ELSEVIER

2012 International Workshop on Information and Electronics Engineering (IWIEE)

The Application Research of OLAP in Police Intelligence Decision System

Liangzhong Shen[a,*], Shihua Liu[a], Shengkai Chen[a], Xiaoyang Wang[b]

[a]*City College of Wenzhou University, Wenzhou and 325000,China*
[b]*College Of Physics and Electronic Information Engineering Wenzhou University, Wenzhou and 325000,China*

Abstract

Aiming at the large amounts of data collected by the public security organs, the technologies of data warehouse and OLAP are used to realize the police intelligence decision system based on SQL Server 2008 platform. The multidimensional analysis results reveal some potential regularity between criminal's action and the cases, so as to help the policemen make correct judgments.

© 2011 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.

Keywords: Police intelligence; OLAP ; Data Warehouse

1. Introduction

With the Golden Shield Project's continuous advance, the public security organs have build different kinds of information systems such as transient population, hotel accommodation, internet bar and crime control etc. Therefore a great deal of data can be collected either directly from the above systems or indirectly from the public communication network which have become important data sources for police intelligence analysis. However the public security Bureau utilizes these data only for basic query and statistics rather than secondary treatment and deep excavation. As a result, the police intelligence's judgments still rely mostly on the policemen's prior experience because lack of effective and scientific analysis model. Especially under the situation of insufficient grass-roots police forces , high criminal cases and increasing difficulty for cracking the criminal case, how to set up data warehouse to integrate

---

* Corresponding author. E-mail address: Johnshen0211@163.com

different system's data effectively, and then use technologies like OLAP(on-line analytical processing) and data mining to realize the police intelligence analysis has become an important issue.

Data warehousing technologies have been widely and successfully deployed in many industries like manufacturing for order shipment and customer support, retail for user profiling and inventory management, financial services for claims analysis, risk analysis, credit card analysis, and fraud detection, telecommunications for call analysis, and healthcare for outcomes analysis etc. Nevertheless these technologies are finally adopted and used in the public security systems. According to the requirement analysis from Wenzhou public security organs, the real-name information of hotel accommodation, Internet bar and flight are integrated and stored in the data warehouse. In order to formulate the corresponding decisions and measures, OLAP is used to find the regularity of various cases and the changing characteristics of social security so as to better ensure social stability and further development.

2. Data warehousing Technologies

2.1. Data Warehouse and OLAP

Compared with traditional OLTP applications, decision support systems place some rather different requirements on database technology, wherein data warehouse and OLAP are the essential elements which have increasingly become the focus of database industry. A data warehouse is a "subject-oriented, integrated, time-varying, non-volatile collection of data that is used primarily in organizational decision making." Data warehouses are targeted for decision supports which contain consolidated data from several heterogeneous operational databases over potentially long periods of time. The data in a warehouse is typically modeled multidimensional in order to facilitate complex analyses and visualization,. For example, in a hotel data warehouse (Fig 1.), date, hotel and passenger might be the dimensions of interest. These dimensions are usually hierarchical, for example dimension date may be organized as a hour-day-month-quarter-semester-year hierarchy, hotel as a hotel-police-area-city-province hierarchy and passenger as a passenger -area-city-province hierarchy.
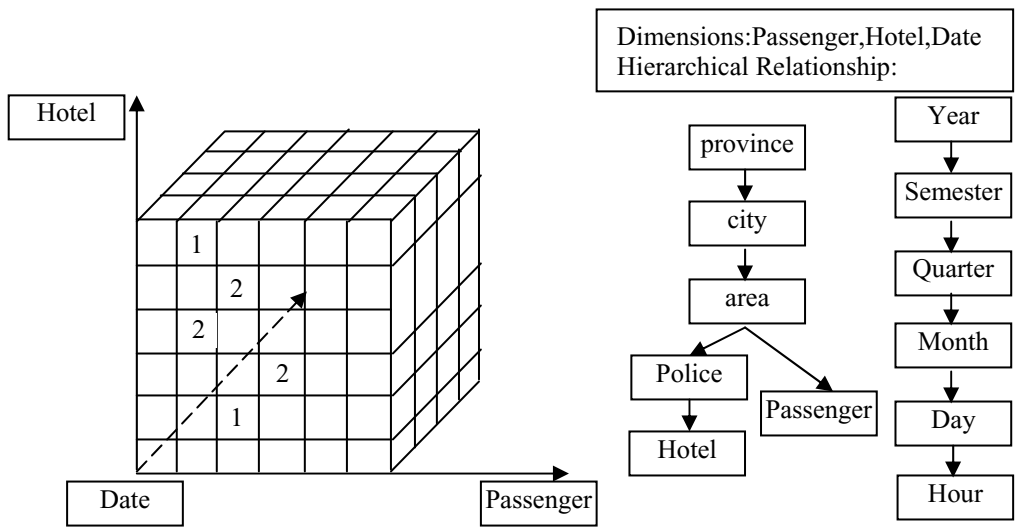

