Volume 5, Issue 9, September 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

An Effective Temporal Mobile Frequent Item Set Find Using Sequence Cluster with Location Affinity Algorithm

S. N. Sharmila, M. Preetha
Research Scholar, Asst. Professor, PGP College of Arts & Science,
Dept of CS& Applications, Tamil Nadu,
India

Abstract- The rapid advance of wireless communication technology Mobile Commerce is not only being widely accepted but also it is being more used as a popular way of business commerce done by portable devices. It is becoming an interesting to find patterns and prediction of mobile user behaviors such as their location and purchase transactions in mobile commerce effectively to provide the service. Here, it provides a more efficient service to the mobile commerce users by applying weighted frequent pattern and periodical pattern for prediction of purchase behavior of mobile users. The Mobile commerce Explorer consists of three major components: 1) Similarity Inference Model SIM for measuring the similarities among stores and items, which are two basic mobile commerce entities considered Here; 2) Personal Mobile Commerce Pattern Mine(PMCP-Mine) algorithm for efficient discovery of mobile users' Personal Mobile Commerce Patterns (PMCPs); and 3) Mobile Commerce Behavior Predictor MCBP for prediction of possible mobile user behavior .and find the closest pattern along with support value. In addition, temporal periodical pattern method is used to find the frequent user behavior in all time intervals of the transaction including the weight of the each item set and support value of the user for an item. Finally, the percentage of precision and recall is measured by comparing the various methods to prove the efficiency of the proposed pattern mining and prediction.

Keywords- Mobile Commerce, User behavior, Similarity, Pattern, Prediction.

I. INTRODUCTION

Mobile users can request services through their mobile devices via Information Service and Application Provider (ISAP) from anywhere at anytime. This business model is known as Mobile Commerce (MC) that provides Location-Based Services (LBS) through mobile phones. MC is expected to be as popular as e-commerce in the future and it is based on the cellular network composed of several base stations. The communication coverage of each base station is called a cell as a location area.

.

1.1 MOBILE TRANSACTION SEQUENCE

The average distance between two base stations is hundreds of meters and the number of base stations is usually more than 10,000 in a city. When users move within the mobile network, their locations and service requests are stored in a centralized mobile transaction database. Fig. 1 shows an MC scenario, where a user moves in the mobile network and requests services in the corresponding cell through the mobile devices. Fig. 1a shows a moving sequence of a user, where cells are underlined if services are requested there. Fig. 1b shows the record of service transactions, where the service S1 was requested when this user moved to the location A at time 5.

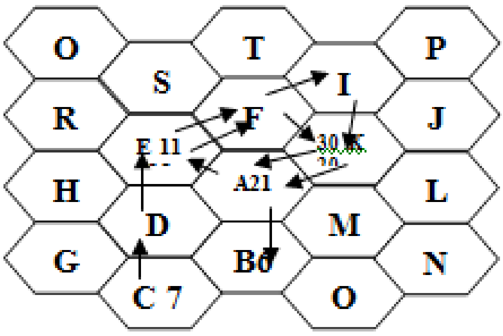

Fig1. An example for a mobile transaction sequence. (a) Moving sequences

Time	Cell	Service
5	A	S_1
6	B	
7	C	S_3
10	D	S_2
11	E	
13	F	S_3S_4
18	I	
20	K	S_5
21	A	S_1
25	E	
28	F	
30	K	S_2

Fig2. An example for a mobile transaction sequence (b) Service sequences.

In fact, there exists insightful information in these data, such as movement and transaction behaviors of mobile users. Mining mobile transaction data can provide insights for various applications, such as data prefetching and servicse recommendations. A mobile transaction database is complicated since a huge amount of mobile transaction logs is produced based on the user's mobile behaviors. Data mining is a widely used technique for discovering valuable information in a complex data set and a number of studies have discussed the issue of mobile behavior mining. The main difference between these literatures is the involved information of proposed patterns. Previous studied by Tseng and Tsui addressed the problem of mining associated service patterns in mobile web networks.Tseng and Lin also proposed SMAP-Mine to efficiently mine users' sequential mobile access patterns, based on the FP-Tree.

