KỶ YẾU HỘI NGHỊ KHOA HỌC CÔNG NGHỆ
QUỐC GIA LẦN THỨ VI

FAIR

NGHIÊN CỨU CƠ BẢN VÀ ỨNG DỤNG CÔNG NGHỆ THÔNG TIN

THỪA THIÊN – HUẾ, 20 – 21 tháng 6 năm 2013

Proceedings of the 6th National Conference on Fundamental and Applied Information Technology Research (FAIR)

NHÀ XUẤT BẢN KHOA HỌC TỰ NHIÊN VÀ CÔNG NGHỆ
Hà Nội 2013

MỤC LỤC

A MOBILITY PREDICTION MODEL BASED ON GROUP BEHAVIORS IN WIRELESS NETWORKS

Thuy Van T.Duong[1], Dinh Que Tran[2], Cong Hung Tran[3]

[1]Department of Information Technology,University of Ton Duc Thang, Ho Chi Minh City

[2]Department of Information Technology, Post and Telecommunication Institute of Technology, Ha Noi

[3]Department of Information Technology, Post and Telecommunication Institute of Technology, Ho Chi Minh City

dttvan@it.tdt.edu.vn, vanduongthuy@yahoo.com

Abstract. *Mobility prediction of users has played a crucial role in the resource management of wireless networks and it has attracted recently a great deal of research interests. However, since most of the current approaches are based on user movement profiles for predicting the next location of mobile users, these techniques may fail to make a prediction for new users or ones with movements on novel paths. In this paper, we propose a prediction model which is based on group mobility behaviors to deal with such the lack of information of individual movement histories. Our proposed prediction approach makes use of clustering techniques in data mining to classify mobility patterns of users into groups, where the mobile users within the same movement group may exhibit similar movement characteristics. Towards this goal, we first determine which movement group the user's current paths belong to and then make use of mobility rules derived from that group to predict the next location for the user. Experiments will be performed to demonstrate that using group mobility behaviors may avoid the lack of extracted mobility rules due to the incompleteness of information on individual movement histories and significantly enhance the accuracy of the mobility prediction.*

Keywords: *movement group, mobility pattern, mobile user, prediction, wireless network*

I. INTRODUCTION

In the recent decades, wireless networks have been widely used due to their convenience, cost efficiency, and ease of integration with other networks. However, the dynamic relocation of mobile users leads to problems of location management and network resource allocation. Among the proposed solutions, mobility prediction has received a great deal of attention since it may detect user location for providing effectively network resources [1- 9]. It means that by using user mobility information, network resources may be reserved prior to the users entering. Mobility prediction has been proposed as a means of enhancing performance including: (*i*) providing flexible usage of limited resource in WLANs due to accurate estimation of resource demands at future time; (*ii*) reducing the handover latency due to resource allocation prior to the actual handover; (*iii*) supporting an effective Call Admission Control scheme to reduce the new Call Blocking Probability and the Handoff Call Dropping Probability.

Unlike most previous approaches [1-3, 8- 9], which have utilized personal profile to predict the next locations visited by a mobile user, our premise is that applying group mobility behaviors to facilitate the predicting in the case of a new user or with movements on novel paths. In reality, even though human movement and mobility patterns have a high degree of freedom and variation, they also exhibit mobility behaviors in groups due to geographic, social and friendship constraints [10, 11]. For instance, mobile users in a university campus network can be categorized as students, graduate students, departmental staffs, lecturers and so on. Students often move to classrooms, laboratories, library, etc whereas departmental staffs spend most of the day near the administrative offices. In order to fully exploit the similar movement characteristics, it is crucial to determine the movement groups, where the mobile users belong to the same group have the same mobility behaviors. In other words, the group mobility behavior reflects the fact that mobile users often behave as groups. Discovering such group mobility behaviors is a key issue for predicting future movement of a new mobile user based on the mobility behaviors of members in a group.

