Top-k web services compositions: A fuzzy-set-based approach
Karim Benouaret, Djamal Benslimane, Allel Hadjali

To cite this version:

Karim Benouaret, Djamal Benslimane, Allel Hadjali. Top-k web services compositions: A fuzzy-set-based approach. ACM - Symp. on Applied Computing (SAC), Mar 2011, Taiwan. pp.1038-1043. hal-00670704

HAL Id: hal-00670704
https://hal.archives-ouvertes.fr/hal-00670704
Submitted on 15 Feb 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Top-k Service Compositions: A Fuzzy Set-Based Approach

Karim Benouaret  
Lyon 1 University  
69622, Villeurbanne, France  
karim.benouaret@liris.cnrs.fr

Djamal Benslimane  
Lyon 1 University  
69622, Villeurbanne, France  
djamal.benslimane@liris.cnrs.fr

Allel Hadjali  
Enssat-University of Rennes 1  
22305, Lannion, France  
allel.hadjali@enssat.fr

ABSTRACT
Data as a Service (DaaS) is a flexible way that allows enter-
prises to expose their data. Composition of DaaS services
provides bridges to answer queries. User preferences are
becoming increasingly important to personalize the com-
position process. In this paper, we propose an approach to
compose DaaS services in the context of preference queries
where preferences are modeled by means of fuzzy sets that
allow for a large variety of flexible terms such as 'cheap', 'af-
fordable' and 'fairly expensive'. The proposed approach is
based on RDF-based query rewritings to take into account
the partial matching between individual DaaS services and
parts of the user query. Matching degrees between DaaS
services and fuzzy preference constraints are computed by
means of different constraints inclusion methods. Such de-
grees express to which extent a service is relevant to the
resolution of the query. A fuzzification of Pareto dominance
is also proposed to better rank composite services by com-
puting the score of services. The resulting scores are then
used to compute the top-k DaaS service compositions that
cover the user query.

Categories and Subject Descriptors
H.3.3 [Information Search and Retrieval]: Query for-
mulation, Search process, Selection process; H.3.5 [Online
Information Services]: Web-based services

General Terms
Measurement, Theory, Performance

Keywords
Fuzzy dominance, Web service, Top-k compositions, fuzzy
preferences queries.

1. INTRODUCTION
Nowadays, companies are increasingly moving towards a
service-oriented architecture for data sharing on the Web

Permission to make digital or hard copies of all or part of this work for
personal or classroom use is granted without fee provided that copies are
not made or distributed for profit or commercial advantage and that copies
bear this notice and the full citation on the first page. To copy otherwise,
to republish, to post on servers or to redistribute to lists, requires prior specific
permission and/or a fee.

SMC’11 / March 21-25, 2011, TaiChung, Taiwan.
Copyright 2011 ACM 978-1-4503-0113-8/11/03 ...$10.00.
Table 1: Example of DaaS Services

<table>
<thead>
<tr>
<th>Service</th>
<th>Functionality</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{11}(x, y)$</td>
<td>Returns the automakers $y$ whose country is $x$</td>
<td>$z$ is cheap, $t$ is short</td>
</tr>
<tr>
<td>$S_{21}(x, y, z, t)$</td>
<td>Returns the cars $y$ along with their prices $z$ and warranties $t$ for a given automaker $x$</td>
<td>$z$ is accessible, $t$ is long</td>
</tr>
<tr>
<td>$S_{22}(x, y, z, t)$</td>
<td>$z$ is expensive, $t$ is a car</td>
<td></td>
</tr>
<tr>
<td>$S_{23}(x, y, z, t)$</td>
<td>$z$ is [9000, 14000], $t$ is [6, 24]</td>
<td></td>
</tr>
<tr>
<td>$S_{31}(x, y, z)$</td>
<td>Returns the power $y$ and the consumption $z$ for a given car $x$</td>
<td>$y$ is weak, $z$ is small</td>
</tr>
<tr>
<td>$S_{32}(x, y, z)$</td>
<td>$y$ is ordinary, $z$ is approximately 4</td>
<td></td>
</tr>
<tr>
<td>$S_{33}(x, y, z)$</td>
<td>$y$ is powerful, $z$ is high</td>
<td></td>
</tr>
<tr>
<td>$S_{34}(x, y, z)$</td>
<td>$y$ is [60, 110], $z$ is [3.5, 5.5]</td>
<td></td>
</tr>
</tbody>
</table>

