SURVEY ON PERSONAL MOBILE COMMERCE PATTERN MINING AND PREDICTION

Ms. Pranjali P. Ghode
Department of Computer Science and Engineering
Sipna College of Engineering and Technology
Amravati, India.

***Abstract*—Data Mining refers to extracting or "mining" knowledge from large amounts of data. Due to a wide range of potential applications, research on mobile commerce has received a lot of interests from both of the industry and academia. Among them, one of the active topic areas is the mining and prediction of users' mobile commerce behaviors. In this paper we focus on Personal Mobile Commerce Pattern Mining and Prediction. Pattern mining is used to discover patterns to represent the relations among items. Prediction is important in intelligent environment; it captures repetitive patterns or activities and also helps in automating activities. This paper gives a brief introduction to various algorithms and a detailed study has been performed. This paper conducts a theoretical analysis study on pattern mining and prediction in mobile commerce.**

Index Terms—Data mining, Mobile commerce, Prediction, FP-growth

I. INTRODUCTION

Data mining refers to extracting or mining knowledge from large amounts of data. Data Mining (DM) uses the powerful software tools to separate important or significant qualities that are previously unknown from databases or data warehouses. Data mining uses information from past data to analyze the outcome of a particular problem or situation that may arise. Data mining works to analyze data stored in data warehouses that are used to store that data that is being analyzed. The advantages of data mining are Marketing/Retailing, Banking/Crediting, Law Enforcement, Researchers. Mobile Commerce, also known as M-Commerce or mCommerce, is the ability to conduct commerce using a mobile device. Mobile Commerce is a new emerging technology with greater scope. Mobile commerce is the buying and selling of goods and services through Wireless handheld devices. Mobile devices mainly smart phones overcome laptops and desktops in many perspectives. Its size, portability, convenience and so on [7]. It is advantage to the customers during purchasing; customers usually carry a mobile device mainly a smart phone than laptops because of its smaller size and portability. Mobile commerce has several applications, in that Localization of products and services plays a major role. It is used to know user locations and the services requested by the user. Knowing users' preferences and surfing habits marketers can send

- User-specific advertising messages
- Location-specific advertising messages

The remainder of the paper is organized as follows, in section II, we discuss about the related work. In section III, we describe the comparison of the techniques. Finally, we summarize our conclusion.

A) Frequent Pattern Mining

Frequent patterns are patterns that appear in a database frequently (e.g. a set of items, such as iPhone and headset, which appear frequently in a transaction data set is a frequent item set). A set is called frequent if its support is no less than a given absolute minimal support [1]. Two measures are used they are, 1.Support and 2.Confidence. In support, the rule holds with support sup in T (the transaction data set) if sup% of transactions contain *X* È *Y*. In confidence, the holds conf in T if conf% of transactions that contain X also contain Y. In the following, we describe the methods for mining frequent item sets.

B) FP-growth algorithm

FP-growth [4] is a well-known algorithm that uses the FP-tree data structure to achieve a condensed representation of the database transactions and employs a divide-and-conquer approach to decompose the mining problem into a set of smaller problems. In essence, it mines all the frequent item sets by recursively finding all frequent 1-itemsets in the conditional pattern base that is efficiently constructed with the help of a node link structure. A variant of FP-growth is the H-mine algorithm [5]. It uses array-based and trie-based data structures to deal with sparse and dense datasets respectively. Patricia Mine [6] employs a compressed Patricia trie to store the datasets. FP growth* [7] uses an array technique to reduce the FP-tree traversal time. In FP-growth based algorithms, recursive construction of the FP-tree affects the algorithm's performance.

C) Apriori algorithm

Apriori employs an iterative approach known as a level-wise search [4],where k-item sets are used to explore (k+1) item sets. The set of frequent 1-itemsets is found by scanning the database to gather the count for each item, and collecting the items that satisfy the minimum support count value. It is a seminal algorithm, which uses an iterative approach known as a level-wise search. It uses the Apriori property to reduce the search space: All nonempty subsets of frequent item set must also be frequent.

