Sergii V. KOSTISHYN[1], Sergii V. TYMCHYK[2], Myhaylo V. BACHYNSKYI[2], Irina V. FEDOSOVA[3], Aynur KAZBEKOVA[4], Wojciech SURTEL[5]

Vinnytsa National Technical University(1), Ternopil Ivan Puluj National Technical University (2), Pryazovskyi State Technical University (3), (Kazakh National Research Technical University after K. I. Satpaev (4), Lublin University of Technology (5)

doi:10.15199/48.2017.05.22

Ways and possibilities of creating medical information systems based on OLAP-technology

Abstract. *This article discusses the principles of creating medical information systems based on the technology of multi-dimensional data sets OLAP. The article presents the essence of this technology and method for use it to build data management technology in medical information systems, as well as the principles of data warehousing.*

Streszczenie. *W artykule omówiono zasady tworzenia medycznych systemów informatycznych opartych na technologii wielowymiarowych zbiorów danych OLAP. Artykuł prezentuje cechy tej technologii i metody wykorzystania jej do budowy systemu zarządzania danymi w medycznych systemach informatycznych, a także zasady magazynowania danych. (**Sposoby i możliwości tworzenia medycznych systemów informatycznych w oparciu o technologię OLAP).***

Keywords: OLAP, medical information systems, databases.
Słowa kluczowe: OLAP, medyczne systemy informatyczne, bazy danych.

Formulation of the problem

Information technology today is one of the most rapidly evolving branches of industries. Daily, new software products are created, old technology are improved, their algorithms and interfaces are changed. Information technology covers more sectors of the economy and provide automation in the areas that were previously inaccessible to this.

One of the vectors of development of information technology is the scope of medicine. Today we are witnessing a total computerization of this area, the introduction of medical information systems in the activities of medical, preventive and rehabilitative health care institutions. The specificity of the healthcare industry in the context of medical information systems (MIS) is as follows:

1) the high cost of error (sometimes, the health or the patient's life);
2) moderate speed (MIS does not require absolute performance);
3) switching between different types of tasks;
4) high value of the reporting;
5) specifics of work in conditions of uncertainty (unknown number of accompanying diagnoses).

All this imposes certain requirements that must be integrated into MIS during its design and development. In turn, all of these requirements affect storage methods, manipulating, analyzing and processing large amounts of information, which are implemented with the help of MIS and its subsystems.

Medical information system, like any other, designed primarily for complex multidimensional data analysis, which creates the necessary conditions for the physician acceptance solutions. In turn, the decision-making system has the properties, which provides to submission of aggregate data for user for various samples. Typically, these aggregate functions are multidimensional (and therefore non-relational) data set and the name of work technology with them is OLAP (On-line Analitical Processing).

Analysis of the literature

For the first time OLAP concept was described in 1993 by Edgar Codd [1], and in 1995 on the basis of the requirements which it had formed, FASMI test was built (Fast Analysis of Shared Multidimensional Information), which elaborated these requirements in relation to the multidimensional data analysis [2].

The OLAP-technologies based on the idea of multi-dimensional data model. Humans thinking is multidimensional by definition. When a person asks questions, it imposes restrictions, thereby forming questions in many dimensions. Therefore, the process of analysis in a multidimensional model is very close to the reality of human thought [2].

There are two main types of analytical data using the OLAP. This MOLAP and ROLAP: MOLAP (multidimensional OLAP), which based on using non-relational database, providing a multi-dimensional storing, processing and presentation of data. Products which related to this class usually have a multi-dimensional database server and information are chosen only from the multi-dimensional structure in the process of analysis.

The multi-dimensional data structure in ROLAP (relational OLAP) is implemented with relational tables.

Naturally, each of the above OLAP has its advantages and disadvantages. For MOLAP is a high capacity, but low efficiency in the processing of large data objects, and for ROLAP – the efficiency of large amounts of data, but low speed when working with multi-dimensional data. NOLAP (Hybrid OLAP) has been developed for overcome the these disadvantages. It provides online analytical processing using both approaches – relational and multi-dimensional [3].