2.1 Preference Queries

Users express their conjunctive preference queries over domain ontologies using a slightly modified version of SPARQL query language. For instance, query $Q_1$ given in Section 1 is expressed as follows:

```
SELECT ?n ?pr ?w ?pw ?co
WHERE {?Au rdf:type AutoMaker ?Au hasNationality 'French'
?C hasPrice ?pr ?C hasWarranty ?w
?C hasPower ?pw ?C hasConsumption ?co
Preferring {?pr is 'affordable', ?w is 'around 18',
?pw is 'normal', ?co is 'medium' }
```

In our approach, user preferences are modeled using fuzzy sets. Formally, a fuzzy set $F$ on a referential $X$ is characterized by a membership function $\mu_F: X \rightarrow [0, 1]$, where $\mu_F(x)$ represents the grade of membership of $x$ in $F$. In particular, $\mu_F(x) = 1$ reflects full membership of $x$ in $F$, $\mu_F(x) = 0$ absolute non-membership and $0 < \mu_F(x) < 1$ partial membership. Namely having $x, y \in F$, $x$ is more preferable than $y$ iff $\mu_F(x) > \mu_F(y)$. If $\mu_F(x) = \mu_F(y)$, $x$ and $y$ are equally preferred. In practice, membership functions are often of trapezoidal form represented by the quadruplet $(A, B, a, b)$. A regular interval $[A, B]$ can be seen as a fuzzy set represented by the quadruplet $(A, B, 0, 0)$. The semantics of all fuzzy terms of Table 1 are available in the URL: http://vm.liris.cnrs.fr:36880/FuzzyTerms/.

2.2 DaaS services

Let us assume that DaaS services are partitioned into different classes of services. A class $S_j$ represents services with the same inputs, outputs, and providing the same functionality but different (fuzzy) constraints. A DaaS service $S_{ji}$ of class $S_j$ is described as a predicate $S_{ji}(X_j, Y_j, Z_j) : = \langle G_j(X_j, Y_j, Z_j), C_{ji} \rangle >$ where:

- $X_j$ and $Y_j$ are the sets of input and output variables of $S_{ji}$, respectively. Input and output variables are also called distinguished variables. They are prefixed with the symbols ‘$’ and ‘?’”, respectively.
- $G_j(X_j, Y_j, Z_j)$ represents the functionality of the service. This functionality is described as a semantic relationship between input and output variables. $Z_j$ is the set of existential variables relating $X_j$ and $Y_j$.
- $C_{ji} = \{C_{ji1}, \ldots, C_{jim}\}$ is a set of constraints expressed as intervals or fuzzy sets on $X_j$, $Y_j$ or $Z_j$ variables.

$X_j$ and $Y_j$ variables are defined in the WSDL description of DaaS services. Functionality $G_j$ and constraints $C_{ji}$ of a DaaS service $S_{ji}$ are added to the standard WSDL descriptions in the form of annotations. The annotations are represented in the form of SPARQL queries. For instance, the following SPARQL query illustrates the functionality and service constraints of the DaaS service $S_{21}$:

```
SELECT ?n ?pr ?w
WHERE {?Au rdf:type AutoMaker ?Au name ?n
?C hasPrice ?pr
?C hasWarranty ?w ?w is 'cheap' ?w is 'short'}
```

2.3 Discovering Relevant Services

Let $Q$ be a conjunctive preference query. We use an efficient RDF query rewriting algorithm to discover the parts
of $Q$ that are covered by each service—recall that in the general case services may match only parts (referred to by $q_j$) of $Q$. The same part $q_j$ is in general covered by one or more services that constitute a class of relevant services and is designated as class $S_j$. A service $S_i \in S_j$ is said to be relevant to a query $Q$ if the functionality of $S_i$ completely or partially matches the part $q_j$ and its constraints match completely or partially the preference constraints of $q_j$ (i.e., the matching degree is strictly higher than 0).

As preference constraints of a query are expressed in the rich fuzzy sets framework, their matching degrees with DaaaS services constraints may differ from one constraints inclusion method (CIM) to another. Each relevant service is then associated with $|M|$ matching degrees (if $M = \{m_1, ..., m_{|M|}\}$ is the set of the used methods). For instance, Table 2 shows the matching degrees between each service $S_i$ from Table 1 and its corresponding component $q_j$ (of the query $Q_i$). The degrees are computed by applying the following CIMs:

- **Cardinality-based method** (CBM) [15]. Let $C$ and $C'$ be two fuzzy constraints, $\text{Deg}(C \subseteq C') = \frac{|C \cap C'|}{|C|}$.
- **Implication-based method** (IBM) [1]. $\text{Deg}(C \subseteq C') = \min_j(\mu_C(x) \rightarrow \mu_{C'}(x))$, where $\rightarrow$ stands for a fuzzy implication. The following IBMs are used for our running example: Godel (G-IBM), Lukasiewicz (L-IBM) and Kleene-Dienes (K-IBM).

