in which the locations are visited. Here, users are classified into different categories. The main drawback with pattern based technique is that the pattern period and periodicity varies from one mobile terminal to other and the locations are included in the profile, without considering the possible reasons for the visit and hence the prediction could be misleading [3]. Larger the size and more up-to-date the data is, it is more likely to make accurate predictions using pattern based techniques. But the restriction on the size of the user's history leads to loss of information, which results in misleading predictions. As the user movement pattern is highly uncertain, assumptions about the user's movement pattern should be made with utmost care. Contextual knowledge based Prediction (CKbP) is proposed to overcome these problems [4]. Contextual knowledge based technique predicts the whereabouts of the MT using knowledge of user's preferences, goals and analyzed spatial information without imposing any assumptions about the availability of users' movements history. User context and spatial conceptual maps are combined for the prediction process. But, the information gathering and evidence extraction are laborious in contextual knowledge based technique. The information required is huge and keeping information up to date is a difficult task. There is a clear need for a technique, which could help to overcome the drawbacks with current prediction techniques.

This paper proposes an Activity based Mobility Prediction algorithm using Markov modeling (AMPuMM) to predict the future location of the MTs. Normally in Pattern based and Contextual knowledge based techniques, the volume of information that needs to be gathered, stored and manipulated is huge. To overcome this problem AMPuMM uses activity patterns to make prediction in contrast to existing techniques that uses the time sequence of cells visited by a user. The transition between activities is modeled as a Markov chain [5] to predict (n+1) $^{th}$ day location using 'n' day's information. In activity based prediction technique proposed in [3] an attempt is made to identify the activity performed by the user in each visited location like home, office, shopping etc using real time traces. Such an approach would give rise to serious privacy concerns. To overcome this problem, the activities are first classified into more abstract groups. The activities are classified based on the regularity of their occurrence as regular and irregular and further, it is classified into short, medium and long based on the length of visit.

Through various surveys it had been established, that majority of user activities occur only at specific locations, for limited time periods in the case of most the users. The number and location of everyday activities that may be performed by a person is limited by the amount of time available and the space-time constraints associated with various obligatory

R.V. Mathivaruni is with Department of Electronics Engineering, Madras Institute of Technology, Chennai-600044, India (email: mathi_varuni@yahoo.co.in)

V.Vaidehi is with Department of Electronics Engineering, Madras Institute of Technology, Chennai-600044, India (email: vaidehi@annauniv.edu)

activities (e.g. work) and joint activities with others. The geographical behavior of individual are found to be stable over longer periods of time while the time-sequence of visits and repeatability has a shorter life span. Hence patterns based on the time sequence in which various locations are visited could have very short life span and accurate predictions may not be possible. Based on these facts this thesis proposes an activity based location prediction scheme that could reflect changes in user's behavior earlier than other pattern based techniques and could provide accurate predictions on MT's future locations [3].

The rest of the paper is organized as follows. Section 2 explains the activity model and the profile generation. Markov based prediction method and path finding algorithms are described in section 3. The simulation work and results are presented in section 4 followed by conclusion and future work in section 5.

## 2. Activity Based Mobility Model

With activity based mobility models [6], future location can be predicted with better accuracy. Activity models accumulate and store information about the mobility behavior of the users in terms of the time sequence in which activities are performed. The activities performed by a user can be broadly classified into navigation activity and location activity. Location activities are typically performed while staying at a location. In this paper a minimum threshold $t_{min}$ is set for an activity to graduate as a location activity. Navigation activities are those associated with physical movement between places where a location activity occurs. Users' everyday activities are observed and recorded into six groups based on the regularity with which it occurs and the duration. Based on these criteria the activities are classified into the following six groups.

| **Type** | **Duration** | **Symbol** |
|---|---|---|
| Regular | Short | RS |
| | Medium | RM |
| | Long | RL |
| Irregular | Short | IS |
| | Medium | IM |
| | Long | IL |

Each day on which the observations are made is segmented into 9 different time slots. For each slot the activity data is gathered and stored separately to make accurate predictions even when the sequence in which activities are carried out is variable. For each time slot, the probability of occurrence of the activity and the probability to move from one activity to next can be extracted from users' trace.

