Journal Of Harmonized Research (JOHR)

Journal Of Harmonized Research in Engineering
3(4), 2015, 162-171

ISSN 2347 – 7393

Original Research Article

EVALUATING NOSQL DATABASES WITH FUZZY DECISION MAKING

Bashar Hindawi, Hazem M.El-Bakry, Mervat Abu-Elkheir

Faculty of Computer and Information Sciences
Mansoura University, Mansoura, Egypt

Abstract: NoSQL databases are widely used to handle and store data for large scale applications. Database system performance is an important quality attribute to develop software and applications because it is related to the other qualitative attributes such as availability, reliability, functionality and so on. There are no tools or software in the market to accurately measure the performance of NoSQL databases. As a result of having various levels of performance within NoSQL databases, it is important to evaluate and compare their performances to identify potential strategies. Because the evaluation process is subject to various degrees of expert opinions and preferences, it is difficult to assign the performance priorities and specify how NoSQL databases can be ranked. We propose a Fuzzy evaluation scheme that provides evaluation degrees with more precise. This scheme depends on conducting the pairwise comparisons between the alternatives in terms certain criterion. This paper implement fuzzy scheme and preliminary results will be showed clearly by total performance for each database. The numerical values are represented in the results to be easier during the ranking. This study allows the developers and system analysts to specify the most suitable database due to the application needs.

Index Terms**—**Consistent Fuzzy Preference Relation (CFPR), fuzzy AHP, Multi criteria Decision Making (MCDM), fuzzy LinPreRa, Fuzzy Linguistic Assessment Variables (FLAV)

Introduction: Because of the increasing need to evolve the performance and enable the scalability, NoSQL technology has emerged as an alternative to relational model or RDBMS. Besides the large number of the products, the growing interest of NoSQL was detected the pressing need to compare and evaluate the underlying technologies and features of NoSQL databases. The purpose of the evaluation process is to facilitate choosing an appropriate database for certain use case.

Although the implementation of the applications

For Correspondence:
hindawibashar@yahoo.com
Received on: December 2015
Accepted after revision: December 2015
Downloaded from: www.johronline.com

and enterprises are not driven by only performance, it is important to inform that all the other quality attributes of NoSQL (Availability, Consistency, reliability scalability etc.) are correlated directly to the performance of databases. Hence, the performance is crucial factor for developers, systems analyst and software engineers to decide which database is more suitable for their enterprises or applications. Therefore, it is necessary to identify the best method to evaluate the various categories of NoSQL databases.

Most of the research works use summary table for the evaluation process to indicate if the database is suitability, high quality attribute or best performance etc. In these researches, the evaluation process is generally based on binary representation e.g. √ vs. nothing, + vs. -, good vs. bad… etc.) [16][17] or different grades of representation scale which has specific range[14]. In practical terms, both representations have weaknesses for several reasons : Firstly, these kinds of representations do not present a clear idea about the real values of the evaluation: Secondly, there is no indicator to inform us which one of these databases are more preferred over the other, when they have the same representation or grade.

To resolve the above weaknesses and provide the developers with a powerful tool, the evaluation process can take different trend from the traditional representations. This tool help stakeholders to determine the priorities and make the decision.

We suggest a new methodology to achieve the evaluation process based on fuzzy reasoning. This method focus on pairwise comparisons between the alternatives to obtain the final results with numerical values. These numerical values represent the final rank for each database that was subjected to the evaluation.

Despite the evaluation process is a part of uncertainty situation, there is no study or research work that adopt conducting the evaluation in accordance with the fuzzy decision making method. The goal of this paper is to fill the gap and aid the developers to specify the suitable database for certain scenario.

The reminder of this paper is structured as follows: In section II, we discuss and compare the previous research works. In section III, we introduce the fundamental concepts of fuzzy decision making. The necessary propositions and equations to be applied in our study are presented in section III. With pseudo notation, the required procedures for evaluating the alternatives and making the decision are presented in section IV. To verify the fuzzy method and examine its efficiency, case study is presented in Section V. The results of applying the procedures on the alternatives are detailed in Section VI. In section VII comparing between the proposed methods for decision making and fuzzy AHP. Some of conclusions and remarks are provided in section IX.

Previous Works

Since 2011, the underlying technologies of NoSQL were prosperous with several databases as center of the large number of studies [15]. There are more than 140 of available NoSQL database as open source and each one of them offers its own set of services [7]. It is impossible to find database has high level of all the quality attributes, where each databases offers trade-offs. For instance, MongoDB provides high degree of reliability, whether it presents worse service with write intensive operation [14]. Therefore, the "one size fits the all" approach which was followed in relational databases would never be applicable on NoSQL.

