Data Mining Assisted Resource Management in Wide WLANs

Thuy Van T. Duong[1(✉)], Dinh Que Tran[2], and Cong Hung Tran[2]

[1] Faculty of Information Technology, University of Ton Duc Thang, No. 19 Nguyen Huu Tho Street, Tan Phong Ward, District 7, Ho Chi Minh, Vietnam
vanduongthuy@tdt.edu.vn, vanduongthuy@yahoo.com

[2] Faculty of Information Technology, Post and Telecommunication Institute of Technology, Ho Chi Minh, Vietnam

Abstract. WLANs are currently being considered for use in the context of a larger geographical area such as a city or a campus due to their convenience, cost efficiency, and ease of integration with other networks. Due to support large numbers of portable devices and their dynamic relocation, wide WLANs must face problems of location management and network resource allocation. In order to solve these challenges, future mobility information of all mobile users in the network is required to accurate estimation of network resource demands at future time toward more efficient network resource management. Therefore, mobility prediction has played a crucial role in the resource management of wide WLANs and it has attracted recently a great deal of research interests. However, since most of the current approaches are based on personal movement profile 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. Experiments will be performed to demonstrate that using group mobility behaviors may significantly enhance the accuracy of the mobility prediction.

Keywords: Movement group · Mobility pattern · Mobile user · Prediction · Wireless network

1 Introduction

WLANs may allow users access Internet applications from where they want and still remain connected to the Internet whilst on the move, using smart phones, iPads, iPods, tablets, laptops and other portable devices, they are no longer limited by the length of the cable. Due to their convenience, cost efficiency, and ease of integration with other networks, WLANs are growing fast. WLANs have often deployed to variety of limited coverage areas, such as coffee shops, buildings, etc before. But now, its coverage area is extended across a larger geographical area such as a city or a campus and called wide WLAN. In fact, some cities such as New York and London have started to offer

P.C. Vinh et al. (Eds.): ICCASA 2013, LNICST 128, pp. 329–338, 2014.
DOI: 10.1007/978-3-319-05939-6_32,

wide WLANs to enable people access the Internet even outside their normal work environment, for example when they ride the train home. Some colleges and universities have also deployed campus-wide WLANs to open higher-learning environments with many educational facilities and enable both students and faculty use portable communication devices as learning and teaching tools. The portable devices such as IP phones, smartphones, iPads, iPods and tablets are light enough to walk-and-talk, so these users often leave their devices on most of the time for accessing Internet applications, for example Voice over IP (VoIP) or realtime multimedia data transmission, etc. Due to campus-wide WLANs provide wireless availability not only in the classroom but also at every corner of the campus - communal spaces, cafeterias, campus retail outlets, libraries, on-campus living accommodation and also temporary or mobile study areas, these portable device users may still remain connect to the Internet during they moving around the campus. For example, Dartmouth College has about 550 APs providing 11 Mbps 802.11 b coverage to the entire campus of about 190 buildings, including all administrative, academic, and residential buildings, and most athletic facilities [1]. Dartmouth College has about 5,500 students, 4,200 undergraduate students and 1,215 full-time professors. Most of the undergraduate students live on campus and use wireless devices on most of the time. That mean, there are a large numbers of mobile and wireless users active in any given day, about 3,500 users in 2003.

Due to support large numbers of portable devices and their dynamic relocation, wide WLANs must face problems of location management and network resource allocation. In reality, whenever a mobile user moves from one cell to another, called *handover* or *handoff*, the network resources must be reallocated for his device at the new cell to continue the service. The required network resources are not available or not sufficient in the new cell force the network to terminate service to the user. Dropping a service in progress is considered to have a more negative impact from the user's perspective than blocking a newly requested service, thus handoff services are normally assigned higher priority over the new services. This requires future mobility information of all mobile users in the network to accurate estimation of network resource demands at future time. Therefore, mobility prediction has received a great deal of attention of research community [2–10] to provide useful mobility information for reserving resources in such cells where the users are likely to be in toward more efficient network resource management.

Among the proposed approaches, data mining has been widely used for mobility prediction based on the past mobility behaviors of WLAN users. Because mobility history hides useful knowledge patterns that describe typical behavior of mobile users, many approaches [2–4, 9, 10] make use of such mined knowledge patterns, which are represented in the form of mobility rules, to predict future movement of mobile users. Although these current studies are effective in predicting next cells for majority of mobile users, most of them are only based on individual mobility behaviors that cause false prediction for new users or ones with movements on novel paths due to the lack of information on personal movement profile.

