Assisting Trustworthiness Based Web Services Selection Using the Fidelity of Websites*

Lijie Wang, Fei Liu, Ge Li**, Liang Gu, Liangjie Zhang, and Bing Xie

Software Institute, School of Electronic Engineering and Computer Science,
Peking University, Beijing 100871, P.R. China
Key Laboratory of High Confidence Software Technologies, Ministry of Education,
Beijing 100871, P.R. China
{wanglj07,liufei08,lige,guliang05}@sei.pku.edu.cn,
{zhanglj06,xiebing}@sei.pku.edu.cn

Abstract. Web services selection aims to choose an appropriate web service among a number of service candidates. The trustworthiness of web services is an important metric for web services selection. Many trustworthiness based web services selecting methods have been proposed in the academic community. However, the fidelity of web service's supporting websites (e.g. the websites providing the service, or referencing the service.), as an important factor for the evaluation of web services' trustworthiness, is often ignored. This leads to that existing methods cannot provide service consumers with a comprehensive view of the web services. In this paper, we propose a method to estimate the fidelity of web services' supporting websites, and present a novel trustworthiness based web services selection approach using our estimation result. A case study conducted in this paper indicates that, by using our approach, we can provide active assistance for service consumers during web services selection.

Keywords: Web Services, Services Selection, Fidelity, Quality of Service.

1 Introduction and Related Work

With the increasing number of available web services on the internet, web services selection, which aims to select an appropriate web service among a bunch of functionally similar service candidates, is becoming a crucial task. Many researches on web services selection were proposed in the academic community in recent several years [3-6]. Recently, trustworthiness is introduced into services selection by many researches [5, 6]. In these researches, the trustworthiness of an individual web service is often acquired by collecting feedback from previous consumers, mainly in the form of rating on specific criteria of the service. However, all these existing researches mainly focus on using the quality or the trust of the individual service to conduct

* This work was supported by the National Basic Research Program of China (973) (SN: 2005CB321805), the Science Fund for Creative Research Groups of China (SN: 60821003), and the National Natural Science Foundation of China (SN: 60803010, 60803011).

** Corresponding author.

L. Baresi, C.-H. Chi, and J. Suzuki (Eds.): ICSOC-ServiceWave 2009, LNCS 5900, pp. 429–436, 2009.

selection of services. A comprehensive evaluation about the service is lacking in these researches.

For a specific web service on the internet, there are always some websites being related with the service. For instance, for a web service, there must be a website providing the executing environment for it, there must be some website(s) to publish the service, and there may be some website(s) providing reference links to the service. All these websites provide support for web services. In this paper, websites providing support for web services are called "***supporting websites***". Based on the different type of support they provide for services, we classify supporting websites into three types:

- ***Service Provider*** *(abbreviated as SP)*: The website providing the executing environment for the service. A service's *SP* is specified by an end point described in the service's WSDL document.
- ***WSDL Hoster*** *(abbreviated as WH)*: The website where the service is published on. Because a service is described by a WSDL document, publishing a service is actually hosting the service's WSDL document onto a website. Thus, we can get a service's *WH* from the URL of the service's WSDL document.
- ***Service Reference*** *(abbreviated as SR)*: The website containing reference links (i.e. hyperlinks) to the service. Because a web service is described by the service's WSDL document, a service's *SR* actually refers to the website containing hyperlinks to the service's WSDL document.

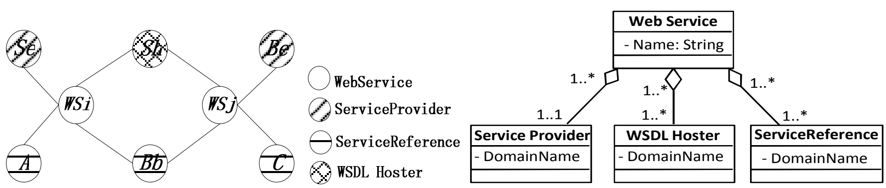


Fig. 1. Web services and supporting websites

The ***fidelity*** of a supporting website is the probability that it may provide valid web service information, it can be calculated using the proportion of valid web service information it provides.

Since there are three types of supporting websites for a web service, a website may play different roles for the same web service (or for different services). Actually, a website has different fidelity when it plays different roles. For example, all of the services provided by website *SiteX* are available, but many service hyperlinks on it are invalid. It means that website *SiteX* has high fidelity as a *SP*, but its fidelity is relatively low as a *SR*. Therefore, we treat a supporting website playing multi-roles as different ones, e.g. we treat '*SiteX*' as two websites, one plays the role as a *SP*; the other one plays the role as a *SR*. As shown in Fig 1, *Be* stands for the website *B* playing the role as a *SP*, *Bb* stands for the website *B* playing the role as a *SR*.