The evaluation process was carried in [16] by comparing three NoSQL databases products as follows: Cassandra, MongoDB and Couchbase. The summery table put sixteen different feature to be evaluated according to NoSQL products. These features are connected to set the quality attributes such scalability, availability, consistency and performance. The evaluation process was achieved according to the binary choice, which mean each one of NoSQL database meet level of certain feature should have the mark √ else nothing. Although the paper presented the most common features of the performance of NoSQL databases, the evaluation process did not specify the databases which is more suitable for certain feature. For example, due to "Support for Sharding" feature, both Cassandra and MongoDB have the same grade (√) of the presentation, but this study avoided explanation which one of the products more support for sharding than the other. On the other hand, since the features of certain product are changing from time to time then, this study may be ineffectual in the next years.

R.Hetch & S.Jabolonski [17] presented a survey to evaluate NoSQL features. Instead of NoSQL products, authors used the underlying techniques such as query possibilities, eventually currency, replication, and partitioning to be evaluated. The paper also highlighted the important features which help to select the suitable database. The results of this study showed potential using the four types of data model(will be stated later) in different use cases scenario such as Document Store provides flexible data model with great query possibilities and Column Family Stores are more suitable for large datasets. Unlike the previous study, this paper evaluated the NoSQL technologies using the binary scale (+, -), where the feature itself is either good or bad at determining the suitability.

J.Lorenco et al [5] presented detailed study about choosing the right NoSQL database for the right job. The study showed set of quality attributes to be evaluated in terms NoSQL databases. The quality attributers included ten criteria which are evaluated relative to seven of NoSQL products. In concerning the performance, the study used only two operations to be evaluated: Read and Write optimized. This study does not interested with the other operations related to performance such using the memory, ability to scale up, concurrency and etc. The study established a comprehensive summery table to indicate which database best fit of quality attribute. 4- Scale representation (good, average, mediocre and bad) is presented to determine the grade of each product in terms of quality attribute.

The previous works based on the literature study, experimental analysis and experts' opinions to evaluate the NoSQL technology whether products or databases. This paper intends to apply the fuzzy decision making method instead of binary representation or scale based to obtain the numerical values which are very closely to the reality.

Fuzzy Decision Making

The process of making the decision with emerging multiple criteria or alternatives is called Multi Criteria Decision Making or MCDM [1]. MCDM methods can help the decision makers or experts to evaluate the criteria and select the best alternatives based on own perspectives [14].

Analytic Hierarchal Process (AHP) [5] is a one of fuzzy decision making methods that was widely applied to rank multiple criteria and choose the best alternative through decision making process, e.g. evaluating the importance of risk factors[9], selection the suppliers [11] and solving the budget allocation problems[13] etc. In fuzzy AHP method, decision matrix is constructed from experts' preferences by answering the pairwise comparisons. If the problem has n of criteria, n*(n-1)/2 of pairwise comparisons are required to be answered. The number of pairwise comparisons directly proportional with number of criteria or alternatives. Having a large number of comparisons may cause a mental confusion for decision makers which results in inconsistent answers. Thus, the comparisons' questions must be reconstructed in order to change or update some the answers. This process may causes wasting time, losing efforts and inefficient method.

To solve the above mentioned problem, reference [16] presented one of the newly advanced MCDM techniques called Consistent Fuzzy Preference Relation (CFPR). CFPR has the ability to provide the preferences for a set of criteria or alternatives with less number of pairwise comparisons. CFPR reduces the number of pairwise comparisons as well as avoid the situation of inconsistent. For n criteria, only (n-1) of questions must be answered as pairwise comparisons within CFPR. The purpose of this process is ensuring the consistency

Although the consistency is a one of the significant concepts to avoid misleading solutions, ensuring the consistency with 100% is a very difficult to be accomplished in practice. Wang and Chen [4] proposed fuzzy linguistic assessment variables (FLAV) to construct decision matrices according to fuzzy linguistic preference relations. The purpose of FLAV is to mitigate the inconsistencies and to avoid the unexpected results.

As a result of various degrees, the experts' preferences are often vague based on the natural language and it is very difficult to be estimated within numerical values. Instead of numerical values or crisp data, the linguistic variables are more adequate for modeling the real life problems.

This study combines between CFPR which proposed by Herrera-Vidman [3] and FLAV which proposed by Wang and Chen [4] to evaluate the performance of NoSQL databases.

