Modelling Financial-Accounting Decisions by Means of OLAP Tools

Diana Elena CODREANU
Universitatea "Constantin Brancoveanu", Piteşti
codreanudia@yahoo.com

At present, one can say that a company's good running largely depends on the information quantity and quality it relies on when making decisions. The information needed to underlie decisions and be obtained due to the existence of a high-performing information system which makes it possible for the data to be shown quickly, synthetically and truly, also providing the opportunity for complex analyses and predictions.
In such circumstances, computerized accounting systems, too, have grown their complexity by means of data analyzing information solutions such as OLAP and Data Mining which help perform a multidimensional analysis of financial-accounting data, potential frauds can be detected and data hidden information can be revealed, trends for certain indicators can be set up, therefore ensuring useful information to a company's decision making.
Keywords**:** analytic proces, data warehousing, databases, decision support systems, OLAP cubes.

1 Introduction

In the current context of the world economy, the performance of an organization is both ensured and conditioned by the quality of the decisions made by its manager. Making the best decision relies upon a large amount of information as well as a complex process of its analysis and synthesis.

When analyzing the global efficiency of an enterprise, the quantity and quality of financial-accounting information in the system of the economic unit at a precise moment is specially important, preferably the information at more and more analytical levels [1].

To an enterprise, the use of financial-accounting information is the raw material when making the best managerial decisions.

A decision is the result of a conscious process of choosing an action direction and getting involved in it which usually implies the allocation of resources. A decision ensues as a consequence of processing certain information and knowledge, and belongs to a person or a group of people who has/have necessary authority and is/are responsible for the efficient use of resources in certain circumstances [2].

OLAP Tools

OLAP is the accronym of On-Line Analytical Processing and, as the words themselves show, it can be said that OLAP within an organization is used to provide easy and fast access to the analytical resources underlying decision-making and management processes.

Specialists think that an OLAP system is an information server which allows quick access to data (atomic and derived data) and complex calculation facilities [3].

OLAP has quickly become the basis of smart solutions and one can also mention the high-performing management in business environments, planning, resource allocation, budget allocation, predictions, financial report drafting, data discoveries and data storehouse reports.

The 60's -70's were hallmarked by the emergence of the concept of On-line Analitycal Processing used to model the financial activities in an organization by means of analytical functions. Thus, in the year 1962, Ken Iverson, in his paper called "*A Programming Language",* described multidimensional language for the first time, namely *APL* (*A programming Language*). The language has been developed by IBM Company and used on mainframes ever since 1962. A large amount of the concepts in this language are still used nowadays and examples are the languages called Adaytum Planning and Lex 2000.

OLAP Council was founded in 1995, made up of companies concerned with developing OLAP products and having as main goal the removal of confusion and the desire for OLAP systems to become more interesting on the market by setting up open standards (OLAP API). OLAP Council has defined OLAP technology as a *”category of software tools which allows analysts, managers and directors to understand the essence of data by the rapid, substantial and interactive access to a wide range of potential visions upon information which has been obtained by the change of primary data so that to reflect an enterprise's real dimensions as perceived and understood by users”* [4].

The analytical functions and management facilities of data were integrated within a language in 1972 called Express Language. Express is still one of the main OLAP technologies in use, as its data model and concepts have remained unchanged.

OLAP provides users with the opportunity to analyze several dimensions on the spot, making sure all the necessary information is available to make the best decisions.

OLAP technology is characterized by a dynamic, multidimensional analysis which helps end-users by means of various activities [5]:

- in-depth analysis (drill-down);
- opportunity to make predictions during various periods of time;
- use of formulae and models for dimensions and rankings;
- extracting a sub-set of data to be viewed;
- rotations within dimensions.

