

2014

# BioTechnology

*An Indian Journal*

FULL PAPER

BTAIJ, 10(23), 2014 [14300-14306]

## Research on knowledge service matching method based on semantic similarity

XingangWang\*, Chengcheng Ge

Qilu University of Technology, Changqing District, Jinan, (CHINA)

E-mail: gechengcheng01@126.com

### ABSTRACT

For the issue that knowledge services support for business process dynamic building, this paper presents a knowledge service model which regards business activities as clues. The model uses Web Service as its media, it describes enterprise knowledge resources uniformly, packages knowledge resources with a certain function to services. It uses a knowledge service matching algorithm based on semantic similarity, considers service quality similarity and service function similarity, sort ad services in accordance with the similarity in descending order, the highest similarity service is the best service. This method improves the recall and precision, better support for business process building.

### KEYWORDS

Knowledge services; Semantic similarity; Service matching; QoS; Service matching.



## INTRODUCTION

With the deepening of knowledge-based, the quantity of knowledge-intensive enterprises is more and more<sup>[1]</sup>, the scale is also growing, knowledge is considered to be the core resource of the modern enterprise, all enterprise economic activities are inseparable from the support of knowledge. Therefore, how to efficiently use knowledge, turn knowledge into real power of enterprise development, has become the key to improve the competitiveness of enterprises.

Different enterprise sectors have different business systems, enterprise knowledge resources scattered in various heterogeneous systems, so it is difficult for concentration share. Web service encapsulates data, messages, behaviors and business logic flow uniformly, shields equipment and environmental the systems uses, thus provides a loosely coupled integrated environment system unrelated to platform and language<sup>[2]</sup>. Therefore, enterprises need to integrate the various sub-business systems, provide a Web-based knowledge service, and transfer knowledge resources to the enterprise in a knowledge service mode which unrelated to platforms and languages. Knowledge services can support for business process management, the combination of knowledge and business processes can effectively solve the problem that knowledge management implementation divorces from enterprise reality. So, enterprises need to provide dynamic Web-based, personalized, proactive knowledge services to enhance knowledge sharing and reuse, support for business processes building.

The rise of knowledge service<sup>[3]</sup>, provides a new idea for enterprise business process building, the paper presents a knowledge service model which regards business activities as clues. The model uses Web Service as its media, it describes enterprise knowledge resources uniformly, packages knowledge resources with a certain function to services. It uses a knowledge service matching algorithm based on semantic similarity, considers service quality similarity and service function similarity, sort ad services in accordance with the similarity in descending order, the highest similarity service is the best service. This method improves the recall and precision, better support for business process building.

## KNOWLEDGE SERVICE MODEL

Knowledge service model regards business activities as clues, organize enterprise knowledge resources (tools, methods, empirical parameters, location, etc.) associated with the business activities, and package into a knowledge-services unit which has a certain function. The business activity each service unit completes is a node in the whole business process, the output of the previous services unit can be input of a latter services unit. To meet complex business needs of users, you can combine service units in certain rules, and deposit combined services into knowledge services library.

Composite service<sup>[4]</sup> recombines services to generate new service according to user needs without changing the service components, in order to bring greater flexibility to the system and implement user customizable personalized service. Composite service potentially shortens the development time and reduces the workload of the development of new applications.

### Knowledge Service Modeling

In this article, knowledge service doesn't mean all kinds of knowledge or document in the usual sense, but enterprise application which can complete a specific function through knowledge-intensive business processes<sup>[5]</sup>. In other words, when a company or an organization organizes its various fragmented knowledge units through a business process, you become a knowledge services, it generates a knowledge service (as shown in Figure 1).