## 3. MARKOV MODEL FOR ACTIVITY TRANSITION

User's movement patterns attain a steady state and are predictable after an initial volatile period. The activity transition is modeled as a Markov chain using the seven states considered for the study. The user activity pattern history is used to develop a model that predicts the user's $n+1^{th}$ day's movements given n days of information. The Markov chain model uses the current activity on the $n+1^{th}$ day to predict the next activity for the day. A= {a1, a2….a7} is the set of all activity states. The transition probability from an activity i to activity j is given by the relative frequency of the sequence $a_i a_j$ where $\{a_i, a_j \in A\}$. This corresponds to the i, j entry of the transition probability matrix.

$$\begin{array}{c} \\ 1 \\ . \\ . \\ 7 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & . & . & 7 \end{array} \\ \begin{bmatrix} p_{11} & \cdot & \cdot & p_{17} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ p_{71} & \cdot & \cdot & p_{77} \end{bmatrix} \end{array}$$

In the matrix P, $p_{17}$ is the transition probability from activity 1 to activity 7 and $p_{71}$ denotes the transition probability from activity 7 to activity 1. The sum of all the transition probabilities in any row is equal to 1

i.e. $\sum_{i=1,j=1}^{j=7} p_{ij} = 1$ .

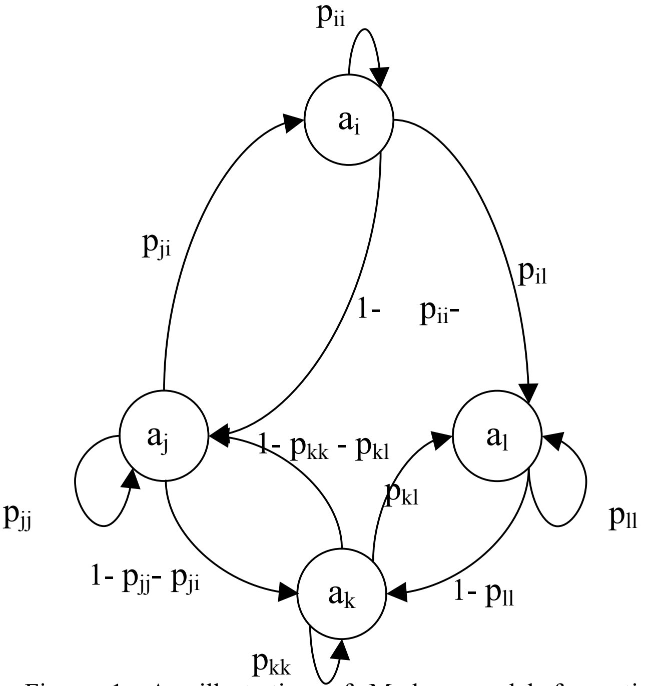


Figure 1: An illustration of Markov model for activity transition with four states

A single transition matrix for the whole day will not give correct predictions for the next probable activity, given the current activity. For example a user might move from activity “home” to activity “workout” every morning, come back to home and again move to activity “office”. So a single activity transition matrix will give equal probability of occurrence for both the activities at any time of the day. In order to overcome this problem and to give a temporal dimension and to aid timed predictions, the day is divided into eight fixed time slots (0-360, 361-480, 481-600, 601-720, 721-840, 841-960, 961-1080 and 1081-1440). Individual activity transition matrices are created and maintained for each of the eight time slots. Figure 1 is illustration of the Markov model for activity transition with 4 activities.

