Hybrid Temporal Sequential Pattern Mining Scheme for Mobile Services

M.Nagaraj¹, A. Venugopal²

¹Department of Computer Applications, Erode Arts and Science College (Autonomous), Bharathiar University, India
²Department of Computer Science, Sree Narayana Guru College, Bharathiar University, India

mdotnagaraj@gmail.com; venugopal.achuthan@gmail.com

ABSTRACT: Researches on Location-Based Service (LBS) have been emerging in recent years due to a wide range of potential applications. One of the active topics is the mining and prediction of mobile movements and associated transactions. Most of existing studies focus on discovering mobile patterns from the whole logs. However, this kind of patterns may not be precise enough for predictions since the differentiated mobile behaviors among users and temporal periods are not considered. In this work, we propose a novel algorithm, namely, Cluster-based Temporal Mobile Sequential Pattern Mine (CTMSP-Mine), to discover the Cluster-based Temporal Mobile Sequential Patterns (CTMSPs). Moreover, a prediction strategy is proposed to predict the subsequent mobile behaviors. In CTMSP-Mine, user clusters are constructed 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). Meanwhile, a time segmentation approach is presented to find segmenting time intervals where similar mobile characteristics exist. To our best knowledge, this is the first work on mining and prediction of mobile behaviors with considerations of user relations and temporal property simultaneously. Through experimental evaluation under various simulated conditions, the proposed methods are shown to deliver excellent performance.

Keywords - Data mining, transportation, mining methods and algorithms, mobile environments.

I. INTRODUCTION

Devices and systems based on mobile technologies are now common place in everyday life. Such devices and systems include cellular telephones and pagers, cordless telephones, two-way radios, baby crib monitors, remote car locking systems, wireless networking systems (including wireless local area networks

© 2016, IJCSMC All Rights Reserved
[LANs], Global Positioning System (GPS)–based locators and maps, and electronic monitoring devices for parolees. Sometimes these devices and systems allow existing activities to be carried out more effectively or efficiently; at other times, they enable entirely new and different activities [6].

The mobile industry has experienced significant growth during the past two decades. Practically the fast expansibility of ubiquitous mobile [10] technology has created a numberless opportunity to gather and extract information from mobile agent systems. Recently a critical problem was exist which is how services and access methods can be provided to mobile users. Mobile services are by definition consumed through a mobile handset, which is defined to mean a pocket-sized device with at least cellular connectivity capabilities. In the above mentioned definition of the mobile industry, network vendors, handset vendors, network operators, virtual network operators, service providers, content aggregators, third party software developers and emerging mobile Internet players are all therefore part of the mobile communications industry [3].

Cell phones have been adopted faster than any other technology in human history [11] and as of 2008, the number of cell phone subscribers exceeds 2.5 billion, which is twice as many as the number of PC users worldwide 1. To capture a slice of this lucrative market, Nokia, Google, Microsoft, and Apple have introduced cell phone operating systems (Symbian, Android, Windows Mobile, and OS X) and open APIs for enabling application development on the cell phones. Recently, cell phones have also attracted the attention of the networking and ubiquitous computing research community due to their potential as sensor nodes for city-wide sensing application [1].

Cellular communication has been experiencing a rapid growth in recent years. Since its introduction in the early 1980s, cellular communication has been evolving from costly service with limited availability toward an affordable alternative to wired telephone service. This wide acceptance of cellular communication has led to the development of a new generation of mobile communication network called personal communications services (PCS), which can support a larger mobile subscriber population while providing various types of services unavailable to traditional cellular systems [21].

Many mobile devices also incorporate low power wireless connectivity protocols, such as Bluetooth, that can be used to identify an individual to other people nearby. We have developed an architecture that leverages this functionality in mobile phones—originally designed for communication at a distance—to connect people across the room. Serendipity is an application of the architecture. It combines the existing communications infrastructure with online introduction systems’ functionality to facilitate interactions between physically proximate people through a centralized server [15].
Technology developments have created new types of services. Mobile services differ from traditional services in their ability to provide service offerings regardless of temporal and spatial constraints. The benefits of mobile services are often summarized in four factors: ubiquity, convenience, localization and personalization that differentiate mobile services from online services [9].

