Study on Port Business Intelligence Platform Based on Data Mart

Lei Huang and Baoyu Zhao

Abstract This paper analyzes the informationization development direction of port enterprise, puts forward a kind of port business intelligence platform based on data mart according to the characteristics of port operations. Clarify the reason of using data mart as business intelligence platform. Establish three-tier architecture of the port business intelligence platform based on data marts. Take Guangzhou Port as an example, this paper designs the star logical model and gives statistical results based on data marts. Finally, describe the current problems of data marts and look into the future of the direction.

Keywords Data mart · Port · Logic model · Business intelligence platform

1 Introduction

With the development of computer theory and information technology, modernization and informationization of management have become important indicators of modern port construction and development. However, as the technologies become widespread, the advantage brought by technology was over-shadowed. For long-term development, port enterprises not only need to improve their service level, efficiency and resource allocation, but also need to provide assistant decision information by mining the existing information produced by the OLTP system (Xing et al. 2007). It's hard to construct an effective data collection which breaks the isolation barriers among different systems, since the port enterprises do not just have a single business system. At the same time, the port information decision support system can only provide limited data inquiring and statistics. How to get the decision support information by fully mining the data generated by the port's

B. Zhao (✉) · L. Huang
School of Economics and Management, Beijing Jiaotong University, 100044 Beijing, China
e-mail: 13120648@bjtu.edu.cn

L. Huang
e-mail: lhuang@bjtu.edu.cn

Z. Zhang et al. (eds.), *LISS 2014,* DOI 10.1007/978-3-662-43871-8_66

existing information system turns to be the focus in the field of port's business intelligence.

The application of new techniques which service to system decision such as data mining, pattern recognition, expert system and multi information fusion, not only need the support of related algorithms, but also need analysis based on large amount of data. Although the traditional database has already achieved great success, it is not enough for decision analysis for lacking a large number of historical information, especially the classification statistics information. So the data warehouse and independent data mart were introduced. Presently, domestic scholars in port data mining research has introduced the concept of data warehouse, but it seems too bloated and run too much ahead of the domestic situation for the single port business process. Based on this situation, this paper puts forward a kind of port business intelligence platform based on data mart.

2 Construction of the Smart Platform of the Port Business

2.1 Architecture of Port Business Intelligence Platform

Port business data mart is a decision maker which integrated the data of the port production business system. Data mart has a big difference to the traditional database in data organization and data processing aspect. According to the definition of the data mart, it collect the data which related to the analysis and decision produced by each business subsystem for the operation of the analytical. So the separation of operational processing and analytical processing should be guaranteed in the design of port business intelligence platform architecture (Chen 2013).

As shown in Fig. 1, the port business intelligence platform is divided into: data buffer layer, data mart layer and application layer.

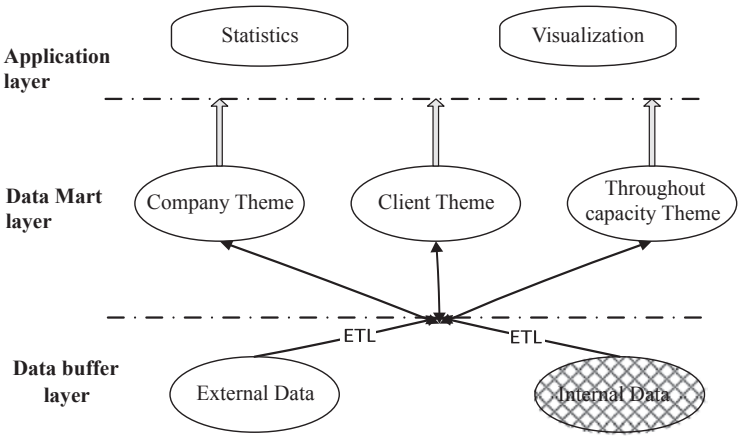

Fig. 1 Ports business data mart three-tier architecture

1. The data buffer layer

Data buffer layer contains the OLTP database and other external data interface which is the source of data for the data mart. Data kept the original form from the business system which may be data files like the port handling capacity report which is released by the national bureau each year or other database systems such as the financial system. External data and metadata summarized through the data buffer layer for the use of the data mart layer.