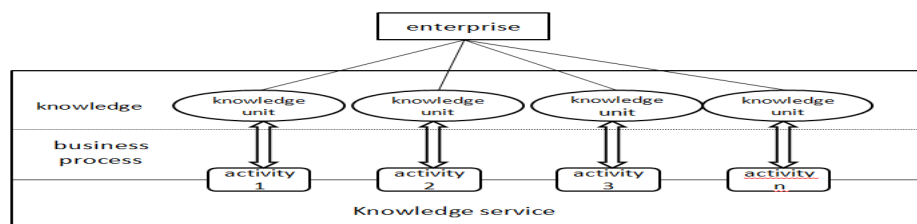


Figure 1 : Formation of knowledge service

### Analysis and present the enterprise knowledge

At first, we need to analysis and present the enterprise knowledge, form knowledge unit. We not only consider the general and domain-independent knowledge features, but also consider the proprietary knowledge characteristics required of the different areas of the proprietary. In order to provide a comprehensive analysis and presentation of enterprise knowledge, we classify knowledge in two dimensions: layer and category, thereby we create a two-dimensional knowledge framework to analysis and present a large number of enterprise knowledge<sup>[6]</sup>.

From the perspective of layer of knowledge, knowledge can be divided into 3 layers: conceptual layer, rule layer and method layer. The conceptual layer is the domain knowledge, it describes the main static information and knowledge objects in special areas; the rule layer is reasoning knowledge, in other words, the basic reasoning steps and rules which use domain knowledge; method layer is task knowledge, it describes the goal an application wants to achieve, how to decomposes task into subtasks and reason to achieve these goals. From the perspective of category of knowledge, we can decompose the

knowledge in each core business function module in accordance with knowledge area, each obtain knowledge area also contains a number of knowledge directory. For example, research function module, covers a few knowledge areas, process design, product design, etc. knowledge in design process also includes design criteria, process models, etc.(as shown in Figure 2)

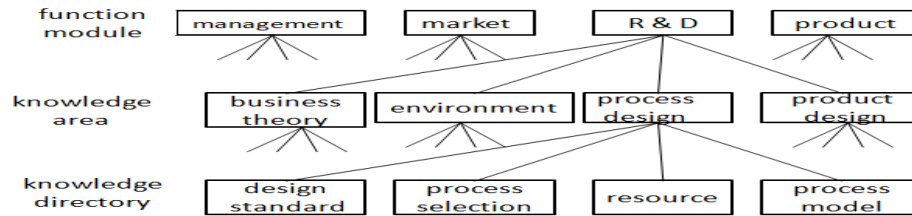


Figure (2). knowledge category

After classifying enterprise knowledge in the two dimensions, we get a two-dimensional knowledge framework as shown in TABLE 1. According to this intellectual framework, we can analysis and present the enterprise knowledge, fill specific enterprise knowledge into each grid of the knowledge framework.

TABLE 1: Knowledge net

layer	category			
	knowledge node K	knowledge node J	knowledge node I	...
concept layer C	(K, C)	(J, C)	(I, C)	...
rule layer R	(K, R)	(J, R)	(I, R)	...
method layer M	(K, M)	(J, M)	(I, M)	...

**Combine these knowledge and specific business activities**

We combine knowledge units and specific business activities. Business process is logically connected by a series of business, and it can complete the specific business objectives. It can be composed of one or more business activities, divides business activity into logically related business activities; then identify the knowledge unit each business activity requires, establish the mapping relation between business activities and knowledge units, then activities have their own knowledge map, the integration of knowledge unit and business process has been completed.

**Package knowledge related to business activities into a knowledge service with unified header and interface**

Based on Web Service technology, use WSDL to describe the knowledge uniformly, use unified header and interface to package the business process which integrate with knowledge units into a knowledge service, then publish knowledge service to UDDI center and store into knowledge service library, it's easy for users to find knowledge service. Knowledge service interface is used to transfer the input and output parameters between user and knowledge service. The header of knowledge service describes its basic information, includes the information knowledge service provider (name, contact information, nature of business) and abstract, keywords, functional areas, access, price, time validity of the knowledge service function, used history records (includes resolved problem, users reviews, etc.).

