# Multi Matchmaking Approach for Semantic Web Services Selection Based on Fuzzy Inference

Zahira Chouiref[1,2], Karim Benouaret[3], Allel Hadjali[2], and Abdelkader Belkhir[4]

[1] Université de Bouira, 10000 Bouira, Algeria
zahira.chouiref@univ-bouira.dz
[2] LIAS, ENSMA, 86360 Chasseneuil-du-Poitou, France
{zahira.chouiref,allel.hadjali}@ensma.fr
[3] LT2C, Université Jean Monnet, 42000 Saint-Etienne, France
karim.benouaret@univ-st-etienne.fr
[4] USTHB, 16000 Algiers, Algeria
kaderbelkhir@hotmail.com

**Abstract.** Selecting services from those available according to user preferences plays an important role due to the exploding number of services. Current solutions for services selection focus on selecting services based either only on non functional features, on context preferences or profile preferences. This paper discusses an improvement of existing services selection approaches by considering both user context and profile. The ultimate aim is to derive maximum profit from available profile and context information of the user by inferring the most relevant preferences w.r.t his/her contextual profile. Linguistic/fuzzy preference modeling and fuzzy inference based approach are used to achieve efficiently a selection process. Some experiments are conducted to validate our approach.

**Keywords:** Web Services Selection, Profile, Context, Preferences, Fuzzy Logic Theory, Fuzzy Inference Rules, Contextual Profile Matching.

## 1 Introduction

Semantic Web services; *SWS* field plays an increasingly important role in enhancing the user interaction in the Web and enterprise search, as well as in providing a flexible solution to the problem of application integration. With the rapid worldwide deployment of offered services on Internet, SWS selection has been an active and fast growing research area. The services selection is a technique which uses functional features *(Input, Output, Precondition, Effect; IOPE)* and non functional features *(Quality of Service; QoS, etc.)* [12], to search Web services from large scale service repositories that fit best the user requirements. Development of methods which would increase research accuracy and reduce research time is one of the main challenges in SWS selection. Most of these approaches focus on satisfying the functional requirements. A service consumer copes with a difficult situation in having to make a choice from a mass of already discovered services satisfying the functional requirements. To discriminate

S. Casteleyn, G. Rossi, and M. Winckler (Eds.): ICWE 2014, LNCS 8541, pp. 440–449, 2014.

such discovered services, the focal point of current SWS selection is on the non-functional aspect of a service [12]. This can be done using *QoS parameters* [1], *context* [8], *preferences* [11], *profile* [5]. These approaches help to improve the service discovery, selection and composition and simplify the management process for non functional attributes of Web services. However, such approaches do not address the issue of: i) Taking into account all the information characterizing the service (offered and requested) often called contextual profile; *CP*, ii) the gradual nature of the parameters related to context/preferences in a human language, iii) deriving new relevant preferences on the basis of user *CP* information by means of fuzzy inference rules. The key concept of the approach is the user/service profile where fuzzy logic theory is used to describe information related to the profile in a faithfully way. The first objective of this paper is to propose a common profile model that can capture all information describing the user and the service. The second objective is to introduce linguistic terms to express preferences and fuzzy rules to model contextual preferences.

The remainder of this paper is structured as follows: Section 2 presents a brief background on fuzzy set theory and provides a survey on existing approaches of SWS selection. In Section 3, we set up a required and provided service model based on *CP*, then we provide an SWS selection framework based on our model. Section 4 describes the query processing in a real case study related to the field of restaurant business and the ranking mechanism as well. In Section 5, an experimental study is described to show the feasibility and effectiveness of our proposal. Finally, Section 6 concludes the paper.

## 2 Related Work

A key issue in service computing is selecting service providers with the best user desired quality. Recently, existing service selection approaches reviewed below are distinguished by the fact that they rely on QoS parameters [1], context information [8], user preferences [2][7][9][11], etc.

