International Journal of Computer Techniques -– Volume 3 Issue 2, Mar- Apr 2016

RESEARCH ARTICLE OPEN ACCESS

Visualization of Association Relationand its Application in Tax Field

Liang Wu[1],MingXin He[2]

[1,2](College of Information Science and Technology, Jinan University, Guang dong, China)

----------**********----------

Abstract:

All things in the world are connected by nature or human activities; All kinds of information systems have recorded data related to the production and life, and stored these data in the database. People are expected to understand the relationship between things that are implicit in the data, including the overall situation and the specific details of a particular relationship. Finding out the potential relationship from the massive data, using the visual technology to present the related network, revealing the correlation hidden in data, have important practical value and significance for commercial and government applications. By means of visualization, the implicit association relation is presented in this paper to help people better understand and explore relevant information. This paper describes the research background, application status and related concepts and technologies of the visualization of association relations; and explores the total solution and implementation technology of visual presentation of the relationship; proposes a successive and interactive "drill down" approach taking network map as the information entrance to explore the general characteristics, the specific relationship and related details of the real world. Finally, get a thorough understanding of the specific issues and subtle insight.

The project in this paper has been applied in the tax department. Show the relationship between taxpayers and tax doubtful information through the analysis of related graphics, comparison, "information drilling" and other means to help the business personnel find problems from the relationship between the interests and risk assessment. It has played an important role in supporting the decision making function for tax assessment, tax inspection and tax management, and had achieved satisfactory results.

Keywords **—association relation; network visualization; information drilling ; tax application**

----------**********----------

I. INTRODUCTION

This paper is based on a pilot project of the tax revenue data integrated application in a Local Taxation Bureau. With the development of social economy, both tax payment units and its trading volume have increased sharply. Tax evasion, tax avoidance methods emerge in an endless stream. The traditional use of corporate financial statements to find loopholes in corporate tax revenue is not only inefficient, but also has little effect. In view of this situation, our innovative use of visualization technology makes full use of the tax system based on the accumulation of a great deal of data records to show a transaction relationship between each enterprises and other units or individuals in a graphical way, which provides effective tools for tax assessment. Different from the traditional visual application, this system not only regard the visual technology as a simple expression of information, but also regard it as a search, analysis, thinking tool; a tool for helping others to analyze complex issues. We will gain insight into the overall relationship by presenting the individual, local graph. We show the

relationship of tax data to help people get real, reliable sources of tax units.

The following content of this article is arranged as follows: The second chapter is the introduction of visualization technology and visualization about correlation. In the third chapter, the paper introduces the framework and the implementation process of the application of the visualization of the association relationship combining with the tax practice project. The fourth chapter shows visualization techniques applied in the results of the tax project. The fifth chapter, as the conclusion, analyzes the essence of the visualization of the relationship combining with visualization technology.

II. VISUALIZATION TECHNIQUES AND VISUALIZATION OF ASSOCIATION RELATION

Visualization technology is a technology that allows people to accept and understand the original data and information, which is based on computer graphics, by generating a visual image[1]. Its appearance has a profound background. In the early twentieth Century, people used the traditional visualization techniques such as charts and statistics to analyze the data. With the development of visualization technology, the visualization technology is divided into four categories according to the object and the purpose: visualization in scientific computing, data visualization, information visualization and knowledge visualization[2].The above four kinds of visualization technology are divided according to the processing object and purpose. The object of scientific visualization and data visualization is the data, the object of information visualization is information, and the object of knowledge visualization is knowledge. They adopt different methods, but they are interrelated. From data to knowledge is a process of more and more abstract. Data is the carrier of information, information is the content of the data, knowledge is the conclusion of information classification, they complement each other, and have a common development. In practical applications, we often need to use a variety of techniques to achieve the desired results.

Visualization of association relation is displaying the relationship between entities hidden in the data in the form of map using visualization technology [3-4] . In the visualization of association relationships, people not only need to observe the overall characteristics of the map, but also need to pay attention to the details of the map, such as the meaning of the node represents, the relationship between the types of edges, the relationship between the node and the node, and so on. Visualization of related relationships is often used to help people to analyze the relationship between things and to think about the benefits of these relationships.

III. VISUALIZATION OF THE RELATIONSHIP BETWEEN THE TAX SYSTEMS

With the continuous refinement of tax source management, local tax bureau had made great achievementsfor tax assessment and tax inspection and other aspects. But the tax management system still has some deficiencies, for example, the system function is not perfect, the lack of application in the tax assessment, the fine management of the tax revenue and so on; data acquisition source is wide, which can not form a unified data query channels; a large number of tax data has not been tested and used, and has not played its role in tax analysis.This paper has carried out the research on the visualization of the association relations, the purpose of which is the implicit relationship of the tax data, using relational mining and D3 visual language to show the relationship of tax system [5] In the following, the paper introduces the specific process of visualization of the relationship between the tax system. Figure 3-1: Flow chart of implicit relationship dominance. Figure 3-2: System logic structure diagram [6] .

