

A SEMANTIC-BASED SERVICE DISCOVERY FRAMEWORK FOR COLLABORATIVE ENVIRONMENTS

Xu, S. Y.* & Raahemi, B.**

* Research Center for E-business and Intelligent Service, Wuhan University of Technology, Wuhan, 430070, China

** Telfer School of Management, University of Ottawa, 55 Laurier Ave., E. Ottawa, ON K1N 6N5, Canada

E-Mail: xsying@gmail.com

Abstract

In recent years, service-oriented and ubiquitous technologies have experienced impressive development. As these services grow rapidly both in scale and type, effective and accurate service discovery methods play an increasingly important role in the search and selection of services that match consumer requirements and preferences. In order to discover the optimum service and enhance the effectiveness of discovered results, a semantic-based service discovery framework, consisting of user model, context model, service model and a service discovery process, was presented in this study. Then the personalized service ontology was introduced to adjust the service search range adaptively on the basis of the service ontology structure and user information. Furthermore, a semantic-based service discovery method was designed in the proposed framework, which enabled names, attributes and relations of services to be more accurately matched and mapped with user preferences. Finally, to evaluate the effectiveness and accuracy of this method, the simulation analysis was conducted based on service ontology, in which information on 102 separate services and 10 scenarios were extracted from actual data. The simulation results show that compared with the keywords-based method, the proposed semantic-based method shows an increase in recall rate, precision and *F*-measure. The simulation results also reveal that the proposed method improves service discovery efficiency and performs well in accuracy. Therefore, collaborative environments considered in service discovery can provide useful and effective guidance to study the service recommendation.

(Received in April 2015, accepted in October 2015. This paper was with the authors 2 months for 1 revision.)

Key Words: Service Discovery, Recommendation, Service Ontology, Similarity, Semantic

1. INTRODUCTION

The rapid development in the Internet-based service industry has attracted researchers to focus on efficient solutions for service discovery. Service discovery has become an important topic in recent studies. A QoS-aware evaluation method was proposed in [1] to improve the efficiency of service discovery based on service related concepts. In order to meet the needs of customers, an efficient recommendation system was proposed in [2], which considered historical usage data instead of text-based analysis. User information offers new possibilities for service discovery. A user-centric service discovery method was proposed for personalized services based on first-hand experience from a European Union-sponsored research project OPUCE [3]. A personalized service ontology-based intelligent service dispatching mechanism was presented to manage customer expectations and behaviours, as well as to improve customer experiences and increase satisfaction in a high performance service ecosystem [4]. For pervasive environments, a proactive service discovery approach based on the essential aspects of service discovery, which are user context and user preference was also proposed in [5]. However, a key challenge is that existing studies on optimized service discovery concentrate mainly on Web Services, providing the best services for customers from the aspect of different service quality attributes, while ignoring differences in the preferences of customers and profits of service providers in the business service environment, and overlooking the phenomenon of changes in context or customer preference over time.

Relationships among user preference, user context, and services exert significant influences on service selection and decision. In this paper, we propose a novel ontology-based service discovery framework that includes models for user preference, user context, and services. We exploit the relations among these models to determine which services satisfy users' requirements. A key problem in service discovery is the calculation of similarities between user input and target services, and subsequently specifying best services based on user preferences, context, and related rules. After the services mashup was performed, composite services were ranked for users as their search results. As the core of our service discovery framework, a semantic-based service discovery method is proposed. This method is based on service ontology and is exploited to calculate the similarity of service name, attributes and relations between user request and service ontology.

The remainder of this paper is organized as follows. Section 2 provides the most important state of the art work in corresponding fields. Section 3 presents the service discovery framework, service ontology, and describes the details of the proposed semantic-based service discovery method. Section 4 analyses simulation results and makes a comparison with the proposed semantic-based service discovery method and keyword-based approaches. Conclusions are summarized in Section 5.

2. STATE OF THE ART

Various studies have attempted to improve the efficiency of methods aimed at discovering services, including online and offline services, social networks, and pervasive services.

Despite ongoing improvements in existing discovery techniques, web services discovery continues to be an issue for users due to the vast amount of web services available on the Internet. Palmieri defined a novel approach towards efficient and robust service discovery in modern ubiquitous and pervasive computing systems [6]. Faci et al. presented a framework for weaving the principles of social networks into web services discovery [7]. However, some issues remain unsolved, such as syntactic-based discovery returning results with poor precision, different structures of current web services resulting in difficulties in capturing how each web service has interacted with peers, and the lack of general structure of services to describe and extract significant topology of relationships between web services.

The development of the Internet of Things, Cloud Computing, and Location-Based Service has given rise to studies on the use of contextual information to realize dynamic service discovery, ubiquitous computing [6], pervasive services [5] and ad hoc networks. Rasch et al. considered how the changes in a user's current context, set of registered services, or set of user preferences could result in updating the current list of discovered services [5]. Lee et al. proposed a model for designing good u-Service ontology and evaluating its quality with respect to dynamic service composition during the design phase [8]. Nardi et al. proposed a commitment-based account of the notion of service captured in a core reference ontology called UFO-S [9]. Fardin et al. analysed and described semantic relationships between web service entities and built a service ontology model based on OWL-DL to successfully build large service directories that operate faster and with lower costs as compared to existing models [10]. However, while the replace ability of devices is considered, the replace ability of services is ignored. Moreover, a comparison analysis during the design phase does not appear to be feasible due to the fact that most existing pervasive service models are applicable only to static applications on a specific small-scale or in a single service domain. These models are not suitable for complex service domains that involve lots of inter-dependent dynamic objects and would possibly provide the user with undesired and unpredictable services. Listed below are the shortcomings exhibited by existing approaches which could impair the efficiency of service discovery:

(a) Most existing service discovery approaches are keyword or syntax-based. While some approaches are semantic-similarity-based, they yield a discovery result of similar services [11]. Users require an intelligent service mashup based on associations between services. Hence, researchers need to investigate new service discovery approaches from other viewpoints, such as semantic relevance [12], contextual awareness, social network [13], or linked prediction.

(b) Studies on service discovery focus mainly on the area of web services. With the development of networks, users require more advanced network services, such as electronic services, cloud services, and business services. Therefore, new service discovery approaches must include efficient and powerful mechanisms for handling complex service discovery.

