A Fuzzy Based Matrix Methodology for Evaluation and Ranking of Data Warehouse Conceptual Models Metrics

Naveen Dahiya[1], Vishal Bhatnagar[2], and Manjeet Singh[3]
[1]Maharaja Surajmal Institute of Technology, C-4, Janakpuri, India
[2]Ambedkar Institute of Advanced Communication Technology and Research, India
[3]YMCA University of Science and Technology, Sector-6, India

Abstract: *The authors present a methodology for ranking data warehouse conceptual models metrics based on opinion of experts using fuzzy inference technique. The fuzzy based approach gives a precise ranking methodology due to its ability to handle imprecise data involved in ranking of metrics and ambiguity involved in expert decision making process. The proposed work aims towards ranking of quality metrics already proposed and validated by Manuel Serrano along certain identified parameters based on expert opinion and evaluation of criteria matrix using permanent function. The results obtained are also compared with the actual experts ranking. The achieved results are better as the imprecise human thinking is taken into consideration during calculation of results to give realistic results.*

Keywords: *Fuzzy, data warehouse, conceptual models, quality metrics, criteria matrix.*

Received October 23, 2014; accepted July 7, 2015

1. Introduction

In modern times, it is the data that plays a crucial role in success or failure of any sort of systems. The data is stored in large repositories called data warehouses. An efficient data warehouse system provides valuable information that can analyze the past trends, relate it to current scenario and predict actual futuristic trends. It is the quality of data warehouse that decides the success or failure of systems. There exists a need to design efficient data warehouse to gain competitive advantage in today's competitive environment.

The design process of data warehouse is a three step process starting from conceptual phase, to logical phase and finally physical design phase. To build an efficient data warehouse system, quality objectives needs to be set from initial design phase to final implementation of data warehouse. It is rightly said that a strong base leads to the foundation of a robust building erected on it.

Several design techniques like starER , dimension-fact modeling , object oriented dimensional modeling have been proposed [29] for conceptual data warehouse design. Various quality metrics have been proposed for quality evaluation of each type of conceptual design techniques. The quality metrics are based on size and structural complexity of conceptual data warehouse models. There exist multiple criteria like understandability, efficiency, effectiveness along which quality of conceptual schemas can be measured using quality metrics. There needs to be a methodology for precise ranking of quality metrics towards building of good quality conceptual models. The ordering of metrics can be one of the major considerations during design of conceptual data warehouse models.

Ranking of metrics along variable criteria leads to multiple-criteria decision making problem. The criteria are defined qualitatively and the significance of quality metrics along the criteria varies according to user requirements, situations and expert opinion. A fuzzy based ranking system should be evolved to deal with imprecise and qualitative (non-numeric) data based on actual human (expert) decision making.

The organization of paper is as follows: section 2 describes the related research carried out in conceptual data warehouse domain and ranking of quality metrics in the same domain. Section 3 gives a overview of the basic concepts used in the paper for ranking quality metrics. Section 4 outlines the research methodology followed in the present study. Section 5 gives the application of proposed fuzzy methodology for ranking of quality metrics using an illustrative example. In section 6 the results obtained are analyzed and a comparison of results obtained is made with other ranking methodology. The paper concludes with section 7 that presents the merits of proposed fuzzy methodology along with future scope for researchers in the connected domain.

2. Related Research

The conceptual schemas provide the base for development of data warehouse systems. An efficient conceptual schema leads to the design of best data warehouse systems in terms of information delivery.

The quality of conceptual schema [10] depends on quality metrics. There must be some mechanism for evaluating the significance of each quality metric towards quality prediction of conceptual schemas. If we have some procedure for ranking the quality metrics in terms of certain identified criteria, then due consideration can be given to certain metrics in design of efficient conceptual data warehouse schemas [11]. Based on the above considerations, our review in the paper is broadly categorized as follows:

- Study of conceptual modeling design techniques and quality metrics for measuring the ability of conceptual schemas in terms of criteria (understandability, effectiveness and efficiency).
- Study of various techniques for ranking of software metrics.

Some of the techniques for design of conceptual multidimensional models are as follows: starER Model Tryfona *et al*. [39], Multidimensional/ER (M/ER) Model Sapia *et al*. [34], CGMD Model Kamble [21], DWDCM Franconi and Sattler [13], REMDM Model Perez *et al*.[31], EGOLD Model Gamel *et al*. [14], OOMD Model Trujillo and Palomar[38], Yet Another Multidimensional Model (YAM2) Abello *et al*. [1], The Dimensional Fact (DF) Model Golfarelli *et al*. [18], MAC Model Tsois *et al*. [40].

A number of quality metrics have been proposed by various authors [7, 17, 24, 30, 32, 35] based on the size and structure of conceptual data warehouse schema. Serrano *et al*. [36] discuss eleven metrics. Singh *et al*. [37] discuss several software quality metrics, their fault proneness along with their empirical validation.

Zhang and Pham [44] conducted an empirical research on data collected from managers, system engineers, programmers and testers of top 13 companies. Based on collected data 32 factors were identified that were involved in every phase of software development. Two techniques namely relative weight method and analysis of variance were used to analyze and rank the identified factors affecting software reliability.

Li and Smidts [26] conducted a study on thirty identified potential factors affecting software design and reliability. The ranking score for each factor was elicited based on expert opinion to identify and rank the factors in terms of their potential significance. Johnson and Yu [20] proposed a software quality model, based on Bayesian Belief Network (BBN), to predict software reliability through analysis of software metrics. Dyba [12] identified and ranked key factors involved in software process based on expert opinion. To estimate cost effectiveness of software model Briand *et al*. [5] proposed an approach based on expert opinion.

In all of the proposed techniques algebraic aggregation has been used to quantify scores of expert opinion with no consideration of uncertainties, ambiguities and biases in human thought process. Also no technique has considered interdependencies of attributes for ranking of attributes. So the results of the above discussed techniques lack accuracy and rigidness. From the study of literature, authors identified the need of a systematic ranking approach that considers uncertainties, ambiguities, biases involved in human thought process and takes into account all possible interdependencies of attributes involved. The only viable solution that can handle analysis of expert opinion in a more better, consistent and flexible way is fuzzy based ranking approach. Fuzzy set theory [2, 3, 4, 7, 8, 9, 15, 17, 19, 23, 24, 25, 27, 30, 31, 32, 33, 35, 41, 42] is new emerging technique finding its application in diverse domains.

3. Preliminaries

The paper aims to present a precise methodology based on fuzzy logic and matrix operations to rank quality metrics of conceptual data warehouse models. As no other technique for ranking of quality metrics for data warehouse conceptual models exists that takes into consideration actual human thought process to give realistic results against certain identified parameters. This is the first paper of its kind to rank conceptual model metrics along identified parameters based on expert opinion. The basics of fuzzy logic, linguistic variables and matrix functions are present subsequently.