A. Data extraction.

The source data of this system is extensive, including tax source management platform if tax bureau, data of warehouse in up layer units and some related database. The first step is to extract these data, and classify these data in accordance with the commonly used tax relations, and then store the data in the database as a source. From a technical point of view, we need to extract large and complex data specified by users from the database or data warehouse, and convert these data into the data we need in interface. There are two ways to extract data from the database, one is the total amount of extraction, one is incremental extraction. Total extraction refers to the extraction of wholly intact tables or views in the data source, and convert them into the format we need to deal with. Incremental extraction is to extract new data, or data needing to be modified or deleted. We use it more. Our data extraction of the specific approach is to determine the data source in the source database and the data tables and fields to be extracted. Then create a storage table and a field as the data receiving object in the target database. In the end, we create an intermediate process to implement the data stream from the source data to the target database. Pay attention to two points when creating the middle course:

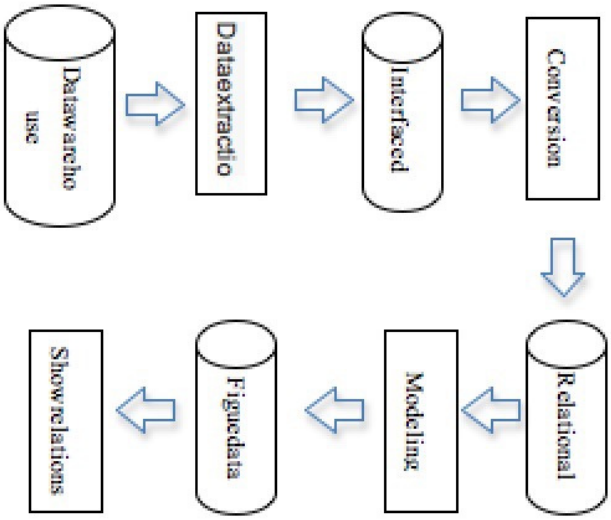

Figure 3-1 Implicit relationship explicit flow chart

1) To maintain the practical storage process of the target database and connect the data source with the transparent network management system.

2) Before data extraction, check that whether the dependent data is ready, to prevent errors due to data dependency. Therefore, we need to determine the operation of the storage process according to the scheduling period and the dependence of the extraction process.

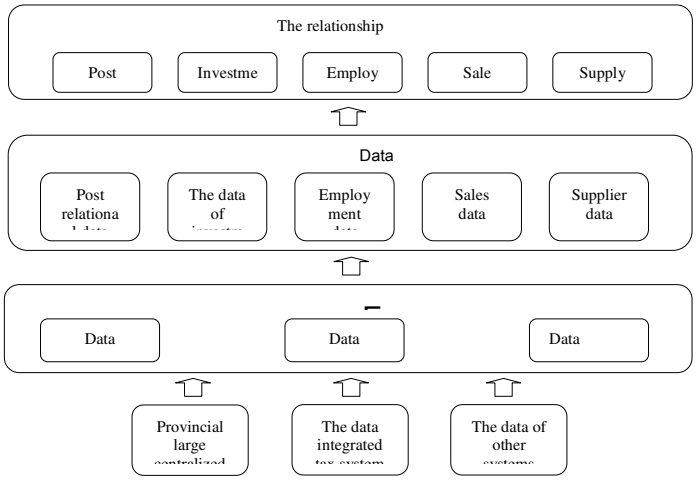

Fig. 3-2 Visualization data processing logic diagram

B. Cleaning conversion.

Due to the wide data source of the tax system, interface data may contain incomplete, erroneous, or duplicate data that are not in compliance with the requirements of the tax. We need to complete, modify or remove the data according to suggestion of tax related personnel.First of all, we need to define the standard relational data structure and set the filter rules,Then write the stored procedure, and find out the data items that do not meet the rules in the interface data,Finally, store the data to the database, delete the data needing to be deleted

according to the recommendations of the business department.

C. Data modeling.

The data can be directly used for modeling, and it contains all the information that we may need to use in the tax analysis. However, users do not need to use all the characteristics of the relationship, they are only concerned about one or several characteristics. In the tax system, we need to observe the relationship between the enterprises. However, there are a lot of different types of trading relationships, such as investment relationships, employment relationships, relationships and supply relationships, and so on. From the tax point of view, the significance of different types of relationship is different. We want to further simplify the data according to user defined correlation characteristics, in order to form the data model we need. This is not only in line with the needs of users, but also can simplify the data, so that the focus of the presentation is more prominent.

