Research of Decision Support System Based on Data Warehouse Techniques

Qingtian Han
Control Engineering Department
Naval Aeronautical and Astronautical University
Yantai, China
e-mail: hbluesky@163.com

Xiaoyan Gao
Information Engineering Department
Shandong Business Institute
Yantai, china
e-mail: swallowgxy@126.com

Abstract— **This paper discusses the concepts of data warehouse technology and its importance for decision support system (DSS). The DSS as one kind smart software system of the important application value, provides, each kind of decision information as well as the commercial question solution for the enterprise. Thus Data warehouse can meet the requirements of the database management subsystems of DSS, and is fitted to form its technology frame. Then we put forth the structure of data warehouse and its main functional components. Along with widespread application of the data warehouse, DSS based on data warehouse arises at the historic moment. Begin with actual demand of the DSS, characteristic and the structure of the data warehouse have been analyzed. DSS based on the data warehouse technology is elaborated, and finally give the application example.**

Keywords- decision support system;data warehouse

I. INTRODUCTION

In recent years, the computer technology extremely rapidly develops in the management, mainly is the management of matter, money, the person. But now DSS makes the computer use in the enterprise's intelligent management, this is a very much significant action. Therefore, we combine DSS of a metal company to realize decision support based on the database technology.

At present China's software enterprise was still at the initial period development phase, although many enterprises in China ran the large information system at present, but many of them are the overseas products, and the systems developed by our country are not certainly many (it is changing more and more many), but also some small enterprises still have not the digitization .And these need the Chinese software public unceasing endeavor [1].

Overseas institutes extremely regard the research of decision analysis theory and the method, such as the international application system analysis research institute (IIASA). Since it is established, decision analysis is regarded as the main direction to develop.

The database is extremely widely applied in the society such as the financial management system, MIS, the banking transaction system and so on which are the typical database application procedures. Now lives cannot leave the database, and cannot leave the database application procedure.

II. DATA WAREHOUSE AND DECISION SUPPORT SYSTEM

A. Data Warehouse

1) The concept of Data Warehouse

The concept of DW is the most early proposed in the 80's intermediate period by IBM Corporation, and in the 90's initial period famous DW expert W.H.Inmon described DW in its work "Building DW" like the following: Integrate, Subject Oriented, Non-Volatile [2].

DW system is that the system complete information distilling to satisfy the users each kind of demand on the foundation of DW through inquires tool and the analysis tool. It composes of DW, DW Management System (DWMS), and DW tools three parts.

2) Characteristics of Data Warehouse

Compares with the traditional database, DW has the following four characteristics:

Face subject. The information in DW is organized according to the subject f, and it provides the information to the process which makes decisions according to the subject.

Integration, the information in the warehouse is extracted from original each isomerism processing database in the organization, after processing, compiling with the reorganization, then form consistent, full information about the entire enterprise.

Relatively stable. Stable refers that once some data input DW, and it will usually be kept long.

Reflects the historical changes. In DW all of data has the time attribute, which recorded the enterprise's information in each stage from past some point to the present. It will be allowed makes the quantitative analysis and the forecast to the information to enterprise's development course and the future tendency, but in operation database it main cares current some time compartment data.

3) Data organization of Data Warehouse

DW is one organization form of data storage, it obtains the primary data from the traditional database, and first it forms the current master data level according to the request of the assistance decisions subject, and then form the summarized information level according to the request of the synthesis decisions. Along with time passing, the current master data level will be transferred the historical data level by the timing control mechanism. Obviously in DW the logical organization data is composed of three or four layers data which is organized by t metadata [3].

978-0-7695-3543-2/09 $25.00 © 2009 IEEE
DOI 10.1109/WKDD.2009.96

B. Decision Support System

DSS is a class of computerized information systems that support decision making activities.

As one kind of burgeoning information technology, DSS provides each kind of decision information as well as solution of commercial question for the enterprises, thus reduces the superintendent to engage the low level information processing and the analysis burden, then makes them to concentrate the work most needs decisions wisdom and experiences, therefore improve the decisions quality and the efficiency. The superintendent needs DSS to help them to constitute the effective decisions.

During the system exploitation process it must inevitably involve the various technology such as computer network and database, data communication, window face intercalating and so on, then realize the data correspondence and the control as well as the entire company data management [4].

DW proposes to take relational database, parallel processing and distributional technical rapid development as the foundation, and the goal is the problem that there is the mass data but the useful information is deficient is settled during development of the information technology. DW technology, the on-line analytical processing (OLAP) technology and data mining are took three kinds of independent information processing technology to appear. But these three kinds of technologies actually have the intrinsic relation. In order to fully display their respective special skills, we may unify them and propose one kind of general DSS skeleton, In other words an entire set operating and implementing solution take DW as foundation, OLAP and DM in close integration (DW+OLAP+DM=DSS)is formed.

