Fuzzy multi-criteria selection procedures in choosing data source

E. Azhir[1], N. Daneshpour[2*] and S. Ghanbari[3]

1. Department of Computer Engineering, Qazvin Azad University, Qazvin, Iran.
2. Department of Computer, Shahid Rajaee Teacher Training University, Tehran, Iran.
3. IRIB Technical Research Center, Tehran, Iran.

Abstract

Technology assessment and selection has a substantial impact on the organization procedures in regard to technology transfer. Technological decisions are usually made by a group of experts, and whereby, integrity of these viewpoints to a single decision can be quite complex. Today, operational databases and data warehouses exist to manage and organize data with specific features, and henceforth, the need for a decision-aid approach is essential. The process of developing data warehouses involves time-consuming steps, complex queries, slow query response rates, and limited functions, which is also true for operational databases. In this regard, fuzzy multi-criteria procedures used to choose efficient data sources (data warehouse and traditional relational databases) based on organization requirements are addressed in this paper. In proposing an appropriate selection framework, the paper compares a triangular fuzzy number (TFN)-based framework and a fuzzy analytical hierarchy process (AHP) based on data source models, business logic, data access, storage, and security. The results obtained show that the two procedures rank data sources in a similar manner and due to an accurate decision-making.

Keywords: *Analytical Hierarchy Process, Data Warehouse, Fuzzy, Multi-Criteria Decision-Making, Operational Database.*

1. Introduction

Today's digital world has accounted for an exponential rise of data. However, data can only be useful if it is appropriately managed and accessed. This rise has caused organizations to reconsider their data management technologies. While some simply enhance their existing operational databases, others consider larger solutions, data warehouses. The evaluation and selection of a suitable solution is a typical decision-making problem.

Data warehouses store large amounts of data using a multi-layer architecture. Data can be integrated from a number of sources including operational databases and/or other distributed data sources of different organizations. Whereas operational database tables are analyzed and designed based on entities, the analysis and design of data warehouse tables are based on the subjects that the system has been designed to analyze. Therefore, the fact and dimension tables of a data warehouse are designed by the operational database tables that take into consideration the subjects and goals of analysis.

Although most data warehouse queries are writable through operational database commands, and also elements of data warehouse are provided by operational databases, some of these queries are not expressed easily using operational database queries, and when they are actually expressed, their performance is extremely low. Other disadvantages of operational databases include formulating complex queries, slow response rates to queries, limited support of a number of functions and operations, and insufficient memory in answering queries. Nonetheless, it should be considered that operational databases are quite adequate for some organizations [1]. When correctly implemented, a data warehouse system enables companies to enjoy its benefits and obtain timely information for decision-making.

In the recent decades, various decision-making methods and algorithms have been designed. More often multi-criteria decision-making methods are proposed, taking the advantages of multiple criteria. Multi-criteria decision-making problems can be classified into two types, one of which exactly measures the weights of the criteria [2], and the other type uses fuzzy multi-criteria decision-making (FMCDM) such that the weight is approximately measured and is expressed in linguistic terms and transformed into fuzzy numbers (first introduced by Zadeh [3]). Furthermore, various multi-criteria decision-making methods including value measurement models (such as Fuzzy AHP), goal, aspiration, and reference models (such as TOPSIS) and outranking methods (such as ELECTRE and PROMETHEE) are implemented [4].

AHP decomposes complicated problems from higher hierarchies into lower ones. Also the analytical network process (ANP) technique is a multi-attribute approach that allows feedback loops among decision elements in the hierarchical or non-hierarchical structures, and is used to determine data sources in a longer term. In both techniques, impreciseness of human judgments can be handled through the fuzzy set theory.

Moreover, a variety of methodologies and frameworks have already been developed for software selection and evaluation. A method used for selecting enterprise resource planning (ERP) systems has been proposed based on the fuzzy set theory and analytical hierarchy process in [5]. Mamghani [6] has applied the analytic hierarchy process, designed for the decisions that require integration of quantitative and qualitative data in order to evaluate and select the anti-virus and content-filtering software. Combination of the fuzzy set theory and hierarchal structure analysis has resulted in a robust selection algorithm, as proposed by Liang and Wang [7].

Furthermore, a comprehensive framework has been developed for evaluation of the software systems based on the multi-criteria decision-making methods using the fuzzy set theory [8]. A new decision-making approach has been proposed by Eldrandaly and Naguib for solving the geographic information system (GIS) software selection by integrating expert systems and multi-criteria decision-making techniques [9]. Blanc and Jelassi [10] have developed a multi-criteria decision methodology for the decision support system (DSS) selection. Philips Wren, Hahn, and Forgionne [11] have applied fuzzy technologies to choose an appropriate ERP. A new model of assessing enterprise implementation readiness of

ERP based on the fuzzy multi-criteria decision-making has been proposed in [12]. Furthermore, a fuzzy multi-criteria decision-making procedure has been proposed in [13] to facilitate the data warehouse system selection with consideration given to both the technical and managerial criteria. A fuzzy set approach has been developed in [14] for the multi-criteria selection of the object-oriented simulation software for analysis of the production systems.

Selection of the most suitable data source system from a set of alternatives on the basis of many criteria is a multi-criteria decision-making problem. We found that using TFN made data collection and interpretation of results easier for the experts and decision makers. In this work, a fuzzy multi-criteria decision-making framework was developed using TFN.

Furthermore, fuzzy AHP, which is a multi-criteria evaluation method evolved from Saaty's AHP [15], has become one of the best known and most widely used multi-criteria decision-making methods. In this case, the criteria that have to be considered are quite a lot, and AHP that can take into consideration the relative priorities of the factors or alternatives and represent the best alternative, in turn, provides a simple and very flexible model for a given problem. In this situation, all levels of details about the main focus can be listed or structured in this method. Computer software can aid decision-makers to apply AHP quickly and precisely. AHP relies on the judgments of experts from different backgrounds, and thus the main focus can be evaluated easily from different aspects. Subsequently, the decision-maker can analyze the elasticity of the final decision by applying the sensitivity analysis.

However, the method has some disadvantages such as scoring the relative importance among related criteria that can be difficult about a certain amount. However, over-simplifying the hierarchy may lose important inter-dependencies among criteria, and over-extending the hierarchy may increase the time complexity for creating pairwise comparison matrices. There is not always a solution to the linear equations since the computational requirements are tremendous, even for a small problem. AHP allows only the triangular fuzzy numbers to be used, and the subjective nature of the modeling process is a constraint of AHP (that means that the methodology cannot guarantee the decisions as definitely true). When the number of levels in the hierarchy increases, the number of pair comparisons also increases, and hence, building

the AHP model takes much more time and effort [16].

Considering the multi-criteria structure of the data source selection problem and the vagueness in real environment, fuzzy AHP is thought to be suitable and simple enough for selecting the best data source. Any level of details about the main focus can be listed or structured in this method. Through this way, the overview of the main problem can be represented very easily.

Based on a case study conducted in this paper and a comparison between the TFN-based fuzzy method and a Fuzzy AHP, the proposed selection framework is shown to reduce both cost and time of data source selection due to a proper decision-making based on organization requirements.

This paper is organized as follows. In the next section, a basic concept is discussed in relation to data warehouse generation. In the third section, the TFN and Fuzzy AHP procedures for data warehouse selection using a case study is described. In the fourth section, application of TFN and Fuzzy AHP is compared, and the results obtained are presented. Finally, conclusions based on this study are presented in the fifth section.

2. Basic concepts; data warehouse generation

The common steps involved in data warehouse generation comprises organizational requirements, physical environment set-up, data modeling, data preprocessing, extraction, transformation, data loading, and reporting. Each of these steps are fully-discussed in this section.

Collection and analysis of business requirements: In this step, based on the meetings held with different groups of data warehouse users, and other known approaches in gathering requirements [17] including interviews, standard templates, requirement prioritization, knowledge of subject areas, and through the review of the existing documents, the analytical requirements of end-users are documented.