Consistent Fuzzy Preference Relation (CFPR)

By FLAV, CFPR provides the experts with all values for representing the various degrees of preference to compare single alternative over the next one [6]. The pairwise comparisons will be subject to decision matrices using additive reciprocal and consistency property [12]. The fuzzy preference relation *P* on set of alternatives *X* is a fuzzy set of $X \times$ with membership function $P: X \times X \rightarrow [0$[8]. The preference relations are presented as n×n matrix, $P = [p_{ij}], \forall$ is interpreted as degree of importance of criterion over

$$p_{ij} = \begin{cases} 0.5, & x_i \text{ and } x_j \text{ are the same important} \\ 1, & x_i \text{ is absolutly important than } x_j \\ > 0.5, & x_i \text{ is more important than } x_j \\ < 0.5, & x_i \text{ is less important than } x_j \end{cases}$$

Only n - of experts' judgments is required to ensure the consistency with n criteria or alternatives [7]. For a set of criteria and alternatives, the following description expresses some of significant propositions [9], [10] that will be applied in this study:

Proposition1: For set of alternatives, $X = \{x1, \dots x$ associated with reciprocal linguistic preference relation, $P = (p$ where $p_{ij} \in [0$, verifies the *additive reciprocal* property, thus, the following equation equivalent:

$$p_{ij} + p_{ji} = 1 \quad \forall i, j \in (1, \dots, n) \tag{1}$$

Proposition 2: For reciprocal fuzzy linguistic preference relation $P = (p$, to be consistent, verifies the *additive consistency*, the following equations are equivalent:

a) $p_{ij} + p_{ji} + p_{ki} = \frac{3}{2} \; \forall i < j < k$ (2)

b) $p_{i(i+1)} + p_{i(i+1)(i+2)} + \dots + p_{(j-1)j} = \frac{j-i+1}{2} \; \forall i < j$ (3)

Mostly, the values of the obtained matrix are not in the interval [0,1], but in the interval $[-c, 1+$. In such case, using the following transformation function [3] to transform the fuzzy numbers within the interval [0,1].

$$f: [-c, 1+c] \rightarrow [0,1], f(x) = \frac{x+c}{1+2c} \tag{4}$$

This transformation function preserves the reciprocity and additive consistency for all elements in decision matrix.

Fuzzy LinPreRa Algorithm

Fuzzy LinPreRa was introduced by Wang and Chen [5] to handle the vague judgments and overcome the inconsistency. This method suggests FLAV to handle the various degrees of experts' judgments. Table 1 shows prototype of FLAV. Triangular Fuzzy Numbers (TFN) represents a range of possibility memberships in distributions which can be effectively used in logic reasoning [8].

Table 1 Fuzzy linguistic assessment variables (FLAV)

Linguistic variables	TFN
Absolutely Important(AI)	$(0, p^M_{AI}, p^R_{AI})$
⋮	⋮
Equally Important(EI)	$(p^L_{EI}, 0.5, p^R_{EI})$
⋮	⋮
Absolutely Not Important (AN)	$(p^L_{AN}, p^M_{AN}, 1)$

To obtain the importance weights for each alternative, this algorithm uses the arithmetic mean (Average) of each row *i* in decision matrix, then normalize the weights of alternatives. The required equations [2] as follows:

$$A_i = \frac{1}{n}\left(\sum_{i=1}^{n} p_{ij}\right) \tag{5}$$

$$W_i = \frac{A_i}{A_1 + A_2 + \cdots A_n} \tag{6}$$

Combining both CFPR and fuzzy LinPreRa can be employed in performance evaluation of NoSQL database. The following pseudo notation presents three algorithms to implement the performance evaluation of NoSQL databases with fuzzy decision making

Algorithm1 Creating Consistent & Complete Decision Matrix

Input: Linguistic Pairwise Comparisons between P_i and P_{i+1}, no_of_alternatives
Output: Consistent Matrix {C_mat(i,j)}
Begin

```
Let P(i,j) is a matrix has experts' opinions
Function C_mat( i:integer , j:integer )
  For i ← 1 to no_of_alternatives Do
    For j ← 1 to no_of_alternatives Do
      If (i = j) then C_mat(i,j) = 0.5
       If (i = j + 1) then
C_mat(i,j) = 1 - P(j,i)
    End For
  End For {proposition 1}
  For i ← 1 to no_of_alternatives Do
    For j ← 1 to no_of_alternatives Do
      For k ← 1 to no_of_alternatives Do
        If (i < j) AND (j < k) Then
C_mat(k,i) =
  1.5 - P(i,j) - P(j,k)
         If (i < j) AND P(i,j) ≠ Null Then
          C_mat(j,i) = 1 - P(i,j)
      End For {proposition 2}
    End For
  End For
End Function
End Begin
```

