A Path Prediction Model to Support Mobile Multimedia Streaming

[1]Apollinaire Nadembega, [1]Abdelhakim Hafid, [2]Tarik Taleb
[1]Network Research Lab, University of Montreal, Canada
{nadembea, ahafid}@iro.umontreal.ca

[2]NEC Europe, Heidelberg, Germany
talebtarik@ieee.org

Abstract – **Along with the recent and ongoing advances in the wireless and mobile access technologies, a wide plethora of mobile multimedia services have emerged. Ensuring an acceptable level of Quality of Service (QoS) is a crucial requirement to allow users enjoy these mobile multimedia services. One means to ensure QoS is to minimize the frequency and magnitude of fluctuations in the mobile multimedia streaming rates during the multimedia service and while users are on the move. For this purpose, there is need for tools to predict a user's long-term movement. In this vein, this paper proposes a Path Prediction Model (PPM) to predict a user's movement path. PPM is based on historical movement trace, current movement data and spatial conceptual maps; it assumes a priori knowledge of the destination. At each road intersection, the probability of selecting the next road segment is evaluated, based on historical data, towards the destination. These probabilities are computed via (a) filtering historical data according to the day of the week (e.g., weekend, holiday) and the time of the day; and (b) applying conditional probability rules taking into account the path used between the origin of movement, current position, and the destination. Simulations are conducted using real-life data to evaluate the performance of the proposed model. Encouraging results are obtained in terms of average prediction accuracy and mitigation of the impact of learning period and the remaining distance to reach the destination on the path prediction performance.**

I. INTRODUCTION

In the last decades, several applications have emerged and have been taking more space and importance in the lives of all of us, such as multimedia applications. Indeed, the exponential growth of mobile telephony, among others, has created the need for new services. These services have specific requirements in terms of bandwidth and Quality of Service (QoS). Thus, wireless networks, in particular cellular networks, should deploy mechanisms to support these requirements. In cellular networks, available data rates may vary noticeably from cell to cell. Thus, a mobile user may experience rate fluctuations when moving between cells (i.e., handoffs [1]); this may be not acceptable for several applications, such as multimedia streaming [3]. The key challenge is to provide a minimum acceptable QoS in each cell visited by the user; this requires prior knowledge of the user's long term movement (e.g., 10 minutes in advance). Indeed, if the network can predict the user's path, then it can provide him/her with the required QoS during the whole session. More specifically, the network will accept the user into the network (e.g., for a multimedia streaming session) only if there are sufficient resources (e.g., bandwidth) in each cell (during the user's residence in the cell) along the predicted path; otherwise, the user will not be allowed to access the network. Thus, path prediction with reasonable accuracy is a key to provide QoS for mobile users [2].

To predict user's mobility/movement, several mobility models have been proposed [4-14]. Most existing mobility models, that are suitable to model real-life mobility of users, are based on historical data of motion or mobility trace files of the mobile users. L. Liao*, et al.* [9] introduced a hierarchical Markov model that can learn/infer user's daily movements through urban communities. They described a system that creates a probabilistic model of user's daily movement patterns using unsupervised learning from raw GPS data. The proposed model works fine when the mobile user's movements follow a specific pattern; however, if there is a slight change in the user "specific" pattern, the prediction accuracy considerably suffers. Since the user's movements are generally uncertain, over time, the applicability of such a model is limited.

H. Jeung*, et al.*[2] proposed an approach that enables prediction of the path of a mobile user whose movement is constrained to a road network. More specifically, the approach predicts, based on historical data, when and where the user will make a turn at an intersection; the prediction is based only on the statistics of turns at intersections. Indeed, it does not take into account, for a given intersection, the user's previous movement, starting from the time it did access the network to the current intersection, the time of the day and the type of the day; without these elements, we believe the prediction accuracy will not be high

W. Wanalertlak*, et al.* [10] presented a scheme, called the Behavior-based Mobility Prediction (BMP), that provides mobility prediction; BMP provides, for a given user at a given cell, an accurate prediction of next cell the user will move to; it makes use of several factors including historical movement data of users and time of the day. BMP does not support path prediction of mobile users from source to destination.