(c) The low level of service relationship exploitation and utilization. The patterns of service composition and aggregation are mainly based on service name similarity, and their results are far from what users require. Service function similarity can be divided into services with similar functions and services with complementary functions. Service social networks can be divided into recommendation network and collaboration network. These categories are worth investigating further to improve service discovery efficiency.

(d) The current service discovery approaches are not fully intelligent, with only a “Pull” mode existing, and no “Push” mode. With the popularity of social networks and the rapid development of network applications, an increasing amount of user interaction has provided applications with large amounts of information that can be converted into intelligence. Thus, how to achieve personalized service discovery and recommendations is a popular topic in service discovery research.

Although existing service discovery models do not fully support intelligent service discovery, these models are well-established and widely used in various areas. As such, we consider them as references during the development phase. We also consider the following components as the main factors facilitating intelligent service discovery:

User information: this information includes both implicit and explicit information that can be generated by analysing and mining user history information, current user information, customer expectation, user satisfaction and preference, and so on [4].

Service information: this information can be obtained from the attributes of a service, such as functional attributes, non-functional attributes, and QoS (availability, reliability, response time, cost), or from the information provided by service providers.

Contextual information: this information includes facts concerning the context, such as location, time, environment, and device status.

3. METHODOLOGY

Although existing keyword-based service discovery approaches could provide users with related services, the probability of failing to find services that satisfy most or the main components of the customer’s requirements is high. In this research, we propose an intelligent service discovery method that provides users with service search results based on a comparison of users’ requirements with advertised services, as well as the characteristics and structure of service semantic features analogous to those of service ontology.

3.1 Framework

We introduce a new service discovery mechanism that can provide intelligent results for user service requests. This mechanism is distinguished by its service ontology, which is built as a services dataset. The method of organizing and classifying available services also results in an improvement in the quality of the service discovery results. Fig. 1 shows the proposed discovery framework. This framework includes three models and a process, as follows.

(a) “User Model”: It is a container for a set of basic user information, user histories information, and user social information. Basic user information is used to form an individual user profile. Users show certain service preferences or habits when consuming or interacting with services, and thus, user preference can be established through the user’s history.

(b) “Context Model”: It is a container for a set of user environment characteristics, including natural environment, social environment, and network environment. A user contextual information usually changes with time, that is, the context model must be updated to meet users’ changing demands for services.

(c) “Service Model”: It is a container for a set of formalized service information useful for saving time and effort when searching for services and their locations. Its management includes personalized services ontology (PSO) generation and PSO updating.

(d) “Service Discovery Process”: It starts with receiving the user’s services request and then performs the steps in Section 3.2 of this paper.

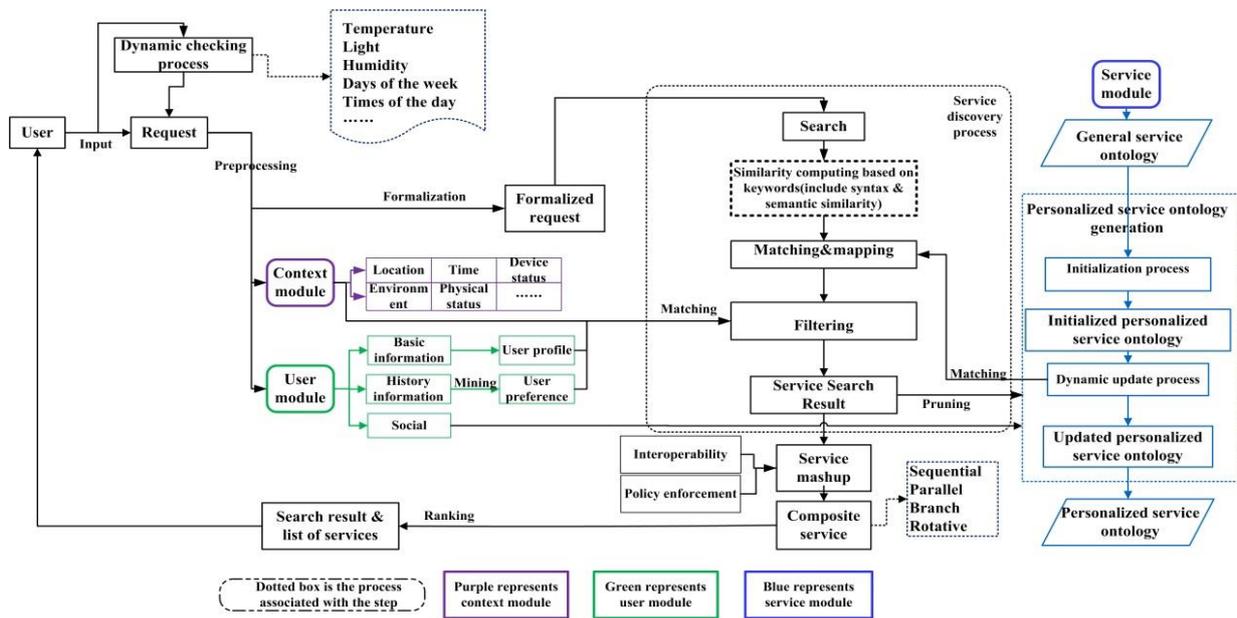


Figure 1: Proposed service discovery framework.

3.2 Service discovery process

As shown in Fig. 2, we propose a semantic-based service discovery method (SSDM) to accurately search and find a target service that satisfies a user’s requirements by generating a service set that reflects the user’s service queries. The service set is applied to service filtering to analyse the extent to which the service context, user preference, and service attributes of different services are matched and thereby improve the identification process of services to satisfy the user’s request. Detailed steps of service discovery process are as follows.

Step 1: User request preprocessing. This step includes the processes of user information extraction, obtaining contextual information, and processing service information. User information is extracted, and user context information obtained and analysed. The services request entered by the user is then formalized. The service request keywords are extended based on WordNet to expand the scope of the search.

Step 2: Service matching and mapping. This step aims to determine user target (and related) services that can satisfy the user’s requirements. First, the results of keyword extension are collated as the relevant service concepts set and then matched with service ontology to provide the first set of candidate services. Next, service attributes (functional and quality attributes, value and object attributes) matching is performed for each service to

achieve the second set of candidate services that could potentially satisfy the user’s service attributes demand. Finally, related services are matched based on service ontology, which contributes to achieving a relevant recommendation.

