Temporal-aware Location Prediction Model Using Similarity Approach

Ghulam Sarwar, Farman Ullah, Sungchang Lee
Department of Information and Communication,
Korea Aerospace University, Goyang, South Korea,
{sarwar,farman,sclee}@kau.ac.kr}

***Abstract*—We propose scalable next location prediction model using temporal, movement behavior and trajectory history information. Most of the existing models use the whole information database and do not consider temporal information for location prediction; lead to scalability and accuracy issues. We consider road topology, multiple predictive factors and temporal information such as day and time to improve the accuracy and scalability issues. We extend the existing Mobility Markov Chain (MMC) model to incorporate *n* previously visited locations have movement factors m_k in the time interval *t* of the day that we coined *nmt-MMC*. We use the whole information database in making the model in offline, and find similarity with the model for location and trajectory prediction in online. Simulation results demonstrate the trade-off between scalability and accuracy, and the effect of Top-N similar trajectories on the accuracy. We found that incorporating movement predictive factors improves the accuracy by approximately 7-percent, and adding both predictive factors and temporal information improve the accuracy by 14-percent.**

Keywords- Temporal-aware; Location Prediction; Trajectory Estimation; Movement Predictive Factors

I. INTRODUCTION

The recent convergence of multimedia, communication networks, location-aware clients and geo-processing has given rise to a new generation of Location-Based Services (LBSs) [4]. Movement of mobile users is the main reason that leads to higher complexity of designing mobile network services. Mobile communication systems need to implement a series of operations such as location management, call admission and resource allocation, adaptive service management etc when a user is moving. The traditional system becomes aware of the movement when the QoS of the user cannot be satisfied. Although a passive approach can reduce complexity of the system design but leads to signaling overhead [4], which is not tolerable in delay sensitive, services such as streaming media. Hence mechanisms for obtaining predictive mobility information have been proposed, mostly having scalability and accuracy issues.

Location prediction provides a longer time available for service preparation and presentation and is needed for delay sensitive multimedia streaming and multi-user session management [5, 6]. However, there are many difficulties in performing active location prediction and management. These include: 1) user movement behavior is random. 2) Although mobile behavior shows some regularity on the basis of vehicle movement, but there is no mature analysis of these characteristics. 3) Adding a terrain feature to the prediction model almost linearly increases the complexity of the model. 4) The existing location prediction mechanisms are mostly based on Bayesian network theory and Markov chains, but research did not focus on similarity-based location prediction which may be a good alternative for simple location prediction.

Considering the mentioned issues, we proposed a new location prediction model based on similarity approach. Similarity based approaches have widely been used for prediction and recommendation in recommender systems. In previous works [2, 3], we proposed scalable hybrid model for online recommender systems. Knowledge about user's location is the basic requirement for location prediction and can be obtained from Global Positioning System (GPS) enabled devices [13]. In this paper, we predict the future location of a user based on observation of his movement behavior and recent trajectory history. In order to enhance the accuracy of location prediction model, we integrate temporal and movement factors to existing MMC model. We model the location prediction database in offline, and find similarity of movement factors and trajectory of a user with the model to predict the next edge at each junction in the network. Based on the predicted next edge, we find trajectory of a user after the time interval T. We use the terms next-location and location interchangeably in this paper.

We organize the paper as follows. Section II discusses related work. We present the system design and location prediction model database (LPM-DB) in section III. Section IV describes the Location prediction model and trajectory estimation. Section V is about the evaluation and simulation of the proposed model, and section VI concludes the paper.

II. RELATED WORKS

Location prediction model estimates the next or future location of a user based on his current location, trajectory history and some movement predictive factors. Current location prediction strategies can be classified into four approaches: cell-based approach, map-based approach, memory-based approach and model-based approach.