A matchmaking algorithm proposed by Adnan et al. [1] is based on tying QoS metrics of Web service with fuzzy words that are used in users request. The aim of the paper is to satisfy user's requirements and preferences regarding only QoS and not all preferences related to non-functional service parameters. In [8], authors proposed the non-functional properties that are related to local constraints which reflect the user preferences and context of the demanded service. However, in real-life systems, context information is naturally dynamic, uncertain, and incomplete, which represents an important issue when comparing the service description with user requirements. This approach, however, is based on context rather than the data of services and could not handle both exact and fuzzy requirements. They do not allow reasoning on context information to determine criteria weights automatically, also the model proposed by the authors does not consider the implicit user preferences. Web services selection based on preferences mainly consider single user's preferences. Benouaret et al [2] introduce a novel concept called collective skyline to deal with the problem of multiple users preferences to select skyline Web services. Chao et al. [4] proposed a framework, which leverages fuzzy logic to abstract and classify the underlying data of Web services as fuzzy terms and rules. The aim is to increase the efficiency of Web services discovery and allow the use of imprecise or vague terms at the level of the search query. Steffen et al [7] model Web service configurations and associated prices and preferences more compactly using utility function policies, and propose flexible and extensible framework for optimal service selection that combines declarative logic-based matching rules with optimization methods, such as linear programming. In [10], a framework of SWS discovery based on fuzzy logic and multi-phase matching is proposed in this work. The first level matchmaking is executed with service capability against the second level matching is executed with service fuzzy information. The authors do not take into account neither the vague information of the user profile (personal information, etc.) nor the vague information of the context.

The presented work links user profile, user context and user preferences and provides suitable selection method that uses inferred preferences in order to have powerful, yet scalable ranking process. These preferences are inferred from the user's contextual profile, this is done by directly applying an efficient inference method based on *an extended modus ponens.*

# 3 Our Model

## 3.1 Service Description Model

We describe a SWS (provided / required) by the following model, which not only supports service capability information ($IO$), but also supports *service profile*, *service context* and presents *service vague information* associated to *preferences.* For the sake of illustration, the following reference example is used.

**Reference Example:** Let U be a user wants to book a hotel in Australia. He sets his preferences and submits the following query $Q$: "return the hotels in Australia preferably {***near*** **to his position**} with {***affordable*** **price**} and {***at least three*** **stars**} and having a **restaurant** with an {***asiatic*** **cuisine**}, knowing that he has a **car** and he is accompanied by his **wife** and his **child**".

**Definition 1 (Advertised Service Description Model/Required Service Description Model).** An advertised Web service (A required Web service) is described by the following model respectively: $SA = \{CA,\ CPA,\ FZA\}/$ $SR = \{CR, CPR, FZR, \theta\}$, where:

- $CA/CR$ is the advertised/required service capability information description, which contains *(NA, DA, FPA)/(NR, DR, FPR, θ)* where $NA/NR$ is the name of the advertised/required service, $DA/DR$ is the functional description of the advertised/required service and $FPA/FPR$ is all functional parameters of the advertised/required service *(IOPE).*

- $CPA = (CPA_1, \ldots CPA_n)$, $CPR = (CPR_1, \ldots CPR_n)$, is a set of non-functional parameters that make up the *CP* of *SA/SR*. The detail context can be explored in section 3.2.
- *FZA/FZR* is a set of linguistic terms which used to describe the vagueness that pervades information containing in *CA/CR* and *CPA/CPR*.
- $\theta(0<\theta<1)$ is a threshold such that if the satisfaction degree of a service w.r.t the query at hand is below this threshold, it is not retrieved.

*Each I/O attribute of request* $a_i$ *is characterized by a set of preferences values* $p_i$*, i.e.* $(a_i,\ p_i)$. For each SA attribute, we can assign crisp constraints (e.g. MinStars) or fuzzy constraints (FZA) (e.g. CheaperPrice). The attributes are self explanatory, which indicates the pereference values which help the user to choose the service that suits its preferences. For each SR attribute, we can assign crisp preferences, or fuzzy preferences (FZR) (e.g. AffordablePrice).