Yun and Chen proposed a novel method of mining mobile sequential patterns. To increase the accuracy of predictions, the moving pathwas taken into consideration in the previous studies. However, mobile behaviors vary among different user clusters or at various time intervals. The prediction of mobile behavior will be more precise if we can find the corresponding mobile patterns in each user cluster and time interval. To provide precise location-based services for users, effective mobile behavior mining systems are required pressingly. Clustering mobile transaction data helps in the discovery of social groups, which are used in applications such as targeted advertising, shared data allocation, and personalization of content services. In previous studies, users are typically clustered according to their personal profiles (e.g., age, sex, and occupation).

However, in real applications of mobile environments, it is often difficult to obtain users' profiles. That is, we may only have access to users' mobile transaction data. To achieve the goal of user clustering without user profiles, the evaluation of the similarities of mobile transaction sequences (MTSs) is needed. Although a number of clustering algorithms have been studied in the rich literature, they are not applicable in the LBS scenario in consideration of the following issues: 1) Most clustering can only process data with spatial similarity measures, while clustering methods with non-spatial similarity measures are required for LBS environments. 2) Most clustering methods request the users to set up some parameters.

However, in real applications, it is difficult to determine the right parameters manually for the clustering tasks. Hence, an automated clustering method is required. Although there exist many non-spatial similarity measures, most of them are used to measure the string similarity. However, the mobile transaction sequences discussed Here include multiple and heterogeneous information such as time, location, and services. Therefore, the existing measures are not applicable directly for measuring the similarity of mobile transaction sequences.

1.2. TIMEINTERVAL SEGMENTATION

The time interval segmentation method helps us find various user behaviors in different time intervals. For example, users may request different services at different times (e.g., day or night) even in the same location. If the time interval factor is not taken into account, some behaviors may be missed during specific time intervals. To find complete mobile behavior patterns, a time interval table is required. Although some studies used a predefined time interval table to mine mobile patterns, the data characteristic and data distribution vary in real mobile applications. Therefore, it is difficult to predefine a suitable interval table by users. Automatic time segmentation methods are, thus, required to segment the time dimension in a mobile transaction database.

OBJECTIVE

- To find the similarity between items sold in stores.
- To find the similarity between stores.
- To find the item similarity matrix.
- To find the store similarity matrix.
- To predict the purchase behavior of users in terms of items and stores.
- To mine and predict of mobile movements and associated transactions
- To propose a prediction strategy to predict the subsequent mobile behaviors.
- To present a time segmentation approach to find segmenting time intervals where similar mobile characteristics exist.

- To propose a novel algorithm, namely, Cluster-based Temporal Mobile Sequential Pattern Mine (CTMSP-Mine), to discover the Cluster-based Temporal Mobile Sequential Patterns (CTMSPs).
- To construct user clusters by a novel algorithm named Cluster-Object-based Smart Cluster Affinity Search Technique (CO-Smart-CAST) and similarities between users are evaluated by the proposed measure, Location-Based Service Alignment (LBS-Alignment).
- To mine and predict of mobile behaviors with considerations of user relations and temporal property simultaneously.

II. LITERATURE SURVEY

2.1 CLUSTERING GENE EXPRESSION PATTERN

"Clustering Gene Expression Patterns" [1] the authors AMIR BEN-DOR, RON SHAMIR, and ZOHAR YAKHINI stated that Recent advances in biotechnology allow researchers to measure expression levels for thousands of genes simultaneously, across different conditions and over time. Analysis of data produced by such experiments offers potential insight into gene function and regulatory mechanisms. A key step in the analysis of gene expression data is the detection of groups of genes that manifest similar expression patterns. The corresponding algorithmic problem is to cluster multi condition gene expression patterns. In the paper they described a novel clustering algorithm that was developed for analysis of gene expression data. They defined an appropriate stochastic error model on the input, and prove that under the conditions of the model, the algorithm recovers the cluster structure with high probability. They also presented a practical heuristic based on the same algorithmic ideas. The heuristic was implemented and its performance is demonstrated on simulated data and on real gene expression data, with very promising results.

