Research Article [Agrawal, 1(3): May, 2012]
ISSN: 2277-9655

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Temporal Association Rule Mining: A Survey In Fuzzy Framework

Deepak Agrawal [1], Manoj Yadav[2], Hitesh Gupta [3]
[1,2,3]Dept. of Computer Science Engineering, Patel College of science & Technology, Bhopal, India (M.P.)
deepakpcst@yahoo.com

Abstract

Temporal data mining generate temporal association rule that encapsulate transaction of item with time that's recorded in temporal data base. Now these days recent research has focused to generate efficient fuzzy temporal association rule and transforming each quantitative value into fuzzy sets using the given membership functions. This paper presents a survey on temporal association rule and fuzzy logic. The Technical constraint of temporal data mining and fuzzy logic are identified and presented.

Keywords—Temporal association rule , fuzzy set , fuzzy temporal association rule.

Introduction

The growing significance of data stream arise in a wide range of complex application has led to the wide study of mining frequent patterns. Mining frequent sets over data streams present attractive new challenges over traditional mining in static databases. Data mining is a generally used for discovering information and knowledge from huge databases, where 'association rules discovery' is one of the most popular technologies. It was introduced by Agrawal in 1993 [1]. It provides information of the type of "if-then" statements. These rules are generated from the dataset and it draw from measurements of the support and confidence of each rule that can show the frequency of occurrence of a given rule. Association Analysis [1, 2, 4, 6] is the process of discovering hidden pattern or condition that occurs frequently together in a given data. Association Rule mining techniques finds interesting associations and correlations among data set. An association rule [1,3,4,5] is a rule, which entails certain association relationships with objects or items, for example the interrelationship of the data item as whether they occur simultaneously with other data item and how often. These rules are computed from the data and, association rules are calculated with help of probability. Mining frequent item-sets is a fundamental and essential problem in many data mining applications such as the discovery of association rules, strong rules, correlations, multi-dimensional patterns, and many other important discovery tasks. The problem is formulated as follows: Given a large database of set of items transactions and all frequent item sets, where a frequent item set is one that occurs in at least a user-specified percentage of the database. This paper has divided into seven sections where the first one is introduction of the area. The second one is data mining with frequent pattern. Third, one has the temporal data mining. Fourth section explains about fuzzy association rule. Fifth section explains fuzzy sets. Sixth section explains fuzzy sets and finally the conclusion of the paper.

Data Mining With Frequent Pattern

Data Mining evolved from a simple taking out of raw data to an analytical process from large amount of data in order to collect knowledge [7]. It can be done in three stages
1: Exploration - data training
2: Architecture Construction and Verification – have the information about choosing the ones that are best appropriate method to be used
3: Deployment – it will focus on using the selected data to be proceeding with the generation of the output results.
A variety of data mining techniques such as, decision trees, association rules, and neural networks are already proposed and become the point of interest from numerous of years.
Currently temporal-based mining becomes point of interest, there are many research has been done in this area, and we will discussed them in the afterwards section. Along with that, idea of Temporal Association Rules (TAR) which includes time expressions into association rules to handle the time series for solving the problem is introduced in [9]. A standard temporal association rule is said to be frequent within its Maximum Common exhibition Period (MCP) if and only if its support and confidence are greater than the required minimum support threshold and minimum confidence respectively [9][11]. The association rules known as Temporal Association rules (TAR) are applicable during specific time periods [8][9][12]. The inclusion of constraints in the rule mining practical applications where utility mining [13][14][15] is likely to be helpful. The objective of utility

ISSN: 2277-9655

mining is to discover high utility item-sets which are attributable to a considerable portion of the total utility [7]

Temporal Data Mining

Temporal data mining involves application of data mining technique on temporal uses data to discover temporal pattern and that deal with temporal behaviour of user/data. Temporal mining is a data mining that built-in time aspects and deal with the extraction of temporal patterns from large data sets. The utility of temporal data mining continues to grow as increasing amounts of temporal data about everyday activities become available.

