PERFORMANCE ANALYSIS OF TEMPORAL MOBILE SEQUENTIAL PATTERNS IN LOCATION BASED SERVICE

M. I. Thariq Hussan
Research Scholar,
Anna University of Technology, Coimbatore,
Tamilnadu, India

Dr. B. Kalaavathi
Professor & Head,
Department of Computer Science & Engineering,
K.S.R Institute for Engineering & Technology,
Tiruchengode, Tamilnadu, India

ABSTRACT

This paper presents an Associative Relational Mobile Transaction Sequential Pattern Mining (ARM-TSPM) to extract the hidden information in location based service environment. Major problem in the cluster analysis is determining the cluster in unlabeled data. It is overcome by Dark Block Extraction (DBE) method to automatically determine the number of clusters but lacks in processing the full-search. To meet this confronts, an indexing scheme is introduced for the latent and energy efficient processing of full-text searches over the wireless broadcast data stream. The data transmitted to the target state is reached with minimal energy consumption but less information are shared to the neighboring nodes regarding the target state.

Various methods have discussed the mobile sequential pattern and clustering in location based service environment. This paper analyzes the current sequential pattern and provides an overview of the emerging mobile sequential pattern based on clustering. Comparisons are done between the various schemes to explain the advantages and limitations. The experimental evaluation of mining cluster based Mobile Sequential Pattern shows the performance analysis based on tuning time, energy usage, accuracy and execution time.

Keywords: *Data Mining, Sequential Pattern Mining, Location based service, Clustering, Relational Mobile Transaction.*

1. INTRODUCTION

The current acceptance of ubiquitous computing technologies by very large portion of the world population facilitate for the first time in human history to detain large scale spatio-temporal data about human motion. In this circumstance, mobile phones play an important role as sensors of human performance since they are characteristically owned by one individual that carries it at all times and are nearly ubiquitously used. Hence, it is no disclosure that most of the quantitative data about human motion are gathered via Call Detail Records (CDRs) of cell phone networks.

When a cell phone makes/receives a phone call, the information concerning the call is logged in the form of a CDR. This information contain initiating and destination phone numbers, the time and date when the call started and the towers used, which gives an estimation of the caller’s geographical location. Such data is very rich and used freshly for numerous claims such as to study user’s social networks, human mobility behaviors and cellular network improvement.

The progression of wireless communication techniques and the attractiveness of devices such as mobile phones, PDA and GPS enabled cellular phones have donated a new business model. Mobile users can request services during their mobile devices via Information Service and Application Provider (ISAP) from everywhere at any time. This business model is known as Mobile Commerce (MC) that provides Location-Based Services (LBS) through mobile phones. MC is predictable to be as popular as e-commerce in the future and it supports on the cellular network collected of several base stations.

The message coverage of each base station is called as a location area. The standard distance between two base stations is hundreds of meters and the number of base stations are usually more than 10,000 in a city. When users move, their location and service requests are stored in a centralized mobile transaction database.

The time interval segmentation method helps to discover various user behaviors in dissimilar time intervals. For example, users may demand for different services at different times (e.g. day or night) even in the similar location. If the time interval is not in use, some behaviors are missed during exact time intervals. To locate absolute mobile behavior patterns, a time interval table is requisite. Although some studies used a predefined time interval table to mine mobile patterns but the data quality and data allocation vary in real mobile applications. Consequently, it is tricky to predefine a suitable interval table by client. Regular time segmentation methods are necessary to segment the time measurement in a mobile transaction database.

Clustering is a significant and expensive ability in the data mining field. For high dimensional data, current research statement of traditional clustering techniques suffers from the problem of determining meaningful clusters due to the curse of dimensionality. A common approach to cope with the curse of dimensionality problem for mining tasks is to decrease the data dimensionality by using the techniques of feature change and feature assortment. The feature transformation techniques such as principal component analysis (PCA) and singular value decomposition (SVD) sum up the data in a less set of dimensions derived from the grouping of the original data attributes.