3.1. Introduction to Fuzzy sets

The concept of fuzziness shows uncertainty, imprecision, ambiguity, inconsistency, vagueness of situations. Zadeh [43] introduced a theory whose objects, fuzzy sets, are the sets with no precise boundaries. The fuzzy sets are capable to show gradual transition from membership to non-membership and vice-versa. The range of membership functions is the unit interval [0, 1]. The membership function of a fuzzy set A is denoted by μ_A,

$\mu_A : X \rightarrow [0,1]$, where X is a universal set

Degree of membership is 0 when the element is not in set; degree of membership is 1 when the element is in the set. A value between 0-1 shows the ambiguity of membership.

3.2. Triangular Fuzzy Membership Functions

The use and application of the membership functions depends on the scenario to which it is applied. The most commonly used is the triangular function [43] due to its ease of use in calculations. So we have used triangular function in this study. The triangular fuzzy membership function [22], denoted as $\wedge$: $x \rightarrow [0, 1]$ is defined as follows by Equations (1, 2, 3) :

$$\wedge\,(x{:}a,b,c) \text{ when } a \le x \le b = \{x - a/x - b\} \tag{1}$$

$$\wedge\,(x{:}a,b,c) \text{ when } b \le x \le c = \{c - x/c - b\} \tag{2}$$

$$\wedge\,(x{:}a,b,c) \text{ otherwise} = \{0\} \tag{3}$$

Where a, b, c are real numbers $a \le b \le c$. The triangular fuzzy function can be represented graphically as shown in Figure 1:

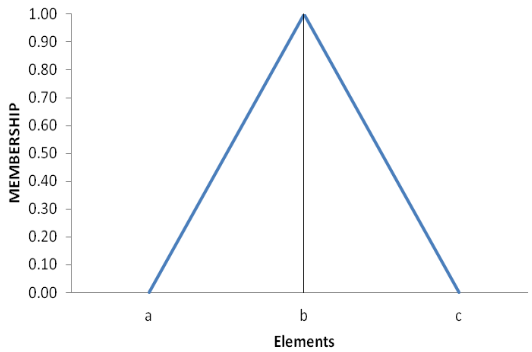

Figure 1. Triangular fuzzy membership function graph.

3.3. Fuzzy Linguistic Variables

Fuzzy logic theory involves the uncertainty and ambiguity in human thought process and quantify it in terms of lingual terms. The natural linguistic terms used in common usage are closer to human perceptions and thoughts than crisp numeric values. A linguistic variable is some non-numeric syllable/term used in natural usage. In this paper, the authors have used various linguistic variables to weight the criteria and rate the metrics.

The weights assigned to specified criteria are evaluated in terms of linguistic variables High (H), Medium (M), Low (L). The membership values for each of the linguistic variables is expressed as H(0.5,0.8, 1), M(0.3,0.5,0.8), L(0,0.3,0.5) as shown in Table 1 and the corresponding membership graph is shown in Figure 2.

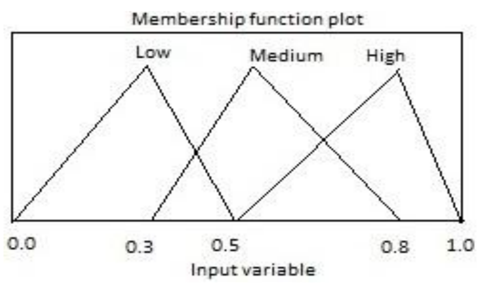

Figure 2. Fuzzy membership graph for weighting criteria.

Table 1. Fuzzy membership values for weights assigned to criteria.

Linguistic Variable	**High(H)**	**Medium(M)**	**Low(L)**
Fuzzy membership	(0.5,0.8, 1)	(0.3,0.5,0.8)	(0,0.3,0.5)

Similarly, the ratings assigned to quality metrics are expressed in terms of linguistic variables Very Good (VG), Good (G), Medium (M), Poor (P), and Very Poor (VP). The triangular fuzzy membership values are assigned to the variables as shown in Table 2 and the membership graph is shown by Figure 3.

Table 2. Fuzzy membership values for rating assigned to quality metrics.

Linguistic Variable	**Very Good(VG)**	**Good(G)**	**Medium(M)**	**Poor(P)**	**Very Poor(VP)**
Fuzzy membership	(0.7,1,1)	(0.5,0.7,1)	(0.2,0.5,0.7)	(0,0.3,0.5)	(0,0,0.3)

3.4. Quality Metrics Ranking Problem With Efficient Fuzzy Solution

The section defines the methodology for ranking quality metrics of conceptual data warehouse model using fuzzy based approach and multi-criteria analysis.

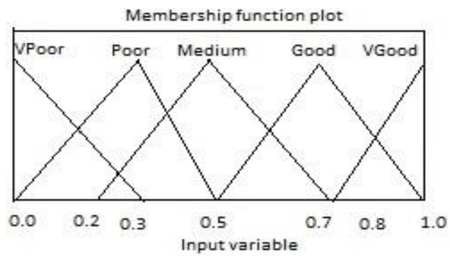

Figure 3. Fuzzy membership graph for rating quality metrics.

The quality metrics ranking problem [16] and its multi criteria fuzzy solution can be defined as:

A team of n experts (E_1, E_2, E_3,…, E_n), has to analyze and grant weights to k criteria (C_1, C_2, C_3,…,C_k) and the ratings to m quality metrics (Q_1, Q_2, Q_3,…, Q_m) for each of the k criteria. Let W_{ij} (i=1, 2, 3,…,k; j=1, 2, 3,…,n) be the weight assigned to criteria C_i by expert Ej. Let R_{ijt} (i=1, 2, 3,…,m; j=1, 2, 3,…,n; t=1, 2, 3,…,k) be the rating given to metric Q_i by expert E_j under criteria C_t.

$$W_i = \frac{1}{n} \otimes (\mathrm{Wi1} \oplus \mathrm{Wi2} \oplus \ldots \oplus \mathrm{Win}) \tag{4}$$

$$R_{ij} = \frac{1}{n} \otimes (\mathrm{Rij1} \oplus \mathrm{Rij2} \oplus \ldots \oplus \mathrm{Rijn}) \tag{5}$$

Where W_i is the average weight of criteria and R_{ij} is the aggregated rating of quality metric Q_i under criteria C_j as shown by Equation (4) and Equation (5).Also for defuzzification (conversion of fuzzy aggregations to crisp scores) [6] we have applied area of centroid method due to ease of application and usage.