State space A corresponds to different activities $A=\{a_i, a_j, a_k, a_l\}$. Transition probability matrix $\mathrm{T(a_i, a_j) = N(a_i} \rightarrow a_j) / \mathrm{N(a_i^{'}} \rightarrow a_j)$ where $\mathrm{N\ (a_i \rightarrow a_j)}$ is the number of times activity aj follows activity ai and $\mathrm{N(a^{'}_i} \rightarrow a_j)$ is the number of times activity aj doesn’t follow ai. Initial state probability vector is defined by Λ (ai) = N (ai) / Σ N (ai’), next state probability vector is defined by A (t) = A (t-1)* T.

## 4. MOBILITY PREDICTION ARCHITECTURE

The User Mobility History (UMH) is derived from UAPs observed over a period of ‘n’ days. The UMH contains two sets of data for location activities for each time slot. The first suite of figures furnishes the probability of transition from one activity to another. The second set of data gives the set of all activity locations of interest visited during the said periods arranged in the order of time-of-visit. It furnishes the information about the cell in which the activity was carried out and other information to identify the exact next activity location in the cases where more than one location is available for a single activity. When updates were made to the activity transition matrix depending upon the number of days and the days on which an activity occur weights are assigned for obtaining the transition probability. By this way the system adapts to changes in pattern much faster while the history is also given due importance. Since weights are assigned based on the occurrences, regular and irregular activities are clearly identified. Regular activities carry more weight when compared to irregular activities. Navigation activities are recorded after proper scrutiny to avoid repetitions. They are identified a series of cells that occur in the trace between to end cells where an activity has been identified. The schematic representation of the proposed mobility prediction architecture is shown in Figure2.

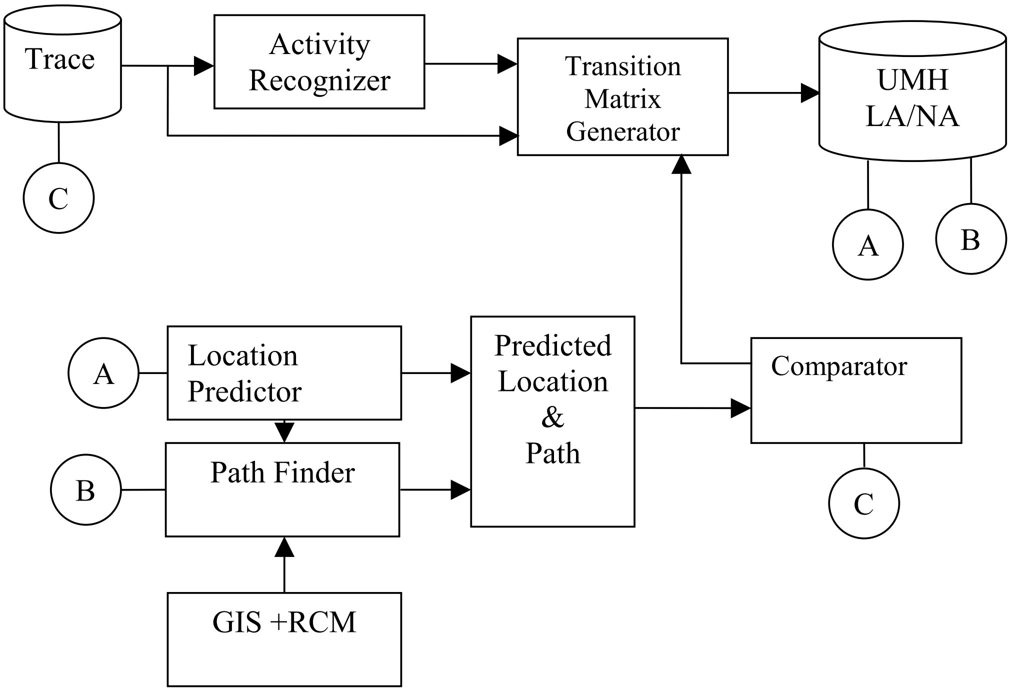


Figure 2 Mobility prediction architecture

### 4.1 Location prediction

Future significant locations of the mobile terminal can be predicted using the UMH when the current location of the mobile terminal is known.

