RESEARCH ARTICLE OPEN ACCESS

Multidimensional Data Model for Health Service Decision-Making Data

Bremang Appah[1], Prof. David Amos[2]

[1](African University of Science and Technology (AUST), Abuja, Nigeria.)

[2] (University of Lorraine, Nancy2, France.)

Abstract:

Decision makers are faced with increasingly stressful environments that are highly competitive, fast-paced, near real-time, overloaded with information, and data distributed throughout the organization. Moreover, every process improvement effort relies on data to provide a factual basis for making decisions throughout the organization. However, data and requirements acquisition enables decision makers to formulate and test working assumptions about a process and develop information that will lead to the improvement of the key quality features of the services being rendered to the public. The acquired data and requirement are used to design multidimensional model (thus data warehouse) for analytical processing to facilitate multidimensional analysis of the data warehouse decisions data that supports basic online analytic processing (OLAP) operations, and the decision data can be viewed from different perspective (i.e., slice-dice, drilling, pivoting etc.). Furthermore, OLAP tool is designed to support multidimensional data analysis, and targeted toward simplifying and supporting interactive data analysis to assist decision-making process and facilitates timely access to consistent data in the data warehouse.

Keywords —multidimensional modeling, data warehouse, multidimensional analysis, OLAP.

I. INTRODUCTION

The healthcare standard lacks an integrated historic electronic record keeping of data on patients' couple with improper design of databases (that is, appropriate data models) and inefficient use of resources due to lack of electronic decision-making data. For example, the manual exploratory of health service data lies in the administrators' ability to analysis the data to extract useful report, spot interesting events and trends to support decisions and hence planning to address policies that would achieve healthcare business goals. However, such manual exploratory and analysis of data is becoming impractical in health service domain as data volumes continue to grow exponentially due to continuous increase in population and health related issues. In turn, this manual exploit of dataset is slow, inefficient, ineffective and highly subjective to rely on it for improved decision making. This time spending on manual exploratoryand analyzing can be avoided when information is process quickly and timely by facilitating the use of multidimensional model to build decision support system. The process of patient medical care under health service may be described as the relationship of healthcare to a specific patient complaint and the numbers, and types of services received over time.In order to reduce the impact of falling standards in our healthcare sector, there is an urgent need for designing appropriate databases with modern technology to enhance data collection, storage and retrieval on the healthcare business services system to find answers for "What next", "goal-seeking" and "sensitivity analysis" scenarios.

Moreover, data and requirement collection improves decision-making by helping decision makers to focus on objective information about what is happening in the process rather than subjective opinion. Moreover, for a data to be collected uniformly for the multidimensional modelthe collector will need to develop a data

collection plan method and the elements of the plan must be clearly or operationally defined to establish the purpose for collecting the data. For example, it needs to note that data and requirements are collected for the purpose of improving the process in health service organization; therefore the collection must conform to the appropriate data model designing plan to suit the purpose of analyzing health service decision-making data from different perspectives. In addressing this problem, a multidimensional data model is chosen to facilitate analyzing health service decision-making data from various perspectives to enhance understanding of the information that come along with the designed. In addition, an implementation of multidimensional model to facilitate analysis in health organization services serve as a guide to uncovering inherent trends and tendencies in patients' historical information (such as patients' demographic and medical records, services received and healthcare resources).For example, multidimensional model (data warehouse)with OLAP tool brings out information that was previously hidden or that would require a cumbersome process of accessing different operational databases in order to obtain such information.

II. Methodology for Multidimensional Model

Data collection improves decision-making by helping decision makers to focus on objective information about what is happening in the process rather than subjective opinion.Moreover, for a data to be collected uniformly for the dimension model the collector must developed a data collection plan method and the elements of the plan should clearly or operationally defined to establish the purpose for collecting the required data and requirement. For instance, the data and requirements are collected for the purpose of improving the process in health service organization; therefore the collection must conform to the multidimensional model design plan. The extracteddecision making data from health service multi sources such as patients' demographic data and medical records files, resources used in treating patients' files, services received files and other procedure documents should meet the data requirements for multidimensional model.The needed benefits from this system will result in an effective discharging of duties by various functional actors in the administrative control to improve healthcare delivery, and efficiently manage the health service resources.

