Advances in Intelligent Systems Research, volume 174
Proceedings of the 8th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2020)

Multi-Criteria Decision Making Approach to Urban Web Service Databases Evaluation Based on the Analytic Hierarchy Process

Andrey Zemtsov*
Department of Computers and Systems
Volgograd State Technical University
Volgograd, Russian Federation
ecmsys@yandex.ru

Tran Dung Khanh
School of Information Technology and Digital Economics
National Economics University
Hanoi, Vietnam
cybertrine@gmail.com

***Abstract*—The article presents a comparison of modern database management systems for the implementation of urban web services. The work describes a technique that allows selecting the optimal database management system according to a set of criteria, including the implementation of priority projects for the creation of urban web services based on the organization of electronic document management: the Situation Center of E-government, the Open Platform, the Unified Personal Account, a various Geographic Information Systems for territorial planning, et al. As a mathematical apparatus, when analyzing a database management system, an Analytic hierarchy process is used. To solve this problem, we examined the six most popular database management systems that implement the relational data model and have similar functionality.**

Keywords—Analytic hierarchy process, multi-criteria decision making, information technology, web services, software selection, database management system, E-government, electronic document management

I. INTRODUCTION

In recent years, the Russian Federation has seen a transition to the provision of public services in electronic form. On July 1, 2012, all regions and municipalities began a gradual transition to electronic interdepartmental cooperation. And by 2018, in accordance with the Decree of President of the Russian Federation No. 601, 70% of all public services should be provided in electronic form. All over the world, it is government initiatives that are the main engine for the development of informatization in the country.

In this regard, the Ministry of Telecom and Mass Communications of the Russian Federation is conducting systematic work aimed at improving the quality and level of accessibility of state and municipal services in electronic form as well as the possibility of receiving them on a «one-stop-shop» principle.

The «one-stop-shop» principle in the provision of public services cover for the exclusion or the maximum possible restriction of the participation of applicants in the collection process from different authorities. Additionally, this principle also applies to the provision of the whole set of authorities of various documents and certificates confirming the applicant's right to receive public services.

In the course of this work, it is necessary to ensure interdepartmental electronic interaction, increase the openness of government authorities, and citizen participation in government decisions in the framework of the Open Government initiative. In addition, it is necessary to increase the efficiency of spending budget funds on information and communication technologies in government authorities. Particular attention in this process is given to the use of the SaaS (software as a service) model.

The implementation of the «one-stop-shop» principle is ensured through the reorganization of existing processes for the provision of public services, which should be expressed in the execution of administrative procedures for the collection of a full package of documents for applicants based on interagency and interdepartmental interaction between executive bodies of state power, subordinate to them and other authorized organizations among themselves.

We believe that e-government, in general, cannot consist solely of information systems owned by the state. Housing and communal services, payment of duties, penalties, etc. carried out by non-governmental organizations. Without the use of information systems of accrual providers, banks, the functioning of e-government cannot exist fully. In this sense, additional services of other organizations that do not replace, but supplement the capabilities of state systems, should be welcomed in every possible way.

The implementation of the provision of public services on a «one-stop-shop» principle implies the availability of a large volume of various data archives as well as the ability to search efficiently. Therefore, the selection of a database management system that should satisfy not only current but also the requirements for public services provided can be presented to them in the nearest future. For example, the scheme of territorial planning of capital construction of gas supply facilities in the Volgograd region is shown in Figure 1. It should be noted that this is far from the only working urban web service.

