Volume : 2 | Issue : 6 | June 2013 • ISSN No 2277 - 8160

Research Paper | Computer Science

Data Pre-Fetching Framework using Sequential Patterns for Mobile User Behavior Mining

D. Priyadharshini PG Scholar, Department of Computer Science and Engineering, Angel College of Engineering and Technology, Tirupur, Tamilnadu-641665.

S.Sundaramoorthy Assistant Professor, Department of Computer Science and Engineering, Angel College of Engineering and Technology, Tirupur, Tamilnadu-641665.

ABSTRACT

Data mining is a widely used technique for discovering valuable information in a complex data set. 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. We propose a novel algorithm with GPS navigation method, 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 proposed to find segmenting time intervals where similar mobile characteristics exist. Here the mining and prediction of mobile behaviors with considerations of user relations and temporal property simultaneously can be predicted easily. Through experimental evaluation under various simulated conditions, the proposed methods are used to deliver excellent performance.

KEYWORDS: Data Mining, Mobile User Behavior, Mining Techniques, Sequential Patterns.

1. INTRODUCTION

Data mining is a widely used technique for discovering valuable information in a complex data set. Data mining (the analysis step of the "Knowledge Discovery in Databases" process, or KDD), a field at the intersection of computer science and statistics, is the process that attempts to discover patterns in large data sets. It utilizes methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Aside from the raw analysis step, it involves database and data management aspects, data preprocessing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating.

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 such as their movements and transactions. The time interval segmentation method helps to find various user behaviors in different time intervals. To find complete mobile behavior patterns, a time interval table is required. In some cases, it is 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 time interval table by users. Automatic time segmentation methods are, thus, required to segment the time dimension in a mobile transaction database.

Here, we propose a novel data mining algorithm named Cluster-based Temporal Mobile Sequential Pattern Mine (CTMSP-Mine) to efficiently mine the Cluster-based Temporal Mobile Sequential Patterns (CTMSPs) of users. Then, novel prediction strategies are proposed to effectively predict the user's subsequent behaviors using the discovered CTMSPs. To mine CTMSPs, we first propose a transaction clustering algorithm named Cluster-Object-based Smart Cluster Affinity Search Technique that builds a cluster model for mobile transactions based on the proposed Location-Based Service Alignment (LBS-Alignment) similarity measure. Based on the produced user clusters and the time interval table, all CTMSPs can be discovered by the proposed method.

To our best knowledge, this is the first work on mining and prediction of mobile sequential patterns by considering user clusters and temporal relations in LBS environments simultaneously. Besides, the proposed CTMSPs provide information including both user clusters and temporal relations.

2. PROPOSED WORK

In proposed system, we review and classify relevant work into three categories: i) Similarity Measures, ii) Mobile Pattern Mining Techniques, and iii) Mobile Behavior Predictions.

i. Similarity Measure: There have been many studies on measuring the similarity between two objects. The first one is based on multiple-level hierarchical structures. The concept of set similarity is to apply Measure to calculate the similarity of two sets. Let S1 and S2 be two sets, the set similarity set similarity (S1; S2) is defined as (1). However, set similarity is not applicable to store similarity in mobile commerce.

ii. Mobile Pattern Mining: Here to propose the TMSP-Mine for discovering the temporal mobile sequence patterns in a location-based service environment. To propose a prediction approach called Hybrid Prediction Model for estimating an object's future locations based on its pattern information. This considers that an object's movements are more complicated than what the mathematical formulas can represent. However, there is no work consider user relations in the mobile pattern mining.

iii. Mobile Behavior Prediction: The mobile behavior predictions can be roughly divided into two categories. The first category is a vector-based prediction that can be further divided into two types: 1) linear models, and 2) nonlinear models. The nonlinear models capture objects' movements with sophisticate regression functions. Thus, their prediction accuracies 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. The second category is a pattern-based prediction.

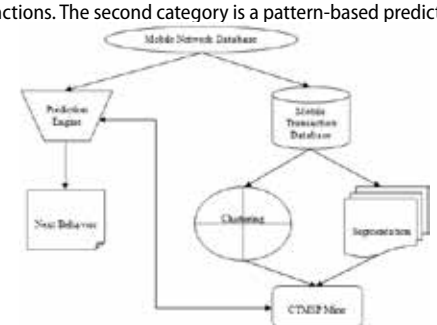


Fig 1: System Architecture

In a mobile transaction database, users in the different user groups may have different mobile transaction behaviors. The first task we have to tackle is to cluster mobile transaction sequences. We proposed a parameter-less clustering algorithm CO-Smart-CAST. In this proposed system, we describe our system design.