Mobile services differ from traditional services in their ability to provide service offerings regardless of temporal and spatial constraints. They are also different from traditional interpersonal services that deliver face-to-face, or from other types of e-services, such as wireless online services, where the service delivery linked to a specific fixed local area network or specific location [7]. With rapid development of the wireless communication techniques, more and more users subscribe various kinds of mobile services in a mobile environment due to the advantage of easy access.

In mobile computing environment, analyzing the moving pattern of the mobile users can help understand the behavior of the users. When a user is roaming in a mobile environment, the information relevant to the movement of the users will be recorded in the system log. From the business standpoint, valuable information regarding the users moving behavior may exist in the system log [10]. To discover the moving pattern of the mobile users from the system log, utilizing the discovered moving behavior of the users could provide personalized services.

II. LITERATURE SURVEY

Agrawal et.al [2] introduced the concept of mining association rule between sets of items in large database of customer transaction. The algorithm used here is Apriori algorithm. This algorithm finds if there is an association between departments in the customer behavior. The main advantage is no redundancy of mining data. The available memory space is managed efficiently. Tseng and Tsui addressed the problem of mining associated service patterns in mobile web networks [22]. Tseng and Lin also proposed SMAP-Mine to efficiently mine users’ sequential mobile access patterns, based on the FP-Tree [24]. Yun and Chen proposed a method of mining mobile sequential patterns [8].

Agrawal et.al [3] considered the problem of discovering association rules between items in a large database of sales transactions. Progress in bar-code technology has made it possible for retail organizations to collect and store massive amounts of sales data called basket data. A record in such data consists of item date and items bought in the transactions. The problem of mining association rules over basket data is analyzed.

In mobile web environments users request services through cell phones from an arbitrary location at any time. A sequence of requests of a user form a location service stream and it is called as behavior
patterns. The behavior patterns carry not only the traversal path but also the service request. A data mining method called SMAP (Sequential Mobile Access Pattern) Mine is used to discover user pattern along with the associated request. Thus Tseng et.al [23] analyzed the user behavior pattern and efficient mining algorithm in mobile web systems.

A new data mining capability for mobile commerce environment is explored by Ching-Huang Yun et.al [5] in a mobile commerce environment. Moving patterns and purchase patterns are taken into consideration and a model for mobile commerce system is built. Association rule mining and path traversal serves as a basic scheme to determine the frequent mobile pattern. By having different priorities on the factors involving large item sets, traversal paths, and orders of purchases for determining mobile sequential patterns from transaction sequence.

Monreale et.al [16] designed a location predictor on trajectory mining. A method to predict the location of the moving object with certain level of accuracy to predict the next location of moving object. The prediction uses previously extracted movement patterns named trajectory patterns which are the representation of behaviors of moving objects. Here T-patterns are used with temporal information to find out the behaviors.

Spatio-temporal databases that manage information about objects moving in two dimensional spaces are important for many applications. Tracking of moving objects which typically can occupy only a single position at a given time. Among several spatio-temporal structures that focus on predictive query processing, the most popular one is TPR tree (Time Parameterized R-tree). Thus Tao Y., et.al [20] formulated a prediction method of moving objects with unknown motion patterns.

Trajectories of moving vehicles and algorithm for mining frequent patterns of trajectory data had been proposed by Ajay kumar Akasapu et.al [4] for both sparse and dense dataset. Many applications track the movement of mobile objects, using location-acquisition technologies such as Global Positioning System (GPS), Global System for Mobile Communications (GSM) etc. and it can be represented as sequences of time stamped locations.

In ubiquitous environment intelligent mobile agents are mandated to communicate with users and it is enabled by capturing interesting user’s behaviour patterns. Temporal mobile access patterns that could discover mobile user’s temporal behaviour patterns associated with location and requested services. The user’s temporal mobile access patterns are stored in memory space. A compact data structure is used to store data in memory space. It consumes less amount of memory. It is applicable to health care system, mobile web system. Thus Seung-Cheol Lee et.al [19] gave an idea about efficient mining in a mobile environment.

