

# Mobility Prediction in Dynamic Grids

Iraky Khalifa<sup>1</sup> & Hala Mohamed Abbas<sup>1</sup>

<sup>1</sup> Faculty of Computers and Information, Computer Science Department, Helwan University, Cairo, Egypt

Correspondence: Hala Mohamed Abbas, Faculty of Computers and Information, Computer Science Department, Helwan University, Cairo, Egypt. Tel: 20-100-890-9016. E-mail: ass.hala@gmail.com

Received: February 1, 2012 Accepted: February 24, 2012 Online Published: May 1, 2012

doi:10.5539/cis.v5n3p90

URL: <http://dx.doi.org/10.5539/cis.v5n3p90>

## Abstract

Mobile grids are one of the emerging grid types. They help to pool the resources of several cooperative mobile devices to resolve a computationally intensive task. Mobility management is a key challenge in mobile grids which includes mechanisms to track and maintain the locations of mobile nodes. Mobility prediction is one of the most important components of mobility management, which defined as proactively tracking the future locations of mobile users. In this paper we introduce a straight forward mobility predictor that can predict the future locations of mobile users and the duration for which the users will remain in those locations.

**Keywords:** mobile grid, mobility management, mobility prediction, prediction accuracy

## 1. Introduction

The problem of computing power moved toward a decentralized approach due to the availability of high speed networks and powerful computers (Buyya, Abramson, & Giddy, 2000). Because of the multi institutional and collaborative nature of science, it leads to the need for accessing computing resources and scientific devices which are globally distributed. As a result, a new architecture known as “The Grid” has been explored for high performance distributed applications. (Foster, Kesselman, & Tuecke, 2001) defined the Grid as “coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations”. Grid computing contains many types of resources such as CPU power, files, storage devices, sensors and many others. These resources may be located in many organizations and distributed among different geographical locations. The basic role for Grid is to facilitate these resources' interactions in order to support large-scale science and Engineering. The challenge for the grid infrastructure is to enable transparent access while allowing resource providers to keep or establish their own local policies.

Recently grid computing has been changed to wireless and mobile grid environments that take into account user mobility by integrating mobile devices into wired grid, but the limitations of the previous mobile devices such as poor performance and the small amount of battery represented a challenge for this integration. Fortunately, due to the rapid increase of users who use mobile devices, this led to growing the capacity of latest mobile devices and hence encourages researchers to utilize it in the grid. Compared to traditional (wired) Grids, mobile Grids have the advantage of exploiting the concepts of mobility and ubiquity in terms of being available any-time, anywhere and by all means. Making use of wireless connection, the mobile Grid environment is available any-time and any-where.

The fundamental requirement for the wireless networks to provide flawless service to the mobile user is that it must know the point of attachment of the mobile device at any point of time, which promoted the mobility management problem. Mobility management in mobile grids is defined as the ability to store and update the location information of mobile users who are served by the system. Mobility prediction is the most important topic in mobility management, it is defined as the ability to predict the next movement of a mobile user who travels between the cells of a wireless network. The predicted movement can then be used to enhance the efficiency of the system by effectively allocating resources to the most probable-to-move cells instead of blindly allocating excessive resources in the cell-neighborhood of a mobile user. This in turn could reduce the latency in accessing the resources and improve resource utilization (Yavas, Katsaros, Ulusoy, & Manolopoulos, 2005). Most of the mobility predictors depend on collecting the user's movement history and partitioning this record to get what is named user mobility patterns. And because the prediction accuracy is the most important factor, it depends not only on the method used for extracting movement patterns but also on the prediction algorithm used. By analyzing real-world traces which provide a flexible and realistic way to extract movement patterns and using

pattern matching approaches, more feasible mobility predictors can be provided. Most of the work that has been done in the field of location prediction uses Markov models (models based on text compression) which provide feasible location predictors but they have some deficiencies, which are introduced in the following:

- The number of transitions exponentially grows with increasing the model order (i.e., the number of history states used in the prediction) which limits their scalability.
- Adding new historical data needs rebuilding the movement patterns.

In this paper we proposed a simple mobility predictor with the ability to model the frequencies of movement history transitions and without the need for a training step like Markov models. Thus, new locations and history traces can be included faster.

This paper is organized as follows. Section 2 presents related work on mobility prediction and how it is used in mobile Grid. Section 3 demonstrates the proposed model. Experimental results are presented in Section 4. Finally, in Section 5 we conclude our paper.

