International Journal of Computer Science and Applications, 2008, Vol. 5, No. 3a, pp 57 - 66

DEVELOPMENT OF A SOLAP PATRIMONY MANAGEMENT APPLICATION SYSTEM: FEZ MEDINA AS A CASE STUDY

ISMAIL SALAM

Faculty of sciences Dhar Mahraz, Department of Maths and Comuter sciences,
P.O. Box 1796, Fez, Morocco
ismailsalam@gmail.com

MOHAMMED EL MOHAJIR

Faculty of sciences Dhar Mahraz, Department of Maths and Comuter sciences,
P.O. Box 1796, Fez, Morocco
m.elmohajir@ieee.ma

ABDESLAM TALEB

ADER (Agency for Dedensification and Rehabilitation of the medina of Fez),
Fez, Morocco

BADREDDINE EL MOHAJIR

Faculty of Technical sciences, Department of computer sciences,
Fez, Morocco

ABSTRACT

It is well known that transactional and analytical systems each require different database architecture. In general, the database structure of transactional systems is optimized for consistency and efficient updates while the database structure for decision-support systems is optimized for complex query analysis and key performance indicators reporting. Spatial data has also become an important ingredient to consider in data warehouses to support spatial-temporal analysis aims.

To monitor their activities and measure the performance of their procedures, today's organizations deploy data warehouses and client tools such as OLAP (On-Line Analytical Processing) to access, visualize, and analyze their integrated, aggregated and summarized data. Since a large part of these data has a spatial component, suitable tools are required to take full advantage of the geometry of the spatial phenomena or objects being analyzed. With this regard, Spatial OLAP (SOLAP) technology offers promising possibilities. Our study focuses on the design of a SOLAP application prototype that can be used to analyze and monitor the restoration requirement and profitability of commercial ancient houses. The Medina of Fez is taken as a case study for developing our prototype. Diagrams and geographical maps are produced for the key indicators. These can be considered as powerful intuitive charts to support visual spatio-temporal analysis.

KEYWORDS

Multidimensional data warehouse architecture; spatial dimension; Spatial On-Line Analytical processing (SOLAP); Geographic Information Systems (GIS).

1. Introduction

The data stored in the data warehouses is the basis of any analysis that guides the organization in its decision-making process (Vuibert, 2001). However, the data is not always exploited to their full potential and richness (Han et al. 2001), i.e., their spatial component, is often unexploited. "Hidden in most data is a geographical component that can be tied to a place: an address, postal code, global positioning system location, (…) region or country" (ArcView, 2000). The simple visualization of this spatial component makes it possible to meet a first need, which make easier the understanding of the phenomenon in question by identifying position within a geographical framework, by looking to its extent on the territory and seeing its distribution on the territory (concentrated, dispersed, by groups, random, regular, etc) (Bédard et al. 2001). GIS are potential candidates for spatial analysis but despite interesting spatial-temporal analysis capabilities, it is recognized that actual GIS systems are not optimally designed to be used to support decision applications and therefore coupling SIG and OLAP solutions should be realized (Bédard et al. 2005). The visualization option makes it possible to discover information non-visible in traditional OLAP tools, that is to say modes of geographical distribution of the phenomenon not following preset territorial borders. In this context, the Spatial OLAP (SOLAP) tools offer promising possibilities. A SOLAP tool can be defined as "a visual platform built especially to support rapid and easy spatial-temporal analysis and exploration of data following a multidimensional approach comprised of aggregation levels available in cartographic displays as well as in tabular and diagram displays" (Bédard, Y 1997).

Our study aims to provide a system that will help decision makers to preserve and promote the architectural patrimony and attractiveness of historical buildings and ancient houses. Main focus is on the officially declared protected spaces that offer historical, aesthetic and cultural values. The medina of Fez (UNESCO World Heritage imperial Moroccan city) is considered for our case study and for which the authorities, ADER agency included, aim to preserve and promote the architectural, urban, ethnological heritage as well as the artistic richness.

The information system prototype developed will help, on the one hand, to manage in an optimal way the incomes generated by the visits to the commercial ancient houses (restaurant, museum …etc), and on the other hand, acting on time and in a efficient way for maintaining and restoring the threatened houses ruins. The objective of this system would thus be to offer relevant key indicators to the decision makers. These indicators reflect a real image of the state of the ancient houses and their profitability rates; profitability can be obtained by recording the increase of the number of visitors in the commercial ancient house (price of ticket of entry, purchases of souvenirs and accessories…etc). Our study doesn't take into account any patrimonial objects others than ancient houses (Riad). We are particularly interested in those where we can organize visits of short duration (museum, restaurant, artisan buildings …).