Let i be the current activity, [tmin, tmax] be the current time slot, $t_{current}$ be the current time, $t_{elapsed}$ be the elapsed duration in the current activity in the time slot, $t_{mean}$, i be the mean activity duration of I in the time slot, and tpredicted be the time until which the mobility of the user is to be determined.

```
If
($t_{\min} \le t_{predicted} \le t_{\max}$ && $(t_{current} + (t_{mean,i} - t_{elapsed})) > t_{predicted}$)
next activity= current activity
else
$t_{current} = (t_{current} + t_{mean,i} - t_{elapsed})$
while ($t_{current} < t_{predicted}$)
Get next activity and $t_{mean,na}$ of next activity from the UMH
next activity (na) = activity with maximum transition probability given the current activity
if$(t_{current} + t_{mean,na} \ge t_{predicted})$
go to location prediction module with series of activity predictions made
else
$t_{current} = t_{current} + t_{mean,na}$
end while
```

By doing this the series of activities that are between current time and the predicted time can be determined which can in turn be used to find the corresponding locations and the path between them. If the $t_{predicted}$ lies in a time slot different from the current time slot, the series of activities in the current time slot, if any is first predicted. Then the activity the user is in at the time $t_{max}$ of the current time slot is taken as the current activity for the tmin of the next time slot and the activity transition matrix of the next time slot is used to predict the series of activities a user will be involved in until tpredicted is reached.

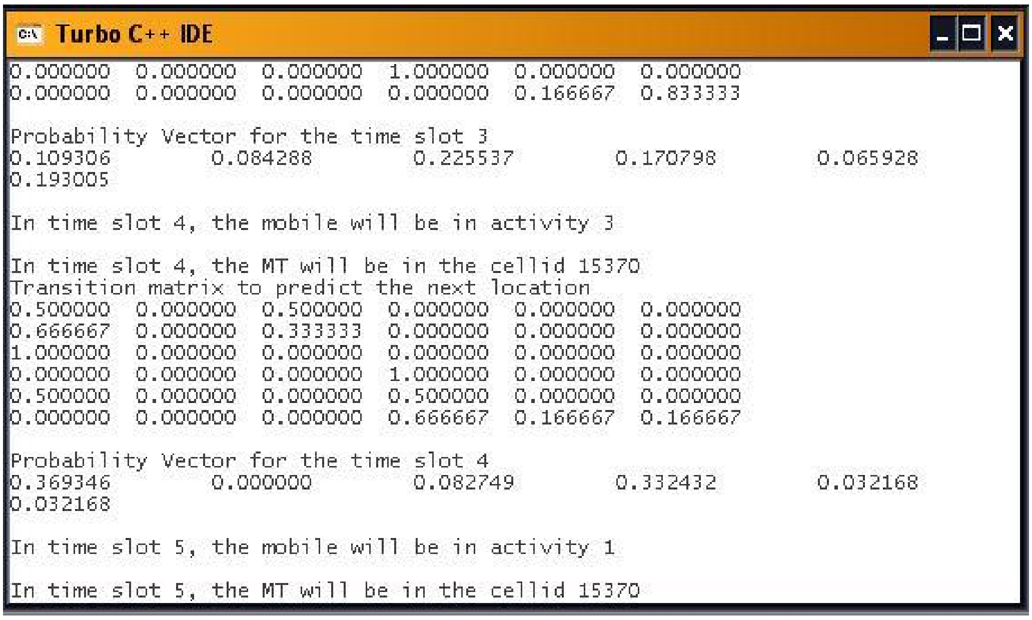
Once the activity is predicted the location where the activity is likely to take place has to be determined. If there is only one location activity with act_id = $a_{predicted}$ within period 'm' then it automatically qualifies as the location of the MT at time tpredicted. If more than one location activity qualifies with act_id = $a_{predicted}$ then (fromx_act_id, fromx_cell_id, freqx) values are used to determine the location. For example the number of shopping locations of interest for a user could be more than one. In case of incidental shopping people prefer places close to their home or office and the choice of shopping location is generally dependant on current location from where the user ventures.