### 3.2 Fuzzy Model to Contextual Profile

**Fuzzy Contextual Profile Modeling.** The information of *CP* can be: static such as *Profile* {*Personal data, etc*}, evolutionary such as *Preferences* {*Colour, Language, etc*} and temporary such as *Context* {*Devices, Localization, etc*}. These pieces of information must be captured to match demands to offers of services, in order to improve the relevance of answers during a selection process. For a given query *X*, we define its *CP* environment $CPE_x$ as a finite set $\{(P_1, P_2, \ldots P_n)\}$ of multidimensional parameters, for instance, {*personal informations, context, preferences, etc*}. Each parameter *Pi* is modeled as a finite set $\{(C_1, \ldots C_m)\}$ of concepts, for instance {*demographic information, spatial context, display mode, etc*}. Each concept *Ci* is modeled as a finite set $\{(C'_1, \ldots C'_t)\}$ of sub concepts for instance {*gender value, family situation, country name , etc*}, and/or a finite set $\{(v_1, \ldots v_t)\}$ of attributes value, for instance {*full screen, postscript format, etc*}. Each concept (sub-concept) is characterized by a set of *preference values.* The attribute value domain $dom(C_i)$ can be expressed by means of: numerical assessments, logical assessments and fuzzy linguistic assessments.

An instantiation of the *CP*, called *CP state*, writes:
$w = (C_1\ is\ v_1 \wedge \ldots \wedge C_k\ is\ v_k),\ k \preceq m$ , where each $C_i \in CPE_x$, $1 \preceq i \preceq k$ and $v_i \subseteq dom\ (C_i)$ (the symbol $\wedge$ denotes a conjunction).

*Example 3.* For instance, *w* may be (family situation is *married*, means of transport is *car*, accompanying people is *wife* and *child*) for the example above.

**Definition 2 (Contextual Profile Preferences).** *A contextual profile preference CPP is a fuzzy rule of the form:* ***if*** $C_1$ ***is*** $v_1 \wedge \ldots \wedge C_m$ ***is*** $v_m$ ***then*** $A_1$ ***is*** $F_1 \wedge \ldots \wedge A_l$ ***is*** $F_l$*, where vi,* $1 \leq i \leq m \leq n$*, stands for a crisp or fuzzy value of the context or the profile parameter* $CP_i$ *and* $F_j$*,* $1 \leq j \leq l$ *represents a fuzzy preference related to attribute* $A_j$*.*

The meaning of *CPP* is that in the *CP* state specified by the left part of the rule, the preference $A_j$ *is* $F_j$ is inferred. From the user profile, one can deduce the following preferences on the searched hotels:

*Example 4.* A user who has a *car*, generally prefers hotels with *parking*. This may be expressed as ($CPP_1$): **if** *means of transport is car* **then** *PreferenceParking is yes.*

**User Preferences Modeling.** Let us now discuss the notion of fuzzy preferences: for instance, *"affordable price"* and *"nearest city"* are primitive terms. A primitive term can be described thanks to fuzzy sets, allowing to obtain for a price and a given distance, the satisfaction levels defined on the interval [0, 1]. As for categorial attributes, the membership functions are modeled as follows: The membership function of *"CuisineStyle"* is modeled by: $\mu_{C\_cuisine}$={1/chinese, 0.9/japanese, 0.8/thaiwanese, 0.7/sushi, 0.7/indonesian, 0.5/vietnamese, 0.3/indian, 0.2/pakistani, 0.1/americain} for chinese cuisine and $\mu_{F\_cuisine}$={1/french, 0.9/belgian, 0.8/mediterranean, 0.7/italian, 0.7/dutch, 0.6/latin, 0.5/german, 0.4/british, 0.2/american} for french cuisine. The membership function of parking is $\mu_{parking} = \{1/yes, 0/no\}$.

## 4 Query Processing

### 4.1 Semantic Web Service Selection Framework

Let $H_D$ be a Hotels database. The desired services should accept {Address} as inputs and return {HotelName, StarsNumber, Price, CuisineStyle} as output for the case of our *reference example*. This query $Q$ is written as follows:
**SELECT** *name-Hotel* **FROM** $H_D$ **WHERE** (Hotel.Price is *affordable* AND Hotel.Stars is *at least 3* AND Hotel.Dist is *near of city* AND Hotel.CuisineStyle is *asiatic*).