Physical environment set-up: Once the business requirements have been collected and the analytical requirements of end-users have been clearly defined, it is necessary to set-up the physical servers and databases. At a minimum, it is necessary to set-up the development and production environments. For this, it is best for the different environments to use distinct applications and database servers. Having different environments is very important since all changes can be tested without affecting the production environment. Notably, development and production can occur during the user activity in the data warehouse.

Data modeling: A logical data model is built based upon the user requirements, and is subsequently translated into the physical data model. Deliverables of this step includes the definition and identification of the data sources as well as the logical and physical data models.

Data preprocessing: Once the fact and dimension tables of the data warehouse are designed based on the goals and subjects of the analysis, potential errors in the operational databases are identified and subsequently removed. Data errors include incorrect, incomplete, redundant, inconsistent or inappropriate structured data.

Extraction, transformation, and data loading (ETL): Kimball [18] has stated that the processes of extraction, transformation, and data loading in data warehouses is time-consuming, and he has even specified that it can take up to approximately 60% of the data warehouse implementation time. His reasoning is based upon the fact that it should extract data from operational databases, transform data into appropriate formats, and finally, load data into target tables. A relational data warehouse is the deliverable of this step.

Front end and report development: Data-source users use a variety of reporting tools to query the data warehouse such as excel and dashboard monitoring.

3. Fuzzy-based decision-making procedures

In this section, after a brief introduction to the basic concepts of fuzzy sets, algebraic operations, triangular fuzzy numbers, linguistic variables, and ranking fuzzy numbers, the proposed fuzzy-based decision-making framework is presented in section 3.1, followed by Fuzzy AHP in section 3.2.

The theory of fuzzy sets, introduced by Zadeh (1965), was developed to describe vagueness and ambiguity in real-world systems. Fuzzy set is a class of objects that defines membership degree through interval [0:1]. Zadeh has defined a fuzzy set A in X as a function $f_A(x)$, which associates each point in X a real number in the interval [0:1], with the value of $f_A(x)$ at x representing the "grade of membership" of x in A. If the degree of the membership of an object is "1", it means that the objects are completely bounded, and an object value of "0" means that the object is not absolutely bounded. Uncertain values are set between 0 and 1. This fuzzy set definition can be used to effectively express the vagueness of many real-world cases. In this paper, triangular fuzzy numbers are used as membership functions, since they are straightforward for decision-makers to use and calculate.

Calculation of the membership functions can also be achieved, considering Zadeh principle, after mapping fuzzy set via a function. For example, assume A_1 and A_2 as two triangular fuzzy numbers, where their addition and multiplication are respectively shown in (1) and (2) [2].

$$A_1 = (c_1, a_1, b_1), A_2 = (c_2, a_2, b_2)$$
$$A_1 \oplus A_2 : (c_1, a_1, b_1) \oplus (c_2, a_2, b_2) = \quad (1)$$
$$(c_1 + c_2, a_1 + a_2, b_1 + b_2)$$

$$A_1 \otimes A_2 : (c_1, a_1, b_1) \otimes (c_2, a_2, b_2) = \quad (2)$$
$$(c_1 * c_2, a_1 * a_2, b_1 * b_2) c_1 \geq 0, c_2 \geq 0$$

The importance weights of various criteria and the rating values of alternative data sources are considered as linguistic terms throughout this paper. In [14], professionals have used linguistic terms such as very low, low, medium, high, and very high for the importance weights of various criteria, which can be expressed via triangular fuzzy numbers (Table 1), and use linguistic terms such as very poor, poor, fair, good, and very good for rating values of alternative data sources (Table 2).

Table 1. Linguistic terms for importance weight of each criterion.

	Very low (VL)	Low (L)	Medium (M)	High (H)	Very high (VH)
Membership function	(0,0,0.3)	(0,0.3,0.5)	(0.2,0.5,0.8)	(0.5,0.7,1)	(0.7,1,1)

Table 2. Linguistic terms for rating.

	Very poor (VP)	Poor (P)	Fair (F)	Good (G)	Very good (VG)
Membership function	(0,0,0.2)	(0,0.2,0.4)	(0.3,0.5,0.7)	(0.6,0.8,1)	(0.8,1,1)

The aforementioned membership functions have been used in different areas such as robot selection [7] and selection of the object-oriented simulation software [14].

Selection of the most suitable data source from a set of alternatives (data warehouse and operational database) on the basis of many criteria creates a multi-criteria decision-making problem.

The alternative set provided to compare and choose is $A = (A_1, A_2, \dots, A_n)$, in which, each alternative A_i has two possibilities: selected and unselected. There is a group of k decision makers $(D_1, D_2, \dots, D_k)$ that evaluate the importance weights of m criteria $(C_1, C_2, \dots, C_m)$, and the appropriateness of n alternatives $(A_1, A_2, \dots, A_n)$ under each one of these m criteria. Let $A_{td}(t = 1, 2, \dots, m; d = 1, 2, \dots, k)$ be the weight given to C_t by decision-maker D_d, and let $R_{dti}(d = 1, 2, \dots, k; t = 1, 2, \dots, m; i = 1, 2, \dots, n)$ be the rating assigned to alternative A_i by decision-maker D_d under criterion C_t. A_{td} and R_{dti} are defined as follow ((3) and (4)):

$$R_{dti} = \left(\frac{1}{k}\right) * \begin{bmatrix} \sum_{d=1}^{k} r_{d11} & \cdots & \sum_{d=1}^{k} r_{dm1} \\ \vdots & \ddots & \vdots \\ \sum_{d=1}^{k} r_{d1n} & \cdots & \sum_{d=1}^{k} r_{dmn} \end{bmatrix} \quad (3)$$

$$A_{td} = \left(\frac{1}{k}\right) * \left(\sum_{d=1}^{k} a_{1d}, \dots, \sum_{d=1}^{k} a_{md}\right) \quad (4)$$

Notably, the fuzzy set theory is applied to give values of matrices ((3) and (4)).

Choosing the most appropriate data source and considering the importance of weights of various criteria and the rating values of alternative data sources will be done by matrix F in (5). In these matrices, $F_1, F_2, \dots, F_m$ show the weight of each data source for selection by the decision-makers to make a choice.

$$F = \left(\frac{1}{m}\right) * (R \otimes A^T) =$$

$$\left(\frac{1}{m}\right) \left(\begin{bmatrix} \frac{\sum_{d=1}^{k} r_{d11}}{k} & \cdots & \frac{\sum_{d=1}^{k} r_{dm1}}{k} \\ \vdots & \ddots & \vdots \\ \frac{\sum_{d=1}^{k} r_{d1n}}{k} & \cdots & \frac{\sum_{d=1}^{k} r_{dmn}}{k} \end{bmatrix} \otimes \begin{bmatrix} \frac{\sum_{d=1}^{k} a_{1d}}{k} \\ \frac{\sum_{d=1}^{k} a_{2d}}{k} \\ \vdots \\ \frac{\sum_{d=1}^{k} a_{md}}{k} \end{bmatrix} \right) \quad (5)$$

$$= (F_1, F_2, \dots, F_m)^T$$

The objective functions are written as a function of decision variables, and express the purpose of the problem and where the decision-maker attempts to maximize or minimize the objective function. In situations where the purpose is maximization, the most suitable option is when their objective function has the highest value. In considering (6) and according to calculations carried out by Chen [19], the maximizing set M is $M = \{(x, f_M(x)) | x \in R\}$:

$$f_M(x) = \begin{cases} \dfrac{(x - x_{min})}{(x_{max} - x_{min})}, & x_{min} \leq x \leq x_{max} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

and the minimizing set $G = \{(x, f_G(x)) | x \in R\}$, considering (7), is defined as:

$$f_G(x) = \begin{cases} \dfrac{(x - x_{max})}{(x_{min} - x_{max})}, & x_{min} \leq x \leq x_{max} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where:

$$x_{min} = \inf S, x_{max} = \sup S,$$
$$S = \bigcup_{i=1}^{n} F_i, F_i = \{x | f_{Fi}(x) > 0\}, i = 1, 2, \ldots, n$$

Furthermore, the right utility value $U_M(F_i)$, left utility value $U_G(F_i)$, and total utility value $U_T(F_i)$ for alternative i are denoted as follow ((8)-(10)).

$$U_M(F_i) = \sup(f_{Fi}(x) \cap f_M(x)), i = 1, 2, \ldots, n \quad (8)$$

$$U_G(F_i) = \sup(f_{Fi}(x) \cap f_G(x)), i = 1, 2, \ldots, n \quad (9)$$

$$U_T(F_i) = \frac{U_M(F_i) + 1 - U_G(F_i)}{2} \quad (10)$$

The alternative with a maximum $U_T(F_i)$ value is the optimal choice in the decision-making problem.