2. Definition and Design

Various information systems exist in different organizations and institutions in Fez (ADER, Municipal, INDH, ONG, …) that encompass large amount of historical information related to the evolution of the state of the medina of Fez. This data is unfortunately not exploited to its full potential as it is stored mainly in production transactional systems. Our objective here is to come up with a decision support system that can help on taking accurate and on-time decisions for the arrangement and the management of historical sites.

The functional specifications we have set for the proposed system are as follows:

- Report on actual architectural state and level of criticality for restoration of the ancient sites
- Report on the income of the commercial sites
- Report on the visitor's spending in the commercial buildings

These key indicators will help on the follow-up of the restoration operations, on estimating the necessary budget by the state of each house, and on monitoring and re-organizing activities. At this stage, the system is obviously not intended to offer a platform for comprehensive management of the Medina heritage.

We have first created an object-relational model where we have grouped relevant information related to application domain like the visit entity with its duration and cost. This model will serve as a basis to the construction of the star schema for our data warehouse. The model should ensure two essential system functions. On the one hand, it should offer the possibility to identify each visit done to the commercial ancient house as well as the visitor spending during this visit. On the other hand, the model should contain most relevant attributes that an operational IS should include to satisfy the analyst queries; like house construction date, address …etc.

The following figure presents our object-oriented model that we have defined within this framework:

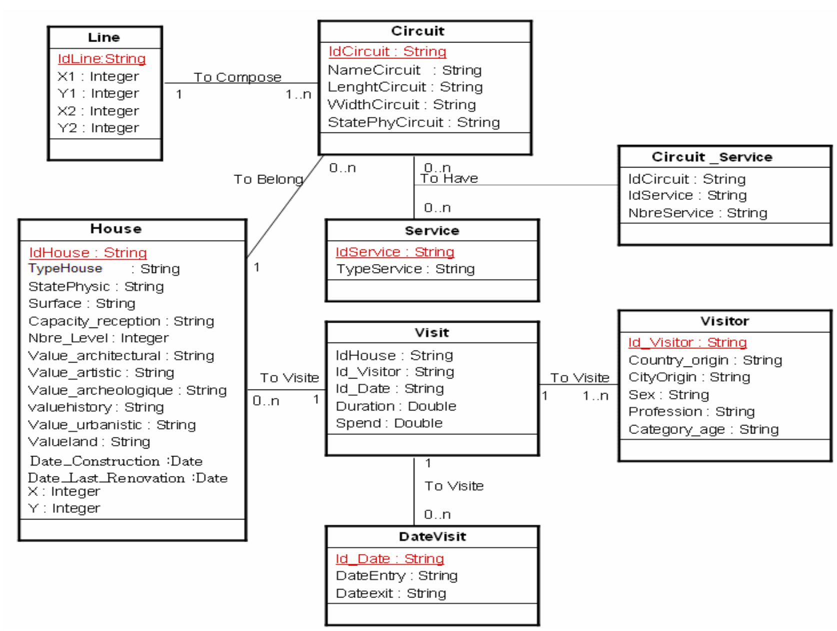

Fig. 1. Object-relational model for our case-study: restoration and management of historical sites of Fez Medina

2.1. Architecture

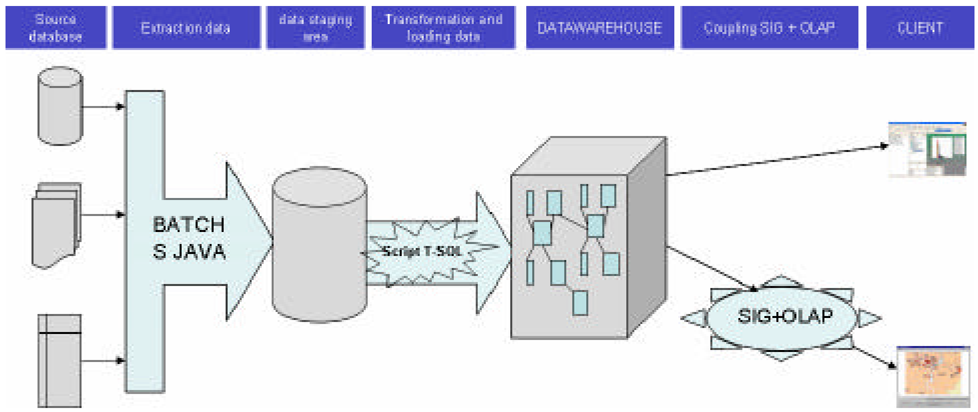

Fig. 2. Global architecture of a decisional system

The figure above represents the different components that make up the architecture of our decision-support system. These components offer a high level of flexibility and scalability for the agency to implement a Business Intelligence solution.