Due to lack of space, we do not present these methods. The service $S_{11}$ covering the component $q_1$ does not have a matching degree because there are no constraints imposed by the user on $q_1$. However, each service covering the component $q_2$ is associated with four (the number of methods) degrees. Each matching degree is formulated as a pair of real values within the range $[0, 1]$, where the first and second values are the matching degrees of the constraints $\text{price}$ and $\text{warranty}$, respectively. Similarly, for the matching degrees of the services covering the component $q_3$, the first and second values represent the inclusion degrees of the constraints $\text{power}$ and $\text{consumption}$, respectively.

Table 2: Matching Degrees between services and fuzzy preference constraints of $Q_i$

<table>
<thead>
<tr>
<th>$S_i$</th>
<th>$q_j$</th>
<th>CBM</th>
<th>G-IBM</th>
<th>L-IBM</th>
<th>K-IBM</th>
</tr>
</thead>
<tbody>
<tr>
<td>S_{11}</td>
<td>q_1</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>S_{21}</td>
<td>q_1</td>
<td>(1.0, 0.57)</td>
<td>(1.0)</td>
<td>(1.0)</td>
<td>(0.80, 0)</td>
</tr>
<tr>
<td>S_{22}</td>
<td>q_2</td>
<td>(0.89, 1)</td>
<td>(0.90, 1)</td>
<td>(0.50, 1)</td>
<td></td>
</tr>
<tr>
<td>S_{23}</td>
<td>q_2</td>
<td>(0.20, 0.10)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
</tr>
<tr>
<td>S_{24}</td>
<td>q_2</td>
<td>(0.83, 0.88)</td>
<td>(0.60, 0.50)</td>
<td>(0.60, 0.50)</td>
<td>(0.60, 0.50)</td>
</tr>
<tr>
<td>S_{31}</td>
<td>q_3</td>
<td>(0.50, 0.30)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
</tr>
<tr>
<td>S_{32}</td>
<td>q_3</td>
<td>(0.79, 0.75)</td>
<td>(0.25)</td>
<td>(0.60, 0.50)</td>
<td>(0.40, 0.50)</td>
</tr>
<tr>
<td>S_{33}</td>
<td>q_3</td>
<td>(0.21, 0.64)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
</tr>
<tr>
<td>S_{34}</td>
<td>q_3</td>
<td>(0.83, 0.85)</td>
<td>(0.50, 0.50)</td>
<td>(0.50, 0.50)</td>
<td>(0.50, 0.50)</td>
</tr>
</tbody>
</table>

### 2.4 Problem Statement

Given a query $Q = q_1, ..., q_n$ where each $q_j$ is a subquery (query component), $q_j$ is a tuple $q_j = (\bar{q}_j, P_{q_j})$, where $\bar{q}_j$ represents $q_j$ without its preferences $P_{q_j}$. Given a set of services classes $\{S_1, ..., S_n\}$ where each class $S_j$ regroups the services that are relevant to a query part $q_j$. Given $|M|$ CIMs to compute the matching degrees between the constraints of relevant services and the user’s preference. The problem we are interested in is how to rank DaaaS services in each class $S_j$ to select the most relevant services and how to rank generated DaaaS service compositions to select the top-$k$ ones that answer the fuzzy query $Q$.

### 3. Fuzzy Dominance and Fuzzy Scores

#### 3.1 Fuzzy dominance v.s. Pareto dominance

Services of the same class $S_j$ have the same functionality, they only differ in terms of constraints, providing thus different matching degrees. Individual matching degrees of services could be aggregated. One method is to assign weights to individual degrees and, for instance, compute a weighted average of degrees. In so doing, users may not know enough to make trade-offs between different relevancies using numbers (average degrees). Users thus lose the flexibility to select their desired answers by themselves. Computing the skyline [3] as a natural solution to overcome this limitation. The skyline consists of the set of points which are not dominated by any other point.

**Definition 1.** (Pareto dominance) Let $u$ and $v$ be two $d$-dimensional points. We say that $u$ dominates $v$, denoted by $u \succ v$, iff $\forall i \in [1,d], u_i \geq v_i \wedge \exists k \in [1,d], u_k > v_k$.

