International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 2 Issue 1, January- 2013

Efficient Mining and Prognostication of User Behaviours in Mobile Commerce

Hepzibha. D
II M.E, Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology ,Coimbatore, India.

Narmadha. R. P.
Assistant Professor, Department of Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India.

Abstract

Mobile computing technology which created an unprecedented opportunity to gather and extract information from mobile agent systems has rapidly expanded in recent days. Among them, one of the active topic areas is the mining and prediction of users' mobile commerce behaviours such as their movements and purchase transactions. In this paper, a novel model is proposed which is called Mobile Commerce Explorer (MCE), for mining and prediction of mobile users' movements and purchase transactions. The MCE framework consists of three major components: 1) Similarity Inference Model (SIM) for measuring the similarities among stores and items; 2) Personal Mobile Commerce Pattern Mine(PMCP-Mine) algorithm for efficient discovery of mobile users' Personal Mobile Commerce Patterns (PMCPs); and 3) Mobile Commerce Behaviour Predictor (MCBP) for prediction of possible mobile user behaviours. Furthermore, in future it can be applied to the MCE model to location-based services, which aims to achieve greater precision in predicting object behaviours.

Keywords — Data mining, mobile agent, electronic commerce.

1. Introduction

Look at a supermarket with a large solicitation of items. The major business decisions that the governing body of the supermarket makes include what sale has to be put on, how coupons are to be designed, how the merchandise are placed on shelves for the purpose of maximizing the profit. The most commonly used approach is to improve the quality of such decisions which becomes the analysis of past data. However, till recent days, the world-wide data is obtained with the method of increasing successive addition during a particular period of time say a day, a week or a month that was available on the system. The advancement in bar-code technology has greatly influenced to store the items purchased on a per transaction basis. These basket data type transactions need not necessarily include items that are bought together at the one point of time. It may also include items purchased by a customer over a certain period of time. A track record of these data generally includes the transaction date and the items purchased in the transaction. Eminent managements view these databases as significant piece of the marketing infrastructure. They are more specific in instituting information-driven marketing processes, managed by database technology which enables marketers to develop and implement, modify marketing agenda and strategies. Here we bring out the trouble of mining sequential patterns over this data. A typical example of such a pattern is that customers usually rent \Mission Impossible", then \Mummy Returns", and then \Transporter ". It is not always necessary that these rentals be consecutive. Customers who tend to rent other videos also support this sequential pattern. Items present in a sequential pattern need not be simple ones. \Bed Spread and pillow cases", followed by \curtains and mattresses", followed by \gult", is an example of a sequential pattern in which the items present are a sets of items.

The increase in the usage of computing for a variety of applications has influenced the importance of database mining at a rapid pace recently. Catalog companies can accumulate sales data from the orders received. The analysis of past transaction data can give valuable information on customer's buying behavior which in turn develops the quality of business decisions such as what sale has to be put on, how coupons are to be designed, how the merchandise are placed on shelves for the purpose of maximizing the profit, how to customize marketing programs, etc. Hence before any meaningful conclusion, it is essential to collect a sufficient amount of sales data. As a result, the quantity of these processed data tends to be large. It is therefore essential to design effective algorithms to perform mining on these data.

Digitized geographic data has lead to the growing demand for the spatial database technology to govern large volume of spatial information. Moreover, mobile users' efficient data management has become more crucial, as the airing use of mobile devices. The important database research areas are the development of *spatiotemporal database* technologies to support moving objects. In the field of researching movable object databases, exploitation of effective indexing techniques is one of the crucial issues and there exist many proposals of *spatio-temporal indexes*. Additionally, there are some proposals from spatio-temporal databases for the extraction of statistical information. Statistics involving spatio-temporal data is not only helpful for the efficient query processing but also in *mobility analysis* to examine the moving patterns of objects from collected spatio-temporal trajectory data. Since collected trajectory data may have huge volume, we need an efficient method to calculate statistics. The Markov chain model in spatio-temporal data analysis is used for analyzing movement tendency of moving objects such as how population moves from a certain region to other regions while a specified period. Using such statistical information, we can estimate with high probability whether an object at some region will move to another region in the next period.