In this paper, we propose a four-phase mobility prediction technique which is able to deal with the lack of information on personal profile. We first make use of data mining techniques to discover frequent mobility patterns (or mobility patterns for short) from personal profile of all mobile users in the coverage region. The second phase is to cluster mobility patterns of all mobile users into movement groups. The third phase determines which group the user's current path belongs to. And then the fourth phase makes use of mobility rules derived from that group to predict the next location for the user. The efficacy of the proposed approach is verified on the synthetic dataset and experimental results show that the group mobility behavior may significantly enhance the accuracy of mobility prediction.

The remainder of this paper is constructed as follows. In Section II, some mobility prediction approaches are introduced. Section III presents a novel model of mobility prediction based on the group mobility behavior. Section IV is devoted to describing experiments and results for evaluating the proposed mobility prediction approach. Finally, Section V draws concluding remarks and further research work.

II. RELATED WORKS

The importance of mobility prediction in improving the performance of location management, smooth handoffs and call admission control has attracted a great deal of interests. Various mobility prediction techniques [1-3, 6-9] have been proposed and most of them are based on individual mobility behaviors which are discovered from personal movement profile. One of the most effective approaches has been proposed by Y. Manolopoulos et. al. [1] which has utilized personal movement history for predicting future movements of mobile users in a PCS (Personal

Communication Systems) network. After mining all frequent mobility patterns from user movement history, the authors extract mobility rules and utilize this knowledge for prediction. According to simulation results, moderate prediction accuracy was achieved, decreasing only minimally with an increase in random mobility. Another approach that has been also based on personal movement profile is presented in [2]. Juyoung Kang and Hwan-Seung Yong [2] have utilized line simplification and clustering for discovering all spatio-temporal patterns from historical trajectories of mobile user. Due to using frequent trajectory patterns in prediction, this approach can forecast accurate locations in the distant future. In [3], the authors have proposed a Moving Pattern mining algorithm to extract frequent moving patterns from individual user movement history for predicting future movement. They firstly represent the spatial location of mobile users in coordinate system with x- and y-axis and then utilize the spatial operation to transform spatial location into area for discovering significant information. After that, they apply time constraints between locations of moving objects to make valid moving sequences. Finally, the frequent moving patterns are extracted from the generated moving sequences and used for predicting. The advantage of these above techniques is that useful knowledge hidden under mobility history of one mobile user will be fully discovered to predict his future movement. However, these approaches may fail to make a prediction for a new user or with movements on novel paths due to the lack of information on personal movement profile.

In order to deal with the problem, some approaches [4, 5], which have been based on group mobility behaviors of mobile users, have been proposed. The basic idea of these approaches is to utilize mobility behaviors in the same movement group for predicting future movement of members in the corresponding group. In [4], Chih-Chien Hung and Wen-Chih Peng have proposed a new approach for predicting future movements based on group movement behaviors. They first utilize PST (Probabilistic Suffix Trees) to capture moving patterns of mobile users. Second, they define a distance function to determine the similarity between two PSTs and then propose a clustering algorithm to partition mobile users with similar moving behaviors into groups using the distance function. Third, for each group, they select one representative PST to predict future movements of users in the same group. The advantage of this approach is that the storage cost is decreased. Another approach, which also considers group of mobile users in mobility prediction, is presented by S. Avasthi and A. Dwivedi [5]. They have first cluster the mobile transaction sequences using the LBS-Alignment to evaluate the similarity of sequences. Second, a time segmentation method is introduced to find the most suitable time intervals. After clustering and segmentation, a user cluster table and a time interval table are generated. Third, the CTMSPs (Cluster-based Temporal Mobile Sequential Patterns) are mined from the mobile transaction database according to the user cluster table and time interval table. Finally, the appropriate CTMSP, which is from the corresponding cluster a user belongs to, is selected to predict future movement of the user. These above approaches focus on group mobility behaviors but do not really pay attention to elimination random movements of mobile users before clustering mobility behaviors into groups. This may affect the quality of clustering process and thus mobility behaviors in groups may be not good for predicting future movements. In conclusion, the drawback of these approaches is that the accuracy of mobility prediction may be affected due to interference from random movements of mobile users.

