ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 12, December 2013

SURVEY ON PERSONAL MOBILE COMMERCE PATTERN MINING AND PREDICTION

S. Jacinth Evangeline, K.M. Subramanian, Dr. K. Venkatachalam

***Abstract*—Data Mining refers to extracting or "mining" knowledge from large amounts of data. 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**.

***Index Terms*—Data mining, Mobile commerce, Prediction**

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 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[9]. 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.

Association rule mining is a popular and well researched method for discovering interesting relations between variables in large databases [12]. It is used to identify strong rules discovered in databases (e.g. Basket data analysis, clustering, and classification). The association rule mining can be of two types:

1. Frequent item sets: The items that frequently occur in the database and satisfies the minimum support count.

2. Generate strong association rules from the frequent item sets: Satisfy the minimum support and minimum confidence based on the rules.

The remainder of the paper is organized as follows, In section II, we discuss about the frequent pattern mining. In section III and IV, we discuss about the Apriori and FP Growth Algorithm. In section V, we briefly review the related work. In section VI, we describe the comparison of the techniques. Finally in section VII, we summarize our conclusion.

II. 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 [10]. 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 $X \cup Y$. 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.

III. APRIORI ALGORITHM

Apriori employs an iterative approach known as a level-wise search [14],where k-itemsets are used to explore (k+1)-itemsets. 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. The resulting set is L1. L2 can be derived from L1, and so on until no more frequent k-itemset can be found [14]. To improve the efficiency Apriori property is used. Apriori property is used to reduce the search space. An simple example for Apriori Algorithm,

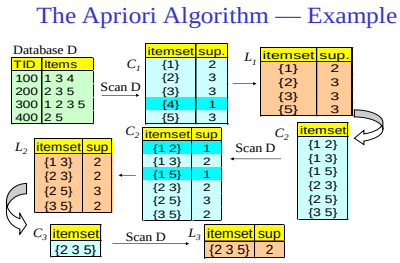


Fig. Example for Apriori Algorithm

IV. FP-GROWTH ALGORITHM

FP-growth supports the divide-and-conquer methology. It compress the database for representing frequent items into a frequent-pattern tree or FP-tree [15],which retains the association item set information. Then it divides the compressed database into a set of conditional databases, each associated with one frequent item or "pattern fragment" and mines each such database separately. Some of the advantage of FP-Growth is only two scans over database, no candidate generation, much faster than Apriori.

V. 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 cant 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 arehierarchied. 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 an 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.

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] TJ_{PF} 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. Web Access Pattern

WAP-Tree is used to access the patterns from pieces of logs. Web access patterns [7] is the sequence of accesses performed by users frequently. Weblogs is like a storage medium, which contains the information about the accesses are recorded, including the URL, origin of request and timestamp. Weblogs is divided into several

pieces by preprocessing technique. With the each piece of web log [7] they can't able to predict the frequent sequential pattern. So WAP-Mine Algorithm is proposed to mine the web access patterns from large set of pieces of web log. WAP-Tree [7] is used to handle the sequences with effective data structure and efficient search space.

E. Temporal Mobile Sequential Patterns

TMSP-Mine for efficiently discovering the Temporal mobile sequential patterns [4] of users in Location based service environments. The moving paths and time intervals in LBS is mined and associated with mobile sequential patterns. The Location prediction strategies [1] are used to predict the next movement of mobile users. The customers moving path, service request and time of request are recorded in mobile transaction log, from this they can predict the mobile sequential pattern of multiple users based on time. To discover the user behavior, time interval factor is used. The performance is efficient and accurate**.**

F. Mining Compressed Frequent Pattern Sets

Sheer size is the challenging issue in frequent pattern mining. Frequent Pattern mining [8] is a pattern (or) a set of items that occurs frequently in a data set. The patterns are clustered with a tightness measure and representative pattern can be selected for each cluster. Several techniques are proposed [8], in that RPglobal is used to mine frequent patterns. RPlocal is much more efficient and quite good compression quality. RPcombine is used to combine both the RPglobal and RPlocal. FPclose is an efficient algorithm to generate all closed itemsets.

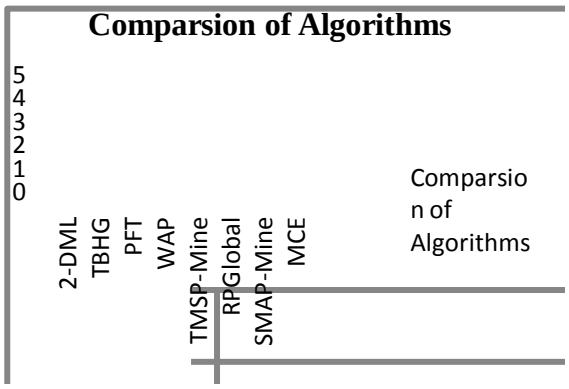
G. SMAP-Mine and SMAP-Tree

SMAP-Mine is used to discover mobile users' sequential movement patterns with the service that was requested. By using the mobile, we can get thevarious kinds of services in any time. To predict the user behavior is an difficult task, so they proposed adata mining method called SMAP-Mine [5], is used to predict the sequential mobile patterns, from that we can efficiently discover the user behavior. The SMAP-Tree is used to combine the access patterns, to mine efficient frequent pattern. Only one scan is needed to mine the frequent pattern**.** The performance is accurate, scalable and efficient.

