ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

A Review on Mining Frequent Patterns on Temporal Data Using Basic Time Cubes

Shweta S. Holkar[1], Sachin B. Takmare[2]

M.E. Student, Department of CSE, Bharati Vidyapeeth's College of Engineering, Kolhapur, India[1]

Professor, Department of Computer Engineering,PCT's A P Shah Institute OBBC Technology, Thane (W), Mumbai, India[2]

ABSTRACT: Data mining is the process of analyse hidden patterns from the data which is very large in amount with different prospect and organize the useful information into the group. One of the most significant and essential application of data mining is the analysis of transactional data. In the previous research many interesting techniques for temporal data mining were proposed and the techniques were shown to be beneficial in many applications. In this paper we mine frequent itemsets on temporal data with basic time cubes. Our focus is to enhance the efficiency of mining frequent itemsets on temporal data. In this paper we are trying to develop/use meta-heuristic algorithm by which it analyse the whole database and accordingly generate time cubes to reduce the user chosen parameter, and with the help of this time cubes generates the certain valid patterns that occur in the temporal database.

KEYWORDS: Data mining, frequent itemset, temporal data, hypercube, basic time cube (BTCs).

I. INTRODUCTION

Data mining is a process used to discover an interesting pattern or retrieve usable data from a large set of any raw data. Analysis of transactional data is one of the application of the data mining. Transactional data always contains a time dimension, a numerical value and also shows one or more objects. The basic thing in the transactional data mining is the capturing of co-occurrence of items or the patterns [2].Temporal database means that contain time stamping information [7]. For example satellite continuously collect images and sensory data. This information is temporal that containing time stamping information. From this temporal data we need interesting patterns that are often related to the specific period of time therefore, the time during which they can be observed is important. If time intervals are different then different patterns can be found.Another example is call detail records are commonly used by police and intelligent services all over the world. Cell phone act as a basic entity in any of the crime scene. The database of call detail records containing time-stamping information. With the help of finding particular records at particular crime time can play an important role in finding the possible suspects; therefore, the time during which they can be observed is important. In this context many studies have been done in recent years are: finding calendar based periodic patterns[8], mining temporal web interesting patterns [9], mining association rule in temporal database [10],mining frequent ordered patterns without candidate generation [11], and sequential association rule[12].

We alsostudied about the limitation of previous system is that it requires basic time cubes as a parameter which may not be optimal and also may not give better results. In this paper we conduct a study on developing an efficient algorithm to obtain frequent valid patterns on temporal data with optimized BTCs.

II. RELATED WORK

Discovering calendar-based Temporal Association rules:
Yingjiu Li, Peng Ning, X.Sean wang, Sushil Jajodia studied about discovering calendar based temporal association rules. This paper studies temporal association rules during time intervals specified by user given calendar schema. Calendar schema is denoted by (year, month, and day).

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

Concept/ Method: In this paper they define a new representation mechanism for temporal association rules on the basis of calendar and identify two classes of temporal association rules: in precise-match during every interval that require the association rules, in Fuzzy-match during most of these intervals require the association rule. These extend well known Apriori algorithm and develops two optimization techniques to get the benefit of the special properties of the calendar-based patterns.
Findings: They identified two classes of temporal association rules: temporal association rules with respect to precise match and temporal association rules with respect to fuzzy match to represent regular association rules along with their temporal patterns. One of the most important feature of their representation mechanism is that corresponding data mining problem require less prior knowledge than the prior methods and hence may discover more unexpected rules. These rules are very easy to understand. Moreover, they also developed two optimization techniques to improve the performance of data mining process. According to their experiment their optimization techniques are quite effective.
Future work: The future work includes two directions: first one would be to explore other meaningful semantics of temporal association rules and second one would be to consider temporal patterns in data mining problems like clustering [3].