3.5. Criteria Matrix

Each of the quality metrics has multiple rating scores corresponding to expert evaluation along several criteria. The crisp scores for each metric are achieved using Criteria matrix. A criteria matrix [16] is aggregation of metric rating along multiple criteria and aggregated relative weights of each criteria. The order of criteria matrix is n×n, where n is the number of criteria for metric evaluation. The diagonal elements of criteria matrix show the aggregated rating of a metric along multiple criteria and the off diagonal elements represent the relative aggregated weights of multiple criteria. Thus, a criteria matrix is a combination of two matrix. One is metric rating matrix and other is relative weight matrix.

- *Metric rating matrix*: This is a diagonal matrix is a n×n matrix, whose elements are the aggregated rankings of a metric evaluated along multiple criteria.

$$\begin{bmatrix} a11 & 0 & \dots & 0 \\ 0 & a22 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & ann \end{bmatrix}$$

- *Relative weight matrix*: This is a n×n matrix, whose diagonal elements are all 0's, and whose off diagonal elements gives the aggregated relative weights of criteria. In mathematical terms, an element a_{ij} of the relative weight matrix equals weight of criteria j divided by weight of criteria I as given by Equation (6).

$$a_{ij} = \frac{Weightj}{Weighti} \tag{6}$$

$$\begin{bmatrix} 0 & a12 & \dots & a1n \\ a21 & 0 & \dots & a2n \\ \dots & \dots & \dots & \dots \\ an1 & an2 & \dots & 0 \end{bmatrix}$$

Thus the criteria matrix which is a combination of metric rating matrix and relative weight matrix is as follows:

$$\begin{bmatrix} a11 & a12 & \dots & a1n \\ a21 & a22 & \dots & a2n \\ \dots & \dots & \dots & \dots \\ an1 & an2 & \dots & ann \end{bmatrix}$$

3.6. Permanent of Matrix

Permanent [28] of a matrix is an important technique for ranking of systems based on multi-criteria evaluation. The permanent is similar to determinant with the only difference that no negative term appears in calculation of permanent. In mathematical terms, permanent [28] is given by Equation (7):

For a square matrix M (order n) = $[m_{ij}]_{1\leq i,j\leq n}$

$$\text{perm}(M) = \sum_{\Pi\in S} \prod_{i\in I} M(i) \tag{7}$$

Where S consists of the group of symmetric elements S_n.

3.7. Expert Opinion Ranking Methodology

We, in this paper, have compared the results of proposed fuzzy methodology with aggregations of expert opinion methodology [26]. The input given for calculation of expert opinion methodology is algebraic aggregation of linguistic membership functional data collected from experts. Ranking problem and its expert opinion solution can be stated as:

A team of m experts (E_1, E_2, E_3,…, E_m), has to analyze and grant weights to l criteria (C_1,C_2,C_3,…,C_l) and the ratings to n quality metrics (Q_1, Q_2, Q_3,…, Q_n) for each of the l criteria. Let r(i,j,k) be the rating given to metric Q_i by expert E_k under criteria C_j and w(j) be the weight evaluated by experts for criteria C_j. The ratings and weights given as input are the mean algebraic aggregation of linguistic membership functional data collected from experts. The aggregated rating R(i) for metric Q_i is calculated using the aggregation mean value function [26] defined as follows by Equation (8):

$$R(i) = \frac{1}{ml} * \sum_{j=1}^{l} \sum_{k=1}^{m} r(i,j,k) * w(j) \tag{8}$$

4. Research Methodology

The research methodology followed by the authors during their current research is shown in Figure 4. The stepwise detail of research methodology shown in Figure 4 is given as follows:

4.1. Identification of Quality Metrics for Conceptual Data Warehouse Models

The research in the paper is carried on the quality metrics proposed by Serrano *et al*. [36] and Dahiya *et al*. [11]. The quality metrics have been theoretically and empirically validated. The detail of metrics is as follows:

- *NFC*: Number of fact classes.
- *NDC*: Number of dimension classes.
- *NBC*: Number of base classes.
- *NC*: Total number of classes which includes fact classes, dimension classes and base classes.
- *RBC*: Number of base classes per dimension class.
- *NAFC*: Number of attributes of the fact class.
- *NADC*: Number of dimension attributes of the dimension classes.
- *NABC*: Number of dimension attributes of the base classes.
- *NA*: Total of all attributes which includes fact class attributes, dimension class attributes and base class attributes.
- *NH*: Number of hierarchy relationships.
- DHP: Maximum depth of the hierarchy relationships.
- *RSA*: Number of fact attributes to the number dimension attributes.
- *NRFD*: Number of relations between fact classes and dimension classes.

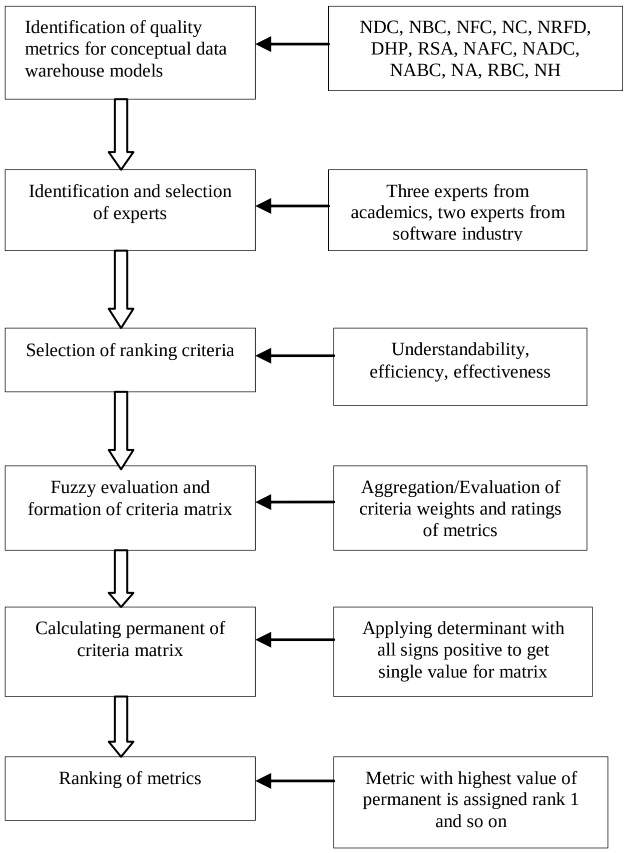

Figure 4. Research methodology.

4.2. Identification and Selection of Experts

One common data collection technique is to prepare questionnaires and conduct a survey based on questionnaire. Various statistical techniques are then applied to data collected. Due to blind nature of statistics, the results vary from one survey to another and cannot be generalized. In the questionnaire survey various possible threats to validity exist like fatigue effects, biased results, motivation effects, learning effects that cannot be avoided. So expert's opinion [44] was identified as the best feasible approach for data collection.

In this study five experts from data warehouse domain having up to date knowledge of technological advances and rich practical hands on experience, with more than 10-20 years of experience were selected and approached. Out of five experts three were from academics and two were from software industry.