Motivated by the limitations of the above approaches, our paper introduces a novel mobility prediction model which utilizes mobility rules discovered from multiple similar users. This approach is not only able to deal with the lack of information on personal profile but also able to avoid interference from random movement of mobile users due to discovering frequent mobility patterns.

III. MOBILITY PREDICTION BASED ON GROUP MOBILITY BEHAVIORS

A. Discovering Mobility Patterns

Even though human movements have a high degree of freedom, no one move randomly all day and every user's movement is based on some regular habits. It's easy to see that the mobility prediction can only work well with regular movements and its accuracy degrades with increasing random movements. Therefore, random movements should be eliminated from the trajectory database as much as possible. This phase takes responsibility for mining frequent mobility patterns (or mobility patterns for short) from trajectory databases of all mobile users in the coverage region. The algorithm for discovering all mobility patterns is the modified version of the Apriori technique [12, 13]. We have used spatial and temporal constraints in order to reduce the number of generated candidates. That is, the mobility pattern and the candidate derived from it must satisfy the neighbor constraint and ascending time constraint. The detail of the mining algorithm has been presented in our previous work [9].

B. Clustering Mobility Patterns into groups

Our purpose in this work is to discover group mobility behaviors of a population of mobile users for predicting future movement of individuals, who are new users or ones with movements on novel paths. To get such group information of mobility behaviors, we first collect mobility patterns of all mobile users and then classify them into groups of similar behaviors. The subsection presents briefly our proposed similarity measure [14], which is the basis for constructing the clustering procedure to identify groups of similar mobility patterns. As presenting in previous work [9], a mobility pattern is defined as a sequence of $<(c_1, t_1), (c_2, t_2), \ldots, (c_n, t_n)>$, where c_i denotes the ID number of the cell to which the user enters at time t_i. It is clear that two consecutive ID numbers c_i and c_{i+1} of a mobility pattern must be the ID numbers of two neighbor cells in the coverage region. The ascending order of elements of the mobility pattern is sorted by using time t as the key.

Suppose that given two mobility patterns $P_a = <(c_{a1}, t_{a1}), (c_{a2}, t_{a2}), \ldots, (c_{an}, t_{an})>$ and $P_b = <(c_{b1}, t_{b1}), (c_{b2}, t_{b2}), \ldots, (c_{bm}, t_{bm})>$.

Definition 1. Let f: S × S → R be a function representing the number of uncommon cells in two mobility patterns P_a and P_b. Then, f is determined by the formula:

$$f(P_a, P_b) = card(\{c_{ai} \mid c_{ai} \notin P_b\}) + card(\{c_{bi} \mid c_{bi} \notin P_a\}) \tag{2.1}$$

The spatial similarity measure can be defined in terms of spatial dissimilarity between two mobility patterns. The more uncommon cells there are in two patterns, the more spatially dissimilar they are.

Definition 2. The spatial similarity measure $D_{space}(P_a, P_b)$ between two mobility patterns P_a and P_b with length n and m, respectively, is defined as follows:

$$D_{space}(P_a, P_b) = \frac{f(P_a, P_b)}{n+m} \tag{2.2}$$

It is clear that if two mobility patterns are the same, the result of $f(P_a, P_a)$ is zero and thus $D_{space}(P_a, P_b) = 0$. Conversely, if the two mobility patterns do not share any cells, the number of uncommon cells in these patterns is $f(P_a, P_b) = (n + m)$, and thus $D_{space}(P_a, P_b) = 1$.

For determining the temporal dissimilarity between the two mobility patterns, we need to calculate the total of temporal difference between the timestamps of the common cells in two patterns. The smaller the total time difference is, the more temporally similar the two patterns are. These statements are formalized as follows:

Definition 3. The temporal similarity measure $D_{time}(P_a, P_b)$ between two patterns P_a and P_b with length n and m, respectively, is given by

$$D_{time}(P_a, P_b) = \frac{1}{k}\sum_{i=1,j=1}^{n,m} \frac{|t_{ai} - t_{bj}|}{\max(t_{ai}, t_{bj})} \quad \text{where } c_{ai} = c_{bj} \tag{2.3}$$

where k is the number of common cells of P_a and P_b.