Extraction of Temporal Association Patterns for Temporal data mining has been productively applied in number of fields including trading, marketing ,social analysis, medical, fraud detection, robotics and assisted design [16]. Because of that explorer's number of efficient algorithms for temporal data model like symbolic time series, symbolic time sequences, symbolic interval series, numeric time series, item set sequences, etc have been proposed.

Temporal Associate Rules

Temporal data mining generally focus on many of practical disciplines such as statistic, meteorology, telecom, finance, temporal pattern recognition, temporal database, and optimization visualization, high performance computing & parallel computing. Earlier association rule mining pay no attention to time constrain of activity, however, the application areas are always changing with time. Temporal association rule overcome from this problem may recommend a number of interpretations, such as

Preceding event (PE) → succeeding event (SE)

The event (E) → PE and SE

Events → coincidental (c)

Traditional association rules have no concept of order, while time implies an ordering. If we could find the associability of time with event, nothing will be hidden to us as the events are associated to each other in the form PE→PtE (present event)→SE. In this way temporal association rule mining is to discover the valuable relationship among the items in the temporal database.

Fuzzy sets

Fuzzy sets are generalized sets which allow for a graded membership of their elements. Usually the real unit interval [0; 1] is chosen as the member-ship degree structure.

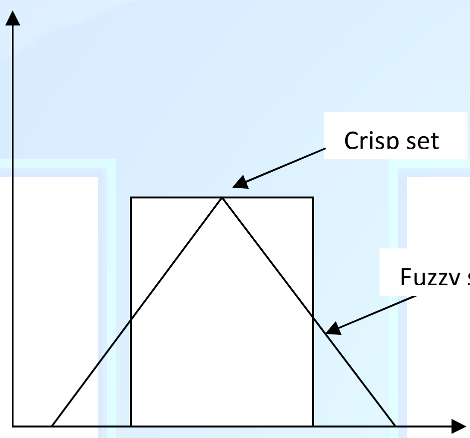

Fig. 1 Boundary of fuzzy set

Fuzzy logic

Fuzzy logic is an extension of multivalve logic. However its uses are quite differing from their objectives. Thus, the fact that fuzzy logic deals with approximate rather than precise modes of reasoning implies that, in general, the chains of reasoning in fuzzy logic are short in length and rigor does not play as important a role as it does in classical logical systems.

Fuzzy Association Rules mining is a novel approach based on classical association rule mining. Whenever we have a data set having a certain range of values then we might face the sharp boundary problem.

Suppose we have three range of age.

F(x) is a function such that

0<x<=30 then f(x) = younger

30<x<=45 then f(x) = middle ager

x>45 then f(x) = older

However in this example, a person aged 44 years would be middle age and a 46 year old would be older where as in reality, the difference between those ages is not that great. Therefore, here is a problem of sharp boundary. As shown in figure 2.

There are some basic approaches to solve the sharp boundary problem.

Quantitative approach

Fuzzy Taxonomic Structures

Approximate Item set Approach

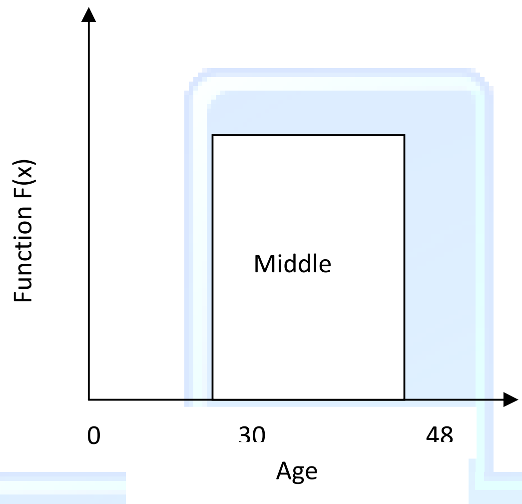

Fig. 2 Example of sharp boundary problem

To resolve the sharp boundary problem by using Quantitative approach divide the variable Age into three fuzzy sets. The fuzzy sets and their membership functions will have to be defined by a domain expert. For easy demonstration, we will just define the borders of the sets and split the overlapping part equally between the so generated fuzzy sets. For an example, we will use the following borders for the fuzzy sets of the variable age: Age.Low={0−33}, Age.Medium={27−48}, Age.High={42−∞}. The generated fuzzy sets is shown in Figure 1 . For all areas having no overlap of the sets,the support will simply be 1 for the actual itemset. If there is an overlap, the membership can be computed by using the borders of the overlapping fuzzy sets. The added support will here always sum up to 1.