4.3. Selection of Ranking Criteria

The quality metrics can be evaluated collectively in Enter terms of several parameters termed as ranking criteria. Three ranking parameters have been considered for ranking the quality metrics of conceptual data warehouse models. The identified parameters [36] are as follows:

- *Understandability*: It is defined as the time taken to understand a conceptual schema and perform tasks (answer questions) based on understanding of the conceptual schema.
- *Efficiency*: It is defined as the number of correct tasks performed per unit time based on the understandability of conceptual schema.
- *Effectiveness*: It is defined as the number of correct tasks performed per total number of tasks based on size and structural complexity of conceptual schemas.

Each of the identified experts was to fill requisite form for assigning weights and ranking of metrics based on their experience and opinion. The forms are as follows:

Criteria	**Expert Opinion**	How do you weight the criteria towards quality evaluation of conceptual data warehouse schemas in terms of linguistic variables H(0.5,0.8,1), M(0.3,0.5,0.8), L(0,0.3,0.5). High(H), Medium(M), Low(L).
Understandability		
Efficiency		
Effectiveness		

Each expert was to fill the above proforma to assign weights to specified criteria in terms of linguistic variables High (H), Medium (M), Low (L). Second form that assigns rating to metrics versus criteria, filled by each of the experts is as follows:

Metrics	**Understandability**	**Efficiency**	**Effectiveness**	How do you rank the metrics along criteria of understandability , efficiency, effectiveness towards quality evaluation of conceptual data warehouse schemas in terms of linguistic variablesVG(0.7,1,1), G(0.5,0.7,1),M(0.2,0.5 ,0.7), P(0,0.3,0.5), VP(0,0,0.3). Very Good (VG), Good (G), Medium (M), Poor (P), Very Poor (VP)
NA				
NADC				
NRFD				
NBC				
NC				
RBC				
NH				
NDC				
NFC				
DHP				
NABC				
NAFC				
RSA				

4.4. Fuzzy Evaluation and Formation of Criteria Matrix

Fuzzy evaluation of expert's opinion is a incremental stepwise process. Firstly, experts evaluate weight of each identified criteria and give ratings to metrics versus criteria in terms of fuzzy linguistic variables. Then the weights and ratings are aggregated. The aggregations are then converted to crisp scores to form a criteria matrix for each of the metrics (combination of weights and ratings).

4.5. Calculating Permanent of Criteria Matrix

A permanent function (determinant with all signs positive) is calculated for each of the criteria matrix build up in the previous step. A permanent gives the single value for the entire criteria matrix.

4.6. Ranking of Metrics

The matrix with the highest value of permanent, calculated in previous step, is ranked to number 1 and subsequently to number 2, 3 and so on.

5. Practical Application

The authors present an example to illustrate the fuzzy based methodology discussed in the above sections. Thirteen quality metrics namely NDC, NFC, NBC, NC, NRFD, DHP, RSA, NH, NAFC, NADC, NABC, NA, RBC are ranked, based on three ranking criteria namely understandability, efficiency and effectiveness.

The weights assigned to three ranking criteria and ratings of thirteen quality metrics versus each ranking criteria are assigned in terms of linguistic fuzzy variables by each of the five experts presented in Tables 3 and 4.

Table 3. Fuzzy membership values and linguistic representation for ranking criteria.

Criteria	E1	E2	E3	E4	E5
Understandability	H(0.5,0.8, 1)	H(0.5,0.8, 1)	H(0.5,0.8, 1)	H(0.5,0.8, 1)	H(0.5,0.8, 1)
Efficiency	H(0.5,0.8, 1)	H(0.5,0.8, 1)	M(0.3,0.5, 0.8)	H(0.5,0.8, 1)	M(0.3,0.5,0.8)
Effectiveness	M(0.3,0.5, 0.8)	H(0.5,0.8, 1)	M(0.3,0.5, 0.8)	L(0,0.3, 0.5)	L(0,0.3, 0.5)

A fuzzy triangular aggregation, as already discussed, is used for further aggregation.

Using fuzzy triangular aggregation, the aggregated weights (W_t) and aggregate ratings (R_{it}) of A_i metric under criteria C_t were calculated as shown in Table 5 and Table 6. For example, the aggregated weight of criteria C_1 (understandability) was calculated as:

$$C_1 = \tfrac{1}{5} \otimes (0.5, 0.8, 1 \oplus 0.5, 0.8, 1 \oplus 0.5, 0.8, 1 \oplus 0.5, 0.8, 1 \oplus 0.5, 0.8, 1)$$

$$= 1/5(2.5, 4.0, 5) = (0.5, 0.8, 1)$$

Table 4. Fuzzy membership values and linguistic representation for quality metrics.