In order to fully exploit the characteristics of mobility patterns, it is crucial for weighted combination of two similarity measures on space and time.

Definition 4. The composition similarity measure between two patterns P_a and P_b is given by

$$D(P_a, P_b) = W_{space} . D_{space}(P_a, P_b) + W_{time} . D_{time}(P_a, P_b) \tag{2.4}$$

in which W_{space} and W_{time} are respectively weights of spatial and temporal similarity measures, such that $W_{space} + W_{time} = 1$.

The purpose of this phase is to classify the set of mobility patterns into groups such that patterns within the same group have a high degree of similarity, whereas patterns belong to different groups have a high degree of dissimilarity. K. Huang et al. [15] and the other works [16] [17] [18] have shown that comparing with conventional approaches, the k-means clustering technique offers several key advantages and most importantly it can be applied to forecast the future. Thus, it has become the most popular clustering algorithm in scientific and industrial applications with large data sets. In order to take advantages of the simplicity and the computational speed, we extend the traditional k-means approach to partition the set of mobility patterns into k clusters. Our alternative clustering algorithm focuses on applying our proposed similarity measures for finding cluster centers and assigning mobility patterns to clusters. The detail of the clustering procedure is presented in **Algorithm 1**. After random initialization for k cluster centers (line 2-4), each pattern in the set of mobility patterns is assigned to the nearest cluster based on the proposed similarity measure (line 5). Then, the center of each cluster will be updated (line 7-9) and further each mobility pattern will be reassigned to the nearest cluster (line 10) by also using our similarity measures. The processes of center updating and pattern reassigning are repeated until no mobility pattern has changed clusters via a test cycle of the whole set of mobility patterns (line 6-11).

C. Determining the movement group of a path

The previous phase takes responsibility for discovering movement groups such that mobile users in the same group have the same movement behaviors. Finding such grouping information of mobile users, based on the similarity among their movement behaviors, play a crucial role in predicting future movement to deal with the lack of mobility behaviors of individuals. Hence, for predicting the future locations of a mobile user, we first determine which movement group his current path belong to and then using mobility rules derived from that group to forecast. Since each cluster of mobility patterns is represented by a cluster center, the movement group of a current path is the cluster such that the similarity between its center and the path is the largest. The process of group determining is outlined as follows:

1. Calculating the similarity between the current path P and each cluster center C_i, $D(P, C_i)$;

2. Choosing C_i such that $D(P, C_i)$ is minimized;

3. Returning the cluster that is represented by C_i.

The detail of the group determining procedure is presented in **Algorithm 2**.

Algorithm 1 Clustering Mobility Patterns

Input: the set of mobility patterns, $L = \{P_1, P_2, \ldots, P_n\}$
Output: k clusters, $C = \{C_1, C_2, \ldots, C_k\}$

```
flag = false;
for each cluster center Ci in C
    Ci = Random(L)
end for
//finding the cluster center X for each mobility pattern Pi such that among cluster centers in C, the similarity
//between X and Pi is the largest and assigning Pi to the cluster of X
flag = Assignation(C,L)
while flag
    for each cluster center Ci in C
        //the new cluster center is the mobility pattern in the cluster such that the sum of all the similarity
        //between it and each remaining pattern in the cluster is maximum.
        Ci = Update(Ci,L)
    end for
    flag = Assignation(C,L)
end while
return C
```

Algorithm 2 Determing Movement Group()

Input: the current path, P
k clusters, $C = \{C_1, C_2, \ldots, C_k\}$
Output: the cluster center, C_i

```
// array S = {S1, S2, ..., Sk}to stores the similarity between P and each cluster center Ci
for each cluster center Ci in C
    Si = D(P,Ci)
end for
//index = Min(Si)
index = 1
i = 2
while i <= k
    if Sindex > Si then
        index = i
    end if
    i = i + 1
end while
return Cindex
```

D. Using mobility rules in groups to predict future movement

Due to geographic and friendship constraints, human movements express similar movement characteristics. It is clear that with good knowledge of groups to which a mobile user belongs, we can derive common behaviors among objects during their moving. Therefore, this phase focuses on discovering mobility rules from multiple users whose mobility behaviors are similar to facilitate the predicting future movement of a new user or with movements on novel paths.