Cell-based approaches [1] divide the geographic area of mobile network into cells and determine the location of a user through Paging. *Lui et al.* [1] proposed a method recording previously visited cell by a user to predict future

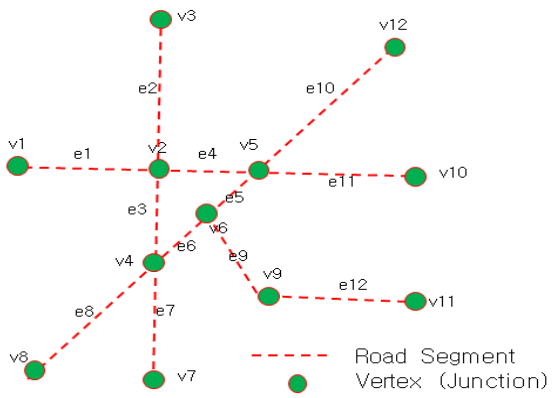

Figure 1. A Simple Road Topology 'G'

$$
\begin{array}{c}
 \\ v1 \\ v2 \\ v3 \\ v4 \\ v5 \\ v6 \\ v7 \\ v8 \\ v9 \\ v10 \\ v11 \\ v12
\end{array}
\begin{array}{c}
\begin{array}{cccccccccccc} v1 & v2 & v3 & V4 & V5 & V6 & V7 & V8 & V9 & V10 & V11 & V12 \end{array} \\
\begin{bmatrix}
0 & e1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
e1 & 0 & e2 & e3 & e4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & e2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & e3 & 0 & 0 & 0 & e6 & e7 & e8 & 0 & 0 & 0 & 0 \\
0 & e4 & 0 & 0 & 0 & e5 & 0 & 0 & 0 & e11 & 0 & e10 \\
0 & 0 & 0 & e6 & e5 & 0 & 0 & 0 & e9 & 0 & 0 & 0 \\
0 & 0 & 0 & e7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & e8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & e9 & 0 & 0 & 0 & 0 & e12 & 0 \\
0 & 0 & 0 & 0 & e11 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & e12 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & e10 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\end{array}
$$

Figure 2. Adjacency Matrix of Road Topology 'G'

location. Although cell-based approaches have played an important role in mobile networks, but they have some inherent shortcomings. They cannot precisely locate a mobile user as the radius of a cell may vary approximately from 150-30000 meters and do not support fine granularity.

On the contrary, map-based approaches [14, 9] determine user location as a point on road using GPS. Reference [14] uses trajectory history of a user for location. Furthermore, it assigns equal probability for a user visiting a location for the first time. Reference [9] collects user choices at different sections of the road and then transforms the information into transition probabilities for future location prediction. Map-based approach can obtain higher accuracy than cell-based approach, but both approaches have poor accuracy when the user mobility behavior is unusual. To resolve the mentioned issues, ref. [8] added predictive factors of user movement behavior but do not consider trajectory history information which has a significant impact on the accuracy. In [10] the authors consider temporal information for the next place prediction but do not consider road network and movement factors. Memory-based approaches use recently visited location history for recommendation [4], and model-based approaches model the whole database and predict future location by comparing current parameters with the model [3, 5]. Ref. [7] considers previously visited location of a user but does not consider the movement factors. Markov chains have widely been used in location prediction [7, 8].

In this paper, we propose a scalable and accurate location prediction model integrating temporal information, multiple prediction factors. We model the whole database in offline to maintain the *nmt*-MMC matrix, and use cosine similarity approach to predict next location and trajectory after time T. Simulation results depict that the proposed

III. System Design And LPM-DB

Before presenting the system design and location prediction model database (LPM-DB), we first give some formal definitions. In a mobile environment, information about a user's location includes two parameters, L and t, indicating that the user is at location L at time t. Prediction period T is the time duration between the current time t and latter time t+1. A trajectory is a sequence of connected road segments. The proposed system maintains the previous 'n' road edges visited by the user with their corresponding movement predictive factors and temporal information. Given L, t, m_k and T, the problem of location prediction is to estimate location L+1 and m_k+1 at time t+1 or t+T.