N. Samaan, *et al.*[11] applied the Dempster-shafer' theory to the knowledge of user's preferences and goals to predict his/her mobility; the work does not make any assumption about the availability of users' movement history. The authors in [12, 13] applied the social theory to the structure of the relationships among individual users to predict their movements while F. Ekman*, et al.* [14] defined mobility models based on daily planned activities; they assumed that users move from home to work, from work to restaurant, from restaurant to work, from work to leisure, and return home in the evening.

In this paper, we propose a novel approach, called Path Prediction Model (PPM), which predicts the path the user will use during is/her movement from source to destination. PPM makes use of historical movement trace, current movement data and spatial conceptual maps; it also assumes a priori knowledge of the destination. More specifically, starting from the user initial position (i.e. where he/she first accesses the network), PPM, at each road intersection, determines the next road segment the user will likely use during his/her movement towards the destination; indeed, it computes the probability, based on historical movement trace and destination, the user will use a road segment and selects

the road segment with the highest probability as the next road segment. PPM repeats the same process until the selection of the last road segment to destination. The predicted user's path consists of the sequence of the selected road segments.

Unlike existing approaches, PPM takes into account the type of the day and the time of the day which we believe are key parameters to consider when dealing with mobility prediction. PPM assumes that the destination is known in advance; it uses among others the direction from the source to the destination to select next intersection (road segment). This allows increasing prediction accuracy (e.g., when a U-turn is necessary to reach the destination) in opposition to existing approaches that use direction from the source to current position. More importantly, PPM is the only scheme (to the best of our knowledge) that allows for predicting, without restrictive assumption (e.g., known specific user pattern [9]), the whole path from source to destination (existing schemes predict only the next intersection/cell).

The remainder of this paper is organized as follows. In section II, we describe the proposed PPM. Section III evaluates the performance of PPM. Section IV concludes the paper.

II. PROPOSED MODEL

In .[15], the authors proposed a framework (see Fig. 1) that assists in avoiding frequent fluctuations in the streaming rates of mobile multimedia services and ultimately ensuring acceptable Quality of Service (QoS). Our proposed PPM approach is implemented at the User's Equipment (UE) Mobility Predictor (MP) as shown in Fig. 1. This model assumes that the destination is known in advance and uses this information in the path prediction.
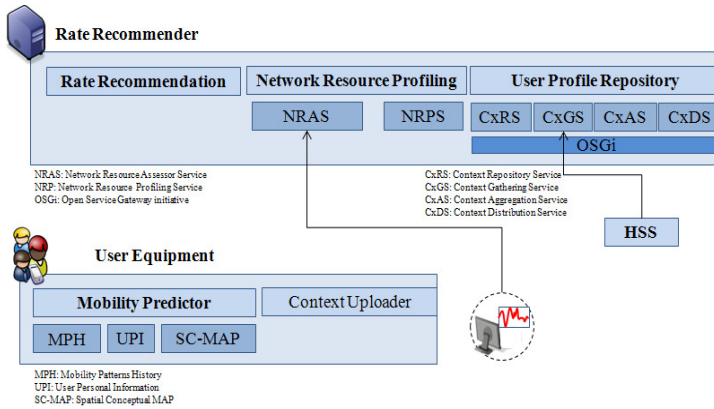

Fig. 1. Key components of RR and mobile terminals (details on the functionality of each unit in the framework are available in [15]).

Note that a UE maintains a database (DB) which records data about users' movements, users' contexts and their living areas. We assume the availability of static data about geographic areas (topology/map of roads), called Navigation Map (NM). NM contains coordinates of road intersections and coordinates of user's Frequently Visited Locations – FVLs – (home, office, shopping mall, etc). We assume that UEs embed a technology, such as GPS, that allows recording the used road segments taking into account day and time. All user possible locations (home, office, road intersection, restaurant, etc) are considered as nodes. Each node is identified by a node ID, latitude and longitude. A road segment is the road portion between two intersections. The table of a road segment consists of edge ID, direction of navigation, node ID 1 and node ID 2. The User Movement Trace (UMT) contains user ID, date, time and node ID that represents user location at date and time. Various algorithms could be used for gathering these data.

