# Group Decision Making in Ontology Matching

Mahdieh Kargar-Ghavi
Department of Computer, Faculty of Engineering
University of Isfahan
Isfahan, Iran
0098 7934079
mahdieh.kargar@eng.ui.ac.ir

Mohammad Reza Khayyambashi
Department of Computer, Faculty of Engineering
University of Isfahan
Isfahan, Iran
0098 7934079
m.r.khayyambashi@eng.ui.ac.ir

## ABSTRACT
Ontology matching tries to establish semantic relations between similar elements in different ontologies to provide interoperability in the semantic web. Dealing with the problem of semantic heterogeneity is a key point in the semantic web environment. The (semi) automatic generating of mappings with respect to uncertainty is a labor intensive and error prone task. While the confidence values of mappings themselves are uncertain, how can the aggregation method between them be certain? How can we model them in a certain manner? This paper introduces a new approach for modeling uncertainty in ontology matching on the basis of fuzzy set theory and then describes an iterative algorithm that exploits group decision making solutions to aggregate opinions of matchers into a group consensus one. Thus matching systems are combined to overcome contradictory and incomplete alignments, so that the quality and accuracy of final alignment will be improved.

## Categories and Subject Descriptors
I.2.4 **[Knowledge Representation Formalisms and Methods]:** Ontologies- I.2.3 [**Deduction and Theorem Proving**]: Uncertainty, "fuzzy," and probabilistic reasoning

## General Terms
Algorithms, Management.

## Keywords
Ontology Matching, Mapping Aggregation, Fuzzy set theory, Decision Making, Semantic Web.

## 1. INTRODUCTION
Progress in information and communication technologies has made a lot of heterogeneous and distributed information available. Data will inevitably come from many different resources and providing interoperability among them is a crucial aspect of the semantic web. Interoperability has been transformed into a major and vital affair in the semantic web. Web designers have been undergoing semantic interoperability problems when using all potential facilities of the web.

Ontology is a form of knowledge representation in the semantic web. Ontologies are everywhere and ontology matching is one of the most plausible solutions to cope with heterogeneity problems in ontological contents. Ontology matching refers to the process of finding relations or correspondences between similar elements of different ontologies in the semantic web. It is a basic issue in many applications such as data integration, data warehouse, e-commerce and semantic query processing. Thus it is necessary to find the mapping between ontologies before processing across them. According to Euzenat et al. [6]: "Many different matching solutions have been proposed so far from various viewpoints, e.g., databases, information systems, artificial intelligence. They take advantage of various properties of ontologies, e.g., structures, data instances, semantics, or labels, and use techniques from different fields, e.g., statistics and data analysis, machine learning, automated reasoning, and linguistics. These solutions share some techniques and tackle similar problems, but differ in the way they combine and exploit their results. As a consequence, they are quite difficult to compare and describe, lacking a uniform framework…"

Nowadays, many matchers are available. Most of them combine the results of the individual matchers into a combined matching system in different ways. Dealing with uncertainty issues remains as a challenge for ontology matching [21] and has been addressed in a number of papers. The (semi) automatic ontology matching brings a degree of uncertainty. Besides uncertainty in the similarity matching method, ontology matching and integration tools are going to produce heuristic knowledge in the form of rules that also have uncertainty integrated with them [4]. This is a natural extension of existing methods. Essences of this paper are modeling uncertainty of mappings on the basis of fuzzy set theory and then aggregating them with a new method to deal with uncertainty issues and improve the matching process. Here the output results of state of the art matching systems are used rather than a single best matcher without any assumptions about them.

Recently, several researches have concentrated on mapping validation regarding the semantics of the ontologies involved in while they maintain the uncertain nature of mappings. A language is proposed by Nagy et al. [18] for representation of and reasoning with uncertain mappings. This approach combines ontology and rule languages with probabilistic reasoning and resolves inconsistencies by using trust probabilities, and also argues about these on a numeric level. Eckert et al. [5] has implemented a method in which the correctness of a correspondence is determined using a learning approach. It uses two information resources: the output of different matchers and the additional information about the nature of the elements to be matched. A classifier is trained on


*WIMS'11*, May 25–27, 2011, Sogndal, Norway.

the outcome of different state of the art matching systems and learned what combination of results from matchers provides the best indication of a correct correspondence. The difference lies in the use of the learning approach, limitation of one-to-one mappings between named concepts and properties, and not supporting 1:n or n:1 or n:m mappings. Another approach that is implemented [17, 19] introduces a new trust management method for resolving conflict among beliefs in similarities. Their solution uses fuzzy voting model to manage the conflicts by applying trust between the map agents. Isaac's approach [10] is another solution for selecting correspondences from a set of matcher-generated mappings. It re-uses an argumentation framework that considers the confidence levels of mappings and consensus between matchers. Another fuzzy approach which implies on fuzzy interpretation of mappings is proposed by Ferrara et al. [7], the great advantage of this solution is that it can present the degree of similarity between ontology element better than similar methods. Moreover, a method is described which computes the minimal sets of conflicting mappings and can be the basis of different validation approaches. The RiMOM (Risk Minimisation based Ontology Mapping) approach, that is based on Bayesian decision theory, considers ontology matching as a decision making problem [22]. Its processing is as follows. First several decisions find the mappings independently in multi-strategy execution. Then the outputs by the independent decisions are combined in strategy combination; thirdly some methods are used to discover the mappings based on the combined results in mapping discovery. This is an iterative process and can take place until no new mappings are discovered. User interaction is needed to refine the obtained mappings. It discoveries 1:1, n:1, 1:null and null:1 mappings.