**Operational process of knowledge service**

The operational process of knowledge service model base on Web Services technology has five main steps: service description, service publication, service discovery, service binding and service call, the specific process is as follows:

Firstly, use WSDL (Web Service Definition Language) to package the knowledge resources with a certain function to knowledge services, and define the properties of these services, includes service descriptions and call interface, generate WSDL document that can be available for network access. Then register and publish services through UDDI center, when enterprises need to use the knowledge service, they can go to find it in UDDI center, UDDI center match the requestor's remand and existing service description then return results,. According to corresponding service interface and service provider location in registration information, enterprises can bind their own applications with the knowledge service, and complete the corresponding business activities with the help of SOAP, HTTP protocol and the corresponding knowledge service unit.

**KNOWLEDGE SERVICE MATCHING BASE ON SEMANTIC SIMILARITY**

There are a wide range of knowledge services in knowledge service library, how to find the service which accurately meet the requirements of user from so many knowledge services, in other words, service matching. It's a key problem for

enterprise knowledge sharing and use knowledge resources to support the process building. Research on traditional service matching based on keywords and syntax has been relatively mature, and the corresponding matching mechanism is relatively simple, but it has limited description capacity, its flexibility and scalability is poor, it's not well deal with service matching. People start to pay more attention to the semantic description with stronger service description<sup>[7]</sup>. In this paper, we adopt the knowledge service matching algorithm based on semantic similarity, the method improves the efficiency of service matching and achieve flexible service discovery mechanism.

In this paper, we comprehensive measure service quality similarity and service function similarity between ad services and request services, sort ad services in accordance with the similarity in descending order, the highest similarity service is the best service. The main properties of service function include input parameters, output parameters and knowledge operation.

Definition 1. The knowledge service is a 5 tuples,  $KS=(General, Input, Output, Operation, QoS)$ .

General means general information, includes the service name, service description, service provider information and so on. Input represents service input, output represents service output, Operation represents knowledge operation, QoS represents quality of service.

**Operation**  $\in \{knowledge\ browse, knowledge\ calculati\ on\}$  Knowledge browse provides browsing service to users, feedback knowledge description to the user, knowledge calculation means to solve service input problem with knowledge, get service output. Knowledge service is different from conventional web services concept. KS not only uses ontology to describe service capability, but also take knowledge operation into its service description. It not only clarifies the knowledge properties of the service, but also increases the available level of knowledge services<sup>[8]</sup>. To express semantic similarity of service capability, this paper introduces the conception of concept semantic similarity.

Definition2. Assume  $C_1$  and  $C_2$  are any two concepts in the knowledge ontology, if there is an inheritance path between  $C_1$  and  $C_2$ , the path length is  $p, S_c(C_1, C_2)$ , the concept semantic similarity between  $C_1$  and  $C_2$ , is defined as:

$$S_c(C_1, C_2) = \frac{1}{1+p}, p \geq 0;$$

Otherwise,  $S_c(C_1, C_2) = 0$

Obviously that  $0 \leq S_c(C_1, C_2) \leq 1$ .

**RKS** and **KS** represent ad services and request services,  $Sim(RKS, KS)$  represents their service similarity,  $Sim_{in}$ ,  $Sim_{out}$ ,  $Sim_{\phi}$  and  $Sim_{QoS}$  represent service input similarity, service output similarity, knowledge operation similarity and service quality similarity.

Definitions3. Service similarity between RKS and KS is:

$$Sim(RKS, KS) = \begin{cases} 0, & \text{if } Sim_{out} \times Sim_{in} = 0 \\ \omega_1 Sim_{out} + \omega_2 Sim_{in} + \omega_3 Sim_{\phi} + \omega_4 Sim_{QoS} & \text{else} \end{cases}$$

In the formula,  $\omega_1, \omega_2, \omega_3, \omega_4$  are all weight

coefficients, weight coefficients are set by user according to the actual situation. As can be seen from the definition 3, knowledge services input similarity and output similarity is a prerequisite to determine whether the service capabilities of RKS is similar to the service capabilities of KS, in case that knowledge services input and output are all similar, RKS and KS service capability matching is possible.