Requirements Analysis

This phase takes place when health service decision making data and requirementsto the dimensional data model had been elicited from transaction processing system and external sources. The health service decision data and requirements analysis is done at the same time as the initial data model in order to facilitate smooth transition from healthcare business source data and requirements to the dimensional data model. The requirements analysis techniques that are used to build an initial dimensional data model (or the data warehouse model) represent the end-user requirements captured during data and requirements elicitation. The health service requirements analysis produces a schematic representation of a data model that information analystscan interpret directly. The results of health service requirements analysis are the primary input for multidimensional modeling (i.e., data warehouse model) once they have passed the requirements validation phase. The scope of work forhealth service requirements analysis for the data warehouse model is summarized and discussed as follows:

- First, determine candidate measures (attributes), facts (central table 'or entity'), and dimensions (other tables 'or entities') including the dimension hierarchies (i.e., how entities are grouped). A query-oriented approach is used here, so measures are determine first (e.g., analysis of the healthcare business queries), and then dimensions associated with measures, followed by facts. This approach flows naturally when the requirements analyst picks up the end-user queries as the first source. Measures are usually associated with several dimensions in the relational model.
- The second stepdetermines granularities. This step formally states the lowest level of detail (i.e. the granularities) used for recording the measure in the dimensional model. The granularity of a measure is determined by the

combination of recording the details of its entire dimension. It was realized that determining the right granularities of measures in the data warehouse model is extreme importance because it basically determines the depth at which end users will be able to perform information analysis using the data warehouse.

- The third step builds the initial dimensional model. The above two steps serves as input for the initial dimensional model leading to the final data model for health service decision-making data. That is, the above based elements (i.e. measures, dimensions and dimension hierarchies and facts.) have to be identified and arranged in the model.

The last step establishes the health service business directory for the elements in the data model. This step of the initial dimensional model is the establishment of health service business directory (metadata) for documentation, which includes definitions of elements in the dimensional model.

Initial Dimensional Modeling

Fig.1 demonstrates initial dimensional modeling as a start process to design the multidimensional data model (that is, dimensional modeling). The user requirements and the source data (including source data models and existing databases models) serves as primary input for the initial dimensional model. This step illustrates a notation technique that is used to schematically document the initial dimensional model. It shows facts (or fact tables) with the measures they represent and dimension hierarchies or aggregation paths associated with the facts. The other part of the initial dimensional model is the health service business directory (metadata) documentation, which includes definitions of elements in the dimensional model that would give a better understanding of the fact tables that will be developed out of initial dimensional modeling. However, metadata defines the content and location of the data (or data model) in the data warehouse relationship between the operational system (databases) and the data warehouse, and health service business views of the data in the data warehouse as accessible to the end-user tools. Metadata is searched by users to find the subject areas and the definitions of the data. For example, in decision support, the pointers required to data warehouse are provided by the metadata. Therefore, it acts as a logical link between the decision support system application and the data warehouse. Figure 1 summarizes the health service context in which initial dimensional modeling is performed and the kind of deliverables that are produced.

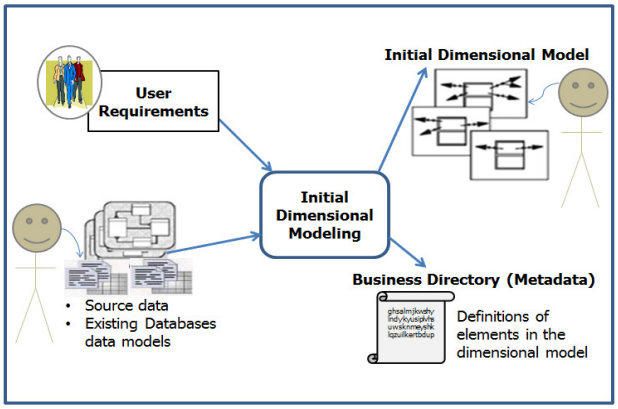

Fig.1: Initial Dimensional Modeling

Multidimensional Data Modeling

The multidimensional modeling or data cube has become one of the suitable data model for the data warehousing environment. A data warehouse is typically a read-only dedicated database system created by integrating data from multiple databases and other relevant information sources [6]. In multidimensional data modeling, data is stored as facts and dimensions, however,facts can be numerical or factual data which represents the activities specific to the health service organization business. The dimension on the other hand represents a single perspective of the data [7] and attributes of the dimension characterizes each dimension. For example, a patient dimension may consist of the attributes such as name, address, gender, age, diagnosis among others. A modeling technique namely "Star Schema or Snowflake models" are used to represent the multidimensional data. A star schema is defined as "a specific type of database design used to support analytical processing that includes a specific set of de-normalized tables" [3].The star schema is used in most data warehouse to represent the

multidimensional data model for easy access to a structured store of quality data that is used to assist in decision-making, however, the approach both simplify and improve the performance of end user queries [1]. Meanwhile decision makers often want to dig deeper into the data once initial findings are made to establish it relevance to support analyzing decision-making data from different perspectives.