To the best of our knowledge, no existing work has extensively introduced this model and used group decision making solution under the fuzzy set theory to improve the quality of existing ontology mappings.

The rest of the paper is organized as follows: Section 2 describes the terminologies. In section 3 the problem of uncertainty, especially in ontology matching context is briefly discussed. It also represents an interpretation of uncertainty on the basis of trapezoidal fuzzy numbers and concerns with describing the approach for solving the problem of aggregation in more detail. In section 4 a numerical example is represented. Finally, section 5 summarizes the paper and concludes with an outlook on future work.

# 2. ONTOLOGY MATCHING

## 2.1 Ontology

The term ontology, although first used in the area of philosophy, has been used by researchers in various areas such as Artificial Intelligence (AI), Information Retrieval (IR), database theory, linguistics, and e-commerce. There are many definitions of ontology for AI research field. Among them, the definition given by Gruber is most common, that is, an ontology is an explicit specification of a conceptualization [8]. An ontology $O$ is a 5-tuple $O=(C, I, R, F, A)$, where $C$ is the set of the concepts, $I$ is the set of individuals or instances, $R$ is the set of relationships defined on the set $C$, $F$ is the set of functions defined on the set $C$ and that return a concept, and $A$ is the set of axioms that constrain the interpretation and well-formed use of the vocabulary in some domains of discourse.

## 2.2 Ontology Matching

Ontology matching tries to establish semantic relations between similar elements in different ontologies to provide interoperability in the semantic web. Ontology matching with respect to uncertainty is a difficult process. The development of ontology matching has become an important field of ontology research [6]. Ontology matching takes a pair of ontologies as an input and creates the semantic correspondence relationships between these ontologies [20].

According to [6] the matching process can be seen as a function $f$ which, from a pair of ontologies to match $o$ and $o'$, an input alignment $A$, a set of parameters $p$ (e.g., weights and thresholds) and a set of oracles and resources $r$ (e.g., WordNet), returns an alignment $A'$ between these ontologies: $A' = f(o, o', A, p, r)$. This can be schematically represented as illustrated in Figure 1[6].

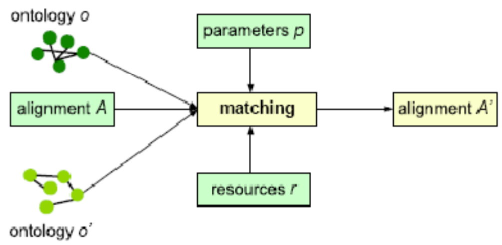


**Figure 1. The matching process**

The output of ontology matching process is called alignment. Alignment is a set of correspondences between two or more ontologies [6]. It can be used in many tasks such as ontology transformation, navigation and query answering on the web, ontology merging, and semantic web service composition.

A correspondence is a 5-uple: <*id, e, e', n, r*> and asserts that the relation $r$ holds between the ontology entities $e$ and $e'$ with confidence $n$, and $id$ is a unique identifier of the given correspondence [6]. The level of confidence $(n)$ is a real number of the interval [0, 1]. These correspondence relationships include equivalent, contain and disjoint relations and so on. Some authors use the term mapping instead. It will be used in this paper accordingly.

Ontology matching is the necessary requirement of semantic interoperability. Until now, the achievements of (semi)automatic ontology matching are very restricted. There are many challenges in this research field, and some problems need to be resolved [3].

# 3. AGGREGATION METHOD

There is no guarantee that two matching systems use the same notion for computing the degree of confidence. For instance, some automatic ontology matching systems use heuristics or the machine-learning techniques and they also use Dempster-Shafer theory, Bayesian Networks, rough set theory and so on to cope with uncertainty. For the detailed descriptions, one can refer to Euzenat et al. [6]. Here mappings are validated in the presence of uncertainty. Different methods for representing uncertainty in the context of various domains of applicability are presented by Klir [13], and they are briefly summarized as Classical set theory, Probability theory, Fuzzy set theory, Fuzzy measure theory, and Rough set theory [11].