© 2016, IJCSMC All Rights Reserved
Ester et.al [13] proposed a density based algorithm for discovering clusters in large spatial databases with noise. Density based clustering called DBSCAN is used to handle local density variation within the cluster. It can detect clusters of different shapes and sizes from large amount of data which contain noise and outliers. Tseng et.al [22] introduced the concept of mining multilevel and location-aware associated service patterns in a mobile web environment. A new data mining method -two-dimensional multilevel (2-DML) association rules mining. Efficient execution and memory consumption of data. In this correspondence, the issue of efficiently mining multilevel and location-aware associated service patterns in a mobile web environment is addressed.

Anant Ram et. al [5] presented a new idea that a density varied DBSCAN algorithm is capable to handle the local density variation within the cluster. A new algorithm DVBSCAN (Density Variation Based Spatial Clustering of Applications with Noise) an enhancement of DBSCAN algorithm is being proposed. This algorithm finds clusters which represent uniform regions without being separated by sparse regions. Experimental results prove that the proposed algorithm satisfies optimized result. Eric Hseuh-chan Lu et.al [12] analysed mining pattern using an algorithm called CTMSP mine. To identify the similarities among users Cluster Object cluster Affinity Search Technique is used. They conducted a series of experiments to evaluate the performance of the algorithm. The performance measures precision and F-measure sounds better for behaviour prediction.

A classifier to mine trajectory data is identified by Sharma et.al [18].The algorithm focussed here is Nearest Neighbour. The experimental investigation shows that classified pattern exhibits the success rate of 98.2. Rama kalaivani et.al [17] proposed an optimization technique to improve the performance of mobile prediction behaviour. This helps in improving the quality of service. Cluster tables are formed to identify similarities of mobile user. The accuracy of prediction process is improved with the results of mining. Hanna et.al [15] presented a new idea that apart from location and requested services time plays a key role in detecting the behaviour of mobile user. Temporal mobile access patterns are compared in terms of complexity and accuracy.

III. PROPOSED METHOD

In this section, we describe our system design. Four important research issues are addressed here:
1. Clustering of mobile transaction sequences.
2. Time segmentation of mobile transaction sequences.
3. Discovery of CTMSPs.
4. Mobile behavior prediction for mobile users.
Spatial data mining techniques are used to mine data values with location information. Temporal analysis is applied to perform time analysis. Location and time based analysis is applied on a variety of mobile service analysis application. Different services are provided under the mobile environment. User behavior analysis is carried out using the sequential pattern mining methods. Location and time factors are used in the sequential pattern mining methods. Clustering techniques are used to group up the transactions based on the transaction relevancy.

Most of existing techniques focus on discovering mobile patterns from the whole logs. However, this kind of patterns may not be precise enough for predictions since the differentiated mobile behaviors among users and temporal periods are not considered. Cluster-based Temporal Mobile Sequential Pattern Mine (CTMSP-Mine) is used to discover the Cluster-based Temporal Mobile Sequential Patterns (CTMSPs). A prediction strategy is proposed to predict the subsequent mobile behaviors. In CTMSP-Mine, user clusters are constructed by Cluster-Object-based Smart Cluster Affinity Search Technique (CO-Smart-CAST). The similarities between users are evaluated by the Location-Based Service Alignment (LBS-Alignment). A time segmentation approach is presented to find segmenting time intervals where similar mobile characteristics exist.

The CTMSP mine model is enhanced with hybrid prediction model (HPM) and recursive motion functions (RMF) scheme to improve the pattern identification and prediction accuracy levels. The system development is planned with Java language and Oracle database.
Problem Definition

Let \( S = < (t_1, l_1, s_1), (t_2, l_2, s_2), \ldots, (t_n, l_n, s_n) > \) be an MTS of a user with length equal to \( n \), where item \((t_i, l_i, s_i)\) represents the user requests’ service \( s_i \) in location \( l_i \) at time \( t_i \) and \( t_i < t_{i+1} \ \forall \ i \geq 1 \leq n \). The ascending order of elements in a sequence is determined, using time as the key [20].

