

Research Article

The Design of Olap-Based Multi-Dimensional Analysis System for Hospital Information

Ruijuan Hu

PLA University of Foreign Languages, Luoyang 471003, China

***Corresponding author**

Ruijuan Hu

Email: huruijuan01@126.com

Abstract: A multi-dimensional analysis of the hospital information system design is proposed based on OLAP technology which realizes OLAP multi-dimensional data analysis. A number of themes are determined in accordance with the specific demands of Puyang People's Hospital of HIS data sources and data warehouse by making use of dimension modeling approach. Logical views of subjects are constructed and multi-dimensional hospital cubes are generated by taking patient-in-hospital as an example which uses star models. Results show the application of OLAP in HIS not only provided intelligent analysis of the ways and means from the perspective of business management for hospital administrators, but also provided new means and methods from the perspective of medical research for the hospital researchers.

Keywords: OLAP, dimensional modeling, multi-dimensional data analysis, SSAS.

INTRODUCTION

The application of computer information technique in medical field has improved the digitization of medical information. Meanwhile, the wide application of electronic medical and case records and the digitization of medical equipment and instrument has accumulated a large number of data resources, which cause that the information in hospital database constantly increases. However, at present the operations on database in most hospitals simply conclude data inputting, updating, searching, counting, deleting and so on, which lack the data integration and analysis. Used OLAP and data mining technology, comprehensive analysis of medical data can search the necessary knowledge for medical staff and express the analysis results to them in a visual and simple way. Also, it will provide objective foundation for their clinical diagnosis, treatment and the study on epidemiology.

BASIC THEORIES OF DATA WAREHOUSE AND OLAP TECHNOLOGY

Basic Concepts of Data Warehouse

W. H. Inmon, the father of the data warehouse defined the data warehouse in study reports of "Recording System", "Atomic Data" and "Decision Support System": A data warehouse is a subject-oriented, integrated, nonvolatile, and time-variant collection of data in support of manager's decisions [1].

The data warehouse contains granular corporate data. Data in the data warehouse is able to be used for many different purposes, including sitting and waiting for future requirements which are unknown today.

OLAP Technology

In 1993, Online Analytical Processing(OLAP) was first put forward by E. F. Codd in "Providing OLAP to User-Analysis". He defined OLAP as an online data access and analysis technology which shares multidimensional information and focuses on certain problem. OLAP can be simply understood as multidimensional database and analysis. It aims at satisfying decision supporting or certain query and report requirements. The core technology is "dimension". Therefore, OLAP can also be considered as a set of multidimensional data analysis tools [2].

As figure 1 shows that the concrete implementation scheme of OLAP is constantly based on C/S three-tier structure. The first layer is data layer, concluding basic database and data warehouse, which is the foundation of OLAP server. It realizes the connection with basic database system and enterprise data consistency and data sharing. The second layer is OLAP server, which finally converts the clients' requirement to analysis actions and uses data in data warehouse to finish the actions. The third layer is

foreground applications, which transfer the results into visual from, such as multidimensional reports, pie chart,

histogram, algorithm, etc. for users.

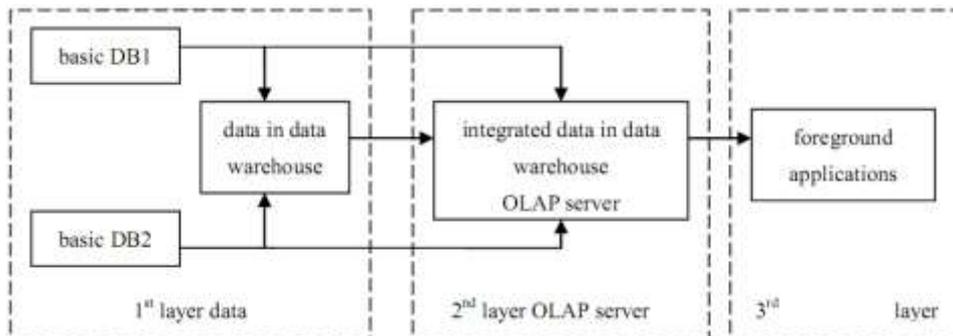


Fig-1: C/S three-tier structure of OLAP

Multidimensional cube is on the 2nd layer, the server layer in the C/S three-tier structure, which can collect data in the data warehouse and save it in a multidimensional way. It is the direct data source which can be operated of OLAP with the foreground applications. Among the data warehouse applications, OLAP applications usually be used as foreground applications, meanwhile, they can also be cooperatively used with data mining tools and statistical analysis tools. OLAP is based on the C/S structure so that it has to operate the basic operable data with a multidimensional and pre-integrated way [3].