### 4.2 Path finder

Once the future location or set of all locations that are likely to be visited by the MT is determined the path taken can be obtained from the UMP. The navigation activity records of the UMP provide the path between locations of interest. Alternatively it is possible to obtain the path which is a sequence of radio cells lying between these locations of interest using the GIS+RCM system that was used in the activity recognition phase. To determine the path when several alternatives exists a shortest path algorithm was used. In this study to obtain the shortest path when alternatives exist we have used Dijiktra's shortest path algorithm.

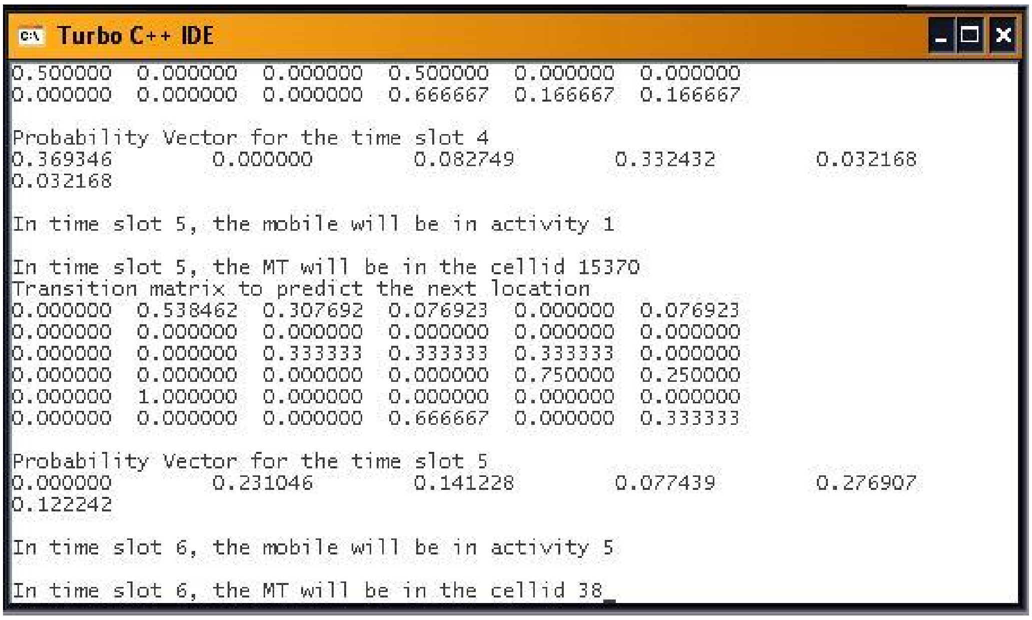
## 5 Simulation and Results

This section summarizes simulation results of the proposed framework. Simulation architecture was constructed to model and predict MT's traveled trajectories. The prediction and path finding algorithms explained earlier were implemented using C++. All activities on working days alone were considered. Radio coverage maps obtained from a major cellular service provider in Chennai, India were used to determine and represent user's location and navigation activities as a sequence of radio cells with their geographic positions. User mobility traces collected in Chennai, India alone were used as radio coverage maps for the other set of data was not available. The location activities were identified and segregated and the UAPs were obtained. The UMH summarizes the activity traces and creates the database with records and attributes based on which predictions were made. The minimum location activity duration $du_{min}$ was taken as 30 minutes. As far as visited locations were concerned the information were maintained for nine fixed time slots (0-360, 361-480, 481-600, 601-720, 721-840, 841-960, 961-1080 & 1081-1440). Using 'n' days old UMH the 'n+1'th day's location and navigation activities were predicted and compared with the original values and error graphs were plotted. The performance of the algorithm, for users chosen randomly from predictable and quasi-predictable predictability groups is provided. Figure 3 and Figure 4 are snapshots of the simulator output showing predicted location for a given time in two consecutive time slots. A weighted update system is followed that helps the system to react quicker to changes in mobility pattern. This helps the scheme to overcome some of the drawbacks associated with regular pattern based prediction techniques. Fig 5 gives the performance improvement in AMPuMM technique