Using UMT, an algorithm extracts User FVL Trace (UFVLT) that contains User ID, date, arrival time, departure time and node ID.

Algorithm 1 presents the pseudo-code for recording data. This pseudo-code is run by UEs. At each time slot t, acceleration a and geographic coordinates (*latitude and longitude*) are measured.

Algorithm 1: Pseudo-code for movement data gathering.

Input : User_id, NM , a_s , d_s

Output : UMT and UFVLT

Variables :
- ✓ Arrival: is a boolean which marks user arrival in the specific position
- ✓ Departure: is a boolean which marks user departure from the specific position
- ✓ Var: is a data set recorder
- ✓ node: is a node

Functions
- ✓ getNode(lat,lon) : it is used to get the node where latitude=lat and longitude=lon from DB.
- ✓ existNode(lat,lon): it is used to check whether the node where latitude=lat and longitude=lon exists in DB.
- ✓ Node(lat,lon) : it is used to create a new node with a generated node_id and latitude=lat and longitude=lon
- ✓ getDate(): it is used to get current date
- ✓ getTime(): it is used to get current time

1. **each** t sec
2.   sample a // acceleration
3.   sample (lat, lon) // position
4.   **if** NM.existNode(lat,lon) // (lat, lon) is into NM
5.   **insert into** *UMT* values (User_id, getDate(),getTime(), NM.getNode(lat,lon))
6.   **if** $|a| < a_s$ **and** Arrival==true
7.   **if** UC.FVLC.existNode(lat,lon)   // (lat, lon) is intoUC.FVLC
8.   node= new Node(UC.FVLC.getNode(lat,lon))
9.   **else if** NM.existNode(lat,lon) // (lat, lon) is into NM
10.   node= new Node(NM.getNode(lat,lon))
11.   **else**
12.   node=new Node(lat,lon)
13.   **insert into** *NM* values (User_id,node)
14.   **end if**
15.   **insert into** *UFVLT* **values** (User_id, getDate(),getTime(),NULL,node)
16.   Departure=True
17.   Arrival=False
18.   var(user_id)=user_id
19.   var(node_id)=node_id
20.   var(time)=time   // keeping some values
21.   **else if** $|a| \geq a_s$ **and** Departure==true
22.   **if** getTime() – var(time) >= d_s
23.   **update** *UFVLT* **set** departure_time=getTime() **where** user_id=var(user_id) **and** node_id=var(node_id)
24.   **else**
25.   **delete from** *UFVLT* **where** user_id=var(user_id) **and** node_id=var(node_id)
26.   **end if**
27.   Departure=False
28.   Arrival=True
29.   **end if**
30. **end each**

When the current position coordinate is a road intersection or a user's FVL, it is recorded together with current timestamp (date and time) in the user's UMT. If the measured acceleration is smaller than a given threshold a_s, the user is deemed not moving. Thus, if this position is recorded in user's context as FVL, it is

recorded with the current date into UFVLT. Time of arrival, recorded at FVL, corresponds to the time when the acceleration falls below a_s.

Table I: User Contextual (UC) information structure.

Personal Context (PC)	Frequently Visited Location Context (FVLC)	Task Context (TC)	Interest Context (IC)	Calendar Context (CC)	Day Type (DT)
• User ID • Name • Age	• Location (NM_node) • Location name • Preferable day • Earliest preferable time • Latest preferable time • Duration • Characteristics • Importance • Frequency	• Task name • Earliest time • Deadline • Duration • Characteristics • Importance • Frequency	• Interest name • Preferable day • Earliest preferable time • Latest preferable time • Duration • Characteristics • Importance • Frequency	• Location (NM_node) • Date • Time • Characteristics	• From (date) • To (date) • Work-day • no workday

Departure time from a FVL is recorded when acceleration exceeds a_s. Note that to designate a location visited by a user as FVL, the user needs to reside at the location for a time period longer than a predefined duration threshold d_s. User Contextual (UC) information is gathered and organized in six categories as shown in Table I. The user context database may be built by having users (1) fill in a questionnaire and explicitly express their interests with regard to different places within their living area, and (2) continuously registering both their tasks and scheduled appointments.