Today, the OLAP-technology intensive introduce into different areas of application, ranging from the analysis of multidimensional data to financial planning and financial consolidation [4]. Until now, unfortunately, there is no terminological data standards, data transmission standards, and the query language of forming cubes. The integration of different sources for the analysis contributes to the novel use of analytical technology.

Main text

To support the operation of the information providing of treatment and rehabilitation process, as a rule, are used specialized databases (DB). MIS database is designed to save a large array of medical information that appears as a result of working software modules of the MIS, according to its functional organization.

DB provides a data repository for all information-intensive modules of MIS. This modules activity is the result

of the accumulation of certain specific information for specific tasks of treatment.

The vast majority of modern MIS built in "client-server" architecture. The practical experience proved the inevitability of such a solution for the creation of an integrated MIS. In addition, much of the existing requirements for the MIS is already implemented in the professional DBMS, which can significantly reduce the time to create a private system [5].

At the same time, storing of information must be associated with not only the reliability but also the convenience of its use. This creates new requirements for structuring the information in the database and processes of its retrieve and display.

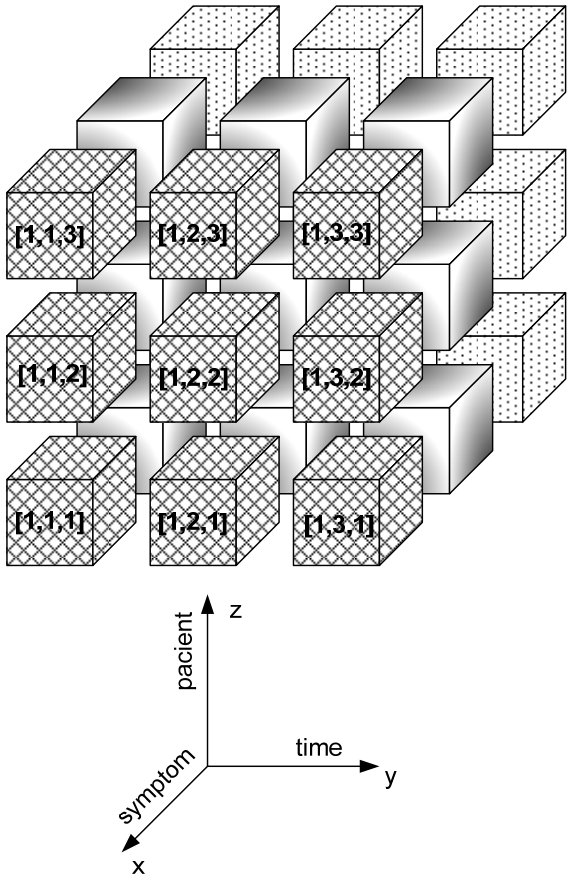

Fig.1. A simple example of OLAP cube for the problems of medical information system

In our opinion, another area of application OLAP-technologies can be medicine, namely the establishment of medical information systems, working with:

a) the multidimensional medical data;
b) the databases of these data, which must have a high performance;
c) the decision support systems by the physician (often under severe time constraints and the information necessary to make the only correct decision).

In this connection, in further consideration of the material, we will use the following definition OLAP [6]: ”OLAP-technology data processing is to produce total (aggregated) information based on large amounts of data, structured according to the principle of multi-dimensional”.

It is necessary to assume that the first step should be a projection of industrial and commercial using of OLAP-technology in the field of medical applications. Thus, the possibility of using OLAP-technology in the design process MIS can be implemented to create information software such systems, one of the main components of which are medical databases.

The next step can be considered an analytical review of the classification and the formation of OLAP-technologies to relation with medical problems. In the future, it is necessary to theoretically justify the use of OLAP-technologies for working with multi-dimensional medical data, to evaluate the feasibility of the development of such technologies, software and information packages.

