

Research on Data Warehouse Application in Prison Government Administration System

Wei YAO¹, Jin-jin CAI², Wei LI¹

¹ school of information science and technology, Agricultural University Of HeBei, Baoding, China

² M&E College Agricultural University Of Hebei, Baoding, China

Email: yaowei@hebau.edu.cn

Abstract: Found on the actual situation in the prison administration information system and analysis on the prison departments, this paper Proposals the integrated architecture and model scheme in the current prison information .The data warehouse system applied in the prison departments was designed. The paper also studied the plan and development methods of data warehouse system. Finally, the outlook and summary on the application of data warehouse in Administrative System of prison was brought forward^{[1][2][3]}

Keywords: Data warehouse; Prison administration information; System architecture; Star schema

1 Introduction

Since 1990s to now, Data Warehouse has been experienced a process from start to rapid development. "The Father of Data Warehouse" Million H. display defined it as data warehouse is a subject oriented, integrated, non-volatile and time variant data acquisition system and decision support system. With the lucubrating of data warehouse, a lot of returns come forth all over the world. Among them, Business Intelligence (BI) proposed by IBM Corporate is the most important reform and breakthrough. Based on data warehouse, methods and products of finding useful information display their talent for the first time in business and gradually grow up. [4]But the application of data warehouse in government administration is still exploring, therefore the feasibility, necessity and particularity await us to discuss. This paper analyzes the particularity of prison administration and then designs the architecture of data warehouse in prison administration and studies the method of exploitation.

2 Case Analysis

As a particular administration department, the prison department remains traditional system of units as well as special function of offender rehabilitation. Take a certain prison for example, striation it preserves management of personnel, finance, document and material. The structure mainly includes some leaders, more than twenty administrative offices and dozen of prison areas. Main administrative offices include the office, material group and prison management office. Each prison area is divided into several sub-prison areas and groups. All of prison inmates are about several thousands. Offender rehabilitation management includes two aspects: one is to quantities the behavior of prisoners in which there are fixed score, rewards and punishment, stubborn and control and evaluation of each week, month and year. The

other is some assistant measures such as preservative files of criminals, changes of criminals, change of rules, the kindred meets and daily consumption management and so on.^[5]

Because this prison had an early start on information, phonological information system, personnel management system, financial management system, IC card system and information integration system have already been used in whole prison. And the information integration system is responsible for affairs management which are not dealt by the other three systems. All of the systems use Windows platform and SQL Server 2000 database environment. Scoring and rehabilitation use C/S mode and notification use B/S mode. Network topology is star schema.

According to the above, we can summarize them as followings:

- Every single criminal has to be scored every single day. Scoring items include production and other three items and their corresponding rules.^[6]
 - At present the historical modification part of scoring sub items only can meet the scoring historical modification temporarily.
 - There are plenty of useful datum among volumes of history datum, such as a criminal's performances in reforming oneself, jail policeman's scoring and banding together. Digging out these deep information in order to control in time. Whether or not using load transfer, the system could judge from the transfer income(G) between the node I & j, G could figure out the formula (1).
- $$G_{ij} = \frac{(C_j - Lh)}{(D_{ij} * m_{ij})} \quad (1)$$
- Data mining can reflect the weak points either of administrative staffs or system function on the other side. These weak points will be improved in time.

The characteristics of data warehouse are not only very useful for solving above problems but also a good supplement for present system. It is a general trend of using data warehouse.^[7]

3 Data Warehouse System Architecture Design

Data warehouse system architecture design include five levels: modeling level, data acquisition level, data storage level, data exhibition level and metadata management level.

3.1 Design Modeling level

This part is mainly used to analysis the data source of existing system and according to theories of modeling to finish the data warehouse frame design. The data warehouse frame include four levels bottom-up: OLTP and data warehouse level, ETL level, data warehouse storage level and OLAP and data mining level.^[8]

- OLTP and data warehouse level include DB—DW and DB—ODS—DW two types. According to the case analysis, jail integrated system can satisfy a part of ODS. To solving this special problem, we can combine this two types together to save datum of four department data as well as to pick up integrated system datum. And node i will,C need to satisfy as follow condition(2). Among them ,Ci means the load value of node i , Cj means node j's, Lh means node threshold.

$$C_{ij} = \min \left\{ \left\lfloor \frac{C_j - C_i}{2} \right\rfloor, Lh - C_i \right\} \quad (2)$$

- ETL level mainly takes charge of extraction of this two different types. For global data ETL can extract and load directly but for four department data ETL still need extraction, clean, conversion and loading.
- For OLAP and data mining level, we can use different mining methods. Therefore it needs a higher expandability via object-oriented method, mining system and display system apart or expansible web plug-in architecture.^[9]

3.2 Data Acquisition Level

Define the ETL project which is from meta service system to data warehouse system, there are many problems such as heterogeneous network, heterogeneous operating system, granularity selection in data, wiping off operational data and designing of data load interface.^{[10][11]} Considering integrated problem, data warehouse adopts SQL Server 2000 data warehouse production as data storage platform. Because the specific extraction data, granularity selection and details data integration are

close related to data warehouse modeling, we will discuss it in chapter 3.

3.3 Data Storage Level

According to estimate the data quantity of data warehouse and the quantity of users' access, choosing software and hardware of data warehouse main platform and then confirm the system configuration. This data warehouse platform can be designed as a independent server which is not only asking a higher hardware configuration, a optimization large capacity storage method but also a higher bandwidth and a rapider processor speed. We choose Net Center as the physical location of data warehouse.^[12]

3.4 Data Exhibition Level

This level can display the results of user's query, OLAP operating or data mining by web server as statistical data or different comparative graphic(histogram, pie).