The three-tier structure separate data, applied logic and user applications, which is beneficial to maintenance and upgrade of the system. If the functions in the system needs updating or adding, there are some of system needs updating.

DESIGNING AND IMPLEMENTATION OF OLAP-BASED MULTIDIMENSIONAL

ANALYSIS SYSTEM FOR HOSPITAL INFORMATION

General system design

Based on the basic theory of data warehouse and OLAP technique, the multidimensional analysis system for hospital information is designed and founded with dimensional modeling method. Source data from the foreground is extracted and transferred by the data introduction platform programs under the arrangement of metadata, after that it is loaded into data warehouse for the analysis data. There are not only detail data but also amount of converging investigation data from different levels and different angles. Using OLAP, multidimensional searching and data mining tools, we can analyze historical data and make reliable decisions. In this article, the HIS relational database of all the electronic medical and case records of the People's Hospital of Puyang City from April, 2014 to March, 2015 is used as data source.

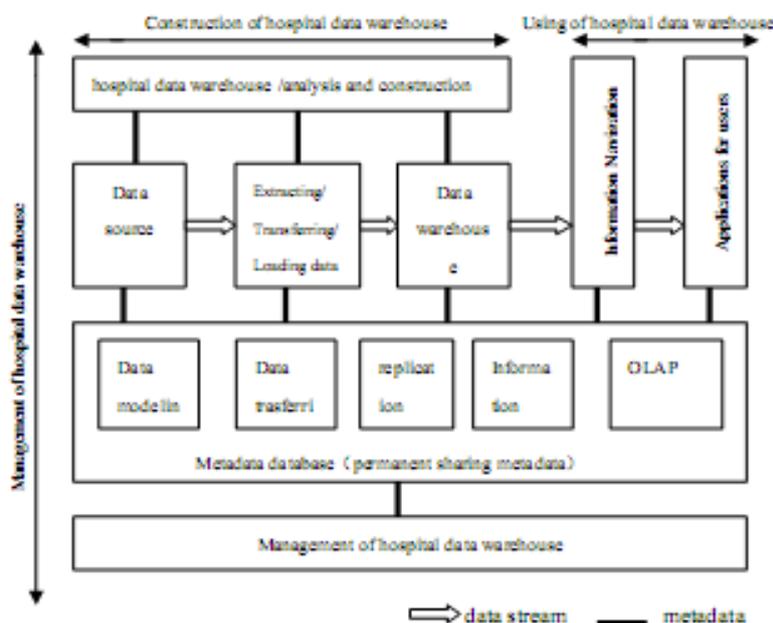


Fig-2: System Structure Graph

The main design steps of the multidimensional analysis system for hospital information:

1. Demand analysis of the multidimensional analysis system for hospital information

The People's Hospital of Puyang City is a general hospital, including outpatient, ward, logistic, administration etc... There are about 40 departments, 700 beds and a complete range of categories. The hospital has developed the automated information management in the whole hospital, which owns a certain complete medical records system. The main HIS data sources are from outpatient registration information, inpatient information (admissions records and the first page of records), outpatient and hospitalization expenses lists, drugs administration information, as well as other auxiliary information such as logistics material management, medical devices management and so on. Through talking with medical staff and managers many times, their diagnostic, analysis and management requirements are known clearly.

2. The concept design of the hospital's multidimensional cube.

According to the demand analysis, determine the system boundary and main subjects: patient information subject, outpatient information subject, drugs information subject, etc.

3. The logical design of the hospital's multidimensional cube.

After completing the concept design and determining the subjects, define information contained in the subjects in detail and begin to design dimensionality and granularity and create fact tables and dimension tables. Based on that, construct subject logic view and create hospital multidimensional data model.

