Research and Application of Data Warehouse on Enterprise Information Management

Bei Wang

Shaanxi Polytechnic Institute, Xianyang, Shaanxi, 712000

Abstract: Today, the rapid development of information technology, enterprises will be brought into the online analysis and processing, data warehouse and data mining information analysis era. Data warehouses have evolved from a simple theoretical research into a highly practical technology in the field of information management and information systems development in just a few years. Most enterprises in the early process of information technology to build a relatively complete information processing system - one on-line transaction processing system, and accumulated a lot of valuable business information. However, due to the limitations of information management system, most of the online transaction processing system can not provide better support for business management decision-making. To this end, many business executives began to set up their own data warehouse for management decision-making, in response to increasingly severe market competition pressure.

Keywords: Data Warehouse, Enterprise, Information Management

1. INTRODUCTION

With the deepening of information research, people's understanding of information is gradually deepening, especially in the study of information science - information theory, the concept of information almost into all disciplines, so the concept of information will be increasingly Attracting the attention of many scientists, the American mathematician Wiener believes that information is the name we use to adapt to the outside world and communicate with the outside world in making this adaptation to the outside world. In general, the information is Information is the sum of all knowledge and reflection, it is universal, objectivity, infinity, dynamic, dependency, metrological, transitive, etc., which is the data that can be affected by the specific activities of people. Characteristics, the universality, objectivity and infinity are the whole characteristics of information collection. Dependency, metrology and transitivity are the characteristics of the elements in the information set. The dynamic is not only the overall characteristics of the information set, The characteristics of the elements in the collection.

2 ENTERPRISE INFORMATION MANAGEMENT SYSTEM

With the rapid development of computer technology, the rapid increase in users and constantly put forward new requirements, people are increasingly satisfied with the database technology to a single data resources as the center, a simple transaction processing, eager to the existing many Data in the database to filter, extract, re-processing, the formation of a comprehensive, the theme of the analysis of the environment to better support the analysis of decision-making, so the data warehouse technology came into being. It is from the development of database technology, only 10 years, data warehouse in the field of foreign information computing, from quietly rising to perfect, and has been widely used.

The concept of Data Warehouse originated in the mid-1980s, the father of the data warehouse, WHInmon in its "Building the Date Warehouse" book defines the concept of data warehouse "data warehouse is a thematic, Integrated, nonvolatile, time-varying set of data used to support management decisions.

Data warehouse technology does not have rigorous mathematical theory and mature model, but more biased towards engineering, according to its work process to analyze, mainly reflected in three aspects: basic principles, online analytical processing (OLAP), data mining (DM) , Online analytical processing and data mining is the data warehouse data analysis means. Data warehouse and database compared to the following main features:

The theme is an abstract concept that abstracts and classifies the data in the enterprise information system at a higher level. In the logical sense, it is a corresponding macro analysis of the enterprise involved in the analysis of the object, is set for a decision problem. The subject-oriented data organization is a complete, unified, consistent description of the data of the analysis object at a higher level. It can complete and unify the data of the enterprises involved in each analysis object, and the data the relationship between. The division of the subject must ensure the independence of each subject. That is to say that each subject should have an independent connotation, a clear boundary. When dividing a theme, you need to ensure that the data you need to analyze the topic can be found in this topic. As the theme is at a higher level of data abstraction, which makes the subject-oriented data

organization can be independent of the data processing logic, easy in this data environment for management decision-making analysis and processing.

Integrated, transaction-oriented operational databases are often associated with certain applications, and databases are independent and often heterogeneous. The data corresponding to each subject field in the data warehouse also has many duplication and inconsistencies in the individual databases. When the data warehouse obtains data from the transaction processing system, it can not load the original data directly into the data warehouse. Therefore, before entering the data warehouse, the data must be extracted, cleaned up on the basis of the system processing, aggregation and collation, to eliminate inconsistencies in the source data. Such as unified data in the field of the same name, data types are inconsistent, inconsistent units and so on.