The projection of industrial applications for medicine does not completely solve the problem of creating OLAP-medical systems and technologies. Therefore, it is necessary to adapt the existing mechanisms in medicine, principles and approaches to OLAP-technology structure to achieve a positive result. It is primarily a question of:

a) the concept and the mechanism of functions realization ”support the adoption of a doctor solutions”;
b) the features of the ”friendly smart interface” construction;
c) ensure the development of the psychological compatibility concept of triad ”the patient – the OLAP database – the doctor”.

As already mentioned [2, 3, 4], OLAP is a data repository, which requires a description of the condition and its specific application in the composition of information support of the MIS, the formation of new approaches and objectives, simulation medical databases, which based on OLAP features.

At the same time it is necessary to clarify their functions and conditions of application, etc. [7]. Data repository (DR) – a special kind of software solutions designed for collect, integrate and analyze data. They are based on principles other than taken in OLAP-systems that can meet the challenges associated with data analysis using DR is more efficient than with other systems. This division provides accounting and analytical information; integration of data from disparate sources; availability of funds in the DR cleaning of the incoming information.

Since the data repository are the basis for the construction of decision support systems, there is a need for concrete definition of the they characteristics [8,9]: objective oriented; integration; binding at the time; immutability.

The main components of the data repository are:

1. Operational data sources.
2. Means the design / development.
3. Transfer and transformation tools.
4. DBMS.
5. Means of access and data analysis.
6. Administration tools.

Today, there are two main approaches to the storage architecture: Corporate Information Factory (CIF) by Bill Inmon data warehouse and data repository with bus-architecture (BUS) by Ralph Kimball [9].

The special features of the architecture of data repository by Bill Inmon can include such characteristics [8]:

1. Using the relational model of atomic data organization and space – for the organization of the total data.
2. Using the ”iterative” or ”spiral” approaches to building of large data repository, i.e. the process of storage construction is not once. This allows, if necessary, makes changes in small blocks of data or program code and eliminating the need programming the large amounts of data in the repository. The same can be said about the potential errors that will be localized within a relatively small cell.
3. Use of the third normal form for the organization of atomic data, which provides a high degree of integrated data design.

4. DR is a project of corporate scale.
5. DR is not a mechanical collection of data marts, and the physical integrity of the object [8].

At the same time, the distinctive features of approach by R. Kimball is as follows [8]:

1. Using the three-dimensional model data organization architecture, named ”Star” (star scheme).
2. Using the two-tier architecture that includes data preparation stage, inaccessible for end users, and the data repository bus architecture (it does not contain a unified physical integrity or centralized data repository).
3. DR with bus architecture has the following characteristics: it is spatial; it includes a data transaction and summary data; it includes data marts, which dedicated to only one subject area or having only one fact table; it may comprise a plurality of data windows within a single database.
4. DR is not a single physical repository (unlike approach by B. Inmon). This ”virtual” repository, a collection of data marts, each of them has a ”star” type architecture.

The optimal structure of a database must be provided on the stage of preliminary design by means of clear execution of optimization rules. The main purpose of the database design is a reduction of stored data redundancy, consequently, saving memory usage, reducing the cost of multiple redundant copies of the update operation and eliminating the possibility of conflicts due to storage in different parts the disputed information.

When designing a database that must contain a significant amount of information for the tasks of OLAP-analyses, two problems are solved:

1. Conversion of domain objects to the formal data model objects, considering a typology of selection tables and encoding information database fields.
2. Enforcement of database queries by placing the data in a structured form in the memory.

Design of databases is part of the software development process, and consists of two levels: logical and physical design.

The functionality of the entire software system will largely depend precisely on this organization databases. Proper design and implementation of database provides a more optimal performance on a range of parameters such as ease of implementation of the code, reliability, speed and economy of computer resources.