There are 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.

Our system has an "offline" mechanism for CTMSPs mining and an "online" engine for mobile behavior prediction. When mobile users move within the mobile network, the information which includes time, locations, and service requests will be stored in the mobile transaction database.

In the offline data mining mechanism, we design two techniques and the CTMSP-Mine algorithm to discover the knowledge. So, we propose the CO-Smart-CAST algorithm to cluster the mobile transaction sequences. In this algorithm, we propose the LBS-Alignment to evaluate the similarity of mobile transaction sequences.

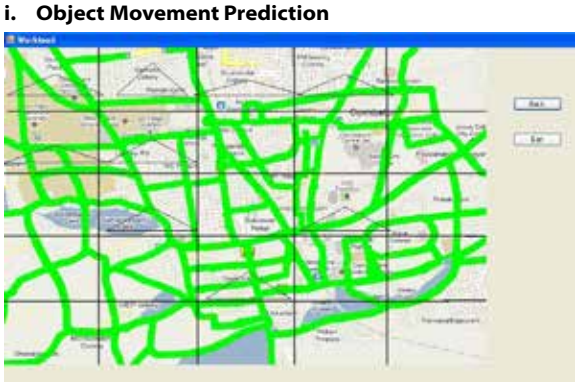
3. EVALUATION

In this section, we conducted a series of experiments to evaluate the performance of the proposed CTMSP-Mine, under various system conditions. Experiments can be divided into three parts: 1) user clustering, 2) time interval segmentation, and 3) precision of prediction.

We compare the proposed CTMSP Mine method with four other methods:

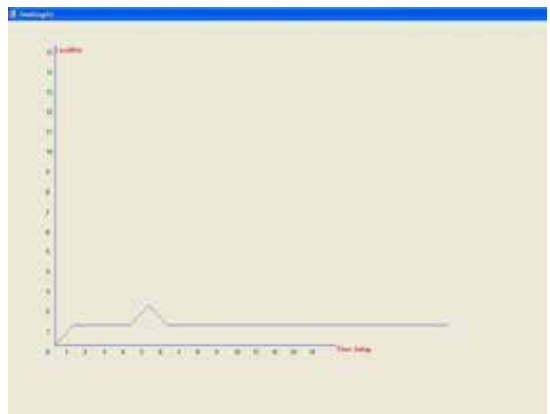
1. Recursive Motion Function.
2. Markov Model.
3. Hybrid Prediction Model.
4. Mobile Sequential Pattern.

i. Object Movement Prediction



ii. Graph Generation

iii. Time Graph



In this series of experiments, our major purpose is to measure the precisions and recalls of mobile behavior predictions by various methods.

4. CONCLUSION

In this work, we have proposed a novel method, named CTMSP-Mine, for discovering CTMSPs in LBS environments. 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 CO-Smart- CAST to form user clusters based on the mobile transactions using the proposed LBS-Alignment similarity measurement. 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.

5. FUTURE WORK

In future work, we will apply our method to real data sets. We will also try to integrate CTMSP-Mine with HPM or RMF as a hybrid scheme and design more sophisticated strategies. In addition, we will apply the CTMSP-Mine to other applications with the aim to enhance precision for predicting user behaviors. Then, also try to take advantage of the Genetic Algorithm (GA) to produce a more suitable time interval table. There may also have chance to use two nonparametric techniques to increasing the predictive precision of the mobile users' behaviors.

REFERENCES

[1]. Eric Hsueh-Chan Lu, "Mining Cluster based Temporal Mobile Sequential Patterns in Location Based Service Environments, "IEEE Transactions on Knowledge and Data Engineering", Jun 2011. | [2]. Wang-Chien Lee, "A Framework for Personal Mobile Commerce Pattern Mining and Prediction, "IEEE Transactions on Knowledge and Data Engineering", May 2012. | [3]. Liang Wang, "Efficient Mining of Frequent Item Sets on Large Uncertain Databases", "IEEE Transactions on Knowledge and Data Engineering", Dec 2012. | [4]. Xiang Lian, "Ranked Query Processing in Uncertain Databases", "IEEE Transactions on Knowledge and Data Engineering", Mar 2010. | [5]. Yunhao Liu, "Mining Frequent Trajectory Patterns for Activity Monitoring using Radio Frequency Tag Arrays", "Ieee Transactions on Parallel and Distributed Systems", Nov 2012. | [6]. Avrilia Floratou, "Efficient and Accurate Discovery of Patterns in Sequence Data Sets", "IEEE Transactions on Knowledge and Data Engineering", Aug 2011.