The fuzzy set theory can be used in a wide range of domains in which information is incomplete or imprecise. Here matching systems are combined with respect to overcome contradictory and incomplete alignments by considering a fuzzy interpretation of mappings and defining each mapping as a trapezoidal fuzzy number and then using a group decision making algorithm. Fuzzy interpretation of a mapping states that the individuals of the first concept belong to the second concept with a certain degree, which is exactly the semantics of fuzzy membership functions [7]. The degree of membership is determined on the basis of strength of the similarity relation. Group decision making approach is used to combine uncertain outputs from multiple ontology matching systems. Here the output results of state of the art matching systems are used rather than a single best matcher without any assumptions about them.

## 3.1 Trapezoidal Fuzzy Number (TFN)

A fuzzy number [12] $A$ in $\Re$ (real line) is a normal TFN [9] (see Figure 2), if its membership function $f_A$: $\Re \rightarrow [0, 1]$ is:

$$f_A(x) = \begin{cases} (x-a)/(b-a), & a \leq x \leq b, \\ 1, & b \leq x \leq c, \\ (x-d)/(c-d), & c \leq x \leq d, \\ 0, & otherwise \end{cases} \quad (1)$$

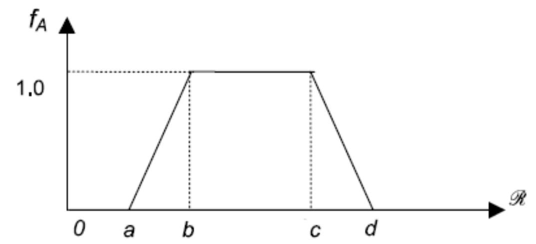


**Figure 2. Membership function of a trapezoidal fuzzy number *A=(a,b,c,d)***

The normal trapezoidal fuzzy number can be denoted by $A= (a, b, c, d)$. The interval $[b, c]$ gives the maximal grade of $f_A(x)$ i.e. $f_A(x) = 1$, $x \in [b, c]$, it is the most possible value of evaluation data. The $a$ and $d$ are the lower and upper bounds of the available area for the evaluation data. New aggregation method will be defined using the distance of two trapezoidal fuzzy numbers.

Cheng [1] ranked fuzzy numbers by using the distance between the centroid point of a fuzzy number and the original point and Chu [2] by using the area between them. In those papers, the centroid point of a normal trapezoidal fuzzy number $A= (a, b, c, d)$ was denoted by $(x, y)$, $x$ is a value on the horizontal axis and $y$ on the vertical axis. The real numbers $x$ and $y$ were calculated by using the following equation:

$$\begin{cases} x = \frac{1}{3}\frac{d^2+c^2-a^2-b^2+dc-ab}{d+c-a-b}, \\ y = \frac{w}{3}\frac{a+2b+2c+d}{a+b+c+d}, \end{cases} \quad (2)$$

Where in normal TFN $w=1$.

## 3.2 Decision Making Process

The process is defined in three phases:

1. Map generated confidence values of each matcher to a new space
2. Transform the mapped numbers of new space to fuzzy numbers
3. Catch common consensus on mappings

Since the imprecision and vagueness enter the alignments of matchers, individual mappings will be represented by fuzzy numbers and the issue will be solved with fuzzy set theory.

For mapping the generated confidence values of each matcher to a new space, at first the results of matching systems, i.e. $n$ alignments of state of the art matchers are selected. The minimum of $n$ is two. Each of them according to its method generates a set of mappings with some confidence values of the interval [0, 1]. Here only matchers that give alignments with equivalence relations must be used. Confidence values of each matcher are reordered in descending order and mapped to a new space of the interval [0, 1]; it is considered that first values are from the interval [*Min, Max*] where *Min* and *Max* are the minimum and maximum amounts between generated confidence values of each matcher. A simple proportion of mathematics is used to map these values to a new space. The values of each matcher are mapped independently of the other matchers because each of them has its own minimum and maximum.

A linguistic rating set is used for the transformation of new confidence values to the normal positive TFNs as follows. It is suggested [16] that the decision-makers utilize the linguistic rating set W= {VL, B.VL & L, L, B.L & M, M, B.M & H, H, B.H & VH, VH}, where VL=Very Low, B.VL & L=Between Very Low and Low, L=Low, B.L & M=Between Low and Medium, M=Medium, B.M & H=Between Medium and High, H=High, B.H & VH=Between High and Very High, VH=Very High, to evaluate the preference of objects versus different management strategy attributes. Herein, they are defined as: VL=(0,0,0,0.2), B.VL & L=(0,0,0.2,0.4), L=(0,0.2,0.2,0.4), B.L & M=(0,0.2,0.5,0.7), M=(0.3,0.5,0.5,0.7), B.M & H=(0.3,0.5,0.8,1), H= (0.6,0.8,0.8,1), B.H & VH=(0.6,0.8,1,1), and VH=(0.8,1,1,1) [15]. Each of them can at least belong to one of these TFNs.

Initial step: For matcher $M_i$, confidence value which is approximated as the normal positive TFN is denoted as follows:

$$\tilde{T}_{ij}^k = \left(a_{ij}^k, b_{ij}^k, c_{ij}^k, d_{ij}^k\right), \quad (3)$$

Where $i$ $(i = 1, 2... n)$ is the index of the matcher, $k$ is the iteration number whose initial value is one $(k=1)$ and $j$ $(j=1, 2...L)$ is the mapping index.