OLAP systems have been included in data-oriented decision-making support systems and here are a few SSD notions since decision support systems (SSD) theory leaves its mark upon the OLAP systems theory:

- In early 70's, the first definition of SSD was stated by Little who asserted that an SSD is *a model relying on a series of procedures for data processing and for assisting a manager during decision making. An SSD should be simple, robust, easy to maintain, adjustable, available for communication etc.* The features identified by Little are still valid at present.
- In the paper entitled ”Foundation of Decision Support Systems”, authors Bonczek and Holsaple defined an SSD as: *”information system made up of three interacting constituents: buffer to users (Management Dialogue), Data Management and Model Management”.*
- In 2002, Power defined an SSD as : *”interactive information system meant to help decision makers use data, documents and models in order to identify and solve problems, and make decisions”.*

OLAP is a way to provide answers to complex inquiries of the data bases. OLAP is part of what is called Business Intelligence along with ETL (Extract, Transform and Load), relational reporting and data mining [6].

Since OLAP tools work with multidimensional data models, perform on-spot complex, analytical inquiries and have a high processing speed, certain specialists have suggested that they should rather be called FASMI (Fast Analysis of Shared Multidimensional Information).

Functional requirements of OLAP systems

OLAP systems must meet the following functional requirements [5]:

– *Dynamic analysis of data* – it implies the existence of various analysis tools as well as multiple dimensions, with an emphasis upon the control of an enterprise's data models. The dynamic analysis of data helps better understand the causes of changes within an organization and it is also used to find solutions.

– *Quick access to data* – OLAP applications use a large amount of data and imply very quick access.
– *Multiple data sources* – most OLAP applications access data in several sources, including external ones and various applications performed in different programmes.
– *Data sources' synchronization* – if the data in an OLAP application come from several data bases, it is possible they are modified in various cycles.
– *Historic analysis* – almost all OLAP applications use time as a dimension and therefore the useful results come from time series analyses. In order for that to happen, it is necessary the data stored in data warehouses or data marts should last at least two, three years.
– *High generalization level* – the decision makers in an organization require the information to be grouped, aggregated and shown as synthetically as possible. As a rule, in order to ensure high efficiency and quick access to data, the latter are merged and aggregated at a high level and decision makers can also view detail levels if they request it.

Erik Thomson divided OLAP systems' functional requirements into two large categories: *logical and physical* [7].

Logical requirements ensure the possibility to process dimension data, structure data and render systems flexible. They are the following:

– Complete structuring of dimensions by ranking – an OLAP system should model the dimensions in an organizational environment according to certain hierarchies and at various levels, starting from the most detailed to the highest, generalized level.
– Efficient calculations and processing – OLAP systems should ensure the accomplishment of complex analyses, data comparison and prediction opportunities.
– Flexibility – the data obtained after processing should be shown according to users. OLAP systems should be flexible in all ways. Flexibility allows a model to be changed by a user without changing (redesigning) the entire system.
– Independence of presentations to a model's structure – an OLAP system should ensure the opportunity of not affecting data structure if a presentation changes.

Physical requirements refer to the access and response time of the system as well as to the multi-user support:

– *Quick and direct access* – OLAP systems should provide support for on-spot analytical requests for large data amounts. OLAP Council believes that the main goal of OLAP systems is to *"provide response duration of five seconds or less, irrespective of the request type or data base size, within a multi-user and shared environment"*.
– *Multi-user support* – due to the fact that the data stored in a data warehouse are accessed by several users, OLAP systems should ensure competitive, shared access to analytical processing.

E.F.Codd used the OLAP notion for the first time in 1993, in an article called "*Providing OLAP (On-Line Analytical Processing) to User-Analysts: An IT Mandate*", and he was subsequently named the "father of OLAP concept". He identified twelve rules that OLAP systems should meet.

E. F. Codd outlined the processing difference between relational and multidimensional models, stating that *"irrespective of how powerful relational systems might be to users, the former are not designed to provide strong data synthesis, analysis and enhancement* *functions, collectively known as the multidimensional analysis of data”.*

Later on, in 1995, the twelve rules set up by E. F. Codd were improved and there were eighteen rules, shown below:

A. Basic features

Rule 1: A multidimensional conceptual image

Codd believed that a user’s image of an organization is multidimensional and therefore the conceptual image of OLAP models should be multidimensional, too, and should rely upon the image or model existing within an organization.