Step 3: Services filtering. Service filtering can be broken down into context information, user preference, and service attributes filtering. It aims to obtain services that fit service filtering regulations, which can be defined by domain experts or based on predefined rules.

More details on the proposed SSDM are described in Section 3.6 of this paper.

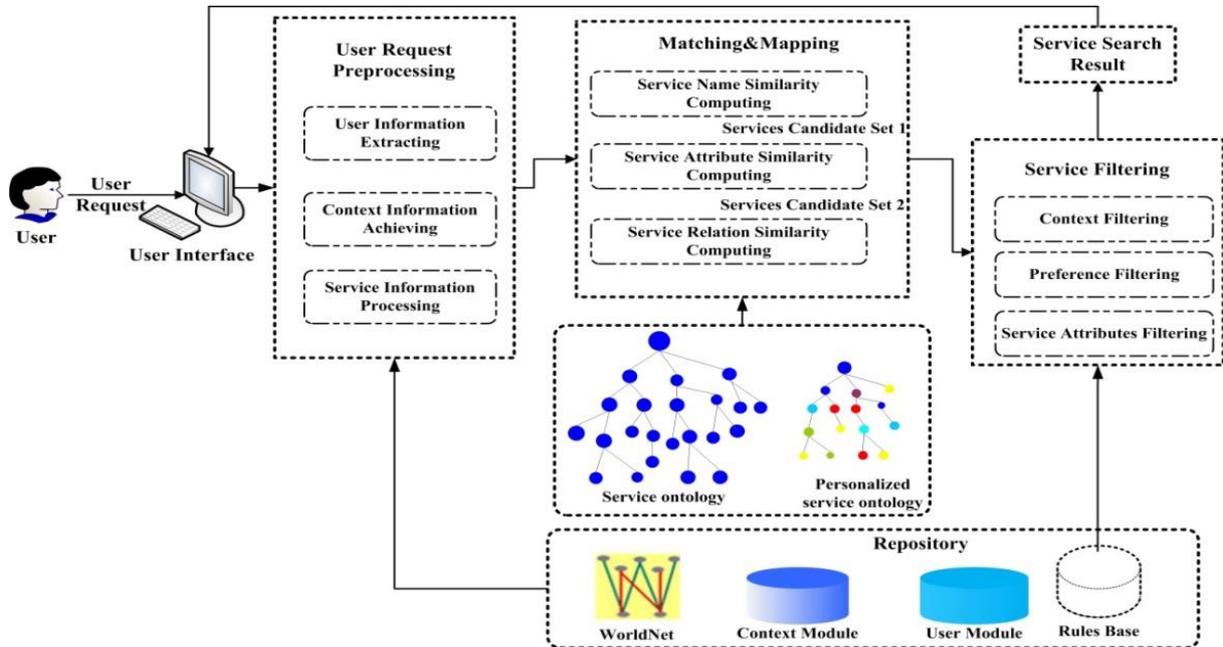


Figure 2: Architecture of Semantic-based Service Discovery Method.

3.3 Service ontology model

The goals of unified service ontology research are to classify and describe service concepts in a services recommendation system and to identify the service ontology attributes relevant to a user’s request. The main challenge that prevents service ontology from achieving its goals is the problem of building a unified and powerful ontology description framework. Hence, in this study, we propose personalized service ontology architecture to avoid this problem which could have a negative effect on the discovered results. Fig. 3 shows the proposed architecture which is composed of three components, namely, (a) the service ontology structure, (b) personalized service ontology initialization, and (c) personalized service ontology updating.

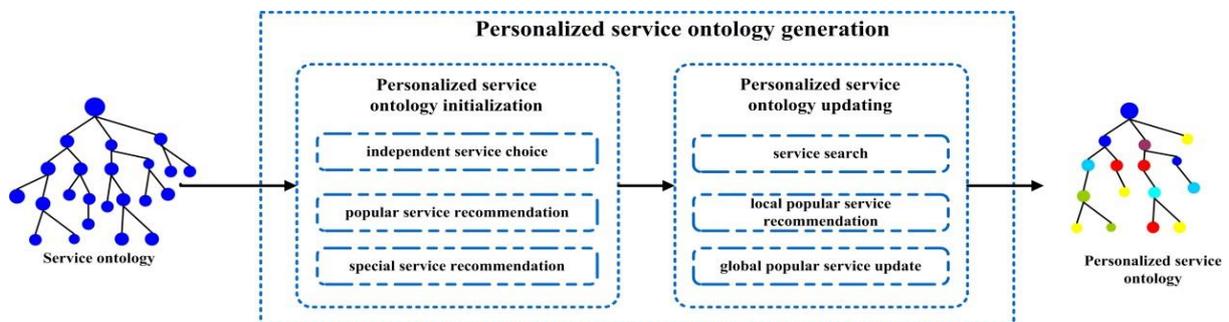


Figure 3: Personalized service ontology architecture.

Service ontology refers to the formal service information that uses ontology to construct components, such as class, attribute, relationship, and instance. Service ontology is defined as

$SO = (SN, SNR, SA, I, X)$, where SN is the services concepts set, SNR is service concepts relationship set, SA is service attributes set, I is service instances set, and X is axiom set.

Service ontology concepts (SN) and relationship (SNR): Service concepts are taken from the service website, while their semantic relationships are obtained from the website definition and standard classification system (eg. UNSPSC, ecl@ss). Semantic relationships include the terms: (a) “kind of”; (b) “equivalent to”; (c) “has part of”; (d) “similar to”, meaning similar service concepts; (e) “disjoint with”; and (f) “related to”.

Service attributes (SA): Service attributes can be classified into Generic and Business Property based on the purpose of the attributes. In the proposed service ontology structure, service name, service description, and service category are Generic Property, whereas service price, service level, and payment method are Business Property. Service attributes can also be classified into Data type Property and Object type Property based on their value types.

Service instance (I) and axiom (X): Instance refers to a specific individual of a service. In this paper, we define instances such as Individuals of Location {1525 Alta Vista DR, Ottawa, ON, Canada, K1G 0G1; 55 Laurier Ave., E. Ottawa, ON, Canada, K1N 6N5} and Individuals of Pay Method {Cash; MasterCard; Visa; PayPal}. X can be generated from the service provider or deduced from service history.