The proposed confidence values of different matchers for one mapping are placed in a set. In other words for $j$-th mapping in $k$-th iteration, $S_j^k$ is defined as:

$$S_j^k = \{\tilde{T}_{ij}^k, i = 1,2, \ldots, n\}. \quad (4)$$

The procedure introduced in [14] will be used to get the common consensus on mappings in an iterative manner:

Step 1: Calculate the centroid point of $\tilde{T}_{ij}^k$ as follows (for simplicity, index of the iteration (i.e. $k$) is not shown in the following equations):

$$\begin{cases} x_{ij} = \frac{1}{3}\frac{d_{ij}^2+c_{ij}^2-a_{ij}^2-b_{ij}^2+d_{ij}c_{ij}-a_{ij}b_{ij}}{d_{ij}+c_{ij}-a_{ij}-b_{ij}}, \\ y_{ij} = \frac{1}{3}\frac{a_{ij}+2b_{ij}+2c_{ij}+d_{ij}}{a_{ij}+b_{ij}+c_{ij}+d_{ij}}. \end{cases} \quad (5)$$

Step 2: For two positive TFNs $\tilde{T}_{ij}^k$ and $\tilde{T}_{rj}^k$ that denote confidence values for the $j$-th mapping from matcher $M_i$ and matcher $M_r$, the intersection of the area reflects the distance between $\tilde{T}_{ij}^k$ and $\tilde{T}_{rj}^k$ both on the horizontal and vertical axis. So the agreement degree $S_{ir}(j)$ can be defined between each pair of matchers $M_i$ and $M_r$ on the same mapping ($j$) as follows:

$$S_{ir}(j) = \frac{min\{x_{ij},x_{rj}\}\, min\{y_{ij},y_{rj}\}}{max\{x_{ij},x_{rj}\}\, max\{y_{ij},y_{rj}\}}, \quad (6)$$

Where $(x_{ij}, y_{ij})$ and $(x_{rj}, y_{rj})$ are the centroid points of $\tilde{T}_{ij}^k$ and $\tilde{T}_{rj}^k$, respectively.

Step 3: Calculate the Matcher's Average agreement degree, $MA_i(j)$, for the mapping $(j)$ as follows:

$$MA_i(j) = \frac{1}{n-1}\sum_{r=1,r\neq i}^{n} S_{ir}(j). \quad (7)$$

Step 4: Calculate the Matcher's Relative Agreement degree, $MRA_i(j)$, for each mapping ($j$) as follows:

$$MRA_i(j) = \frac{AM_i(j)}{\sum_{i=1}^{n} AM_i(j)}. \quad (8)$$

Step 5: User can assign a weight to each matcher that indicates the user's trust level to a matcher or relative importance of each matcher. For a group of matchers, the power of each one is extensively different; because each of them acts on different aspects of two ontologies such as concepts, instances, structure, etc., to extract mappings.

So beside exploring the interdependencies (i.e. $MRA_i(j)$) between different mappings, $\xi_i$ $(i = 1, 2... n)$ can be defined as the weight of matcher $M_i$, where $\xi_i$ satisfies:

$$\xi_i \geq 0 \;\; and \;\; \sum_{i=1}^{n} \xi_i = 1. \quad (9)$$

Calculate the Matcher's Consensus degree Coefficient, $MCC_i(j)$, for each mapping ($j$) using the following equation:

$$MCC_i(j) = \beta\xi_i + (1-\beta)MRA_i(j), \quad (10)$$

Where $\beta \in [0, 1]$. According to [14] if $\beta = 0$, the decision maker refuses to consider the weight of matchers, and if $\beta = 1$, the decision maker ignores to consider the opinions of matchers.

Step 6: Calculate the consensus opinion $\tilde{T}_{mj}^k$ for each mapping by using:

$$\tilde{T}_{mj}^k = \sum_{i=1}^{n}(CMC_i(j) \,.\, \tilde{T}_{ij}^k), \quad (11)$$

Where $\tilde{T}_{mj}^k$ for the $j$-th mapping is denoted as follows:

$$\tilde{T}_{mj}^k = (a_{mj}^k, b_{mj}^k, c_{mj}^k, d_{mj}^k). \quad (12)$$

After determination $\tilde{T}_{mj}^k$ if the process reaches a stable state, the algorithm terminates. The calculated $\tilde{T}_{mj}^k$ is considered as an approximation of confidence value for $j$-th mapping; Otherwise for each mapping, difference from the consensus opinion $\tilde{T}_{mj}^k$, is calculated as follows:

$$(a_{mj}^k - a_{ij}^k, b_{mj}^k - b_{ij}^k, c_{mj}^k - c_{ij}^k, d_{mj}^k - d_{mj}^k). \quad (13)$$

The information $c_{mj}^k - c_{ij}^k$ for each mapping is sent to each matcher and it goes to the next iteration. This information is sent in the form of alignment. According to Figure 1 a matcher gets an alignment as input beside the two ontologies.