Source Systems: A data source system is the operational or legacy system of record whose function is to capture the transactions of the business

Data Staging Area: The data staging area is the portion of the data warehouse restricted to extracting, cleaning, matching and loading data from multiple legacy systems.

Data Warehouse Database: The warehouse is no special technology in itself. The data warehouse database is a relational data structure that is optimized for distribution. It collects and stores integrated sets of historical, non-volatile data from multiple operational systems and feeds them to one or more data marts.

2.2. Multidimensional Model

The entity-relationship data model is normally used in the development of On-Line Transaction Processing (OLTP) systems. The entity-relationship diagram leads to the construction of the operational production database. This can however not serve as a platform for extracting key indicators on which are based the organization performances. Indeed, the entity-relationship model normalizes data and incorporates many table joins increasing by this the query time responses. This limitation is surmounted with the generation of the multidimensional model that is more suitable with the functional and performance requirements of OLAP systems.

The most popular multidimensional model for relational OLAP (ROLAP) is certainly the star schema (Kimball, R. 1996). This model is centered on a fact table containing measures with related dimension tables, which characterize these facts. Each dimension has a number of attributes used for selection or grouping. A dimension is usually organized in hierarchies supporting different levels of data aggregation as well as multiple

inheritances. The snowflake schema is a variant model where the hierarchies in the dimensions are explicit following normalized tables. (Pesersen et al. 2001) analyzed 14 multidimensional data models including star and snowflake models and showed that these models do not support requirements such as multi dimension in each dimension, non strict hierarchies, handling change and time and handling different levels of granularity. They defined an extended multidimensional data model for these requirements; their model is adapted for imprecise data.
Spatial dimension: Describes the representation of the territory surface, it could comprise specific members, but that would restrict the cartographic representation of the data at one moment given according to only valid territorial cutting to this moment.
For our case study of the medina of Fez, the multidimensional model contains a central fact table (Fact_Visite) containing the bulk of the data, with no redundancy and connected to a set of smaller attendant tables (dimension tables), like time, house, visitor and circuit.

2.1.1. Star schema diagram

The star schema we built for our case study relates to the subject of managing the patrimony of the Medina of Fez. The system intends to raise alerts for prompt architectural restoration and also to monitor the income of the commercial buildings generated by the visits. The decision maker or anyone responsible of the general management of the Medina will be able to know for example the most visited houses, who are the visitors by category of age, gender (male/female), the actual state of the buildings and if they require immediate restoration, revenue generated by each commercial building to classify criticality of restoration …etc. We have therefore the following measurements in the fact table:

- The incomes in DH spent by a visitor (Spending).
- Duration in hour of a visit.

In the object-relational model that we established, the class “Visit” contains the duration and the price spent by a visitor. The spending of a visitor during a visit in a commercial ancient house is the price of the ticket paid at the entry and the amount of the purchases and the consumption of services sold inside the commercial houses (purchase of a product, taking photographs…). The duration is the time spent by a visitor inside the house and it is calculated by the difference of the entry time and the exit time.

Dimensions on which these measures can be projected are:

- “Time” dimension is inherent to any star schema model. Indeed, in our study we calculate the duration and the costs of the visits. The attributes chosen for the “Date” dimension are the year, quarter, month and day.
- “House” dimension describes the ancient house. The attributes of this dimensions are id_house that identify each ancient house, Type_house that determine the type of the house (hotel, restaurant, museum.) , etat_house describes the state of the house (very good, good, average, degraded, very degraded), Area, number of

floors, capacity, architectural, artistic, urban and archaeological values of the house (Large - Average - Low - Without), date_condtuction represent date of construction, date_last_renovation represent date of last renovation and finally the spatial coordinates X and Y to determine the exact localization of the house.