3.1. Fuzzy-based multi-criteria decision-making framework using TFN

A generalized fuzzy multi-criteria decision-making procedure is depicted in figure 1 as an activity diagram. Each level of this activity diagram is described as follows.
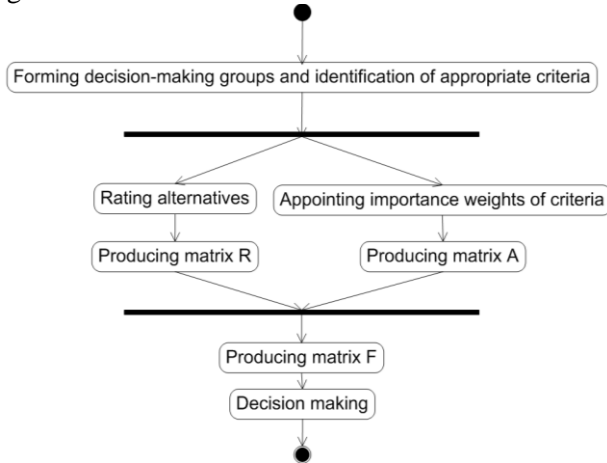

Figure 1. Activity diagram of a fuzzy-based decision-making procedure for selection of a data source.

Forming decision-making groups for analysis of data sources and identification of appropriate

criteria: Due to different advantages and disadvantages of the data sources, different organizations based on their requirements may decide to create data warehouse or use existing data sources (those that do not create data warehouse). Therefore, several groups are formed in the organization to investigate possible data storage alternatives, organization requirements, and influencing factors in order to determine appropriate selection criteria.

Delphi is a technique of committee research, developed by Dalkey and Helmer [20]. It is a widely used and accepted methodology for achieving convergence of opinion concerning real-world knowledge solicited from experts within certain topic areas. Assuming the skills of experts are more important than their quantity. Delphi is used in this paper to choose the committee experts. First of all, lists of the names of professionals of the organization have been grouped by their profession. Professionals of each group were rated by their level of competency, and subsequently, the organization research team chose six experts. The six data source experts have an average of 9/5 years of experience in business.

The main purpose of choosing criteria is to provide a vision for a better decision-making. At first, the initial list of objectives including major and minor objectives of the organization is drawn. These are the organization objectives that have been organized in a hierarchal structure. In this paper, the hierarchal structure of objectives was developed during a two-week period, whereby three questionnaires were sent to the chosen experts.

It is important to note that too many criteria may lead to conflict of evaluations and consequently unsuccessful decision-making. For this reason, in this study, at first 30 objectives were expressed in the committee. Subsequently, fuzzy Delphi was used to ensure the selection of appropriate criteria, and as a result, 15 criteria were chosen. The chosen criteria were discussed and approved by four data warehouse professionals of two reputable organizations. Figure 2 shows the proposed criteria that are derived by committee decisions and studies conducted in [19,20] and are suitable for decision-making to choose the most suitable data source in a hierarchical structure.
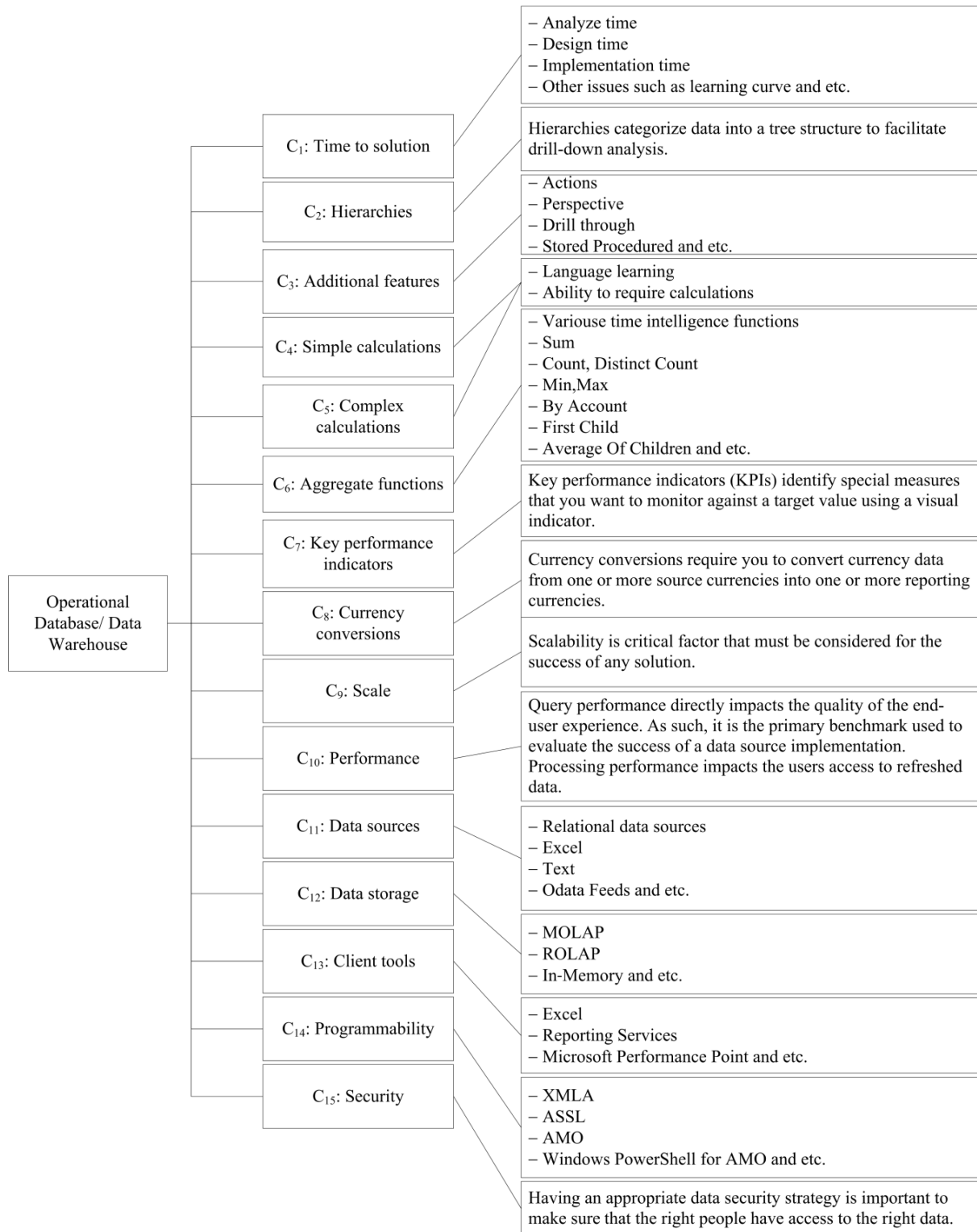
Figure 2. Criteria used for decision-making.

Rating alternatives and producing matrix R activities: In these activities, data sources are rated by the fuzzy method. This rating and assigning of importance weights of criteria and producing matrix A activities of figure 1 can be carried out simultaneously. The rating alternatives under the figure 2 criteria are defined according to the standard characteristics of the operational databases and data warehouses defined by their vendors, conducted studies, and selected expert experiments using the terms in table 2. The ratings acquired by the six experts are expressed as follows.