After sending information, each matcher is asked to determine a new approximation for every mapping with respect to the consensus opinion. The new values, similar to the first step are displayed in the form of positive TFNs. Notice that algorithm has gone into the next iteration ($k+1$). Now with the new mappings, the algorithm returns to the initial step, replaces $k$ with $k+1$ and repeats. It will continue until the algorithm's termination condition is satisfied.

As mentioned in Step 6 when $\tilde{T}_{mj}^k$ for each mapping reaches a stable state, algorithm terminates and evaluated TFN is considered as an approximation of confidence value for $j$-th mapping. The two following methods are suggested for process stability and termination condition.

The first method: A natural number $t$ must be defined by the user as the number of iterations, and when $k=t$, the process reaches a stable state and the algorithm terminates. This method is very simple, but in some cases it isn't accurate. Therefore, this method requires the user to understand the nature of the matchers.

The second method: A $\varepsilon$ value is determined by the user. Whenever the difference between matcher's opinions $(\tilde{T}_{ij}^k)$ and consensus opinion $(\tilde{T}_{mj}^k)$ for each mapping is less than the amount of $\varepsilon$, it is said the process is in a stable state; $\delta$ is a simple difference function for TFNs that mentioned in equation (13). So the algorithm should terminate when the following condition is satisfied for all mappings:

$$\delta(\tilde{T}_{mj}^k, \tilde{T}_{ij}^k) < \varepsilon\,. \quad (14)$$

As you can see in the termination condition of algorithm, calculation of the second method is much more complicated than the first method, but the calculation accuracy is higher in this method.

# 4. NUMERICAL EXAMPLE

Consider an ontology matching problem evaluated by four matching systems under a special criterion and alternative for a pair of ontologies $O_1$ and $O_2$ (see Figure 3) [6].

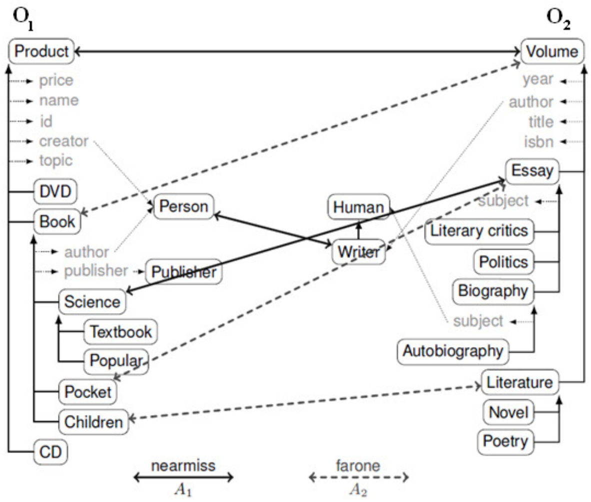


**Figure 3. Two class alignments between the fragments of two ontologies**

The output alignments of the four systems, $M_1$=Nearmiss, $M_2$=Farone, $M_3$=Align1 and $M_4$=Align2 are given in the form of similarity matrices as follows (see Table 1). The similarity matrix rows belong to ontology $O_1$ and its columns belong to ontology $O_2$. The two alignments, Align1 and Align2, are derived from Nearmiss and Farone alignments respectively. Now, we intend to combine these alignments using group decision making aggregation method. Inspecting the proposed method function is going to be considered in the following.

**Table 1. Similarity Matrices of the four alignments**

Nearmiss

| $O_1$ \ $O_2$ | Human | Essay | Volume | Writer | Literature |
|---|---|---|---|---|---|
| Person | 0.0 | 0.0 | 0.0 | **0.8** | 0.0 |
| Science | 0.0 | **0.8** | 0.0 | 0.0 | 0.0 |
| Book | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Product | 0.0 | 0.0 | **1.0** | 0.0 | 0.0 |
| Pocket | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Children | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

Farone

| $O_1$ \ $O_2$ | Human | Essay | Volume | Writer | Literature |
|---|---|---|---|---|---|
| Person | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Science | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Book | 0.0 | 0.0 | **1.0** | 0.0 | 0.0 |
| Product | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Pocket | 0.0 | **0.8** | 0.0 | 0.0 | 0.0 |
| Children | 0.0 | 0.0 | 0.0 | 0.0 | **0.8** |

Align1

| $O_1$ \ $O_2$ | Human | Essay | Volume | Writer | Literature |
|---|---|---|---|---|---|
| Person | 0.0 | 0.0 | 0.0 | **0.8** | 0.0 |
| Science | 0.0 | **0.8** | 0.0 | 0.0 | 0.0 |
| Book | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Product | 0.0 | 0.0 | **1.0** | 0.0 | 0.0 |
| Pocket | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Children | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

Align2

| $O_1$ \ $O_2$ | Human | Essay | Volume | Writer | Literature |
|---|---|---|---|---|---|
| Person | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Science | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Book | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Product | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Pocket | 0.0 | **0.6** | 0.0 | 0.0 | 0.0 |
| Children | 0.0 | 0.0 | 0.0 | 0.0 | **0.6** |

At first generated confidence values of each matcher are reordered in descending order and mapped to a new space of the interval [0, 1]. The mentioned linguistic rating set is used for the transformation of new confidence values to the normal positive TFNs. The results are shown in the first column of Table 2 using equation (3). Since, the zero elements of the similarity matrix is over, some elements of matrix with non-zero values will be shown explicitly and the $T_i(k,p)$ is the symbol of zero elements.