- "Visitor" dimension describes the visitor. We chose as attributes for this dimension city and country of origin, gender, profession, and age.
- "Circuit" dimension describes the road where the ancient house is located. Its attributes are Id_Circuit as circuit identifier, name, state, width and length of the circuit, and the service located along this circuit (Bank, Police, Pharmacy, Hotel).
- "Line" sub-dimension of the "Circuit" dimension contains spatial attributes that determine the exact location.

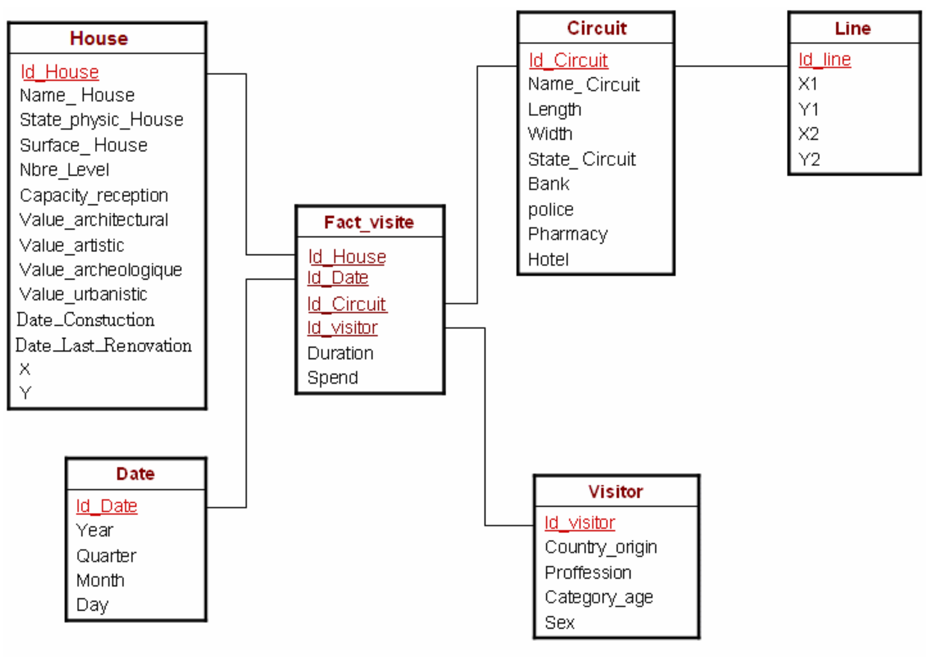

Fig. 3. Star schema of the data warehouse related to our case study.

3. Prototype implementation

The proposed system has been developed using MS SQL Server and analysis Services. The multidimensional structure has been stored in a multidimensional OLAP (MOLAP) database. The data (raw and aggregated data) in MOLAP is stored as series of multidimensional arrays. MOLAP database respond faster to multidimensional queries since the structures are highly denormalized and stored in RAM.

Otherwise, our SOLAP prototype has been designed to support the spatial-temporal exploration of multidimensional structures using Proclarity that includes software for business intelligence management, MapInfo as software of geographical information system and Intergraph Geomedia software component (MAPX). We developed an application in visual basic to drive Proclarity functionalities to perform OLAP operators on the descriptive data and Geomedia to navigate on spatial dimension and display charts which was developed in MapInfo.

Our SOLAP application interface allows the generation of diagrams that depict specific measures on selected dimensions. Cubes can be generated with the measures projected on prescribed set of dimensions (axis of analysis). The available operators like drilling (up and down), roll up and aggregating can also be applied to explore and display data under different angles. Data contained in tabular or graphs like histograms (Fig. 5; Fig.6) can be exported to an ASCII text file to be used for other purposes.
In addition to the above type of data representation, our application allows a cartographic representation of any measures for the spatial dimension. Figure 6 shows an example of a map that can be generated. The values are based on artificial data that has been generated to test and validate our prototype.