Rating according to criterion C_1 (solution time): Design and implementation of data warehouse is in the form of a multi-layer architecture, where the first layer is composed from data sources such as relational databases, flat files, and other sources.

After designing the tables of a relational data warehouse, the extracted and transformed data of the operational databases is loaded into this data warehouse. Henceforth, to design a data warehouse, in addition to analysing tables of data sources, it is necessary to create fact, dimension, and multidimensional cube in business intelligence (BI) environment.

Although comprehensive data modelling and sophisticated analytics are important benefits of data warehouse, they often come with the trade-off of longer development cycles. Based on the average comments of the six experts and the provided definitions, the quantitative values (0.61,0.81,0.95) and (0.1,0.3,1) are assigned to criterion C_1 for the operational databases and data warehouse accordingly.

Rating according to criterion C_2 (hierarchies): Creation of data warehouse hierarchies such as standard, parent-child, and ragged [22] is easily available by BI wizard features, while creation of these hierarchies in the operational databases is more difficult and requires complex queries. Henceforth, based on the average comments of the six experts and the provided definitions, the quantitative values (0.73,0.93,1) and (0.31,0.51,0.7) are assigned to criterion C_2 for the data warehouse and operational database.

Rating according to criterion C_3 (additional features): Data warehouse additional features, which are provided in table 3, are easily available by BI environment, although in operational databases, some of these features are much more complex. For example, a translation feature is not available in operational databases, and drill-through and write-back is more difficultly provided and requires writing complex queries. Linguistic terms of these six experts is transformed into fuzzy numbers and then assigned to the values (0.4,0.6,0.8) and (0.7,0.9,1) for criterion C_3 in operational database and data warehouse.

Table 1. Additional features.

Additional features	Data warehouse	Database
Actions	✓	✓
Perspective	✓	✓
Drill through	✓	Complex solutions
Stored procedures	✓	✓
Write-back	✓	Complex solutions
Translations	✓	✗

Rating according to criterion C_4 (simple calculations): Calculations in the data warehouses and operational databases depend on both the learning data source language and the ability of the language to handle the required queries. As multi-dimensional expressions language (MDX) is more complicated in the data warehouses in comparison with operational databases languages such as structured query language (SQL) and excel functions, it is not recommended for simple calculations. After comparing these two data sources by six experts and due to the ease of operational database calculation language learning and ability of operational database in processing simple queries, the fuzzy values (0.4,0.6,0.8) and (0.53,0.73,0.9) are assigned to criterion C_4.

Rating according to criterion C_5 (complex calculations): MDX calculation languages are more applicable to complex calculations than operational databases languages such as SQL, excel functions, oracle and so forth, as in the majority of cases they will fail in processing such complex calculations. Henceforth, based on the average comments of the six experts and the provided definitions, the quantitative values (0.7,0.9,1) and (0.05,0.18,0.35) are assigned to criterion C_5 for the data warehouse and operational database.

Rating according to criterion C_6 (aggregate functions): As shown in table 4, data warehouse, in addition to providing aggregate functions that is supported by operational databases, provides other aggregate functions. Due to the storage of pre-computed aggregate functions in data warehouse multi-dimensional cube cells that improves the efficiency of query retrieves, and based on the average comments of the six experts, the quantitative values (0.6,0.93,1) and (0.1,0.3,0.5) are assigned to criterion C_6 for the warehouse and operational database.

Table 2. Aggregate functions.

Aggregate functions	Data warehouse	Database
Sum	✓	✓
Count, Distinct Count	✓	✓
Min, Max	✓	✓
None	✓	✗
By account	✓	✗
Average of children	✓	✗
First child, Last child	✓	✗
First non-empty	✓	✗
Last non-empty	✓	✗
STDEV	✗	✓
VAR	✗	✓

Rating according to criterion C_7 (key performance indicators (KPIs)): KPIs that are provided by each data source and expressed in the third level of figure 2 are shown in table 5. KPIs of operational

database are provided by components such as SQL reporting services and performance point services. However, while these are not available as a wizard in an operational database environment, the KPIs are provided in the data warehouse BI environment. Due to the higher capability of the data warehouse in criterion C_7, difficulty of preparation of KPIs by operational database, and average comments of the six experts, the quantitative values (0.7,0.9,1) and (0.05,0.18,0.18) are assigned to criterion C_7 for the data warehouse and operational database.

Table 3. Key-performance indicators.

key performance indicators	Data warehouse	Database
Actual	✓	Complex solutions
Goal	✓	Complex solutions
Status	✓	Complex solutions
Trend	✓	Complex solutions
Graphical indicators	✓	Complex solutions

Rating according to criterion C_8 (currency conversions): In data warehouse, the BI Wizard can be used to create the MDX currency conversion calculations that are optimized to support multiple source and reporting currencies. In an operational database currency, conversions can be made by SQL queries. These currency conversions are easier to perform in data warehouses. Based on the average comments of the six experts and the capability of data warehouse in currency conversions, the quantitative values (0.6,0.93,0.1) and (0.05,0.21,0.41) are assigned to criterion C_8 for the data warehouse and operational database.

Rating according to criterion C_9 (large scale): Data warehouses have the ability of extending into large scales (multi-terabyte) without affecting query performances. Data warehouses provide extensive capabilities to manage the most complex and largest-scale BI challenges. Sophisticated models and complex business logic can be implemented in data warehouses. On-disk data storage, pre-calculated aggregates, and in-memory caching enable multi-dimensional models to grow to multi-terabyte scale and to provide fast query responses.

However, operational databases that only have one processor and are without suitable indexing, partitioning, and tables normalization are not able to extend into large scales. This feature of operational database and data warehouse is compared. In this comparison, the scaling ability of data warehouse based on average comments of six experts is represented by fuzzy numbers (the quantitative values (0.6,0.93,1) and

(0.15,0.35,0.41) are assigned to criterion C_9 for the warehouse and operational database).

Rating according to criterion C_{10} (performance): data warehouses provide a variety of mechanisms to accelerate query performance including aggregations, caching, indexed data retrieval, and data compression. In addition, one can improve query performance by optimizing the design of dimension attributes, cubes, and MDX queries. As a result of data compression and storing pre-calculated aggregate functions in multi-dimensional cube cells, I/O will reduce. A disadvantage of operational databases is that they cannot compress data, although in certain situations, through indexing, caching, and storing pre-calculated aggregate function response rates to queries can be reduced.

Performance of data warehouse and operational database based on C_{10} is compared. This comparison is achieved by taking the average comments of decision-makers and converting them to fuzzy numbers and then assigning them to the quantitative values (0.7,0.9,1) and (0.1,0.3,0.5) for the data warehouse and operational database, respectively.

Rating according to criterion C_{11} (data sources): A number of data sources supported by data warehouse and operational database are expressed in the third level of figure 2 and table 6. Henceforth, the data warehouse and operational databases are compared based on the linguistic terms of six decision-maker experts. (The linguistic terms of these six experts are transformed into fuzzy numbers, and then assigned to the average values (0.38,0.71,0.73) and (0.05,0.21,0.41) for criterion C_{11} in operational database and data warehouse.)

Rating according to criterion C_{12} (data storage): In a data warehouse, data can be stored in multi-dimensional online analytical processing (MOLAP), relational online analytical processing (ROLAP) and hybrid online analytical processing (HOLAP) data architectures. In MOLAP, the data is stored on disk in an optimized multi-dimensional format. In ROLAP, the data is stored in the source relational database. Due to optimized storage, multi-dimensional indexing, and caching, query response time is fast in MOLAP but in MOLAP, on-disk size of data compared to data stored in the relational database is smaller.

The ROLAP tools access the data in a relational database, and generate SQL queries to calculate information at the appropriate level as and when requested by end-users.

As ROLAP places no limitation on the quantity of data and since each ROLAP report is essentially a

SQL query (or multiple SQL queries) in the relational database, the query time can be long if the underlying data size is large. HOLAP combines features of the MOLAP and ROLAP data architectures.

Table 4. Supported data sources.