The main problem we are addressing in this work is formulated as follows: Given a user’s current mobile transaction sequence \( S \) and the current time \( t_c \), our goal is to develop a framework to predict the subsequent mobile behaviors. We aim to predict the subsequent mobile behaviors using not only \( S \) and \( t_c \) but also all the mined CTMSPs. The problem of CTMSPs mining is formulated as follows: Given a mobile transaction database \( D \) containing a large number of mobile transaction sequences of users and a specified support threshold \( \delta \), the problem is to discover all the CTMSPs existing in the database. In this work, we propose the CTMSP-Mine algorithm and the behavior prediction mechanism for solving this problem.

IV. RESULTS AND DISCUSSION

1. Input Design

Input design is the link between the information system and the users and those steps that are necessary to put transaction data in to a usable form for processing data entry. The activity of putting data into the computer for processing can be activated by instructing the computer to read data from a written printed document or it can occur by keying data directly into the system. The designs of input focusing on controlling the amount of input required controlling the errors, avoid delay extra steps, and keeping the process simple. The input design considers the input data, input medium, user interface, messages, validation and error handling factors.

The mobile service and location prediction system is designed up with user friendly and interactive forms which enable the user to operate the application with ease of use. The input forms are highly designed with data validation, data integration and consistency with databases and application logic. The users are directed with standard messages and alerts which enables them to feed the data with accuracy.

The mobile service management system is designed with a set of input forms to collect user input data values. The subscriber entry form is used to register the subscriber details. The page entry form is designed to update new page details. The service register form is designed to collect new service details. The location details are collected under the location entry form. The page access entry form is designed to fetch web page request details. The service access entry form is designed to fetch the service request details. The automatic log generation form is designed to update access log details automatically by the system. The location based alignment and time slot update operations are carried out using different input
forms. The transaction sequence form is designed to prepare transaction sequences. The clustering process is designed in a separate form. The prediction process is designed under the service prediction form. The input forms are designed with validation and list selection based models.

Output Design

The mobile service and location prediction system is designed with a set of output forms to show the intermediate and final output data values. Designing computer should proceed in well thought out manner. The term output means any information produced by the information system whether printed or displayed. When analyst design computer output they identified the specific output that is needed to meet the requirement. Computer is the most important source of information to the users. Output design is a process that involves designing necessary outputs that have to be used by various users according designing necessary outputs that have to be used by various users according to requirements. Efficient intelligent output design should improve the system relationship with the user and help in decision making. Since the reports are directly required by the management for taking decision and to draw the conclusion must be simple, descriptive and clear to the user. Options for outputs and forms are given in the system menus.

The subscriber list form is designed to show the registered subscribers under the service provider. The page list and service list forms are designed to show the registered pages and services under the provider. The location details are displayed under the location list form. The access log list form shows the list of service and page requests submitted by the users. The optimized data values are listed in a separate form. The transaction sequences are listed under transaction list form. The cluster results and pattern details listed in separate forms. The prediction results are produced under the prediction process form.

i. System Testing

The mobile service and location prediction system is tested with its operations and results. System testing is the important stage in the software development life cycle. The functionalities and requirements are analyzed in the testing phase. The testing includes unit testing, integration testing and acceptance test methods. The test cases are used to verify the testing results. Test cases are prepared with reference to the customer requirements. The mobile service management model is tested using unit testing and user acceptance testing methods.

ii. Unit Testing

The mobile service management system is tested module by module. Unit testing is applied to check the operations of the modules. The module level test tasks are completed in the development stage. Each and every form is tested separately. The input and output data values are verified in the unit testing
process. Data update, delete and list operations are tested in the unit testing process. The log management module is designed with a set of sub modules such as subscriber, page, service and locations. The subscriber entry form is tested with validation methods. The update data values are verified under the subscriber table. In the same way the page, service and location entry forms are verified. The data pre-process, clustering process and prediction modules are also tested with their results.

iii. Acceptance Testing

The mobile service and location prediction process is verified under the client environment. The acceptance testing is carried out to verify the system fulfills all the user requirements or not. The system requirements and their implementation details are verified in the acceptance test. The problems and their solutions are also analyzed in the acceptance test. The system is passed into the implementation phase after the getting the acceptance from the users. The client requirements are verified under the acceptance testing process with service prediction process. The service and location prediction system fulfils all the client requirements.