Initial step: There has no relative importance of each matching systems, i.e., $\beta = 0$.

Step 1: The centroid points of each mapping are calculated using equation (5) (see second column of Table 2).

**Table 2. The mappings and their centroid points in each alignment**

| $T_{1j}$ | $(x_{1j},y_{1j})$ |
|---|---|
| $T_1(1,4)=(0.3,0.5,0.8,1.0)$ | (0.6500,0.5000) |
| $T_1(2,2)=(0.3,0.5,0.8,1.0)$ | (0.6500,0.5000) |
| $T_1(4,3)=(0.6,0.8,1.0,1.0)$ | (0.8444,0.5098) |
| $T_1(k,p)=(0.0,0.0,0.0,0.2)$ | (0.0667,0.3333) |

| $T_{2j}$ | $(x_{2j},y_{2j})$ |
|---|---|
| $T_2(3,3)=(0.6,0.8,1.0,1.0)$ | (0.8444,0.5098) |
| $T_2(5,2)=(0.3,0.5,0.8,1.0)$ | (0.6500,0.5000) |
| $T_2(6,5)=(0.3,0.5,0.8,1.0)$ | (0.6500,0.5000) |
| $T_2(k,p)=(0.0,0.0,0.0,0.2)$ | (0.0667,0.3333) |

| $T_{3j}$ | $(x_{3j},y_{3j})$ |
|---|---|
| $T_3(1,4)=(0.3,0.5,0.8,1.0)$ | (0.6500,0.5000) |
| $T_3(2,2)=(0.3,0.5,0.8,1.0)$ | (0.6500,0.5000) |
| $T_3(4,3)=(0.6,0.8,1.0,1.0)$ | (0.8444,0.5098) |
| $T_3(k,p)=(0.0,0.0,0.0,0.2)$ | (0.0667,0.3333) |

| $T_{4j}$ | $(x_{4j},y_{4j})$ |
|---|---|
| $T_4(5,2)=(0.6,0.8,1.0,1.0)$ | (0.8444,0.5098) |
| $T_4(6,5)=(0.6,0.8,1.0,1.0)$ | (0.8444,0.5098) |
| $T_4(k,p)=(0.0,0.0,0.0,0.2)$ | (0.0667,0.3333) |

Step 2: The agreement degrees between each pair of matching systems are given by equation (6) on each mapping:

$$S_{11} = S_{22} = S_{33} = S_{44} =$$

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

| | | | | | |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 |

$S_{12} = S_{21} =$

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 1.0000 | 1.0000 | 0.0684 | 1.0000 | 1.0000 |
| 1.0000 | 0.0684 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.0516 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.0516 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 0.0684 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0684 | 1.0000 |

$S_{13} = S_{31} =$

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

$S_{14} = S_{41} =$

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 1.0000 | 1.0000 | 0.0684 | 1.0000 | 1.0000 |
| 1.0000 | 0.0684 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.0516 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 0.0516 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0516 | 1.0000 |

$S_{23} = S_{32} =$

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 1.0000 | 1.0000 | 0.0684 | 1.0000 | 1.0000 |
| 1.0000 | 0.0684 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.0516 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.0516 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 0.0684 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0684 | 1.0000 |

$S_{24} = S_{42} =$

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.0516 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 0.7549 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.7549 | 1.0000 |

$S_{34} = S_{43} =$

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 1.0000 | 1.0000 | 0.0684 | 1.0000 | 1.0000 |
| 1.0000 | 0.0684 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.0516 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 0.0516 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0516 | 1.0000 |

Step 3: The average agreement degrees of the four matching systems are given using equation (7) as follows:

$MA_1 =$

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 1.0000 | 1.0000 | 0.3789 | 1.0000 | 1.0000 |
| 1.0000 | 0.3789 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.6839 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.3677 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 0.3733 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.3733 | 1.0000 |

$MA_2 =$

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 1.0000 | 1.0000 | 0.3789 | 1.0000 | 1.0000 |
| 1.0000 | 0.3789 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.0516 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.3677 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 0.2972 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.2972 | 1.0000 |

$MA_3 =$

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 1.0000 | 1.0000 | 0.3789 | 1.0000 | 1.0000 |
| 1.0000 | 0.3789 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.6839 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.3677 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 0.3733 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.3733 | 1.0000 |