3.5 Meta data Management Level

It is mainly used to manage metadata including mapping from logical to physical model, data access granting and user security control in data warehouse running process.

- Metadata management includes data warehouse's building, running and maintaining though the whole life cycle. During the building process, there are three types of metadata: business metadata, technique metadata and operation metadata.^[13]
- Data warehouse user's access right is designed according to their department and distinction. Global users can query even modify historical data while other users take some or certain access rights.
- For users' access right control, system will identify users' IP, digital certificate, name and password four aspects to safeguard data security of enterprise data warehouse.

Summarizing the above five levels, the author design the data warehouse as follows,as in Figure 1.

4 Data Warehouse Design Modeling

The key parts of data warehouse modeling are enterprise model analysis, conceptual model analysis and logical model analysis. Besides these three parts, there are physical model design and data loading interface design. Partied from contents there are subject, data table segmentation and merging data organization, data append, measure and storage period.

4.1 Enterprise Model Analysis and Design

This part unite the above two parts and analyzes according to different jail departments' needs. Determine

and partitioning subject. Each departmental subject is free-running. Therefore different subject fields can be identified by different subjects.

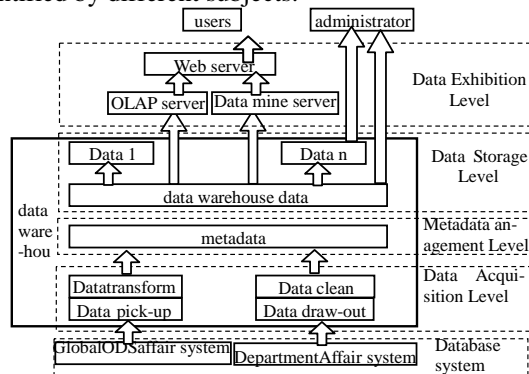


Figure1 Data warehouse Architecture

4.2 Logical Model Design

The work of logical model design is to finish dimensionality analysis, granularity level analysis, data merging strategy, relational schema definition and increasing derived field.

- Logical model design is the core of data warehouse modeling. Take scoring check for example, according to the actual demand of every week, month and year records and department level, we can design dimensionality as 3D which is misplaced by snowflake-like shape dimensionality table combined fact table: time, criminal and scoring items.
- The data hasn't display the whole snowflake-like shape and just show the scoring check part. That means as an crucial dimensionality department has close relationship with criminals.
- Because details data usually saved for at least one year, we use vertical separation triple granularity strategy. In specific, synthesis data of date, month and year of time dimensionality, and at the same time transversely part data into time-varying, sometime-varying and almost unchanged data.
- Relational schema definition is similar to traditional transaction data warehouse field definition but increasing dimensionality item and some derived fields such as time, department, same item amount and same item average.^[14]

4.3 Data Loading Interface Design

The crucial part of physical model design is to choose appropriate index strategy.

At present there are B-Tree index, Bitmap index, Bitwise index, Generalized index and Join index so on. Join index is to link index items of fact table and dimensionality table and save the result. It is used the most widely. Linking fact table index and dimensionality table index, we can get join index.

5 Conclude Remarks

This data warehouse designed by feasible technique aims at prison government administration and controls well in implementation cost. The most important is this data warehouse increases stability and integrity of data and operation. Besides above, base on the data warehouse, added by information mining and knowledge discovery and some techniques of AI, prison operation can be supervised effectively. The application of the data warehouse is strong guarantee of modern prison administration work.

References

- [1] W. H. Inmon. Building the Data Warehouse [J]. Wellesley, 1991.
- [2] Jane Zhao. Designing Distributed Data Warehouses and OLAP Systems [J]. ISTA2005, 2005.
- [3] Wang Q B. Research and Implementation of Web Page Display for Handhold Intelligent Terminal Based on Information Extraction [D]. East China Normal University 2006 (in Chinese).
- [4] Li J. Research and Implement of Network Information Extraction Oriented to The Mobile Platform [D]. Harbin Institute of Technology 2006 (in Chinese).
- [5] XU Qian, CHANG Gui-ran, YU Ke-qing. Application and Research on the visualization technology in knowledge discovery, paper volume of Qingdao-Hongkong international computer conference in 1999 [C].
- [6] ZHOU Hai-yan, GUO Jian-zhong, WANG Jia-yao, Tentative Study of Knowledge Discovery and Data Visualization Technologies [J]. Journal of Information engineering university, 2002, 3(04): 78-80.
- [7] DUAN Xiao-jun, DU Xiao-yong, YI Dong-yun. The Visual datamining technology and application [J]. Computer Application, 2000, 20(1): 54-56.
- [8] Guy Eddon, Henry Eddon. The technology of component program with VB 6.0. [M]. Beijing: Hope Electron Press, 2001. 30~51.
- [9] QIU Zhong-pan. Development manual of Visual Basic-SQL Server [M]. Electron Industry Press. 2000. 1-115, 233-244.
- [10] Jiawei Han, Micheline Kamber. Data Mining Concepts and Technique [M]. Beijing: Advanced Education Press, Morgan Kaufmann Publishers, 2001. 229-236.
- [11] WANG Jia-Hua, LIU Bing, JIANG Zao. Visualization technology in knowledge discovery [J]. supplement of Computer Science. 2000.10
- [12] ZHU Jian-qiu, CAI Wei-jie. The visual datamining [M]. Electron Industry Press. 2004. 50-85.
- [13] ZHANG Yin-kui, Datamining Principle [M]. China Machine Press. 2003.88-96R. Agrawal, T. Lmielinski, and A. Swami, Database mining: A performance perspective [J]. In IEEE Trans, 1993.