In real life, the fidelity of supporting websites provides valuable reference for a service consumer to judge the services at hand during web services selection. For instance, it is natural that a service consumer tends to choose a web service whose supporting websites have higher fidelity. The fidelity of supporting websites is an important factor for evaluation of the trustworthiness of services. It provides a comprehensive view of the services. However, existing evaluation mechanisms often ignore this factor and do not provide service consumers with a comprehensive view of web services. In this paper, we propose a method to assess the fidelity of supporting websites, and present a novel trustworthiness based web services selection approach using the fidelity of the service's supporting websites.

The rest of this paper is organized as follows. In section 2, we propose an approach to assessing the fidelity of supporting websites. The fidelity is used to assess the trustworthiness of web services in section 3. In section 4, we conduct a case study to evaluate our work. Section 5 draws the conclusion and some future work.

2 Assessing the Fidelity of Supporting Websites

2.1 Modeling the Relationship between Supporting Websites

When a service consumer uses a web service (represented as *s*), the typical sequence of supporting websites that he/she visits is from a *SR* of *s* to a *WH* of *s*, and then from the *WH* to the *SP* of *s* (the solid arrow shown in Fig 2). In other words, there exists linkage relation between the three types of supporting websites of a service. Z. Gyöngyi *et al.* argued that the trust of a website contributes to that of its linked websites [2]. The fidelity here also reflects a facet of trust for websites. We argue that there exists a fidelity transition relationship between the supporting websites of a service, along the sequence mentioned above, shown as the dotted arrow in Fig 2. The fidelity of a *SR* website contributes to that of its linked *WH* websites; the fidelity of a *WH* website contributes to that of its linked *SP* websites.

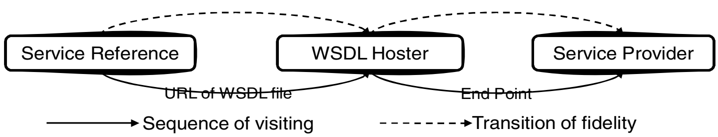


Fig. 2. Relationship between the supporting websites of a service

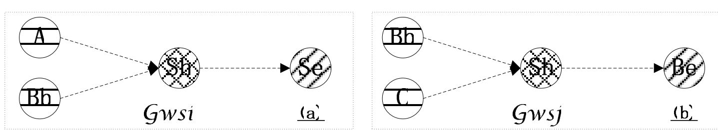


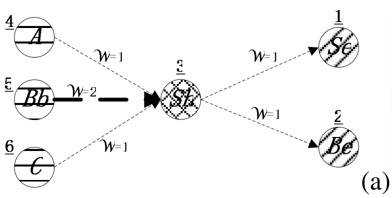
Fig. 3. Fidelity transition relationship between supporting websites of service *WSi* and *WSj* shown in Fig 1, respectively

We model the supporting websites of a web service (represented as *s*) and the fidelity transition relationship between them as a graph $G_s = (V_s, E_s)$ consisting of a set V_s of N_s sites (i.e. vertices, N_s is the number of the supporting websites of service *s*) and a set E_s of directed links (i.e. edges) that represent the fidelity transition relationship between the two linked websites. Thus, the fidelity transition relationships between the supporting websites of service *WSi* shown in Fig 1 are modeled as the graph shown in Fig 3(a).

The fidelity transition relationships between all of the supporting websites of all web services are modeled by combining the models of each individual service in the following way: $G=(V,E)$, where $V=\bigcup_{k=1}^{n} V_{ws_k}$, n is the number of web services,

$$E=\{<p,q,w_{pq}> \mid \exists E_{ws_i}, <p,q> \in E_{ws_i} \wedge w_{pq}=CountOf(<p,q>)\};$$

w_{pq} is the weight of edge <*p, q*>, and actually it is the number of the services co-supported by *p* and *q*. The weight of an edge reflects the closeness degree of the relationship between the two linked websites. According to the modeling method, the fidelity transition relationships between the supporting websites of the two web services shown in Fig 1 are modeled as a graph shown in Fig 4(a).