Fig. 1 gives an overview of the selection framework. The matching engine must match the list of services with the input, output specified by the user. The main steps of the framework are:

**First Search Filter.** The result returned by *step 1* and *2* of Fig. 1 is $\sum_Q$, the list of services that correspond to the desired *number of stars*, *price*, *asiatic cuisine* and *distance to user*. Note that the preferences expressed on these attributes are mandatory. The system will computes the distance or similarity between all the concepts vector of the query and those of the Web services by means of suitable measures of similarity. In presence of different types of attributes, for each attribute, an adequate similarity measure is used as the following functions:
$S_{p1}$, $S_{p2}$, $S_{p3}$ and $S_{p4}$ represents respectively functions that compute the degrees of satisfaction of the *distance*, *Price*, *Number of Stars* and *Cuisine Style* of the SWS at hand w.r.t the fuzzy set modeling the user preference on attribute *distance*, *Price*, *Stars* and *Cuisine*. For instance, $\mu_{near}(8) = 0.4$, $\mu_{affordable}(45) = 0.5$, $\mu_{stars}(4) = 1$ and $\mu_{F\_Cuisine}(italian) = 0.7$.

*It is worth noticing the all the functions* $(S_{p1}, S_{p2}, S_{p3}, S_{p4})$ *provide degrees that belong to the same scale [0,1]. This property of commensurability allow aggregating them, in a convenient way, to obtain an overall score of an SWS.*

**Overall Matching** $Score_1$**.** After calculating the different individual matching values of each attributes of a service, one way to obtain an overall matching score is to aggregate these individual matching values using a *T-norm* operator (such the *min operator*) as follows:
$Score_1 = \top(S_{p1}, S_{p2}, S_{p3}, S_{p4}) = \min(S_{p1}, S_{p2}, S_{p3}, S_{p4})$.
Now, if the $Score_1$ of a service is higher than $\theta$, then this service is added to $\sum_Q$ list of services that satisfying request's attributes.

**Second Search Filter.** In this phase, not only *I/O* parameters has to be considered but also contextual profile information such as accompanying people, age of child, etc. This means that service contextual profile should contribute to the development of an advanced search strategy. The steps (3, 3', 4, 5, 5', 6) of the Fig. 1 illustrate this strategy and are summarized in the following: (i) Infering a set of relevent preferences and their semantics from the fuzzy rules base $B^{CPP1}$, regarding the user *CP* state $w$, then augment the query by the inferred preferences. To achieve this, we make use of a *knowledge based model* described bellow. (ii) Calculation of the satisfaction of the result provided by the first filter w.r.t. to the inferred preferences.

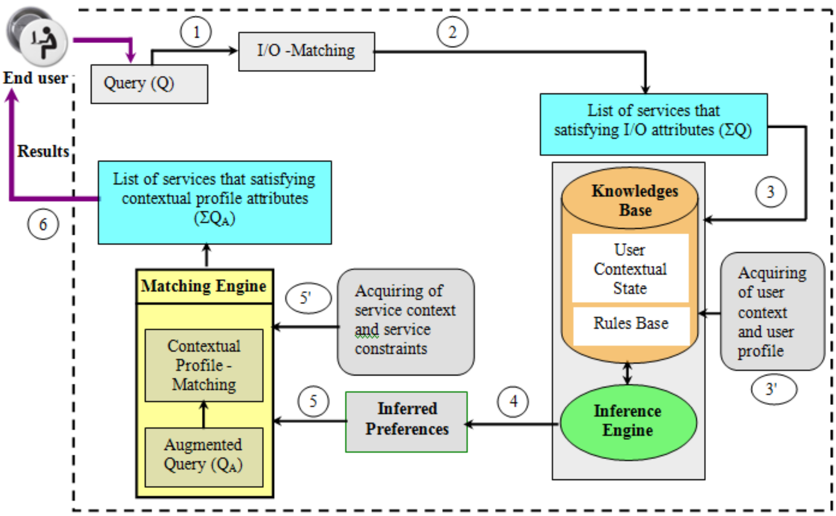



**Fig. 1.** Semantic Web Services Selection Framework

[1] Is a set of contextual profile preferences that can be built from users' experiences in the domain considered.