$MA_4 =$

| | | | | | |
|---|---|---|---|---|---|
| 1.0000 | 1.0000 | 1.0000 | 0.3789 | 1.0000 | 1.0000 |
| 1.0000 | 0.3789 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.6839 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 0.3677 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 0.2861 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.2861 | 1.0000 |

Step 4: The matcher's relative agreement degrees are calculated using equation (8) as follows:

$MRA_1 =$

| | | | | | |
|---|---|---|---|---|---|
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.3252 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2807 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2807 | 0.2500 |

$MRA_2 =$

| | | | | | |
|---|---|---|---|---|---|
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.0245 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2235 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2235 | 0.2500 |

$MRA_3 =$

| | | | | | |
|---|---|---|---|---|---|
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.3252 | 0.2500 | 0.2500 | 0.2500 |

| | | | | | |
|---|---|---|---|---|---|
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2807 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2807 | 0.2500 |

$MRA_4 =$

| | | | | | |
|---|---|---|---|---|---|
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.3252 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2151 | 0.2500 | 0.2500 | 0.2500 | 0.2500 |
| 0.2500 | 0.2500 | 0.2500 | 0.2500 | 0.2151 | 0.2500 |

Step 5: Omit it for $\beta = 0$; $MCC_i = MRA_i$.

Step 6: The consensus opinion of the four matching systems is shown below using equation (11).

$a_m =$

| | | | | | |
|---|---|---|---|---|---|
| 0.0000 | 0.0000 | 0.0000 | 0.1500 | 0.0000 | 0.0000 |
| 0.0000 | 0.1500 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | 0.0147 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | 0.3000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.1961 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.1961 | 0.0000 |

$b_m =$

| | | | | | |
|---|---|---|---|---|---|
| 0.0000 | 0.0000 | 0.0000 | 0.2500 | 0.0000 | 0.0000 |
| 0.0000 | 0.2500 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | 0.0196 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | 0.4000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.2838 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.2838 | 0.0000 |

$c_m =$

| | | | | | |
|---|---|---|---|---|---|
| 0.0000 | 0.0000 | 0.0000 | 0.4000 | 0.0000 | 0.0000 |
| 0.0000 | 0.4000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | 0.0245 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | 0.5000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.3939 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3939 | 0.0000 |

$d_m =$

| | | | | | |
|---|---|---|---|---|---|
| 0.2000 | 0.2000 | 0.2000 | 0.6000 | 0.2000 | 0.2000 |
| 0.2000 | 0.6000 | 0.2000 | 0.2000 | 0.2000 | 0.2000 |
| 0.2000 | 0.2000 | 0.2196 | 0.2000 | 0.2000 | 0.2000 |
| 0.2000 | 0.2000 | 0.6000 | 0.2000 | 0.2000 | 0.2000 |
| 0.2000 | 0.5509 | 0.2000 | 0.2000 | 0.2000 | 0.2000 |
| 0.2000 | 0.2000 | 0.2000 | 0.2000 | 0.5509 | 0.2000 |

If all confidence values that belong to the set of one mapping (i.e. $S_j^k$ ) are identical, the combined result is the common estimation. Also, the result of the method will not depend on the order with which matcher's opinions are combined. If there is no intersection for any *i, r (i, r =1, 2... n),* the consensus opinion $\tilde{T}_{mj}^k$ can also be derived.

In the second method of algorithm's termination condition, in special situations, a matcher may provide opinions different from others, insist on its opinion and repeat it in various iterations. In this case, a significant difference between group consensus opinion and this matcher is created and the process will not be stable. If one matcher's estimation is far from the other matchers, its estimation is of little importance. In cases where such a condition is caused, it is recommended to use heuristic methods like omitting inappropriate opinions from the considered set of one mapping.

## 5. CONCLUSINS

Clearly, Ontology matching with respect to uncertainty is a difficult task. In this paper a new model was considered to represent uncertainty in ontology matching context, which demonstrated the notion of uncertainty better. A new similarity aggregation method on the basis of Fuzzy set theory was used to aggregate the individual opinions of matchers into a group consensus one under group decision making solutions. Weaknesses and strengths of this approach had been argued in the paper. Combining two ontology matching systems A and B with complementary strengths/weaknesses, and computing a consensus using the procedure presented here obviously yielded a result that is better than A where A is weaker than B, and worse than A where A is stronger than B.

Thus Matching systems were combined to overcome contradictory and incomplete alignments, so that the quality and accuracy of final alignment was improved. In this case, any assumptions about the employed state of the art matchers were not made. This makes the approach more widely applicable.

For future work, improving the current aggregation approach and implementing a framework especially for complex mappings and dealing with inconsistencies that are produced by different matchers can be continued. Evaluating whether such a combined result is useful, or determining good weights (algorithm step 5) will be subjected to future research. On the other hand, investigating the uncertainty issues in ontology matching and using different uncertainty theories to deal with various situations in ontology matching will be considered.