In order to address the issue, this phase is first interested in generating a set of mobility rules from the movement group which the current path belongs to. Each generated rule has a confidence value and only rules which have confidence values higher than the predefined confidence threshold $conf_{min}$ are selected for predicting future locations of the current path. The process of generating mobility rules (line 3 in **Algorithm 3**) has been presented in our previous work [9]. Then, we match the current path to the head of rule and find the best matching using the time constraint (line 5-12).

Algorithm 3 Predicting Future Locations()

Input: the current path of the user, $P = <(c_1, t_1), (c_2, t_2), \ldots, (c_{i-1}, t_{i-1})>$
k clusters, $C = \{C_1, C_2, \ldots, C_k\}$
minimum confidence threshold, $conf_{min}$
Maximum predictions made each time, m

Output: Set of predicted cells, *Pcells*

```
PCells = ∅ // Initially the set of predicted cells is empty
C_index = GroupDetermination(P,C)
R = RulesGeneration(C_index, conf_min)
l = 0
for each rule r: <(a_1,t'_1), (a_2,t'_2) ..., (a_j,t'_j)> → <(a_j+1,t'_j+1), ..., (a_t,t'_t)> in R
    // check all the rules in R find the set of matching rules
    if <(a_1,t'_1), (a_2,t'_2) ..., (a_j,t'_j)> is contained by
                     P = <(c_1, t_1), (c_2, t_2),..., (c_i-1, t_i-1)> and a_j = c_i-1
        r.score = ScoreCalculation(r)
        MatchingRules = MatchingRules ∪ r
        //Add the (a_j+1, r.score) tuple to the Tuples array
        TupleArray[l] = (a_j+1, r.score)
        l = l+1
    end if
end for
// Now sort the Tuples array in descending order
TupleArray = Sort(TupleArray)
l = 0
// Select the first m elements of the Tuples array
while l < m && l < TupleArray.length
    PCells = PCells ∪ TupleArray[l]
    l = l+1
end while
return PCells
```

IV. EXPERIMENTAL EVALUATION

In this paper, we also use two following measures [9] to analyze the predicting results:

• Recall: The number of correct predictions divided by the total number of requests. The recall counts the "no prediction" case as an incorrect prediction.

• Precision: The number of correct predictions divided by the total number of made predictions. The precision skips the "no prediction" case.

A. Generating Synthetic Dataset

In the scope of this paper, our experiments are conducted with a dataset generator which describes the movement behaviors of mobile users around the coverage region as in **Figure 1**. At first, we manually generate 5 movement behaviors which are represented in sequences, called the set of initialized patterns. For each movement behavior, approximately 100 trajectories are automatically generated by following procedure. Each trajectory is a movement around the graph G in Figure 1. So, trajectories are generated by inserting some points $q_i=(c_i,t_i)$ into the movement behavior which satisfy two constraints. The first one is the network constraint, i.e., consecutive cells in the same trajectory are neighbor in graph G and the second one is the temporal constraint which satisfies the ascending order of timestamps in trajectories. All of 100 generated trajectories are assigned the same label. Repeating the procedure for each movement behavior, we obtain a trajectory dataset which consists of 500 trajectories and is called *DS500*. We can generate different datasets by using different sets of initialized movement behaviors.