Though, the transformed features have no intuitive meaning any more and thus the resultant clusters are hard to understand and examine. On the other hand, the feature selection methods decrease the data dimensionality by annoying to select the majority applicable attributes from the original data attributes. In such way, only a scrupulous subspace is chosen to determine the clusters. On the other hand, in many real data sets, clusters may be entrenched in varying subspaces and thus the feature selection advance the information of data points clustered in a different way.

2. LITERATURE REVIEW

The progression of wireless communication techniques and the reputation of mobile devices such as mobile phones, PDA and GPS enabled cellular phones are giving up a new business model [3] wherein the energy usage and latency measured are major concerns. An indexing scheme introduced for the energy and latency efficient dispensation of full-text searches. Inverted list-style indexing method [5] where overturned lists are placed in front of the data on the wireless channel. In order to decrease the latency overhead planned a two-level indexing method which adds an additional level of index structure to the essential inverted list-style index.

Energy savings in WSNs developed in the information controlled transmission power (ICTP) alteration [6] where nodes with more information use higher broadcast powers than those that are less revealing to share their target state information with the neighboring nodes.

Differentiation mobile behaviors among the users and temporal periods are not measured, so a prediction strategy is proposed [1, 4] to envisage the subsequent mobile users. In CTMSP-Mine, user clusters are created by a novel algorithm named Cluster-Object-based Smart Cluster Affinity Search Technique (CO-Smart-CAST) and comparison between users are evaluated by Location-Based Service Alignment (LBS Alignment).

High-dimensional data space is typically very sparse and significant clusters found in lower dimensional subspaces [11]. ORSC (Arbitrarily ORiented Synchronized Clusters), a novel effectual and competent method to subspace clustering enthused by synchronization [9]. Synchronization is a basic fact common in nature, competent of scheming even highly complex processes such as attitude arrangement in a group. Control of complex processes is realized by simple procedure based on connections between objects.

Clustering develop one-dimensional correlations and is thus a more dense representation of the data set. A new cluster adaptive distance [12] bound based on unraveling hyper plane boundaries of Voronoi clusters to harmonize our cluster based index. This bound facilitate efficient spatial filtering with a comparatively small preprocessing storage overhead. Major problems in cluster analysis are decided the number of clusters in unlabeled data, which is an essential input for the majority clustering algorithms. A new method called Dark Block Extraction (DBE) for robotically approximation the number of clusters in unlabeled data sets [2] is investigated using numerous common image and signal dispensation subspace techniques.

Subspace clustering is an emergent task [8] which aims at distinguish clusters entrenched in subspaces. Novel subspace clustering model to discover [7] the clusters based on the relative region compactness in the subspaces, where the clusters are observe as regions whose densities are comparatively high as compared to the region densities in a subspace. Based on this idea, dissimilar density thresholds are adaptively determined to find out the clusters in dissimilar subspace cardinalities.

A temporal data clustering framework via a weighted clustering ensemble of multiple partitions formed by initial clustering analysis on dissimilar temporal data representations. Proposed a novel weighted consensus function steer by clustering corroboration criteria to reunite initial partitions to candidate accord partitions from diverse perspectives, As a result, the proposed weighted clustering ensemble algorithm supply an effective enabling technique for the shared use of diverse representations.

3. METHODOLOGIES

The different work involved in “Performance analysis of temporal mobile sequential patterns in location based service” is:

3.1 Temporal Mobile Sequential Pattern in Service Environments

Researches on Location-Based Service (LBS) have been budding in recent years due to a wide range of potential applications. One of the vigorous topics is the mining and prediction of mobile movements and linked communication. Most of obtainable studies focus on ascertaining mobile patterns from the whole logs. However, this variety of patterns may not be accurate enough for predictions since the discriminate mobile behaviors among users and temporal periods are not measured.

A novel algorithm namely Cluster-based Temporal Mobile Sequential Pattern Mine (CTMSP-Mine) to find out the Cluster-based Temporal Mobile Sequential Patterns (CTMSPs) is established. Moreover, a prediction strategy is proposed to forecast the subsequent mobile behaviors. In CTMSP-Mine, user clusters are created by a novel algorithm named Cluster-Object-based Smart Cluster Affinity Search Technique (CO-Smart-CAST) and comparison between users are appraise by the proposed measure Location-Based Service Alignment (LBS-Alignment).