2.2 EFFICIENT MINING OF USER BEHAVIORS BY TEMPORAL MOBILE ACCESS PATTERNS

In the paper "Efficient Mining of User Behaviors by Temporal Mobile Access Patterns" [2] the authors Seung-Cheol Lee, Juryon Paik, Jeewoong Ok, Insang Song and Ung Mo Kim stated that ubiquitous computing offers various kinds of dynamic services to the mobile users with versatile devices at anytime and anywhere. In ubiquitous environment intelligent mobile agents are mandated to communicate with users and it is enabled by capturing interesting user's behavior patterns. Existing mining methods have proposed frequent mobile user's behavior patterns statistically based on requested services and location information. In this case some problems are caused because it was not considered that the mobile user's dynamic behavior patterns are usually associated with temporal access patterns. Therefore, ubiquitous computing provides the dynamic services with timely manner when user wants to get useful services information on time. Here, they proposed a novel data mining method, namely temporal mobile access patterns that can efficiently discover mobile user's temporal behavior patterns associated with location and requested services. Furthermore, they presented a novel data structure T-Map to store the temporal mobile access patterns. The advantage of our data structure compactly stores the user's behavior pattern according to location and service information in memory. Even though the information data sets require large shared memory when they are stored, their approach still provides fast access and consume less memory than other methods. The proposed technique Here works especially well for context-awareness data sets in mobile agent systems.

The rapid expansion of ubiquitous mobile computing technology has created an unprecedented opportunity to gather and extract information from mobile agent systems. Therefore, effective modeling the behavior patterns of users in mobile agent systems is becoming very important. Effective modeling the behavior items of users in mobile agent systems benefits not only the users in smart access by caching or perfecting but also the mobile service providers in financial profit like advertising.

2.3 FINDING SIMILARITY OF MOVING OBJECT TRAJECTORIES USING EDP

In the paper "Robust and Fast Similarity Search for Moving Object Trajectories" [26] the authors Lei Chen, M. Tamer ¨ Ozsu and Vincent Oria stated that An important consideration in similarity-based retrieval of moving object trajectories is the dentition of a distance function. The existing distance functions are usually sensitive to noise, shifts and scaling of data that commonly occur due to sensor failures, errors in detection techniques, disturbance signals, and different sampling rates. Cleaning data to eliminate these is not always possible. Here, we introduce a novel distance function, Edit Distance on Real sequence (EDR) which is robust against these data imperfections. Analysis and comparison of EDR with other popular distance functions, such as Euclidean distance, Dynamic Time Warping (DTW), Edit distance with Real Penalty (ERP), and Longest Common Subsequences (LCSS), indicate that EDR is more robust than Euclidean distance, DTW and ERP, and it is on average 50% more accurate than LCSS. They also developed three pruning techniques to improve the retrieval efficiency of EDR and show that these techniques can be combined effectively in a search, increasing the pruning power significantly. The experimental results confirmed the superior efficiency of the combined methods. With the growth of mobile computing and the development of computer vision techniques, it has become possible to trace the trajectories of moving objects in real life and in videos. A number of interesting applications are being developed based on the analysis of trajectories.

For example, using a remote sensing system, and by mining the trajectories of animals in a large farming area, it is possible to determine migration patterns of certain groups of animals. In sports videos, it is quite useful for coaches or sports researchers to know the movement patterns of top players. In a store surveillance video monitoring system, finding the customers' movement patterns may help in the arrangement of merchandise. Since existing similarity measures can not readily be used to retrieve trajectories, in the paper, they introduced a novel distance function that addresses the peculiarities of trajectories, and they discussed the retrieval efficiency issues relative to this distance function. The major contributions are the following:

1. A novel distance function, Edit Distance on Real sequence (EDR), to measure the similarity between two trajectories. EDR is based on edit distance on strings, and removes the noise effects by quantizing the distance between a pair of elements to two values, 0 and 1. Seeking the minimum number of edit operations required to change one trajectory to another offers EDR the ability to handle local time shifting. Furthermore, assigning penalties to the unmatched parts improves its accuracy. Through a set of objective tests on benchmark data, they showed that EDR is more robust than Euclidean distance, DTW, and ERP, and more accurate than LCSS when it is used to measure the similarity between trajectories that contain noise and local time shifting.
2. Developed three pruning techniques – mean value Q-grams, near triangle inequality, and trajectory histogram – to improve the retrieval efficiency of EDR. Unlike the pruning methods proposed for LCSS [26, 27] or DTW [28], these pruning methods do not require setting constraints on warping length (or matching region) between two trajectories, and therefore, offer users more flexibility.
3. They showed how to combine the three pruning methods to significantly reduce the number of false candidates. Furthermore, they developed different variations of the three pruning methods and compare their performance in terms of pruning power and speedup ratio and we show the superior searching efficiency of the combined methods.

2.4 EFFICIENT DATA MINING FOR PATH TRAVERSAL PATTERNS

In the paper "Efficient Data Mining for Path Traversal Patterns" [29] the authors Ming-Syan Chen and Jong Soo Park explored a new data mining capability that involves mining path traversal patterns in a distributed information-providing environment where documents or objects are linked together to facilitate interactive access. Their solution procedure consists of two steps.

First, they derived an algorithm to convert the original sequence of log data into a set of maximal forward references. By doing so, they could filter out the effect of some backward references, which are mainly made for ease of traveling and concentrate on mining meaningful user access sequences. Second, they derived algorithms to determine the frequent traversal patterns, i.e., large reference sequences from the maximal forward references obtained. Two algorithms are devised for determining large reference sequences; one is based on some hashing and pruning techniques, and the other is further improved with the option of determining large reference sequences in batch so as to reduce the number of database scans required. Performance of these two methods is comparatively analyzed. Sensitivity analysis on various parameters is conducted. UE to the increasing use of computing for various applications, the importance of database mining is growing at a rapid pace recently. Progress in bar-code technology has made it possible for retail organizations to collect and store massive amounts of sales data. Catalog companies can also collect sales data from the orders they received. It is noted that analysis of past transaction data can provide very valuable information on customer buying behavior, and thus improve the quality of business decisions (such as what to put on sale, which merchandises to be placed together on shelves, how to customize marketing programs, to name a few).

It is essential to collect a sufficient amount of sales data before any meaningful conclusion can be drawn there from. As a result, the amount of these processed data tends to be huge. It is hence important to devise efficient algorithms to conduct mining on these data. Note that various data mining capabilities have been explored in the literature. One of the most important data mining problems is mining association rules [30], [31], [32], [33]. For example, given a database of sales transactions, it is desirable to discover all associations among items such that the presence of some items in a transaction will imply the presence of other items in the same transaction.

Also, mining classification is an approach of trying to develop rules to group data tuples together based on certain common features. This has been explored both in the AI domain [34], [35] and in the context of databases [36], [37], [38]. Mining in spatial databases was conducted in [39]. Another source of data mining is on ordered data, such as stock market and point of sales data. Interesting aspects to explore from these ordered data include searching for similar sequences [40], [41], e.g., stocks with similar movement in stock prices, and sequential patterns [42], e.g., grocery items bought over a set of visits in sequence. It is noted that data mining is a very application-dependent issue and different applications explored will require different mining techniques to cope with. Proper problem identification and formulation is therefore a very important part of the whole knowledge discovery process. In the paper, they explored a new data mining capability which involves mining access patterns in a distributed information-providing environment where documents or objects are linked together to facilitate interactive access.

Examples for such information-providing environments include World Wide Web (WWW) [43] and on-line services where users, when seeking for information of interest, travel from one object to another via the corresponding facilities (i.e., hyperlinks) provided. Clearly, understanding user access patterns in such environments will not only help improve the system design (e.g., provide efficient access between highly correlated objects, better authoring design for pages, etc.) but also be able to lead to better marketing decisions (e.g., putting advertisements in proper places, better customer/user classification and behavior analysis, etc.). Capturing user access patterns in such environments is referred to as mining traversal patterns in the paper.

