ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)

Vol. 2, Special Issue 10, March 2016

AN EFFECTIVE FILTERING DYNAMIC TRANSACTION DATASETS USING SERIESWITH LOCATION BASE SERVICE MODEL

S.Nanthini[1],R.Sivaraj[2]

M.E, Department of CSE, Velalar College of Engineering and Technology, Thindal, Erode, Tamilnadu1
HOD, Department of CSE, Velalar College of Engineering and Technology, Thindal, Erode, Tamilnadu2

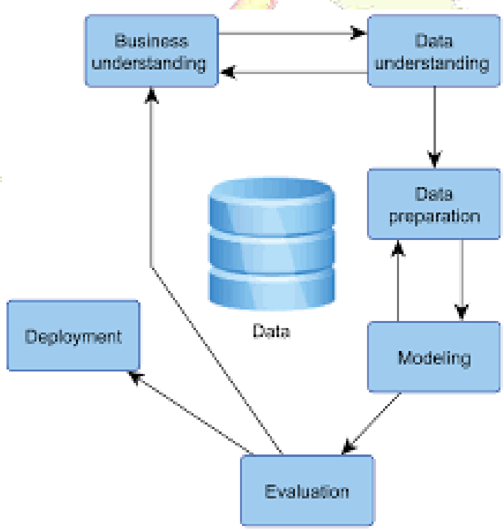
ABSTRACT

Location-based services have been widely adopted in many systems. Existing works employ a pull model or user-initiated model, where a user issues a query to a server which replies with location-aware answers. We propose an R-tree based index by integrating textual descriptions into R-tree nodes. Our method can support both conjunctive queries and ranking queries. We discuss how to support dynamic updates efficiently.

Keywords: Location-based services, frequent items, Similarity Measures, Cluster-based Temporal Mobile Sequential Pattern Mine.

1. INTRODUCTION

Data mining is an interdisciplinary subfield of computer science. It is the computational process of discovering patterns in large **data** sets ("big **data**") involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems.



Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data set. These tools can include statistical models, mathematical algorithm and machine learning methods. Consequently, data mining consists of more than collection and

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

managing data, it also includes analysis and prediction. Classification technique is capable of processing a wider variety of data than regression and is growing in popularity. Data mining refers to extracting or "mining" knowledge from large amounts of data. The term is actually a misnomer. Remember that the mining of gold from rocks or sand is referred to as gold mining rather than rock or sand mining. Thus, data mining should have been more appropriately named "knowledge mining from data," which is unfortunately somewhat long. "Knowledge mining," a shorter term may not reflect the emphasis on mining from large amounts of data. Nevertheless, mining is a vivid term characterizing the process that finds a small set of precious nuggets from a great deal of raw material.

2. CTMSP TECHNIQUES

Clustering mobile transaction data helps in the discovery of social groups, which are used in applications such as targeted advertising, shared data allocation, and personalization of content services. In previous studies, users are typically clustered according to their personal profiles (e.g., age, gender, and occupation).However, in real applications of mobile environments, it is often difficult to obtain users profiles. That is, we may only have access to users mobile transaction data. To achieve the goal of user clustering without user profiles, the evaluation of similarities of mobile transaction sequences (MTSs) is required. The proposed system develops a novel algorithm, namely, Cluster-based Temporal Mobile Sequential Pattern Mine (CTMSP-Mine), to discover the Cluster-based Temporal Mobile Sequential Patterns (CTMSPs). Since, a prediction strategy is proposed to predict the subsequent mobile behaviors, in CTMSP-Mine, user clusters are constructed by a novel algorithm named Cluster Affinity Search Technique (CAST) and similarities between users are evaluated by the proposed measure, Location-Based Service Alignment (LBS-Alignment). At the same time, a time segmentation approach is presented to find segmenting time intervals where similar mobile characteristics exist. The project considers mining and prediction of mobile behaviors with considerations of user relations and temporal property simultaneously.

3. RELATED WORKS

Cao.X et al . The location-aware keyword query returns ranked objects that are near a query location and that have textual descriptions that match query keywords. This query occurs inherently in many types of mobile and traditional web services and applications, e.g., Yellow Pages and Maps services. Previous work considers the potential results of such a query as being independent when ranking them. However, a relevant result object with nearby objects that are also relevant to the query is likely to be preferable over a relevant object without relevant nearby objects. They are proposes the concept of prestige-based relevance to capture both the textual relevance of an object to a query and the effects of nearby objects. Based on this, a new type of query, the Location-aware top-k Prestige-based Text retrieval query, is proposed that retrieves the top-k spatial web objects ranked according to both prestige-based relevance and location proximity. A straightforward approach to computing the query is to adapt the algorithms for computing Personalized Page Rank Vectors (PPVs) to computing the PR scores of all objects in the spatial object graph, to use an R-tree for computing spatial distances, and then to combine the two scores. However, this solution is expensive.