Meanwhile, a time segmentation approach is obtainable to find segmenting time intervals where similar mobile individuality survives. To our finest acquaintance, this is the first work on mining and prediction of mobile behaviors with considerations of user relations and temporal property concurrently. Through experimental evaluation under various simulated circumstances, the proposed methods are shown to transport excellent performance.

3.2 Number of Cluster determined in Unlabeled Data Sets

One of the major problems in cluster analysis is the purpose of the number of clusters in unlabeled data which is an essential contribution for most clustering algorithms. A new method called Dark Block Extraction (DBE) for mechanically estimating the number of clusters in unlabeled data sets is investigated, which is based on an obtainable algorithm for Visual Assessment of Cluster Tendency (VAT) of a data set using numerous common image and signal processing techniques. Our DBE method is nearly automatic depending on just one easy-to-set parameter.

3.3 Processing Full-Text Searches in a Wireless Broadcast Stream

In wireless mobile computing environments, distribution is an effective and scalable technique to distribute information to an enormous numeral of clients in which the energy usage and latency are measured major concerns. An indexing scheme for the energy and latency efficient processing of full-text searches over the wireless broadcast data stream is elucidated. Even though a lot of access methods and index structures proposed in the past for full-text searches, all of them are embattled for data in disk storage not in the wireless broadcast channels.

For full-text investigated on a Wireless Broadcast Stream (WBS) introduces a naive, inverted list-style indexing method where overturned lists are located in front of the data on the wireless channel. In order to diminish the latency overhead, a two-level indexing method which adds a different level of index structure to the basic upturned list-style index is proposed. In addition, initiate a replication strategy of the index list and index tree to additionally get better the latency performance. The performance of the proposed indexing scheme with admiration to the latency and energy usage measures analyzed and shows the optimality of index replication.

3.4 Sensor Selection for Collaborative Target Tracking

For target tracking applications, wireless sensor nodes provide precise information since they can be deployed and operated near the occurrence. These sensing strategies have the occasion of association among themselves to improve the target localization and tracking accuracies. An energy-efficient shared target tracking example is developed for WSNs.

A Mutual-Information-based Sensor Selection (MISS) algorithm is accepted for contribution in the fusion process. MISS permits the sensor nodes with the uppermost mutual information about the goal state to broadcast data so that the energy consumption is condensed while the preferred target position estimation correctness is met. In addition, a novel approach to energy savings in WSN is devised in the information-controlled transmission power (ICTP) adjustment, where nodes with more information use superior broadcast powers than those that are less revealing to share their target state information with the neighboring nodes. Simulations make obvious the performance gains offered by MISS and ICTP in terms of power consumption and target localization accuracy.

3.5 Subspace Clustering for High-Dimensional Data

Instead of finding clusters in the full feature space, subspace clustering is a developing task which aims at detecting clusters embedded in subspaces. The majority of previous work in the literature is density-based approaches where a cluster is observed as a high-density region in a subspace. However, the recognition of dense regions in previous work lacks of considering a hazardous problem called the density divergence problem which make use of a compactness threshold to determine the dense section in all subspaces that acquire the serious loss of clustering accurateness in diverse subspace cardinalities.

To tackle the density divergence problem, a novel subspace clustering model to determine the clusters based on the relative region densities in the subspaces where the clusters are scrutinize as regions whose densities are comparatively high as compared to the region densities in a subspace. Based on this idea the different density thresholds are adaptively determined to discover the clusters in dissimilar subspace cardinalities. Due to the infeasibility of pertaining previous techniques in this novel clustering model, an innovative algorithm referred to as DENCOS (DENsity Conscious Subspace clustering) to believe a divide-and-conquer scheme to capably discover clusters rewarding different density thresholds in diverse subspace cardinalities. As authorize by our widespread experiments on different data sets, DENCOS can determine the clusters in all subspaces with high quality and the efficiency of DENCOS outperforms prior works.