Access to and processing of operational data is generally done on a single, one-way basis. The data in the operating environment is generally about to be updated, and the data in the data warehouse is usually loaded and accessed in bulk, and is not updated in the general sense. The data warehouse stores historical data for a fairly long period of time, a collection of snapshot snapshots at different points, and derived data based on these snapshots for statistics, synthesis, and reorganization, not data that is processed online. Therefore, once the data into the data warehouse is rarely updated, is relatively stable.

Data warehouse data is relatively stable for the application, that is, the user does not analyze the data processing operations, but not that the data from the data warehouse will never change. The data in the data warehouse is periodically updated over time, and the data generated in the operational database system is extracted at regular intervals, converted into the data warehouse after conversion, and the past version of the data is retained in the data In the warehouse. With the change of time, the data is integrated at a higher level of integration to meet the requirements of trend analysis. These data are deleted from the data warehouse when the data exceeds the storage period of the data warehouse, or when the analysis is no longer useful.

3 ENTERPRISE INFORMATION MANAGEMENT SYSTEM ARCHITECTURE

Enterprise-class data warehouse architecture is generally a three-tier architecture: the bottom is the data storage management, storage and management of data from the source business database; the middle layer is the analysis model server, used to build OLAP and DM data analysis model; The upper layer is a front-end tool and application layer, to provide users with the final data analysis tools. Based on the three-tier architecture, the system user interface is divided into four parts: data acquisition layer, data storage processing layer, system service layer and system interface layer. The system information management system architecture is based on the architecture of the enterprise information management system.

The bottom of the enterprise information management system is the data acquisition layer, which mainly collects the relevant data of each department of the enterprise. The departments will keep the new data in the local database every day, and then transfer the data to the data warehouse through the network connection. Processing layer of the data warehouse to extract the information to extract, clean, convert, gather and other processing to establish a global data warehouse, through the analysis of the necessary information to generate data mart; system business layer using data warehouse analysis mining results , The corresponding business analysis and decision management; system user interface layer is presented to the user's friendly management interface, different users can be authorized to facilitate the operation of business operations.

The data warehouse DW can be represented as a 6-tuple, ie DW = <DS, DD, SD, AD, MD, PS>. Where DS is the data source of the data warehouse and is a collection of database tables OST and EDS (such as electronic tables, text, etc.) in the operating system; DD is the detailed data set in the data warehouse, the detail data Is the core of the data warehouse, and its data comes directly from DS, which is the same enterprise-wide pattern of data sets; SD = {data set DM collection ∪ data collection of cubic DC}, where DM is integrated by DD for enterprise departments AD is the archived data (rarely accessed data); MD is the metadata of the data warehouse, which defines the data warehouse (the data warehouse), the data warehouse, the data warehouse, the data warehouse, the data warehouse, the data warehouse, the data warehouse, the data warehouse, the data warehouse, the data warehouse, the data warehouse, the data warehouse PS is the set of data used to generate and manage data at all levels in the data warehouse, responsible for data extraction, conversion and loading procedures, from the details of the data, the data warehouse, the data warehouse, and the data warehouse. Generate data mart and data cube assemblies .

According to the requirements of the information management system and the design of the architecture, the data warehouse is required to pretreat the consistency of the original data. The data of the preprocessing is excavated. The core part of the data analysis adopts the rough set and genetic algo-rithm, RSGA algorithm); the analysis of mining data after the friendly presentation. Data warehouse workflow shown in Figure 2, which mainly by the pretreatment, RSGA algorithm and post-processing 3 parts. The preprocessing part transforms the initial

data information into the decision table form, clarifies the condition attribute and decision attribute, and transforms the attribute language description to the real field. If the condition attribute contains the continuous value attribute, it needs to be discretized. Postprocessing is to check the generated decision information to remove redundant information, resulting in general and concise decision information. The composition of RSGA algorithm is mainly composed of rough set analysis and decision information mining based on genetic algorithm. The main function of rough set analysis in data mining is to realize data reduction, rough set decision information acquisition, and genetic algorithm based on data mining to optimize the rough decision information. The advantage of RSGA algorithm is that it is effective to deal with inconsistent data, generate deterministic decision information, satisfy decision information of given confidence; the second is to cover all examples to cover each decision class. Third, the use of genetic algorithms, making the excavation of the decision-making information is simple and effective.

