OLAP CONCEPT IN PUBLIC SERVICE QUALITY IMPROVEMENT

BEČEJSKI-VUJAKLIJA Dragana, (SCG) - SUKNOVIĆ Milija, (SCG) - DELIBAŠIĆ Boris, (SCG) - PANTELIĆ Ognjen, (SCG)

ABSTRACT
In all countries of the world, especially countries in transition, public services have problems with limited financial resources (cities are developing faster than their budget). Because of that, resources should be used very rationally and carefully, in order to solve the most urgent problems and for the well-being of the majority of taxpayers. For solving this kind of problems nowadays various methods and software are available, thus their use demands reliable up to date data. A model of gathering and distributing analytical information, based on OLAP cubes, is suggested in this essay, which can give the public services management the basis for using different decision support systems for work quality improvement and directing resources towards solutions of the key problems (in traffic, communal services, supplying and other public services duties).

KEYWORDS
OLAP, information gathering and distributing model, public services, transition

INTRODUCTION
When developing a business system, different data is being gathered. It is data about the enterprise business process and its relationships with the surrounding. It is quite obvious that the development will be more successful and easier if the gathered data migrates into useful information that can be used as guidelines system development. The business subject (public service) can then be more flexible and reactive to its inner and outer surrounding with accurate data about every client and taxpayer.

Collecting data in an appropriate database (DB), through development phases of the company is characterized by huge amount of data. During time the database becomes a data warehouse (DW) because for strategic decisions making historical data is used so a DW represents a good solution for responding to the decision makers demands. The strategic level of decision making is characterized with unstructured problems so a DW can be used as the basis for business intelligence.

DW, as a modern concept, should therefore be used to group information according to the business functions in the enterprise for better business analysis. OLTP (online transaction processing) is then changing towards OLAP (online analysis processing).

CURRENT SYSTEM ANALYSIS
Most IS in the public service sector were developed during the 80-ies. During the time, the system was upgraded to adapt to new demands. The system has strong OLTP possibilities, but still lacks with OLAP possibilities. The data are not in a form ready for analysis and is not integrated; it can be found on different external and internal sources etc.

The project documentation is in most cases the most significant source of information about the system. The OLPT system documentation holds vital information about the system. This is often neglected by system developers so the new system often has data that does not satisfy the users. When developing a system this is the most important phase, because the purpose and the goal of the developing system are being defined.
The phases of a DW implementation are:
- Current system analysis,
- Selection of interesting data for analysis,
- Cleansing and reducing data,
- Data extraction in a temporally database,
- Choosing fact table, dimension tables and appropriate schemas,
- Choosing of aggregations, percentages of aggregations and storage methods, and
- Forming and using of the cube.

The first phase in DW development is finding false, incomplete and wrong data. It is necessary to improve that data or to delete it. The next phase is finding data that has an inappropriate format. It is necessary to correct the format. The reduction of data is often being data because not all data is important for analysis.

**DATA EXTRACTION IN A TEMPORALY DATABASE**

After reducing and cleansing data, the data is collected and put into a temporally database which will serve as a basis for DW development. (Picture 1.). When the OLTP DB is well designed this step is not necessary

![](image)

**Picture 1. Data transformation package for data transformation**

**CHOOSING FACT AND DIMENSION TABLES**

A fact table should be a table that possesses the most detailed data. Ideally, it should possess many relationships towards dimension tables, much data for analysis and should be thin in order to use small amounts of memory.

After the fact and dimension table selection, it is necessary to choose an appropriate schema. During DW design three types can be used:
- Star schema,
- Snowflake,
- Single table schema.

The criteria for choosing a schema are:
Every DW could theoretically be implemented using each schema, but is in real problem solving done with much attention. When solving complex problems, like public service data analysis, the most appropriate schema is the snowflake schema. An example is shown on Picture 2. The shown schema contains 11 dimension tables. The next step in DW design is measure selection. The measures are calculated cells from other cells in the table. Various arithmetic operations can be used for defining a measure. A measure is a cell that is being used for data analysis.