Unlike most previous approaches [2–4, 9, 10], 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 new users or ones 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 [11, 12]. 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 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 from histories of all users in the network. The second phase is to cluster discovered mobility patterns into movement groups. The third phase determines which group the user's current path belongs to. And the last 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. Section 2 describes a typical wide WLAN architecture and the mobility model in this network. In Sect. 3, a novel model of mobility prediction based on the group mobility behavior is investigated and proposed. Section 4 is devoted to describing experiments and results for evaluating the proposed mobility prediction approach. Finally, Sect. 5 draws concluding remarks and further research work.

2 Modeling Mobility in Wide Wireless Local Area Network

In a typical wide WLAN, such as a academic campus WLAN or a city WLAN, the radio coverage region is partitioned into many potentially overlapping cells. Each cell is covered by a wireless access point (AP) which is base station (BS) for the wireless network. APs transmit and receive radio frequencies for wireless devices such as laptops, personal digital assistants (PDAs), iPads, iPods, tablets, IP phones or other smartphones to communicate with. All APs in the same building are connected through a switch, called Access Controller (AC), which is used to control all the APs in the building for performing their jobs. Every building's AC is wired to the campus back-bone network which is managed by some servers [1]. Figure 1 shows a typical wide WLAN architecture.

It is assumed that the radio coverage region is presented by a hexagonal shaped network as in Fig. 2. Each hexagon is a cell which is served by a AP in the communication space. The mobile users can travel around the coverage region. In order to illustrate the mobility model of WLAN users, we use an unweighted directed graph $G = (V, E)$, where the vertex-set V is the set of cells in the coverage region and the

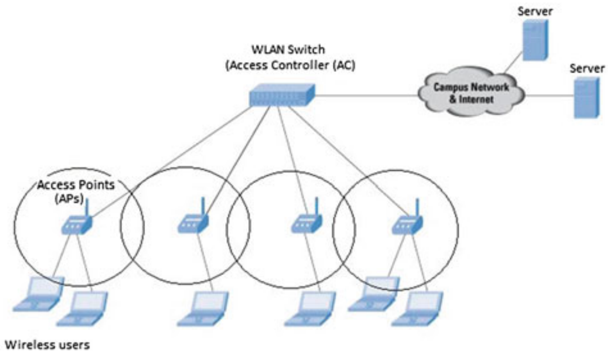

Fig. 1. A typical campus-wide WLAN architecture

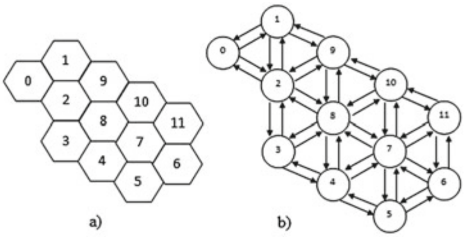

Fig. 2. The coverage region (a) and corresponding graph G (b)

edge-set E represent the adjacence between pairs of cells. These bidirected edges illustrate the fact that WLAN users may move from one cell to another directly and vice versa.

In this paper, we focus on WLAN users who use portable devices such as IP phones, smartphones, iPads, iPods or tablets. This is because that these users leave their devices, which are light enough to walk-and-talk, on most of the time. That means, they show a more mobile characteristic while connected to the network than WLAN users who use laptops or other heavy devices. On the other hand, such portable device users not only connect to the Internet but also use various services such as VoIP and realtime multimedia data transmission, e.g., streaming audio and video. Hence, they may frequently change their associated APs for network access, depending on their locations. Every time a wireless user roams from one AP to another, a syslog message is recorded thereby it is possible to collect sequences of associated APs for each WLAN user. That means, the handoff history of all WLAN users is maintained in the network. The main purpose of this paper is studying WLAN users' handoff history for mobility prediction which has played a crucial role in the resource network management. The following is the formalization for mobility patterns of WLAN users.

Let c be the ID number of the cell to which the mobile user connected at a predefined timestamp t, a point is defined as follows:

Definition 1. *Let C and T be two sets of ID cells and timestamps, respectively. The ordered pairs p = (c, t), in which* $c \in C$ *and* $t \in T$*, is called a point. Denote P to be the set of all points* $P = C \times T = \{(c, t) \mid c \in C \text{ and } t \in T\}$.

Two point $p_i = (c_i, t_i)$ and $p_j = (c_j, t_j)$ are said to be equivalent if and only if $c_i = c_j$ and $t_i = t_j$. Point $p_i = (c_i, t_i)$ is defined to be earlier than point $p_j = (c_j, t_j)$ if and only if $t_i < t_j$, and it is denoted as $(c_i, t_i) < (c_j, t_j)$ or $p_i < p_j$.