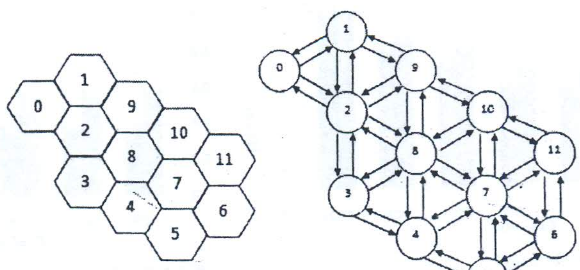

Fig. 1. The coverage region (a) and the corresponding directed graph G (b)

B. Evaluating the Proposed Prediction Model

This section is devoted to studying how effective the proposed mobility prediction approach is. We first mine the trajectory dataset *DS500* to discover all mobility patterns which have support values higher than the predefined minimum support threshold. Second, all mobility patterns are clustered into *k* groups and then generating mobility rules for each group of mobility patterns. For evaluating, we use a test set *TS10* which consists of 7 trajectories with label 1, 2 trajectories with label 2 and 1 trajectory with label 3. The next location of each trajectory in *T10* is predicted and the results are used to calculate precision and recall. Varying *k* from 2 to 10 on 1 incremental steps, we obtain experimental results as in **Figure 2**.

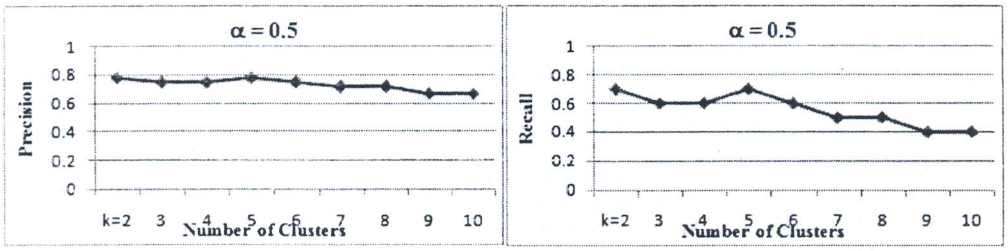

Fig. 2. The effect of the generated number of clusters *k* on prediction accuracy

Figure 2 shows that the recall decreases slightly whereas precision is not affected as the number of clusters *k* increases from 2 to 5. However, when *k* is larger than 5, although the precision is still not affected, the recall is strong reduced. That is due to that the probability of having some no-prediction cases becomes higher as *k* increases.

By considering the experimental result, the unchangeability of the precision values as *k* increases implies that our clustering procedure is good. This is because that each mobility pattern has assigned to the nearest group and further mobility patterns within the same group have a high degree of similarity, so the best matched rules of the current path are always found, even when *k* increases. Moreover, reducing the recall as *k* increases suggest that using personal profile to predict his future locations will return no-predictions and thus affect prediction accuracy if he is a new user or his current paths are novel. Therefore, it is necessary to group mobility behaviors of mobile users according to their similarity for reducing the probability of having some no-prediction cases.

As another results of this experiment, reducing the time cost of computation as *k* increases implies that we should classify the set of mobility patterns into as many groups as possible such that the prediction accuracy is still guaranteed. In this experiment, the best value of *k* is 5, which is the same as the number of movement groups of the dataset *DS500*.

C. Comparing with Prediction Models based on Personal Profile

The question now is that whether or not using group mobility behaviors to predict future movement. In order to answer this question, we perform two prediction approaches:

– *Case 1*: The prediction model based on group mobility behaviors.

– *Case 2*: The prediction model based on individual mobility behaviors [7].

For generating dataset of *Case 2*, we randomly select 50 trajectories with label 1 from the dataset *DS500* and name *DS50*. In *Case 1*, we also use *DS500* and fix the number of clusters at 5. For comparing the prediction accuracy of two approaches with respect to both precision and recall, we use the test set *TS10* and obtain experimental results as in **Figure 3**. The prediction accuracy in the case based on group mobility behaviors is always better than that in the case based on individual mobility behaviors. This is because that if the current path is novel, *Case 2* will return no-prediction whereas *Case 1* may return correct prediction. In addition, Case 1 and Case 2 will have the same results with movements on old paths due to the similar mobility patterns have been classified into the same group.

