Chapter 1
A New Approach to Generate Hospital Data Warehouse Schema

Nouha Arfaoui
Institut Supérieur de Gestion de Tunis, Tunisia

Jalel Akaichi
Institut Supérieur de Gestion de Tunis, Tunisia

ABSTRACT

The healthcare industry generates huge amount of data underused for decision making needs because of the absence of specific design mastered by healthcare actors and the lack of collaboration and information exchange between the institutions. In this work, a new approach is proposed to design the schema of a Hospital Data Warehouse (HDW). It starts by generating the schemas of the Hospital Data Mart (HDM) one for each department taking into consideration the requirements of the healthcare staffs and the existing data sources. Then, it merges them to build the schema of HDW. The bottom-up approach is suitable because the healthcare departments are separately. To merge the schemas, a new schema integration methodology is used. It starts by extracting the similar elements of the schemas and the conflicts and presents them as mapping rules. Then, it transforms the rules into queries and applies them to merge the schemas.

INTRODUCTION

The healthcare industry is considered as one of the world's largest, fastest-developing and most information-rich industries (Foundation, 2006). It generates huge amount of data related to patients, drugs, doctors, etc. The collected data plays a crucial role to ensure complex statistical analysis. It is used to calculate the measurement and key performance indicators that are vital for the organization to be more agile, flexible and fluent (Mike, 2014).

The continuous development, the difficulties related to the collection of the data in the healthcare organization for the analysis, the reduction of the computing cost and the explosion of the healthcare data make the use of Hospital Data Warehouse (HDW) an efficient solution to well exploit the collected data to make good decisions.

The DW is defined as *a subject-oriented, integrated, non-volatile and time-variant collection of data in support of management's decisions* (Inmon, 2005). It allows the end-users to self-service their needs (Inmon, 2005). It may provide information to users in areas ranging from research to management (Sen and Jacob, 1998). It facilitates the storage, enhances timely analysis and increases the quality of area time decision making processes (Sahama and Croll, 2007). It offers one space to store the global truth to enable healthcare analysis such as identifying quickly the causal relationship of diseases. It aggregates, then, the data from clinical and financial systems into one repository.

The top-down starts from the description of the needs of all the users to construct the schema corresponding to the entire DW (Malinowski and Zimanyi, 2008). The bottom-up constructs the global schema of DW starting from the different schemas of Data Mart (DM) (Malinowski and Zimanyi, 2008). The hybrid approach takes advantages of the two previous approaches (Malinowski and Zimanyi, 2008). It has the speed and the user-orientation of the top-down and the integration enforced by a DW of the bottom-up.

The healthcare centers are composed by different departments such as accident and emergency, anesthetics, cardiology, diagnostic imaging, general surgery, maternity departments, neurology, Pharmacy, etc. The departments record their own data, they are still stand along, they do not communicate with other health care centers and they do not share their documents with others (Dutta, 2013). Starting by designing the schemas of the Hospital Data Mart (HDM) one for each healthcare department, then, generating the HDW schema using the bottom-up approach is very suitable and profitable in such case.

The DM is defined *as a flexible set of data, ideally based on the most atomic (granular) data possible to extract from an operational source, and presented in a symmetric (dimensional) model that is most resilient when faced with unexpected user queries* (Kimball and Ross, 2002). It is accessed directly by end users, and its data is structured in a way that is easy for users to understand and use (Moody and Kortink, 2000).

In this work, we mix two approaches: hybrid and bottom-up. The first one is used to generate the HDM taking into consideration the healthcare staffs' requirements and the existing data sources. An assistant system is introduced, at this level, to facilitate to the users the specification of their needs. The second approach is applied to build the HDW from HDM schemas. In the two steps, we use a new schema integration methodology to ensure the automatic generation of the schemas.

Starting by generating HDM helps to resolve the different problems that arise within each department because of the various care practices, data types and definitions, the perceived incompleteness of clinical information systems, the type of information that the medicine and healthcare need (Mul et al., 2012) which make the creation of the HDW in one step a very hard task.

As working hypothesis, it is proposed to present the schemas as star or snowflake because they are the most used models and they are easy to understand (Lee and Ling, 1997), (Levene and Loizou, 2003). Concerning the data sources, it is proposed to deal with Entity-Relationship (ER) database because it adopts the more natural view that the real world consists of entities and relationships; it incorporates some of the important semantic information and it can achieve a high degree of data independence (Chen, 1976).

24 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/a-new-approach-to-generate-hospital-data-warehouse-schema/246236

Related Content

Data Quality and Critical Events in Ventilation: An Intensive Care Study
Filipe Portela, Manuel Filipe Santos, António da Silva Abelha, José Machadoand Fernando Rua (2020).
Hospital Management and Emergency Medicine: Breakthroughs in Research and Practice (pp. 112-121).
www.irma-international.org/chapter/data-quality-and-critical-events-in-ventilation/246242

A Multi-Agent Based Modeling and Simulation Data Management and Analysis System for the Hospital Emergency Department
Manel Saad Saoud, Abdelhak Boubetraand Safa Attia (2020). *Hospital Management and Emergency Medicine: Breakthroughs in Research and Practice (pp. 27-44).*
www.irma-international.org/chapter/a-multi-agent-based-modeling-and-simulation-data-management-and-analysis-system-for-the-hospital-emergency-department/246237

A Soft Computing Approach for Data Routing in Hospital Area Networks (HAN)
 Rakheeand M. B. Srinivas (2020). *Hospital Management and Emergency Medicine: Breakthroughs in Research and Practice (pp. 71-83).*
www.irma-international.org/chapter/a-soft-computing-approach-for-data-routing-in-hospital-area-networks-han/246239

The Evolution of Comorbidities in Hospital Administrative Databases: A 15-Year Analysis
Alberto Freitas, Isabel Garcia Lemaand Altamiro Costa-Pereira (2020). *Hospital Management and Emergency Medicine: Breakthroughs in Research and Practice (pp. 100-111).*
www.irma-international.org/chapter/the-evolution-of-comorbidities-in-hospital-administrative-databases/246241

Surgery Operations Modeling and Scheduling in Healthcare Systems
Fatah Chetouaneand Eman Ibraheem (2020). *Hospital Management and Emergency Medicine: Breakthroughs in Research and Practice (pp. 352-366).*
www.irma-international.org/chapter/surgery-operations-modeling-and-scheduling-in-healthcare-systems/246255