P (I) <min_sup => I is not frequent

P (I+A) <min_sup => I+A is not frequent either Antimonotone property – if a set cannot pass a test, all of its supersets will fail the same test as well.

II. RELATED WORK

A) 2-DML Association Rule Mining

2-DML is used to discover location aware service patterns in mobile web environments [2]. It is mainly used to understand the behavior of mobile users. In existing, they are collecting the communication logs of customers, with that we can't able to discover the services requested in that location. So they proposed a data mining mechanism called 2-DML_T1L1 [2] which can efficiently discover the associated service patterns based on two kinds of hierarchies, the location and service are hierarchied. 2-DML_T1L1 is more efficient in the execution and better memory saving. The disadvantage is that, if the Source logs are very large there will produce a high overload.

B) Tree Based Hierarchical Graph

Mining the Interesting locations and travel sequences in a given geospatial region. GPS Trajectories is for representing people's location histories [1].TBHG is used to mine the multiple individuals' location histories. In existing, user wants to browse each GPS trajectory one by one; it will be a difficult one. So they are creating geo-related web communities, in that they can upload GPS logs. Based on TBHG, they proposed a, HITS based model [1] to know the users travel experience and interesting location within a region. The advantage is that to understand the surroundings and also to plan their journey efficiently.

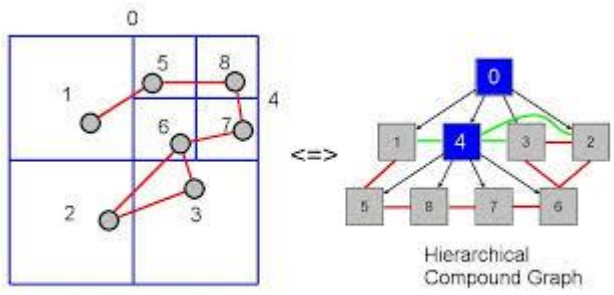

Figure 1- Tree Based Hierarchical Graph

C) Pattern Family Technique

They proposed an efficient model, called mobile sequential patterns. Mobile sequential pattern [3] takes both the moving and purchase patterns of the customers. In existing, they are collecting the various knowledge from various users (or) customers; with this the sequential pattern can't be determined. Three algorithms are proposed, in that Pattern family technique [3] TJPF is used to generate the frequent sequential patterns efficiently. Sequential Pattern mining is the mining of frequently occurring ordered events or subsequences as patterns. The main goal is to discover the user behavior. It is more efficient in execution and consumes much more memory.

D) Temporal Mobile Sequential Patterns

TMSP-Mine for efficiently discovering the temporal mobile sequential patterns [4] of users in Location based service environments. TMSP-Mine algorithm can be divided into three main.

Steps: 1) Large Transaction Generation, 2) Large Transaction Transformation, and 3) TMSP-Mine Algorithm. The details will be described as follows. In order to ensure the transactions of mobile sequential pattern is large, we determine the large transactions for each time interval and cell. Maximal large transaction sequences by mapping table. The main advantages are: 1) it can let a number of service sets represent to symbols, and 2) it can reduce the services of transactions that less than minimal support. Because our method considers time slots, segmenting each mobile transaction sequence to several segment in accordance with time interval table. Each transaction of segment transforms into unique id based on mapping table in mining algorithm, we utilize two-level tree named *TMSP-Tree*. The node in the upper lever stores the transaction sequences, and the leaf node stores the mobile sequential patterns. Beside the leaf node contains a time slots table to store the pattern of each time slot. To discover the user behavior, time interval factor is used. The performance is efficient and accurate**.**

E) SMAP-Mine and SMAP-Tree

SMAP-Mine is used to discover mobile users' sequential movement patterns with the service that was requested. By using the mobile, the input to the SMAP-Mine algorithm is the log of mobile access patterns, which is obtained by integrating both of movement log and service request log. For the SMAP-Mine algorithm, two phases are included, namely (i) construction of SMAP-Tree, and (ii) mining of sequential mobile access patterns. The purpose of constructing SMAP-Tree is to aggregate the access patterns into the memory in a compact form so that the mining of frequent patterns can be done efficiently. The main merits of SMAP-Tree are (1) only one physical database scan is needed to mine all of the frequent patterns, and (2) the SMAP Tree is compact so that it can be loaded into memory for efficient processing. Only one scan is needed to mine the frequent pattern**.** The performance is accurate, scalable and efficient.