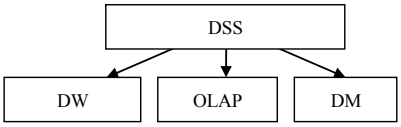

Figure 1. Compose of DSS

This implementation plan which comprehensive settles the question and makes the decision-maker use OLAP or DM tools to carry on multi-dimensional analysis about refined the information. DW technology is used in the data memory and organization, which has solved data inconsistent in DSS system; OLAP provides on-line analysis model of the data for the user; data mining devotes to the knowledge automatic diagnosis. Using three technical unions, thus implements the effective question solution decisions.

III. Design of Decision Support System

DSS is deepening of MIS application concept, and it is the system which develops on the foundation of management information. But until now what is DSS still arguing, in other words that until now there is not a recognition definition. Roughly there are two kinds of extreme views: One kind thinks so long as the system has certain supports to the decisions is DSS"; other kind think DSS is interactive computer-based system which can help the decision-maker using data and model to solve non- structure question".

According to the preceding view, nearly all systems are DSS, the data processing system can provide data to decision-maker, and the data has certain help to decisions, thus it also is DSS. Until now in the article saw the introduction of DSS, the majority is not DSS because most of them do not help to solve the non- structure problems. Some are not interactive; some databases become the model storehouse and they are not entire.

A. The reliability of the data in the Data Warehouse

In recent years, there has been a great deal of interest and investment in DW, data mining, and various tools to capture corporate business intelligence. More often than not, the priority is on "out-of-the-box" solutions and rapid development technology. While it is possible to implement a corporate DW relatively quickly, it is not necessarily true that we have cleaned data to extract and use. In fact, there is a greater probability that the data in the warehouse is not reliable or properly integrated. This is most likely due to the fact that the operational data was not reliable in the first place.

Now organizations contemplate building a corporate DW and those that have experienced failed DW projects keenly focus on the quality (integrity) of data in their existing operational systems.

What is vital to remember is that practical and robust DW is not about the technology itself. They do not implement the analytical and reporting requirements of the company by installing database engines and polished applications. The tools are certainly worth serious consideration, but they are the lesser part of any workable solution. The workable solution must essentially be about the data and the corporate business rules relating the data elements.

In charge of building the corporate DW, the very first and essential step should be to engage the experienced data architects, business analysts, and subject matter experts. By starting there and by keeping the primary goal in mind (reliable data in the DW), we have started on the right track. As always, we should expect many sayers and those who impatiently wait for the DW to magically and rapidly materialize.

For the majority of companies building DW, the basic motivation is to integrate unreliable operational data created by their poorly designed databases and poorly interfaced application systems. Most DW builders, however, usually start with a different intention: to obtain business intelligence and enable corporate online analytical analysis or reporting. This is not possible without dependable data. When it comes to defining and invoking proper data cleansing and transformation functions, they must ultimately also be experienced in resolving the issue of unreliable corporate data.

If the systems consistently annul and fragment the business rules or if the business rules are not well defined, then no amount of work performed in the staging area

(preparing, formatting and cleansing data for DW usage) will successfully result in clean data.

We will soon find that even the cleaned data in the DW cannot be directly and simply fed back into the operational systems. By the time we realize the number of data integrity problems, it will already have cost the corporation a lot of money.

We may think that corporate business intelligence cannot be derivable from fragmented and often ill-defined data structures or application systems. In the absence of proper attention to enterprise data model and data integrity issues, many DW builders are in effect, operating by such a principal. Undertake one of the most valuable and resource intensive corporate entities. The DW we have a perfect opportunity to introduce and establish corporate data standards, incorporate enterprise data architecture, clean the data, and thus maximize return on investment [5].

B. Decision Support System with Data Warehouse

DW completes extract, integration and synthesis of the decisions subject data. OLAP realizes thematic-dimensional data analysis and data mining is used to excavate information that is useful to the decisions support from database and DW. In DSS, the data analysis and decisions main strut technology is OLAP technology and data mining technology.

OLAP permits analysis man browse DW and carry on the multi-dimensional analysis by interactive way and can promptly propose information from mutative data and not too complete data the movement which relates with enterprise management. OLAP is established on the foundation of the premonition and supposition of some kind of knowledge deeply concealed in data. Questions are asked by analysis staffs, from top to bottom to search for the accuracy of confirmation supposition. And it helps users to handle enterprise information.