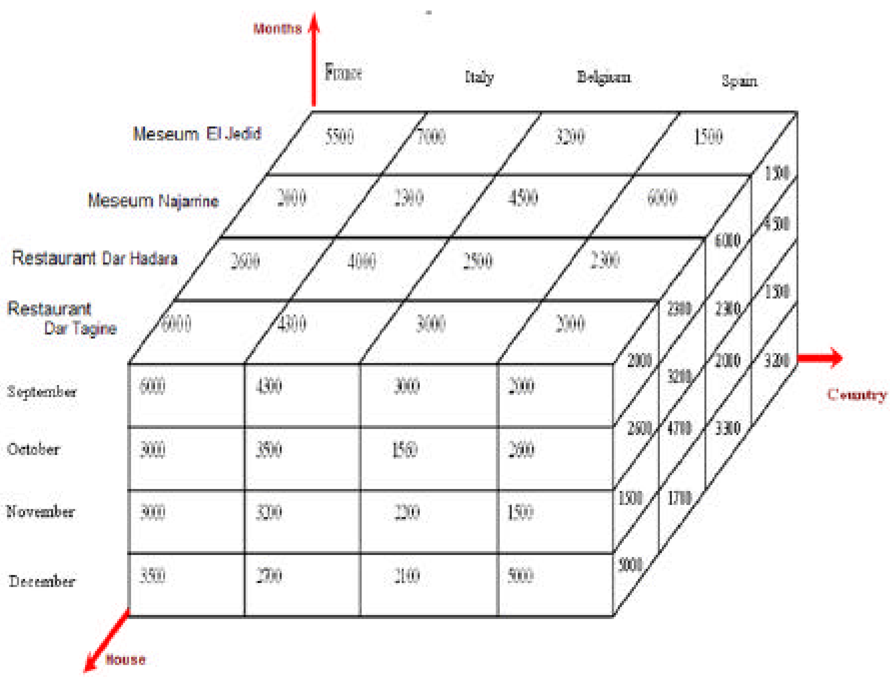

Fig. 4. Example of cubic representation of factual measures.

Figure 4 is an example of cubic representation data of our model, with 3 axes of analysis: month (time), house and country of origin for visitors, and “visitor spending” as indicator.

Figure 5 and 6 show respectively a spreadsheet and a histogram that have been generated to analyze the spending of the visitor for the different commercial ancient houses and their street for the month of September 2006.

Measures: Depense

HOUSE	STREET	septembre
Medersa Es Sahrij	Sidi Youssef	1 200
Medersa EL Attarine	Attarine	1 800
Tsetaouine	Attarine	900
Medersa Mesbahia	Derb Seba	1 500
Kairouine	Attarine	1 040
Mderssa sefarine	Sefarine	320
Mdersa Cheratine	Charatine	1 950
Zaouia Mly Driss II	Nejarine	1 000
El jedid	Najarine	740
Dar Tagine	Ross Rih	600
Kissaria	Attarine	1 050
Mosque Er Rsif	Place Er Rsif	330
Mosque Andalous	Derb Seffah	250
Dar Hadara	Ziate	700
Mdersa El Oued	Derb Lmdersa	330
Mosque Sidi Lemlili	Andalous	300

Fig. 5. Fez visitors' spending (in thousands Moroccan DH) in ancient houses in September 2006.

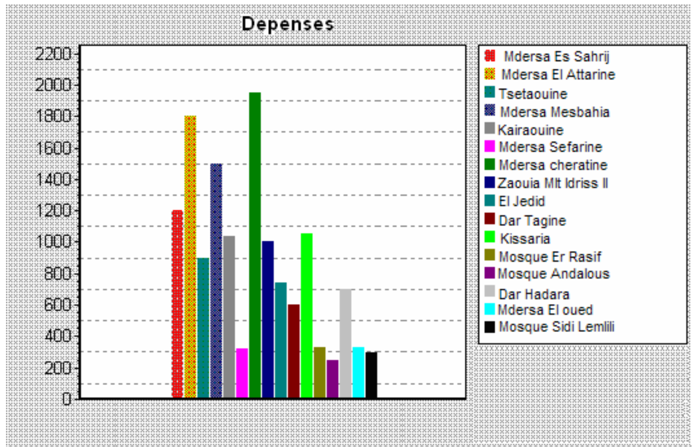

Fig. 6. Visitors' spending (in thousands DH) in ancient houses in September 2006.

Figure 7 can be considered as a geographical map component where we have generated cartographic representation of the spatial dimension.