Algorithm 2: Establishing the Transformation Matrix

Input: Consistent and Complete Matrix {C_mat(i,j)}
Output: Transformation Matrix {T_mat (i,j)}
Begin

```
Let v is a constant and represent the max violent in C_mat(i,j)
Function T_mat( i: integer, j: integer)
   For i ← 1 to no_of_alternatives Do
     For j ← 1 to no_of_alternatives Do
       T_mat(i,j) = (C_mat(i,j)+ v)/1+(2*v))
     End For
   End For
  End Function
End Begin
```

Algorithm3: Assigning Weights for each alternatives using LinPreRa

Input: Transformation Matrix {T_mat(i,j)}
Output: Fuzzy Weights
Begin

```
 Let AvgRow(i) represents the average of elements in row (i) at matrix T_m(i,j)
 For i ← 1 to no_of _alternatives Do
    For j ← 1 to no_of_alternatives Do
       AvgRow(i) ← ∑ [T_mat(i,j)] / no_of_alternatives
         {Average for each rwo in matrix}
    End For
 End For
 For i ← 1 to no_of_alternatives Do
   Wi= AvgRow(i) / Sum of all the elements of
                 Avg Row (i)
 End For
End Begin
```

Performance Evaluation with Fuzz Decision Making: In our study, the evaluation process is primarily depend on the experts' opinions or group decision making. This group examine the databases due to the following means:

- Using the YCSB* to ev4aluate and compare some significant operations of NoSQL Databases.
- The wide experience of the experts in this field

The group decision making consist of five experts with more than five years' experience in large scale applications and academic lecturers in NoSQL databases. Four criteria are related to the performance of NoSQL databases will be evaluated as follows [8]:

1- Read intensive performance (C1)
2- Write intensive performance (C2)
3- Ability to scale(C3)
4- Using memory efficiency (C4)

* YCSB provide opportunity to compare and evaluate some operations. This framework is divided to two parts: 1.Data generator 2. Operations testing center. YCSB can be obtained from https://labs.yahoo.com/

The NoSQL databases are divided in four types[18][19][20] according to the storage and data model. These types are classified according to the fact that each one of these types offer different solutions e.g. Column Store Database is a good choice when improving the writing operations, whereas Document Store database is more oriented towards read operation [16]. The four different categories of NoSQL databases are showed as follows [8]:

A- Key values Store (A1)
B- Document Store (A2)
C- Column Family Store (A3)
D- Graph Databases (A4)

Now, to perform the evaluation process of NoSQL databases in terms the performance, we must determine the three levels : the goal, the criteria, the alternatives as follow:

The goal: Performance evaluation of NoSQL databases.

The criteria: Represent the four factors previously mentioned (read performance (C1), write performance (C2), ability to scale (C3) using memory efficiency (C4)).

The alternatives: Represent the data models or categories of databases (Key Value Stores(A1), Document Stores(A2) , Column Family Stores (A3), Graph Database (A4)).

Fig1 shows the hierarchical structure of the criteria and alternatives.

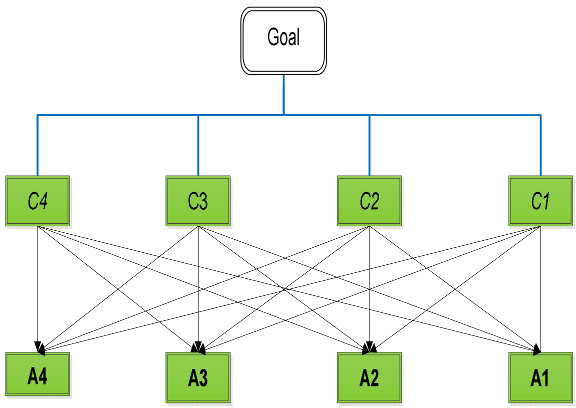

Fig 1. Hierarchal structure of criteria and alternatives

The decision procedures for achieving the goal are illustrated as follows: Five experts k $(k = 1, 2, ..$ provide their preferences according to Table 2.

Table 2: Fuzzy linguistic assessment variables

Linguistic variables	TFN
Very Poor(VP)	(0,0,0.1)
Poor(P)	(0,0.1,0.3)
Medium Poor(MP)	(0.1,0.3,0.5)
Medium(M)	(0.3,0.5,0.7)
Medium Good(MG)	(0.5,0.7,0.9)
Good(G)	(0.7,0.9,1)
Very Good(VG)	(0.9,1,1)

Table 3. Shows the original experts' preferences to evaluate the alternatives for the four criteria.