As an illustration from the example in Figure 2, the star schema model that consists of a central table (fact table: Admission) and other three dimension tables (dimension: Time, Patient, and Diagnosis), which directly link to it. The star schema data model for patient admission, the fact or central table (thus Admission) contains the keys (thus TimeKey, PatientKey, and DiagnosisKey) and measurements (thus NumberOfDays and Value). Each of these dimension tables contains the attributes related to each dimension. For example, the dimension table, Patient, has the attributes (PatientKey, PatientName, and PatientCategory). Multidimensional expression (MDX) queries then use to query from central table (thus Admission) and other dimension tables (thus Time, Patient, and Diagnosis) within the star schema model with constraints on the data to return required information.

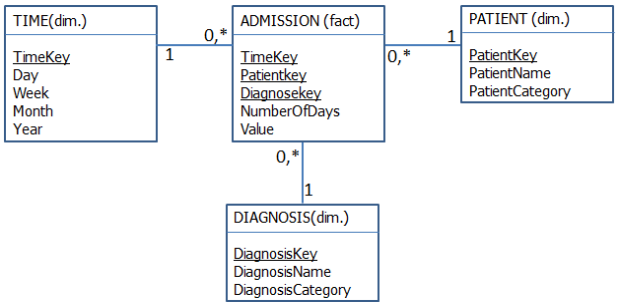

Fig. 2: Example of Patient Admission Multidimensional Data Model

Importance of Multidimensional Data Modeling

The goal of the data model is to make sure that all data objects required by the data warehouse are completely and accurately represented. In turn, the information contained in the data model will be used to define the relational tables, primary and foreign keys, store procedure and triggers. For example, the health service data warehouse design process includes planning and analysis, conceptual design, data model tool to create the logical and physical design, and implementation. As another example, the logical design ensures that all healthcare business requirements, definitions, and rules are supported whereasthe physical design of the data warehouse ensures optimal performance in the planning of indexes, relationships, data types and properties. The data model acts as documentation for the final health service data warehouse and supportsquery and reporting, multidimensional analysis, and data mining such as predictive model and link analysis. In data warehouse, quality and content are more important than retrieval respond time, however, structure and understanding of the data for access and analysis by healthcare business users like administrators is a base criterion in modeling for health service data warehousing. Data warehousing is also more concerned with data transformation, aggregation, sub-setting, controlling, and other process-oriented tasks described above [9]. In the nutshell, the health service data warehouse model also requires information about both the healthcare business source data that will be used as input and how that data will be transformed and flow to the target data warehouse.

III. Multidimensional Analysis Technique

Multidimensional analysis extends the capabilities of query and reporting in data model that is rather than submitting multiple queries; data is structured to enable fast and easy access to answers to the questions that are typically asked by healthcare business end users. For example, the data is structured in the data warehouse to include answers to the question; for example, “How many of each of pregnant women with complication was delivered safely on a particular day with assistance of a particular doctor and by the help of particular logistic?” Each separate part of that query is called a dimension. However, by pre-calculating answers to each subquery within the larger context, many answers can be readily available because the results are not recalculated with each query; they are simply accessed and displayed on a display

platform. For example, by having the results to the above query, one would automatically have the answer to any of the subqueries. That is, end user can retrieve the answer to the subquery:"How many patients with a particular sickness were cured with assistance of a particular doctor". Having the data categorized by these different dimensions make it easier to understand, particularly by healthcare business-oriented end users of the data. These dimensions can have individual entities or a hierarchy of entities such as patient, doctor, logistic, and department among others.

The multidimensional analysis enables end users in the health service to look at a large number of interdependent factors involved in a healthcare business problem and to view the data in complex relationships. That is, end usersin the healthcare business are interested in exploring the health service decision-making data at different levels of detail, which is determined dynamically. For example, the complex relationships can be analyzed through an iterative process that includes drilling down to lower levels of detail or rolling up to higher levels of summarization and aggregation. Moreover, multidimensional analysis provides the end user with integrated data and an efficient interface for rapidly manipulating views and levels of aggregation.

Data Analysis Technique	Type of Analyst	Description
Query and Reporting	Analyst Driven	▪ Analyst must pose those questions to receive an answer. Create and execute queries based on hypotheses.
Multidimensional Analysis	Analyst Assisted	▪ The data is structured to include answers to the question. ▪ Create and execute queries based on hypotheses.