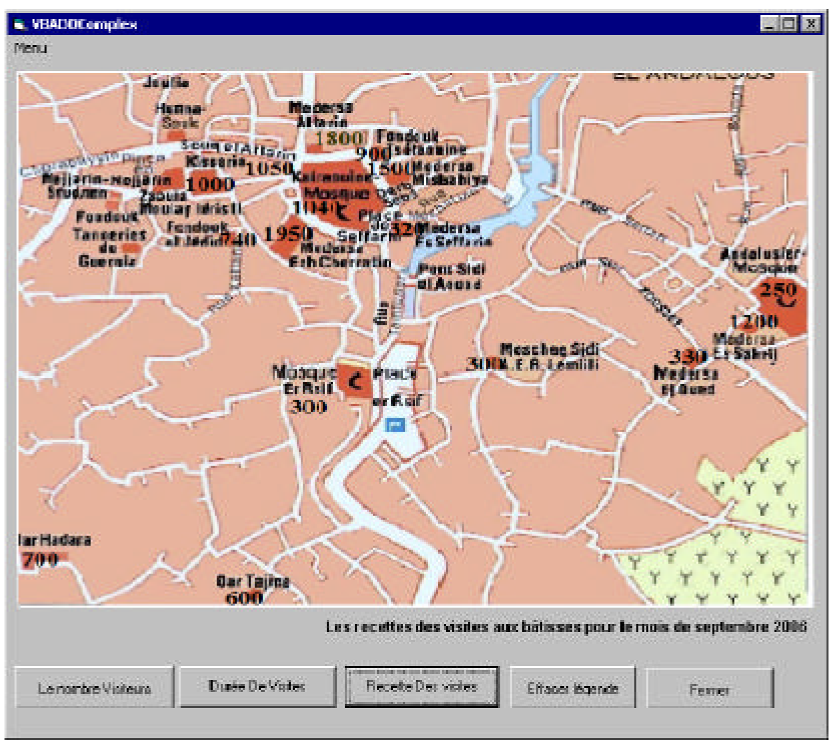

Figure 7: Ancient houses heritage management application (SOLAP tool : ProClarity, Mapinfo Mapx and Visuel Basic) : Presentation of visitors' spending (in thousands Moroccan DH) in various ancient houses for the month of September.

strategically important information system, the spatial data warehouse will play an important role in the new generation of spatial information systems since it provides a unified view of integrated spatial data from heterogeneous spatial databases, cleansed stores, integrated and transformed data, and facilitates multiple dimensional spatial data analysis. Spatial OLAP is a highly advantageous data exploration facility in spatial data warehouse. It provides fast, flexible, and multidimensional options for spatial data analysis. However, great challenges still exist for efficient implementation of such mechanisms in large spatial databases. Our application prototype is an attempt to show the interest of combining GIS and OLAP. It has indeed helped us to generate, in addition to the classical diagrams and views, visual maps that are more intuitive and efficient. We plan as future work to elaborate further the functional specifications of our SOLAP system to respond more realistically to the needs of a patrimony management organisation. We intend also to collect realistic data for our ETL process to assess the system performance and the quality of the reports and the maps generated.

References

[1] ArcView, 2000. GIS Brochure (Booklet)
[2] Proclarity Corporation is a software company specializing in business intelligence http://proclarity.com/
[3] Bédard, Y. 1997, Novembre. Spatial OLAP, Vidéoconférene. 2e Forum annuel sur la R-D, Géomatique VI: Un monde accessible, Montréal, Canada.
[4] Bédard, T. Merrett, J. Han, 2001. Fundamentals of Spatial Data Warehousing for Geographic Knowledge Discovery, Geographic Data Mining and Knowledge Discovery, Taylor & Francis, Research Monographs in GIS, Chap. 3, p. 53-73.
[5] Bédard, Y., Merrett, T., Han, J. 2001. Fundamentals of spatial data warehousing for geographic knowledge discovery. In H. Miller & J. Han (Eds.), Geographic Data Mining and Knowledge Discovery. London: Taylor & Francis.
[6] Bédard, Y., Proulx, M.J., Rivest, S., 2005. Enrichissement du OLAP pour l'analyse géographique: exemples de réalisation et différentes possibilités technologiques. In: Bentayeb, F., Boussaid, O., Darmont, J., Rabaseda, S. (Eds.), Entrepôts de Données et Analyse en ligne, RNTI B_1. Paris: Cépaduès, pp. 1–20.
[7] Han, J., Kamber M., 2001. Data Mining: concepts and Techniques, Morgan Kaufmann Publisher, Inc, 2001.
[8] Kimball R., 1996 The Data Warehouse Toolkit. J. Wiley and Sons.
[9] Kimball R. (Vuibert, 2001) « Entrepôts de données, Guide pratique du concepteur de data warehouse ».
[10] Pedersen T.B., Jensen C., Dyreson C., 2001 A foundation for capturing and querying complex multidimensional data, Information Systems 26 383-423.