Definition4. Assume  $RKS_{in}$  has  $s$  parameters,  $KS_{in}$  has  $t$  parameters, if  $s \geq t$ , and for any parameter in  $KS_{in}$ , there exists at least a parameter in  $RKS_{in}$  that the concept semantic similarity the two parameters is greater than 0, maybe there are several such parameters, use  $SC_{inj}$  to represent the maximum value of the concept semantic similarity, then

knowledge service input similarity is:  $Sim_{in}(RKS, KS) = \frac{1}{t} \sum_{j=1}^t SC_{inj}$  Otherwise  $Sim_{in}(RKS, KS) = 0$ .

Definition5. Assume  $RKS_{out}$  has  $m$  parameters,  $KS_{out}$  has  $n$  parameters, if  $m \leq n$ , and for any parameter in  $RKS_{out}$ , there exists at least a parameter in  $KS_{out}$  that the concept semantic similarity the two parameters is greater than 0, maybe there are several such parameters, use  $SC_{outi}$  to represent the maximum value of the concept semantic similarity, then

knowledge service input similarity is:  $Sim_{out}(RKS_{out}, KS_{out}) = \frac{1}{m} \sum_{i=1}^m SC_{outi}$

Otherwise  $Sim_{out}(RKS, KS) = 0$ .

According to definition 3 and 4, only service output of KS covers the service output of RKS, service input of RKS covers the service input of KS, service capabilities of KS can satisfy the requirement of RKS.

Definition6. Knowledge Operation Similarity

$$\text{Sim}_p(RKS, KS) = \begin{cases} 1 & \text{If } RKS.\text{Operation}=KS.\text{Operation} \\ \frac{1}{2} & \text{If } RKS.\text{Operation}=KC \text{ and } KS.\text{Operation}=KB \\ 0 & \text{Otherwise} \end{cases}$$

In the above formula, KC represents knowledge calculation, KB represents knowledge browse.

### Service quality similarity

QoS (Quality of Service) is important criteria of Web Service performance evaluation, it describes service non-functional attributes explicitly and expresses the ability to meet users' requirements of a service<sup>[9]</sup>. There are a lot of factors affect QoS, such as response time, service price, performance, reputation, availability, reliability, security, throughput, concurrent processing capability and so on. In this paper we describe QoS in three aspects, they are response time, service

price and reputation. And we use  $\text{Sim}_{R\_time}$ ,  $\text{Sim}_{price}$  and  $\text{Sim}_{reputation}$  to represent response time similarity, service price similarity and reputation similarity, then service quality similarity between RKS and KS is:

$$\text{Sim}_{QoS}(RKS, KS) = \omega_1 \text{Sim}_{R\_time}(RKS, KS) + \omega_2 \text{Sim}_{price}(RKS, KS) + \omega_3 \text{Sim}_{reputation}(RKS, KS)$$

$\omega_1$ ,  $\omega_2$  and  $\omega_3$  are the corresponding weights of each QoS attributes, and  $\sum_{i=1}^3 \omega_i = 1$ .

### Response time similarity

Response time refers to the time interval between the service requester sends service request and obtain service response, including the delay time and processing time. Assuming the response time history is  $t_1, t_2, \dots, t_n$ , then average

response time  $t$  can be obtained by the following equation:  $t = \frac{1}{n} \sum_{i=1}^n t_i$ , assuming the response time threshold of service request is  $t_r$ , then response time similarity as follows:

$$\text{Sim}_{R\_time}(RKS, KS) = \begin{cases} 1 & t \leq t_r \\ \frac{t_r}{t} & t > t_r \end{cases}$$

### Service price similarity

Service price refers to the costs of knowledge services, assuming the price set in the request service is  $P_{RKS}$ , and service price is  $P_{KS}$ , then service price similarity as follows:

$$\text{Sim}_{price}(RKS, KS) = \begin{cases} 1 & P_{KS} \leq P_{RKS} \\ \frac{P_{RKS}}{P_{KS}} & P_{KS} > P_{RKS} \end{cases}$$

### Reputation Similarity

Reputation is evaluation criteria of knowledge service reputation, it reflects the users' evaluation to knowledge

services<sup>[10]</sup>. Reputation is a real number between [0, 1], it is calculated as follows:

$$S\_reputation = \begin{cases} 0 & T = 0 \\ \sum_{i=1}^n m_i / n & T = 1 \end{cases}$$