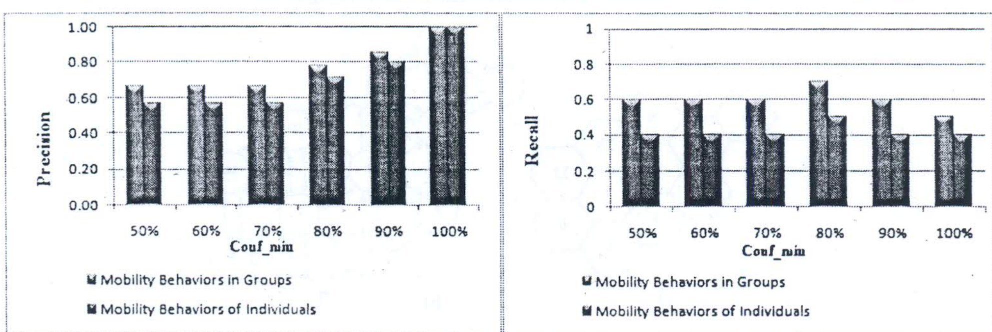

Fig. 3. The difference of prediction accuracy between the two models

In conclusion, we can say that group mobility behaviors may facilitate the predicting future locations of a new user or with movements on novel paths and further contribute considerably to the improvement in prediction accuracy.

V. CONCLUSIONS

In this paper, we have presented the technique for predicting future movement of mobile users in wireless networks, which is based on clustering technique for group behaviors. Our approach may deal with the lack of information on personal profile which is the drawback of making use of conventional prediction techniques. We first discover mobility rules from multiple users whose mobility behaviors are similar to each other. And then, we make use of mobility rules derived from the movement group to which the current path belongs for predicting user next locations. In order to demonstrate the necessity and effectiveness of the proposed approach, we have conducted experiments to evaluate the approach with respect to both precision and recall measures as well as to compare the approach with the other ones. We are currently constructing a real life dataset to verify the efficacy of the proposed approach. These research results will be presented in our future work.

VI. REFERENCES

[1] G. Yavas, D. Katsaros, O. Ulusoy and Y. Manolopoulos: A data mining approach for location prediction in mobile environments. Journal of Data and Knowledge Engineering, Vol. 54, pp. 121-146 (2005).

[2] Juyoung Kang and Hwan-Seung Yong: A Frequent pattern based prediction model for moving objects. International journal of computer science and network security, Vol. 10, No. 3, pp. 200-205 (2010).

[3] J. W. Lee, O. H. Paek and K. H. Ryu: Temporal moving pattern mining for location-based service. Journal of Systems and Software 73, Elsevier, pp. 481-490 (2004).

[4] C. Hung and W. Peng: Clustering object moving patterns for prediction-based object tracking sensor networks, Proc. of the 18th ACM conference on Information and knowledge management, pp. 1633-1636, USA (2009).

[5] S. Avasthi and A. Dwivedi: Prediction of Mobile User Behavior Using Clustering. International Journal of Scientific and Research Publications, Vol. 3, Issue 2, pp. 1-5 (2013).

[6] Andre E. Bergh and Neco Ventura: Prediction AssisteMd Fast Handovers for Mobile IPv6. IEEE MILCOM 2006 Unclassified Technical Sessions, Washington D.C. (2006).

[7] T. V. Duong, D. Q. Tran and C. H. Tran: Spatiotemporal Data Mining for Mobility Prediction in Wireless Network, Proc of the National Conference of Fundamental and Applied IT Research (FAIR), Dongnai, Vietnam, pp. 224-235 (2011).

[8] W. Wanalertlak, B. Lee, C. Yu, M. Kim, S.M. Park and W. T. Kim: Behavior-based mobility prediction for seamless handoff in mobile wireless networks, Springer Science + Business Media, Wireless Netw, Vol 17, pp. 645-658, (2011).

[9] T. V. Duong and D. Q. Tran: An Effective Approach for Mobility Prediction in Wireless Network based on Temporal Weighted Mobility Rule, International Journal of Computer Science and Telecommunications, Vol. 3, Issue 2, pp. 29-36 (2012).