## 6. REFERENCES


[1] Cheng, C. H. 1998. A new approach for ranking fuzzy numbers by distance method. *Fuzzy Sets and Systems 91,* 95**,** 307-317.

[2] Chu, T. C. and Tsao, C. T. 2002. Ranking fuzzy numbers with an area between the centroid point and original point. *Comput. and Math. with Applications,* 43**,** 111-117.

[3] Ding, Y. and Foo, S. 2002. Ontology research and development Part 1 - A review of ontology generation[J]. . *Journal of Information Science,* 28(2)**,** 123-136.

[4] Doan, A., Madhavan, J., Dhamankar, R., Domingos, P. and Halevy, A. 2003. Learning to match ontologies on the semantic web. *The VLDB Journal — The International Journal on Very Large Data Bases,* 12**,** 303 - 319

[5] Eckert, K., Meilicke, C. and Stuckenschmidt, H. 2009. Improving ontology matching using meta-level learning. *Proceedings of the 6th european semantic web conference (ESWC-09),* . Heraklion, Greece, ACM, New York, 158-172.

[6] Euzenat, J. and Shvaiko, P. 2007. *Ontology matching*, Springer-Verlag

[7] F Ferrara, A., Lorusso, D., Stamou, G., Stoilos, G., Tzouvaras, V. and Venetis, T. 2008. Resolution of conflicts

among ontology mappings: a fuzzy approach. *In Proceedings of the 3rd International Workshop on Ontology Matching (OM-2008), Collocated with the 7th International Semantic Web Conference (ISWC-2008),* . Karlsruhe, Germany.

[8] Gruber, T. R. 1993. A translation approach to portable ontology specifications. *Knowledge Acquisition,* 5(2)**,** 199-220.

[9] Heilpern, S. 1997. Representation and application of fuzzy numbers. *Fuzzy Sets and Systems 91***,** 259-268.

[10] Isaac, A., Trojahn, C. A., Wang, S. and Quaresma, P. 2008. Using quantitative aspects of alignment generation for argumentation on mappings. *Proceedings of the ISWC 2008 Workshop on Ontology Matching.* Karlsruhe, Germany.

[11] Isukapalli, S. S. 1999. Uncertainty analysis of Transport-Transformation models *New Brunswick Rutgers, New Jersey State University.* New Brunswick, New Jersey ,United States of America, Rutgers.

[12] Kaufmann, A. and Gupta, M. 1991. *Introduction to Fuzzy Arithmetic Theory and Applications,* London, International Thomson Computer.

[13] Klir, G. J., Ayyub, B. M. and Gupta, M. M. 1994. The many faces of uncertainty. *Uncertainty Modeling and Analysis: Theory and Applications, Elsevier Science,North-Holland Elsevier Science, Cambridge, UK,* 3-19.

[14] Li, Q. and Yang, J. 2008. Aggregation of fuzzy opinions with an area between the centroid point and the original point under group decision making *Fuzzy Systems, 2008. FUZZ-IEEE 2008.(IEEE World Congress on Computational Intelligence), 163-167.*

[15] Liang, G.-S., Chou, T.-Y. and Han, T.-C. 2005. Cluster analysis based on fuzzy equivalence relation. *European Journal of Operational Research,* 166**,** 160–171.

[16] Liang, G. S. and Wang, M. J. 1991. A fuzzy multi-criteria decision-making method for facility site selection. *International Journal of Product Research,* 29(11), 2313-2330.

[17] Nagy, M. and Vargas-Vera, M. 2009. Reaching consensus over contradictory interpretation of semantic web data for ontology mapping *IEEE 5th International Conference on Intelligent Computer Communication and Processing(ICCP).* Cluj-Napoca ,Romania, 63-66.

[18] Nagy, M., Vargas-Vera, M. and Motta, E.. 2007. DSSim - managing uncertainty on the semantic web. *ISWC+ASWC Workshop on Ontology Matching.* Busan, Busan, 160-169.

[19] Nagy, M., Vargas-Vera, M. and Motta, E. 2008. Multi-agent conflict Resolution with Trust for Ontology Mapping. *Intelligent Distributed Computing, Systems and Applications.* Springer Berlin / Heidelberg, 275-280.

[20] Rahm, E. and Bernstein, P. A. 2001. A survey of approaches to automatic schema matching. *The VLDB Journal — The International Journal on Very Large Data Bases* 10**,** 334 - 350

[21] Shvaiko, P. and Euzenat, J. 2008. Ten Challenges for Ontology Matching. *Proceedings of the OTM 2008 Confederated International Conferences, CoopIS, DOA, GADA, IS, and ODBASE 2008. Part II on On the Move to Meaningful Internet Systems.* Monterrey, Mexico, Springer-Verlag.

[22] Tang, J., Li, J., Liang, B., Huang, X., Li, Y. and Wang, K. 2006. Using Bayesian decision for ontology mapping. *Journal ofWeb Semantics,* 4(1)**,** 243-262.