Database	Data warehouse
Microsoft SQL Server	SQL Server relational databases
Microsoft SQL Server Analysis Services for MDX, DMX,	Teradata relational databases
Microsoft PowerPivot, and tabular models	Informix relational databases
Microsoft Windows Azure SQL Database	IBM DB2 relational databases
SQL Server Parallel Data Warehouse	Sybase relational databases
Oracle	Other relational databases (OLE DB provider or
SAP Net Weaver BI	ODBC driver)
Hyperion Essbase	
Microsoft SharePoint List	
Teradata	
OLE DB	
ODBC and XML	

In operational databases, information storage and retrieval is achieved through SQL queries. Based on the average comments of the six experts and the capability of data warehouse in information storage and retrieval, the quantitative values (0.7,0.9,1) and (0.15,0.35,0.60) are assigned to criterion C_{12} for the data warehouse and operational database.

Rating alternatives according to criterion C_{13} (client tools): As shown in the third level of figure 2, Excel reporting services and Microsoft performance point are supported as data warehouse client tools. Furthermore, these tools are also supported in operational databases. After linguistic comparison of these two data sources by the six experts, the fuzzy values (0.66,0.88,1) and (0.61,0.81,0.95) are assigned to criterion C_{13} accordingly.

Rating according to criterion C_{14} (programmability): A variety of application programming interfaces (APIs) that can be used to develop and manage operational database objects such as tables, SQL scripts, and data warehouse objects such as cubes, dimensions, measures groups, and MDX scripts are tabulated in table 7. Considering the provided comparison given in table 7, data warehouses have more functionality than operational databases. After linguistic comparison of these two data sources by the six experts, the fuzzy values (0.7,0.9,1) and (0.4,0.6,0.8) are assigned to criterion C_{14} for the data warehouse and operational database.

Rating according to criterion C_{15} (security): Having an appropriate data security strategy is important to certify that the right people have access to the right data.

Table 5. Programmability.

Data warehouse	Database
XMLA	XML
ASSL	ADO.NET
ADOMD.NET	
MSOLAP	
AMO	
Windows PowerShell for AMO	

Both data warehouse and operational databases offer a set of robust capabilities that satisfy a broad range of security requirements. However, there are subtle differences in their capabilities, which are important to understand before choosing the data source experience that will best meet the security needs. In a data warehouse project, the concept of dimension data security can be used to manage row-level access. In operational database project, a row-level security can be implemented by granting access to rows in a table. In a multi-dimensional project, cell-level security can be implemented to restrict access to a particular cell or group of cells. Cell-level security is not provided in an operational database. Since security of access level in data warehouse is higher than operational database and based on the collected comments, the quantitative values (0.7,0.9,1) and (0.2,0.4,0.6) are assigned to criterion C_{15}.

Appointing importance weight of criteria and producing matrix A activities: In order to create a data warehouse in organization, six experts determine importance weights of figure 2 criteria based on their wide experience and knowledge about requirements of organization. First, the organization requirements in creating a data warehouse are collected, as shown in table 8.

Table 8. Organization requirements to create data warehouse.

Criterion	Organization requirements
C_1	Importance of solution time for organization
C_2	Massive creation of hierarchies and fast retrieval of them based on the nature of organization's analyzes
C_3	Stored Procedures , Actions, Drill through
C_4	60 percent of the organization's calculations need simple queries
C_5	40 percent of the organization's calculations need complex queries
C_6	Organization uses common aggregate functions in its process
C_7	As organization do not require monitoring then key performance indicators is not required either
C_8	Organization do not require currency conversions
C_9	in order to facilitate organization's process, data source must have the ability of extending into gigabyte scales
C_{10}	Organization need to reduce response time of quires and process
C_{11}	Organization need to create data sources based on Microsoft Sql Server relational data sources
C_{12}	Saving data in memory
C_{13}	Probability of using excel client tool
C_{14}	Organization should not need to use different programming environments
C_{15}	Row level security

3.2. Fuzzy AHP

Fuzzy AHP is one of the most commonly used methods for solving multi-criteria decision-making problems. After decomposing the complex problem into a hierarchical structure, the pair-wise comparison matrices are calculated as follow:

$$A=(aij) = \begin{bmatrix} 1 & \frac{w1}{w2} & \cdots & \frac{w1}{wn} \\ \frac{w2}{w1} & 1 & \cdots & \frac{w2}{wn} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{wn}{w1} & \frac{wn}{w2} & \cdots & 1 \end{bmatrix}, \qquad (11)$$

Weight calculation: $w_i = \frac{\sum_{i=1}^{n} aij}{n}$

for all $j = 1, 2, \ldots, n$

In this study, the extent fuzzy AHP is utilized, which was originally introduced by Chang (1996). The value of fuzzy synthetic extent with respect to the object is defined as:

$$S_i = \sum_{j=1}^{n} M_{gi}^{j} \otimes \left[\sum_{i=1}^{n} \sum_{j=1}^{m} M_{gi}^{j} \right]^{-1} \qquad (12)$$

Where $M_{gi}^{j}(j = 1, \ldots, m)$ are TFNs and goal set G = $\{g_1, g_2, g_3, \ldots, g_n\}$

To obtain fuzzy summation of rows, the fuzzy addition operation of m extent analysis values for a particular matrix is performed such as:

$$\sum_{j=1}^{m} M_{gi}^{j} = \left(\sum_{j=1}^{m} l_j, \sum_{j=1}^{m} m_j, \sum_{j=1}^{m} u_j \right) \qquad (13)$$

and to obtain S_j, the following multiplication is performed such as:

$$S_j = \sum_{j=1}^{m} M_{gi}^{j} \otimes \left[\sum_{j=1}^{m} \sum_{j=1}^{m} M_{gi}^{j} \right]^{-1} =$$
$$\left(\sum_{j=1}^{m} l_j \otimes \sum_{i=1}^{n} l_i, \sum_{j=1}^{m} m_j \otimes \sum_{i=1}^{n} m_i, \sum_{j=1}^{m} u_j \otimes \sum_{i=1}^{n} u_i \right) \qquad (14)$$

As $M_1 = (l_1, m_1, u_1)$ and $M_2 = (l_2, m_2, u_2)$ are two triangular fuzzy numbers, the degree of possibility of M2 ≥ M1 is defined as:

$$V(M_2 \geq M_1) = \begin{cases} 1, & \text{if} M_2 \geq M_1 \\ 0, & \text{if} l_1 \geq l_2 \\ \frac{l_1 - u_2}{(m_2 - u_2) - (m_1 - l_2)}, & \text{otherwise} \end{cases} \qquad (15)$$

The degree possibility for a convex fuzzy number to be greater than k convex fuzzy M_i (i = 1,2,3, \ldots, k) numbers can be defined by:

$$V(M \geq M_{1,} M_2, \ldots, M_k) = \min V(M \geq M_i), \qquad (16)$$
$$i = 1,2,3, \ldots, k$$

Finally, $W = (\min V(S_1 \geq S_k), \min V(S_2 \geq S_k), \ldots, \min V(S_n \geq S_k))^T$ is the weight vector for $k = 1,2, \ldots, n$.

The importance weights of various criteria and the rating values of alternative data sources are considered as linguistic terms, which can be expressed via triangular fuzzy numbers (Table 9).

Table 9. Membership function of linguistic scale.

Linguistic	Scale of fuzzy number
Perfect	(8,9,10)
Absolute	(7,8,9)
Very good	(6,7,8)
Fairly good	(5,6,7)
Good	(4,5,6)
Preferable	(3,4,5)
Not bad	(2,3,4)
Weak advantage	(1,2,3)
Equal	(1,1,1)

Importance weights of criteria based on organizations requirements are assigned by six groups and fuzzy summation of pair-wise comparison matrix is calculated, as shown in table 10.

Table 10. Fuzzy summation of rows (sum of l, m, and u values).