Fuzzy set theory has been used more and more habitually in intellectual systems because of its simplicity and similarity to human reasoning [18, 20, 26]. Numerous fuzzy data mining algorithms for inducing rules from given sets of data have been designed and used to good effect with specific domains [17, 19, 21, 22, 25]. As to fuzzy temporal data mining, since fuzzy calendar algebra could help users describe temporal requirements in fuzzy temporal calendars easily, Lee proposed two temporal patterns that were fuzzy temporal association rules and fuzzy periodic association rules based on fuzzy calendar algebra [23]. Based on Lee's approach, Zhuo et al. introduced a relativity-based interest measure value for mining fuzzy calendar-based temporal association rules [24]. However, those fuzzy data mining approaches didn't take item lifespan into consideration. Although Lee proposed two algorithms for discovering fuzzy temporal association rules and fuzzy periodic association rules by using fuzzy calendar algebra [23], lifespan of each item still didn't be considered.

Conclusions

Temporal Association Rules (TAR) includes time expressions into association rules to handle the time series for solving the problem and applicable only during specific time periods. The objective of temporal mining is to discover temporal item-sets which are attributable to a considerable portion of the demand and today intelligent systems used fuzzy set because of its simplicity and similarity to human reasoning. Where data mining with fuzzy set inducing rules from given sets of data have been designed and used to good effect with specific domains. We will present a frame work for fuzzy temporal data mining that generate effective fuzzy temporal association rule.

Acknowledgment

ISSN: 2277-9655

I would like to say thanks to my guide **"Mr. Manoj Yadav"** who gives their knowledge and time in order to complete this paper. This paper will never complete without the support of faculty members of CSE department of **Patel College of Science & Technology.**

References

[1] R. Agrawal, T. Imielinski, and A. Swami, "Mining Association rules between sets of items in large databases ", In proceeding of the 1993 ACM SIGMOD International Conference on management of data, pages 207-216, Washington, DC, May 26-281993.

[2] R Srikant, Qouc Vu and R Agrawal. "Mining Association Rules with Item Constrains". IBM Research Centre, San Jose, CA 95120, USA.

[3] Ashok Savasere, E. Omiecins ki and Shamkant Navathe "An Efficient Algorithm for Mining Association Rules in Large Databases". Proceedings of the 21st VLDB conference Zurich, Swizerland,1995.

[4] R. Agrawaland R, Shrikanth, "Fast Algorithm for Mining Association Rules". Proceedings Of VLDB conference, pp 487 – 449, Santigo, Chile, 1994.

[5] Arun K Pujai "Data Mining techniques". University Press (India) Pvt. Ltd. 2001

[6] J. Han and M. Kamber, "Data Mining: Concepts and Techniques" ,Morgan Kaufman, San Francisco, CA,2001.

[7] Yao H., Hamilton, H. J. and Butz, C. J., "A Foundational Approach to Mining Itemset Utilities from Databases", In Proceedings of SIAM International Conference on Data Mining, pp. 482-486, 2004.

[8] Yingjiu Li, Peng Ning, Xiaoyang Sean Wang, Sushil Jajodia, "Discovering Calendar-Based Temporal Association Rules", In Proceedings of the 8th International Symposium on Temporal Representation and Reasoning, pp. 111–118, Cividale del Friuli, Italy,June 2001.