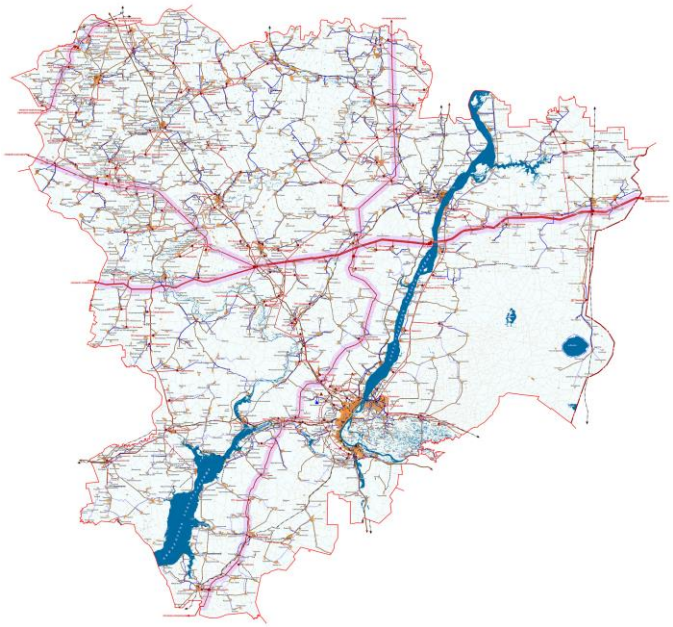

Fig. 1. The scheme of territorial planning of capital construction of gas supply facilities in the Volgograd region.

The above scheme of territorial planning of gas supply facilities is narrow-profile but contains a large number of existing and under construction objects of various types, such as gas pipelines, gas distribution stations, inter-settlement gas pipelines, underground gas storages. All these objects have many parameters for describing them in databases.

II. MULTI-CRITERIA DECISION MAKING

It is usually considered that the chosen or the best is an acceptable solution that most fully satisfies the desires, interests, or goals of the person making the decision. The desire of a decision-maker to achieve a specific goal is often expressed in mathematical terms in the form of maximizing (or minimizing) a certain numerical function defined on some set of feasible solutions X.

However, in more complex situations, one does not have to deal with one, but several functions of this kind [1]. This will happen when the phenomenon, object, or process being studied is examined from different points of view and the corresponding function is used to formalize each point of view. If the phenomenon is studied dynamically in stages, then for the assessment of each stage it is necessary to introduce a separate function [2]. In this case, several objective functions $f_1, f_2, \dots, f_m, m \geq 2$ must also be taken into account.

The above numerical functions form a vector criterion $f = (f_1, f_2, \dots, f_m)$. A vector criterion f takes values in the space R^m of m-dimensional vectors, which is called the criterion space or the rating space, and any $f(x) = \big(f_1(x), f_2(x), \dots, f_m(x)\big) \in R^m$ of the vector criterion f for a certain $x \in X$ is called the vector estimate of a possible solution x. All possible vector estimates form the set of estimating vectors $Y = f(X) = \{y \in R^m | y = f(x), x \in X\}$.

Usually, a one-to-one correspondence can be established between the sets of possible solutions X and the corresponding set of estimating vectors Y, i.e. each possible solution can be associated with a certain estimating vector, and vice versa, with each estimating vector we associate a certain possible solution.

The selection problem, which includes the set of feasible solutions X and the vector criterion f, is usually called the multicriteria problem or the multicriteria optimization problem.

III. ESTABLISH AN ANALYTIC HIERARCHY PROCESS BASED MODEL ON EVALUATION DATABASE MANAGEMENT SYSTEMS

The Analytic hierarchy process [3-5] is designed to solve multicriteria problems with a finite set of estimating vectors. Its application is based on expert information on the relative importance of the criteria in the form of a matrix of pairwise comparisons. The Analytic hierarchy process is used to evaluate the quality of software. The software quality evaluation models and their computing algorithms are presented in [6] and [7].

The Analytic hierarchy process is applied to solve problems that involve complex criteria [8]. To select the optimal database management system for the above-described problem of creating and developing urban web services based on the organization of electronic document management, we define a list of criteria by which we will compare them:

1. The data model. This group includes the used data model and the provided data types.

2. Features of architecture and functionality. This group includes scalability, independence of the environment in which it operates, and network capabilities. By scalability, we mean here the ability of the system to cope with an increase in workload.

3. Performance. This is one of the main criteria for choosing a database management system [9-11]. This group includes the Transaction Processing Performance Council (TPC-C and TPC-H Results List [12]) rating, architecture parallelization potential, and query optimization capabilities.