Criterion	Fuzzy summation of rows
C_1	(33,55,79)
C_2	(28.25,45.5,66)
C_3	(24.95,38.83,55)
C_4	(23.95,36.83,53)
C_5	(34,54,78)
C_6	(18.53,28.83,42)
C_7	(6.04,10.30,19.23)
C_8	(4.057,5.20,9.3)
C_9	(15.48,27.72,42.7)
C_{10}	(26.25,44.53,64.25)
C_{11}	(19.85,33.56,50)
C_{12}	(8.49,12.64,23.65)
C_{13}	(9.70,19.92,32.58)
C_{14}	(6.52,13.27,23.08)
C_{15}	(4.73,7.63,13.76)

The weight vector from table 10 is calculated and normalized, as shown in table 11.

Table 11. Normalized weight calculation.

Possibility Degree of Si *vs.* Sk														Degree of possibility	Normalized weight
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**0.1164**
0.901	1	1	0.910	1	1	1	1	1	1	1	1	1	1	0.901	**0.1049**
0.809	0.915	1	0.817	1	1	1	1	1	1	1	1	1	1	0.809	**0.0942**
0.782	0.887	0.972	0.790	1	1	1	1	0.900	1	1	1	1	1	0.782	**0.0910**
0.991	1	1	1	1	1	1	1	1	1	1	1	1	1	0.991	**0.1153**
0.643	0.751	0.840	0.869	0.648	1	1	1	0.767	1	1	1	1	1	0.643	**0.0748**
0.178	0.267	0.345	0.371	0.171	0.510	1	0.550	0.292	0.442	1	1	1	1	0.171	**0.0199**
0	0	0	0	0	0.111	0.688	0.181	0	0.068	0.564	0.375	1	1	0	**0**
0.639	0.743	0.828	0.856	0.644	0.981	1	1	0.758	1	1	1	1	1	0.639	**0.0744**
0.889	0.989	1	1	0.898	1	1	1	1	1	1	1	1	1	0.889	**0.1035**
0.738	0.842	0.926	0.953	0.745	1	1	1	1	1	1	1	1	1	0.738	**0.0859**
0.285	0.379	0.460	0.487	0.282	0.621	1	1	0.655	0.551	1	1	1	1	0.282	**0.0328**
0.474	0.576	0.662	0.690	0.476	0.822	1	1	0.847	1	1	1	1	1	0.474	**0.0552**
0.014	0.092	0.162	0.186	0	0.327	0.874	1	0.380	0.267	1	1	1	1	0	**0**
0.277	0.373	0.455	0.483	0.273	0.622	1	1	0.657	1	1	1	1	1	0.273	**0.0318**

Similar procedures are carried out to calculate relative importance weight of alternatives with respect to each selection criterion.

4. Evaluation and comparison of data sources by two fuzzy procedures

4.1. Evaluate data sources by TFN based decision-making framework

The average of these six expert comments based on fuzzy values of table 2 that are used to create matrix R, and is denoted in rating alternatives and producing matrix R activities, is calculated and tabulated in table 12.

Table 12. Average comments of six experts.

Criterion	Database	Data warehouse
C_1	(0.61,0.81,0.95)	(0.1,0.3,1)
C_2	(0.31,0.51,0.7)	(0.73,0.93,1)
C_3	(0.4,0.6,0.8)	(0.7,0.9,1)
C_4	(0.53,0.73,0.9)	(0.4,0.6,0.8)
C_5	(0.05,0.18,0.35)	(0.7,0.9,1)
C_6	(0.1,0.3,0.5)	(0.6,0.93,1)
C_7	(0.05,0.18,0.18)	(0.7,0.9,1)
C_8	(0.05,0.21,0.41)	(0.6,0.93,1)
C_9	(0.15,0.35,0.41)	(0.6,0.93,1)
C_{10}	(0.1,0.3,0.5)	(0.7,0.9,1)
C_{11}	(0.38,0.71,0.73)	(0.05,0.21,0.41)
C_{12}	(0.15,0.35,0.6)	(0.7,0.9,1)
C_{13}	(0.61,0.81,0.95)	(0.66,0.88,1)
C_{14}	(0.4,0.6,0.8)	(0.7,0.9,1)
C_{15}	(0.2,0.4,0.6)	(0.7,0.9,1)

Importance weights of criteria of figure 2 based on organization requirements are assigned by six experts, as shown in table 12. A hierarchy approach is applied to determine the importance of the weights of criteria. Weights of second level criteria of figure 2 are calculated based on third-level criteria. For example, based on organization requirements of table 8, value of C_2 is determined, as shown in table 12. The organization needs to create more hierarchies such as standard and parent-child hierarchies, and then retrieves them fast based on the nature of organization analysis. Based on the organization requirements and the comment of one of the decision-makers (D_1), standards and parent child's hierarchies sub-criteria are assigned to a very high term and transformed to the fuzzy number (0.7, 1, 1). As ragged hierarchy is less required by organizations, it is assigned to a very low term and transformed to the fuzzy number (0, 0, 0.3). The average of aforementioned sub-criteria is (0.46, 0.66, 0.76) for C_2.

After determining the importance weights of criteria based on organization requirements in table 13 using Matrix (4), matrix A_1 is created and two columns of table 12 are inserted in matrix R.

Producing matrix F and final decision-making activities: Decision-making in regard to creating data warehouse and/or using the existing operational databases, based on the organization requirements, is achieved through matrix F in (17).

Table 6. Importance weight of each criterion based on requirements of organization.

Criterion	Decision-makers						Average of comments
	D_1	D_2	D_3	D_4	D_5	D_6	
C_1	(0.7,1,1)	(0.63,0.9,0.7)	(0.36,0.66,0.86)	(0.73,0.93,0.46)	(0.83,0.93,0.53)	(0.63,0.9,1)	(0.55,0.83,0.90)
C_2	(0.46,0.66,0.76)	(0.38,0.83,46)	(0.33,0.56,0.83)	(0.66,0.76,0.46)	(0.83,0.93,0.53)	(0.4,0.56,0.76)	(0.44,0.60,0.79)
C_3	(0.35,0.6,0.71)	(0.43,0.66,0.8)	(0.28,0.43,0.68)	(0.55,0.68,0.35)	(0.35,0.6,0.71)	(0.28,0.4,0.65)	(0.34,0.63,0.54)
C_4	(0,0.25,0.55)	(0.35,0.6,0.9)	(0.35,0.6,0.9)	(0.35,0.6,0.9)	(0.5,0.7,1)	(0.6,0.85,1)	(0.35,0.60,0.87)
C_5	(0.6,0.85,1)	(0.6,0.85,1)	(0.35,0.6,0.9)	(0.6,0.85,1)	(0.6,0.85,1)	(0.7,1,1)	(0.57,0.83,0.98)
C_6	(0.36,0.55,0.21)	(0.53,0.68,0.23)	(0.21,0.39,0.54)	(0.3,0.45,0.66)	(0.3,0.45,0.66)	(0.23,0.53,0.68)	(0.24,0.45,0.62)
C_7	(0,0.3,0.5)	(0,0,0.3)	(0,0.3,0.5)	(0.2,0.5,0.8)	(0.2,0.5,0.8)	(0,0.15,0.4)	(0.06,0.29,0.55)
C_8	(0,0,0.3)	(0,0,0.3)	(0.1,0.4,0.65)	(0.1,0.25,0.55)	(0.1,0.25,0.55)	(0,0,0.3)	(0.05,0.15,0.44)
C_9	(0.2,0.5,0.8)	(0.2,0.5,0.8)	(0.2,0.5,0.8)	(0.2,0.5,0.8)	(0.5,0.7,1)	(0.2,0.5,0.8)	(0.25,0.45,0.83)
C_{10}	(0.7,1,1)	(0.73,0.93,0.46)	(0.46,0.73,0.93)	(0.73,0.93,0.46)	(0.7,1,1)	(0.3,0.4,0.7)	(0.51,0.76,0.91)
C_{11}	(0.41,0.58,0.11)	(0.5,0.7,1)	(0.11,0.41,0.58)	(0.2,0.5,0.8)	(0.5,0.7,1)	(0.5,0.7,1)	(0.32,0.57,0.82)
C_{12}	(0.72,0.95,0.47)	(0.4,0.67,0.9)	(0.2,0.5,0.8)	(0.5,0.7,1)	(0.5,0.7,1)	(0.5,0.7,1)	(0.42,0.66,0.94)
C_{13}	(0.27,0.5,0.72)	(0.1,0.4,0.75)	(0.2,0.5,0.8)	(0.27,0.5,0.72)	(0.1,0.4,0.75)	(0.2,0.5,0.8)	(0.19,0.46,0.75)
C_{14}	(0,0,0.3)	(0,0,0.3)	(0.2,0.5,0.8)	(0.2,0.5,0.8)	(0,0.3,0.5)	(0,0,0.3)	(0.06,0.21,0.50)
C_{15}	(0.1,0.4,0.65)	(0.35,0.6,0.9)	(0,0,0.3)	(0.1,0.25,0.55)	(0.1,0.4,0.65)	(0.1,0.25,0.55)	(0.12,0.31,0.60)

As expressed, objective functions are written by the decision variables and express the problem statement. Here, decision-makers try to maximize or minimize the objective functions.