In the above formula,  $T$  represents the authenticity of service description, its value is 0 or 1,  $m_i$  represents a customer evaluation of the service in the history,  $n$  represents the customer sample number we select from the history. Service requester expects the service reputation to be higher. Reputation similarity is in inverse proportion to request service reputation, and is proportional to ad service reputation. Reputation similarity is:

$$Sim_{reputation}(RKS, KS) = \begin{cases} 0 & R_{RKS} > R_{max} \\ \frac{R_{avg}}{R_{RKS}} & R_{avg} < R_{RKS} \leq R_{max} \\ 1 & R_{RKS} \leq R_{avg} \end{cases}$$

In the above formula,  $R_{max}$  represents the maximum value of ad services reputation,  $R_{avg}$  represents the average value of ad services reputation,  $R_{WS}$  represents request services reputation.

**Service matching algorithm**

First, the service requester does keyword matching in UDDI, calculate the similarity between user keywords and service keywords, reserve services which its similarity is greater than a certain threshold and filter out irrelevant services, obtain an effective ad service set. Second, compute services input similarity and services output similarity between request services and ad services, reserve services those knowledge service inputs and outputs are all similar. Then calculate services quality similarity and knowledge operation similarity between request services and ad services. Calculate the overall similarity of ad services and sort ad services in accordance with the similarity in descending order. At last, return services list composed of services those their similarity reaches a specified threshold, the highest similarity service is the best service.

**EXPERIMENTS**

Take "demand forecasting" as an example to illustrate the service matching process, after receiving user requests, query in UDDI center, we get ad service set like this. We use to represent request service, input parameters, output parameters and knowledge operation of request service and ad services set as shown in TABLE 2:

**TABLE 2 : Service function parameters**

knowledge	Input	Output	Operation
RKS	price, market share, pre-sales, economic	demand, optimal inventory	knowledge
KS1	prices, competitors average price, pre-sales,	demand, demand fluctuation range	knowledge
KS2	price level, pre-sales, seasonal index	order expected value, max order	knowledge
KS3	previous production, demand, current	inventory level, safety stock	knowledge
KS4	price, market share, pre-sales, seasonal	order quantity, optimal inventory	knowledge

According to definition4, definition5 and definition6, we can get services input similarity, services output similarity and knowledge operation similarity between request services and ad services, the results are represented in TABLE.3 below:

**TABLE 3: Service function similarity**

knowledge service	Input	Output	Operation
KS1	0.95	0.85	1
KS2	0.85	0.6	0.5
KS3	0.7	0.7	0.5
KS4	0.9	0.9	1

According to the response time similarity formula, price formula and reputation similarity formula we introduce earlier in this article, we get QoS similarity parameters of ad services, as shown in TABLE 4:

**TABLE 4: Service quality similarity**

Service	Response time	Price	Reputation
KS1	0.8	0.6	0.7
KS2	0.7	0.7	0.6
KS3	0.6	0.6	0.5
KS4	0.8	0.8	0.7

The weights of response time, price and reputation are 0.5, 0.3 and 0.2. We can obtain service quality similarity for each service:

$$\text{Sim}_{QoS}(\mathbf{RKS}, \mathbf{KS1}) = 0.5 \times 0.8 + 0.3 \times 0.6 + 0.2 \times 0.7 = 0.72$$

$$\text{Sim}_{QoS}(\mathbf{RKS}, \mathbf{KS2}) = 0.5 \times 0.7 + 0.3 \times 0.7 + 0.2 \times 0.6 = 0.68$$

$$\text{Sim}_{QoS}(\mathbf{RKS}, \mathbf{KS3}) = 0.5 \times 0.6 + 0.3 \times 0.6 + 0.2 \times 0.5 = 0.58$$

$$\text{Sim}_{QoS}(\mathbf{RKS}, \mathbf{KS4}) = 0.5 \times 0.8 + 0.3 \times 0.8 + 0.2 \times 0.7 = 0.78$$