Table 3. Original preferences for alternatives

C1	*E1*	*E2*	*E3*	*E4*	*E5*	
A1	MG	G	G	MG	VG	A2
A2	G	G	VG	G	MG	A3
A3	MP	M	P	M	M	A4
C2	*E1*	*E2*	*E3*	*E4*	*E5*	
A1	G	VG	VG	G	MG	A2
A2	P	M	P	VP	P	A3
A3	M	MG	MG	MG	M	A4
C3	*E1*	*E2*	*E3*	*E4*	*E5*	
A1	MG	MG	M	G	G	A2
A2	MP	M	P	M	MG	A3
A3	G	MG	MG	MG	MP	A4
C4	*E1*	*E2*	*E3*	*E4*	*E5*	
A1	G	VG	M	MG	G	A2
A2	P	M	VP	P	MP	A3
A3	G	M	M	MG	MG	A4

Regarding the first criterion (C1), table 4 shows decision matrices for the five experts' opinions as follows:

Table 4. Experts' judgments for C1

		A_1	A_2	A_3	A_4
$E_{1}=$	A_1	(0.5,0.5,0.5)	(0.5,0.7,0.9)	P_{13}	P_1
	A_2	P_{21}	(0.5,0.5,0.5)	(0.7,0.9,1)	P_{24}
	A_3	P_{31}	P_{32}	(0.5,0.5,0.5)	(0,0.1,0.3)
	A_4	P_{41}	P_{42}	P_{43}	(0.5,0.5,0.5)
		A_1	**A_2**	**A_3**	**A_4**
$E_{2}=$	A_1	(0.5,0.5,0.5)	(0.7,0.9, 1)	P_{13}	P_{14}
	A_2	P_{21}	(0.5,0.5,0.5)	(0.7,0.9,1)	P_{24}
	A_3	P_{31}	P_{32}	(0.5,0.5,0.5)	(0.3,0.5,0.7)
	A_4	P_{41}	P_{42}	P_{43}	(0.5,0.5,0.5)
		A_1	**A_2**	**A_3**	**A_4**
$E_{3}=$	A_1	(0.5,0.5,0.5)	(0.7,0.9,1)	P_{13}	P_{14}
	A_2	P_{21}	(0.5,0.5,0.5)	(0.9,1,1)	P_{24}
	A_3	P_{31}	P_{32}	(0.5,0.5,0.5)	(0,0,0.1)
	A_4	P_{41}	P_{42}	P_{43}	(0.5,0.5,0.5)
		A_1	**A_2**	**A_3**	**A_4**
$E_{4}=$	A_1	(0.5,0.5,0.5)	(0.5,0.7,0.9)	P_{13}	P_{14}
	A_2	P_{21}	(0.5,0.5,0.5)	(0.7,0.9,1)	P_{24}
	A_3	P_{31}	P_{32}	(0.5,0.5,0.5)	(0.3,0.5,0.7)
	A_4	P_{41}	P_{42}	P_{43}	(0.5,0.5,0.5)
		A_1	A_2	A_3	A_4
$E_5=$	A_1	(0.5,0.5,0.5)	(0.5,0.7,0.9)	P_{13}	P_{14}
	A_2	P_{21}	(0.5,0.5,0.5)	(0.5,0.7,0.9)	P_{24}
	A_3	P_{31}	P_{32}	(0.5,0.5,0.5)	(0.3,0.5,0.7)
	A_4	P_{41}	P_{42}	P_{43}	(0.5,0.5,0.5)

To obtain the aggregated experts' opinions within single decision matrix, the linguistic averaging factor proposed by [7] should be applied. Table 5 shows inconsistent decision matrix for all experts' preferences of C1

Table 5. Inconsistent decision matrix

C_1	**A1**	**A2**	**A3**	**A4**
A1	(0.5,0.5,0.5)	(0.58,0.78,0.94)	P_{13}	P_{14}
A2	P_{21}	(0.5,0.5,0.5)	(0.7,0.88,0.98)	P_{24}
A3	P_{31}	P_{32}	(0.5,0.5,0.5)	(0.18,0.32,0.5)
A4	P_{41}	P_{42}	P_{43}	(0.5,0.5,0.5)

According to proposition 1 and proposition 2, the complete decision matrix will be obtained as a result of applying the reciprocal additive property and additive consistency property.