4 APPLICATION OF DATA WAREHOUSE IN ENTERPRISE INFORMATION MANAGEMENT

Data warehouse can strengthen the enterprise's ability to manage information, change the management of enterprises, so that the process of enterprise decision-making more scientific and rapid, to bring huge business benefits, the establishment of a wide range of competitive advantage. Data warehouse technology in enterprise information management applications mainly in the following aspects. Data warehouse can strengthen the management of information management capabilities, change the business management, so that the process of enterprise decision-making more scientific and rapid, to the enterprise to bring huge benefits, the establishment of lasting competitive advantage. Specific performance in the following areas. Improve enterprise information management capabilities. With the industry competition and user needs tend to diversify # personalized, companies must change the management concept, according to market demand for product production and sales, which requires a lot of data and information analysis, these data are highly integrated , And some are details, some years ago, or even a few decades ago, and the traditional database generally only retain the current details of the data, so the traditional database can not do this task, you must use the data warehouse to store the data. Help enterprises to establish a good customer relationship. Modern business competition is more and more intense, customer groups are growing, service requirements are getting higher and higher, so customer relationship management in the enterprise management becomes more and more important, but only by hand is difficult to complete. Through the data warehouse in the customer & product type & regional and sales channels and other data for comprehensive analysis, the market can be broken down for different customers to provide different services to help enterprises better for existing customer service, Make these customers more satisfied. Improve the level of enterprise decision-making. The data warehouse system is oriented to decision analysis, has strong data synthesis capabilities and complex data analysis capabilities, can provide timely and accurate data and information to help decision makers quickly make the right decisions, so that they make the market development Accurate evaluation, so as to improve the competitiveness of enterprises. To encourage enterprises to reorganize the business process. The internal structure of the enterprise is loose, and each department is one of the main problems of the long-term troubled enterprises. It seriously affects the development of the enterprise itself. Construction of data warehouse is not only to build an information system, but also the need for sector consolidation, integration, re-division of functions, the need for business processes to re-planning, so as to play the role of data warehouse.

5 CONCLUSION

In today's increasingly information technology, customers and business information has become an enterprise's most important resources. In the enterprise based on the data warehouse information management system, not only conducive to the development of enterprise information, but also help business managers in a timely manner to develop the right decision-making, improve business efficiency and enhance the enterprise in the fierce market competition advantage. Based on the deep analysis of the requirements of enterprise information management system, this paper constructs and designs the enterprise information management system based on data warehouse, and improves the decision data mining, user's use and system security. But with the continuous needs of users digging, the system's function to be strengthened, the overall performance of the system also need to be further improved, it will be the direction of the next step of the work.

REFERENCES

[1] Yang Jun. Data Warehouse Data Modeling [J]. Journal of Yancheng Institute of Technology (Natural Science Edition). 2004 (02)

[2] Jiang Bin, Zhao Guangrong. Data mining technology based on data warehouse [J]. Journal of Shandong TV University 2004 (02)

[3] Song Yueshan. Data warehouse technology research and application [J]. Computer Engineering and Applications. 2003 (33)

[4] Zhang Zhongping, Li Rong, Guo Lili. Online

analysis and analysis of the analysis and analysis [J]. Journal of Computer Applications. 2003 (08)
[5] Zhuang Xiaoqing, Xu Lizhen, Dong Yisheng. Data Clearing and Its Application in Data Warehouse [J] .Application Research of Computers. 2003 (06)
[6] Li Changshu, Tian Feng. Securities company data warehouse solution [J] .Computer Engineering. 2002 (05)
[7] Xiong Zhongyang, Zhang Yufang, Wu Zhongfu. Data Warehouse Data Loading Technology [J]. Journal of Chongqing University (Natural Science Edition). 2002 (02)