In the following, we describe how our proposed PPM scheme predicts a path using the above described database. At a road intersection, user chooses a new road segment/intersection to continue his/her trip according to the navigation zone (Fig. 2). S and D denote the source/origin of the trip and the destination respectively.
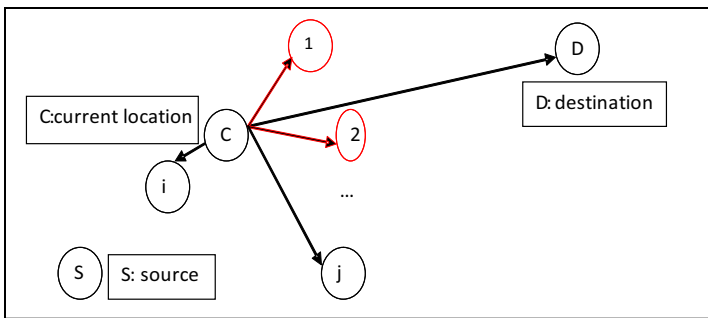

Fig. 2. next road segments: {1,2}: An examplr

Let j be an element of the set of adjacent nodes to the current node c and θ_j^c is the value of the angle formed by jcD. We apply a direction function o to select potential next road intersection/road segment.

$$o : [0,180] \rightarrow [0,1]$$

$$o(\theta) = 1 - \frac{\theta}{180}$$

Thus, we compute the degree of the deviation of current position (c) to each its adjacent positions (j) as follows:

$$o(\theta_j^c) = 1 - \frac{\theta_j^c}{180} \tag{1}$$

The selected adjacent nodes are those that belong to the following set:

$$\Psi = \bigcup_j \left\{ j \mid o(\theta_j^c) \geq \alpha \right\} \tag{2}$$

where $\alpha = o(\theta_{seuil})$ and θ_{seuil} is a predefined threshold. At position c, the next road intersection to be visited will be selected from Ψ. For better understanding, let us use the example shown in Fig. 2; using direction to the destination, we compute $\Psi = \{1, 2\}$ using Equation (1); in this case $\alpha = \frac{3}{4}$.

The selection of a node from Ψ as a next intersection, is performed using conditional probability that uses historical user movement trace (UMT) given the source and the destination of the trip. In our case, the path from the source to the current position and the destination are known. To calculate the conditional probability at day d and time of the day t_c (Timestamp), we use historical data recorded in UMT. We consider that $t = t_c \pm \Delta t$

where t_c and Δt denote the current time (when the prediction is taking place) and a given time interval respectively. To select i from Ψ as the next road intersection, we compute the weights of each position/element in Ψ (using Equation 3) and choose the position with the biggest weight.
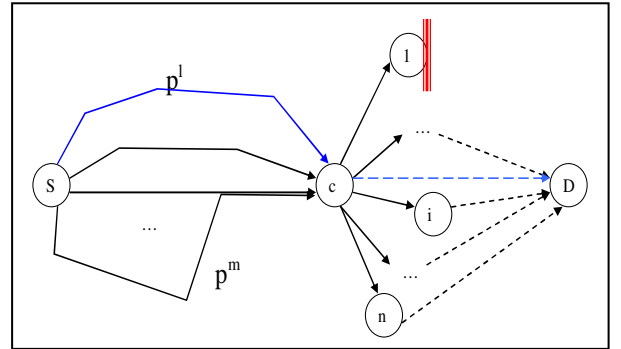

Fig. 3. Evaluation of conditional probability

The weight of i, in Ψ, is computed as follows:

$$w(i) = proba_{d,t}(i \mid p^l \xrightarrow{i} D) \times o(\theta_i^c) \times pos(i) \tag{3}$$

where

$$pos(i) = \begin{cases} 1 & if \ possible \quad to \ reach \ D \ through \ i \\ 0 & if \ non-possible \ to \ reach \ D \ through \ i \end{cases} \tag{4}$$

We made use the function o, in computing the weight, to give more weight to intersections whose directions are more oriented towards the destination. For better understanding of the computation of the weight (see Equation 3), let us consider the example shown in Fig. 3. The figure shows the different paths that can be used to reach the current position c from the source S (p^1 to p^m). For the current trip, user used path p^l to reach current position c. At this current position c, Fig. 3 shows all the options that can be used to reach the destination D; next intersection 1 is not an option (e.g., dead end road) and thus, $pos(1)=0$.