Test Cases

The mobile service management system is tested with a collection test case. The test cases are checklist for the testing process. The facts that are considered in the testing process are listed in the test cases. These test cases are used in the testing process. The test cases are prepared by the designed with the details collected from the client. All testing tasks are performed with reference to the test cases. The mobile service management system is uses the following test constraints:

- Data validation and entry verification on subscriber, page, service and location entry forms.
- Log data update process verification.
- Automatic log generation process verification.
- Check the location and time rearrangement process.
- Clustering process verification.
- Check the transaction sequence conversion process.
- Analyze the pattern extraction process.
- Verify the result of service and location prediction process.

System Implementation

i. Implementation procedure

After proper testing and validation, the question arises whether the system can be implemented or not. Implementation includes all those activities that take place to convert from old system to new. The new system may be totally new replacing an existing manual or automated system, or it may be a major
modification to an existing system. In other case, proper implementation is essential to provide a reliable system to meet institute requirements.

ii User Training

A well-designed, if not operated and used properly could fail. Training the users is important, as if not done well enough could prevent the successful implementation of an information system. Through the systems development life cycle the user has been involved. By this stage the analyst should possess an accurate idea of the users they need to be trained. They must know what their roles will be, how they can use the system and what the system will do and will not do. Both system operators and user need training. During their training, they need to be given a trouble-shooting list that identifies possible and remedies for the problem. They should be advised of the common mal functions that may arise and how to solve them.

iii Operational Documentation

Once the implementation plan is decided, it is essential that the user of the system is made familiar and comfortable with the environment. Education involves right atmosphere & motivating the user. A documentation providing the whole operations of the system is being developed. The system is developed in such a way that the user can work with it in a well consistent way. The system is developed user friendly so that can work the system from the tips given in the application itself. Useful tips and guidance is given inside the application itself to help the user. Users have to be made aware that what can be achieved with the new system and how it increase the performance of the system. The user of the system should be given a general idea of the system before he uses the system.
Fig. 3 Location Based Alignment Details

Fig. 5 Prediction Process
V. CONCLUSION

The availability of common, basic interactive elements increases the user transfer of learning between devices and services and improves the overall usability of the entire interactive mobile environment. Such a transfer becomes even more important in a world of ubiquitous mobile telecommunication devices and services. Simplifying the learning procedure for end-users will allow the reuse of basic knowledge between different terminal devices and services and lead to a faster and easier adoption of new technologies.

In this system, we have proposed a novel method, named CTMSP-Mine, for discovering CTMSPs in LBS environments. Furthermore, we have proposed novel prediction strategies to predict the subsequent user mobile behaviors using the discovered CTMSPs. In CTMSP-Mine, we first propose a transaction clustering algorithm named COSmart-CAST to form user clusters based on the mobile transactions using the proposed LBS-Alignment similarity measurement. Then, we utilized the genetic algorithm to generate the most suitable time intervals. To our best knowledge, this is the first work on mining and prediction 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, our GA-based method obtains 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 our proposed methods are efficient and accurate under various conditions.

Location and time based analysis is applied on a variety of mobile service analysis application. User behavior analysis is carried out using the sequential pattern mining methods. Cluster-Based Temporal Mobile Sequential Patterns (CTMSP) algorithm is used for the sequential pattern mining process. The integrated CTMSP model identifies the user behaviors. Clustering techniques are used for the pattern extraction process. Mobile sequential patterns are identified with temporal and spatial information. The system also predicts the user movements. Time segmentation based approach.

VI. Future Work

The mobile access sequence pattern mining system is enhanced with service prediction and movement prediction models. The service prediction identifies the future service calls for the users. In the same way user future movements can be predicted by the system. The system can be enhanced with the
following feature directions. The system can be enhanced to detect group moving patterns. The system can be improved to predict base station placement process. Service replica assignment can be integrated with the system. The system can be improved to handle bandwidth allocation.

REFERENCES