**Knowledge Based Model.** We rather propose here to specify contextual preferences by means of *fuzzy gradual rules*. The use of gradual rules for deriving preferences with respect to the user *CP* leads to refining the search results and returns the best ones to the user.

***Inference Process:*** We assume available $B^{CPP} = \{CPP_1, ..., CPP_m\}$ a fuzzy rules base modeling a set of contextual profile preferences. Each rule is of the form **if** $C_1$ **is** $v_1 \wedge ... \wedge C_m$ **is** $v_m$ **then** $A_1$ **is** $F_1 \wedge ... \wedge A_l$ **is** $F_l$ (see Definition 2.). Now, given a user's *CP* state, i.e., *CP*=($C_1$ **is** $v_1 \wedge ... \wedge C_k$ **is** $v_k$), one can derive relevant preferences to the user by using a fuzzy inference schema, called the *Generalized Modus Ponens; GMP* [6]. In a simple case, from the rule: **if** *C* is *V* **then** *A* is *F* and the fact: *C* is *V'*, where *V*, *F* and *V'* are gradual predicates modeled thanks to fuzzy sets, the *GMP* allows inferring the preference *A is F'* and the fuzzy semantics of *F'*. See [6] for more details.

*Example 6.* The preference *Parking is Yes* is inferred by applying the rule (**If** user has a *car* **then** preferenceParking is *Yes* ) and the fact (user has a *car*). Note that *car* and *yes* are fuzzy sets represented by (car/1)(yes/1) respectively.

*Example 7.* In our booking hotels example, assume available the rules base $B^{CPP}$ of Table 1. The preference *Hotel is Animated'* is inferred by applying the rule (**If** age is *young* **then** hotel is *animated* ) and the fact (age is *about_26*) where *young* and *about_26* are fuzzy sets represented by (0,0,25,27) and (24, 26, 26, 28) and *animated* is expressed using a qualitative rating such as: *low*, *medium*, *high* and *very high* whose semantics are given by triangular membership function. The semantics of *Animated'* is calculated from the semantics of *Animated, young* and *about_26*, see [3].

Our model entirely leverages the knowledges base presented in Table 1 to derive new relevant preferences. A rule-based representation which includes different possible preferences for our reference example is proposed.

***Augmented Query Process:*** Once the user preferences are inferred, we offer the system the possibility of augmenting the query in order to refine the selection process. Then, we have a final user query $Q^A$, the augmented query of $Q$ writes: $Q^A = \{C_1 \wedge C_2 ... \wedge C_n \wedge P_1 \wedge P_2 \wedge ... P_m\}$.

*Example.* $Q^A$ = {*Price* ∧ *Stars* ∧ *Distance* ∧ *CuisineStyle* ∧ ***Parking*** ∧ ***Hair and Beauty Service*** ∧ ***Sauna*** ∧ ***Games and Activities*** ∧ ***Kid Friendly Menu*** ∧ ***Childrens Highchairs*** ∧ ***Family*** ∧ ***Animated'***}, where the infered preferences appear in bold.

***Similarity Computing:*** in this phase, the system computes the similarity between inferred user preferences and services' constraints by means of fuzzy semantics associated w.r.t these preferences. To accomplish this phase, the system will evaluate the similarity degree between all the inferred preferences from $B^{CPP}$ and the services' constraints of $\sum_Q$ by using the following functions: Now,