Detailed system design

1. Theme determination

Puyang People's Hospital, including many departments, large amount of information, we only outpatient, inpatient, pharmacy, three departments of multidimensional data analysis. According to the guiding ideology of cube "multidimensional data model to organize" around the central theme, the scattered in various departments of the data according to different themes re organization, and in accordance with the multi-dimensional cubic structure form are stored in different sub cube cube.

According to the needs of the hospital, after repeated communication with the relevant personnel to determine the theme of the hospital information system: patient information, patient information, drug information. When the subject is well defined, the multidimensional cube logic model of the hospital information multidimensional analysis system is basically formed.

2. Dimension and granularity design

First of all, the design of the general dimension, the general dimension of shared and private dimension two. According to the analysis, the dimension is belong to several common cube: the time dimension by each cube are common; the doctor information dimension by patients and patients with cube cube have set up shared dimension, referred to as the doctor was in patients with multi dimension; inspection dimension of the hypercube and the cube of hospitalized patients a prescription by patients; dimension table and cube cube are common in hospitalized patients. The above dimensions are designed to share. The dimension tables such as the drug storage dimension and the outpatient category dimension are respectively designed by each cube.

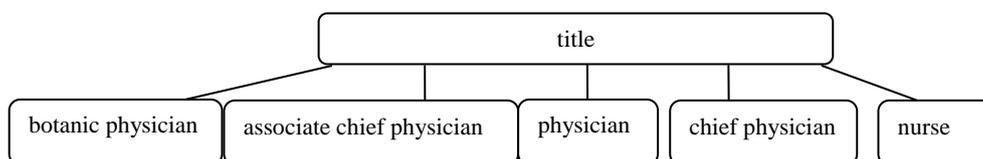


Fig-3: Concept hierarchy of title

Members of age dimension, dimension cost dimension values are numeric, more value, for convenience and easy to understand and realize the discretization of these dimensions, namely digital

generalization, thus saving storage space, generalized the age dimension concept hierarchy as shown in figure 4.

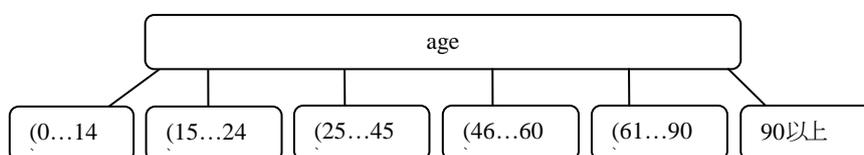


Fig-4: A discrete classification of age dimension

Granularity is one of the most important aspects of designing a data warehouse. Granularity refers to the level of refinement or synthesis of data stored in a data warehouse. Analysis of the vast majority of hospital information query system in the cube are based on a comprehensive data to a certain extent, only a few queries related to the details, in order to take into account the details of the query and the requirements of quick response, the system of different degrees of aggregation of data, aggregated data according to MOLAP structure stored in the data cube. However, leaf level data is still stored in the OLTP database, thus forming a data structure of a multi granularity.

3. Fact table and dimension table design

According to the design of the theme and the dimension and granularity of the design, the system on the theme of the fact table and the detailed design of the table [4].

Fact table:

Inpatient information (medical record number, serial number, name, sex ID, age, telephone number, address, date of hospitalization, date of discharge, doctor ID, bed number, diagnosis ID, operation ID, length of stay, hospitalization expenses)

Dimension table:

- Hospitalization time (Hospitalization date, year, quarter, month, day)
- Diagnostic information (diagnostic ID, diagnostic name, diagnostic hierarchy, diagnostic status)
- Operative information (Operative ID, type of operation, level of operation, mode of operation)
- Doctor (doctor ID, name, gender, title)
- Gender (gender ID, gender)
- Age (age ID, age)

All fact tables have two or more than two key words. The dimension table has many columns or properties that describe the rows of the dimension table. The dimension table is the entrance to the fact table, and the rich dimension attributes provide rich analysis and cutting ability. Dimension provides the user with the interface of data warehouse. Based on the above analysis and design, we use the star and snowflake architecture to construct data marts. Star model, the structure is characterized by: the central table is a large fact table, surrounded by a small set of attached tables, called the dimension table. The dimension table around the fact table displayed in the ray, like the stars explode, so called star structure. Figure 5 is a logical model of the patient's topic. The patient's information is similar to the logical view of the subject.