Note that although some efforts have elaborated upon analyzing the user behavior [44], [45], [46], there is little result reported on dealing with the algorithmic aspects to improve the execution of traversal pattern mining. This can be in part explained by the reason that these information-providing services, though with great potential, are mostly in their infancy and their customer analysis may still remain in a coarser level such as user occupation/age study.

Note that the problem of finding large reference sequences is similar to that of finding large item sets for association rules [47], where a large item set is a set of items appearing in a sufficient number of transactions. However, they are different from each other in that a reference sequence in mining traversal patterns has to be consecutive references in a maximal forward reference whereas a large item set in mining association rules is just a combination of items in a transaction.

Although trimming the transaction database as it proceeds to later passes, algorithm FS is required to scan the transaction database in each pass. In contrast, by properly utilizing the candidate reference sequences, the second algorithm devised, referred to as selective-scan (SS) algorithm, is able to avoid database scans in some passes so as to reduce the disk I/O cost involved. Specifically, algorithm SS has the option of using a candidate reference set to generate subsequent candidate reference sets, and delaying the determination of large reference sets to a later pass when the database is scanned. Since SS does not scan the database to obtain a large reference set in each pass, some database scans are saved. It is noted that, although the concept of selective scan was used for mining association rules, its implementation and performance implication are different when it is employed for mining path traversal patterns.

Experimental studies are conducted by using a synthetic workload that is generated based on referencing some logged traces, and performance of these two methods, FS and SS, is comparatively analyzed. It is shown that the option of selective scan is very advantageous and algorithm SS thereby outperforms algorithm FS in general. Sensitivity analysis on various parameters is also conducted.

2.5 A DENSITY-BASED ALGORITHM FOR DISCOVERING CLUSTERS IN LARGE SPATIAL DATABASES WITH NOISE

In the “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise” [50] the authors Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu stated that Clustering algorithms are attractive for the task of class identification in spatial databases.

However, the application to large spatial databases rises the following requirements for clustering algorithms: minimal requirements of domain knowledge to determine the input parameters, discovery of clusters with arbitrary shape and good efficiency on large databases. The well-known clustering algorithms offer no solution to the combination of these requirements. In the paper, they presented the new clustering algorithm DBSCAN relying on a density-based notion of clusters which is designed to discover clusters of arbitrary shape. DBSCAN requires only one input parameter and supports the user in determining an appropriate value for it. They performed an experimental evaluation of the effectiveness and efficiency of DBSCAN using synthetic data and real data of the SEQUOIA 2000 benchmark. The results of their experiments demonstrate that (1) DBSCAN is significantly more effective in discovering clusters of arbitrary shape than the well-known algorithm CLARANS, and that (2) DBSCAN outperforms CLARANS by a factor of more than 100 in terms of efficiency.

III. PROBLEM FORMULATION

Clustering mobile transaction data helps in the discovery of social groups, which are used in applications such as targeted advertising, shared data allocation, and personalization of content services. In previous studies, users are typically clustered according to their personal profiles (e.g., age, gender, and occupation).

However, in real applications of mobile environments, it is often difficult to obtain users' profiles. That is, we may only have access to users' mobile transaction data. To achieve the goal of user clustering without user profiles, the evaluation of similarities of mobile transaction sequences (MTSs) is required. The existing system consists of a novel framework, called Mobile Commerce Explorer (MCE), for mining and prediction of mobile users' movements under the context of mobile commerce. The MCE framework consists of three major components:

1) Similarity Inference Model (SIM) for measuring the similarities among stores and items, which are two basic mobile commerce entities considered in this paper;
2) Personal Mobile Commerce Pattern Mine (PMCP-Mine) algorithm for efficient discovery of mobile users' Personal Mobile Commerce Patterns (PMCPs); and
3) Mobile Commerce Behavior Predictor (MCBP) for prediction of possible mobile user behaviors. Drawbacks of the existing

- Discovering mobile patterns from the whole logs are not precise enough for predictions.
- Differentiated mobile behaviors among users and temporal periods are not considered.
- Request the users to set up some parameters which is difficult to determine.
- Not applicable in the LBS (Location Based Service) scenario.
- Process data with spatial similarity measures, while clustering methods with non-spatial similarity measures are required for LBS environments

IV. PROPOSED SYSTEM

The proposed system develops a novel algorithm, namely, Cluster-based Temporal Mobile Sequential Pattern Mine (CTMSP-Mine), to discover the Cluster-based Temporal Mobile Sequential Patterns (CTMSPs).