Stock Trading Rule Based on Temporal data mining:
Asadullah Al Galib, Mahbub Alam, Nowshad Hossain, Rashedur M Rahman proposed and implemented the STRDTM(Stock Trading Rule Discovery by temporal mining) algorithm. In this research they had focused upon temporal data mining with time dimension approach. This approach has led them to discover the sequential continuous patterns. The patterns act as rules that enable the user to determine the occurrence of an event on a particular stock transaction day.
Concept/ Method: The stock database containing the events that are frequent. They developed event generation algorithm which maps the frequent events to certain integers or in a bit stream. After that these generated events are then transformed into single sequence. Then frequent sequence generator generates the frequent sequence from the k single sequences using appropriate joins. These frequent sequences is input for the next algorithm that is rule generation where all the frequent sequences are filtered so that all the sequence that fail to satisfy the minimum confidence set by the programmer. Here they use two different kinds of joins *ExpJoin ()* and *ConcatJoin()* in order to generate the next possible frequently occurring sequences.
Findings: The rules generated by the algorithm are accurate to a certain event compared to MACD approach. The rules generated by the algorithm tested on the two dataset with 60% accuracy and 90% accuracy respectively. On an average the rules were found to be 70% accurate. These rules suggest the suitable “sell” situation either “buy” situation.
Future work: They would like to test their algorithm with different time windows rather than a fixed one and also drive the rules that will change sturdily with respect to stock market behaviour [4].
A New Framework of Mining Association Rules with Time-Windows on Real-Time Transaction Database:
YiYong Xiao, Renqian Zhang and Ikou Kaku propose a more general form for association rule is the Association Rule with Time-Windows (ARTW). This paper dedicates efforts to discover the part-time association rules in real time transactional database. They focus on the time attribute of transaction. Part-time association rules occurred just in specific time period, so the difficulty of these task is time-windows may be multiple, aperiodic with arbitrary length and not specified by the user.
Concept/Method: This paper conducts a study on developing an efficient method to mine the part-time association rules and their related time-window from real-time transaction database.
They present some new notation like association rule with time-windows (ARTW) and frequent itemset with time – windows (FITW). Association rule with time-windows (ARTW), to express the association rule occurred in multiple and arbitrary time period and discover frequent itemset with time-windows (FITW), which means the itemsets that are only frequent in their own time windows. And then, they propose new framework for mining ARTMs based on minsup, minconf and minwin. They developed an efficient algorithm TW-Apriori for efficiently generating FITWs. Here TW Means time –window and new generation of FITWs is only produced in their parent’s intersecting time-window.
Finding: In this study mining part-time association rules that occur in multiple, arbitrary and aperiodic time-windows. Based on experiments new framework and the new algorithm performed on two databases-a synthetic benchmark database and a real database. The result is very valuable for market management. The efficiency of the proposed algorithm is also proven feasible; it can be finish the calculation within one minute.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

Future Work: The time-windows of an ARTW produced by TW-Apriori algorithm in some senses are not really arbitrary, other issues on this topic such as dealing with noise, extracting useful information from ARTWs etc. This weakness is to be improved in future work [5].

Mining temporal profiles of mobile Application for usage Prediction:
Liao, Po-Ruey Lei, Tsu-Jou Shen, Shou-Chung Li, Wen-Chih Peng design an AppNow widget that is able to predict users' Apps usage in year. Due to the usage of mobile application increases on smart phones, user can install many Apps to make their life easier. In this paper they design an AppNow Widget that is able to predict the users' Apps usage. Therefore, user could simply execute app from the widget. The main idea of this paper is constructing the temporal profile which identifies the relationship between the apps and their usage time. However, the average number of apps in each device exceed to 150. Users will spend an increasing amount of time looking for and launching the apps they want to use. To deal with this problem they designed AppNow widget, by which dynamically predict the uses' apps usage.
Concept/Method: The AppNow widget is consist of three components, the usage logger, the temporal profile and the apps predictor. First the usage logger records every App start time. Then, the temporal profiles are built by applying Discrete Fourier Transform and exploring usage periods and specific time. Finally, the system calculates the usage probability at current time for each App and shows the list of Apps with highest probability.
Findings: The AppNow widget is developed on Android based smart phone. The system automatically logs the App usage behaviour and updates the temporal profile. When the widget becomes active the predictor displayed the four Apps regarding current time. User can directly execute the Apps by touching the App icon in the AppNow widget.
Limitation: In this they neither move nor re-organize the placement of Apps in users' smart phone or devices, when the AppNow widget cannot provide the correct Apps, the user can still find them in their original place [6].

III. EXISTING SYSTEM

Ghorbani Mazaher, Abessi Masoud [1] developed an algorithm for finding frequent itemset on temporal data by presenting time hierarchy concept in mining process. First they present time cubes (TC), a new notation to consider time hierarchy in the mining process.
Concept/Method: They developed an efficient algorithm by extending the well-known Apriori Algorithm. They first present time cubes (TC) a new notation to considerthe time hierarchy approach in the mining process. First they partition the database into small fragments according to time cubes. Then calculate support and density value. They used two procedure to find frequent itemsets first: find large 1 itemsets, second generate candidate k itemsets.
Findings: They developed an efficient algorithm for mining frequent itemsets along with their temporal patterns. The new feature of their proposed algorithm is notation of Time Cubes (TC).It enable to find different types of temporal patterns. Also introduce new density threshold to solve the problem of overestimating the time period.
Limitation: Limitation of their algorithm is the initial basic time cubes and α parameters should be given by the user. If user chosen time cubes are not optimal then it may affects the result. And also it is necessary to tune α parameter after running the algorithm several times.

IV. PROPOSED SYSTEM

Our contribution of paper is to mine frequent itemset on temporal data with optimize BTCs to provide more satisfactory result. We will try to enhance this proposal by developing an efficient algorithm which is extension of apriori algorithm for mining frequent itemset on temporal data and use meta-heuristic algorithm for analyse the transactional data, so that it creates the basic time cubes accordingly.

The proposed system defines a framework for mining frequent itemsets on temporal data with BTCs. Input will be the transactional database and output will be the frequent itemsets with BTCs.