Fig1. Hotel data warehouse and corresponding dimensions

OLAP is an information analysis processing technology based on data warehouse which can enable the user observe the data from multidimensional throughout the operation of slice, dice, rollup/drill down

and rotate. Among them, slice is to choose a group of dimension member from one dimension of the multidimensional array; dice is to choose some area dimension member from one dimension of the multidimensional array; roll up takes the low layer detail data to sum up the high layer collection data in one dimension or reduces dimension, while drill down is opposite, it goes deep into detail data from collection data to observe or increases new dimension; rotate is to change the displayed dimension's direction. Therefore, the combination of data warehouse and OLAP can effectively solve the problems of how to deal with a mass of data in the decision support systems.

2.2. architecture of data warehousing

Since a data warehouse is used for decision making, it is important that the data from multiple sources should be correct. It is inevitable that when different data are integrated into the data warehouse, there is a high probability of errors and anomalies. Therefore, tools for data extraction, data cleaning, data integration and finally data load are required. When data are loaded into the data warehouse, they may be organized into several data marts according to the subjects analyzed. Data in the warehouse and data marts are stored and managed by one or more warehouse servers, which present multidimensional views of data to a variety of front end tools: query tools, analysis tools, and data mining tools. The architecture of data warehousing is as Fig 2:
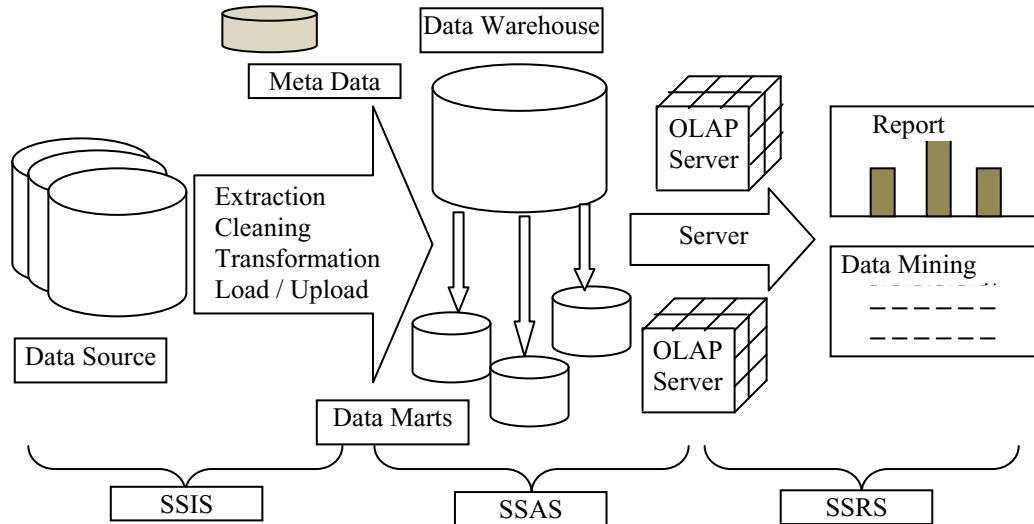

Fig. 2 architecture of data warehousing

SQL Server 2008 is used to develop the police intelligence system because it covers not only relational database management service, but also integrated service, analysis service and reporting service. Among them, integration service can help us to integrate data from different heterogeneous data sources, namely providing function of data extraction, transformation and load; analysis services provide us with function of OLAP and data mining which can help us to analyze the current situation and predict the future trend; reporting services provide us with various forms of data report and graphical display of the analysis result.