## 2. Related Work

This section divides the related work according to the two main research topics covered by this paper: Mobility prediction and Using mobility models in mobile grids.

### 2.1 Mobility Prediction

The main goal of mobility prediction is to facilitate continuous access to grid resources irrespective of users' mobility. It also aims to enable mobile devices to easily interact with resources in the grid. Although many different mobility prediction techniques have been proposed, these techniques can be broadly classified into the following three categories: stochastic techniques, data mining techniques, and pattern matching techniques. First, stochastic techniques provide mobility predictions using probabilistic models. These techniques provide means to describe users' movements by assuming certain topographies of areas. Different mobility prediction models have been introduced in (Issac, Hamid, & Tan, 2010) such as random walk mobility model. (Nandeppanavar, Birje, Manvi, & Shridhar, 2010) used the normal walk mobility model to predict users' locations and their direction of movement and evaluated the impact of user mobility on task execution time and task failure rate of a mobile grid system. Second, data mining techniques use a database to track and characterize the long-term mobility patterns, which are then used to predict locations of mobile users. A considerable amount of work has been done in mobility prediction using pattern mining, in (Yavas, Katsaros, Ulusoy, & Manolopoulos, 2005) a three phase algorithm was proposed by mining user mobility patterns from the history of user's trajectories, generating the association rules and then using these rules to predict the next location. Monreale, Pinelli, Trasarti and Giannotti (2009) proposed a prediction method using trajectory pattern mining. In trajectory pattern mining the learning process depends not only on the movement history of an individual object, but on the movement history of all available objects in a certain area. In (Vu, Ryu, & Park, 2009) a new prediction method was introduced which provides a higher accuracy level. This method used sequential pattern mining to discover frequent patterns which was considered an improvement to (Yavas, Katsaros, Ulusoy, & Manolopoulos, 2005). Another algorithm called Knowledge Grid based Mobility Pattern Mining was proposed in (Sakthi, & Bhuvaneswaran, 2009) in which data mining techniques have been integrated with grid technologies to use distributed data mining for large data sets in order to find hidden valuable information. A hybrid between stochastic and pattern mining was introduced in (Issac, Hamid, & Tan, 2010) to forecast the future path of a mobile user.

Although, these techniques provide efficient mobility predictors, they require large storage capacities and fast processors to analyze long-term mobility behavior. Pattern based approaches (history based approaches) provide feasible location predictors, since they rely on the fact that the probability of the future outcome is based on the current and past outcomes. The most popular history based technique is Markov Models, Typically, a Markov mobility predictor performs the following two operations, the first operation is to maintain a collection of past locations of the mobile users, while the second operation is to predict future locations based on the value of conditional probability that matches the past locations of the mobile users. In (Song, Kotz, Jain, & He, 2004) several O(k) Markov predictors have been evaluated. They proved that the prediction accuracy improved by using fallback optimization which recursively uses the result of the O(k - 1) predictor when the O(k) predictor does not produce a prediction. An improved version of that predictor has been introduced in (Sun & Blough, 2007) which used not only past movement histories in the prediction but they used user's future information to enhance prediction accuracy. The results showed that integrating future knowledge with the predictor resulted in a significant performance gains.

## 2.2 Using Mobility Models in Mobile Grids

Considering user mobility in grid environments has established new challenges in the area of resource management. It includes many problems such as job scheduling, replication and replica management. Several studies have researched on the effect of integrating mobility prediction schemes with resource management solutions in mobile Grid. (Farooq & Khalil, 2006) proposed a generic mobility model to predict the time for which a user will remain in a specific domain. The learning process of this model based on the user's behavior in the past to compute the average mobility and the number of jobs that can be executed during user's availability. (Ghosh, Roy, & Das, 2010) proposed a node mobility prediction framework to formulate a cost effective job scheduling model based on a predetermined pricing strategy for mobile grid. In this model, a scheduler decides the price per unit resource by the game theory and allocates jobs to mobile resources with cheaper cost. This model can minimize the total price, which a system pays to resources to complete jobs, but cannot assure the QoS of a grid computing service.

In replica management aspect, there are few studies in the field of replica placement. (Huang & Chen, 2006) proposed replica placement model that considers Group mobility. Gossa, Janecek, Hummel, Gansterer and Pierson (2008) proposed a mobility predictor and investigated its effect in the performance of a replica placement algorithm.