Pareto dominance is not always significant to rank-order points that, however, seem comparable from a user point of view. To illustrate this situation, let $u = (u_1, u_2) = (1, 0)$ and $v = (v_1, v_2) = (0.90, 1)$ be two matching degrees. In Pareto order, we have neither $u \succ v$ nor $v \succ u$, i.e., the instances $u$ and $v$ are incomparable. However, one can consider that $v$ is better than $u$ since $v_2 = 1$ is too much higher than $u_2 = 0$, contrariwise $v_1 = 0.90$ is almost close to $u_1 = 1$. This is why it is interesting to fuzzify the dominance relationship to express the extent to which a matching degree (more or less) dominates another one. In line with the general fuzzification dominance approach discussed in [8], we define below a fuzzy dominance relationship that relies on particular membership function of a graded inequality of the type **stronger larger than**.

**Definition 2.** (Fuzzy dominance) Given two $d$-dimensional points $u$ and $v$, we define the fuzzy dominance to express the extent to which $u$ dominates $v$ such as:

$$\text{deg}(u \succ v) = \sum_{i=1}^{d} \mu_{\succ u}(u_i, v_i)$$

(1)

Where $\mu_{\succ u}(u_i, v_i)$ expresses the extent to which $u_i$ is more or less (strongly) greater than $v_i$. $\mu_{\succ u}$ can be seen as a monotone membership function defined as:

$$\mu_{\succ u}(x, y) = \begin{cases} 1 & \text{if } x - y > \lambda + \varepsilon \\ 0 & \text{if } x - y \leq \varepsilon \\ \frac{x - y}{\lambda} & \text{otherwise} \end{cases}$$

(2)

Where $\lambda > 0$, i.e. $\succ$ is more demanding than the idea of 'strictly greater'. We should also have $\varepsilon \geq 0$ in order to ensure that $\succ$ is a relation that agrees with the idea of 'greater' in the usual sense.

Let us reconsider the previous instances $u = (1, 0), v = (0.90, 1)$, with $\varepsilon = 0$ and $\lambda = 0.2$. We have $\text{deg}(u \succ v) = 0.25$ and $\text{deg}(v \succ u) = 0.5$. This is more significant than $|u \succ v| = |v \succ u| = 0$ provided by Pareto dominance, where $|u \succ v| = 1$ if $u \succ v$, 0 otherwise. In the following sections, we use the defined fuzzy dominance to compute scores of services and compositions.
3.2 Associating fuzzy score with a service

It is well known that under a single matching degree method (mono criteria), the dominance relationship is unambiguous. When multiple CIM are applied (multi-criteria), resulting in different matching degrees for the same couple of constraints, the dominance relationship becomes uncertain. The model proposed in [10], namely probabilistic skyline overcomes this problem. Contrariwise, Skoutas et al. show in [11] the limitations of the probabilistic skyline to rank Web services and introduce the Pareto dominating score of individual services. We generalize this score to fuzzy dominance and propose a fuzzy dominating score (FDS). An FDS of a service $S_{ji}$ indicates the average extent to which $S_{ji}$ dominates the whole services of its class $C_j$.

**Definition 3.** (Fuzzy dominating score of a service) The fuzzy dominating score ($FDS$) of a service $S_{ji}$ in its class $C_j$ is defined as:

$$FDS(S_{ji}) = \frac{1}{(|S_j| - 1)|C_j|} \sum_{h=1}^{M} \sum_{k \in S_j} \sum_{r=1}^{M} deg(S_{ji} > S_{jk})$$

where $S_{ji}$ is the matching degree of the service $S_{ji}$ obtained by applying the $k^{th}$ CIM. The term $(|S_j| - 1)$ is used to normalize the FDS score and make it in the range $[0...1]$. Table 3 shows the fuzzy dominating scores of the DaaS services of our running example.

3.3 Associating fuzzy score with a composition

Different DaaS service compositions can be generated from different $C_j$ service classes to answer a user query. To rank such generated compositions, we extend the previous FDS definition to service composition and associate each composition with an $FDS$. The $FDS$ of a composition $C$ is an aggregation of different FDSs of its component services.