Definition 2. *The trajectory of the mobile user is defined as a finite sequence of points* $<p_1, p_2, \ldots, p_k>$ *in* $C \times T$ *space, where* $p_j = (c_j, t_j)$ *are points for* $1 \le j \le k$ *and ID cells of two consecutive points must be neighbors in the coverage region. A trajectory composed of k elements is denoted as a k-pattern.*

Note that the ascending order of a mobility pattern's points is sorted by t. For example $<(c_1, t_1), (c_2, t_2), (c_3, t_3), (c_4, t_4)>$ is a 4-pattern, where $t_1 \le t_2 \le t_3 \le t_4$.

3 Mobility Prediction Based on Group Mobility Behaviors

3.1 Discovering Frequent 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 [13, 14]. We have used spatial and temporal constraints 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 [10].

3.2 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 frequent mobility patterns which are discovered from histories of all mobile users and then classify them into groups of similar behaviors. The subsection presents briefly our proposed dissimilarity measure [15], which is the basis for constructing the clustering procedure to identify groups of similar mobility patterns.

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 3. *Let* $f: S \times S \rightarrow R$ *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:*

$$f(P_a, P_b) = card(\{c_{ai} | c_{ai} \notin P_b\}) + card(\{c_{bi} | c_{bi} \notin P_a\}) \quad (1)$$

The spatial dissimilarity 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 4. *The spatial dissimilarity 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:*

$$D_{space}(P_a, P_b) = \frac{f(P_a, P_b)}{n + m} \quad (2)$$

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 less temporally dissimilar the two patterns are:

Definition 5. *The temporal dissimilarity measure $D_{time}(P_a, P_b)$ between two patterns P_a and P_b with length n and m, respectively, is given by*

$$D_{time}(P_a, P_b) = \frac{1}{k} \sum\nolimits_{i=1, j=1}^{n,m} \frac{|t_{ai} - t_{bj}|}{\max(t_{ai}, t_{bj})} \text{ where } c_{ai} = c_{bj} \quad (3)$$

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 dissimilarity measures on space and time.

Definition 6. *The composition dissimilarity measure between two patterns P_a and P_b is given by*

$$D(P_a, P_b) = W_{space}.D_{space}(P_a, P_b) + W_{time}.D_{time}(P_a, P_b) \quad (4)$$

in which W_{space} and W_{time} are respectively weights of spatial and temporal dissimilarity 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. We extend the traditional k-means approach to partition the set of mobility patterns into *k* clusters in order to take its advantages of simplicity and computational speed [16–19]. Our alternative clustering algorithm focuses on applying our proposed dissimilarity measures for finding cluster centers and assigning mobility patterns to clusters. After initialization for *k* cluster centers, each pattern in the set of mobility patterns is assigned to the nearest cluster based on the proposed dissimilarity measure. Then, the center of each cluster will be updated and further each mobility pattern will be reassigned to the nearest cluster 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.

3.3 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. Due to both cluster center and path are mobility patterns, the similarity between them are also measured by Definition 6. 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.

3.4 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. Then, we match the current path to the head of rule and find the best matching using the time constraint.

4 Experimental Evaluation

4.1 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* which is produced by a dataset generator into 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 trajectories with label 3 from the dataset *DS500*.

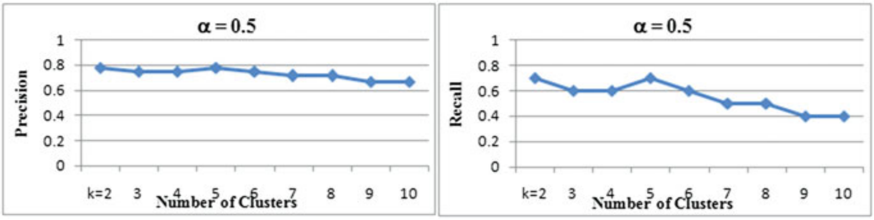

Fig. 3. The effect of the generated number of clusters k on prediction accuracy

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 Fig. 3.

Figure 3 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*.

4.2 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 [8, 10]

For generating dataset of *Case 2*, we randomly select 50 trajectories with label 1 from the dataset *DS500* to create a set of trajectories of the same mobile user and

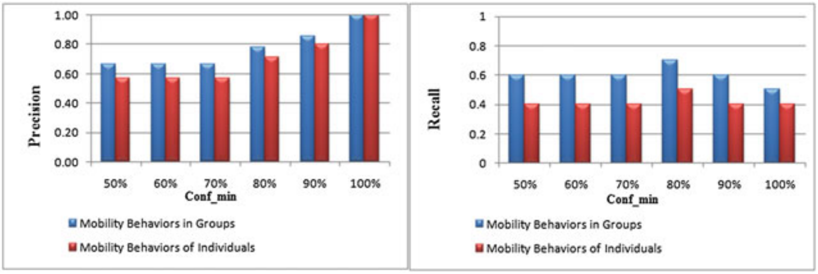

Fig. 4. The difference of prediction accuracy between the two models

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 Fig. 4. 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.