To provide a high-precision mobile commerce behavior predictor (MCBP), personal mobile pattern mining must be emphasized. Inorder to overcome the prediction failure problem, similarities of stores and items into the mobile commerce behavior prediction can be incorporated. Therefore, in this paper, a model namely Mobile Commerce Explorer (MCE), to mine and predict mobile users' trend and transactions under the circumstance of mobile commerce is proposed. The MCE model generally include three major components: 1) Similarity Inference Model for calculating the similarities among stores and items; 2) Personal Mobile Commerce Pattern Mine (PMCP-Mine) algorithm for effectively discovering the mobile users' Personal Mobile Commerce Patterns (PMCPs); and 3) Mobile Commerce Behavior Predictor (MCBP) for predicting the possible mobile user characteristics. This is the first work (to our best knowledge) that facilitates mining and prediction of mobile users' commerce behaviors that may suggest stores and items not known to a user previously. Finally, through evaluation, our proposals deliver an excellent performance in terms of predicting efficient user behavior. The contributions and advantages of this paper are

- The MCE model proposes the approach for mining and prediction of mobile commerce behavior.
- A novel model SIM proposed for measuring the similarities among stores and items automatically from a mobile transaction database.
- A PMCP-Mine algorithm for mining PMCPs from a mobile transaction database is proposed to better understand the personal mobile behaviors.
- MCBP prediction technique is used for accurate prediction of the mobile behaviors which include the trends and transactions of a user which is based on the SIM and PMCP-Mine.

2. Related Works

Reviews and classification of relevant previous studies can be divided into three categories in this section: 1) measuring similarity, 2) techniques for mobile pattern mining, and 3) behavior predictions for mobile.

2.1. Similarity Measure:

There have been many studies on measuring the similarity between two objects. Multiple-level hierarchical structure was the first one in this category. In[1], Lu first proposes the concept of multiple-level hierarchical structure in data mining. In[2], Han and Fu propose the multiple-level association rules mining. Representing the hierarchical relations of items in terms of taxonomy is incorporated in this study. In [3], Tseng and Tsui, initially uses the multiple-level hierarchical concept to mine related service patterns in mobile web environments. The items in the same level are regarded as similar items, based on the structure. The relations between the items in the different levels are not known. Sequence alignments is the second one. SimRank is proposed to iteratively compute the similarities between objects in [4] by Jeh and Widom. Two objects are similar if they are related to same type of objects is the idea behind SimRank. The efficiency of SimRank is improved, in [5], Yin et al., by developing the hierarchical structure called SimTree which reduces the cost of computation and the storage of similarities among objects but still identify the relationships between objects. A pattern distance measure based on set similarity (SET) between two association patterns is proposed in [6], Xin et al. To apply Jaccard Measure to calculate the similarity of two sets is the basic concept of set similarity. Let S1 and S2 be two sets, the set similarity set_similarity(S1; S2). However, set similarity is not applicable to store similarity in mobile commerce. For example, there are two stores 1 and 2 which only provides Bread and Butter, respectively. The similarity of store 1 and store 2 should not be 0, since Bread and Butter belong to the same category.

2.2. Mobile Pattern Mining:

Now-a-days, the usage of data mining techniques to discover useful rules/patterns from WWW, transaction databases [7], [8], [9] and mobility data [10] have been discussed in a number of studies. Mining association rules [7] are proposed to find important items in a transaction database. To mine the association rules[8], Agrawal and Srikant, proposes the Apriori algorithm. The DHP algorithm to improve the performance of an association rule mining was proposed in [11], Park et al. An algorithm named WAP-Mine to effectively identify the web access patterns in web logs which uses a tree-based data structure without candidate generation was proposed by Pei et al. In [9], sequential pattern mining was first introduced to search sequential patterns within transaction databases. Considering the relation between location and service, in[12], Chen et al., proposed the path traversal patterns for mining web user behaviors. Tseng and Tsui, [3] first study the problem of mining associated service patterns in mobile web environments. Tseng et al proposed SMAP-Mine for efficiently mining users' sequential mobile access patterns, based on the FP-Tree. Lee et al., propose T-MAP [10] to efficiently find the mobile users' mobile access patterns in distinct time intervals. Mobile Sequential Pattern [13] was proposed by Yun and Chen to take moving paths into consideration and add the moving path between the left hand and the right hand in the content of rules. In[3], Tseng et al., propose the TMSP-Mine for discovering the temporal mobile sequence patterns in a location-based service environment. A prediction approach called Hybrid Prediction Model for estimating an object's future locations based on its pattern information was proposed by Jeung. An object's movements are more crucial than what the mathematical formulas can represent is what is considered in this paper. There is no work considering the user relations in the mobile pattern mining.