**Table 1.** Knowledges Base

| Knowledges Base | | |
|---|---|---|
| **Rule** | **Rules Base** | **Facts Base** |
| R1 | **If** user has car **then** preferenceParking is Yes | Car |
| R2 | **If** accompanying people is wife **then** facilities is hair and beauty service | Married |
| R3 | **If** accompanying people is wife **then** facilities is sauna | Wife |
| R4 | **If** accompanying people is child **then** facilities is games and activities | Child |
| R5 | **If** accompanying people is child **then** facilities is kid friendly menu | |
| R6 | **If** accompanying people is child **then** facilities is childrens highchairs | |
| R7 | **If** accompanying people is wife **then** theme is family | |
| R8 | **If** age is young **then** hotel is animated | |

assume that the initial query $Q$ is only augmented by the inferred preferences: *Parking* and *Animated'*. To compute the satisfaction of each hotel h $\in \sum_Q$ w.r.t such preferences, we use:
$S_{p5}$: for each h $\in \sum_Q$, we have $\mu_{parking}(h) = 1$ if *h.parking = yes*, 0 otherwise, and $S_{p6}$: the fuzzy semantics of the predicate *Animated'* is computed by means of the combination/projection principle [3] (see example 7). Then, for each h $\in \sum_Q$, the degree of satisfaction is $\mu_{animated'}(h)$.

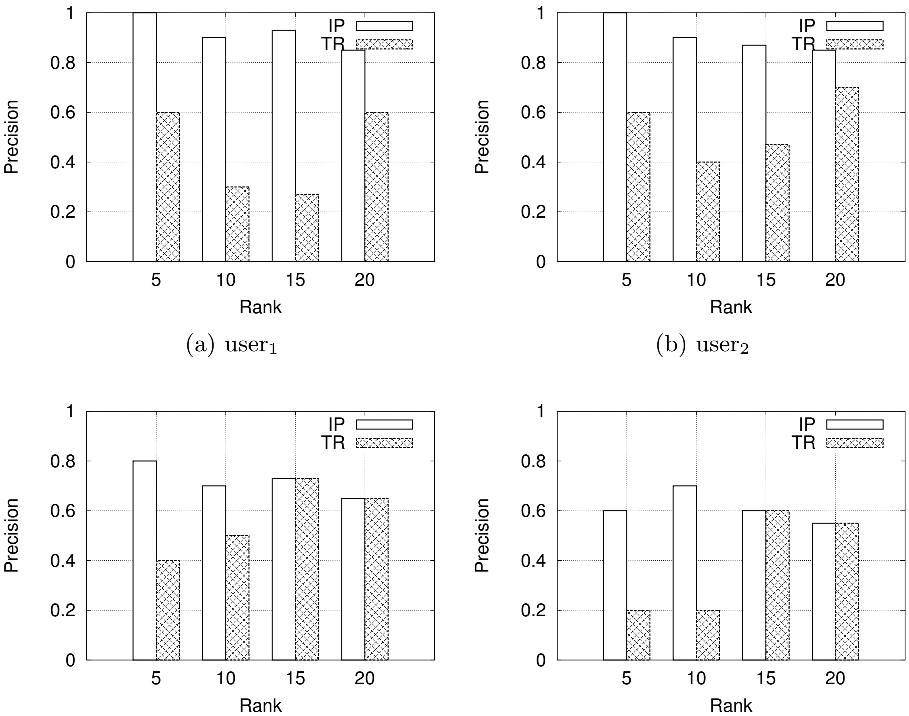
**Overall Matching** $Score_2$.
$Score_2 = \top(Score_1, S_{p5}, S_{p6}) = min(Score_1, S_{p5}, S_{p6})$.

**Overall Score S.** We use an aggregation function of $Score_1$ and $Score_2$ to compute the overall score $S$. Now, to give priority to the initial preferences w.r.t. to inferred preferences (IP) we make use of the following formula (where $\alpha \in ]0, 1]$ is the priority of IP). Then S is given by:
$S = min(Score_1, max(Score_2, 1 - min(Score_1, \alpha)))$
Finally, the user can select the top-k answers or the answers whose score $S$ is greater than a given threshold.

## 5 Experimental Evaluation

The main purpose of this evaluation is to compare the effectiveness of our proposed selection framework (referred to as IP for inference process) with the traditional frameworks that do not use the inference process (referred to as TR for traditional). We perform a case study, due to the limited availability of public services. We created a set of 100 synthetic restaurant service descriptions, and we involved different users to conduct our experiments. However, due to lake of space we only report results regarding 4 users.