over the CKbP technique. This improvement is achieved as the motivation for each move has also been considered. Updating new values and correcting the errors further train the system that shows improvement in system performance. This performance improvement is shown in fig 6. Here, performance is valued by the hit ratio. A hit refers to a correct prediction that matched with the actual location and an incorrect future location prediction is referred to as a miss. These results are observed after training the system for 25 days.



```
0.000000  0.000000  0.000000  1.000000  0.000000  0.000000
0.000000  0.000000  0.000000  0.000000  0.166667  0.833333

Probability Vector for the time slot 3
0.109306        0.084288        0.225537        0.170798        0.065928
0.193005

In time slot 4, the mobile will be in activity 3

In time slot 4, the MT will be in the cellid 15370
Transition matrix to predict the next location
0.500000  0.000000  0.500000  0.000000  0.000000  0.000000
0.666667  0.000000  0.333333  0.000000  0.000000  0.000000
1.000000  0.000000  0.000000  0.000000  0.000000  0.000000
0.000000  0.000000  0.000000  1.000000  0.000000  0.000000
0.500000  0.000000  0.000000  0.500000  0.000000  0.000000
0.000000  0.000000  0.000000  0.666667  0.166667  0.166667

Probability Vector for the time slot 4
0.369346        0.000000        0.082749        0.332432        0.032168
0.032168

In time slot 5, the mobile will be in activity 1

In time slot 5, the MT will be in the cellid 15370
```

Figure 3 Snapshot of the simulator output showing predicted location for slot 4



```
0.500000  0.000000  0.000000  0.500000  0.000000  0.000000
0.000000  0.000000  0.000000  0.666667  0.166667  0.166667

Probability Vector for the time slot 4
0.369346        0.000000        0.082749        0.332432        0.032168
0.032168

In time slot 5, the mobile will be in activity 1

In time slot 5, the MT will be in the cellid 15370
Transition matrix to predict the next location
0.000000  0.538462  0.307692  0.076923  0.000000  0.076923
0.000000  0.000000  0.000000  0.000000  0.000000  0.000000
0.000000  0.000000  0.333333  0.333333  0.333333  0.000000
0.000000  0.000000  0.000000  0.000000  0.750000  0.250000
0.000000  1.000000  0.000000  0.000000  0.000000  0.000000
0.000000  0.000000  0.000000  0.666667  0.000000  0.333333

Probability Vector for the time slot 5
0.000000        0.231046        0.141228        0.077439        0.276907
0.122242

In time slot 6, the mobile will be in activity 5

In time slot 6, the MT will be in the cellid 38_
```