A) *Architecture:*

Among different possible architecture, we can simply consider a client/server architecture having mobile clients equipped with GPS and fixed servers. Furthermore, we consider that there is no mobile to mobile communications. The clients provide required information to the server at the start of every time period for resource and service management.

B) *Location prediction model database:*

The LPM-DB maintains two categories of data. One category includes the road network. The second category includes the historical trajectories information of a user along with their predictive movement factors such as velocity, angular velocity, and acceleration etc and temporal information such as time and day.

1) Road Topology & Adjacency matrix: A road topology consists of edges 'E' and vertices 'V' where edges are road segments and vertices are intersection as shown in Fig. 1, and can be modeled as graph G<V, E>. An edge may have multiple attributes such as length, width, speed limit, direction etc. Fig. 2 shows the adjacency matrix for road the road network.

ndx	Edge t-n (Vel,AC,Aw)	…	Edge t (Vel,AC,Aw)	Edge T+1	Time Ti	Day Di
1	e11(8,12,13)	…	e12(12,32,24)	e6	T1	D1
…	…	…	…	…	….	…
n	e5(35,23,23)	…	e16(23,23,21)	e2	T4	D2

T1=7-9am	T3=5-8pm
T2=9am-5pm	T4= 8pm-7am
D1= Week days,	D2=Weekend

Figure 3. Adjacency Matrix of Road Topology 'G'

2) *Historical Trajectories with Movement Factors and Temporal Information:* A MMC can be represented either as a graph or a transition matrix. In matrix representation, a row represents the current edge of a user and a column represents the destination edge. Standard MMCs are memoryless, and prediction of next edge/location only depends on the current edge/location. The memoryless property of MMC has a negative impact on the accuracy. To address the issue, ref [7] presents the concept of using the n previously visited location for prediction. However, the authors did not consider road network, temporal information and multiple movement predictive factors that can further improve the accuracy of future location prediction. We extent the existing MMC model to nmt-MMC by integrating the previously n visited road segments, movement factors and temporal information to improve the accuracy of the location prediction model as shown in Fig 3. We cluster the trajectories based on their temporal information in offline, and find trajectories and movement factors similarities in online to improve the scalability of the system.

IV. Location Prediction Model

This section describes the next location and trajectory estimation model. LPM utilizes information of the adjacency matrix of the road network G, historical trajectories of a user with temporal and movement behavior information, time period T and current location to perform prediction. Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

A) *Sojourn Time in next Junction*

The average sojourn time 'S_t' of a user 'u' in junction 'j' on edge 'e' in the time interval 't' of day 'd' can be calculated as in (1).

$$S_T(u,j,e,t,d) = \frac{\sum_{i=1}^{n} d_i(T_{left} - T_{reached})}{n} \quad (1)$$

B) *Next edge after the Junction*

First of all, we determine the cluster to which the user belong. Then, we use cosine similarity approach to select Top-N similar trajectories from

the cluster based trajectory and movement similarity as given in (2) and (3).

$$Comb_{sim} = Traj_{sim} + Movement_{sim} \quad (2)$$

$$sim(u,h,t,d) = \frac{\sum_{i=i}^{n}\sum_{k=1}^{K}(V_{uktd}) * (V_{hktd})}{\sqrt{\sum_{i=1}^{n}\sum_{k=1}^{K}(V_{uktd})^2 * \sum_{i=1}^{n}\sum_{k=1}^{K}(V_{hktd})^2}} \quad (3)$$

Where 'n' is the number of previous edges, 'k' is the number of predictive factors, 'u' is the current trajectory, 'h' is historical trajectory, and 't' and 'd' are temporal information. After selecting Top-N similar trajectories, we consider the

1. Initialize $T_k = T$; % T is the time period
2. Calculate:
   a. t_i = time to reach next junction using equation (7).
      % next can be find from V to V transition/adjacency matrix
   a. S_t = average sojourn time in next junction using equation (1)
3. If ($t_i > T_k$)
   a. Calculate distance covered in T_K (as discussed in section-4.3)
   b. User location after time T can be calculated as
      Current position +distance covered in time T_K
   a. Exit
   else
   a. $T1 = t_i + S_t$
   b. If ($T1 > T_k$)
      1) Calculate user location as in step 3-b
      2) exit
      else
      1) $T_k = T_k - T1$
      2) Predict next edge using equation (2 & 3)
      3) Consider start of predicted edge as current position
      4) Goto step 2

Figure 4. Algorithm for next location prediction after time T

mode of the next edge of the similar trajectories as the prediction.