Rule 2: Intuitive control of data

OLAP systems should allow data control operations (drill down, drill up, drill across operations), as many OLAP tools permit the intuitive control of data (for example, Microsoft OLAP, Express etc.).

Rule 3: Accessibility

OLAP systems should provide a single logical image of the data in an organization. Within an OLAP model, data sources should be transparent to users. E. F. Codd believed that even users can be a source of data.

Rule 4: Varied data sources

An OLAP system should have the ability to work with data stored in multidimensional data warehouses (MOLAP) or in relational data warehouses (HOLAP).

Rule 5: OLAP analysis models

There are four analysis models that OLAP tools should hold: explanatory, direct, contemplative and forming. Therefore, OLAP tools should at least ensure the drafting of standardized reports, analyses such as *”what happens if...?”,* drill-down/roll-up and slice/dice operations.

Rule 6: Client/server structure

An OLAP system should provide the client/server structure, allowing the access of users by means of a client and aiming at performing multidimensional processing by means of a specialized server.

Rule 7: Transparency

OLAP systems should permit the access to heterogeneous data sources which must be transparent to users and include buffers to various client-tools, such as: table calculation sheet tools, text editors etc.

Rule 8: Multi-user support

OLAP systems should provide competitive, shared access to data, at the same time making sure the data are true and safe.

B. Special features

Rule 9: Data denormalization

When data are processed within an OLAP environment, the external data used as sources should not be affected. OLAP tools process large data collections upgraded periodically and in order to do that, they should be able to persistently connect to external data sources, in order to ensure the synchronization between external sources and data cubes. OLAP systems are generally separated from source systems and that is why the connections serve as transformation functions which indicate the way table data or table calculation sheet data are turned into multidimensional data.

Rule 10: Storage of results generated by OLAP systems

The data subjected to analysis have to be stored and processed separately from relational sources or the folders they come from. This condition should be met as a consequence of the differences between operational data and the ones meant for decision-making support.

Rule 11: Control of missing values

The *”spreading”* term has been used with the meaning of missing value, inapplicable value and zero value. The first two are invalid data (nule data). The third, where spreading means the existence of several zero values, is a special case of the way in which a large number of repeated values are stored, zero value here. Yet, zero value is as valid as any other number. The confusion emerged because there are many zero values in OLAP applications, where there are also a lot of missing and non-valid data. The techniques for the physical improvement of many repeated values' storage are similar and sometimes identical with the ones used to physically improve the storage of numerous missing and non-valid data. However, missing and non-valid values are not valid data. They cannot be treated the same as any other value. Therefore, special techniques are necessary in such circumstances [8].

Rule 12: The way to treat missing values

To any analysis of any data series, either multidimensional or not, the accuracy of calculations is a major requirement. Treating spread data is very important and often debated upon in the field of data bases. The two types of data (missing and non-valid) should be however treated individually, as they affect calculations in different ways [8].

C. The way to show data

Rule 13: Flexibility of reports

The data which are subject to analysis should be accessibly shown to users, so that they can arrange their data according to various dimensions on available axes.

Rule 14: Performance of reporting

The performance of reporting should not be influenced by data's dimension or organization manner. However, there are two factors which affect the performance of reporting: the way calculations are made (they can be calculated before or during an inquiry) and the place where calculations are made (client/server). It can be stated that the importance of these factors is higher than the dimension of a data base or the number of dimensions.

Rule 15: Automatic adjustment of physical level

OLAP systems should automatically change the physical structure of data bases, according to the type of logical model and to the data amount.

D. Control of dimensions

Rule 16: Generic dimensionality

Dimensions should be structurally and operationally even. It implies that dimensions should allow multiple ranking as well as all types of multidimensional operations (adding/deleting a member, adding/deleting a hierarchy, changing a member or hierarchy etc.).