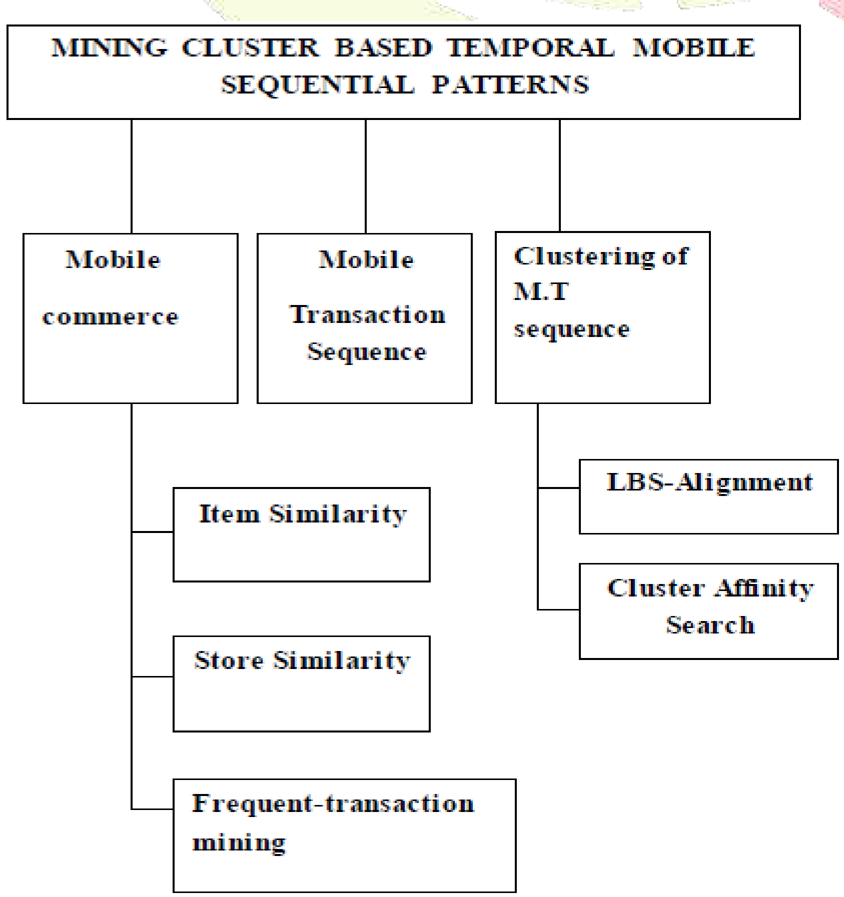
ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)

Vol. 2, Special Issue 10, March 2016

Ooi.B.C With the proliferation of geo-positioning and geo-tagging, spatial web objects that possess both a geographical location and a textual description are gaining in prevalence, and spatial keyword queries that exploit both location and textual description are gaining in prominence[4]. However, the queries studied so far generally focus on finding individual objects that each satisfy a query rather than finding groups of objects where the objects in a group collectively satisfy a query. To address the need for such collective answers to spatial keyword queries,we assume a database of spatial web objects and then consider the problem of how to retrieve a group of spatial objects that collectively meet the user‟s needs given as location and a set of keywords: 1) the textual description of the group of objects must cover the query keywords, 2) the objects are close to the query point, and 3)the objects in the group are close to each other.

Approximation Algorithm

Show that the first subproblem is NP -complete by a reduction from the Weighted Set Cover (WSC) problem in Lemma. The reduction in the proof is approximation preserving. Thus, the approximation properties of the WSC problem carry over to the problem. The IR-tree is essentially an R-tree extended with inverted files. Each leaf node in the IR-tree contains entries of the form (o, o.λ, o.di), where o refers to an object in datasetD, o.λ is the bounding rectangle of o, and o.di is an identifier of the description of o. Each leaf node also contains a pointer to an inverted file with the keywords of the objects stored in the node. An inverted file index has two main components. A vocabulary of all distinct words appearing in the description of an object. A posting list for each word t that is a sequence of identifiers of the objects whose descriptions contain t.



ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)

Vol. 2, Special Issue 10, March 2016

4.ALGORITHM IMPLEMENTATION

a) Item Similarity

In this module, the items are added with their score values. Then similarity between two items are found out using the given similarity formula. Likewise similarity matrix is found out for all the items in the list.

b) Store Similarity

In this module, the stores are added with their score values. Then similarity between two store are found out using the given similarity formula. Likewise similarity matrix is found out for all the stores in the list.

c) Clustering of mobile transaction sequences

In a mobile transaction database, users in the different user groups may have different mobile transaction behaviors. The first task to tackle is to cluster mobile transaction sequences. In this module, a parameter-less clustering algorithm called CAST is proposed. Before performing the CAST, a similarity matrix S is to be generated, based on the mobile transaction database. The entry Sij in matrix S represents the similarity of the mobile transaction sequences i and j in the database, with the degrees in the range of [0, 1]. A mobile transaction sequence can be viewed as a sequence string, where each element in the string indicates a mobile transaction. The major challenge to tackle is to measure the content similarity between mobile transactions. The LBS-Alignment algorithm is proposed, which can obtain the similarity. LBS-Alignment is based on the consideration that two mobile transaction sequences are more similar, when the orders and timestamps of their mobile transactions are more similar. CAST algorithm is used to cluster the users.