In problems where the goal is maximization, a suitable alternative is the one whose objective functions have a maximum measure. The membership functions of fuzzy numbers F_1 (operational database) and F_2 (data warehouse) are calculated in (17).

$$F = \left(\frac{1}{15}\right).(R \otimes A_1) =$$

$$\left(\frac{1}{15}\right).\begin{bmatrix}(0.61,0.81,0.95)(0.31,0.51,0.7)(0.4,0.6,0.8) & \cdots & (0.61,0.81,0.95) & (0.4,0.6,0.8) & (0.2,0.4,0.6) \\ \vdots & \ddots & & \vdots & \\ (0.1,0.3,1) & (0.73,0.93,1)(0.7,0.9,1) & \cdots & (0.66,0.86,1) & (0.7,0.9,1)(0.7,0.9,1)\end{bmatrix} \otimes \begin{bmatrix}(0.55,0.83,0.90) \\ (0.44,0.60,0.79) \\ (0.34,0.63,0.54) \\ (0.35,0.60,0.87) \\ (0.57,0.83,0.98) \\ (0.24,0.45,0.62) \\ (0.06,0.29,0.55) \\ (0.05,0.15,0.44) \\ (0.25,0.45,0.83) \\ (0.51,0.76,0.91) \\ (0.32,0.57,0.82) \\ (0.42,0.66,0.94) \\ (0.19,0.46,0.75) \\ (0.06,0.21,0.5) \\ (0.12,0.31,0.6)\end{bmatrix} = \quad (17)$$

$$((0.07,0.24,0.46) \quad , \quad (0.15,0.38,0.69))$$
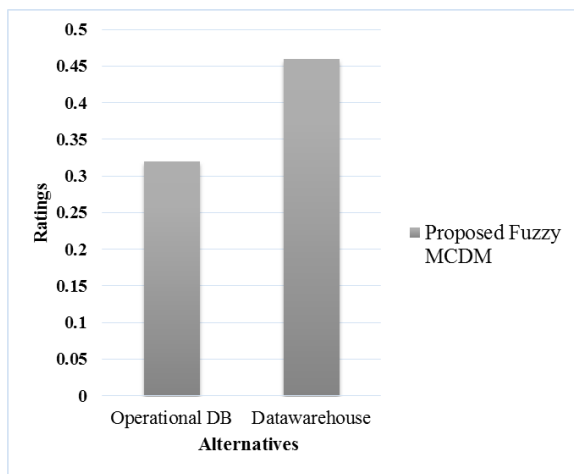
Figure 3 is derived from calculation of F_1 and F_2.

Figure 2. Rating with TFN-based decision-making framework.

Hence, it is clear that the most appropriate data source is F_2. Thus the committee can be comfortable in recommending alternative F_2 as the most suitable data source based on the organization requirements.

4.2. Comparison between TFN-based decision-making approach and fuzzy AHP

Amongst the multi-criteria decision-making methodologies, fuzzy AHP, due to its compatibility and paired comparisons, provides efficient results. In fuzzy AHP, information is decomposed into a hierarchy of alternatives and criteria for decision-making. In this section, the proposed framework is implemented using fuzzy AHP, and is such to determine its significance, it is compared with the results obtained from the fuzzy-based decision-making approach.

As figure 4 shows, similar to the developed FMCDM method, the highest weights are for the complex calculations, running time, and performance criteria. As figure 5 shows, rankings of models by the developed FMCDM method and the fuzzy AHP method present the same results.
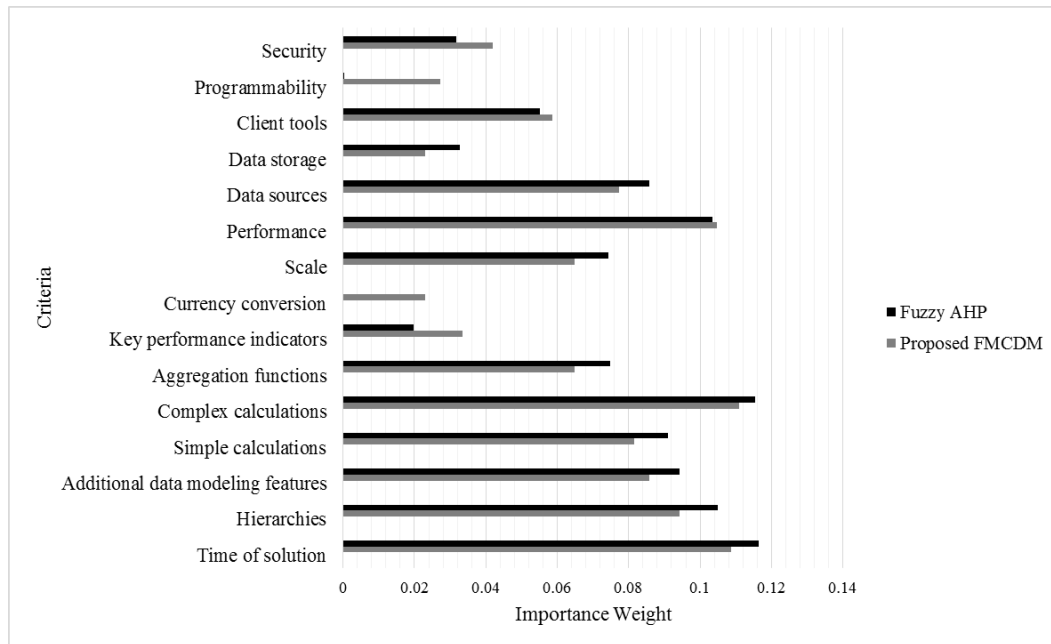

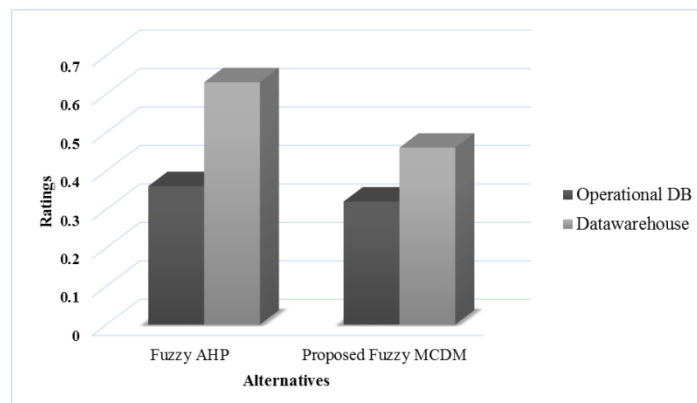**Figure 4. Importance weights of criteria for selected data source.**

Figure 3. Rating alternatives with FMCDM (TFN-based) and Fuzzy AHP methods.

Furthermore, sensitivity analysis graphs allow for ranking the alternatives to be examined, whilst changing the criteria's importance. For example, the need of the organization to perform complex calculations (C_5) is 11%; as a result data warehouse has more capability rather than operational database for this criterion.