Rule 17: Unlimited aggregation dimensions and levels

Codd thinks that a maximum number of 15-20 dimensions should be used. In practice, there are many other requirements and limitations of OLAP tools, so that the issue of the maximum dimension number can become a minor, unsignificant requirement [5].

Rule 18: Operations between non-restrictive operations

OLAP systems, by means of the control language, should allow operations among various dimensions, without any restriction.

OLAP cube

OLAP technology uses data structures called cubes which are organized in multidimensional data bases. The process of defining the structure of cubes is called multidimensional modelling, the same as the design of a data warehouse structure. Since an OLAP cube is used to extract the information needed in the decision making within a data warehouse, a cube's constituents are similar to a data warehouses's.

Multidimensional data warehouses are optimized data structures used to exploit the data stored in data warehouses and for OLAP analyses.

One can say that multidimensional modelling is the core of OLAP technology. Multidimensional modelling helps one show the results of certain economic activities closely related to one or several factors that have been part in their make-up.

Specialized literature includes several definitions of the ”cube” concept of which one is that of SQL Server Book Online, which states that a cube is a

”multidimensional data structure that contains dimensions and measures”.

A cube can be seen as a data sub-set within a data warehouse where the existing data have multidimensional structures.

Aggregated and hierarchical data are stored within a cube. For instance, if the sales of a specific product are inquired during a certain year term, an OLAP cube helps classify the sales of the respective product by months, weeks and days.

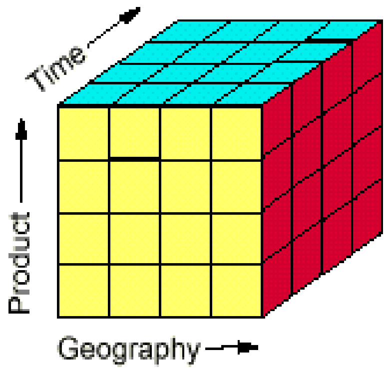

Fig. 1: OLAP Cub

A cube has the following constitutive elements:

- Dimensions are the features according to which the analysis of an activity’s measures is made [9]. Dimensions are the main elements used to define a cube. Dimensions are those which allow data to be viewed and they may contain a single data viewing criterion. Practically, dimensions are those which make an ”in-depth” analysis (drill-down) of data, starting from a lower detail level and gradually reaching more and more detailed levels. For example, time, or geographic dimension (such a dimension helps describe the organizational structures of a company from the perspective of its placement).

Dimensions are twofold: private (Private Dimensions) and public (Shared Dimensions). Shared dimensions are made independently from any cube whereas private dimensions are created along with a decisional cube and they are saved within the respective cube’s warehouse. One can choose private dimensions when a dimension implements the logical viewing of data which is available just for the respective cube. In all the other circumstances, it is recommended that shared dimensions should be defined by means of which the data can be viewed uniformly.

A dimension is characterized by the following elements:

– *Dimension name;*

– *Dimension structure* (also known as attributes): it is the set of elements that characterize a dimension;

– *Dimension members:* the values of dimension attributes;

– *Dimension levels:* a level emerges as a consequence of aggregating the members of a dimension that have a feature in common;

– *Dimension ranking:* the tree structure of father-child levels of the respective dimension.

- Measures are the elements that help describe the indicators meant for analysis (for example, the value of a company’s expenses, the value of a company’s loans etc.). When the measures of a cube are defined, it is necessary to take account of the following [10]:

– Measures should reflect numeric values;

– Measures should be stored in the cube’s fact values;

– Measures should be associated to each level in each dimension.

It is compulsory that for each fact table within a data warehouse at least one measure should exist.

- Facts
- Fact tables
- Dimensional tables

From the architectural perspective, cubes can be organized in two ways:

a. *Star:* just like a data warehouse, a star-shaped cube has all its dimensional tables directly connected with the fact

table. In most cases, the starlike
b. is more frequently used as it has advantages such as: it is more easily maintained and provides higher performance at the same time.
c. *Snowflake:* within such a scheme, a dimension is represented by several dimensional tables. Unlike the star structure, the snowflake one has a more readable dimensional model when the amount of data is too large.