D. Using visualization technology, data model is converted to graphics.

In many visualization technology, the system uses the d3.js visual language to achieve graphics display. There are two reasons:

1) d3 is a visual language based on the browser, it is more convenient for users to use

2) d3 is a data driven language, according to the data changes, it can complete real-time graphics changes, which greatly improves the flexibility of the system of human-computer interaction.

The purpose of visualization is to display the information in a graphical way. However, it is more important to facilitate people to explore the useful information contained in the data.[7-8].So the following points should be paid attention to when converting graphics:

1) The figure is simple and clear. In this paper, the force directed placement algorithm is used to generate the graph. Force directed placement is a kind of layout algorithm based on cost function, which is proposed by Eads in 1984, and it is gradually improved in the future[9].Itmodeled a graphic to be tension and repulsion of a physical system, so as to avoid overlapping nodes overlap.

2) Pay attention to human computer interaction. In the analysis of the relationship between enterprises, we are concerned about not only the overall characteristics of the network map itself, we also need to clearly understand the specific information represented by each node in the graph. But showing all the details in a picture will lead to confusion of information. This is not conducive to the observation and contrary to the user's aesthetic. We take the way of human-computer interaction to show the basic information directly on the node, hide important information on the menu, to ensure that users can not only grasp the overall characteristics of the map, but also can easily understand the details of each node.

3)Show the relationship layer by layer, and the user can edit, delete, and implement other operations. Layered display can use a limited screen to show attractive information, at the same time, it allows users to delve into one relationship, and find the problem of relationship in a clear and orderly way. After editing and expanding nodes, users will get a unique "mind map". We need to store the graph in order to reproduce the thinking process of the corresponding conclusion, at the same time, it is convenient for other users to have a secondary analysis based on the results.The figure is simpleThe figure is simpleThe figure is simple.

IV. SYSTEM FUNCTION IMPLEMENTATION AND RESULTS DISCUSSION

This system mainly includes the search engine and the display and processing of related relations. The search engine is the portal of the whole system. Relational presentation is the

graph of relationship, taking a unit as the center, using its corresponding unit as a node. The purpose of graphics processing is to facilitate the analysis of the relationship, it mainly includes graphics scaling, graphics filtering, add notes and other functions.

E. *Relational graph portal—Search Engines*

When tax unit or name and other information is input by the user in the search engine, the system will find out the input keyword association graph, financial statements, tax records, and so on from various data sources, and display the search results below the search box, and provide hyperlinks to the search results corresponding detailed information. It sets the information of multiple data sources, and solves the problem of the information island of the tax system.Figure 4-1 shows the system search engine function display.

Figure 4-1 Tax management system search engine function display

F. *Graphical presentation and processing of relational graphs*

The system generates the corresponding data model according to the user's search condition, and the graph of the relationship is displayed in a complex network. In the graph, we can express the information of relevant units or relations by the shape, size, color and the length, thickness, color and other information of the node. How to use the limited space in the map to show more effective information, is a constant improving exploratory practice.

G. *Graphic detail information presentation*

In order to make the graph more concise, the graphics unit tax of all information and sources units between the transaction details are not displayed directly. When the user moves the mouse to the connection, The system will automatically display details of the transaction between the two units, as shown in Figure 4-2. When the mouse is moved to the node, the menu will pop up, you can see the details of the node related information according to the menu tips.

H. *Questionable information display*

The system can find out some doubtful points according to the predetermined rules. And the corresponding node name is set to the red to alert user. The user can view the details of the questionable information via the node menu, as shown in Figure 4-3.

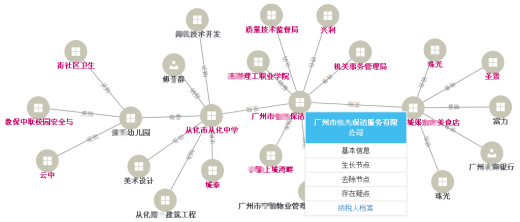

Figure 4-2 The relationship between tax unit

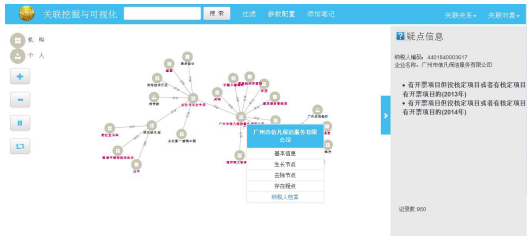

Figure 4-3 Node doubt information

I. *Graphics zoom*

Graphics in this project is divided into two forms, one is not to change the size of the node itself, only stretch it in accordance with the ratio, or reduce the length of the node to facilitate the user to highlight the relationship between the nodes and get inspired. The other is that the size and length of the node and the connection are enlarged or reduced according to the corresponding proportion to prevent some information is not clear, because the graphics are too cumbersome, or some nodes can not be completely displayed in the display, because the graphics are too large.