Now, the complete decision matrix is available after applying the whole calculations on fuzzy preference relation matrix. Table 6 shows decision matrix with all elements of fuzzy numbers

Table 6. Complete decision matrix

	A_1	**A_2**	**A_3**	**A_4**
A1	(0.5,0.5,0.5)	(0.58,0.78,0.94)	(0.78,1.16,1.42)	(0.45,0.98,1.32)
A2	(0.06,0.22,,0.42)	(0.5,0.5,0.5)	(0.7,0.88,0.98)	(0.38,0.7,0.88)
A3	(-0.42,-0.16,0.22)	(0.02,0.12,0.3)	(0.5,0.5,0.5)	(0.18,0.32,0.5)
A4	(-0.32,0.02,0.54)	(0.12,0.3,0.62)	(0.5,0.68,0.82)	(0.5,0.5,0.5)

As noted, some of fuzzy numbers in the above decision matrix are outside the interval [0,1], therefore transformation functions must be applied. Table 7 shows the transformation matrix with applying the equations 4.

Table7. Transformation matrix

C_1	A_1	A_2	A_3	A_4
A_1	(0.5,0.5,0.5)	(0.54,0.65,0.74)	(0.65,0.86,1)	(0.48,0.76,0.95)
A_2	(0.26,0.35,0.46)	(0.5,0.5,0.5)	(0.61,0.71,0.76)	(0.43,0.61,0.71)
A_3	(0,0.14,0.35)	(0.24,0.29,0.39)	(0.5,0.5,0.5)	(0.33,0.4,0.5)
A_4	(0.05,0.24,0.52)	(0.29,0.39,0.57)	(0.5,0.6,0.67)	(0.5,0.5,0.5)

The weights of alternatives are calculated by equation (5), (6) and using the center of gravity (COG) we gain the crisp values (defuzzified) of evaluation. Table 8 shows the final results of evaluation the alternatives of C1 including importance weights and defuzzified numbers .

Table 8. Weights of alternatives C1

Alternatives of (C1)	Average (Ai)	Weight (Wi)	Defuzzification (Di)
A1	(0.543,0.693,0.798)	(0.226,0.346,0.500)	0.357
A2	(0.45,0.543,0.608)	(0.187,0.271,0.381)	0.279
A3	(0.268,0.333,0.435)	(0.111,0.166,0.273)	0.183
A4	(0.335,0.433,0.565)	(0.139,0.346,0.354)	0.280

Similarly, repeat the previous steps for obtaining the importance weights of alternatives for C2, C3 and C4. Tables 9 presents the decision matrices with the final evaluations of alternatives for each one of the remaining criteria.

Table 9. Weights of alternatives for C2,C3,C4

alternative	Average (Ai)	Weight (W_i)	Defuzzification (D_i)
A1	(0.465,0.648,0.808)	(0.177,0.320,0.568)	0.355
A2	(0.22,0.313,0.455)	(0.084,0.155,0.320)	0.186
A3	(0.455,0.578,0.688)	(0.173,0.286,0.483)	0.314
A4	(0.283,0.485,0.677)	(0.107,0.240,0.476)	0.244
A1	(0.460,0.605,0.758)	(0.185,0.302,0.502)	0.330
A2	(0.383,0.458,0.573)	(0.154,0.229,0.380)	0.254
A3	(0.410,0.518,0.626)	(0.165,0.259,0.414)	0.280
A4	(0.258,0.420,0.535	(0.103,0.210,0.354)	0.222
A1	(0.398,0.590,0,766)	(0.151,0.295,0.561)	0.336
A2	(0.245,0.358,0.50)	(0.093,0.179,0.366)	0.213
A3	(0.458,0.59,0.708	(0.174,0.295,0.518)	0.329
A4	(0.265,0.463,0.665)	(0.100,0.231,0.486)	0.272

The Evaluation Results:

The four categories of NoSQL databases that have used in the evaluation are framed and organized in hierarchy structure as shown in Fig 1. To rank the databases and evaluate the performance of NoSQL databases, five academic experts were given their opinions to prefer alternative over the next alternative. Proposition 1 and 2 have applied to obtain the consistent fuzzy preference relation matrix. As a result to existence the values which are outside the interval [0,1], fuzzy numbers were transformed by transformation functions. The importance weights for all alternatives have computed using the equations (5) and (6). The evaluation degrees or defuzzification of the alternatives are shown in the last columns of the Tables 8 and 9. Graphically, the share of each one of databases (alternatives) in terms of certain criterion are shown in Fig. 2-a,Fig. 2-b,Fig 2-c and Fig. 2-d.