To implement the optimal structure of the database, which operates within the MIS with integrating OLAP-technology, the following principles have been formulated:

1. The principle of "reasonable redundancy" implies a certain level of duplication of data that represent the greatest importance for the therapeutic and rehabilitation process.
2. The principle of "fixing the facts" – facts of the process of registration of data modifications in the database based on the login, the time and the content of the modification.
3. The principle of "limited autonomy" – the accumulation of data in the package during the absence of communication with the server and the sending this package when the connection appears.
4. The principle of "typing of project" – definition of the limited typology of attributes in the database design stage.
5. The principle of "facilitation ties" – reducing the number of connections between the entities and their replacement by SQL-queries, which have a higher speed and leads to a simplification of the database structure.
6. The principle of "personalization" – an introduction to database structures for the distribution of data areas to which access is possible from a certain user groups (the patient has the right to access information only of himself, the doctor has access to any information, e. t.);
7. The principle of "the territorial delineation" – all information is divided into 2 areas – local and global. The first is on the user's local machine and use them in only their work, the second – important for the entire system and is located on the server machine. This principle is introduced to conserve network resources.

These principles are related to, first of all, the specifics of medical information systems, and take into account the work with data that are vital information.

An important issue is to protect against unauthorized access and errors that could lead to distortion or their loss of integrity; the impact of unpredictable risks. The solution to these problems is to develop new principles of organization of MIS based on OLAP-technology, new models and analytical data processing functions and feature of the language syntax in the implementation of the OLAP-principles. It is necessary to provide the sufficient level of health information reliability. This can be realized through the creation of algorithmic and software for minimize errors MIS, the ranking of information and analytical functions of the MIS on the access level, perfection technological elements and approaches in the design of information systems [10].

The principle of ROLAP-clients is preliminary description of the semantic layer, which hides the physical structure of the source data [11]. This data source can be: local tables, relational databases, and other structures. User can manipulate them for clear targets in terms of subject area to create cubes and analytical interfaces.

The working principle of the OLAP server-client is other. It consists in the fact that when you create cubes user manipulates the physical description of the database.

The formation principles of the OLAP clients allow to propose them for working with health data. It will provide the intelligent interaction during the construction of medical complexes, form a list of possible operations on the OLAP, propose the data processing algorithms and the formation of medical and statistical data using OLAP model [12].

Today, there are problems which are associated with the implementation of the OLAP technology elements in the structure of MIS (such as automation equipped working place (AEWP) of medical specialists). This is due to the lack of scientific publications that are related to non-industrial and non-commercial using OLAP-technology. The closest example is the AIS "Electronic deanery ED ++" [13], which is a hardware and software, organizational and administrative system of data collection and processing system. This system is associated with the learning process and operates in real time.

It is noted that new OLAP technology and "data warehouse" received application in the system of "ED ++". This is gives managers the opportunity to own, without the help of programmers, to obtain analytical data. At the same time they are using easy and intuitive interface.

For AEWP, OLAP-technology of data analysis is especially interesting because it allows to quickly get the information you need in any dimension of the cube and construct different reports. Thus, functional tasks of AEWP from the point of view of OLAP can be formulated as follows:

1. Organization of work with the client database. This task includes working with a database that is in the making, editing and saving relevant medical information relating

to the subject area specialist doctor or receptionist, as well as the formation of the corresponding statistical reports based on previously entered data.
2. Centralization of medical background information. This task can be attributed to the integration of AEWP reference information used by the doctor in his medical practice and is intended to build support mechanism of decision-making of this problem.

Developed in the article principles of database construction for the MIS based on the use of OLAP technology correspond to the basic principles, which ensuring its integrity and compatibility with the other components of the system. In short, they can be summarized as follows:

1. "Openness and accessibility". Database should provide access to information from a variety of software tools, without the use of extraneous export and import tools.
2. "Compatibility". Information, which stored in the database, should provide all the necessary forms for reports.
3. "Continuity". The database provides the use of software tools, which use own directories and data formats (sometimes inconsistent and heterogeneous on execution).
4. "The optimal structure". The database should be fully excluded duplication of information, and its redundancy must be reduced to the minimum necessary.
5. "Extensibility". With the development of MIS and the inclusion of new types of AEWP, the database should allow to store any kind of information.