3. Police intelligence decision system

3.1. System requirement analysis

The hotel is a very complicated place where a lot of floating population come and go. Some floating criminals prefer the place around the hotel to commit the offense, or directly choose the hotel as their first choice for the commission of the offense like drug abuse, gambling, prostitution etc. For the past few years, with the development of information technology, almost all the hotels require the passenger to register when check in. Therefore the above registered data can be analyzed to provide some clues for some cases. In fact, there are many analysis points for hotel accommodation. For example: ① gender, as some illegal and criminal activity can only be committed by specific sex; ② native place, as some illegal and criminal activity has a very strong regional features; ③ hotel position, as hotel position is mostly of careful consideration by the suspect; ④ hotel's star-level, as specific criminal activity will usually choose the corresponding hotel; ⑤ stay frequency, as some criminal activity need check in hotel frequently; ⑥ check in time, as some criminal activity have specific time and their register time are usually abnormal; ⑦ ex-prisoner, check if the person is a ex-prisoner or not.
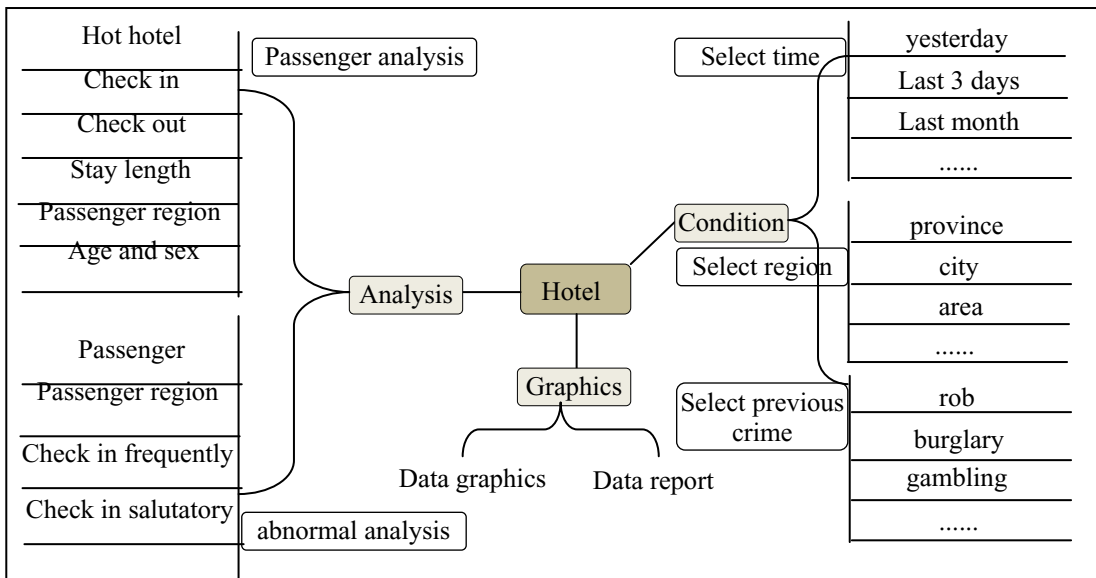

Fig 3. Demand analysis of Hotel part

According to the system requirement from Wenzhou public security organs, the questions interested are as follows: ① what are the common characteristics of those criminals reside in the hotel? ② What kind of hotel the criminals often reside in?③ What is the relationship between a case and the time of check in and check out? ④ what is the amount of hotel occupancy for each hotel every day? ⑤ What timeframe is the live peak? ⑥ what is the amount of ex-prisoner occupancy for each hotel every month? ⑦ the length of time that different ex-prisoner stays in different hotel? In summary, the requirement analysis of hotel part can be described is as Fig 3:

The requirement analysis of internet bar and flight is similar to the hotel. If the above three or more places are considered, the passenger's identity card number can be used to tract his location with real-time data updates, so that we can master a chain of evidence to predict the his future possible action.