4. Requirements for the working environment. This group includes the minimum hardware requirements and supported platforms.

5. Features of application development. It is worth considering the possibility of using the Internet environment, multilingual support, and design tools.

6. Reliability is one of the main criteria for choosing a database management system [13, 14]. Reliability has many definitions, which include the safety of information in the event of system failures and the protection of data from unauthorized access.

A comparative analysis based on the above criteria will allow us to rationally choose a database management system for urban web services. The Analytic hierarchy process is based on a pairwise comparison of all database management systems for each of the above criteria, at the output, we get several matrices of pairwise comparisons of alternatives.

Using the Analytic hierarchy process will allow us to structure the problem of the decision in the form of a hierarchy or network. In the simplest form, the hierarchy is built from the top (the objective), through intermediate levels-criteria (the technical and economic parameters) to the lowest level, which is a set of alternatives.

After a hierarchical presentation of the problem, the priorities of the criteria are established and each of the alternatives is evaluated according to the criteria. In the Analytic hierarchy process, the elements of a task are compared in pairs concerning their effect on a characteristic common to them.

The system of pairwise comparisons allows you to get a result that can be represented as an inverse symmetric matrix. The element of the matrix is the intensity of the manifestation of the element of the hierarchy relative to the element of the hierarchy, estimated on an intensity scale. There are reasons for preferring one of the factors, but they are not convincing enough [15].

The following database management systems will be considered as alternatives: DB2 (D_1), Firebird (D_2), MySQL (D_3), Microsoft SQL Server (D_4), Oracle (D_5), and PostgreSQL (D_6). All selected database management systems are suitable for analysis and comparison since they implement a relational data model. Based on the results of the analysis, we will create a pairwise comparison matrix.

For the resulting pairwise comparison judgment matrix, the following indicators are calculated:

1. The priority vector is determined by summing the elements of each row and dividing each sum by the sum of all matrix elements.

2. The main eigenvalue λ_{max} is determined by the sum of the product of the sum of each column by the priority vector;

3. The Consistency Index CI shows the deviation from the consistency and is determined as $CI = \frac{\lambda_{max}-n}{n-1}$, where n is the dimension of the matrix ($n = 6$ in our case).

4. The consistency ratio CR is calculated as $CR = \frac{CI}{RI}$, where RI is the tabular value for a matrix of a given order and in our case is 1.24. The CR value is acceptable if less than or equal to 0.10.

IV. Establish a judgment matrix and measure consistency

To quantify the importance of each level of indicators, we will evaluate indicators based on the famous Saaty fundamental 9-point ratio scale as shown in Table 1 to obtain a judgment matrix for each dimension. The judgment matrices of relevant importance are shown in Tables 2, 4, 6, 7, 8, and 9.

TABLE 1. Ratio scale of the database management systems criteria

Intensity of importance	Definitions	Explanations
1	Equal	Two factors contribute equally to the objective
3	Moderate	Convincing evidence for one factor over another
5	Strong	There are reliable data or logical judgments about the preference of one of the factors
7	Very strong	Convincing evidence for one factor over another
9	Extreme	Evidence of one factor over another is highly convincing
2, 4, 6, 8	Intermediate values	Compromise is needed

W is the priority vector and define the relative weights or importance of each alternative concerning the criterion.

TABLE 2. A judgment matrix of relevant importance under the data model

	D_1	D_2	D_3	D_4	D_5	D_6	W
D_1	1	1	1	1	1	1	0.12
D_2	1	1	2	1/3	4	3	0.22
D_3	1	1/2	1	1/2	4	2	0.18
D_4	1	3	2	1	5	2	0.28
D_5	1	1/4	1/4	1/4	1	1/3	0.06
D_6	1	1/3	1/2	1/2	3	1	0.12

The main eigenvalue $\lambda_{max} = 6.61$, $CI = 1.24$, $CR = 0.09$. The value of the consistency ratio is within the normal range.