3.4 Personalized service ontology initialization

The personalized service ontology (PSO) initialization process eliminates services that users are not interested in and preserves what users are interested in, as well as other relevant recommendations. The process is shown in Fig. 3. A user decision on the preference of a particular service is partly decided by the user himself/herself and partly affected significantly by the user’s social network (such as Top N services selected by the best user’s friends group) and service providers (such as service promotion). In this paper, we use the effects of the user’s degree of interest, degree of service concern of other users, and degree of services provider’s recommendation to calculate the Service Popularity Degree (SPD).

User Interest Degree (UID): UID refers to the proposed quantitative measure of a user’s interest in a service. To calculate, we take a weighted average of these data for the UID , as shown in Eq. (1), $C_o T(s_j)$ is the total consumption times that user u_i consumed service s_j ;

$\sum_{\lambda=1}^n B_{\lambda} t(s_j)$ is the length of browsing time of s_j , and λ is the total length of browsing time of

s_j ; $CS(s_j) = \begin{cases} 1 & u_i \text{ collected } s_j \\ 0 & \text{other} \end{cases}$, is the collection status of s_j (because only two collection statuses exist,

namely, collected (1) or not (0)); $A_{UID} > 0$, $B_{UID} > 0$, $C_{UID} > 0$, are the corresponding standardization factors, and α_{UID} , β_{UID} , γ_{UID} (where $\alpha_{UID} + \beta_{UID} + \gamma_{UID} = 1$) are the corresponding coefficients. In this paper, the function of standardization factors balances the forces of each statistical datum. These standardization factors are determined by experts.

$$UID_i(s_j) = \frac{\alpha_{UID} \cdot C_o T(s_j) \cdot A_{UID} + \beta_{UID} \cdot \frac{\sum_{\lambda=1}^n B_{\lambda} t(s_j)}{\lambda} \cdot B_{UID} + \gamma_{UID} \cdot CS(s_j) \cdot C_{UID}}{A_{UID} + B_{UID} + C_{UID}} \quad (1)$$

Service Concern Degree (SCD): SCD denotes the proposed quantitative measure for a concern degree of a service from the viewpoint of services, as shown in Eq. (2), $ST(s_j)$ refers to the total number of times service s_j was consumed; $C_{\chi R}(s_j)$ is the consumer review of service s_j , χ is the evaluation times of s_j ; $NUI(s_j)$ is the number of users interested in service s_j ; $A_{SCD} > 0$, $B_{SCD} > 0$, $C_{SCD} > 0$ are corresponding standardization factors; and α_{SCD} , β_{SCD} , γ_{SCD} (where $\alpha_{SCD} + \beta_{SCD} + \gamma_{SCD} = 1$) are corresponding proportion coefficients.

$$SCD(s_j) = \frac{\alpha_{SCD} \cdot ST(s_j) \cdot A_{SCD} + \beta_{SCD} \cdot \frac{\sum_{k=1}^n C\chi R(s_j)}{\chi} \cdot B_{SCD} + \gamma_{SCD} \cdot NUI(s_j) \cdot C_{SCD}}{A_{SCD} + B_{SCD} + C_{SCD}} \quad (2)$$

Service Recommendation Degree (SRD): *SRD* refers to the proposed quantitative measure for a recommendation degree of a service from the viewpoint of the service provider. *SRD* is the weighted average of these statistical data, as shown in Eq. (3), *DPP*(s_j) is the direct service price promotion value of service; *SPP*(s_j) is the service point promotion value of s_j ; *FSP*(s_j) is the free service promotion value of service s_j , $A_{SRD} > 0$, $B_{SRD} > 0$, $C_{SRD} > 0$ are corresponding standardization factors; and α_{SRD} , β_{SRD} , γ_{SRD} (where $\alpha_{SRD} + \beta_{SRD} + \gamma_{SRD} = 1$) are their corresponding proportion coefficients.

$$SRD(s_j) = \frac{\alpha_{SRD} \cdot DPP(s_j) \cdot A_{SRD} + \beta_{SRD} \cdot SPP(s_j) \cdot B_{SRD} + \gamma_{SRD} \cdot FSP(s_j) \cdot C_{SRD}}{A_{SRD} + B_{SRD} + C_{SRD}} \quad (3)$$

Service Popularity Degree (SPD): *SPD* represents user service popularity degree, which is derived by weighted averaging of *UID*, *SCD*, and *SRD*. As shown in Eq. (4), $\alpha_{SPD} > 0$, $\beta_{SPD} > 0$, $\gamma_{SPD} > 0$ (where $\alpha_{SPD} + \beta_{SPD} + \gamma_{SPD} = 1$) are the corresponding proportion coefficients.

$$SPD_i(s_j) = UID_i(s_j) \cdot \alpha_{SPD} + SCD(s_i) \cdot \beta_{SPD} + SRD(s_j) \cdot \gamma_{SPD} \quad (4)$$

3.5 Personalized service ontology updating

Modifications of the nodes in our proposed personalized service ontology are based on three types of data, namely, (a) user service search history, which is employed to mine the degree of service preference of users; (b) local popular service ranking, which is generated from users' social networks; and (c) global popular service ranking, which is generated from the recommendation of the service platform provider. After collecting the required data, we can perform the required operations on the service node.

Add Operation is used to add new services to the personalized service ontology. As the search activities of users continue to grow, we can analyse the information from the users' service search history. In this operation, we compare users' interests with a threshold and take action on the services node accordingly. The specific steps are given as follows.

Step 1: Evaluate *SPD* of service concept *S*, and compare it with threshold θ . If $SPD < \theta$, terminate the process; if $SPD \geq \theta$, go to step 2.

Step 2: Determine whether service concept *S* exists in PSO or not. If yes, terminate the process; if not, go to step 3.

Step 3: Determine whether service concept *S* is a leaf node of PSO or not. If yes, add service concept *S* to its father node; if not, go to step 4.

Step 4: Determine relationship between service concept *S* and child nodes of its father in PSO. If they are brother, add service concept *S* to its father node; if it is a father-son relationship, add service concept *S* to its father node and demote S_1 to child node of *S*.

Delete Operation is used to delete a service node from the personalized service ontology. In this operation, we compare users' interests with a threshold. The operation may be time consuming but it provides a more concise ontology than the general service ontology described earlier. The specific steps are given as follows.

Step 1: Evaluate service popularity degree of service concept *S*, and compare it with threshold θ . If $SPD \geq \theta$, terminate the process; if $SPD < \theta$, go to step 2.