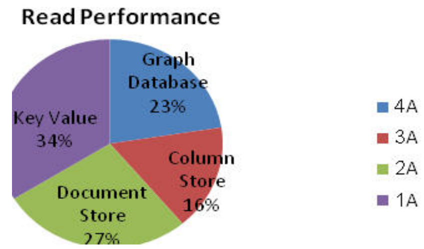

Fig. 2-a Percentage of each alternative with C1

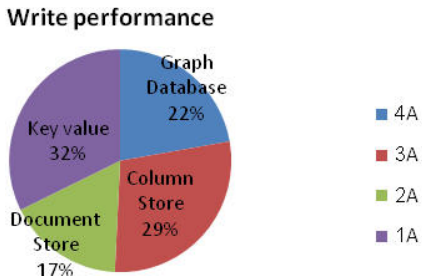

Fig. 2-b Percentages of each alternative with C2

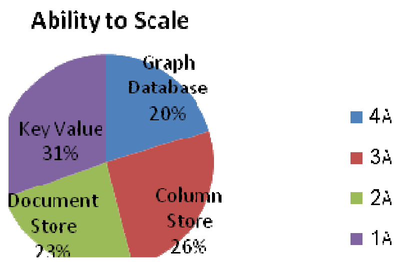

Fig. 2-c Percentages of each alternative with C3

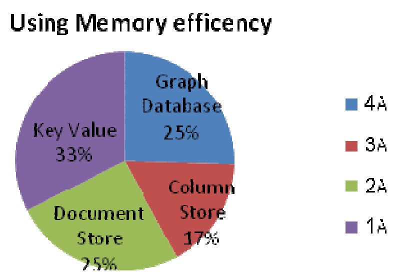

Fig.2-d Percentages of each alternative with C4

Regardless the kind of the operation or features, the overall evaluation for the NoSQL databases is showed in Fig 4. This figure describes the total performance by summing the evaluation degrees (defuzzification) of each one of NoSQL databases then calculating the average for each NoSQL databases.

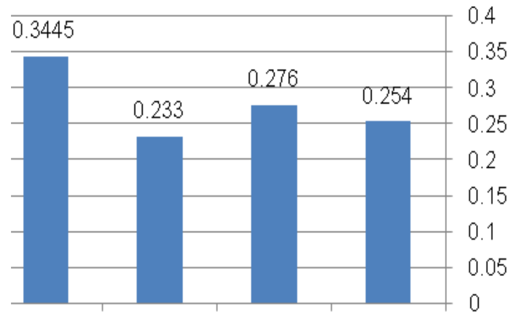

Fig.3 Total Performance evaluation of Databases

The total evaluation shows that Key Value databases have the high optimization in performance as a result to the speed of read and write operations as well as the efficiency in using the RAM. Column Store Databases have a good performance compared to Graph Database and Document Store. Document Store has the worst performance due to the evaluation degrees. It is important to remember that the database which has few or hopeless evaluation does not mean it is not suitable. It just mean that is not the best when comparing to the others.

Evaluation of Decision Models:

This study constructed the decision matrices and evaluated the results depending on CFPR and fuzzy LinPreRa. Data collection process has done by group decision making or experts as a result of comparing an alternative with the other alternative. This process facilitated giving the opinions as well as overcame the inconsistency by reducing the pairwise comparisons. Table 10 shows number of pairwise comparisons for each criterion in both fuzzy AHP and fuzzy LinPreRa.

Table 13. The pairwise comparisons for two methods

Name of Criteria	No.of alternatives	Fuzzy AHP n(n-1)/2	Fuzzy LinPreRa (n-1)
Read performance	4	6	3
Wright performance	4	6	3
Ability to scale	4	6	3
Using the memory efficiency	4	6	3
Summing of Comparison		**24**	**12**

Comparing to fuzzy AHP, The number of pairwise comparisons of fuzzy LinPreRa can be reduced by %50 or 12 times. Fig 4 shows number of pairwise comparisons in two different methods when evaluating the performance of NoSQL databases. Increasing the pairwise comparisons leads to increasing the human interventions, thus reducing the reliability of fuzzy decision models.

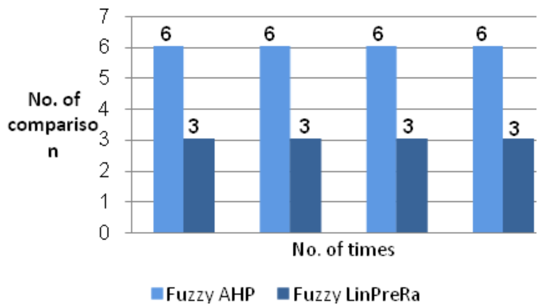

Fig 4. No of comparison of Fuzzy AHP and LinPreRa

Conclusion

This study adopted new model of MCDM namely fuzzy LinPreRa to evaluate the performance of NoSQL Databases.

In addition to their practical experience, five Experts used Cloud Serving Benchmark from Yahoo (YCSB) as open source program to analyze the difference performances from of NoSQL databases.