Figure 4 Snapshot of the simulator output showing predicted location for slot 5

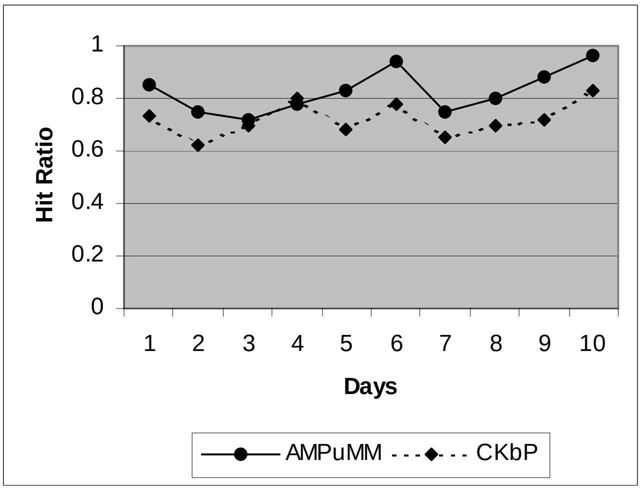


Fig 5 Performance Improvement

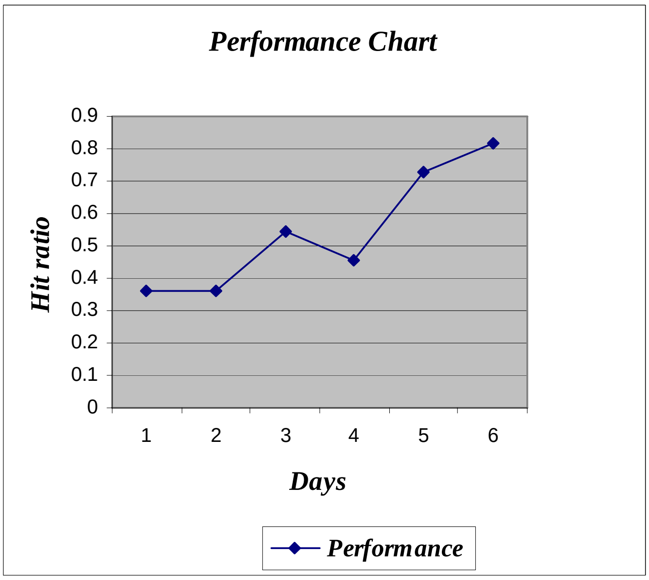


Fig 6, Performance Improvement

## 6 Conclusions and Future work

The Activity based Mobility Prediction using Markov Modeling (AMPuMM) is free from maintaining large user history. By periodically updating the transition matrix and state probability vector, the system performance could be improved. With this AMPuMM technique, 82.6% accuracy is achieved, as the motivation for each move is also included. Some times, the MT's present whereabouts, which will be used to generate the user profile, can't be found because of signal loss. To fix this problem, in future, GPS traces could be used to find the MT's location [7]. Using the longitude and latitude information extracted from the GPS traces, the present cell id could be found. Thus, the GPS trace information could be converted into GSM trace format. Solely using GPS information would also lead to a problem of shadowed region [8]. Sensors could be fixed inside the buildings to overcome this problem. But, this needs to facilitate each building with sensors. To avoid this, for the shadowed regions, the GSM traces could be used. By combining the GPS and GSM trace information, the system performance could be improved.

## References

[1]. Jingyuan zhang, "Location Management in Wireless Cellular Networks", The Computer Journal, Page 20 of 22, January 2005.

[2]. Alejandro Quintero, "A User Pattern Learning Strategy for Managing Users' Mobility in UMTS Networks", IEEE Transactions On Mobile Computing, vol. 4, no. 6, pp. 552-556, 2005.

[3]. M.S.Sricharan, V.Vaidehi, P.P.Arun, "An Activity based Mobility Prediction Strategy for Next Generation Wireless Networks", IEEE, IFIP, Wireless Optical and communication Networks, 2006.

[4]. Nancy Samaan and Ahmed Kormouch, "A Mobility Prediction Architecture Based on Conceptual Knowledge and Spatial Conceptual Maps", IEEE Transactions on Mobile Computing, vol 4, no 6, pp. 537-551, 2005.

[5]. Ramesh R. Sarukkai "Link Prediction and Path Analysis using Markov Chains", 9th International World Wide Web Conference (WWW9) Refereed Paper, 2000.

[6]. Kari Laasonen, "Clustering and Prediction of Mobile User Routes from Cellular Data", Basic Research Unit, HelAsinki Institute for Information Technology Department of Computer Science, University of Helsinki, 2005.

[7]. L. Liao, D. Fox, and H. Kautz, "Extracting Activities and Significant Places from GPS Traces", Proceedings of the International Symposium of Robotics Research (ISRR), 2005.

[8]. Zainab R.Zaidi and Brian L.Mark, "Real-time Mobility Tracking Algorithm for Cellular Networks Based on Kalman Filtering", IEEE Transaction on Mobile computing, Vol. 4, No. 2, March/April 2005