F) Mobile Commerce Explorer

A novel framework called MCE which is mainly used to predict the user behaviors such as their movements and purchase transaction. It is used to predict the behavior of individual users. The mobile network database transformation maintains detailed store information which includes the location. The mobile users' moves between the stores and purchase the items and all these information are stored in mobile transaction database. PMCP-Mining uses the PMCP-Tree to predict the frequent mobile transactions. The MCE Framework [3] has three factors,

- Similarity Inference Model (SIM) is used to measure the similarities between stores and items.
- PMCP-Mine Algorithm is used to efficient

discovery of mobile users from PMCPs

- Mobile Commerce Behavior Prediction (MCBP) is used to predicting the mobile user behavior and movements

III. PROPOSED METHOD

The comparisons of various techniques are discussed as follows,

Techniques (or) Algorithm	**Input**	**Advantages**	**Parameter used**	**Threshold value**
2-DML Association Rule Mining	Integrated Datasets	Execution efficiency and Memory	Network nodes, no.of users,	40%
Tree Based Hierarchical Graph	107 Datasets	Understand the surroundings and plan their	Stay point detection, Clustering	2%
Pattern Family Technique	Datasets with 200 to 1000k	Efficient and Memory saving	Average path length, no.of items, weight	1.5% to 0.25%
TMSP-Mine	Mobile Transaction log	Efficient and Accurate	Network nodes, Time interval and	0.6%
SMAP-Mine	Real datasets	Accuracy, execution, efficiency and	Network nodes, mobility	0.001 to 0.005%
Mobile Commerce Explorer	Synthetic datasets	Accurate, efficient	Mesh network, no.of users,	0.1%

IV. CONCLUSION

This paper gives a brief introduction to various algorithms and a detailed study has been performed. This paper conducts a theoretical analysis study on pattern mining and prediction in mobile commerce. A brief discussion of those techniques is summarized. The advantages and limitations of pattern mining and prediction techniques are summarized with reference to various issues related to mobile commerce. Predicting the user behavior is an important issue. To efficiently predict the user behavior frequently pattern mining is use.

REFERENCES

[1] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, ―Mining Interesting Location and Travel Sequences from GPS Trajectories, Proc. Int'l World Wide Web Conf., pp. 791-800, Apr. 2009.

[2] V.S. Tseng and C.F. Tsui, ―Mining Multi-Level and Location- Aware Associated Service Patterns in Mobile Environments, IEEE Trans. Systems, Man and Cybernetics: Part B, vol. 34, no. 6, pp. 2480-2485, Dec. 2004.

[3] C.H. Yun and M.S. Chen, ―Mining Mobile Sequential Patterns in a Mobile Commerce Environment, IEEE Trans. Systems, Man, and Cybernetics, Part C, vol. 37, no. 2, pp. 278-295, Mar. 2007.

[4] V.S. Tseng, H.C. Lu, and C.H. Huang, ―Mining Temporal Mobile Sequential Patterns in Location-Based Service Environments, Proc. Int'l Conf. Parallel and Distributed Systems, pp. 1-8, Dec. 2007.

[5] V.S. Tseng and K.W. Lin, ―Efficient Mining and Prediction of User Behavior Patterns in Mobile Web Systems, Information and Software Technology, vol. 48, no. 6, pp. 357-369, June 2006.

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

[7] J. Pei, J. Han, X yan and H.cheng "Mining Compressed frequent pattern sets".