2. The data mart layer

Data mart layer is the core of the whole port data mart system which add the data extracted, filtered and washed by the data buffer layer grouped by topic, sorted by time, continue to increase and realized by the multidimensional data tables and relational database. This layer need of careful planning, design and standardized interface which guarantee the smooth realization of the new mart to dock with the original data mart to achieve the overall effect is best in the future development. The establishment of the data mart group is a complicated project but may shock the data warehousing industry when it was realized.

3. Application layer

The application of data marts is decision-making application. In the design of this layer, mainly for query and statistical analysis of data in the data mart, read policymakers concerned with the task of historical task of the historical data inversion or visualization deduction. In the development process, the data layer provides a unified access query interface, the application of this layer designed jointly by data mart implementers and users.

2.2 Data Model of Port Business Data Mart

Dimensional Modeling is a kind of logical design technique which use an intuitive standard framework to show the data cube and provides high-performance data access methods. It can produce a predictable and understandable design of the data mart for the user. Every unit of the cube is a measurement of the interest to the users, which can be described by a set of attributes called metrics or facts property. The metric is built based on the fundamental of the dimensions which is used to describe the particle size. Dimensional model is a star schema, also called star model which can be divided into two parts, facts Table and dimension table. The key of the model is to determine the link between the fact table and various data dimension and a lot of pre-treatment for each dimension, such as in accordance with pre-dimensional statistics, classification, sorting etc (Wang 2011).

According to analysis of the port production operations, the topic of analyzing the export volume is chosen to explain the design of the logic model which using the fact table of the storehouse ID to link various data dimensions, and establish an

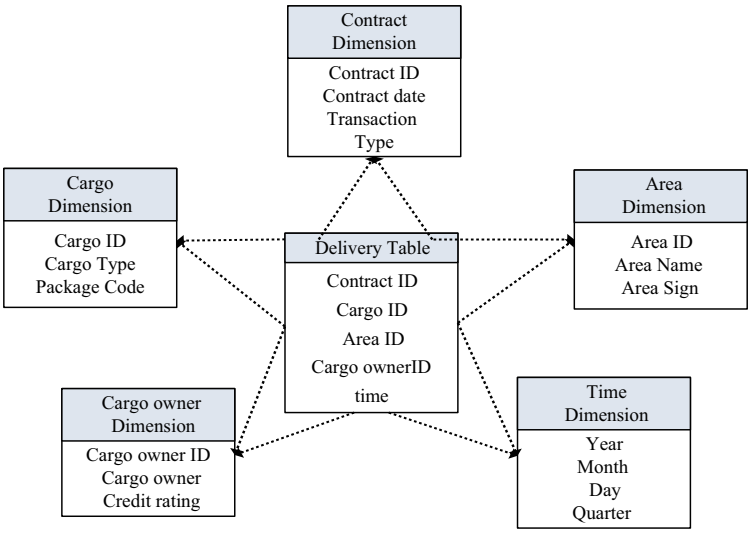

Fig. 2 Export volume analysis of the theme of the star model

export traffic statistics star model using contract dimension, regional dimension, yard maintenance, the time dimension etc. As shown in Fig. 2:

The middle of the fact table is outbound order table which linked to the various related dimension of quantitative information. The figure in the dimension table can be extended to become a snowflake model allows the analyst to flexible use of some analytical methods, such as slices, rotating, drill down to the change of observation data and trends with extremely high flexibility to analyze the data (Zhang and Qin 2013).