Let $C = \{S_{i1}, ..., S_{in}\}$ be a composition of $n$ services. Let also $d = d_1 + ... + d_n$ be the number of user preference constraints where $d_i$ is the number of constraints involved in the service $S_{ji}$. The $FDS$ of $C$ is then computed as follows:

$$FDS(C) = \frac{1}{d} \sum_{j=1}^{n} d_j \cdot FDS(S_{ji})$$

It is important to note that not all compositions are valid. A composition $C$ of services is valid if (i) it covers the user query $Q$, (ii) it contains one and only one service from each service class $S_j$, and (iii) it is executable. A composition is said to be executable if all input parameters necessary for the invocation of its component services are bound.

4. **TOP-K DAAS SERVICE COMPOSITIONS**

4.1 An Efficient Generation of Top-k Compositions

A straightforward method to find the top-$k$ compositions that answer a query is to generate all possible compositions, compute their scores, and return the top-$k$ ones. However, this approach results in a high computational cost, as it needs to generate all possible compositions, whereas, most of them are not in the top-$k$. In the following, we provide an optimization technique to find the top-$k$ compositions. This technique allows eliminating relevant services $S_{ji}$ from their classes $S_j$ before generating the compositions, i.e. services that we are sure that if they are composed with other ones, the obtained compositions are not in the top-$k$. The idea is: we first compute the score of each service in its class, then only the best services in each class are retained, after that we compose the retained services, finally, we compute the score of the obtained compositions and return the top-$k$ ones. To this end, we introduce the following theorem.

**Theorem 1.** Let $C = \{S_{i1}, ..., S_{in}\}$ be a composition. Let top-k($S_{ji}$) (resp. top-k(compositions)) be the top-$k$ services of the class $S_j$ (resp. the top-$k$ compositions). Then, $\exists S_{ji} \in C; S_{ji} \notin \text{top-k}(S_j) \Rightarrow C \notin \text{top-k}(\text{compositions})$

This theorem means that the top-$k$ sets of the different service classes $C_j$ are sufficient to compute the top-$k$ compositions that answer the query at hand.

The fourth column of Table 3 shows the top-$k$ ($k = 2$) services in each service class using the FDS scores. Thus, relevant DaaS services that are not in the top-$k$ of their classes are eliminated. They are crossed out in Table 3. The other DaaS services are retained. The top-$k$ DaaS service compositions are generated from the different top-$k$ $S_j$ classes. Table 4 shows the possible compositions along with their fuzzy scores and the top-$k$ compositions of our running example.

4.2 Top-k DaaS Service Compositions Algorithm

The algorithm, hereafter referred as $TKDSC$, computes the top-$k$ DaaS service compositions according to the fuzzy scores. The algorithm proceeds as following.

**Step.1** compute the matching degrees (lines 1-13). Each class whose services cover a query component is added to the list of relevant classes. If its services touch the query’s user preferences, we compute its different matching degrees according to the number of methods.
Algorithm 1 TKDSC

Require: Q a preference query, S a set of service classes, M = \{m_1, ..., m_{|M|}\} a set of methods, k ∈ N, ε, λ
1: for all S_j in S do
2: \( S \leftarrow \text{random}(S_j, 1) \);
3: if \( \exists q_j \in Q; \text{cover}(S_j, q_j) \) then
4: \( R \leftarrow R \cup S_j \);
5: if \( P_{ij} \neq \emptyset \) then
6: for all \( S_{ij} \in S_j \) do
7: for all \( m \) in \( M \) do
8: \( \text{MatchingDegree}(C_{ji}, P_{ij}, m) \);
9: end for
10: end for
11: end if
12: end if
13: end for
14: for all \( S_j \) in \( R \) do
15: if \( P_{ij} = \emptyset \) then
16: \( \text{top-k}, S_j \leftarrow \text{random}(S_j, k) \);
17: else
18: for all \( S_{ij} \in S_j \) do
19: \( \text{ServiceScore}(S_{ij}) \);
20: end for
21: \( \text{top-k}, S_j \leftarrow \text{top}(k, S_j) \);
22: end if
23: end for
24: \( C \leftarrow \text{ComposeServices}(\text{top-k}, S_{j1}, ..., \text{top-k}, S_{jm}) \);
25: for all \( C \) in \( C \) do
26: \( \text{CompositionScore}(C) \);
27: end for
28: return \( \text{top}(k, C) \);

Step 2 eliminating less relevant services (lines 14-23). For each class whose services do not touch the user preferences, we select randomly \( k \) services since they are all equal with respect to user preferences. Otherwise (i.e., its services touch the user preferences), we first compute the score of its services and then retain only the top-k ones.

Step 3 returning top-k compositions (lines 24-28). First, we compose services from only the retained ones, i.e., the top-k in each class. Then we compute the score of generated compositions and finally we provide the user with the top-k ones.