The working of proposed system is described with following blocks.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

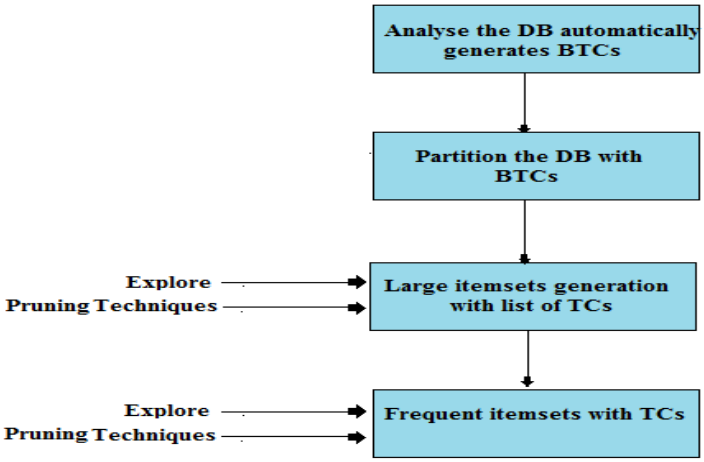

Figure 1: Block diagram of proposed system

1. Analyse the Database and Automatically generate basic time cubes:
   The transactional dataset will be the input here. To reduce the user chosen parameter that is time cube, we want to generate optimized time cubes automatically, so that if time cubes are optimized it gives us batter and valid pattern.

2. Partition the Database with BTCs:
   With the help of automated generated time cubes partitioning of database will be done.

3. Large itemsets Generation with the list of Time cubes:
   It takes the partitioned data. Support and density is calculate for generating large itemsets with time cubes. Support value shows how frequently the items appear in the dataset. Support value is calculate for each itemsets and density value ensure that all the time cubes should be dense.

4. Frequent itemsets Generation with BTCs:
   It takes large itemsets and list of time cubes and also use candidate generation algorithm for mining frequent itemsets with basic time cubes.

In this way we have proposed our framework for mining frequent itemsets on temporal data with automatically generated basic time cubes.

.

V. CONCLUSION

The purpose of this survey paper was to get the deep knowledge about the researches made in recent years on mining frequent itemsets along with their temporal patterns. With this survey we provide general overview on temporal mining. On comprehensive study about the research papers of various publications we came into conclusion that frequent itemset mining on temporal data can be done with the help of new concept Time cubes (TCs).

In this survey paper we studied how we can achieve our goal by using the concept of new notation TC (time cube) which present the time hierarchy in mining process and also how we optimize these BTCs to obtain better result.

REFERENCES

[1] Ghorbani Mazaher, Abessi Masoud, “A new methodology for mining frequent itemset on temporal data”, in IEEE Transactions on Engineering Management, vol. 64, pp. 566-573, Nov 2017.
[2] R. Agrawal, T. Imieli´nski, and A. Swami, “Mining association rules between sets of items in large databases,” ACMSIGMOD Rec.,

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

vol. 22, no.2, pp. 207–216, 1993.
[3] Yingjiu Li, Peng Ning, X.S. Wang, and S.Jajodia, "Discovering calendar-based temporal association rules",Data Knowl. Eng., vol.44, pp. 193-218, 2003.
[4] Asadullah Al Galib, Mahbub Alam, Nowshad Hossain, and Rashedur M Rahman, "Stock Trading Rule Based on Temporal data mining", International Conference on Electrical & Computer Engineering(ICECE), pp. 566-569, 2010.
[5] YiYong Xiao, Renqian Zhang and Ikou Kaku, "A New Framework of Mining Association Rules With Time-Windows On Real-Time Transaction Database", International Journal of Innovative Computing, Information and Control(ICIC), vol. 7, pp. 3239-3253. 2011.
[6] Zhung-Xun Liao, Po-Ruey Lei, Tsu-Jou Shen, Shou-Chung Li, Wen-Chih Peng, "Mining temporal profiles of mobile Application for usage Prediction", IEEE 12th Conference on Data Mining Workshops, 2012.
[7]. T. Mitsa, Temporal Data Mining. Boca Raton, FL, USA: CRC Press, 2010.
[8]. A. K. Mahanta, F. A. Mazarbhuiya, and H. K. Baruah, "Finding calendarbased periodic patterns",Pattern Recognit Lett, vol. 29, no. pp. 1274–1284, 2008.
[9]. Xianwei Hu, Ying yin, Bin Zhang, "Mining Temporal Web Interesting patterns"International Conference on Computational Intelligence and Security, IEEE, pp. 227-231,2007.
[10]. Xinfen Ye and J.A. Keane, "Mining association rules in temporal databases",IEEE International Conference on Systems, Man, and Cybernetics(Cat. No.98CH36218), pp. 2803-2808, 1998.
[11]. Cong-Rui, Zhi-Hong Deng, "Mining Frequent Ordered Patterns without Candidate Generation", Forth International Conference on Fuzzy Systems and Knowledge Discovery, vol.1, 2007.
[12]. R. Agrawal, R. Srikant, "Mining sequential patterns," in Proc. IEEE 11th Int. Conf. Data Eng. Pp. 3-14, 1995.