C) *Time to reach next junction*

Knowing the current position of a user, his next junction can be determined from the adjacency matrix of the road network G. We use 'Haversine Formula' [11] to determine the distance between current location and next edges as given in the following formulas.

$$a = \sin^2\left(\frac{\Delta\emptyset}{2}\right) + \cos(\emptyset_1) * \cos(\emptyset_2) * \sin^2\left(\frac{\Delta\delta}{2}\right) \quad (4)$$

$$c = 2 * atan2\left(\sqrt{a}, \sqrt{(1-a)}\right) \quad (5)$$

$$dist = R * c \quad (6)$$

$$Time_{nj} = \frac{dist}{Avg\_Vel} \quad (7)$$

φ = latitude. σ = longitude (in radians) and R = radius of earth = 3171 km. The average velocity of a user can be calculated from the trajectories database as ratio of road length to the average time taken.

D) *Algorithm for Location after time T*

Fig.4 shows the algorithm for next location prediction after time T.

I. SIMULATION RESULTS

We performed simulation in MATLAB in order to evaluate the performance of the proposed temporal aware location prediction model. We used Geolife-dataset consist

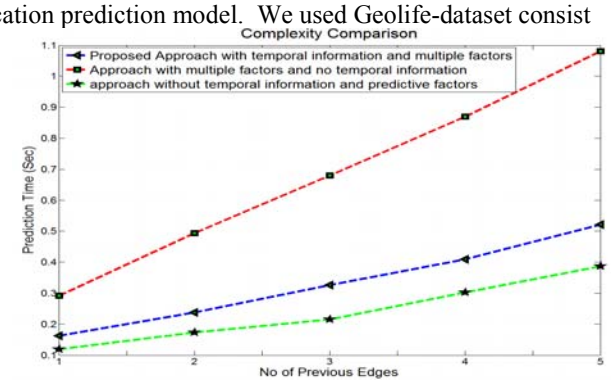

Figure 5. Complexity Comparison

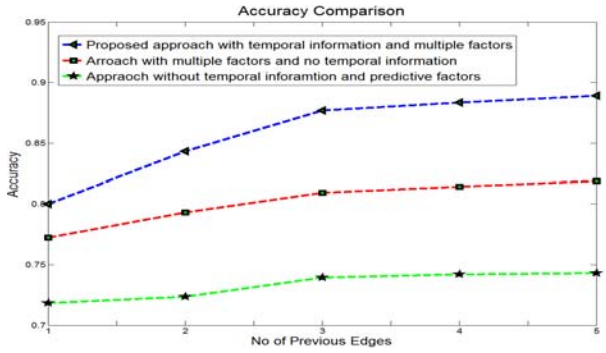

Figure 6. Accuracy Comparison

of mobility traces collected from 2007 to 2011 using GPS-enabled devices [12]. The dataset includes 182 users. We select some users, visualized their trajectories and estimate road junctions and movement behavior. We compared the model with an existing approach which does not consider temporal information and movement factors. Fig. 5 and Fig. 6 show the trade-off between the complexity and accuracy of the proposed temporal aware location prediction model with existing approaches.