## 3. The Proposed Mobility Predictor

Our proposed predictor includes two phases: a preprocessing phase and the actual prediction algorithm. In the preprocessing part, raw history data traces have been converted to a form used in the prediction phase. The following sections present the two components.

### 3.1 Preprocessing

To be able to predict future locations, past locations should be determined and recorded. Any positioning system can be used to determine location as long as it provides enough information for determining the users' location with temporal information such as the arrival time. In this paper locations have been determined using IEEE 802.11 wireless access points. With the advent of location systems based on IEEE 802.11 wireless access points such as Place Lab, and public collections of wireless usage data such as CRAWDAD (Kotz & Henderson, 2005) and UCSD (McNett & Voelker, 2005) it is now feasible to apply prediction algorithms to larger groups of people over indoor and outdoor locations. In this paper we have used The UCSD dataset as used by (Burbey, 2011).

The UCSD dataset consists of polls recorded every 20 seconds of 275 PDAs over an 11-week period, resulting in a file with more than 13 million records. Because of the large amount of data, the first step is to reduce the data to retain the useful features and dispose of noise. We performed the following steps on the UCSD dataset.

1-The large file has been broken to separate files for each user.

2-Records where the device sensed an access point but did not associate will be dropped.

3-Continuous records will be merged into sessions, where each session consists of a starting time, the associated access point, and the duration of the session.

### 3.2 Prediction Algorithm

The location traces resulted from the preprocessing stage used as the base to the prediction part. We used a comparison strategy to find mobility patterns, and then these patterns have been used to predict the next locations of a mobile user. Figure 1 introduces the prediction of the next  $N$  locations of a mobile user  $X$ , it is performed as follows:

1-Constructing a set  $SX$  that contains the past  $i$  locations and the current location  $X_C$ .

2-Matching the sequence  $SX$  with all history sequences to generate the history set  $SH$  which contains the past  $i$  locations, the current location and the next  $N$  locations.

3-Adding the resulted sequences to the prediction set  $Plist$ .

4-Checking if  $Plist$  is empty (i.e., no match was found), then the session with the shortest duration is removed from the past  $i$  locations. The history set is searched again until  $Plist$  contains at least one historical sequence.

5-Predicting the next  $N$  locations by calculating the probability for each of the resulted locations according to how many times it appears in the history set.

```

Input: pre-processed history set
 $SX = \{X_{c-t}, X_{c-t+1}, \dots, X_c\}$ 

Output: prediction of next locations

 $Plist := \text{NULL}$ 

repeat
  for each location sequence in the history set  $SH$  do
    if  $SX$  matches  $SH(H_{c-t}, \dots, H_c)$  then
      add  $SH = \{H_{c-t}, H_{c-t+1}, \dots, H_c, \dots, H_{c+N}\}$  to  $P$ 
    end
  end
  if  $Plist == \text{NULL}$  then
    remove the shortest session
  end
until  $Plist \neq \text{NULL}$ ;

Choose location with highest relative frequency

```

Figure 1. Mobility prediction algorithm in pseudo code

#### 4. Experimental Results

As mentioned previously, the initial analysis of the large UCSD dataset involves preprocessing the data to extract useful features. As part of the Wireless Topology Discovery (WTD) project at UCSD, 300 freshmen delivered PDAs and used for collecting movement traces. The resulting dataset consists of two files, the access point locations and the trace data. The data was collected during an 11 week trace period from 9/22/2002 to 12/8/2002. While a student's device was powered on, WTD sampled and recorded the following information every 20 seconds for all access points (APs) that it could sense across all frequencies not just the access point the wireless card was associated with at the time:

- USER\_ID: Unique identifier assigned to the user
- SAMPLE\_TIME: The time the sample was taken by the WTD software
- AP\_ID: Unique identifier assigned to the detected AP
- SIG\_STRENGTH: Strength of AP signals received by device
- AC\_POWER: Whether the device used AC power (1) or battery (0)
- ASSOCIATED: Whether the device was associated with this AP (1) or not (0)

Since WTD recorded the above information for all APs sensed during a sample, if a user device saw three APs in one sample, there were three entries recorded for that sample (which differ only in the AP detected, signal strength, and associated flag). Figure 2 presents the format of the original dataset.