To construct a pairwise comparisons matrix according to the features of architecture and functionality, we will perform a comparison of database management systems the taking into account the maximum possible amount of stored data for each of them as shown in Table 3.

TABLE 3. Size limits comparison

Database management system	Database size limit	Table size limit	Table row-size limit
DB2	512 Tb	512 Tb	32677 b
Firebird	131 Tb	2.5 Tb	64 Kb
MySQL	∞	256 Tb	64 Kb
Microsoft SQL Server	524258 Tb	524258 Tb	∞
Oracle	∞	4 Gb × physical block size	8 Kb
PostgreSQL	∞	32 Tb	1.6 Tb

A trigger is a stored procedure in a database that is called whenever a table row is inserted, modified, or deleted. Triggers verify any changes for correctness before these changes are accepted. A stored procedure is a compiled program that is stored on a server and can be executed by a client. A stored procedure contains a few SQL statements and some manipulative or logical processing. Since stored procedures are performed directly on the database server, higher performance

is provided than when performing the same operations using the database client.

By the criterion of the possibility of using triggers and stored procedures, all alternatives are identical. All database management systems support triggers, procedures, and functions.

TABLE 4. A JUDGMENT MATRIX OF RELEVANT IMPORTANCE UNDER FEATURES OF ARCHITECTURE AND FUNCTIONALITY

	D_1	D_2	D_3	D_4	D_5	D_6	W
D_1	1	1	2	1/8	1/3	1/6	0.07
D_2	1	1	2	1/3	1	1/2	0.10
D_3	1/2	1/2	1	1/4	1	1/2	0.06
D_4	8	3	4	1	5	3	0.41
D_5	3	1	1	1/5	1	1/2	0.11
D_6	6	2	2	1/3	2	1	0.22

The main eigenvalue $\lambda_{max} = 6.58$, $CI = 0.11$, $CR = 0.09$. The value of the consistency ratio is within the normal range.

Currently, many different methods and benchmarks ratings are used to estimate the performance of database management systems. The most authoritative tool is the TPC analysis conducted by the Transaction Processing Performance Council [9]. This is due to the presence of universal benchmarks for transaction processing. In addition to evaluating performance in the framework of TPC analysis, the ratio of the number of requests processed over a time interval to the cost of the entire system is given. The most interesting is the implementation of unequal access to the tables, and this is evaluated only by the TPC-C performance analysis, therefore, its results are given in Table 5. Performance is measured in tpmC units representing the number of transactions per minute. Cost refers to the unit cost of one transaction, defined as the ratio of price and performance.

TABLE 5. RESULTS OF TPC-C PERFORMANCE ANALYSIS

Database management systems	**Performance, tpmC**	**Cost, USD**
IBM DB2 9.5	1 200 011	0.69
Microsoft SQL Server 2005 Enterprise Edition x64	661 475	1.16
Oracle Database 11g Standard	631 766	1.08

TABLE 6. A JUDGMENT MATRIX OF RELEVANT IMPORTANCE UNDER THE PERFORMANCE

	D_1	D_2	D_3	D_4	D_5	D_6	W
D_1	1	4	5	3	4	5	0.39
D_2	1/4	1	2	1/3	1/3	2	0.10
D_3	1/5	1/2	1	1/3	1/2	1	0.06
D_4	1/3	3	3	1	2	3	0.21
D_5	1/4	3	2	1/2	1	2	0.15
D_6	1/5	1/2	1	1/3	1/2	1	0.06

The main eigenvalue $\lambda_{max} = 6.41$, $CI = 0.08$, $CR = 0.06$. The value of the consistency ratio is within the normal range.

To construct a pairwise comparisons matrix by the criterion of system requirements, we will analyze the operating systems supported by database management systems.