Architecture of OLAP systems

OLAP systems have a basic architecture structured in three main constituents:

a. Data base – is the data source used for an OLAP analysis. A relational data base may be used as a data base to ensure facilities for multidimensional storage, structure multidimensional data base, data warehouse etc.

b. OLAP server – is the one managing the multidimensional structure of data, simultaneously accomplishing the connection between a data base and an OLAP client.

c. OLAP client – is represented by the applications which ensure the exploitation of data and are also a support for generating results (graphs, reports etc.).

The following graph shows the architecture of OLAP systems which varies according to the way of data storage and to the type of their processing, yet seen in broad terms, three data levels can be identified: the level of data sources, the level of OLAP server and the level of data show or of the buffer to users.

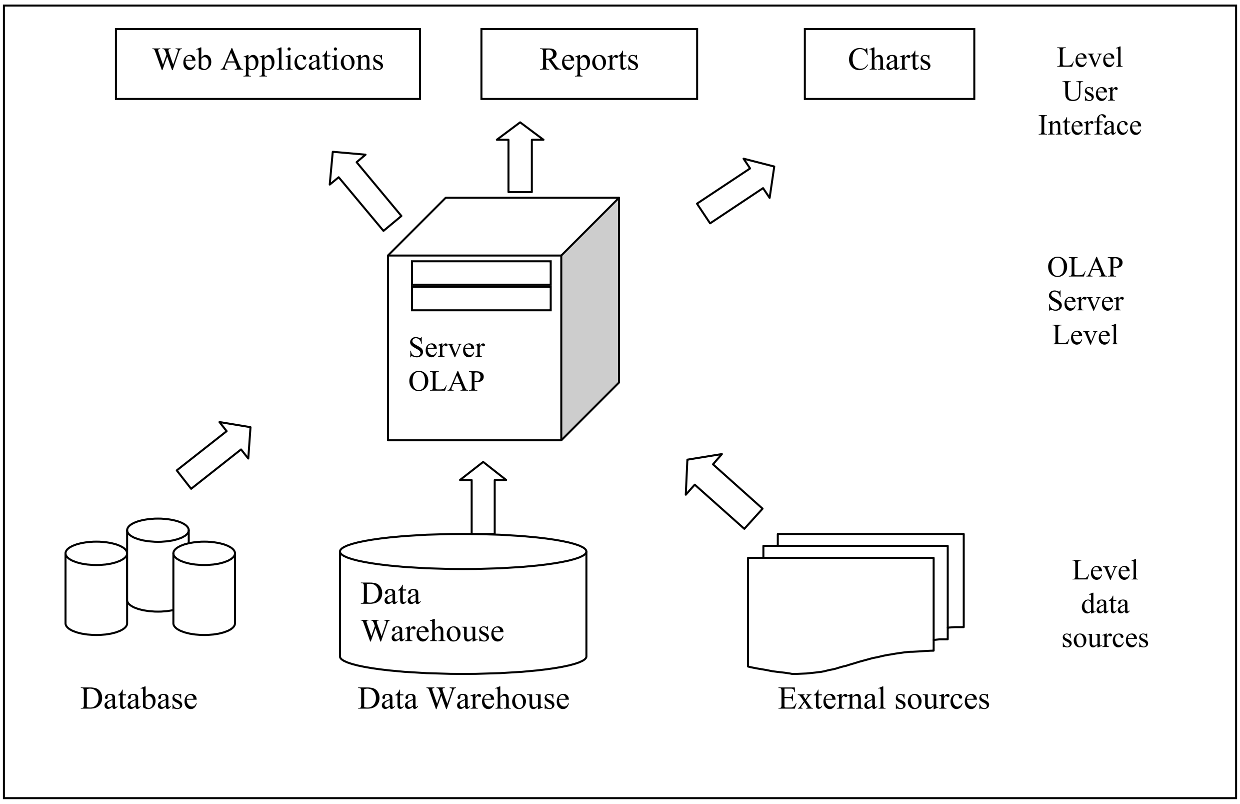

Fig 2: OLAP Architecture

According to the type of data base used for analyses, other versions of OLAP systems have emerged, namely:

– MOLAP systems
– HOLAO systems
– ROLAP systems

MOLAP systems (Multidimensional OLAP) – are seen as traditional solutions for multidimensional analyses. They store both basic and aggregated data in a multidimensional data base which is called a cube and they are used as efficient tools for the operations during analyses as well as during complex calculations. Analyzing the space they cover on a disk along with the time they take in order to reply to complex inquiries, MOLAP cubes can be said to be the best performing.

Among the advantages of MOLAP systems are the ones below:

- Relational tables are not suited to multidimensional data;
- Multidimensional matrices allow the efficient storage of multidimensional data;
- SQL language is not adequate to multidimensional operations.

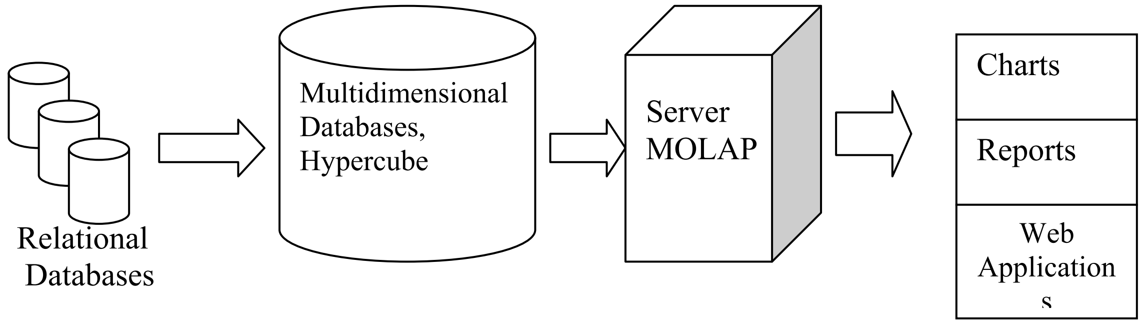

Fig 3: MOLAP Architecture

ROLAP (Relational OLAP) – here, the organization of the data meant for analysis is done in the form of a relational data base. Storing the basic and also the aggregated data is done within a relational data base and memorizing the dimensions' scheme is done within the ROLAP cube.

ROLAP helps exploit large amounts of data and sometimes data base management systems provide facilities typical of ROLAP. As a comparison between ROLAP and MOLAP, it can be said that ROLAP systems are much slower. Practically, a disadvantage of ROLAP systems is the time they take when responding to informational inquiries which can sometimes be quite long when it is about complex inquiries with multiple data sources.

***HOLAP (Hybride OLAP)* –** can be called a compromise variant between the first two versions, which tries to combine the advantages provided by MOLAP and ROLAP, in order to ensure that users receive the best solution from the performance perspective. As to HOLAP, basic data are memorized in relational data bases whereas aggregated data bases are memorized in the HOLAP cube.

Relationship between OLAP and Data Warehouse

OLAP systems have been included in data-oriented decision support sytems, although they are hybrid decision support sytems, because they use simple analytical techniques (data's multidimensional analysis) in the analysis of large amounts of data. Most specialists in the field believe that data warehouses and OLAP tools provide the necessary support to turn companies' large amounts of data into information which is useful to decision makers.

A data warehouse mostly relies on the processes that help ensure the substantiality, truthfulness and validity of data, whereas OLAP systems rely on analytical requirements, modelling processes and necessary calculations. Bill Inmon, who was named "the father of data warehouse concept", emphasized the idea that such a data warehouse mainly aims at ensuring the substantiality and truthfulness of data used.

At present, due to the harsh competitions on the market, decision makers need information that should be provided constantly, rapidly and easily, information which can be given by the analytical techniques supplied by OLAP tools, data warehouses and Web facilities, too.