3.2. Design of multidimensional data model

Multidimensional data model is the most popular data warehouse model. The structure of multidimensional data set is defined by the measurement value and dimensionality which come from the data source tables. Each multidimensional data set model includes at least one fact data table and one or more dimension table. Star model and snowflake model are the two common multidimensional data set model. Most data warehouses use a star schema which consists of a single fact table and several dimension tables to represent the multidimensional data model. Each tuple in the fact table consists of a foreign key to each of the dimension table that provides its multidimensional coordinates, and stores the numeric measures for those coordinates. Each dimension table consists of columns that correspond to attributes of the dimension. Because star schemas do not explicitly provide support for attribute hierarchies, snowflake schemas is usually used to provide a refinement of star schemas where the dimensional hierarchy is explicitly represented by normalizing the dimension table. This leads to advantages in maintaining the dimension table. According to the above requirements analysis, the corresponding multidimensional data model is as Fig 4:
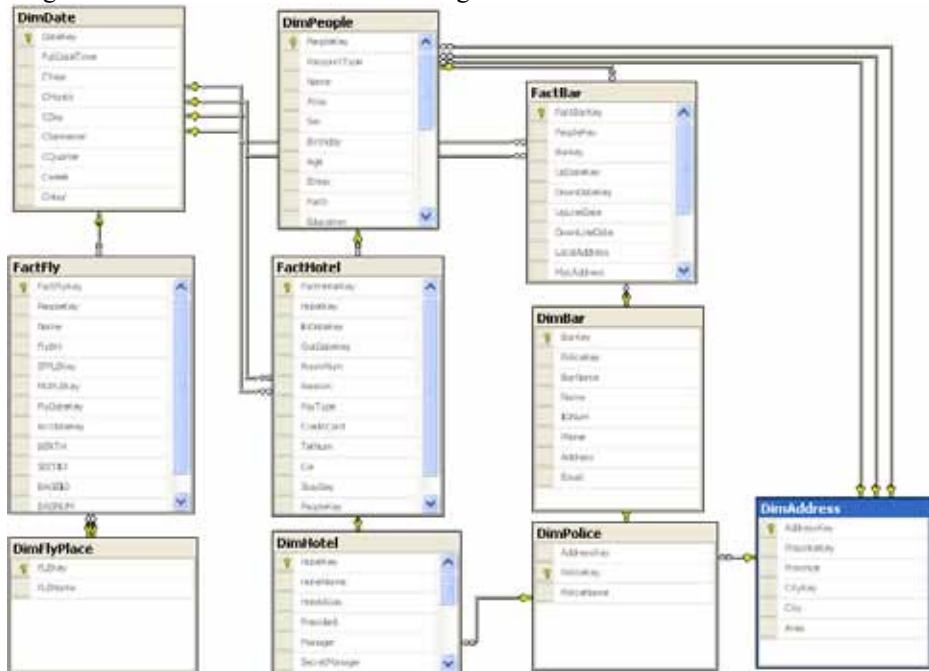

Fig. 4. Multidimensional data model

3.3. Realization of multidimensional data model

Once the multidimensional data model is ascertained, SSAS can be used to realize multidimensional analysis model. Through SSAS, multidimensional analysis data can be browsed and the operations like roll up, drill down and rotation can be performed in accordance with different requirement. The OLAP analysis of hotel accommodation and relationship between hotel and cases are respectively as Fig 5 and Fig6:

Fig 5. OLAP Analysis of Hotel Accommodation

Fig 6. OLAP Analysis between hotel and cases

4. Conclusion

In conclusion, the police intelligence decision system based on SQL Server 2008 platform can not only quickly and efficiently integrated heterogeneous data sources, but also help to explore the dynamic crime trend and crime characteristic, master the security distribution between different regions from a macroscopic view. Therefore, the technologies of data warehouse and OLAP used here are of great importance to improve the comprehensive ability of leadership and decision-making.

Acknowledgements

References

[1] Inmon, W.H., Building the Data Warehouse. John Wiley, 1992.

[2] Surajit Chaudhuri , meshwar Dayal . An Overview of Data Warehousing and OLAP Technology. ACM Sigmod Record, March 1997, P517-526

[3] CHEN Guang-yu, ZHU Jie, the police intelligent system construction method and application cases based on SQL Server, Microcomputer Information, 2011 Volumn 27(2),P135-137

[4] WANG Fu-shan, PAN Dong-jing, Application research of OLAP technology. Journal of Dezhou University, 2006, 22(5):96-99

[5] Shen Liangzhong,Liu Shihua,Chen Shengkai, Design and Implementation of ETL in Police Intelligence Data Warehouse System, 2011 International Conference on Frontiers of Manufacturing and Design Science