4. PERFORMANCE RESULT

In this section, it demonstrates the performance analysis of various locations based services through experiments by examining the mobile sequential pattern factors and clustering based metrics. It is measured in terms of

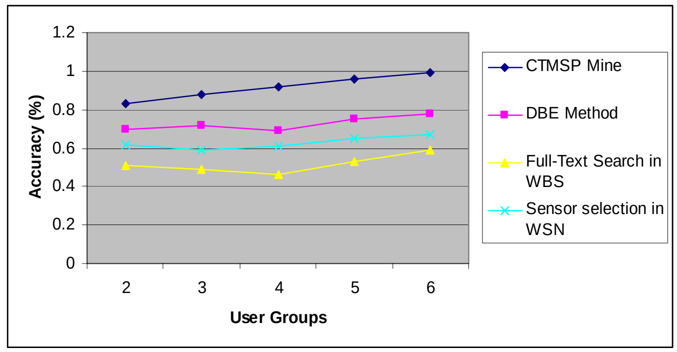
i) Accuracy
ii) Execution time
iii) Energy Usage
iv) Tuning Time

4.1 Accuracy

Fig. 4.1 plots the user group with accuracy using different sequential pattern

User	Accuracy (%)			
	CTMSP Mine	**DBE Method**	**Full-Text Search in WBS**	**Sensor Selection in WSN**
2	0.83	0.70	0.51	0.62
3	0.88	0.72	0.49	0.59
4	0.92	0.69	0.46	0.61
5	0.96	0.75	0.53	0.65
6	0.99	0.78	0.59	0.67

schemes. This result shows that as the group increases and the accuracy also increases



dramatically.

Table 4.1 User Group Vs Accuracy

Fig 4.1 User Group Vs Accuracy

Accuracy is measured in percentage (%). Accuracy is in higher ratio in CTMSP Mine scheme compared with DBE method, Full-Text Search in WBS and Sensor Selection in WSN. In this experiment, CTMSP Mine produces better result than other schemes.

4.2 Execution Time

Fig 4.2 shows execution time of various sequential pattern schemes. Particularly our analysis relies greatly on CTMSP Mining process.

Size of Database	Execution Time (sec)			
	CTMSP Mine	Full-Text Search in WBS	Sensor Selection in WSN	DENCOS Method
50 K	2	8	12	11
75 K	4	14	26	27
100 K	8	20	42	51
125 K	13	32	55	64
150 K	17	45	67	79

Table 4.2 Size of Database Vs Execution Time

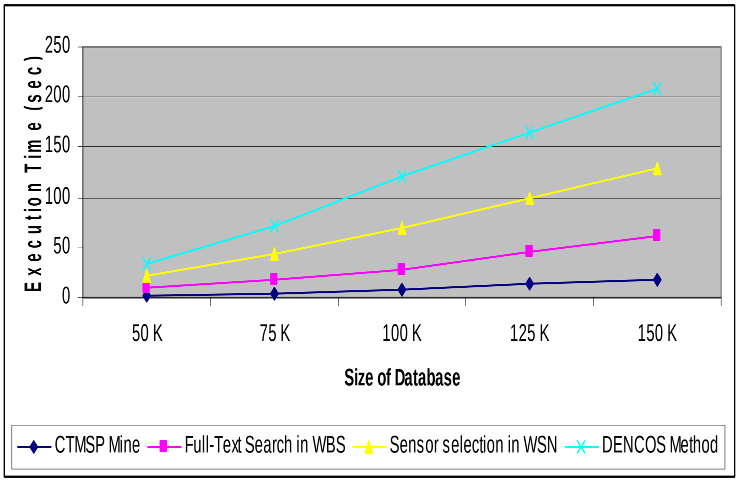


Fig 4.2 Size of Database Vs Execution Time

To check the performance of information retrieval from the database, setup a test which considers the execution time of the CTMSP Mine on database with payloads ranging in size from 50 Kilo bytes to 150 Kilo bytes. From the Fig 4.2 it can be seen that the CTMSP Mine is faster in execution when compared to the other existing system.