Metrics		Understandability	Efficiency	Effectiveness
NA	E1	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	G(0.5,0.7, 1)
	E2	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	G(0.5,0.7, 1)
	E3	G(0.5,0.7, 1)	G(0.5,0.7, 1)	M(0.2,0.5,0.7)
	E4	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	P(0,0.3,0.5)
	E5	G(0.5,0.7, 1)	P(0,0.3,0.5)	P(0,0.3,0.5)
NADC	E1	G(0.5,0.7, 1)	M(0.2,0.5,0.7)	P(0,0.3,0.5)
	E2	G(0.5,0.7, 1)	P(0,0.3,0.5)	P(0,0.3,0.5)
	E3	M(0.2,0.5,0.7)	P(0,0.3,0.5)	VP(0,0,0.3)
	E4	P(0,0.3,0.5)	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)
	E5	M(0.2,0.5,0.7)	P(0,0.3,0.5)	P(0,0.3,0.5)
NRFD	E1	G(0.5,0.7, 1)	G(0.5,0.7, 1)	G(0.5,0.7, 1)
	E2	VG(0.7,1,1)	G(0.5,0.7, 1)	M(0.2,0.5,0.7)
	E3	G(0.5,0.7, 1)	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)
	E4	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	G(0.5,0.7, 1)
	E5	G(0.5,0.7, 1)	G(0.5,0.7, 1)	M(0.2,0.5,0.7)
NBC	E1	VG(0.7,1,1)	G(0.5,0.7, 1)	G(0.5,0.7, 1)
	E2	VG(0.7,1,1)	VG(0.7,1,1)	VG(0.7,1,1)
	E3	G(0.5,0.7, 1)	VG(0.7,1,1)	VG(0.7,1,1)
	E4	VG(0.7,1,1)	G(0.5,0.7, 1)	G(0.5,0.7, 1)
	E5	VG(0.7,1,1)	VG(0.7,1,1)	VG(0.7,1,1)
NC	E1	VG(0.7,1,1)	VG(0.7,1,1)	VG(0.7,1,1)
	E2	VG(0.7,1,1)	G(0.5,0.7, 1)	VG(0.7,1,1)
	E3	VG(0.7,1,1)	VG(0.7,1,1)	G(0.5,0.7, 1)
	E4	VG(0.7,1,1)	G(0.5,0.7, 1)	VG(0.7,1,1)
	E5	VG(0.7,1,1)	VG(0.7,1,1)	G(0.5,0.7, 1)
RBC	E1	G(0.5,0.7, 1)	G(0.5,0.7, 1)	G(0.5,0.7, 1)
	E2	G(0.5,0.7, 1)	M(0.2,0.5,0.7)	P(0,0.3,0.5)
	E3	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)
	E4	M(0.2,0.5,0.7)	G(0.5,0.7, 1)	P(0,0.3,0.5)
	E5	G(0.5,0.7, 1)	M(0.2,0.5,0.7)	G(0.5,0.7, 1)
NH	E1	VG(0.7,1,1)	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)
	E2	G(0.5,0.7, 1)	P(0,0.3,0.5)	P(0,0.3,0.5)
	E3	M(0.2,0.5,0.7)	G(0.5,0.7, 1)	G(0.5,0.7, 1)
	E4	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	P(0,0.3,0.5)
	E5	M(0.2,0.5,0.7)	P(0,0.3,0.5)	M(0.2,0.5,0.7)
NDC	E1	G(0.5,0.7, 1)	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)
	E2	M(0.2,0.5,0.7)	P(0,0.3,0.5)	P(0,0.3,0.5)
	E3	P(0,0.3,0.5)	M(0.2,0.5,0.7)	P(0,0.3,0.5)
	E4	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	P(0,0.3,0.5)
	E5	M(0.2,0.5,0.7)	G(0.5,0.7, 1)	M(0.2,0.5,0.7)
NFC	E1	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	G(0.5,0.7, 1)
	E2	M(0.2,0.5,0.7)	P(0,0.3,0.5)	M(0.2,0.5,0.7)
	E3	P(0,0.3,0.5)	G(0.5,0.7, 1)	P(0,0.3,0.5)
	E4	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	P(0,0.3,0.5)
	E5	P(0,0.3,0.5)	P(0,0.3,0.5)	P(0,0.3,0.5)
DHP	E1	M(0.2,0.5,0.7)	P(0,0.3,0.5)	P(0,0.3,0.5)
	E2	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	P(0,0.3,0.5)
	E3	G(0.5,0.7, 1)	G(0.5,0.7, 1)	M(0.2,0.5,0.7)
	E4	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)
	E5	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)
NABC	E1	VP(0,0,0.3)	M(0.2,0.5,0.7)	M(0.2,0.5,0.7)
	E2	P(0,0.3,0.5)	P(0,0.3,0.5)	P(0,0.3,0.5)
	E3	M(0.2,0.5,0.7)	P(0,0.3,0.5)	VP(0,0,0.3)
	E4	P(0,0.3,0.5)	VP(0,0,0.3)	P(0,0.3,0.5)
	E5	P(0,0.3,0.5)	M(0.2,0.5,0.7)	P(0,0.3,0.5)
NAFC	E1	VP(0,0,0.3)	M(0.2,0.5,0.7)	VP(0,0,0.3)
	E2	VP(0,0,0.3)	VP(0,0,0.3)	M(0.2,0.5,0.7)
	E3	P(0,0.3,0.5)	P(0,0.3,0.5)	P(0,0.3,0.5)
	E4	VP(0,0,0.3)	P(0,0.3,0.5)	P(0,0.3,0.5)
	E5	VP(0,0,0.3)	P(0,0.3,0.5)	M(0.2,0.5,0.7)
RSA	E1	VP(0,0,0.3)	M(0.2,0.5,0.7)	VP(0,0,0.3)
	E2	VP(0,0,0.3)	VP(0,0,0.3)	P(0,0.3,0.5)
	E3	VP(0,0,0.3)	VP(0,0,0.3)	P(0,0.3,0.5)
	E4	P(0,0.3,0.5)	P(0,0.3,0.5)	VP(0,0,0.3)
	E5	VP(0,0,0.3)	VP(0,0,0.3)	VP(0,0,0.3)

Table 5. Aggregated weights for criteria ranking.

Criteria	Understandability	Efficiency	Effectiveness
W_t	0.5,0.8,1	0.42,0.68,0.88	0.22,0.48,0.72

Likewise the aggregated rating of metric A_1 (NA) under criteria C_1 (understandability) was calculated as:

$$A_{11} = \tfrac{1}{5} \otimes (0.2, 0.5, 0.7 \oplus 0.2, 0.5, 0.7 \oplus 0.5, 0.7, 1 \oplus 0.2, 0.5, 0.7 \oplus 0.5, 0.7, 1)$$

$$= 1/5(1.6, 2.9, 4.1) = (0.32, 0.58, 0.82)$$

The results of aggregated ratings are shown in Table 6.

Table 6. Aggregated rating for quality metrics.

Metrics	Understandability	Efficiency	Effectiveness
NA	0.32,0.58,0.82	0.22,0.5,0.72	0.24,0.5,0.74
NADC	0.28,0.54,0.78	0.08,0.38,0.58	0.04,0.28,0.5
NRFD	0.48,0.72,0.94	0.38,0.62,0.88	0.32,0.58,0.82
NBC	0.66,0.94,1	0.62,0.88,1	0.62,0.88,1
NC	0.7,1,1	0.62,0.88,1	0.62,0.88,1
RBC	0.38,0.62,0.88	0.32,0.58,0.82	0.24,0.5,0.74
NH	0.36,0.64,0.82	0.18,0.46,0.68	0.18,0.46,0.68
NDC	0.22,0.5,0.72	0.22,0.5,0.72	0.08,0.38,0.58
NFC	0.12,0.42,0.62	0.18,0.46,0.68	0.12,0.42,0.62
DHP	0.26,0.54,0.76	0.22,0.5,0.72	0.12,0.42,0.62
NABC	0.04,0.28,0.5	0.08,0.32,0.54	0.04,0.28,0.5
NAFC	0,0.06,0.34	0.04,0.28,0.5	0.08,0.32,0.54
RSA	0,0.06,0.34	0.04,0.16,0.42	0,0.12,0.38

The crisp scores of these aggregated values are then calculated using methods described in section above and shown in Table 7.

Table 7. Values of crisp scores for rating quality metrics.