This will provide the necessary flexibility to the database operation and flexibility throughout the MIS.

Thus, OLAP-technologies and specialized DB in medicine are relevant and in demand. The results of the work of such information products can be widely used in all areas of medical information systems. Focus on tools for analysis of large amounts of data increases the value of OLAP-technology, especially for tasks analytics, forecasting, and reporting.

Conclusions

Application of OLAP-technologies for creation of information support MIS requires the development of:

- new principles and approaches to the evaluation of the implementation effectiveness;
- new criteria for the evaluation of information security, which is built on this new principles;
- methodology of comparative analysis of medical information systems.

***Authors**: PhD Sergii V. Kostishyn, Vinnytsa National Technical University, 95 Khmelnytske shose, Vinnytsia, Ukraine, 2102, PhD Sergii V. Tymchyk, Ternopil Ivan Puluj National Technical University, Ruska str. 56, Ternopil, Ukraine, 46001, Prof. Myhaylo V. Bachynskyi, Ternopil Ivan Puluj National Technical University, Ruska str. 56, Ternopil, Ukraine, 46001, PhD Irina V. Fedosova, Pryazovskyi State Technical University, Universytets'ka 7, Mariupol, Ukraine, 87500, Aynur Kazbekova, Kazakh National Research Technical University after K. I. Satpaev, Satpayev Street 22, 050013, Almaty, Kazakhstan; PhD Wojciech Surtel, Lublin University of Technology, Institute of Electronics and Information Technology, Nadbystrzycka 38a, 20-618 Lublin, E-mail: w.surtel@pollub.pl*

REFERENCES

[1] Codd E.F., Codd B.S., Salley C.T., Providing OLAP to User-Analysts: An IT Mandate, *Ann Arbor, Mich.: Codd amp; Associates,* (1993).
[2] Sergeeva I.G., OLAP-technology, *Information technology*, (2016), Access to the page.: http://www.sergeeva-i.narod.ru/inform/page7.htm
[3] OLAP-technology, *KAI Development*, (2012), Access to the page: http://kaidev.ru/Pages/Article.aspx?p=OlapAbout
[4] Kudrjavcev Ju., OLAP-technology: overview of tasks and research, *Biznes-informatika*, (2008), n. 1, 66–70
[5] Gusev S.D., Technological approaches and principles of information systems, *Problems of the development and implementation of information systems in health and CHI: Proceedings of the Inter-Regional Conference, Krasnoyarsk,* (2000), 200–218.
[6] IT Definitions and Glossary, *WebCite*, (2012), Access to the webpage: http://www.webcitation.org/65AkPdIv3.
[7] Pozharnenkov C., OLAP-technology and data repository, *"Connect" Connect World*¸ (2003). n.7, 47–51.
[8] Technology of date repository, *InterSoft Lab*, (2011), Access to the page: http://www.olap.ru/desc/microsoft/storage.asp.
[9] Inmon W.H., Building the Data WareHouse, 4th eddition *Wiley*, (2005).
[10] Mashkov V., Smolarz A., Lytvynenko V., Gromaszek K., The problem of system fault-tolerance*, Informatyka Automatyka Pomiary w Gospodarce i Ochronie Środowiska,* (2014), n. 4, 41–44.
[11] The main approaches to data architecture, *InterSoft Lab*, (2005), Access to the page: Http://www.iso.ru/print/rus/document6082.phtml.
[12] SMART. Description of work with the OLAP-cube, *ZAO "M-Lodzhiks"*, (2004), Access to the page: http://www.smart.su/Demo/ SMARTDescriptionRus.
[13] Romanov V.P., Kuleshova E.A., Sinelnikov I.B., AIS "Electronic dean's office" and corporate knowledge management problem of the university*, "Information technology in education": conference "ITO–2003"*, (2013), Access to the page: http://ito.edu.ru/2003/IV.html