2.3. Mobile Behavior Prediction:

Mobile behavior predictions can be roughly divided into two categories. Vector-based prediction is the first category which can be further divided into two types: 1) linear models [14] and 2) nonlinear models. Capturing objects' movements with sophisticated regression functions are nonlinear models and thus predicting accuracies that are higher than those of the linear models. Recursive Motion Function (RMF) is the most accurate prediction method in the literature based on regression functions. Pattern-based prediction comes under the second category. A Markov Model (MM) is derived in [15], Ishikawa et al., which generates Markov transition probabilities from one cell to another for predicting the next cell of the object. However, these methods can only predict the next spatial locations of objects. SMAP-Min has been proposed to discover sequential mobile access rules and predict the user's next locations and services. In [13], Yun and Chen, propose the Mobile Sequential Pattern to predict the next mobile behaviors. Collaborative Filtering (CF) is an efficient idea that may be applied to the prediction of user's behavior. This can be divided into two types: 1) user-based collaborative filtering and 2) item-based collaborative filtering. The first type is based on the behaviors of similar users. For example, if Ben and Bobby are similar based on their trends or transactions, then we refer to the behaviors of Bobby to predict the next behavior of Ben. Still, it is not necessary that their behaviors are always similar even if the two users are very similar. In the second type, the prediction concept is based on user behavior associated with similar items. For example, suppose that Cheese and Ghee are similar based on their item categories or properties. Thus if a user has purchased Ghee, then, when we try to predict the next behavior of this user, the next behavior after the user purchases Cheese is to be referred. It is still difficult to find the similarity between items. Collaborative filtering techniques depend on users' ratings on items to predict the purchase behavior of the user.

3. System Framework

The proposed MCE model consists of three modules, 1) a mobile database, 2) a data mining mechanism, and 3) a behavior prediction engine (See Fig. 1).

Detailed store information which includes locations are maintained in the mobile network database. Our system has an "offline" mechanism for similarity inference and PMCPs mining, and an "online" engine for mobile commerce behavior prediction. Mobile transaction database stores the mobile information which includes user identification, stores, and item purchased, when mobile users tend to move between the stores. The offline data mining mechanism develops the SIM model and the PMCP-Mine algorithm to discover the store/item similarities and the

PMCPs. The online prediction engine proposes a MCBP which is based on the store and item similarities as well as the mined Personal Mobile Commerce Patterns. When a mobile user purchases items among different stores, then the next purchase will be predicted based on the mobile user's identification and recent mobile transactions. The model is used to assist with the prediction of next movement and transaction.

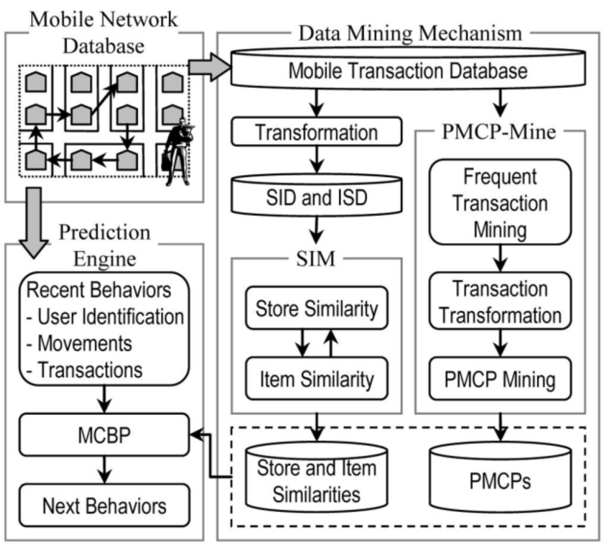

Fig 1.The mobile commerce explorer framework

3.1. Similarity Inference Model

The crucial task in our model is to identify the similarities between stores and items. The solution for this problem may be solved by using store and item category ontology. The store or item ontology need not necessarily match with the mobile transaction database. The main goal is to automatically calculate the store and item similarities from the mobile transaction database, which in turn captivates mobile users' moving and transactional behaviors. The database will have the following information available: 1) In a given store, we will know the items that are available for sale; 2) In a given item, we will know the stores that sell this item. This information can support us to understand which stores or items are similar. We take a note that the users usually purchase similar items in particular stores, then these stores may be weighed as similar. For example, user may purchase burgers, Pizzas, or Cokes in Dominos and Pizza Hut, we weigh them as similar stores. A parameter-less data mining model is proposed namely Similarity Inference Model to manage the task of calculating store and item similarities.

3.2. Discovery of PMCPs

Description of the PMCP-Mine algorithm to extract the personal mobile commerce patterns efficiently is discussed in this section. The PMCP-Mine algorithm is inspired by the TJPF algorithm [13] which is an Apriori-like algorithm. Still, it is noticed that the TJPF algorithm do not take into account user identification, which is essential for finding out personal mobile behaviors. Bottom-up fashion is performed in the PMCP-Mine algorithm. We first discover frequent transaction behaviors in a single store, e.g., {Starbucks, Latte}. Then, these single patterns can be joined to form compound patterns e.g., {Hang Ten, clothes, Starbucks, Latte}. Eventually, the PMCP-Mine algorithm helps in obtaining the complete mobile commerce patterns. The algorithm is divided into three main categories: 1) Frequent-Transaction Mining. A pair of store and items indicating frequently made purchasing transactions is known as Frequent-Transaction. All Frequent-Transactions for each user is first discovered in this phase. 2) Mobile Transaction Database Transformation. By deleting least recently purchased items, the original mobile transaction database can be reduced which is based on all Frequent-Transactions. The main objective is to improve the database scan efficiency for pattern support counting. 3) PMCP Mining. In this phase, mining all patterns of length k from patterns of length k_1 in a bottom-up fashion is discussed.