Step 2: Determine whether service concept *S* is a leaf node of PSO or not. If yes, delete service concept *S* from its father node; if not, delete service concept *S* from its father node, and upgrade its child node as direct child node S_1 to its father node.

3.6 Semantic-based service discovery method

The SSDM aims to identify services according to similarities and relations between semantic features. Ontological mapping can be used as a model for this process because a service has a structure based on ontology that represents the semantic characteristics and elements of a service in our model. However, the process of service matching and mapping faces two problems. One is that service attributes have an important effect on user decisions and the method used to calculate attribute similarity. Another is that the system for using service relations between service nodes to judge service similarity is open to question. In this study, a comprehensive similarity approach that uses service name, service attributes, and the service relationship generated from SO or PSO is proposed as the criteria for comparing service similarity. As shown in Fig. 4, this process includes the subsequent key steps:

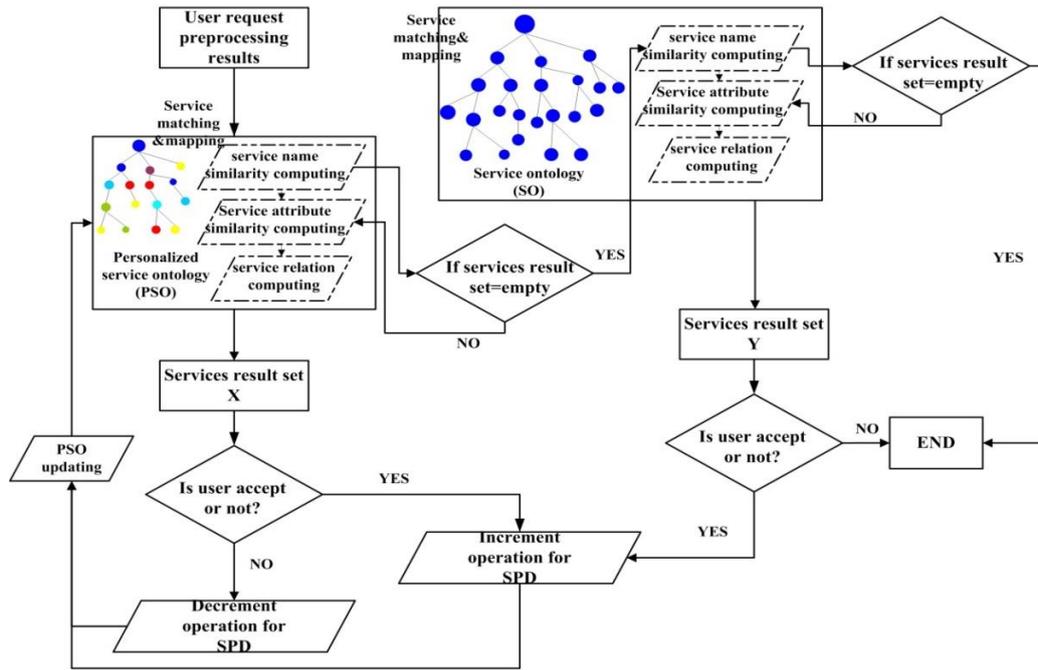


Figure 4: Semantic-based service discovery process.

Step 1: Service name similarity calculation: The Jaccard coefficient is used to evaluate service name similarity. As shown in Eq. (5), SN_O and SN_N represent the service name in ontology and based on users' request. SNS_O represents the service name similarity between SN_O and SN_N , c_o^i is the name of concept i in SN_O , and c_N^j is the name of concept j in SN_N .

$$SNS_O(SN_O, SN_N) = \frac{1}{n} \sum_i \left\{ \arg \max_j \frac{|c_o^i \cap c_N^j|}{|c_o^i \cup c_N^j|} \right\} \quad (5)$$

Step 2: Service attribute similarity calculation: Different attributes require different measurements because of the different kinds and types of service attributes. In this model, a service attribute matrix (SAM) is generated to solve the problems associated with service attribute similarity computing, as shown in Eq. (6).

$$SAM = \begin{bmatrix} sav_{11} & sav_{12} & \cdots & sav_{1j} \\ sav_{21} & sav_{22} & \cdots & sav_{2j} \\ \vdots & \vdots & \vdots & \vdots \\ sav_{i1} & sav_{i2} & \cdots & sav_{ij} \end{bmatrix} \quad (6)$$

In *SAM*, rows represent service attributes, columns represent attribute values, and if the value of a service attribute sa_1 is v_2 , $sav_{12} = 1$, and $sav_{11} = sav_{13} = \dots = sav_{1j} = 0$. The Pearson correlation coefficient is used to calculate service attribute similarity, as shown in Eq. (7), A stands for the common attributes set of service S_O in service ontology and new service S_N , $sav_{O,a}$ refers to the value of attribute a of service S_O , $\overline{sav_O}$ denotes the average value of service S_O attributes; $sav_{N,a}$ is the value of attribute a of service S_N , and $\overline{sav_N}$ refers to the average value of service S_N attributes.

$$SAS_O(SA_O, SA_N) = SAS_O(SAM_O, SAM_N) = \frac{\sum_{a \in A} (sav_{O,a} - \overline{sav_O})(sav_{N,a} - \overline{sav_N})}{\sqrt{\sum_{a \in A} (sav_{O,a} - \overline{sav_O})^2} \sqrt{\sum_{a \in A} (sav_{N,a} - \overline{sav_N})^2}} \quad (7)$$

Step 3: Service relation similarity calculation: In this paper, we use the distance between the nodes in the service ontology as the criteria to calculate the similarities in the services relationship. Based on the Wu-Palmer measure, we present a service relationship similarity calculation method, as shown in Eq. (8), SRS_O stands for the service relationship similarity between S_O and S_N ; $lca(S_O, S_N)$ represents the latest common ancestor between S_O and S_N ; $dep(lca(S_O, S_N))$ denotes the distance between $lca(S_O, S_N)$ and root nodes; $len(S_O, S_N)$ is the distance between S_O and S_N .

$$SRS_O(SR_O, SR_N) = \frac{2dep(lca(S_O, S_N))}{len(S_O, S_N) - \frac{len(S_O, lca(S_O, S_N)) \cdot len(S_N, lca(S_O, S_N))}{(len(S_O, lca(S_O, S_N)) + len(S_N, lca(S_O, S_N)))^2} + 2dep(lca(S_O, S_N))} \quad (8)$$

Step 4: Service similarity calculation: Service similarity is derived by averaging similarities in *SNS*, *SAS* and *SRS*, as shown in Eq. (9), where SS_O represents the similarity between service S_O in service ontology and new service S_N input by the user's service search.

$$SS_O = \frac{SNS_O(SN_O, SN_N) + SAS_O(SA_O, SA_N) + SRS_O(SR_O, SR_N)}{3} \quad (9)$$

Step 5: Service similarity ranking: The accuracy and efficiency of service discovery is improved by applying similarity ranking to the matching and mapping results, which enabled SSDM to extract high similarity services based on a threshold adjusted according to the user.