3 Implementation of the Smart Platform of the Port Business

The port business data mart use the On-Line Analytical Processing (OLAP) and use the Oracle Data Integrator (ODI) to extract, convers and load the data and use the UML modeling tools to design the logical model and the physical design of database using Power Designer 11. The front display using open source tools–BIRT Report as a customized demand design tool to display the ports business volume of exports according to company multi-dimensional, wherein X represents the name of the company (Com Name), Y coordinate display is the type of goods (Cargo ID), Z export volume (Cargo out AMT). As you can see from Fig. 3 the coal and products of XinSha companies in the first 5 months of 2013 got the largest shipments which is the same to the actual situation, but to ensure that new company shipments at

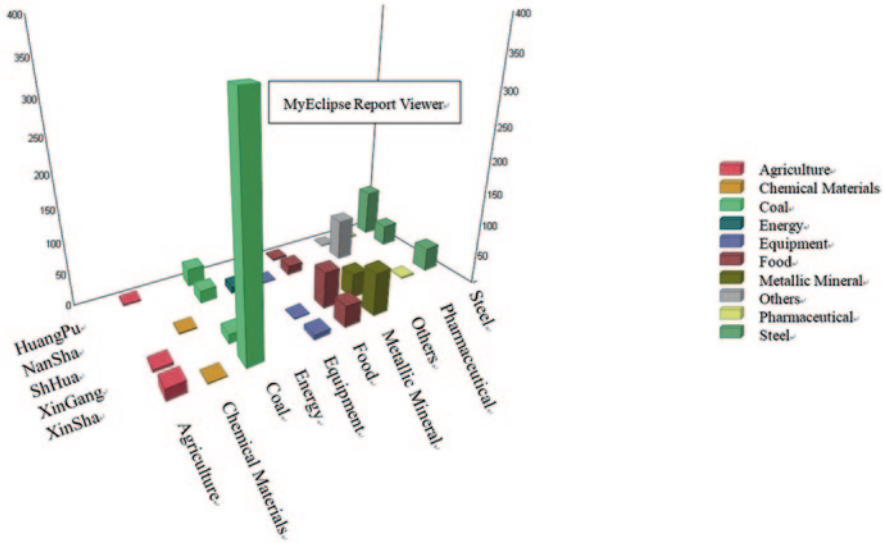

Fig. 3 Statistics show of exports of goods

the same time, the group should take measures to improve the business volume of exports of other companies, in order to improve the group's income level.

This paper based on the most common port business, the export traffic analysis, using the BIRT Report which is same as the original OLTP system as the presentation tool, but make the waiting time arise from 3.125 s to 0.124 ms, shows the advantages of data mart in query speed. What's more, the date sources used for the query based on the same business systems which have a better data environment than other isomerism data environment. So, when analyzing multiple data source and complex data mining analysis tasks, the contribution to the intelligence business platform from date marts will be enormous.

The design and usage of the Multidimensional online process cubic is the main point in the process of the date analysis. In the Oracle database the multidimensional online analytical processing cubes (Oracle Cube) of exports can be established, which can improve the multidimensional analysis ability of port production data mart. The example is just a simple one to achieve the theme on the port business intelligence platform. Multidimensional online analytical processing cubes based on data marts can analyze production data of future.

In the data show level, the report shows is just one form. Using port data mart can also create visual themes. Harbor Yard situation is shown by data marts and visualization tools. After the port data mart creation and maintenance become mature, research will focus on the establishment of data marts group in order to establish the data warehouse. And optimize mining algorithms in order to improve port productivity, increase yields of ports and yard occupancy.

4 Conclusion

This paper proposes to establish a data mart based on the port OLTP system which is used to establish the port business intelligence platform and provides a feasible and inexpensive solution for port's future decision-making system relies on data. The paper describes the advantages of data mart and looks ahead the future development of data mart group.

The solution of data mart still exist: data redundancy, redundancy, scalability and low flow, non-integration and other issues need to be further studied. However, with the continuous development of computer technology, we have reason to believe that the data mart and data warehouse technology will be widely used in more and more fields.

Acknowledgements This research was supported by "Research of Logistics Resource Integration and Scheduling Optimization" under the National Natural Science Foundation 71132008.

References

Chen R (2013) Design architecture of data mart in the development of OLAP reports. Chin E-commer 1:81–81

Wang J (2011) Step-by-step establishment of data marts integrated enterprise-level data warehouse. Comput Knowl Technol 7(11):7581–7582

Xing Y, Wang Y, Pan MX (2007) Design and implementation of port production data warehouse

Zhang S, Qin H (2013) Designs of structures and modules based on local data marts. Comput Sci 40(11A):281–283, 303