4.3 Energy Usage

No. of Queries	Energy Usage (Joule)			
	CTMSP Mine	DENCOS Method	Sensor Selection in WSN	Full-Text Search in WBS
10	0.6	1.6	1.0	1.3
50	0.4	1.4	0.9	0.8
100	0.5	1.0	0.8	0.7
150	0.3	1.2	0.5	1.2
200	0.2	0.9	0.4	1.4

Table 4.3 No. of Queries Vs Energy Usage

The above table (Table 4.3) describes the energy usage of CTMSP Mining method with the various existing system. The energy usage is measured in terms of Joule with query count.

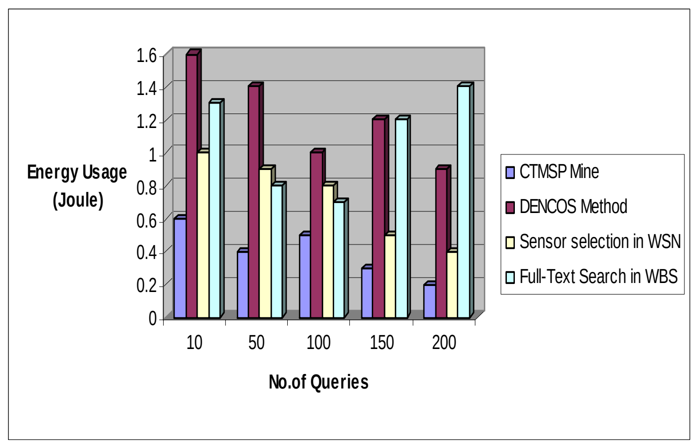


Fig 4.3 No. of Queries Vs Energy Usage

Fig 4.3 describes the energy usage while requesting the queries. In the CTMSP Mine the variance in consuming the energy would be 20-25% less when compared to other sequential pattern schemes.

4.4 Tuning Time

Repeatability	Tuning Time (sec)			
	DBE Method	**Sensor Selection in WSN**	**CTMSP Mine**	**DENCOS Method**
5	1190	1035	1010	915
10	925	845	850	705
15	730	730	700	645
20	650	815	550	555
25	435	625	635	430

Table 4.4 Repeatability Vs Tuning Time

Table 4.4 illustrates the tuning time of four schemes. Comparison result of the CTMSP Mine with other method is measured. When the repetition increases the tuning time also increases drastically.

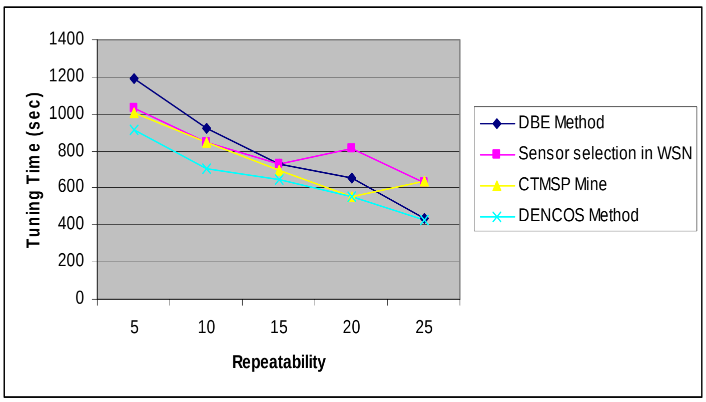


Fig 4.4 Repeatability Vs Tuning Time

Fig 4.4 describes the tuning time. It is measured in terms of seconds. The DBE and DENCOS method testing is 20% less in tuning the query result when compared to the CTMSP Mine in location based service environment and sensor selection in WSN.