TABLE 7. A JUDGMENT MATRIX OF RELEVANT IMPORTANCE UNDER THE SYSTEM REQUIREMENTS

	D_1	D_2	D_3	D_4	D_5	D_6	W
D_1	1	1/3	1/4	1/4	1	1/3	0.05
D_2	3	1	1	1	4	3	0.23
D_3	4	1	1	1	4	3	0.25
D_4	4	1	1	1	4	3	0.25
D_5	1	1/4	1/4	1/4	1	1/2	0.05
D_6	3	1/3	1/3	1/3	2	1	0.15

The main eigenvalue $\lambda_{max} = 6.18$, $CI = 0.04$, $CR = 0.03$. The value of the consistency ratio is within the normal range.

To construct a pairwise comparisons matrix by the criterion of the features of application development, it is necessary to evaluate the time costs for installing and configuring the database, backup, and recovery, and also the ongoing maintenance of the database.

TABLE 8. A JUDGMENT MATRIX OF RELEVANT IMPORTANCE UNDER THE FEATURES OF APPLICATION DEVELOPMENT

	D_1	D_2	D_3	D_4	D_5	D_6	W
D_1	1	1	1	1/6	1	1	0.09
D_2	1	1	1	1/5	1	1	0.11
D_3	1	1/2	1	1/4	1	1	0.09
D_4	6	5	4	1	6	3	0.45
D_5	1	1	1	1/4	1	1	0.01
D_6	1	1	1	1/3	1	1	0.01

The main eigenvalue $\lambda_{max} = 6.16$, $CI = 0.03$, $CR = 0.02$. The value of the consistency ratio is within the normal range.

The urban web services operates with secret data, various methods of protection are used to prevent unauthorized access: identification, authentication, authorization, protection against attacks, data encryption, et al.

To construct a pairwise comparisons matrix by the criterion of reliability, we will analyze the security subsystem functionality supported by database management systems.

TABLE 9. A JUDGMENT MATRIX OF RELEVANT IMPORTANCE UNDER THE RELIABILITY

	D_1	D_2	D_3	D_4	D_5	D_6	W
D_1	1	4	4	3	1/2	1/3	0.20
D_2	1/4	1	1	1/4	1/5	1/5	0.01
D_3	1/4	1	1	1/4	1/5	1/5	0.04
D_4	1/3	4	4	1	1/2	1/2	0.16
D_5	2	5	5	2	1	1	0.25
D_6	3	5	5	2	1	1	0.27

The main eigenvalue $\lambda_{max} = 6.46$, $CI = 0.09$, $CR = 0.07$. The value of the consistency ratio is within the normal range.

To construct a pairwise comparisons matrix where the criterion of performance and reliability are of the greatest importance compared to other criteria. The numbering of the criteria corresponds to the numbering of the main criteria

defined at the stage of establishing an analytic hierarchy process based model on evaluation database management systems. The judgment matrix of relevant importance under the main criteria are shown in Table 10.

TABLE 10. A JUDGMENT MATRIX OF RELEVANT IMPORTANCE UNDER THE MAIN CRITERIA

	1	**2**	**3**	**4**	**5**	**6**	***W***
1	1	1	1/6	1/4	1/2	1/5	0.05
2	1	1	1/6	1/3	1/2	1/5	0.05
3	6	6	1	2	5	2	0.36
4	4	3	1/2	1	1/2	1/2	0.15
5	2	2	1/5	2	1	1/2	0.12
6	5	5	1/2	2	2	1	0.25

The main eigenvalue $\lambda_{max} = 6.46$, $CI = 0.09$, $CR = 0.07$. The value of the consistency ratio is within the normal range.