USER_ID	SAMPLE_TIME	AP_ID	SIG_STRENGTH	AC_POWER	ASSOCIATED
123	09-22, 00:00:00	359	8	0	0
123	09-22, 00:00:00	363	5	0	0
123	09-22, 00:00:00	365	11	0	1
191	09-22, 00:00:00	355	31	0	1
101	09-22, 00:00:00	353	8	1	0
101	09-22, 00:00:00	362	30	1	1
129	09-22, 00:00:00	369	31	0	1
156	09-22, 00:00:00	360	19	1	1
184	09-22, 00:00:02	352	29	0	1

Figure 2. A portion of the UCSD dataset

To be able to practically work with such a large amount of data (over 371 Megabytes), the first step towards data reduction was to parse the large file into separate files, one for each of the 275 users. Further reduction was achieved by using only the records where the device was associated with an access point, dropping the records where the device sensed an access point but did not associate. Continuous records were combined to record sessions, with a starting time, the associated access point and the duration of the session. A sample of the resulting file for user 200 is shown in Figure 3.

User	Start Time of Session	Associated Access Point	Duration of Session
200	26/09/2002 07:18:27 pm	93	0:01:20
200	26/09/2002 07:19:47 pm	48	2:03:45
200	26/09/2002 09:23:32 pm	352	0:01:00
200	26/09/2002 09:24:32 pm	354	2:18:33
200	26/09/2002 11:43:05 pm	352	0:54:00
200	30/09/2002 02:52:20 pm	352	0:00:22
200	02/10/2002 10:01:36 pm	352	0:01:42
200	03/10/2002 09:34:59 am	352	3:08:47

Figure 3. Sample of records for User #200

To evaluate our proposed algorithm, a set of experiments have been performed. The predictions resulted in one of the following: a correct prediction, an incorrect prediction, or a prediction miss. A correct prediction results when the predicted location is equivalent to the actual next location of a user, otherwise it is considered an incorrect prediction. A prediction miss occurs when the predictor fails to make a prediction. The prediction accuracy can be calculated by dividing the number of correct predictions by the number of predictions made when prediction misses are ignored. We evaluated our algorithm considering the current and the last two positions and compared it with a third order Markov predictor.

The performance of the two predictors is shown in Figure 4, we analyzed the prediction accuracy for fractions of users and the results showed that the proposed predictor has greater prediction accuracy than markov model. Figure 5 showed that the implemented predictor has an average performance than the other one for various tests. Additionally we plan to investigate the effect of considering the locations with less probability on the accuracy and not just the highest probability prediction.

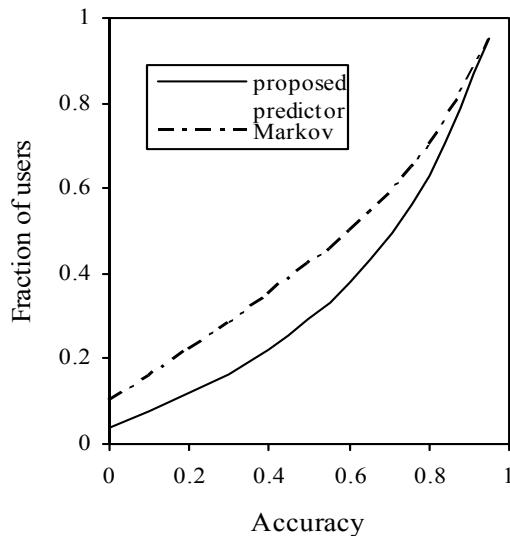


Figure 4. Comparison of prediction schemes

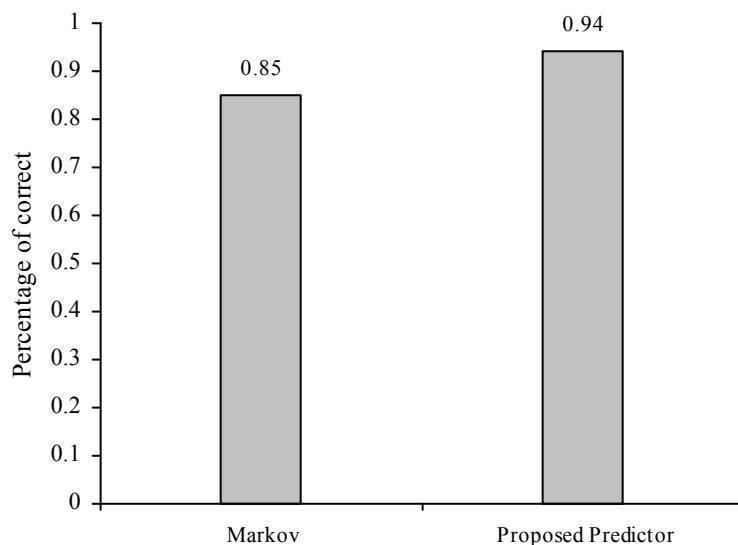


Figure 5. Average accuracy

## 5. Conclusion

This paper presented a mobility predictor based on real world traces. The predictor is straight forward and does not need a training step such as the Markov models. In addition, it is not restricted to a specific type of mobility traces and can be applied to different mobility patterns. Our immediate future work consists of integrating this predictor into a mobile grid system to enhance replica placement techniques.