## RESULT AND DISCUSS

Assume that the weights of input similarity, output similarity, knowledge operation similarity and service quality similarity are 0.3, 0.3, 0.2 and 0.2, according to definition 3, the overall similarity of each service were:

$$\text{Sim}(\mathbf{RKS}, \mathbf{KS1}) = 0.3 \times 0.85 + 0.3 \times 0.95 + 0.2 \times 1 + 0.2 \times 0.72 = 0.884$$

$$\text{Sim}(\mathbf{RKS}, \mathbf{KS2}) = 0.3 \times 0.85 + 0.3 \times 0.6 + 0.2 \times 0.5 + 0.2 \times 0.68 = 0.671$$

$$\text{Sim}(\mathbf{RKS}, \mathbf{KS3}) = 0.3 \times 0.7 + 0.3 \times 0.7 + 0.2 \times 0.5 + 0.2 \times 0.58 = 0.68$$

$$\text{Sim}(\mathbf{RKS}, \mathbf{KS4}) = 0.3 \times 0.9 + 0.3 \times 0.9 + 0.2 \times 1 + 0.2 \times 0.78 = 0.896$$

Consider of service quality similarity and service function similarity, we obtain overall similarity of all service. Sort ad services in accordance with the similarity in descending order, the order is KS4、KS1、KS3、KS2. So for this example, the best matching service is KS4.

## CONCLUSIONS

In the context of the knowledge economy era, in order to implement enterprise knowledge resource sharing and improve reuse rate, this paper presents a knowledge service model which regards business activities as clues, the model packages knowledge resources with a certain function to services base on Web Service, provided services to users in the form of knowledge services. This paper uses a knowledge service matching algorithm based on semantic similarity, considers service quality similarity and service function similarity, select the highest similarity service as the best service. This method improves the recall, precision and the ability to meet customer requirements. Further research work will be aimed at specific service composition algorithm, in order to satisfy personalized service and complex user needs, support for business process building.

## ACKNOWLEDGEMENT

I'd like to show my deepest gratitude to my friend Dr. Geng Yushui, a respectable and resourceful scholar, who has provided me with variable guidance in every stage of the writing of this thesis. Without his enlightening instruction and patience, I could not complete my thesis, His keen and vigorous academic observation enlightens me not only in this thesis but also in my future study.

Last but not least, I'd like to thank all of my friends, especially my three lovely roommates, for their encouragement and support.

## REFERENCES

- [1] Xu Shibo, Wu Xinglong. The knowledge services characterization. Journal of the China Society for Scientific and Technical Information, 2005, 25(10): 206- 207.
- [2] Wang Yang, Liu Xiaobing. Huang Xuewen, A semantic BOM knowledge integration framework oriented cloud manufacturing. Application Research of Computers, 2013, 30(7): 2068-2071.
- [3] Chen Yingqun. The main characteristics and development trend of knowledge service. Library Journal of Henan, 2002, 12(6): 41-44.
- [4] Deng Shuiguang, Wu Zhaohui. A survey of Web Service composition method. Sciencepaper Online, 2008, 3(2): 79-84.
- [5] Fan Yushun, Huang Chen. Knowledge service grid and its application in the manu-facturing network. Computer Integrated Manufacturing Systems, 2005, 11(04): 467-474.
- [6] Liu Jianwei, Yang Lufeng. Comparative study of knowledge representation. Computer Systems and Applications, 2011, 20(3): 242.
- [7] Wu Jian, Wu Zhaohui, Li Ying. Web Service discovery based on ontology and similarity of words. Chinese Journal of Computers, 2005, 28(4): 595-602.
- [8] Lin Lanfen, Gao Peng. A Knowledge Service-Based Model of Collaborative Manufacturing Process Planning for Networked Manufacturing. Journal of Computer-aided Design and Computer Graphics, 2005, 17(9): 2085-2092.
- [9] Hao Mei, Wang Daoping, Feng Xiaodong. Research on knowledge services matching agile supply chain oriented based on ontology and QoS. Application Research of Computers, 2010, 27(7): 2556-2559.
- [10] Sun Qibo, Yang Fangchun. Reputation Evaluation Approach in Web Service Selection. Journal of Software, 2012, 23(6): 1350-135