Since, a prediction strategy is proposed to predict the subsequent mobile behaviors, in CTMSP-Mine, user clusters are constructed by a novel algorithm named Cluster Affinity Search Technique (CAST) and similarities between users are evaluated by the proposed measure, Location-Based Service Alignment (LBS-Alignment).

At the same time, a time segmentation approach is presented to find segmenting time intervals where similar mobile characteristics exist. The project considers mining and prediction of mobile behaviors with considerations of user relations and temporal property simultaneously.

The proposed system has following advantages,

- Predicts the subsequent user mobile behaviors effectively.
- Generate the most suitable time intervals for time segmentation.
- Mines and predicts the mobile behaviors with considerations of user relations and temporal property simultaneously.
- Suitable for Location-Based Service Environments.

The project contains the following

1. Item Similarity
2. Store Similarity
3. Clustering of mobile transaction sequences.
4. Time segmentation of mobile transaction sequences.
5. Discovery of CTMSPs.

Item Similarity

In this module, the items are added with their score values. Then similarity between two items are found out using the given similarity formula. Likewise similarity matrix is found out for all the items in the list.

Store Similarity

In this module, the stores are added with their score values. Then similarity between two store are found out using the given similarity formula. Likewise similarity matrix is found out for all the stores in the list.

Clustering of mobile transaction sequences

In a mobile transaction database, users in the different user groups may have different mobile transaction behaviors. The first task to tackle is to cluster mobile transaction sequences. In this module, a parameter-less clustering algorithm called CAST is proposed.

Before performing the CAST, a similarity matrix S is to be generated, based on the mobile transaction database. The entry Sij in matrix S represents the similarity of the mobile transaction sequences i and j in the database, with the degrees in the range of [0, 1]. A mobile transaction sequence can be viewed as a sequence string, where each element in the string indicates a mobile transaction. The major challenge to tackle is to measure the content similarity between mobile transactions. The LBS-Alignment algorithm is proposed, which can obtain the similarity. LBS-Alignment is based on the consideration that two mobile transaction sequences are more similar, when the orders and timestamps of their mobile transactions are more similar. CAST algorithm is used to cluster the users.

Segmentation of Mobile Transactions

In a mobile transaction database, similar mobile behaviors exist under some certain time segments. Hence, it is important to make suitable settings for time segmentation so as to discriminate the characteristics of mobile behaviors under different time segments.

A new time segmentation method is proposed to automatically obtain the most suitable time segmentation table with common mobile behaviors. The algorithm below shows the procedure of the proposed time segmentation method, named Get Number of Time Segmenting Points (GetNTSP) algorithm.

Discovery of CTMSPs

In order to mine the cluster-based temporal mobile sequential patterns efficiently, we proposed a novel method named CTMSP-Mine to achieve this mining procedure. In CTMSP-Mine, both factors of user cluster and time interval are taken into account such that the complete mobile sequential patterns can be discovered. The entire procedures of CTMSP-Mine algorithm can be divided into three main steps:

1) Frequent-Transaction Mining,
2) Mobile Transaction Database Transformation, and
3) CTMSP Mining.

1) Frequent-Transaction Mining

In this phase, the frequent transactions (F-Transactions) are mined in each user cluster and time interval by applying a modified Apriori algorithm.

2) Mobile Transaction Database Transformation

In this phase, F-Transactions are used to transform each mobile transaction sequence S into a frequent mobile transaction sequence S'. According to Table 3, if a transaction T in S is frequent, T would be transformed into the corresponding F-Transaction.

Otherwise, the cell of T would be transformed into a part of path.