IV. CONCLUSION

In this paper, we proposed temporal aware location prediction model using similarity approach for the next location and trajectory estimation. The proposed system integrates temporal information, multiple movement predictive factors, and previous trajectories to improve the accuracy of location prediction model. We compared the approach with an existing approach using Geolife dataset. The simulation results depict that using only movement factors improve model accuracy by 7-percent, and incorporating both the temporal and movement factors improve accuracy of the prediction model by 14-percent. We used Top-N similar trajectories to recommend next edge in the road network. The results predict that considering more than three similar trajectories for recommendation has a negative impact on the accuracy. However, such observation depend historical trajectories in the dataset. Moreover, increasing the value of previous trajectories for prediction above 3 does not seem to bring improvement due to the cost

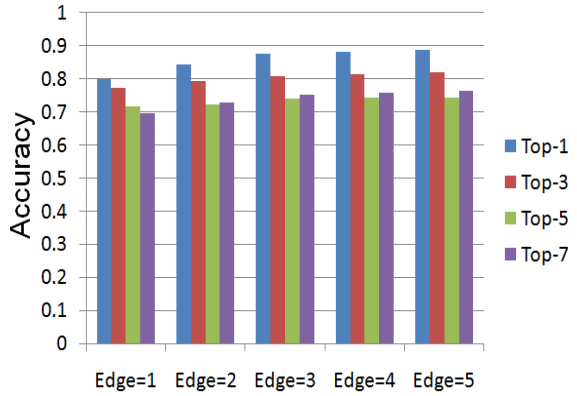

Figure 7. Accuracy Comparison

of overhead in terms of computation and storage requirements an existing approach using Geolife dataset.

REFERENCES

[1] Liu. G.Y., Maguire. G.Q, and Jr., "A predictive mobility management algorithm for wireless mobile computing and communications," Universal Personal Communications. 1995. Record., 1995 Fourth IEEE International Conference on, IEEE, pages 268--272, 1905.

[2] F. Ullah et al., "Hybrid recommender system with temporal information," Information Networking (ICOIN), 2012 International Conference on, pages 421--425, 2012.

[3] F. Ullah, G. Sarwar, and S. Lee, "N-Screen Smart TV Contents Personalization System using Hybrid Adaptive Similarity Approach," ICUFN, 2013, IEEE International Conference on, 2013.

[4] Stojanovic. D.H, Djordjevic. Kajan, and S. J., "Developing location-based services from a GIS perspective," Telecommunications in Modern Satellite, Cable and Broadcasting Service, 2001. TELSIKS 2001. 5th International Conference on, pages 459--462, 2001.

[5] Kang-won. Lee, and Bong-jun ko, "Adaptive Server Selection for Large Scale Interactive Online Games," ACM Int'l Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV}, pages 152--157, 2004.

[6] G. Sarwar, F. Ullah, C. Yoon, and Sungchang Lee, "Control Architecture for N-Screen Based Interactive Multi-Vision System," Journal of the IEEK TC, vol. 50, N0. 6, June 2013, 72-81.

[7] Sebastien at al., "Next Place Prediction using Mobility Markov Chains," Proceedings of the First Workshop on Measurement, Privacy, and Mobility, series = {MPM '12}.

[8] Yucheng Zhang and et al., "Location Prediction Model Based on Bayesian Network Theory," Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE, pages 1--6, 2009.

[9] Wee-Seg Soh et al., "Dynamic Bandwidth Reservation in Cellular Networks Using Road Topology Based Mobility Predictions,"

[10] Gatmir Motahari et al., "Time-Clustering-Based Place Prediction for Wireless Subscribers," Networking, IEEE/ACM Transactions on (Volume:PP , Issue: 99), 2012.

[11] http://mathforum.org/library/drmath/view.51879.html

[12] Y. Zhou et al., "Understanding mobility based on GPS data," 10th international conference on Ubiquitous Computing, pages 312--321, 2008.

[13] P. bahlr, A. balachandran, and at el., "Satisfying the Need of Ubiquitous Connectivity and Location Services," IEEE Personal Communication Magazine (PCS), Vol 6}, 2001.

[14] Hassan A. Karimi, and Xiong Liu, "A Predictive Location Model for Location-Based Services," ACM GIS'03 Conference, 2003