Metrics	Understandability	Efficiency	Effectiveness
NA	0.5733	0.48	0.4933
NADC	0.5333	0.3466	0.2733
NRFD	0.7133	0.6266	0.5733
NBC	0.8666	0.8333	0.8333
NC	0.9	0.8333	0.8333
RBC	0.6266	0.5733	0.4933
NH	0.6066	0.44	0.44
NDC	0.48	0.48	0.3466
NFC	0.3866	0.44	0.3866
DHP	0.52	0.48	0.3866
NABC	0.2733	0.3133	0.2733
NAFC	0.1333	0.2733	0.3133
RSA	0.1333	0.2066	0.1666
Criteria	0.7666	0.66	0.4733

The criteria matrix is formed for each quality metric and the value of permanent for each criteria matrix is calculated. For example the criteria matrix for metric NA is constructed as follows:

$$\begin{bmatrix} 0.5733 & 0.8609 & 0.6174 \\ 1.1615 & 0.48 & 0.7171 \\ 1.6196 & 1.3944 & 0.4933 \end{bmatrix}$$

The value of permanent obtained using criteria matrix is then used to rank the quality metrics. The calculated rank values and rank of each quality metric is shown in Table 8.

Table 8. Ranking values and rank of quality metrics.

Metrics	Ranking Values	Rank
NA	3.6815	6
NADC	3.0704	11
NRFD	4.1686	3
NBC	5.1340	2
NC	5.1906	1
RBC	3.8696	4
NH	3.6539	7
NDC	3.3875	9
NFC	3.2783	10
DHP	3.4823	8
NABC	3.7627	5

6. Result Analysis and Comparison

The quality metrics have been ranked in accordance to their significance towards predicting the understandability, efficiency and effectiveness of conceptual data warehouse schemas.

As can be seen from Table 8, the metrics with higher value of permanent are ranked higher in order. The metric NC has been ranked as first owing to its high score for three criteria namely understandability, efficiency and effectiveness. The metric with the highest value of permanent is having a greater impact and is therefore ranked high. The calculation of permanent has already been discussed. The metric NC is followed by NBC with a slight difference in values of permanent i.e., 5.1340 for NBC and 5.1906 for NC. The metric NBC is followed by NRFD with a score of 4.1686. The successive metrics in order are RBC, NABC, NA, NH, DHP, NDC, NFC, NADC with a score of 3.8696, 3.7627, 3.6815, 3.6539, 3.4823, 3.3875, 3.2783, 3.0704. The score for permanent of these metrics vary in fractions showing their relatively similar significance on the quality of conceptual models. The metrics NAFC and RSA have lowest ranks due to their low score of permanent. This way we can categorize the metrics into 4 groups based on their values of permanent.

G1 = [NC, NBC]

G2 = [NRFD]

G3 = [RBC, NABC, NA, NH, DHP, NDC, NFC, NADC]

G4 = [NAFC, RSA]

The results of proposed fuzzy based approach are compared with the results based on expert opinion [26] can be seen from Table 9.

Table 9. Comparison and analysis with other technique.

Metrics	Ranking Values based on proposed fuzzy method	Rank	Ranking Values based on expert opinion	Rank	Group
NC	**5.1906**	**1**	**0.5447**	**1**	G1
NBC	**5.134**	**2**	**0.5361**	**2**	
NRFD	**4.1686**	**3**	**0.4102**	**3**	G2
RBC	**3.8696**	**4**	**0.3637**	**4**	G3
NABC	3.7627	5	0.1815	11	
NA	3.6815	6	0.3295	5	
NH	3.6539	7	0.3208	6	
DHP	3.4823	8	0.2990	7	
NDC	3.3875	9	0.2825	8	
NFC	3.2783	10	0.2583	9	
NADC	3.0704	11	0.2552	10	
NAFC	**2.7307**	**12**	**0.1433**	**12**	G4
RSA	**2.5106**	**13**	**0.1056**	**13**	

The input data given to ranking based on expert opinion is given in Table 10. The values in Table 10 are calculated using algebraic aggregation as discussed in previous sections.

Table 10. Input to rank based on expert opinion.

Metrics		Understandability	Efficiency	Effectiveness
NA	E1	0.466	0.466	0.733
	E2	0.466	0.466	0.733
	E3	0.733	0.733	0.466
	E4	0.466	0.466	0.266
	E5	0.733	0.266	0.266
NADC	E1	0.733	0.466	0.266
	E2	0.733	0.266	0.266
	E3	0.466	0.266	0.100
	E4	0.266	0.466	0.466
	E5	0.466	0.266	0.266
NRFD	E1	0.733	0.733	0.733
	E2	0.900	0.733	0.466
	E3	0.733	0.466	0.466
	E4	0.466	0.466	0.733
	E5	0.733	0.733	0.466
NBC	E1	0.900	0.733	0.733
	E2	0.900	0.900	0.900
	E3	0.733	0.900	0.900
	E4	0.900	0.733	0.733
	E5	0.900	0.900	0.900
NC	E1	0.900	0.900	0.900
	E2	0.900	0.733	0.900
	E3	0.900	0.900	0.733
	E4	0.900	0.733	0.900
	E5	0.900	0.900	0.733
RBC	E1	0.733	0.733	0.733
	E2	0.733	0.466	0.266
	E3	0.466	0.466	0.466
	E4	0.466	0.733	0.266
	E5	0.733	0.466	0.733
NH	E1	0.900	0.466	0.466
	E2	0.733	0.266	0.266
	E3	0.466	0.733	0.733
	E4	0.466	0.466	0.266
	E5	0.466	0.266	0.466
NDC	E1	0.733	0.466	0.466
	E2	0.466	0.266	0.266
	E3	0.266	0.466	0.266
	E4	0.466	0.466	0.266
	E5	0.466	0.733	0.466
NFC	E1	0.466	0.466	0.733
	E2	0.466	0.266	0.466
	E3	0.266	0.733	0.266
	E4	0.466	0.466	0.266
	E5	0.266	0.266	0.266
DHP	E1	0.466	0.266	0.266
	E2	0.466	0.466	0.266
	E3	0.733	0.733	0.466
	E4	0.466	0.466	0.466
	E5	0.466	0.466	0.466
NABC	E1	0.100	0.466	0.466
	E2	0.266	0.266	0.266
	E3	0.466	0.266	0.100
	E4	0.266	0.100	0.266
	E5	0.266	0.466	0.266
NAFC	E1	0.100	0.466	0.100
	E2	0.100	0.100	0.466
	E3	0.266	0.266	0.266
	E4	0.100	0.266	0.266
	E5	0.100	0.266	0.466
RSA	E1	0.100	0.466	0.100
	E2	0.100	0.100	0.266
	E3	0.100	0.100	0.266
	E4	0.266	0.266	0.100
	E5	0.100	0.100	0.100

It can be seen that the results of proposed fuzzy methodology are consistent with the results based on expert opinion. The ranking of six highlighted metrics in Table 9, namely NRFD, NBC, NC, RBC, NAFC, RSA are exactly same for two approaches. The ranks of NA, NADC, NH, NDC, NFC and DHP differ by one which are more or less same for both approaches. The ranking for NABC shows variation with a rank of 5 in proposed approach and 11 in expert opinion approach owing to fractional differences in the values of their permanent. So the results obtained by proposed methodology are consistent with the results based on expert opinion. The comparison of proposed methodology with expert opinion approach along certain parameters is shown in Table 11. It can be seen from Table 11 that results of the fuzzy based approach are better than the results of aggregate expert opinion along seven identified criteria. The fuzzy approach is more accurate due to consideration of ambiguous human thoughts in the calculations as compared to algebraic aggregation.