4. RESULT ANALYSIS AND DISCUSSION

In this section we used simulation experiments with actual data to evaluate the effectiveness and accuracy of the proposed SSDM. We improved the credibility of the experiments by collecting information on 102 separate services from services websites and treating them as our experimental data set.

We first used the OWL language to model the information and then saved the service information into the "service ontology" defined in the Protégé platform. We implemented our SSDM using JDK1.7 and Protégé 3.4, and ran the method on a personal computer with an Intel Core 2.40 GHz processor and 2 Gb of RAM under a Windows operating system. In Table I, **R** refers to Rank, **ES** refers to Expert Search, **KM** refers to Keywords-based Service Discovery Method, **SM** refers to Semantic-based Service Discovery Method, and the number after "Expert search" or "Keywords-based search" service discovery results (eg. [S1](1)) refers to its order.

Table I: Cases 1 to 10.

Case	Method	Service Discovery Results	SNS	SAS	SRS	SS	R
1	ES	[S1](1);[S2](2);[S4](3);[S6](4)	---	---	---	---	---
	SM	[S2]	0.706	1	1	0.902	1
		[S4]	0.706	1	1	0.902	1
		[S6]	0.706	1	1	0.902	1
		[S10]	0.282	1	0.532	0.605	4
		[S11]	0.282	1	0.532	0.605	4
KM	[S1](1);[S2](2);[S3](3);[S4](4);[S5](5);[S6](6); [S7](7);[S8](8);[S9](9)	---	---	---	---	---	
2	ES	[S6](1);[S5](2)	---	---	---	---	---
	SM	[S5]	0.806	1	1	0.924	1
	KM	[S1](1);[S2](2);[S3](3);[S4](4);[S5](5)	---	---	---	---	---
3	ES	[S2](1);[S3](2);[S8](3)	---	---	---	---	---
	SM	[S2]	0.656	1	1	0.832	1
		[S3]	0.656	1	1	0.832	1
		[S8]	0.656	1	1	0.832	1
KM	[S1](1);[S2](2);[S3](3);[S4](4);[S5](5);[S6](6); [S7](7)	---	---	---	---	---	
4	ES	[S1](1);[S2](2);[S3](3);[S4](4);[S5](5)	---	---	---	---	---
	SM	[S2]	0.743	1	1	0.914	1
		[S4]	0.743	1	1	0.914	1
		[S3]	0.365	1	0.738	0.796	3
		[S5]	0.365	1	0.738	0.796	3
KM	[S1](1)	---	---	---	---	---	
5	ES	[S1](1);[S2](2);[S4](3);[S9](4);[S6](5);[S7](6); [S11](7);[S12](8)	---	---	---	---	---
	SM	[S1]	0.815	1	1	0.951	1
		[S9]	0.815	1	1	0.951	1
		[S6]	0.815	1	1	0.951	1
		[S2]	0.815	1	1	0.951	1
		[S4]	0.706	1	1	0.903	5
		[S7]	0.645	1	0.734	0.902	6
		[S11]	0.282	1	0.532	0.605	7
[S12]	0.282	1	0.532	0.605	7		
KM	[S1](1);[S2](2);[S3](3);[S4](5);[S5](6);[S6](7); [S7](8);[S8](9);[S9](9);[S10](10)	---	---	---	---	---	
6	ES	[S1](1);[S2](2);[S3](3);[S4](4);[S5](5);[S6](6)	---	---	---	---	---
	SM	[S4]	0.713	1	1	0.897	1
		[S5]	0.713	1	1	0.897	1
		[S3]	0.713	1	1	0.897	1
		[S1]	0.282	1	0.532	0.605	4
		[S2]	0.282	1	0.532	0.605	4
		[S6]	0.282	1	0.532	0.605	4
		[S7]	0.187	1	0.541	0.583	5
KM	[S1](1);[S8](2);[S9](3);[S10](4);[S2](5);[S3](6)	---	---	---	---	---	
7	ES	[S10](1);[S11](2);[S12](3);[S1](4);[S6](5);[S9](6); [S3](7);[S4](8)	---	---	---	---	---

Case	Method	Service Discovery Results	SNS	SAS	SRS	SS	R
	SM	[S10]	0.721	1	1	0.916	1
		[S11]	0.721	1	1	0.916	1
		[S12]	0.721	1	1	0.916	1
		[S1]	0.282	1	0.532	0.605	4
		[S5]	0.282	1	0.532	0.605	4
		[S6]	0.282	1	0.532	0.605	4
		[S9]	0.282	1	0.532	0.605	4
	KM	[S1](1);[S2](2);[S3](3);[S4](4);[S5](5);[S6](6); [S7](7);[S8](8);[S9](9)	---	---	---	---	---
8	ES	[S1](1);[S2](2);[S3](3);[S5](4);[S8](5);[S9](6); [S4](7)	---	---	---	---	---
	SM	[S8]	0.721	1	1	0.916	1
		[S3]	0.721	1	1	0.916	1
		[S2]	0.706	1	1	0.902	3
		[S1]	0.282	1	0.532	0.605	4
		[S5]	0.282	1	0.532	0.605	4
	KM	[S1](1);[S2](2);[S3](3);[S4](4);[S5](5);[S6](6); [S7](7)	---	---	---	---	---
9	ES	[S1](1);[S2](2);[S4](3);[S5](4);[S6](5);[S8](6); [S10](7);[S11](8)	---	---	---	---	---
	SM	[S6]	0.706	1	1	0.902	1
		[S1]	0.706	1	1	0.902	1
		[S4]	0.713	1	1	0.897	2
		[S5]	0.713	1	1	0.897	2
		[S2]	0.713	1	1	0.897	2
		[S8]	0.282	1	0.532	0.605	6
		[S10]	0.187	1	0.541	0.583	7
[S11]	0.187	1	0.541	0.583	7		
	KM	[S1](1);[S2](2);[S3](3);[S4](4);[S5](5);[S6](6); [S7](7);[S8](8);[S9](9)	---	---	---	---	---
10	ES	[S10](1);[S11](2);[S12](3);[S2](4);[S3](5);[S7](6); [S8](7)	---	---	---	---	---
	SM	[S10]	0.706	1	1	0.9	1
		[S7]	0.706	1	1	0.902	1
		[S12]	0.706	1	1	0.902	1
		[S2]	0.706	1	1	0.902	1
		[S3]	0.713	1	1	0.897	5
		[S11]	0.713	1	1	0.897	5
		[S8]	0.282	1	0.532	0.605	7
	KM	[S1](1);[S2](2);[S3](3);[S4](4);[S5](5);[S6](6); [S7](7);[S8](8);[S9](9)	---	---	---	---	---