OLAP solutions use aggregations that store previous calculated data of various user queries. The optimal aggregation percentage choice is also a hard decision for the system designer. A high aggregation percentage improves the performance of user queries but uses far more disc memory.

DATA MINING AND DECISION MAKING

Data from business processes forms an important part of the knowledge of an enterprise, which can be used after an appropriate preparation for management decisions. Therefore a large need for tools exists, which support the evaluation and use of enterprise data.

Data Mining enables automatic evaluation of huge volumes of data using data mining algorithms. It is clear that the algorithms themselves are not sufficient for decision making, but that a holistic process of the knowledge discovery is necessary.
Knowledge discovery in data bases (KDD) has the goal of finding patterns and making them usable. Thus, the use of KDD does not prove as trivially in practice and is knowledge and time-consuming. Despite the multiplicity of available tools it is difficult to select the suitable. The complexity of the KDD - process is so extensive and the available knowledge that it demands is too high, despite the attempts for the simplification, so it is rather a task for experts. Tacit knowledge plays a large role in that process.

The KDD - process can be clarified on the basis of Picture 3. This model gives however only one framework for the execution of the knowledge discovery. The task of the user is to illustrate the individual steps of the general model on concrete problems. In each step thereby many possible solutions are available. There are necessary, for example, several data preprocessing steps, different analysis methods must be combined, and finally from the analysis of existing data new knowledge is generated that can be used for strategic as well as operational problem definitions.

Within the KDD process stages of the data preprocessing and reinforcement can be examined (compare [7], [8] and [2]). Another approach is the question of the systematic operational sequence of the entire KDD process.

**GENERATE AND USING THE CUBE**

The cube generates on client or server. Relevant factors that have influence over selection of storage location of the cube are: *size of cube, number of cube users, performances of client and server computers and system flow capabilities*. The cube created on OLAP server can be used with different client tools.
Work with OLAP cube in Microsoft Excel

Before starting to work in Excel surrounding, some actions had to be preformed. Work tables based upon the analytical database are called Pivot tables. After the database name is selected and connected, Excel sheet should look like on the Picture 4.

In so-called pivot table, there are several fields:

- **Data Items**: used for entering measures of the cube.
- **Row Fields**: used for entering dimensions that have to be rows in desired report. One must take care of entering dimensions sequence in this area. Sorting and presenting in report carry out according to dimensions from left to right. So, in this area, several dimensions can be represented.
- **Column Fields**: used for entering dimensions that have to be columns in desired report. One must take care of entering dimensions sequence in this area. Sorting and presenting in report carry out according to dimensions from up to down. In this area, several dimensions can be represented.
- **Page Fields**: Used for presenting summarized data for one dimension. When dimension is placed on this field, then we can filter. This field is used to get results of specific dimension. In this area, several dimensions can be represented.

It is important that dimensions are used for narrowing the domain of measurement (aggregations) and in that way we can find desired information. OLAP cube is built in purpose to create easy queries and reporting, which in classical databases you would need much more time and far more knowledge to ask a query and generating a report.
CONCLUSION
Through simple demonstration it can be seen that storage data are very inflected solutions for end-users, which through tools that use every day, like excel, can with user define queries explore database more efficiently then other tools from OLTP surrounding.

Major advantage of this approach of information and knowledge reveal in databases is that user doesn’t have to know terms of relation model and complex query language.

This method of data analysis as support for decision making can be applied in every segment of public and local governments. Our experiences on the level of city of Belgrade confirm this, although, we are obliged to note that basic prerequisite for successful implementation is computer literacy of the public government employees.

LITERATURE

CONTACT ADRESS
Prof. Dragana Bečeješki-Vujaklija, Ph.D., (draganab@fon.bg.ac.yu)
Prof. Milija Suknović, Ph.D.,
Ass. Boris Delibašić, M.Sc, (boris@bos.org.yu),
Ass. Ognjen Pantelić
Faculty of Organizational Sciences University of Belgrade,
Jove Ilića 154, 11000 Belgrade, Serbia and Montenegro