Table 11. Comparison based on various parameters.

S. No.	Parameters	Proposed fuzzy methodology	Expert opinion approach
1	Number of computations	Proportional to number of attributes (N) i.e. experts, metrics and criteria	Proportional to number of attributes (N) i.e. experts, metrics and criteria
2	Weight matrix	Fuzzy aggregation	Algebraic aggregation
3	Rate matrix	Fuzzy aggregation	Algebraic aggregation
4	Criteria matrix	Fuzzy aggregation	Algebraic aggregation
5	Consideration of all possible interdependencies of variables	Yes	No
6	Rank of metrics	Yes	Yes
7	Accuracy	More due to fuzzy base approach	Lesser due to algebraic approach

7. Result Conclusion and Future Research

The authors proposed a fuzzy based ranking methodology to rank quality metrics of conceptual data warehouse models along criteria of understandability, efficiency and effectiveness. The opinion of experts was taken in terms of fuzzy linguistic variables to assign weights to criteria and ratings to metrics. A criteria matrix of ratings and rankings was formed for each of the metrics.

The permanent of criteria matrix was calculated to rank the metrics. A comparison was also made with other methodology to validate the results of calculation. The results of fuzzy based approach are more reliable, accurate as compared to expert opinion approach due to the consideration of ambiguity, imprecision prevalent in human thought process and consideration of all interdependencies of attributes by the use of permanent function.

The proposed work can be further extended by discovery of more criteria, more metrics for quality evaluation and then applying the proposed fuzzy methodology. Also more number of experts from diverse domains and having wide experience can be consulted for generalization and validation of results. A broad comparison can be made with other ranking methodologies to measure the accuracy of results obtained using the proposed fuzzy based approach.

References

[1] Abello A., Samos J., and Saltor F., "Yam2: a Multidimensional Conceptual Model Extending UML," *Information Systems*, vol. 31, no. 6, pp. 541-567, 2006.

[2] Aliev R. and Huseynov O., "Fuzzy Geometry-Based Decision Making with Unprecisiated Visual Information," *International Journal of Information Technology & Decision Making*, vol. 13, no. 5, pp. 1051-1073, 2014.

[3] Aslolamy A. and Qureshi R., "The Proposal of a Qualification based Approach to Teach Software Engineering Course," *The International Arab Journal of Information Technology*, vol. 12, no. 2, pp. 169-175, 2015.

[4] Bailador G. and Trivino G., "Pattern Recognition Using Temporal Fuzzy Automata," *Fuzzy Sets and Systems*, vol. 161, no. 1, pp. 37-55, 2010.

[5] Briand L., Freimut B., and Vollei F., "Assessing the Cost Effectiveness of Inspections by Combining Project Data and Expert Opinion," *in Proceedings of 11th International Symposium on Software Reliability Engineering*, San Jose, pp. 246-258, 2000.

[6] Buckley J. And Chanas S., "A Fast Method of Ranking Alternative Using Fuzzy Numbers," *Fuzzy sets and Systems*, vol. 30, no. 3, pp. 337-338, 1989.

[7] Calero C., Piattini M., Pascual C., and Serrano M., *Towards Data Warehouse Quality Metrics*, International Workshop on Design and Management of Data Warehouses, 2001.

[8] Chen T., "Data Construction Process and Qualiflex-based Method for Multiple-Criteria Group Decision Making with Interval-Valued Intuitionistic Fuzzy Sets," *International Journal of Information Technology & Decision Making*, vol. 12, no. 3, pp. 425-467, 2013.

[9] Cochran J. and Chen H., "Fuzzy Multi Criteria Selection of Object Oriented Simulation Software for Production System Analysis," *Computers and Operations Research*, vol. 32, no. 1, pp. 153-168, 2005.

[10] Dahiya N., Bhatnagar V., and Singh M., "Effective Data Warehouse for Information Delivery: a Literature Survey and Classification," *International Journal Networking and Virtual Organisations*, vol. 12, no. 3, pp. 217-237, 2013.

[11] Dahiya N., Bhatnagar V., and Singh M., "Enhancing Consistency of Conceptual Data Warehouse Models," *International Journal of Computational Systems Engineering*, vol. 2, no. 1, pp. 11-24, 2015.

[12] Dyba T., "An Instrument for Measuring the Key Factors of Success in Software Process Improvement," *Empirical Software Engineering*, vol. 5, no. 4, pp. 357-390, 2000.

[13] Franconi E. and Sattler U., "A Data Warehouse Conceptual Data Model for Multidimensional Aggregation," *in Proceedings of Workshop on Design and Management of Data Warehouses*, Heidelberg, pp.1-13, 1999.

[14] Gamal N., Galal-Edeen G., and Bastawissy E., "Towards a Generic Conceptual Model for Data warehouses," *in Proceedings of 5th International Business Information Management Association*, Cairo, pp. 262-269, 2005.

[15] Garg R., Gupta V., and Agrawal V., "Quality Evaluation of Thermal Power Plants by Graph Teeoretical Methodology," *International Journal of Power and Energy Systems*, vol. 27, no. 1, pp. 42-48, 2007.

[16] Garg R., Sharma K., Nagpal C., Garg R., and Kumar R., "Ranking of Software Engineering Metrics by Fuzzy based Matrix Methodology," *Software Testing, Verification and Reliability*, vol. 23, no. 2, pp. 149-168, 2013.

[17] Genero M., Poels G., and Piattini M., "Defining and Validating Metrics for Assessing the Understandability of Entity-Relationship Diagrams," *Data & Knowledge Engineering*, vol. 64, no. 3, pp. 534-557, 2008.

[18] Golfarelli M., Maio D., and Rizzi S., "The Dimensional Fact model: a Conceptual Model for Data Warehouses," *International Journal of Cooperative Information Systems*, vol. 7, no. 2, pp. 215-247, 1998.

[19] Harinarayan V., Rajaraman A., and Ulmann J., "Implementing Data Cubes Efficiently," *in Proceedings of the SIGMOD International Conference on Management of Data*, Quebec, pp. 205-216, 1996.

[20] Johnson G. and Yu X., "Objective Software Quality Assessment," *in Proceedings of Nuclear Science Symposium*, Seattle, pp. 1691-1698, 1999.