[10] E. Cho, S. A. Myers, J. Leskovec: Friendship and Mobility: User Movement In Location-Based Social Networks, Proc. of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 1082-1090, USA (2011).

[11] D. Somayeh: Exploring Movement Similarity Analysis of Moving Objects, Doctor thesis, University of Zurich, Switzerland (2011).

[12] Xiaobai Yao: Research Issues in Spatio-temporal Data Mining. Workshop on Geospatial Visualization and Knowledge Discovery, University Consortium for Geographic Information Science, Virginia (2003).

[13] Shashi Shekhar, et al.: Spatial and Spatiotemporal Data Mining: Recent Advances, http://www-users.cs.umn.edu/~shekhar/research/sdm-chapter-ngdm-v2-2008.pdf (2008).

[14] T. V. Duong, D. Q. Tran and C. H. Tran. A weighted combination similarity measure for mobility patterns in wireless networks. International Journal of Computer Networks and Communications, Vol. 4, No.3, pp. 21-35 (2012).

[15] K. Huang, T.H.Yu and T. Kao: Analyzing structural changes using clustering techniques. International Journal of Innovative Computing, Information and Control, Vol. 4, No.5, pp. 1195-1202 (2008).

[16] M. Ranjan, D. P. Anna, P. G. Arka: A systematic evaluation of different methods for initializing the K-means clustering algorithm. IEEE Transctions on knowledge and data engineering. Available at http:/ / www. public. iastate. edu/ apghosh/ _les/ IEEEclust2. pdf (2010).

[17] J. A. Hartigan and M. A. Wong. A k-means clustering algorithm. Applied Statistics, vol. 28, pages 100-108, 1979.

[18] P. Berkhin. A survey of clustering data mining techniques. Grouping Multidimensional Data In Grouping Multidimensional Data, pp. 25-71 (2006).

MỘT MÔ HÌNH DỰ BÁO DI CHUYỂN DỰA TRÊN HÀNH VI DI CHUYỂN THEO NHÓM TRONG MẠNG KHÔNG DÂY

Dương Thị Thùy Vân, Trần Đình Quế, Trần Công Hùng

TÓM TẮT— *Dự báo sự di chuyển của người dùng di động đóng vai trò quan trọng trong quản lý tài nguyên của mạng không dây nên đã thu hút được sự quan tâm của nhiều nhóm nghiên cứu. Tuy nhiên, hầu hết cơ chế dự báo sự di chuyển hiện nay đều dựa trên lịch sử di chuyển của cá nhân để dự báo vị trí kết nối kế tiếp của người dùng di động nên có thể dẫn đến dự báo sai cho người dùng mới hoặc người dùng di chuyển trên những đường đi mới. Trong bài báo này, chúng tôi đề xuất một mô hình dự báo sự di chuyển dựa trên hành vi di chuyển theo nhóm nhằm khắc phục sự thiếu thông tin của lịch sử di chuyển cá nhân như vậy. Cơ chế dự báo của chúng tôi sử dụng kỹ thuật gom nhóm trong khai thác dữ liệu để phân các mẫu di chuyển của người dùng di động vào các nhóm sao cho những người thuộc cùng nhóm di chuyển thì có đặc điểm di chuyển tương tự nhau. Với mục tiêu này, trước hết chúng tôi xác định đường đi hiện tại của người dùng thuộc nhóm nào và sau đó sử dụng luật di chuyển của nhóm đó để dự báo vị trí kết nối kế tiếp cho họ. Kết quả thực nghiệm chứng minh rằng việc sử dụng hành vi di chuyển theo nhóm có thể tránh được sự thiếu luật di chuyển do lịch sử di chuyển của cá nhân không đầy đủ và do đó nâng cao đáng kể độ chính xác dự báo.*

Từ khóa— *nhóm di chuyển, mẫu di chuyển, người dùng di động, dự báo, mạng không dây.*