We also evaluated the effectiveness of our service discovery method using the following evaluation methods: (1) the ratio of success discovery and recall rate; (2) the ratio of accuracy discovery and precision; and (3) *F*-measure. Simultaneously, we conducted a comparison analysis between the SSDM and keyword-based method, based on the precision, recall rate, and *F*-measure and evaluated the effectiveness of services searched under the two methods. In Table II and Figs. 5 to 7, **KR** refers to keywords-based recall, **KP** refers to keywords-based

precision, **KF** refers to keywords-based F -measures, **SR** refers to semantic-based recall, **SP** refers to semantic-based precision, and **SF** refers to semantic-based F -measures.

Table II: Comparison of search results.

	KR	KP	KF	SR	SP	SF
Case1	0.667	0.444	0.5	0.833	1	0.909
Case2	0.667	0.286	0.4	1	1	1
Case3	0.2	1	0.333	0.8	1	0.889
Case4	0.75	0.6	0.667	0.875	1	0.933
Case5	0.625	0.833	0.714	0.75	0.857	0.8
Case6	0.625	0.556	0.588	0.75	0.857	0.8
Case7	0.625	0.714	0.667	0.714	1	0.833
Case8	0.75	0.667	0.706	1	1	1
Case9	0.5	0.2	0.286	0.5	0.5	0.667
Case10	0.571	0.444	0.5	0.583	1	0.737
AVG	0.598	0.574	0.536	0.781	0.971	0.857

This experiment included seven service domains for which services with similar contents were selected using both a semantic-based and a keyword-based method from 102 services. The results were compared to validate the accuracy and performance of the SSDM. For example, in Test Case 1, the user request for “Ottawa hotel” was based on the filtering rule: [rule1:(?a pa:has Payment of ? American Express)]. Table I, Case 1 shows three kinds of search results, where five services are discovered through the semantic-based method and ranked according to their service similarity. Cases 2 to 10, user requests for “Ottawa maple”, “Puzzle game”, “Ticket booking”, “Styling hair”, “Chinese food”, “Pet hospital”, “Courier services”, “Move services” and “Physiotherapy”, based on different filtering rules, each method yields a list of possible related services, as shown in Table I.

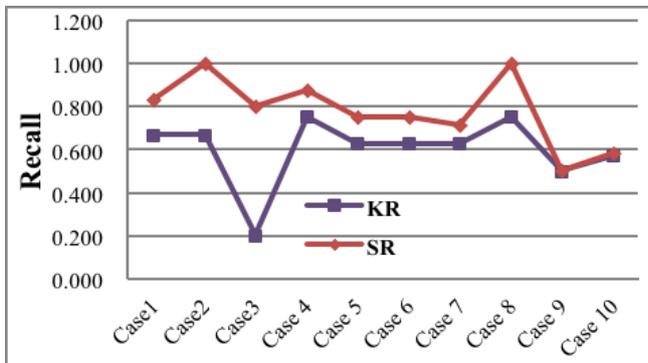


Figure 5: Recall rate.

In the experiment and the mechanism presented in this paper, 10 scenarios were selected as search subjects, based on which similar services were identified from 102 services. Table I shows that the search results based on the “Ottawa hotel” request include nine services identified through the keywords-based method, five by the SSDM, and six by experts. Six of the services are common across all three methods, while one found by the experts is not found by this method. The accuracy, recall rates, and F -measure of the SSDM are 1, 0.833, and 0.909, respectively. For the keywords-based method, the accuracy, recall rates, and F -measure are 0.444, 0.667, and 0.5, respectively. For a comparison of the precision, recall, and F -measure of the two methods, the same method was repeated for 10 different scenarios. The results of this comparison are shown in Table II and Figs. 5 to 7.

Results from the 10 experiments illustrate that the average recall rate of the keywords-based method is 0.598 and that the average recall rate of the SSDM is 0.781. The recall efficiency with SSDM is 31 %. These results indicate that SSDM can discover more related services than the keywords-based method. The average precision rate of the keywords-based method is 0.574, whereas that of the SSDM is 0.971; thus, precision efficiency was improved by over 69 %, proving that SSDM can generate higher customer satisfaction based on a strong relationship between service request and discovery.

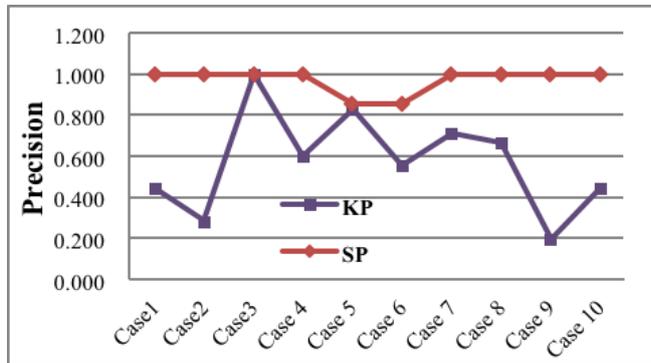


Figure 6: Precision rate.

The SSDM has better recall rate (Fig. 5), precision rate (Fig. 6), and *F*-measure (Fig. 7) than keywords-based method, thereby indicating that SSDM has higher accuracy and effectiveness based on the features of the service (service name, attribute, and relation). Hence, considering its high precision rate, this method can effectively expand users' queries and meet their needs.