Figure 1 depicts some results obtained for queries with 2, 3 and 4 subqueries (query components). These results show that the number of subqueries adds only a slight increase in the execution time of the top-k algorithm. However, this execution time is increased significantly when the number of relevant services per class is large.

5. SYSTEM ARCHITECTURE

Figure 2 presents our implemented top-k DaaS service compositions system. The system consists of the three major modules: Annotation Module, Interactive Query Formulation Module and Top-k Service Compositions Module.

The Annotation Module allows service providers to annotate WSDL description files of services with fuzzy sets to represent linguistic terms and with SPARQL queries to represent the RPV and constraints of services. This annotation is implemented by adding a new element called ‘rdfQuery’ to the XML Schema of WSDL as in WSDL-S approach.

The RDF Query Rewriter implements an efficient RDF query rewriting algorithm (RDF Query Rewriter) to identify the relevant services that match (some parts of) a user query. For that purpose, it exploits the RPVs in the service description files. The Service Locator feeds the Query Rewriter with services that most likely match a given query. The Top-K Compositions component computes (i) the matching degrees of relevant services, (ii) the fuzzy dominating scores of relevant services, (iii) the top-k services of each relevant service class and (iv) the fuzzy compositions scores to return the top-k compositions. The top-k compositions are then translated by the composition plan generator into execution plans expressed in the XPDL language. They are executed by a workflow execution engine; we use the Sarvasti execution engine from Google.

6. RELATED WORK

Preferences in Web service selection/composition have re-
received much attention in the service computing community during the last years. Taking user preferences into account allows to rank candidate services/compositions and return only the best k ones to the user. Hereafter, we review some works for ranking and selecting Web services. Authors in [9] propose a method to rank semantic web services. It is based on computing the matching degree between a set of requested NFPs (Non-Functional Properties) and a set of NFPs offered by the discovered Web services. NFPs cover QoS aspects, but also other business-related properties such as pricing and insurance. Semantic annotations are used for describing NFPs and the ranking process is achieved using some automatic reasoning techniques that exploit the annotations. Wang et al. [12] propose a system for conducting qualitative Web service selection in the presence of incomplete or conflicting user preferences. The paradigm of CP-nets is used to model user preferences. The system utilizes the history information (stored in the system) of users to amend the preferences of active users, thus improving the results of service selection. In [13], the authors present a new ranking method, called ServiceRank. It considers the QoS aspects as well as the social perspectives of services (such as how they invoke each other via service composition). A service that provides good QoS and is invoked more frequently by others is more trusted by the community and will be assigned a higher rank.

The works the most related to our proposal are those proposed in [11, 14]. In [11], Skoutas et al. address ranking and clustering Web services and propose methods based on the notion of dominance; these methods apply multiple matching criteria without aggregating the match scores of individual services parameters. The notion of Web service dominance introduced relies on the one of uncertain dominance discussed in [10]. Two distinct scores are defined for assessing when a service is objectively interesting: (i) the dominated score of a service to indicate the average number of services by which it is dominated (ii) the dominating score of a service to indicate the average number of services that it dominates. Algorithms for computing the top-k Web services according to each score are provided. To achieve a balance between dominated and dominating scores, a third ranking score for a service that combines both, is introduced.

In [14], Yu and Bouguettaya address the problem of uncertainty inherent in the Quality of Web Service (QoWS) and compute the skylines from service providers (referred to as service skylines). A service skyline can be regarded as a set of service providers that are not dominated by others in terms of QoWS attributes that interest all users. To this end, a concept called p-dominant service skyline is defined. A provider S belongs to the p-dominant skyline if the probability that S is dominated by any other provider is less than p. The authors provide also a discussion about the interesting of p-dominant skyline w.r.t. the notion of p-skyline proposed in [10]. They show also that the former, on one hand, avoids requiring from users to transform personal preferences into numerical weights and, on the other hand, will not miss any ‘really good’ providers (i.e., the ones that provide consistently high QoWS).

7. CONCLUSIONS

In this paper, we addressed the top-k service compositions problem to answer fuzzy preference queries. The key notion of our approach is the fuzzy dominance relationship introduced to measure the extent to which a vector of matching degrees more or less dominates another one. This kind of fuzzy dominance is then used to both rank DaaaS services in their classes and DaaaS compositions in order to compute the top-k service compositions. Although we have focused on the DaaaS services to answer a query, the proposed approach could be applied in other cases to compute, for instance, the top-k QoS-based service compositions.

8. REFERENCES