[9] Tarek F. Gharib , Hamed Nassar , Mohamed Taha , Ajith Abraham, "An efficient algorithm for incremental mining of temporal association rules", Data & Knowledge Engineering, Vol. 69, No. 8, pp. 800-815, 2010

[10] Cheng-Yue Chang, Ming-Syan Chen, Chang-Hung Lee, "Mining General Temporal Association Rules for Items with Different Exhibition Periods", In Proceedings of the 2002 IEEE International Conference on Data Mining (ICDM'02), pp. 59–66, Maebashi City, Japan, December

[11] Jen-Wei Huang, Bi.-Ru. Dai, Ming-Syan Chen, "Twain: Two-End Association Miner with Precise Frequent Exhibition Periods", ACM Transactions on Knowledge Discovery from Data, Vol. 1, No. 2, August 2007.

[12] Chang-Hung Lee, Ming-Syan Chen, Cheng-Ru Lin, "Progressive partition miner: an efficient algorithm for mining general temporal association rules", IEEE Transaction on Knowledge and Data Engineering, Vol.15, No.4, pp. 1004–1017, 2003.

[13] Ying Liu, W. K. Liao, and A. Choudhary, "A Fast High Utility Itemsets Mining Algorithm", In Proceedings of the 1st international workshop on Utility-based data mining, pp. 90 – 99, Chicago, Illinois, 2005

[14] Yao H., Hamilton H. J., Geng L., "A Unified framework for Utility based Measures for Mining Itemsets", In Proceedings of the Second International Workshop on Utility-Based Data Mining, Philadelphia, Pennsylvania, 2006.

[15] S.Shankar, T.Purusothaman, "Utility Sentient Frequent Itemset Mining and Association Rule Mining: A Literature Survey and Comparative Study", International Journal of Soft Computing Applications, No. 4, pp.81-95, 2009.

[16] Mathieu Guillame-Bert, James L. Crowley , "New Approach on Temporal Data Mining for Symbolic Time Sequences Temporal Tree Associate Rules" IEEE International Conference on Tools with Artificial Intelligence , 2011

[17] C. C. Chan and W. H. Au, "Mining fuzzy association rules," The Conference on Information and Knowledge Management, Las Vegas, pp. 209-215, 1997.

[18] P. A. Heng, T. T. Wong, Y. Rong, Y. P. Chui, Y. M. Xie, K. S. Leung and P. C. Leung, "Intelligent inferencing and haptic simulation for Chinese acupuncture learning and training," IEEE Transactions on Information Technology in Biomedicine, Vol. 10, No. 1, pp. 28-41, 2006.

[19] T. P. Hong, C. S. Kuo and S. C. Chi, "Trade-off between time complexity and number of rules for fuzzy mining from quantitative data," International Journal of Uncertainty, Fuzziness and Knowledge-based Systems, Vol. 9, No. 5, pp. 587-604, 2001.

[20] H. Ishibuchi and T. Yamamoto, "Rule weight specification in fuzzy rule-based classification systems," IEEE Transactions on Fuzzy Systems, Vol. 13, No. 4, pp. 428-435, 2005.

[21] C. Kuok, A. Fu and M. Wong, "Mining fuzzy association rules in databases," SIGMOD Record, Vol. 27, No. 1, pp. 41-46, 1998.

[22] Y. C. Lee, T. P. Hong and W. Y. Lin, "Mining fuzzy association rules with multiple minimum supports using

ISSN: 2277-9655

maximum constraints", Lecture Notes in Computer Science, Vol. 3214, pp. 1283-1290, 2004.

[23] W. J. Lee, "Discovery of fuzzy temporal and periodic association rules" National Sun Yat-Sen University: Department of Electrical Engineering. 2008.

[24] C. Zhuo, L. Jiahui, and L. Chen, "A Fuzzy Calendar-based Algorithm for Mining Temporal Association Rules and Its Application", The International Conference on Fuzzy Systems and Knowledge Discovery,pp28-33, 2009.

[25] S. Yue, E. Tsang, D. Yeung and D. Shi, "Mining fuzzy association rules with weighted items," The IEEE International Conference on Systems, Man and Cybernetics, pp. 1906-1911, 2000.

[26] L. A. Zadeh, "Fuzzy sets," Information and Control, Vol. 8, No. 3, 1965, pp. 338-353.

ISSN: 2277-9655