J. Hide and display of some objects

After finding out the relationship of a unit, the user is often only concerned about one or a few records and uses these as clues to deeply drill relationship between enterprises, and find the tax doubt of tax unit. At this time, remove other nodes. The system provides two ways for removing nodes. One is to select a node that needs to be displayed or hidden in accordance with the type of node or the type of relationship or click the "filter" button to customize the type of node that needs to be displayed. The other is the user takes the initiative to use the right button to delete a node in a menu or, in the graph, add a relation graph taking the node as the center. Figure 4-4 is the sample graph of hidden nodes in accordance with the relationship.

Figure 4-4 Sample graph of hidden nodes in accordance with the relationship.

K. Add notes

Graphic results are equivalent to a pair of mind map. It contains the thinking process of the conclusion of the target. It can be used as evidence and facilitates the use of existing results for other users to carry out a secondary study. System provides notes saving function. Click on the Save button and add the name and description of the note, and save the snapshot notes of the current graph and current central node is the indexes, as shown in Figure 4-5.

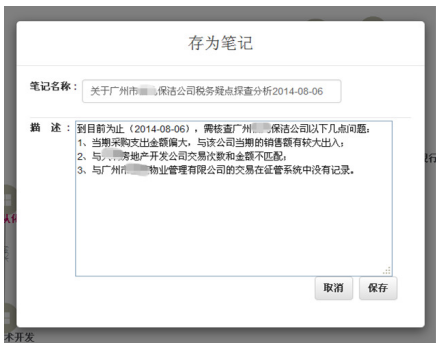

Figure 4-5 Add notes

V. CONCLUSIONS

Visualization transforms the data into a graph, which gives us a greater insight into the contents of the data. [10] However, with the growth of the volume of business, the amount of data is more and more huge, we just focus on the data itself is far from enough, the hidden relationship between the data can often provide us with more abundant and effective information. The explicit relation of implicit relation refers to transform the relationship between the storage data and the data into a specific relationship model, then use the visualization technology to display it. The essence of the implicit relationship is to show the relationship between the real world and the individual in the database in a graphical way, so that users can get the specific details from the local to the specific object. For example, tax uses investment, transaction, tax, office and other relations between the data and other statistics to obtain statistical information on tax statistics. These doubts are convenient for users to carefully study, investigate and compare these analysis to explore the authenticity of transactions tax object, so as to promote the tax assessment.

REFERENCES

[1] Chen Jianjun, Yu Zhiqiang, Zhu Yun. Data visualization technology and its application[J]. Infrared and Laser Engineering, 2001, 30(5): 339-342

[2] Yu Weihe. Research and implementation of visualization technology of association relation in personal data space[D]. Master thesis. Liaoning: Northeastern University, 2010

[3] Liu Yuqin, CengJianxun, Wang Lixue. Design and implementation of the visual system of academic association[J]. Library and Information Service, 2014, 58(5): 75-81

[4] Zhou Jie, Liu Yuqin, CengJianxun. The visual method of the relationship between academic research subjects and research contents[J]. New technology of library and information service, 2012, 226(11): 92-97

[5] WooYun knowledge base. Visual analysis of large data data using d3.js [EB/OL](2013-12-19). http://drops.wooyun.org/tips/823

[6] Huang Zhiqiang, Hu Zhaoming. Design of user centered data visualization scheme[J]. Information exploration, 2015, 207(1): 66-70

[7] Ren Lei, Du Yi, Ma Shuai, Zhang Xiaolong, Dai Guozhong. Overview of big data visibility analysis[J]. Journal of software, 2014, 25(9): 1909-1936

[8] Iribarren-Maestro I, Lascurain-Sanchez M L, Sanz-Casado E. Are Multi-authorship and Visibility Related? Study of Ten Research Areas at Carlos III University of Madrid [J]. Scientometrics, 2009, 79(1): 191-200

[9] Mike Bostock. D3.js[EB/OL] (2013). http://d3js.org/

[10] Heer J, Bostock M. Declarative language design for interactive visualization[J]. Visualization and Computer Graphics, IEEE Transactions on, 2010, 16(6): 1149-1156