d) Segmentation of Mobile Transactions

In a mobile transaction database, similar mobile behaviors exist under some certain time segments. Hence, it is important to make suitable settings for time segmentation so as to discriminate the characteristics of mobile behaviors under different time segments. A new time segmentation method is proposed to automatically obtain the most suitable time segmentation table with common mobile behaviors. The algorithm below shows the procedure of the proposed time segmentation method, named Get Number of Time Segmenting Points (GetNTSP) algorithm.

e) Discovery of CTMSPs

In order to mine the cluster-based temporal mobile sequential patterns efficiently, we proposed a novel method named CTMSP-Mine to achieve this mining procedure. In CTMSP-Mine, both factors of user cluster and time interval are taken into account such that the complete mobile sequential patterns can be discovered. The entire procedures of CTMSP-Mine algorithm can be divided into three main steps:

1) Frequent-Transaction Mining,
2) Mobile Transaction Database Transformation, and

3) CTMSP Mining.

1) Frequent-Transaction Mining

In this phase, the frequent transactions (F-Transactions) are mined in each user cluster and time interval by applying a modified Apriori algorithm.

2) Mobile Transaction Database Transformation

In this phase, F-Transactions are used to transform each mobile transaction sequence S into a frequent mobile transaction sequence S'. According to Table 3, if a transaction T in S is frequent, T would be transformed into the corresponding F-Transaction. Otherwise, the cell of T would be transformed into a part of path.

3) CTMSP Mining

In this phase, all the CTMSPs are mined from the frequent mobile transaction database. Frequent 1-CTMSPs are obtained in the frequent-transaction mining phase. In the mining algorithm, we utilize a two-level tree named Cluster-based Temporal Mobile Sequential Pattern Tree (CTMSP-Tree). The internal nodes in the tree store the frequent mobile transactions, and the leaf nodes store the corresponding paths. Moreover, every parent node of a leaf node is designed as a hash table which stores the combinations of user cluster tables and time interval tables.

5. CONCLUSION

In this thesis, a novel method named CTMSP-Mine is proposed, for discovering CTMSPs in LBS environments. Furthermore, novel prediction strategies are proposed to predict the subsequent user mobile behaviors using the discovered CTMSPs. In CTMSP-Mine, first a transaction clustering algorithm is proposed named CO-Smart-CAST to form user clusters based on the mobile transactions using the proposed LBS-Alignment similarity measurement. Then, the time segmentation algorithm is utilized to generate the most suitable time intervals. To our best 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, the algorithms obtain 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 the proposed methods are efficient and accurate under various conditions.

The application works well for given tasks in windows environment. Any node with .Net framework installed can execute the application and identifies the best site. The underlying mechanism can be extended to any / all kind of web servers and even in multi-platform like Linux, Solaris and more. The system is planned to extend the services can be given as input to IBM architecture also. The system eliminates the difficulties in the existing system. It is developed in a user-friendly manner. The system is

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)

Vol. 2, Special Issue 10, March 2016

very fast in applying algorithm. This software is very particular in predict the subsequent mobile behaviors.

6. REFERENCES

1. Cao.X et al.,(2010) „Retrieving top-k prestige-based relevant spatial web objects" Proc. VLDB Endowment, vol. 3, No. 1, pp. 373–384.
2. Cong.G,(2011) „Reverse spatial and textual k nearest neighbor search," in Proc.
ACM SIGMOD Int. Conf. Manage. Data, pp. 349–360.
3. Hadjieleftheriou.M, (2010) "Approximate string search in spatial databases," in Proc. Int. Conf. Data Eng., pp. 545–556.
4. Mrs. K.Maheswari et al.,(2015) "Traffic Incident Detection Using Weighted Durable Quries" International Journal On Engineering Technology and Sciences, Volume 2 Issue 4,pp55-59.
5. Ooi.B.C,(2011) „Collective spatial keyword querying", in Proc. ACM SIGMOD Int. Conf. Manage.Data, pp. 373–384.
6. Ooi.B.C(2010) „Locating mapped resources in web 2.0," in Proc. Int. Conf. Data Eng., pp. 521–532.
7. Yiu.M.L, (2011) „Efficient continuously moving top-k spatial keyword query processing," in Proc. Int. Conf. Data Eng., pp. 541–552.
8. Zhou.L, (2012) „Seal: Spatio-textual similarity search," Proc. VLDB Endowment,vol. 5, no. 9, pp. 824–835.
9. Zhong.R, (2012) „ Location-aware instant search," in Proc. 21st ACM Int. Conf. Inf. knowl. Manage., 2012, pp. 385–394.