One contribution in this evaluation is obtaining the quantitative results instead of binary representation or point scale ranging as it has stated in previous works. The evaluation degrees (diffuzification) give the developers a comprehensive understanding to distinguish the highest importance compared to the others from the viewpoint of experts or group decision making. This paper can be a reference about the performance evaluation of NoSQL databases

The fuzzy LinPreRa has significant advantages over the fuzzy and conventional AHP because it has the ability for reducing the number of pairwise comparisons to nearly the half and avoiding the lack of consistency. This study proved that CFPR simply and practically provides the solutions to fuzzy decision making problem.

Another contribution in this study is enabling the developers or system analysts to specify the right choice for their large scale applications while building the databases.

Finally, this study must be embraced by organizations, corporations and private sectors as powerful tool to help in evaluating the criteria and decision making.

References

[1] N. Hanimah, K. Abdullah, and H. Mohamed, "Decision Making Models based on Consistent Fuzzy Preference Relations," Science and Engineering Research vol.3, no. 1, ,pp. 838-843, Jun 2011

[2] T. Hsieh, Lu, S.T. Lu. and G.H. Tzeng, "Fuzzy MCDM Approach for Planning and Design Tenders Selection in Public Office Building", International Journal of Project Management, pp.573-584. 2004.

[3] F. Herrera, F. Chiclana and M. Luque, "Some Issues on Consistency of Fuzzy Preference Relations," European Journal of Operational Research 154, pp. 98-109. 2012.

[4] T. C. Wang and Y. H. Chen, "Applying fuzzy linguistic preference relations to the improvement of consistency of fuzzy AHP," Information v Sciences,.vol. 137, no. 5. pp28-32, 2008.

[5] T.L. Saaty, "The Analytic Hierarchy Process, Planning, Priority Setting and Resource Allocation," McGraw-Hill, New York, 1980.

[6] L. Lin and R. Yeh, "A perceptual measure of mobile advertising using fuzzy linguistic

preference relation," Iranian journal of fuzzy systems Vol. 10, No. 5,pp. 14-22, Sep 2013.
[7] Y. H. Chen, "Applying Consistent Fuzzy Linguistic Preference Relations to E-Commerce System Selection," International Conference on Fuzzy Systems and Knowledge Discovery, IEEE computer society, vol. 12, no.4, pp. 252-256, May 2012.
[8].C. Leung, D. Cao, "On consistency and ranking of alternatives in fuzzy AHP," European Journal of Operational Research vol. 12 no. 90 pp.102–113. May 2013.
[9] S.T Lu, S.H Yu,D.S Chang, "Using the Fuzzy Linguistic Preference Relation Approach for Assessing the Importance of Risk Factors in a Software Development Project," Hindawi Publication Corporation, Mathematical problems in Engineering, vol. 126, no.7, pp. 232-256, 2013.
[10] T.C.Wang, Y.H.Chen, "Applying consistent fuzzy preference relation mechanisms to partnership selection," , vol. 25, no.9, pp. 49-56, May 2010.
[11] J. Rezaei and R. Ortt, "Multi-Criteria supplier segmentation using fuzzy preference relation based AHP," Europan Journal of operating research ,vol. 6, no.3,pp. 75-84, 2013.
[12] S.Lue, Jiun-Yi,and Dong-Chang, Development of Management Support System for Property Management Enterprises, Open Journal of Social Sciences, pp.127-132, 2014.
[13] M.Arora , M.Syamala, A Fuzzy-AHP Approach to solve Multi Criteria Budget Allocation Problem, IJCSET, July 2011
[14] J.Ricardo Lourenco, et al., "Choosing the right NoSQL database for the job: a quality attribute evaluation," Journal of Big Data, Springer open journal, 2015
[15] NoSQL- http://www.nosql.org

[16] Mcknight Consulting Group,"NoSQL EVALUATOR’S GUIDE," Global Service, 2015. Available from: http://www.mcknightcg.com/nosql-evaluators-guide
[17] R.Retch and S.Jablonski,"NoSQL Evaluation, Use Case Oriented Survey," International Conference on Cloud and Service Computing, vol. 98, no 1-4 pp. 336-341, 2011.
[18] Han, J., "Survey on NOSQL Databases," Proceedings 6th International Conference on Pervasive Computing and Applications, pp. 363- 366, 2012.
[19] Leavitt, N "Will NoSQL Databases Live up to Their Promise?," Computer Magazine, vol. 43, pp. 12- 14, 2010.
[20] Fonseca A, VuA, GrmanP "Evaluation of NoSQL databases for large-scale decentralized micro blogging," University Catalunya, 2013.