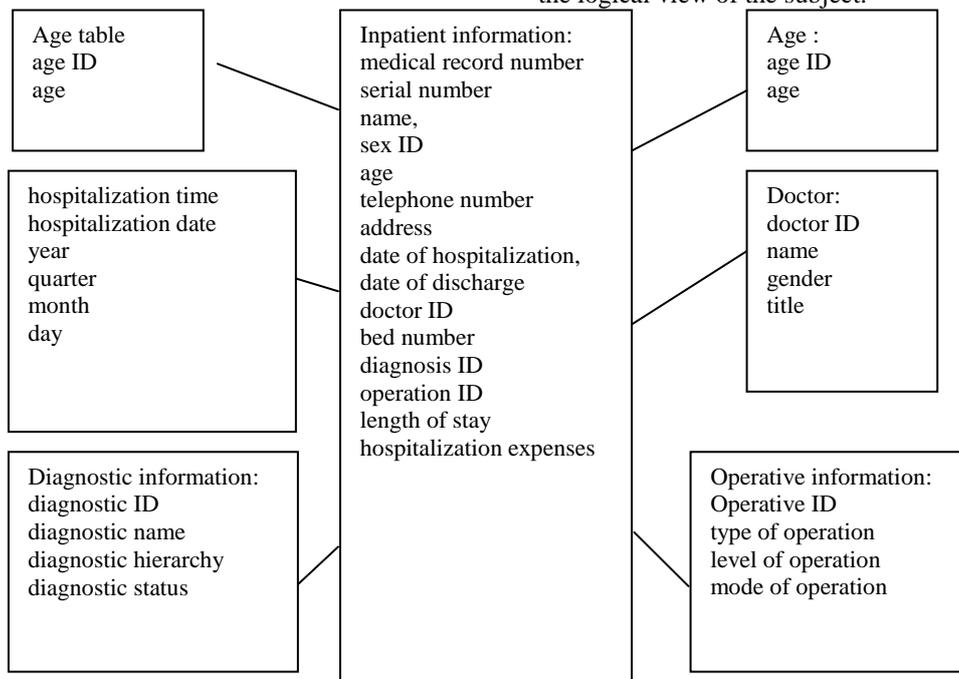


Fig-5: a logical model of the patient's topic

Establishment of multidimensional data model for hospital information

SQL Server 2012 data mining tool SSAS can be used to generate the patient's topic cube. Multidimensional data analysis can be performed according to the generated cube: for example, the data of the inpatient cost data cube is determined according

to the date of hospitalization, the name of the doctor and the operative information.

According to the formation of cube can also be used for slicing and dicing, drilling and rotating operation analysis. Multidimensional data analysis provides flexibility for users to observe data from

different perspectives. For example, the hospital can be from the time of admission, discharge time, the type of outpatient service and the patient's occupation, age, hospitalization costs and other different perspectives to observe the work of the hospital.

You can also use the MDX statement to query analysis, using MDX analysis can be used to complete all kinds of OLAP operation, simplifies the analysis process and improve the efficiency of analysis is to obtain more in-depth information, medical statistical analysis provides new ways and means of access to medical information more convenient, information available more abundant.

CONCLUSION

From the hospital information management and comprehensive analysis of many aspects of the medical data point of view, this paper researches on data warehouse and OLAP technology and applying in HIS, which establishes multi-dimensional data model and helps decision-makers analyze data query and analysis from multiple aspects, so as to draw the valuable conclusion.

REFERENCES

1. Lebdaout I, Oorhanou G, Elhajji S. An integration adaptation for real-time Datawarehousing. *International Journal of Software Engineering and Its Applications*. 2014;8(11):115-28.
2. Shan W. *Data warehouse and data analysis tutorial*. higher education press. 2012;8.
3. Nmon WHI, Wang ZH. *Data warehouse (the original book fourth edition)*. mechanical industry press. 2006;8.
4. Raisinghani MS. Adapting data modeling techniques for data warehouse design. *Journal of Computer Information Systems*. 2000;40(3):73-77.
5. Neumayr B, Schütz C, Schrefl M. Semantic Enrichment of OLAP Cubes: Multi-dimensional Ontologies and Their Representation in SQL and OWL. *On the Move to Meaningful Internet Systems: OTM 2013 Conferences*. 2013;8185(Suppl.1):624-641.