The conditional probability is defined as follows:

$$proba_{d,t}(i|p^l \xrightarrow{i} D) = \frac{proba_{d,t}(p^l \xrightarrow{i} D)}{proba_{d,t}(p^l \rightarrow D)} \qquad (5)$$

It represents the probability that i is next intersection, towards D, given that path p^l was used to reach current position c (Fig. 3).

The probability that p^l will be used to reach D is computed as follows:

$$proba_{d,t}(p^l \rightarrow D) = \frac{frequency_{d,t}(p^l \rightarrow D)}{\sum\limits_{y=1}^{m} frequency_{d,t}(p^y \rightarrow D)} \qquad (6)$$

where m denotes the total number of used paths to reach D through current position c.

The probability that i is next intersection given current position c, t, p^l was used to reach c and destination is D is computed as follows:

$$proba_{d,t}(p^l \xrightarrow{i} D) = \frac{frequency_{d,t}(p^l \xrightarrow{i} D)}{\sum\limits_{p=1}^{m}\sum\limits_{j=1}^{n} frequency_{d,t}(p^p \xrightarrow{j} D)} \qquad (7)$$

where n and m denote the total number of adjacent nodes to the current node c and the total number of used paths to reach D through current position c respectively.

III. PERFORMANCE EVALUATION

To evaluate the performance (i.e., prediction accuracy) of our proposed PPM scheme, we developed a program in java that uses the historical data recordings in the database; we use the MIT media laboratory's database available from the Reality Mining Project [16]. The subjects from this project consisted of students and staff (94 persons) at a major university during the months between September 2004 and June 2005; we used the first thirty days to run our simulation. The database contains information on the cell tower a user is connecting to and the corresponding time of residence in the cell coverage area.

We run our simulations, using only these subjects who do not reside on campus; the objective is to consider subjects with high level of movements in large geographic areas; this will allow for better evaluation of our proposed approach. To take into account different types of subjects according to the motion predictability, we defined 3 groups (not at all predictable, somewhat predictable and very predictable) and we identified ten subjects (from the database) per group for simulation. Table II lists up the values of the parameters used in the simulations.

Table II: simulation parameters.

Parameters	Fixed values	Parameters	Fixed values
d_s	15 minutes	Δt	2 hours
θ_{seuil}	90^0		

We measured the prediction accuracy for different scenarios. The prediction accuracy is computed as follows:

$$A_s = \frac{n}{\frac{(n_{est} + n_{used})}{2}} \qquad (8)$$

where n_{est} denotes the number of road segments that were returned/predicted by PPM, n_{used} the number of road segments that were actually used, and n denotes the number of common road segments (i.e., intersection between the set of predicted segments and the set of used segments).

In all the scenarios, we used ¾ of the total duration of data collection to learn users' habits (as learning phase) and the path is predicted using the source as the departure position except where the learning duration and the distance between the source and destination are used as performance metrics. The simulation results are shown in Figs. 4-7.