Table 1: Data Analysis Technique and Analyst Role

It can be inferred that the end user can start by viewing the total revenue for the hospital and drill down to view the revenue by department, therapeutic services, support services, and finally by pharmacy.Or, the end user could start at pharmacy and roll up through the different levels to finally reach total revenue of the hospital.Therefore if the objective is to perform multidimensional data analysis, a dimensional data model is more appropriate. This type of analysis requires that the data model support a structure that enables fast and easy access to the data on the basis of any of numerous combinations of analysis dimensions. For example, an administrator may want to know how many of specific ailments were reported by patients on a specific day, in a specific department, and how much it cost for receiving treatment. Then for further analysis an administrator may want to know how many departments the patient visited for that ailment, how much it cost for receiving treatment, on a specific day. These two questions require similar information, but one viewed from a patient perspective and the other viewed from a departmentperspective. Moreover, the multidimensional analysis requires a data model that will enable the data to easily and quickly be viewed from many possible perspectives, or dimensions. For example, since a number of dimensions are being used, the model must provide a way for fast access to the data. (Summarize in Table 1)

Moreover, pivoting in the data can also be used here, which is a data analysis operation whereby the end user takes a different viewpoint than typical on the results of the analysis, and changing the way the dimensions are arranged in the result [4].

IV. CONCLUSIONS

The patient admission multidimensional data model is implemented in SQL server, however the SQL server is considered in this work because it is modern software that can store huge amount of data and information via a special database known as data warehouse. In this server, data are store easily, retrieve data or information speedily and executes the queries conveniently by the use of SQL query language. However, relations are created by manually drag and drop features in the SQL server.

This research work deep into approach for analysis of data warehouses for health service decision making data known as online analytical processing

(OLAP), An OLAP tools focus on providing multidimensional data analysis, which is essential to SQL in computing summaries and breakdowns along many dimensions that would enable health service decision-making data to be viewed from different dimension. The OLAP usually supports multidimensional data model (as data warehouse) for easy and fast access to a structured store of quality data that facilitates exploring the data at different levels of detail (that is, structured to dig deeper into the decision-making data). The data warehouse uses online analytical processing tool to facilitate information access, retrieval and interactive information analysis to assist improve decision making and hence planning in the health service.The multidimensional analysis continues in viewing the decision-making data or observations until no more drilling down or rolling up is performed on the data model. This support system will enable the health service administrators' to get together at any time and determine that they need certain types of healthcare data to track strategic goals without scrambling for data to be converted into information in order to perform reporting and analytics on patients, services received, and resources control management.

REFERENCES

[1] S. Chaudhuri,&U. Dayal, (1997). Overview of data warehousing and OLAP technology.

[2] E. F. Codd, S. B. Codd, and C. T.Salley (1993). "Providing OLAP to User-Analysts". E. F. Codd and Associates.

[3] M. Corey and M. Abbey (1999) "SQL Server 7 Data Warehousing", McGraw-Hill, Berkeley.

[4] ExecutionMih (2010). Dimensional model schemas-star, snow-flake, constellation. Retrieved from http://www.executionmih.com/data-warehouse/star-snowflake-schema.php.

[5] G. Fayyad,P. Piatetsky-Shapiro,R. Smyth, and Uthurusamy(1996). pp. 514–560. Menlo Park, Calif.: AAAI Press.

[6] Inmon W. H. (2007). Building the Data Warehouse (4th Edition), John Wiley and Sons, Inc., New York.

[7] M. Mohania, S. Samtani, J. Roddick, Y. Kambayashi(2007). "Advances and research direction in data warehousing technology", Retrieved from http://dl.acs.org.au/index.php/ajis/article/view/287

[8] Surajit, Chaudhuri and Umeshwar, Dayal, (1996). "An Overview of Data Warehousing and OLAP Technology".

[9] Shim, J. P. et. al. (2002). Past, Present, and Future of Decision Support Technology. 33(2002), pp. 111-126. Retrieved from www.elsevier.com/locate/dsw.

[10] J. Srivastava and P. Chen. (1999). Warehouse Creation – A Potential Roadblock to Data Warehousing, IEEE Transactions on Knowledge and Data Engineering, Vol.11: 1.

[11] K. Shams &M. Farishta(2001). Data warehousing: Toward knowledge management. Topics in Health Information Management, 21(3), pp. 24-32.

[12] Rajan Vohra, Nripendra Narayan Das and Meghna Sharma (2008). "Designing a Model Driven Decision Support System for Primary Health Care Managers", International Journal of information Technology and knowledge Management, Vol.1, No.2.

[13] Wickramasinghe N. & Geisler E. (2008). Encyclopedia of health care information systems,information science.