## References

- Burbey, I. E. (2011). *Predicting future locations and arrival times of individuals*. Retrieved from [http://www.scholar.lib.vt.edu/theses/available/etd.../Burbey\\_IE\\_D\\_2011\\_2.pdf](http://www.scholar.lib.vt.edu/theses/available/etd.../Burbey_IE_D_2011_2.pdf)
- Buyya, R., Abramson, D., & Giddy, J. (2000). An architecture of a resource management and scheduling system in a global computational Grid. *HPC Asia. May*(14-17), 283-289, Beijing, China.
- Farooq, U., & Khalil, W. (2006). A generic mobility model for resource prediction in mobile grids, *In Proceedings of the International Symp. on Collaborative Technologies and Systems*.

- Foster, I., Kesselman, C., and Tuecke, S. (2001). The anatomy of the grid: enabling scalable virtual organizations. *International Journal of Supercomputer applications*, 15(3). <http://dx.doi.org/10.1177/109434200101500302>
- Ghosh, P., Roy, N., & Das, S. K. (2010). Mobility aware efficient job scheduling for single-class grid jobs in a generic mobile grid architecture. *Future Generation Computer Systems*, 26, 1356-1367. <http://dx.doi.org/10.1016/j.future.2009.05.003>
- Gossa, J., Janecek, A. G., Hummel, K. A., Gansterer, W. N., & Pierson, J. (2008). Proactive replica placement using mobility prediction. In *proceedings of the 9th International Conference on Mobile Data Management Workshops*.
- Huang, J. L., & Chen, M. S. (2006). On the effect of group mobility to data replication in Ad Hoc Networks. *IEEE Transactions on Mobile Computing*, 5, 492-507.
- Issac, B., Hamid, K., & Tan, C. E. (2010). Hybrid mobility prediction of 802.11 infrastructure nodes by location tracking and data mining. *Journal of IT in Asia*, 3, 9-24.
- Kotz, D., & Henderson, T. (2005). CRAWDAD: A Community Resource for Archiving Wireless Data at Dartmouth. *IEEE Pervasive Computing* 4, 12-14. <http://dx.doi.org/10.1109/MPRV.2005.75>
- Monreale, A., Pinelli, F., Trasarti, R., & Giannotti, F. (2009). WhereNext: A location predictor on trajectory pattern mining. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp.637-645.
- Nandeppanavar, A. S., Birje, M. N., Manvi, S. S., & Shridhar. (2010). Mobility impact on performance of mobile grids. *International Journal of Computer Science and Information Security*, 7(1), 106-111.
- Nett, M. M., & Voelker, G. M., (2005). UCSD Wireless Topology Discovery Trace. Support for the UCSD Wireless Topology Discovery Trace was provided by DARPA Contract N66001-01-1-8933.
- Sakthi, U., & Bhuvaneswaran, R. S. (2009). Mobility prediction of mobile users in mobile environment using knowledge grid. *International Journal of Computer Science and Network Security*, 9(1), 303-309.
- Song, L., Kotz, D., Jain, R., & He, X. (2004). Evaluating location predictors with extensive Wi-Fi mobility data. In *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 2, pp.1414-1424.
- Sun, M. H., & Blough, D. M. (2007). Mobility Prediction Using Future Knowledge. In *Proceedings of the 11th ACM MSWIM international conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*.
- Vu, T. N., Ryu, K. H., & Park, N. (2009). A method for predicting future location of mobile user for location-based services system. *Computers & Industrial Engineering*, Elsevier Science Ltd, 57, 91-105.
- Yavas, G., Katsaros, D., Ulusoy, O., & Manolopoulos, Y. (2005). Data mining approach for location prediction in mobile environments, *Data & Knowledge Engineering*, Elsevier Science Ltd, 54, 121-146.