ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 3 , March 2017

Time Line Based ATM Transaction Clustering

Arpita P.R, Jisha Mariam Jose

P.G. Student, Department of Computer Science, New Horizon College Of Engineering, Bangalore. India
Assistant Professor, Department of Computer Science, New Horizon College Of Engineering, Bangalore, India

ABSTRACT Organizations and firms are capturing increasingly more data about their customers, suppliers, competitors, and business environment. Most of this data is multiattribute (multidimensional) and temporal in nature. Data mining and business intelligence techniques are often used to discover patterns in such data; however, mining temporal relationships typically is a complex task. This paper propose a new data analysis and visualization technique for representing trends in multi attribute temporal data using a clustering based approach. This paper introduce Cluster-based Temporal Representation of Event Data, a system that implements the temporal cluster construct, which maps multiattribute temporal data to a two-dimensional directed graph that identifies trends in dominant data types over time. This paper present temporal clustering-based technique, discuss its algorithmic implementation and performance, demonstrate applications of the technique by analyzing data on wireless networking technologies and baseball batting statistics, and introduce a set of metrics for further analysis of discovered trends.

I. INTRODUCTION

Business intelligence applications represent an important opportunity for data mining techniques to help firms gather and analyze information about their performance, customers, competitors, and business environment. Knowledge representation and data visualization tools constitute one form of business intelligence techniques that present information to users in a manner that supports business decision-making processes. In this paper, we develop a new data analysis and visualization technique that presents complex multi attributes temporal data in a cohesive graphical manner by building on well-established data mining methods. Business intelligence tools gain their strength by supporting decision-makers, and our technique helps the users leverage their domain expertise to generate knowledge visualization diagrams from complex data and further customize them.

Organizations and firms are capturing increasingly more data, and this data is often transactional in nature, containing multiple attributes and some measure of time. For example, through their websites, e-commerce firms capture the click stream and purchasing behavior of their customers, and manufacturing companies capture logistics data (e.g.,on the status of orders in production or shipping information).One of the common analysis tasks for firms is to determine whether trends exist in their transactional data. For example, a retailer may wish to know if the types of its regular customers are changing over time, a financial institution may wish to determine if the major types of credit card fraud transactions change over time, and a website administrator may wish to model changes in website visitors' behavior over time. Visualizing and analyzing this type of data can be extremely difficult because it can have numerous attributes (dimensions).Additionally, it is often desired to aggregate over the temporal dimension (e.g., by day, month, quarter, year, etc.) to match corporate reporting standards.

The approach that we take in the paper for addressing these types of issues is to mine the data according to specific time periods and then compare the data mining results across time periods to discover similarities .. Clusters encapsulate similar data points and identify common types of customers. Note that in this example, we used only two dimensions (age and income) for more intuitive visualization. In many real-life applications, the number of dimensions could be much higher, which further emphasizes the need for more advanced trend visualization capabilities. It is a mapping of the multidimensional temporal data into an intuitive analytical construct that we call a temporal cluster . As will be discussed in the paper, these graphs contain important information about the relative proportion of common transaction types within each time period, relationships and similarities between common transaction types across time periods, and trends in common transaction types over time.In summary, the main contribution of this paper is the development of a novel and useful approach for visualization and analysis of multi attribute transactional data based on a new temporal cluster graph construct, as well as the implementation of this approach as the Cluster-based Temporal Representation of event Data system. The rest of the paper is organized as follows: This paper provides an overview

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 3 , March 2017

of related work in the temporal data mining and visualization research streams. This paper introduces the temporal cluster construct and describes the technique for mapping multi attribute temporal data . This paper discusses the algorithmic implementation of the proposed technique as the Dendrogram algorithm and includes Clustering of various transactions

II. IMPLEMENTATION

Module description:

A. Transformation of data's from excel sheet

Needed data is collected from net and stored in the excel sheet. Data is stored into
Dataset through import and export data wizard present in the sqlserver.To import data we have to follow the instruction before start .Go to start menu->All programs->MS sql server 2000->import and export data->DTS Import/export wizard displays

Click next, select data source as Microsoft Excel 97-2000 and database name as
your database name to which you have to import .click next twice then Select Source Table and views display in which you have to type the destination table name.

B. Creation Partition of dataset:

Separating of data present in the dataset according to the Transcation name and storing the values in separate tables. In this module, we present a new data mining technique for identifying and visualizing trends in multi attribute temporal data. We build on both temporal data mining techniques and visual data exploration techniques and develop a tool that provides the user with the ability to interact with temporal cluster data visualization. Temporal cluster use hierarchical and explore temporal data and provide interactive filtering and zooming capabilities for visualization.
.