Fig. 4 plots the accuracy of PPM for different degrees of users' predictability. The graph shows that regardless of the degrees of predictability, the accuracy of PPM exceeds 70%. This good performance is due to our method of evaluating the weight of the next road segment to use. It is reinforced by our two ingenious techniques for the evaluation of probabilities: filtering the historical data according the type of days of week and the time of the days to compute conditional probability. For subjects with lowest predictability degree, PPM reaches an average of 81.1%. It reaches 92.0% for subjects with highest predictability and 97.9% for others subjects.
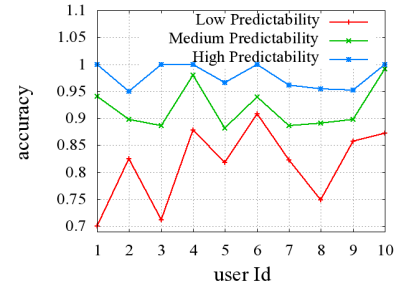

Fig. 4. Accuracy of PPM

Fig. 5 highlights the impact of learning phase duration on prediction accuracy. In this experiment, the duration of the learning phase is expressed in percentage of the total period of data collection. We observe a small increase (< 15%) in prediction accuracy for the simulated learning phase durations; this shows that the impact of the learning period duration is low. For 10%, the performance of PPM exceeds 71%. We note that our PPM needs just 60% (about 15 days) of collection data duration (historical movement data) to attain its best performance. This good performance is indeed attributable to the adoption of direction function to the destination.
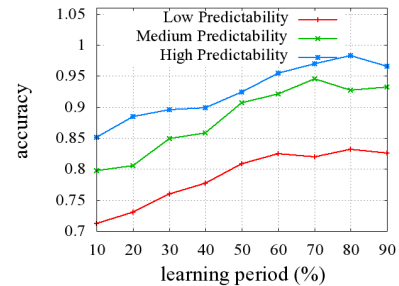

Fig. 5: Accuracy of PPM versus learning phase duration

Fig. 6 exhibits the impact of portion of the road already travelled on the performance of PPM. In this scenario, the path is computed using distinct locations as the departure position. These locations represent different percentages of the path from source to destination. Fig. 6 shows that our PPM performance is not significantly affected, especially for the subjects with highest predicta-

bility (< 3%) due, we believe, to the use of the direction function (see Equation X). For subjects with lowest predictability, the accuracy gap is 16% while it is 9% for those with intermediate predictability.
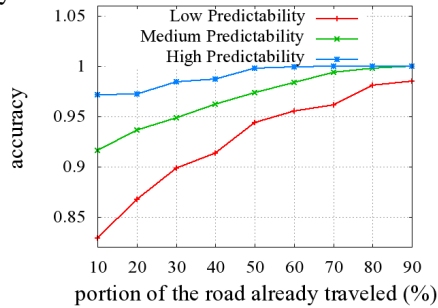

Fig. 6: Accuracy of PPM versus the remaining distance to reach the destination

Fig. 7 shows the impact of the prediction area size on the performance of PPM; the amount of cells between a source and a destination is used as the evaluation criteria (X-axis).

The performance of PPM decreases rapidly for the subjects with lowest predictability degree and very slowly for the other two predictability degree. However, it remains bigger than 45% when the distance between source and destination exceeds 20 cells. For a distance smaller than 4 cells, PPM performance exceeds 90% regardless of the degree of predictability. Knowing the destination in advance increases the performance of the subjects with highest and intermediate predictability.
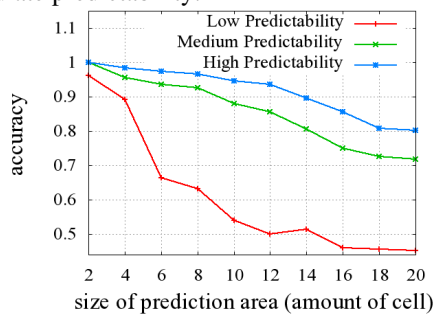

Fig. 7: accuracy of PPM versus the size of prediction area

IV.   CONCLUDING REMARKS

In this paper, we propose a new Path Prediction Model (PPM) based on historical movement trace, current movement data and spatial conceptual map when a destination is known in advance.
We use the deviation from the current position to the given destination to select a subset of possible road intersections/road segments to be visited by the subject. Based on the historical data, we evaluate the probability of each adjacent road intersections/road segment to be visited when subject is at a road intersection. This probability is based on the rule of conditional probability. Also, the historical data is filtered according to the type of days and the time of this day before its use for evaluation. Unlike existing approaches that do not use both current and historical information about movement and do not consider the time of the day and the type of the day, PPM provides good path prediction accuracy.