3) CTMSP Mining

In this phase, all the CTMSPs are mined from the frequent mobile transaction database. Frequent 1-CTMSPs are obtained in the frequent-transaction mining phase. In the mining algorithm, we utilize a two-level tree named Cluster-based Temporal Mobile Sequential Pattern Tree (CTMSP-Tree). The internal nodes in the tree store the frequent mobile transactions, and the leaf nodes store the corresponding paths. Moreover, every parent node of a leaf node is designed as a hash table which stores the combinations of user cluster tables and time interval tables.

V. CONCLUSION

In this paper, a novel method named CTMSP-Mine is proposed, for discovering CTMSPs in LBS environments. Furthermore, novel prediction strategies are proposed to predict the subsequent user mobile behaviors using the discovered CTMSPs. In CTMSP-Mine, first a transaction clustering algorithm is proposed named CO-Smart-CAST to form user clusters based on the mobile transactions using the proposed LBS-Alignment similarity measurement. Then, the time segmentation algorithm is utilized to generate the most suitable time intervals. To our best of mobile behaviors associated with user clusters and temporal relations. A series of experiments were conducted for evaluating the performance of the proposed methods. The experimental results show that CO-Smart-CAST method achieves high-quality clustering results and the proposed CBSS strategy obtains highly precise results for user classification. Meanwhile, the algorithms obtain the most proper and correct time intervals. For behavior prediction, CTMSP is shown to outperform other prediction methods in terms of precision and F-measure. The experimental results demonstrate that the proposed methods are efficient and accurate under various conditions.

REFERENCE

[1] A. Ben-Dor and Z. Yakhini, "Clustering Gene Expression Patterns," J. Computational Biology, vol. 6, no. 3, pp. 281-297, July 1999.

[2] S.C. Lee, J. Paik, J. Ok, I. Song, and U.M. Kim, "Efficient Mining of User Behaviors by Temporal Mobile Access Patterns," Int'l J. Computer Science Security, vol. 7, no. 2, pp. 285-291, Feb. 2007.

[3] R. Agrawal, C. Faloutsos, and A. N. Swami. Efficient similarity search in sequence databases. In Proc. 4th Int. Conf. of Foundations of Data Organization and Algorithms, pages 69–84, 1993.

[4] R. Agrawal, G. Psaila, E. L. Wimmers, and M. Za¨ıt. Querying shapes of histories. In Proc. 21th Int. Conf. on Very Large Data Bases, pages 502–514, 1995.

[5] C. Faloutsos, M. Ranganathan, and Y. Manolopoulos. Fast subsequence matching in time-series databases. In Proc. ACM SIGMOD Int. Conf. on Management of Data, pages 419–429, 1994.

[6] L. Chen, M. T. ¨ Ozsu, and V. Oria. Robust and efficient similarity search for moving object trajectories. In CS Tech. Report. CS-2003-30, School of Computer Science, University of Waterloo.

[7] Y. Zhu and D. Shasha. Warping indexes with envelope transforms for query by humming. In Proc. ACM SIGMOD Int. Conf. on Management of Data, pages 181–192, 2003.

[8] B-K Yi, H. Jagadish, and C. Faloutsos. Efficient retrieval of similar time sequences under time warping. In Proc. 14th Int. Conf. on Data Engineering, pages 23–27, 1998.

[9] S Kim, S. Park, and W. Chu. An indexed-based approach for similarity search supporting time warping in large sequence databases. In Proc. 17th Int. Conf. on Data Engineering, pages 607–614, 2001.

[10] E. Keogh. Exact indexing of dynamic time warping. In Proc. 28th Int. Conf. on Very Large Data Bases, pages 406–417, 2002.

[11] J. S. Boreczky and L. A. Rowe. Comparison of video shot boundary detection techniques. In Proc. 8th Int. Symp. on Storage and Retrieval for Image and Video Databases, pages 170–179, 1996.

[12] M. Vlachos, G. Kollios, and D. Gunopulos. Discovering similar multidimensional trajectories. In Proc. 18th Int. Conf. on Data Engineering, pages 673 – 684, 2002.