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

5 Conclusions

In this paper, we have presented a novel approach for predicting future movement of mobile users in wide WLANs, which is based on clustering technique for group behaviors. The proposed 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 the next point-of-attachments of users. 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 model based on individual mobility behaviors. We are currently constructing a real life dataset which is collected from campus-wide WLAN of Dartmouth College to verify the efficacy of the proposed approach. Moreover, such the real dataset enable us to compare the proposed approach with other ones which are also based on group mobility behaviors. These research results will be presented in our future work.

References

1. Kotz, D., Essien, K.: Analysis of a Campus-wide Wireless Network. In: Proceedings of the 8th Annual International Conference on Mobile Computing and Networking (ACM MOBICOM'02), pp. 107–118 (2002)
2. Yavas, G., Katsaros, D., Ulusoy, O., Manolopoulos, Y.: A data mining approach for location prediction in mobile environments. J. Data Knowl. Eng. **54**, 121–146 (2005)
3. Kang, J., Yong, H.-S.: A Frequent pattern based prediction model for moving objects. Int. J. Comput. Sci. Netw. Secur. **10**(3), 200–205 (2010)
4. Lee, J.W., Paek, O.H., Ryu, K.H.: Temporal moving pattern mining for location-based service. J. Syst. Softw. **73**, 481–490 (2004). (Elsevier)
5. Hung, C., Peng, W.: Clustering object moving patterns for prediction-based object tracking sensor networks. In: Proceedings of the 18th ACM conference on Information and knowledge management, pp. 1633–1636, USA (2009)
6. Avasthi, S., Dwivedi, A.: Prediction of mobile user behavior using clustering. Int. J. Sci. Res. Publ. **3**(2), 1–5 (2013)
7. Bergh, A.E., Ventura, N.: Prediction assisted fast handovers for mobile IPv6. IEEE MILCOM 2006 Unclassified Technical Sessions, Washington D.C. (2006)
8. Duong, T.V., Tran, D.Q., Tran, C.H.: Spatiotemporal data mining for mobility prediction in wireless network. In: Proceedings of the National Conference of Fundamental and Applied IT Research (FAIR), pp. 224–235, Vietnam (2011)
9. Wanalertlak, W., Lee, B., Yu, C., Kim, M., Park, S.M., Kim, W.T.: Behavior-based mobility prediction for seamless handoff in mobile wireless networks. Wirel. Netw. **17**, 645–658 (2011). (Springer Science + Business Media)
10. Duong, T.V., Tran, D.Q.: An effective approach for mobility prediction in wireless network based on temporal weighted mobility rule. Int. J. Comput. Sci. Telecommun. **3**(2), 29–36 (2012)
11. Cho, E., Myers, S.A., Leskovec, J.: Friendship and mobility: user movement in location-based social networks. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1082–1090, USA (2011)
12. Somayeh, D.: Exploring movement similarity analysis of moving objects. Doctor thesis, University of Zurich, Switzerland (2011)
13. Yao, X.: Research issues in spatio-temporal data mining. In: Workshop on Geospatial Visualization and Knowledge Discovery, University Consortium for Geographic Information Science, Virginia (2003)
14. Shekhar, S., Vatsavai, R.R., Celik, M.: Spatial and Spatiotemporal Data Mining: Recent Advances. Symposium, Next Generation of Data Mining, pp. 549–584 (2008)
15. Duong, T.V., Tran, D.Q., Tran, C.H.: A weighted combination similarity measure for mobility patterns in wireless networks. Int. J. Comput. Netw. Commun. **4**(3), 21–35 (2012)
16. Huang, K., Yu, T.H., Kao, T.: Analyzing structural changes using clustering techniques. Int. J. Innovative Comput. Inf. Control **4**(5), 1195–1202 (2008)
17. Ranjan, M., Anna, D.P., Arka, P.G:. A systematic evaluation of different methods for initializing the K-means clustering algorithm. IEEE Trans. Knowl. Data Eng. **38**, 522–537 (2010)
18. Hartigan, J.A., Wong, A.: A k-means clustering algorithm. Appl. Stat. **28**, 100–108 (1979)
19. Berkhin, P.: A survey of clustering data mining techniques. Grouping Multidimensional Data in Grouping Multidimensional Data, pp. 25–71 (2006)