[21] Kamble A., "A Conceptual Model for Multidimensional Data," *in Proceedings of 5th Asia-Pacific Conference on Conceptual Modelling*, Wollongong, pp. 29-38, 2008.

[22] Kaufmann A. and Gupta M., *Fuzzy Mathematical Models in Engineering and Management Science*, Elsevier Science Publisher, 1998.

[23] Khatatnech K. and Mustafa T., "Software Reliability Modeling Using Soft Computing Technique," *European Journal of Scientific Research*, vol. 26, no. 1, pp. 154-160, 2009.

[24] Kpodjedo S., Ricca F., Galinier P., Gueheneuc Y., and Antoniol G., "Design Evolution Metrics for Defect Prediction in Object Oriented Systems," *Empirical Software Engineering*, vol. 16, no. 1, pp. 141-175, 2011.

[25] Lau H., Wong C., Lau P., Pun K., Chin K., and Jiang B., "A Fuzzy Multi Criteria Decision Support Procedure for Information Delivery in Extended Enterprise Networks," *Engineering Applications of Artificial Intelligence*, vol. 16, no. 1, pp. 1-9, 2003.

[26] Li M. and Smidts C., "A Ranking of Software Engineering Measures based on Expert Opinion," *IEEE Transactions on Software Engineering*, vol. 29, no. 9, pp. 811-824, 2003.

[27] Mallick B., Garg D., and Grover P., "Constraint Based Sequential Pattern Mining: A Pattern Growth Algorithm Incorporating Compactness, Length and Monetary," *The International Arab Journal of Information Technology*, vol. 11, no. 1, pp. 33-42, 2014.

[28] Marcus M. and Minc H., "Permanents," *American Mathematics*, vol. 72, pp.577-591, 1965.

[29] Mishra D., Yazici A., and Basaran B., "A Case Study of Data Models in Data Warehousing," *in Proceedings of 1st International Conference on the Applications of Digital Information and Web Technologies*, Ostrava, pp. 314-319, 2008.

[30] Moody D., "Theoretical and Practical Issues in Evaluating the Quality of Conceptual Models: Current State and Future Directions," *Data & Knowledge Engineering*, vol. 55, no. 3, pp.243-276, 2005.

[31] Perez J., Berlanga R., Aramburu M., and Pedersen T., "A Relevanceextended Multi-Dimensional Model for a Data Warehouse Contextualized with Documents," *in Proceedings of 8th ACM International Workshop on Data Warehousing and OLAP*, Bremen, pp. 19-25, 2005.

[32] Piattini M., Genero M., and Jimenez L., "Metrics based Approach for Predicting Conceptual Data model maintainability," *International Journal of Software Engineering and Knowledge Engineering*, vol. 11, no. 6, pp. 703-729, 2001.

[33] Ping-Feng P., Chen-Tung C., and Wei-Zhan H., "Applying Linguistic Information and Intersection Concept to Improve Effectiveness of Multi-criteria Decision Analysis Technology," *International Journal of Information Technology & Decision Making*, vol. 13, no. 2, pp. 291-315, 2014.

[34] Sapia C., Blaschka M., Hofling G., and Dinter B., "Extending the ER Model for the Multidimensional Paradigm," *Advances in Database Technologies*, Singapore, pp. 105-116, 1998.

[35] Schuff D., Corral K., and Turetken O., "Comparing the Understandability of Alternative Data Warehouse Schemas: an Empirical Study," *Decision Support Systems*, vol. 52, no. 1, pp. 9-20, 2011.

[36] Serrano M., Trujillo J., Calero C., and Piattini M., "Metrics for Data Warehouse Conceptual Models Understandability," *Information and Software Technology*, vol. 49, no. 8, pp. 851-870, 2007.

[37] Singh Y., Kaur A., and Malhotra R., "Empirical Validation of Object Oriented Metrics for Predicting Fault Proneness Models," *Software Quality Journa*l, vol. 18, no. 1, pp. 3-35, 2010.

[38] Trujillo J. and Palomar M., "An Object Oriented Approach to Multidimensional Database Conceptual Modeling (OOMD)," *in Proceedings of the 1st ACM International Workshop on Data Warehousing and OLAP*, Washington, pp. 16-21, 1998.

[39] Tryfona N., Busborg F., and Christainsen J., "StarER: a Conceptual Model for Data Warehouse Design," *in Proceeding of the 2nd ACM International Workshop on Data Warehousing and OLAP*, Kansas City, pp. 3-8, 1999.

[40] Tsois A., Karayiannidis N., and Sellis T., "MAC: Conceptual Data Modeling for OLAP," *in Proceedings of the International Workshop on Design and Management of Data Warehouses*, pp. 1-5, 2001.

[41] Wang J. and Lin Y., "A Fuzzy Multi Criteria Group Decision Making Approach to Select Configuration items for Software Development," *Fuzzy Sets and Systems*, vol. 134, no. 3, pp. 343-363, 2003.

[42] Xia M. and Xu Z., "A Novel Method for Fuzzy Multi-Criteria Decision Making," *International Journal of Information Technology & Decision Making*, vol. 13, no. 3, pp. 497-519, 2014.

[43] Zadeh L., "Fuzzy Sets," *Information and Control*, vol. 8, no. 3, pp. 338-353, 1965.

[44] Zhang X. and Pham H., "An Analysis of Factors Affecting Software Reliability," *The journal of Systems and Software*, vol. 50, no.1, pp. 43-56, 2000.

Naveen Dahiya received his B.E. in Computer Science and Engineering from Maharshi Dayanand University, Rohtak, Haryana, India in 2003 and M.Tech in Computer Engineering from Maharshi Dayanand University, Rohtak, Haryana, India in 2005 and Ph.D from Y.M.C.A. University of Science and Technology, Faridabad, Haryana, India. He is working as an Associate Professor and head in Computer Science and Engineering Department at Maharaja Surajmal Institute of Technology, C-4, Janak Puri, New Delhi, India. His Research interests include database systems, data warehouse and data mining.

Vishal Bhatnagar receieved his B.Tech degree from Nagpur University in Nagpur, India in 1999 and M.Tech in Information Technology from Punjabi University, Patiala, India in 2005, and Ph.D from Shobhit University in 2010. He is Associate Professor in Computer Science and Engineering Department at Ambedkar Institute of Advanced Communication Technologies and Research (Government of Delhi), GGSIPU, Delhi, India. His research interests include database, advance database, data warehouse and data mining.

Manjeet Singh received his M.Tech degree in Computer Science and Engineering from Guru Jhambeshwar University, Hissar, India and Ph.D from Maharshi Dayanand University, Rohtak, India in 2009. He is working as Associate Professor in Computer Engineering Department at Y.M.C.A University of Science and Technology, Faridabad, Haryana, India. His research interests include database systems,natural language processing, and compiler construction.