3.3. Mobile Commerce Behavior Predictor

In this section, how to use the discovered PMCPs to predict the users' future mobile commerce behaviors that include movements and transactions are described. In the available pattern-based prediction models, exact matching is based on the pattern selection strategy, i.e., the similarity between various locations is treated as 0. These may lead to prediction failures when there is no available pattern to match. Integrating the similarities of stores and items that are obtained from SIM into the mobile commerce behavior prediction is performed to overcome the above issue. MCBP is proposed which calculates the similarity score of every PMCP. It is calculated from users' recent mobile commerce behavior by taking store and item similarities into account. Three ideas are considered in MCBP: 1) PMCPs with high similarity to the user's recent mobile commerce behavior are considered as prediction knowledge; 2) A greater effect on next mobile commerce behavior predictions can be obtained from more recent mobile commerce behaviors; and 3)

PMCPs with higher support provide greater confidence for predicting users' next mobile commerce behavior.

4. Conclusion

Research on mobile commerce has received a lot of interests from both the industry and academics. Mining and prediction of users' mobile commerce behaviours such as their movements and purchase transactions is one of the active topic areas among them. In this paper, we have proposed a novel framework, namely mobile commerce explorer (MCE) for mining and prediction of mobile users' movements and transactions in mobile commerce environments. The three major techniques in the MCE model are: 1) Similarity Interference Model for measuring the similarities among stores and items; 2) Personal Mobile Commerce Pattern(PMCP)-Mine algorithm for effectively discovering mobile users' PMCPs; and 3) Mobile Commerce Behaviour Predictor(MCBP) for predicting possible mobile user behaviours is proposed in this paper. This is the first work that facilitates mining and prediction of personal mobile commerce behaviours that may recommend stores and items previously unknown to a user. The experimental results show that the MCE model achieves a very high precision in mobile commerce behaviour predictions.

5. References

[1] Y. Lu, "Concept Hierarchy in Data Mining: Specification, Generation and Implementation," master's thesis, Simon Fraser Univ., 1997.

[2] J.Han and Y. Fu, "Discovery of Multiple-Level Association Rules in Large Database," Proc. Int'l Conf. Very Large Data Bases, pp. 420-431, Sept. 1995.

[3] 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.

[4] G. Jeh and J. Widom, "SimRank: A Measure of Structural-Context Similarity," Proc. Int'l Conf. Knowledge Discovery and Data Mining, pp. 538-543, July 2002.

[5] X. Yin, J. Han, and P.S. Yu, "LinkClus: Efficient Clustering via Heterogeneous Semantic Links," Proc. Int'l Conf. Very Large Data Bases, pp. 427-438, Aug. 2006.

[6] 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.

[7] 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.

[8] R. Agrawal and R. Srikant, "Fast Algorithm for Mining Association Rules," Proc. Int'l Conf. Very Large Databases, pp. 478-499, Sept. 1994.

[9] R. Agrawal and R. Srikant, "Mining Sequential Patterns," Proc. Int'l Conf. Data Eng., pp. 3-14, Mar. 1995.

[10] 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.

[11] J.-S. Park, M.-S. Chan, and P.S. Yu, "An Effective Hash Based Algorithm for Mining Association Rules," Proc. ACM SIGMOD Conf. Management of Data, pp. 175-186, May 1995.

[12] M.S. Chen, J.-S. Park and P.S. Yu, "Efficient Data Mining for Path Traversal Patterns," IEEE Trans. Knowledge and Data Eng., vol. 10, no. 2, pp. 209-221, Apr. 1998.

[13] 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.

[14] J.M. Patel, Y. Chen, and V.P. Chakka, "Stripes: An Efficient Index for Predicted Trajectories," Proc. ACM SIGMOD Conf. Management of Data, pp. 635-646, June 2004.

[15] Y. Ishikawa, Y. Tsukamoto, and H. Kitagawa,"Extracting Mobility Statistics from Indexed Spatio-Temporal DataSets," Proc. Workshop Spatio-Temporal Database Management, pp. 9-16, Aug. 2004.