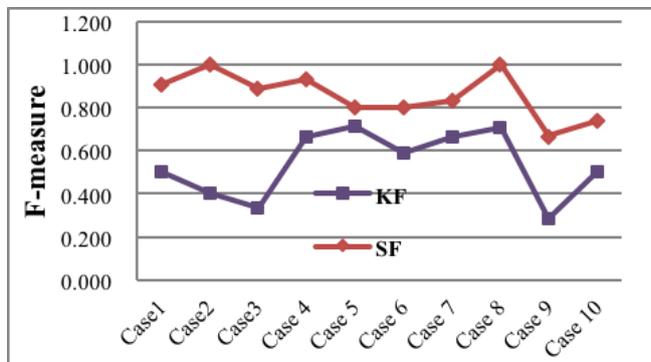


Figure 7: *F*-measure.

5. CONCLUSIONS

In this paper, the service discovery framework under collaborative environments was analysed and studied. User preference, service environment and service feature were considered to design a service ontology whose architecture was further updated with user request as the research foundation. In the research process, SSDM was presented and achieved good results. The main conclusions are shown as follows: (1) SO and PSO support semantic information matching during the service discovery process; (2) SSDM solves the problem of service similarity computing; (3) Compared with the keywords-based method, SSDM improves the accuracy and effectiveness of the service discovery results. In other words, results of SSDM can satisfy user request preferably. In future works, we would like to explore service ontology extension and establish a more functional service discovery system to make greater progress.

ACKNOWLEDGEMENT

This work was supported by the Fundamental Research Funds for the Central Universities (WUT:2015VI022), the National Natural Science Foundation of China (71172043), the National Key Technology R&D Program (2012BAH93F04), and the Humanity and Social Science Youth Foundation of Ministry of Education of China (14YJCZH165).

REFERENCES

- [1] Iordache, R.; Moldoveanu, F. (2014). QoS-aware web service semantic selection based on preferences, *Procedia Engineering*, Vol. 69 (24th DAAAM International Symposium on Intelligent Manufacturing and Automation), 1152-1161, doi:[10.1016/j.proeng.2014.03.104](https://doi.org/10.1016/j.proeng.2014.03.104)
- [2] Chan, N. N.; Gaaloul, W.; Tata, S. (2012). A recommender system based on historical usage data for web service discovery, *Service Oriented Computing and Applications*, Vol. 6, No. 1, 51-63, doi:[10.1007/s11761-011-0099-2](https://doi.org/10.1007/s11761-011-0099-2)
- [3] Yu, J.; Sheng, Q. Z.; Han, J.; Wu, Y.-B.; Liu, C.-F. (2012). A semantically enhanced service repository for user-centric service discovery and management, *Data & Knowledge Engineering*, Vol. 72, 202-218, doi:[10.1016/j.datak.2011.10.005](https://doi.org/10.1016/j.datak.2011.10.005)
- [4] Hsieh, Y.-H.; Yuan, S.-T.; Kuo, R.-L. (2011). A PSO-based intelligent service dispatching mechanism for customer expectation management, *Expert Systems with Applications*, Vol. 38, No. 10, 12128-12141, doi:[10.1016/j.eswa.2011.03.007](https://doi.org/10.1016/j.eswa.2011.03.007)
- [5] Rasch, K.; Li, F.; Sehic, S.; Ayani, R.; Dustdar, S. (2011). Context-driven personalized service discovery in pervasive environments, *World Wide Web*, Vol. 14, No. 4, 295-319, doi:[10.1007/s11280-011-0112-x](https://doi.org/10.1007/s11280-011-0112-x)
- [6] Palmieri, F. (2013). Scalable service discovery in ubiquitous and pervasive computing architectures: A percolation-driven approach, *Future Generation Computer Systems*, Vol. 29, No. 3, 693-703, doi:[10.1016/j.future.2012.08.004](https://doi.org/10.1016/j.future.2012.08.004)
- [7] Faci, N.; Maamar, Z.; Abdeldjelil, H.; Benslimane, D. (2011). Towards a framework for weaving social networks principles into web services discovery, *Proceedings of the 11th Annual International Conference on New Technologies of Distributed Systems (NOTERE)*, 8 pages, doi:[10.1109/NOTERE.2011.5957974](https://doi.org/10.1109/NOTERE.2011.5957974)
- [8] Lee, M.; Lee, J.-W.; Moon, N.; Park, S. S. (2011). Functionality evaluation of ubiquitous service ontology for dynamic service composition, Park, J. J.; Chao, H.-C.; Obaidat, M. S.; Kim, J. (Eds.), *Computer Science and Convergence (Lecture Notes in Electrical Engineering)*, Vol. 114, 475-485, doi:[10.1007/978-94-007-2792-2_45](https://doi.org/10.1007/978-94-007-2792-2_45)
- [9] Nardi, J. C.; de Almeida Falbo, R.; Almeida, J. P. A.; Guizzardi, G.; Pires, L. F.; van Sinderen, M. J.; Guarino, N.; Fonseca, C. M. (2015). A commitment-based reference ontology for services, *Information Systems*, Vol. 54, 263-288, doi:[10.1016/j.is.2015.01.012](https://doi.org/10.1016/j.is.2015.01.012)
- [10] Fardin, A. M.; Naser, N. B.; Ali, N. M. (2012). Empower service directories with knowledge, *Knowledge-Based Systems*, Vol. 30, 172-184, doi:[10.1016/j.knosys.2012.01.010](https://doi.org/10.1016/j.knosys.2012.01.010)
- [11] Liu, W.; Zhao, L.-F.; Deng, J.-Q. (2013). Research on an adaptive algorithm of service selection in pervasive computing, *Journal of Digital Information Management*, Vol. 11, No. 6, 482-488
- [12] Morcillo, J.; Romero, B. P. (2014). Business development management services for energy efficient processes, *DYNA*, Vol. 89, No. 2, 154-158, doi:[10.6036/5861](https://doi.org/10.6036/5861)
- [13] Du, J.-S.; Lu, J.-T.; Guo, Y. (2015). Relationship between technological diversification of social network and technological innovation performance: Empirical evidence from China, *Science Technology and Society*, Vol. 20, No. 1, 60-88, doi:[10.1177/0971721814561388](https://doi.org/10.1177/0971721814561388)