Fig. 2 shows the precision of IP and TR at various ranks for 4 different users. Observe that IP has consistently better precision than TR since IP includes into the ranking process inferred preferences that are interesting for users. See also that, for $user_1$ and $user_2$ IP has an almost perfect precision, while the precision of TR is mediocre. Moreover, for $user_3$ and $user_4$ IP and TR have similar precision at rank 15 and rank 20. The reason is that the increase of the rank may increase the probability that similar services belong to the top-$k$ list of both approaches.




(a) $\text{user}_1$ (b) $\text{user}_2$

(c) $\text{user}_3$ (d) $\text{user}_4$

**Fig. 2.** Precision at Rank ($\theta = 0.5$, $\alpha = 0.5$)

## 6 Conclusion

This paper has proposed an SWS selection framework based on user (fuzzy) preferences. The goal of our work is to enhance accuracy of SWS search results by using multi matching level to compute similarity between advertised SWS and users request and taking into account the user preferences that may be inferred from users profile. We also showed the interest of using gradual rules for representing contextual preferences and deriving new preferences that are relevant to the user. The proposed framework makes the integration of user contextual profile and fuzzy inference rules techniques into the selection process. Some experiments are done to show the feasibility and the precision of our proposal.

## References

1. Al Rabea, A.I., Al Fraihat, M.M.: A new matchmaking algorithm based on multi-level matching mechanism combined with fuzzy set. Journal of Software Engineering and Applications 5(3) (2012)
2. Benouaret, K., Benslimane, D., HadjAli, A.: Selecting skyline web services for multiple users preferences. In: ICWS, pp. 635–636 (2012)

3. Bouchon-Meunier, B., Dubois, D., Godo, L., Prade, H.: Fuzzy sets and possibility theory in approximate and plausible reasoning. In: Fuzzy Sets in Approximate Reasoning and Information Systems (1999)
4. Chao, K.M., Younas, M., Lo, C.C., Tan, T.H.: Fuzzy matchmaking for web services. In: 19th International Conference on Advanced Information Networking and Applications, AINA 2005, vol. 2, pp. 721–726. IEEE (2005)
5. Chouiref, Z., Belkhir, A., Hadjali, A.: Advanced profile similarity to enhance semantic web services matching. International Journal of Recent Contributions from Engineering, Science & IT (iJES) 1(1), 1–13 (2013)
6. Hadjali, A., Mokhtari, A., Pivert, O.: A fuzzy-rule-based approach to contextual preference queries. In: Computational Intelligence for Knowledge-Based Systems Design, pp. 532–541. Springer (2010)
7. Lamparter, S., Ankolekar, A., Studer, R., Grimm, S.: Preference-based selection of highly configurable web services. In: Proceedings of the 16th International Conference on World Wide Web, WWW 2007, pp. 1013–1022. ACM, New York (2007), http://doi.acm.org/10.1145/1242572.1242709
8. Reiff-Marganiec, S., Yu, H.Q.: An integrated approach for service selection using non-functional properties and composition context. In: Handbook of Research on Service-Oriented Systems and Non-Functional Properties: Future Directions, pp. 165–191 (2011)
9. Skoutas, D., Alrifai, M., Nejdl, W.: Re-ranking web service search results under diverse user preferences. In: VLDB, Workshop on Personalized Access, Profile Management, and Context Awareness in Databases, pp. 898–909 (2010)
10. Su, Z., Chen, H., Zhu, L., Zeng, Y.: Framework of semantic web service discovery based on fuzzy logic and multi-phase matching. Journal of Information and Computational Science 9, 203–214 (2012)
11. Wang, H., Shao, S., Zhou, X., Wan, C., Bouguettaya, A.: Web service selection with incomplete or inconsistent user preferences. In: Baresi, L., Chi, C.-H., Suzuki, J. (eds.) ICSOC-ServiceWave 2009. LNCS, vol. 5900, pp. 83–98. Springer, Heidelberg (2009)
12. Yu, H.Q., Reiff-Marganiec, S.: Non-functional property based service selection: A survey and classification of approaches (2008)