Simulations show that PPM is able predict paths with high accuracy. For subjects with lowest predictability degree, PPM reaches an average of 84.5%. It reaches 96.5% for subjects with highest predictability and 90.6% for subjects with intermediate predictability degree. In the future, we plan to propose a Handoff Time Estimation Model (HTEM) based on the Path Prediction Model (PPM).

REFERENCES

[1]   P. S. Prasad and P. Agrawal, "Mobility prediction for wireless network resource management," in *Proc. of the 41st Southeastern Symposium on System Theory, 2009. SSST 2009.* , Tullahoma, Tennessee, USA, March 2009, pp. 98-102.

[2]   H. Jeung*, et al.*, "Path prediction and predictive range querying in road network databases," *The VLDB Journal,* vol. 19, pp. 585-602, 2010.

[3]   A. Aljadhai and T. F. Znati, "Predictive mobility support for QoS provisioning in mobile wireless environments," *in IEEE Journal on Selected Areas in Communications,* vol. 19, pp. 1915-1930, October 2001.

[4]   F. Bai and A. Helmy, "A survey of mobility models," in *Wireless Ad Hoc and Sensor Networks, Kluwer Academic Publishers*, 2004.

[5]   J. Markoulidakis*, et al.*, "Mobility modeling in third-generation mobile telecommunications systems," *in IEEE Personal Communications,* vol. 4, pp. 41-56, August 1997.

[6]   J. Kim and A. Helmy, "Poster abstract: the challenges of accurate mobility prediction for ultra mobile users," *SIGMOBILE Mob. Comput. Commun. Rev.,* vol. 13, pp. 58-61, July 2009.

[7]   S. Michaelis and C. Wietfeld, "Comparison of User Mobility Pattern Prediction Algorithms to increase Handover Trigger Accuracy," in *Proc. of the 63rd IEEE Vehicular Technology Conference (VTC 2006-Spring)*, Melbourne,Victoria, Australia, May 2006, pp. 952-956.

[8]   I. Butun*, et al.*, "Impact of mobility prediction on the performance of Cognitive Radio networks," in *2010 Wireless Telecommunications Symposium (WTS)*, Tampa, FL, USA, April 2010, pp. 1-5.

[9]   L. Liao*, et al.*, "Learning and inferring transportation routines," *Artificial Intelligence,* vol. 171, pp. 311-331, April 2007.

[10]   W. Wanalertlak*, et al.*, "Behavior-based mobility prediction for seamless handoffs in mobile wireless networks," *Wirel. Netw.,* vol. 17, pp. 645-658, April 2011.

[11]   N. Samaan and A. Karmouch, "A mobility prediction architecture based on contextual knowledge and spatial conceptual maps," *in IEEE Transactions on Mobile Computing,* vol. 4, pp. 537-551, Nov.-Dec. 2005.

[12]   M. Musolesi and C. Mascolo, "Designing mobility models based on social network theory," *in ACM SIGMOBILE Mobile Computing and Communications Review,* vol. 11, pp. 59 - 70, July 2007.

[13]   D. Yuan*, et al.*, "Experimental analysis of user mobility pattern in mobile social networks," in *2011 IEEE Wireless Communications and Networking Conference (WCNC)*, Cancun, Quintana Roo, Mexico, March 2011, pp. 1086-1090.

[14]   F. Ekman*, et al.*, "Working day movement model," in *Proc. of the 1st ACM SIGMOBILE workshop on Mobility models for Networking Research (MobilityModels'08)*, Hong Kong, Hong Kong, China, May 2008, pp. 33–40.

[15]   Tarik Taleb*, et al.*, "Mobility-Aware Streaming Rate Recommendation System," presented at the IEEE Global Communications Conference (GLOBECOM 2011), Houston, Texas, USA, Dec. 2011.

[16]   N. Eagle*, et al.*, "Inferring Social Network Structure using Mobile Phone Data," *Proceedings of the National Academy of Sciences (PNAS),* vol. 106, pp. 15274-15278, September 2009.