This design proposal first divides the DSS into many modules according to the function, then on the foundation of each module function independent, and carrying on the independent development, translation, and debugging to each module. Finally, the application system architecture is formed such as figure 2.

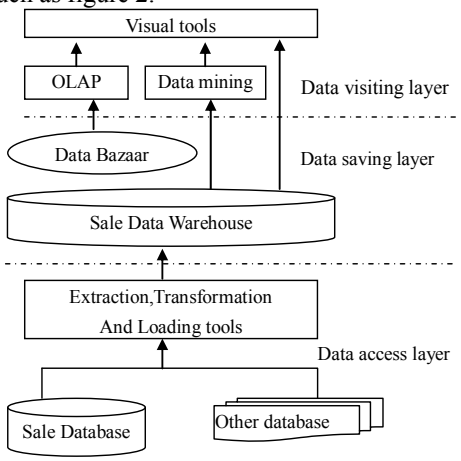

Figure 2. Sales data warehouse architecture

DW data originates from many kinds of information sources. And there are many useful information and knowledge which is unknown for the user. In order to let enterprise accurately and promptly make management decisions, but also needs one kind tool which bases on computer and intellectualized information technology to excavate each kind of knowledge which buries in the data. Therefore, data mining technology arises at the historic moment and receives recognition day by day. Data mining which is discovered and withdraws its hideaway information from DW is one kind of new technology. It establishes above DW foundation and locates to the tabletop, then establishes the new service model.

DW is the core of DSS, so construction of DW should be put on an important position, but DW construction's cost (manpower, physical resource, financial resource and time) generally is very big, but data bazaar is one kind of department progression DW which has small scale and specific application to the department and its development cost is inexpensive. Therefore, it is said to many enterprises, may start to construct DW from the establishment of data bazaar, and divides three steps to carry on: First, establish a few of dummy data fair of departments. Second, establish data fair of each department. Third, establish the enterprise progression DW. While constructing data fair, it must be planed and designed by the overall situation viewpoint and avoid to appear the situation that it is not easy smoothly to transit which when the third stage evolve to DW, even redesign creates the waste and the loss.

As a result of appearance of DW, it creates a very good condition for data distilling of DSS, thus DW may act the role as DSS database. As a result of the decisions' needs, concentrative and unified database is not able to meet the needs. So the collected data is calculated beforehand and put in DSS database. As a result of the decisions analysis needs, the collected data cannot satisfy the purpose, but also must examine historical data. DW is data management method that has both collected data and historical data. DW also may withdraw data from different type isomerism database and then process and put it to DW. The different isomerism source database possibly is many kinds of internal databases which are created by the development history of enterprise information system. It also possibly is exterior each kind of database. Along with the network technology development, it is inevitable to extract data by the exterior many kinds of databases.

DW does not like the traditional database which faces the service level. It mainly faces the high level application and carries on the decisions support. Thus in fact, it is expansion of DSS database. DW not only has the ability that the data generally is processed and collected; moreover it has the ability of deeply process data and data mining.

Anyhow, DW can support the massive data storage, fast parallel processing; preserve history sequence data, and can carry on the trend analysis; supports integration and synthesis of data; supports inquiry of the direct users. So DW is the ideal DSS database.

IV. CONCLUSIONS

This article conduct the research about establishing a DSS based on the Client/Server structure, the decisions support and application of DW is advantageous to the system development and the resources use. DW can support the massive data storage, fast parallel processing; preserve history sequence data, can carry on the trend analysis; support data integration and synthesis; supports the inquiry of the direct users, therefore DW is the ideal DSS database. DW completes extract, integration and synthesis of the decisions subject data. OLAP realizes thematic-dimensional data analysis and data mining is used to excavate information that is useful to the decisions support from database and DW. In DSS, the data analysis and decisions main strut technology is OLAP technology and data mining technology.

REFERENCES

[1] R.E. Johnson. Components, frameworks, patterns. ACM SIGSOPT Software Engineering Notes [A]. Proceedings of the 1997 symposium on Symposium on software reusability [C]. 1997.5, Volume 22 Issue 3.

[2] P. Freeman A Perspective on Reusability [J]. IEEE Computer Society Press, 1987:2-3.

[3] D.Mcllroy.Mass Produced Software Components [A].BuxtonJ Metaled. Software Engineering Concepts and Techniques [C]. 1968 NATO Conference of Software Engineering. New York : Petrocelli/Charter, 1969:88-98

[4] C. Jones. Assessment and Control of Software Risks[J].Prentice Hall PTR. 1993: 14-15.

[5] Szyperski C. Component Software: Beyond Object-oriented Programming [M]. Cambridge, MA: Addison-Wesley Publishing Company, 1997