H. Mobile Commerce Explorer

MCE is mainly used to predict the user behaviors such as their movements and purchase transaction. GPS [1] is used to record the user movement history. Its goal is to predict the behavior of individual users. The mobile network database 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 [13] has three factors,

- *Similarity Inference Model (SIM)*- measure the similarities among stores and items.
- *PMCP-Mine Algorithm*- efficient discovery of mobile users.
- *Mobile Commerce Behavior Prediction (MCBP)*- predicting the mobile user behavior.



Graph: Comparsion of Algorithms

VI. COMPARSION OF VARIOUS TECHNIQUES

The comparsions 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 efficient	Network nodes, no.of users, services	40%
Tree Based Hierarchical Graph	107 Datasets	Understand the surroundings and plan their journey efficiently	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%
Web Access Pattern	Web logs	Scalability, efficient search space and efficient data structure	Label and Count	75%
TMSP-Mine	Mobile Transaction log	Efficient and Accurate	Network nodes, Time interval and Fitness function	0.6%
Mining Compressed Frequent Pattern sets	Accident and chess datasets	Efficient and good compression quality	Running time and Representative pattern	0.4%
SMAP-Mine	Real datasets	Accuracy, execution, efficiency and scalaibilty	Network nodes, mobility events, no.of different services	0.001 to 0.005%
Mobile Commerce Explorer	Synthetic datasets	Accurate, efficient	Mesh network, no.of users, no.of items	0.1%

VII. CONCLUSION

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 used.

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 Patterns," Int'l J. Computer Science Security, vol. 7, no. 2, pp. 285-291, Feb. 2007.

[7] J. Pei, J. Han, B. Mortazavi-Asl, and H. Zhu, "Mining Access Patterns Efficiently from Web Logs," Proc. Pacific Asia Conf. Knowledge Discovery and Data Mining, pp. 396-407, Apr. 2000.

[8] D. Xin, J. Han, X. Yan, and H.Cheng, "Mining Compressed Frequent-Pattern Sets," Proc. Int'l Conf. Very Large Data Bases,pp. 709-720, Aug. 2005.

[9] U. Varshney, R.J. Vetter, and R. Kalakota, "Mobile Commerce: A New Frontier," Computer, vol. 33, no. 10, pp. 32-38, Oct. 2000.

[10] R. Agrawal, T. Imielinski, and A. Swami, "Mining Association Rule between Sets of Items in Large Databases," Proc. ACM SIGMOD Conf. Management of Data, pp. 207-216, May 1993.

[11] Reshma Cherian, R.Bright Gee Varghese, "Individual Mining And Prediction Of Patterns For Imporving Mobile Commerce," International Journal of Computer Applications, volume 66-No.3, March 2013.

[12] www.cis.temple.edu/~vasilis/courses/cis664/slides/Associations.ppt

[13] Eric Hsueh-Chan Lu, Wang-Chien Lee and Vincent S.Teng, "A Framework for Personal Mobile Commerce Pattern Mining and Prediction," IEEE Transactions on Knowledge and Data Engineering, vol.24, no.5, May 2012.

[14] J. Han, J. Pei, and Y. Yin, "Mining Frequent Patterns without Candidate Generation," Proc. ACM SIGMOD Conf. Management of Data, pp. 1-12, May 2000.

[15] J. Han and M. Kamber, Data Mining: Concepts and Techniques,second ed. Morgan Kaufmann, Sept. 2000.

Jacinth Evangeline S received B.E degree in computer science and Engineering from Nandha Engineering College, Erode. She is currently pursuing her M.E degree in computer science and Engineering in Erode Sengunthar Engineering College. She has published 3 papers in national conferences. Her Research interest includes Data Mining and Database Management System.

Subramanian K.M received his B.Sc. degree in Computer Science from Cheran Arts & Science College, Kangayam (Bharathiar University), Tamilnadu. The M.C.A degree from Bharathidasan University, Tiruchirappalli, and M.E. Degree in Computer Science & Engineering from Mahendra Engineering College, Salem. His research interests include Data Mining, Data Warehousing, Web Mining and Neural Networks. At present he is working as a Assistant Professor – Selection Grade-II, Department of Computer Science & Engineering, Erode Sengunthar Engineering College, and pursuing his Ph.D., degree in Anna University, Chennai, India. He has published 3 international journals and 2 national journals. He is a ISTE Life member and CSI Institutional member.

Dr. Venkatachalam K received B.E degree in Electronics and Communication Engineering from Bharathiar University, Coimbatore, M.Tech degree in Electronic and Communication System- First class with Distinction and Universty 2nd rank from Pondicherry University and Ph.D degree in Information Communication Engineering from Anna University, Gunidy and Presently he is working as Professor in Department of Electronics and Communication Engineering, Velalar College of Engineering and Technology, Erode. His Area of Specialization includes Wireless Mobile Communication, Optical Networks, Wireless Embedded Networks and Low Power VLSI Design.