5. CONCLUSION

This paper discussed the various methods of temporal mobile pattern based on sequential pattern mining factors and cluster based metrics. Comparisons are made to explain the advantages and limitations of different temporal schemes. Performance analyses of these schemes are evaluated through the experiments. Experimental results demonstrate that some of the schemes support sequential pattern mining metrics and some of the schemes support cluster based mining. Various schemes are examined and their performance is evaluated on four criteria: Execution Time, Energy Usage, Tuning Time and Accuracy. From the experimental results Mining Cluster-based Temporal Mobile Sequential Pattern (CTMSP) in location based service environment performs well in three criteria Accuracy, Execution Time and Energy Usage compared with DBE Method, Sensor Selection in WSN, DENCOS Method and Full-Text Search in WBS.

REFERENCES

[1] Eric Hsueh-Chan Lu., Vincent S. Tseng., Philip S. Yu., “Mining Cluster-Based Temporal Mobile Sequential Patterns in Location-Based Service Environments,” IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 6, JUNE 2011.

[2] Liang Wang., Christopher Leckie., Kotagiri Ramamohanarao., and James Bezdek., “Automatically Determining the Number of Clusters in Unlabeled Data Sets,” IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 21, NO. 3, MARCH 2009.

[3] Ms. V. C. Belokar., Mr. P.S. Kulkarni., “MOBILE SEQUENTIAL PATTERN MINING IN LOCATIONBASED SERVICE ENVIRONMENT,” International Journal of Electronics and Computer Science Engineering.

[4] Shakina.S., J.Rosaline Nirmala., “Mining and Prediction of Mobile User Behavior in Location Based Service Environment,” International Conference on Computing and Control Engineering, 2012.

[5] Yon Dohn Chung., Sanghyun Yoo., and Myoung Ho Kim., "Energy- and Latency-Efficient Processing of Full-Text Searches on a Wireless Broadcast Stream," IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 22, NO. 2, FEBRUARY 2010.

[6] Tolga Onel., Cem Ersoy., Hakan Delic., "Information Content-Based Sensor Selection and Transmission Power Adjustment for Collaborative Target Tracking," IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 8, NO. 8, AUGUST 2009.

[7] Yi-Hong Chu., Jen-Wei Huang., Kun-Ta Chuang., De-Nian Yang., Ming-Syan Chen., "Density Conscious Subspace Clustering for High-Dimensional Data," IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 22, NO. 1, JANUARY 2010.

[8] Vidal, R., "Subspace Clustering," IEEE Signal Processing Magazine, Volume: 28, Issue: 2, 2011.

[9] Junming Shao., Plant, C., Qinli Yang., Bohm, C., "Detection of Arbitrarily Oriented Synchronized Clusters in High-Dimensional Data," IEEE International Conference on Data Mining , 2011.

[10] Yun Yang and Ke Chen, "Temporal Data Clustering via Weighted Clustering Ensemble with Different Representations," IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 2, FEBRUARY 2011.

[11] Elhamifar, E.,Vidal, R., "Sparsity in unions of subspaces for classification and clustering of high-dimensional data," Communication, Control, and Computing, Annual Allerton Conference on 2011.

[12] Sharadh Ramaswamy., Kenneth Rose., "Adaptive Cluster Distance Bounding for High-Dimensional Indexing," IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 6, JUNE 2011.

AUTHORS PROFILE

M. I. Thariq Hussan is working as a Head of the Department for Computer Science and Engineering in CMS College of Engineering, Namakkal, Tamilnadu, India. He received his ME degree from KSR College of Engineering, Tiruchengode. He is now pursuing Ph.D in Anna University of Technology, Coimbatore. His area of interest is Data Mining, Data Warehousing and Wireless Communication. He is doing research in data mining. He has presented papers in 19 international/national conferences and attended 30 seminars/workshops/FDP. He has 2 years of industry and 10 years of teaching experience.

Dr. B. Kalaavathi received B.E (Computer Science and Engineering), M.Tech.(Computer Science and Engineering) and Ph.D (Computer Science and Engineering) in 1993, 2000 and 2010 respectively. Since 1995, she has been the faculty of Computer Science and Engineering Department. She has 17 years of teaching experience at various levels and now she is working as a Professor and Head in the Department of Computer Science and Engineering in K S R Institute for Engineering and Technology, Tiruchengode, India. She is interested in Mobile Computing and Data Mining. She is a life member in ISTE and a member of CSI.