To calculate overall composite weights we select all the priority vectors of the alternatives according to all criteria and multiply the eigenvector matrix by a priority vector.

$$\begin{matrix} 0.12 & 0.07 & 0.39 & 0.05 & 0.09 & 0.2 \\ 0.22 & 0.10 & 0.10 & 0.23 & 0.11 & 0.01 \\ 0.18 & 0.06 & 0.06 & 0.25 & 0.09 & 0.04 \\ 0.28 & 0.41 & 0.21 & 0.25 & 0.45 & 0.16 \\ 0.06 & 0.11 & 0.15 & 0.05 & 0.10 & 0.25 \\ 0.12 & 0.22 & 0.06 & 0.15 & 0.10 & 0.27 \end{matrix} \times \begin{matrix} 0.05 \\ 0.05 \\ 0.36 \\ 0.15 \\ 0.12 \\ 0.25 \end{matrix} = \begin{matrix} 0.21 \\ 0.10 \\ 0.09 \\ 0.24 \\ 0.14 \\ 0.14 \end{matrix}$$

Thus, the weights of all considered alternatives were distributed as follows: Microsoft SQL Server – 0.24, DB2 – 0.21, Oracle – 0.14, Postgre SQL – 0.14, Firebird – 0.10, MySQL – 0.09. This multi-criteria decision making approach to urban web service databases evaluation based on the Analytic hierarchy process is revealed that the best database management system option is Microsoft SQL Server.

V. CONCLUSION

Chief Information Officers should identify the potential benefits of various elements in terms of performance, profitability, reliability, and flexibility since issues of particular importance are not so much cost-effective as verified and reasonable technical solutions.

The proposed method can be used when choosing the most suitable database management systems according to a set of criteria. The accuracy of the result will depend on the number of selected criteria and the competence of the expert comparing the database management systems according to the selected criteria.

This approach for the urban web service databases evaluation appears to enable the structured and systematic way of selecting the most appropriate database management systems.

REFERENCES

[1] B. Roy. Multicriteria Methodology for Decision Aiding, Springer Science & Business Media, 2013.

[2] V. Belton, T. Stewart. Multiple Criteria Decision Analysis: An Integrated Approach, Springer Science & Business Media, 2002.

[3] T. L. Saaty. A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15, pp. 234-281, 1977.

[4] T. L. Saaty. Decision Making for Leaders, New York: Van Nostrand Reinhold, 1982.

[5] R.W. Saaty. The analytic hierarchy process: What it is and how it is used, Mathematical Modelling 9, 3-5, pp. 161-176, 1987.

[6] C. Chang, C. Wu, H. Lin. Integrating fuzzy theory and hierarchy concepts to evaluate software quality, Software Quality Journal, 16 (2), pp. 263-276, 2008.

[7] C. Gao, J. Ma, Z. Liu, X. Ma. An approach to quality assessment for web service selection based on the analytic hierarchy process for cases of incomplete information, Science China Information Sciences, 58 (12), 2015.

[8] B. Feizizadeh, M. S. Roodposhti, P. Jankowski, T. Blaschke. A GIS-based Extended Fuzzy Multi-Criteria Evaluation for Landslide Susceptibility Mapping, Computers & Geosciences 73, pp. 208-221, 2014.

[9] B. Mbaiossoum, L. Bellatreche, S. Jean. Towards performance evaluation of semantic databases management systems, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 7968 LNCS, pp. 107-120, 2013.

[10] I. Hababeh, I. Khalil, A. Khreishah. Designing high performance web-based computing services to promote telemedicine database management system, IEEE Transactions on Services Computing, 8 (1), pp. 47-64, 2015.

[11] B. Raza, Y. Kumar, A. Malik, A. Anjum, M. Faheem. Performance prediction and adaptation for database management system workload using Case-Based Reasoning approach, Information Systems, 76, pp. 46-58, 2018.

[12] Transaction Processing Performance Council, Retrieved May 20, 2020, from http://www.tpc.org/downloaded_result_files/tpcc_results.xls.

[13] T. Thüm, S. Apel, C. Kästner, I. Schaefer, G. Saake. A classification and survey of analysis strategies for software product lines, ACM Computing Surveys, 47 (1), 2014.

[14] S. Pareek, N. Sharma, A. Geetha Mary. Fault Tolerance in Distributed Database Management Systems - Improving reliability with RAID, 2019 Innovations in Power and Advanced Computing Technologies (i-PACT), 2019.

[15] J. Rezaei. Best-worst multi-criteria decision-making method, Omega (United Kingdom), 53, pp. 49-57, 2015.