Data warehouse and relational database functionalities differ in a number of situations. For example, due to the weakness of the data warehouse in data sources criterion, if the organization reduces this criterion's importance from 78% to 23%, then instead of data warehouse, operational database will be selected.

5. Results and discussion

The use of data warehouses and operational databases in different situations was shown, Operational databases possess a number of disadvantages including the requirement of writing complex queries, slow response rates to queries, limitations in supporting functions and operations, and insufficient memory for answering queries. However, data warehouses with multi-dimensional structures have distinctive functionalities compared to the operational databases. As shown in this study, in data sources, query optimization plays a key role in performance, where each one of these data sources, for each criterion possesses different

abilities. One of the reasons for the organizations failing to create suitable data sources without extra cost and time is the lack of suitable or optimal decision-making. In this study, an effective framework was proposed based on the organization requirements and comparison of the characteristics of data sources such as their models, business logic, data access and storage and security for suitable decision-making for deciding either the generation of a data warehouse or for the use of the existing operational databases. Fuzzy Delphi is used to ensure the selection of appropriate criteria, and as a result, 15 criteria were chosen.

In this paper, the authors proposed an alternative fuzzy-based multi-criteria decision-making model and compared it with Fuzzy AHP. The efficiency of this method was proved through a case study. The results obtained showed that complex calculations (0.116) and solution time (0.116) have higher weightings. Finally, the results obtained show that the decisions made through this framework are the same when implemented using the fuzzy-based decision-making and fuzzy AHP methods. The results of ranking by means of FAHP and the proposed MCDM method show that operational databases perform best, followed by data warehouses.

References

[1] Daneshpour, N. & Abdollahzadeh Barforoush, A. (2007). AUT-QPM: the framework to justify data warehouse systems. 3th International Computer Engineering Conference on Smart Applications for the Information Society, Cairo, Egypt, 2007.

[2] Hwang, C. L. & Yoon, K. (1981). Multiple Attribute Decision Making: Methods and Application. Berlin, Heidelberg. New York: Springer-Verlag.

[3] Zadeh, L. A. (1965). Fuzzy sets. Information and Control. vol. 8, no. 8, pp. 338-353.

[4] Wang, J. W., Cheng, C. H., & Cheng, H. K., (2009). Fuzzy Hierarchical TOPSIS for Supplier Selection. Applied Soft Computing, vol. 9, no. 1, pp. 377-386.

[5] Lan, H. E. & Congbo, L. I. (2009). Method for selecting ERP system based on fuzzy set theory and analytical hierarchy process. Global Congress on Intelligent Systems, Xiamen, China, 2009.

[6] Mamaghani, F. (2002). Evaluation and selection of antivirus and content filtering software. Information Managment and Computer Security, vol. 10, no 1, pp. 28-32.

[7] Liang, G. S. & Wang, M. J. J. (1993). A fuzzy multi-criteria decision making approach for robot selection. Robotics and Computer-Integrated Manufacturing, vol. 10, no 4, pp. 267-274.

[8] Vijayalakshmi, S., Zayaraz, G. & Vijayalakshmi, V. (2010). Multicriteria decision analysis method for evaluation of software architectures. International Journal of Computer Applications, vol. 1, no 25, pp. 33-38.

[9] Eldrandaly, K. & Naguib, S. (2013). A knowledge-based system for GIS software selection. International Arab Journal of Information Technology, vol. 10, no. 2, pp. 152-159.

[10] LeBlance, L. A. & Jelassi, M. (1989). DSS software selection: a multiple criteria decision methodology. Information and Management, vol. 17, no. 1, pp. 49-65.

[11] Phillips, G. E., & Hahnb, E. D. & Forgionne, G. A. (2004). A multiple-criteria framework for evaluation of decision support systems. Omega International Journal of Managment Science, vol. 32, no. 4, pp. 323-332.

[12] Huang, H., Cheng, H., Wang, Q. I. & Shao, Y. (2007). A research of the implementation readiness model for ERP. 6th Wuhan International Conference on E-Business, Wuhan, China, 2007.

[13] Lin, H. Y., Hsu, P. Y. & Sheen, G. J. (2007). A fuzzy-based decision-making procedure for data warehouse system selection. Expert systems with Applications, vol. 32, no. 3, pp. 939-953.

[14] Cochran, J. K. & Chen, H. N. (2005). Fuzzy multi-criteria selection of object-oriented simulation software for production system analysis. Computers & Operations Research, vol. 32, no. 1, pp. 153-168.

[15] Saaty, T. L. (1980). The analytic hierarchy process. McGraw-Hill, New York.

[16] Oğuztimur, S. (2011) Why Fuzzy Analytic Hierarchy Process Approach for transport problems? Available:http://wwwsre.wu.ac.at/ersa/ersaconfs/ersa11/e110830aFinal00438.pdf

[17] Pressman, R. (2009). Software Engineering: A Practitioner's Approach. McGraw-Hill, Science/ Engineering/ Math.

[18] Kimball, R. (1996). The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses. Wiley.

[19] Chen, S. H. (1985). Ranking fuzzy numbers with maximizing set and minimizing set. Fuzzy Sets and Systems, vol. 17, no. 2, pp. 113-129.

[20] Dalkey, N. C. & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. Managment Science, vol. 9, no. 3, pp. 458-467.

[21] Russo, M., Ferrari, A. & Webb, C. (2012). Microsoft SQL Server 2012 Analysis Services: The BISM Tabular Model. Microsoft Press.

[22] LeBlance, P. (2013). Microsoft SQL Server 2012 Step by Step. Microsoft Press.

روش‌های تصمیم‌گیری چندمعیاره فازی برای گزینش منابع داده

الهام آژیر¹، نگین دانشپور²،*٭ و شیرین قنبری³

¹ دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلای قزوین، قزوین، ایران.

٭²، دانشکده مهندسی کامپیوتر، دانشگاه تربیت دبیر شهید رجایی، تهران، ایران.

² مرکز تحقیقات فنی سازمان صدا و سیمای جمهوری اسلامی ایران، تهران، ایران.

چکیده:

ارزیابی و گزینش فناوری مناسب در سازمان‌ها، تأثیر قابل‌توجهی در فرآیندهای انتقال فناوری آنها دارد. انتخاب فناوری مناسب در سازمان‌ها تصـمیم مهمی است که از سوی گروهی از خبرگان صورت می‌گیرد، و بنابراین یکپارچه‌سازی این دیدگاه‌ها دشـوار اسـت. امـروزه، سـازمان‌هـا بـرای مـدیریت و سازمان‌دهی داده‌های خود از پایگاه داده‌های عملیاتی و یا تحلیلی استفاده می‌کنند که به‌دلیل ویژگی‌های متمایز هر مدل از جمله مـدت زمـان توسـعه، سطح پیچیدگی پرس‌وجوها، مدت زمان پاسخگویی و توابع متفاوت، رویکرد مناسبی برای گزینش آنها با توجه به نیاز سازمان مـورد نیـاز اسـت. در ایـن راستا، روش‌های تصمیم‌گیری چندمعیاره فازی برای گزینش منبع داده‌ای مناسب (پایگاه داده تحلیلی و پایگاه داده عملیاتی رایج) بـر اسـاس نیازهـای سازمان مورد بررسی قرار گرفته‌اند. در این مقاله فرآیند تصمیم‌گیری مبنی بر اعداد فازی مثلثی (TFN) و فرآیند تحلیل سلسله مراتبـی فـازی (Fuzzy AHP) با استفاده از معیارهای مدل داده، منطق کسب‌وکار، دسترسی به داده، ذخیره‌سازی و امنیت بررسی و مقایسه شده‌اند. نتایج نشان مـی‌دهـد کـه هر دو فرآیند منابع داده‌ای را به ترتیب مشابه و صحیحی انتخاب کرده‌اند.

کلمات کلیدی: فرآیند تحلیل سلسله مراتبی، پایگاه داده تحلیلی، فازی، تصمیم‌گیری چندمعیاره، پایگاه داده عملیاتی.