C. Dendrogram sorting:

Apply denogram extract algorithm to sort the value present in the partition. DENDRO_EXTRACT starts at the root of the dendrogram and traverses the dendrogram by splitting the highest numbered node (where the nodes are numbered according to how close they are to the root, as numbered in Fig. 7) in the current set of clusters until k clusters are included in the set. MaxCl represents the highest element in the current cluster set CurrCl. It is easy to see that because of the specific dendrogram structure, it is always the case that MaxCl ¼ DendrogramRooti _ jCurrClj Furthermore, the dendrogram data array maintains the successive levels of the hierarchical solution in order; therefore, replacing MaxCl by its children MaxCl:Left (leftchild) and MaxCl:Right (right child) is sufficient for identifying the next solution level in the dendrogram. DENDRO_EXTRACT is linear in time complexity OðkiÞ, which provides for the real-time extraction of cluster solutions.

Implementation is the most crucial stage in achieving a successful system and giving the user's confidence that the new system is workable and effective. Implementation of a modified application to replace an existing one. This type of conversation is relatively easy to handle, provide there are no major changes in the system. Each program is tested individually at the time of development using the data and has verified that this program linked together in the way specified in the programs specification, the computer system and its environment is tested to the satisfaction of the user. The system that has been developed is accepted and proved to be satisfactory for the user. And so the system is going to be implemented very soon. A simple operating procedure is included so that the user can understand the different functions clearly and quickly.

Initially as a first step the executable form of the application is to be created and loaded in the common server machine which is accessible to all the user and the server is to be connected to a network. The final stage is to document the entire system which provides components and the operating procedures of the system.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 3 , March 2017

In module given input and expected output:

Partition of dataset:

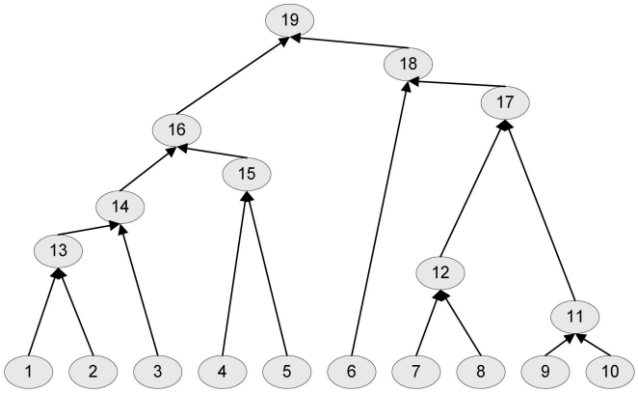
Given input : Result set (collection of unrelated data)

Expected output: Clustered data (separating and grouping relevant data)

Dendrogram sorting:

Given input : clustered column value

Expected output: Sorted sequence fort the given column



Using above algorithm we sort all the value present in each column using above algorithm. here we calculate time for each table using current time of the system .thus calculated time is displayed in separate table.

III. ADVANTAGES

- In our project we use DENDROGRAM Data structure for storing and Extracting cluster solutions generated by hierarchical clustering algorithms
- Calculations are made using Tree structure
- Efficiency is considerably increased.
- N is user defined

IV. CONCLUSION

By harnessing computational techniques of data mining, we have developed a new temporal clustering technique for discovering, analyzing, and visualizing trends in multi attribute temporal data. The proposed technique is versatile, and the implementation of the technique as our system gives significant data representation power to the user domain experts have the ability to adjust parameters and clustering mechanisms to fine-tune trend clusters .Efficiency is increased because of the usage of Dendrogram Algorithm for Clustering. Transactions are clustered using the tree structure which increases the efficiency of the project.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 3 , March 2017

REFERENCES

[1] Cay S. Hortsman, Gary Cornell , Core Java™ 2 Volume I – Fundamentals 7th Edition Pearson Education – Sun Microsystems,. vol. 98, no. 11, pp. 1947–1960, 2010
[2] Cay S. Hortsman, Gary Cornell, Core Java™ 2 Volume II – Advanced Pearson Education – Sun Microsystems,. , pp. 326–33,2011
[3] SPD, Eric Freeman, Elisabeth Freeman Head First Servlets & JSP O'Reilly –, vol. 52, no. 12, pp. 5406–5425, 2006
[4] HertSchildt, Herbert Schildt JAVA 2: The Complete Reference SPD,. vol. 52, no. 2, pp. 489–509, 2006
[5] Joshua Bloch. , Effective Java – Programming Language Guide Pearson Education – Sun Microsystems, pp. 465–48,2010.