To ensure fast access to information necessary for managers, analysts in an interactive and flexible OLAP instrumetele use. We can say that completes OLAP and Data Warehouses are mutually OLAP making huge amounts of data stored in data warehouses, but the information necessary and useful to decision makers.

Several technologies are used to analyze the financial-accounting data stored in a data warehouse, among which the most frequent is OLAP technology (from the perspective of actual use and of the existing software support).

OLAP technology helps use the financial-accounting data stored in a data warehouse in an efficient way during on-line data analyses, supplying a fast reply to complex inquiries. OLAP multidimensional model along with specific aggregation techniques ensure the organization of large data series which allow an easy and prompt interpretation. OLAP provides data analysts with the necessary flexibility and work speed when underlying decisions in real time.

The connection between OLAP technology and data warehouses can be defined as follows: ”whereas data warehouses furnish data management, OLAP implements technologies that turn the data into strategic information” [9].

Using OLAP technology helps data warehouse users by several advantages, of which [11]:

- quick execution of inquiries – as it gives one the possibility to keep within a cube some values calculated before an inquiry, such as aggregated values;
- meta-data-based inquiries – an example is MDX Language which ensures the native generation of search criteria for inquiry results;
- calculation formulae similar to the ones used in table calculation applications – one of the advantages of a table processor is the fact that its users can generate formulae by the addresses of the cells where the necessary calculation values are stored. The use of such a formula is easy within the OLAP environment, too, since the address of any cube cell can be used in formulae.

The importance of a warehouse is not the amount of data, but the quality of stored information. Data warehouses were named *”data jailhouses”* by Aaron Zornes, a famous analyst.

Conclusions

Making the best decisions either helping to solve certain problems occurred in an organization’s economic-financial activities or to ensure an organization’s good running, depends on the quantity and quality of information provided by an economic information system.

Managers have always needed correct and rapid information in order to make the best decisions. Data have always been at hand, yet there has been an impediment when extracting and processing them which have been quite slow. Thus, in order to get rid of such an impediment, data have been organized into multidimensional cubes which have been easy to ”rotate” and ”slice”.

OLAP systems are helpful to managers who, after the process of data analyses, can make the best decisions related to the good running of their companies. OLAP systems help make predictions of revenues and expenses, analyses of sales, and identify the progress of financial indicators. OLAP technology helps managers have their own perspectives upon data in general and financial-accounting data in particular.

References

1. C. Crecană, *Profitability of Small and Medium Entreprises,* Editura Economică, Bucharest, 2000.
2. F.G. Filip, *Computer Aided Decision: Decision makers, basic methods and tools associeted*, Editura Tehnică, Bucharest, 2005.
3. M. Velicanu, M. Munteanu, *Characterization of Current OLAP Systems*, Economic Informatics Review, no.3(19)/2001.
4. The OLAP Council Definitions, January 1995, www.olapcouncil.org
5. I. Lungu, A. Bâra, *Executive Information Systems,* Editura ASE, Bucharest, 2007
6. www.marketwach.ro/artico/816/O_abordare_practică_a_bazelor_de_date_OLAP
7. E. Thomson, *OLAP Solitions: Building Multidimensional Information Systems*, John Wiley&Sons, New York, 2002, second edition
8. M. Muntean, *Introduction to OLAP Technology: Theory and Practice,* Editura ASE, Bucharest, 2004
9. Gh. Popa and others, *Microsoft SQL Sever,* Editura Economică, Bucharest*, 2006*
10. Appdev Products Company LLC, SQL Server 2000: OLAP Cubes and Queries Profesional Skills Development
11. R. Jacobson and others, *Microsoft SQL Server 2005, Analysis Services Step by Step*, Microsoft Press, 2006

Codreanu DianaElena is a graduate of ”Constantin Brâncoveanu” University of Piteşti, Faculty of Management Marketing in Economic Affairs of Rm. Valcea, class 2004. At present, she is a Ph.D. candidate at the Academy of Economic Sciences in Bucharest, specialization in Economic Informatics.