(a)

$$TM=\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

(b)

Fig. 4. Combined model of fidelity transition relationship between supporting websites for the web services shown in Fig 1

The adjacency matrix representation ***TM*** corresponding to the graph G in Fig 4(a) is shown in Fig4(b) (the number labeled beside the node in Fig4(a) is the id of the corresponding website).

2.2 Assessing the Fidelity of Supporting Websites

We divide the assessment of fidelity for supporting websites into two steps: 1) Initialization, i.e. initialize the fidelity for each supporting website according to the service information provided by the website itself; 2) Transition, i.e. utilize fidelity transition between related supporting websites to get the final fidelity for each website.

2.2.1 Initializing the Fidelity of Supporting Websites

For a supporting website *e*, we use $F_0(e)$ to represent the initial fidelity of *e*. We leverage the proportion of valid web service information a supporting website provides to initialize its fidelity.

Actually, the fidelity of supporting websites with different roles, i.e. *SP*, *WH*, and *SR*, has different meanings. The fidelity of a *SP* stands for the probability that it may provide available web services. The fidelity of a *WH* stands for the probability that it may host valid WSDL documents. The fidelity of a *SR* stands for the probability that it may provide valid hyperlinks to web services. Therefore, the initialization of the fidelity for supporting websites with different roles is different:

(1) For a supporting website with the role as a *SR*, we use the proportion of valid service hyperlinks it provides to initialize its fidelity. We set a service hyperlink to be valid only if service consumers could obtain the linked WSDL document which should confirm to the WSDL schema and could reach the referenced service successfully, otherwise invalid.
(2) For a supporting website with the role as a *WH*, we use the proportion of valid WSDL documents it hosts to initialize its fidelity. We set a WSDL document to be valid only if it confirms to the WSDL schema and service consumers could reach the declared service successfully, otherwise invalid.
(3) For a supporting website with the role as a *SP*, we use the proportion of available services it provides to initialize its fidelity. We send an empty testing SOAP message to the service's endpoint listed in the WSDL file and analyze the returned HTTP status codes to verify whether the service is active.

2.2.2 Utilizing Fidelity Transition between Supporting Websites

Based on the initial fidelity, we use the model described in section 2.1 to implement fidelity transition according the following three principles:

1) The fidelity propagated from a supporting website to its linked websites should be attenuated in some degree with respect to the source website's fidelity. We leverage a dampening coefficient $\beta(0<\beta<1)$ to achieve such attenuation;
2) The more targets a website links to, the less fidelity propagated from this website to each target, i.e. the fidelity propagated from a website would be in inverse proportion with the out-degree of the source website;
3) Given a supporting website *s*, the closer the relationship between a linked website and *s* is, the more fidelity propagated from *s* to the target website. We leverage the weight of edges in the model described in section 2.1 to follow this principle.

Input: ***TM*** adjacency matrix of fidelity transition graph G
F_0 initial fidelity vector produced in section 2.2.1
P, H, R the set of supporting websites with roles as , respectively
α, β α is the reservation factor, β is the dampening coefficient, $\alpha+\beta=1$
Output: ***F*** vector of fidelity for each supporting website after transition
Begin
(step1) **For** each in ***H***
$F( ) = \alpha \cdot F_0( )$ //each website reserves a part of its initial fidelity
For each in ***R*** //the fidelity of a is influenced by the fidelity of the linking
$F( ) += \beta \cdot F_0( ) \cdot [TM(,) / \sum_{=1} TM(,)]$
(step2) **For** each in ***P***
$F( ) = \alpha \cdot F_0( )$ //each website reserves a part of its initial fidelity
For each in ***H*** //the fidelity of a is influenced by the fidelity of the linking
$F( ) += \beta \cdot F( ) \cdot [TM(,) / \sum_{=1} TM(,)]$
return ***F*** //return the result
End

Fig. 5. Fidelity transition algorithm

The algorithm in Fig 5 illustrates the process of fidelity transition between supporting websites. Actually, according to the model described in section 2.1, fidelity only propagates from *SR* websites to *WH* websites (labeled as step 1), then from *WH* websites to *SP* websites (labeled as step 2). The transition strategies for step 1 and step 2 are similar, due to the limited space, we only explain step 1 in detail here.

Given a *WH* website (represented as *h*), its fidelity contains two parts: 1) one part is the reserved part of its initial fidelity (i.e. $\alpha \cdot \boldsymbol{F_0}(h)$). 2) The other part is the fidelity propagated from its linking *SR* websites (i.e. $\beta \cdot \boldsymbol{F}_0(i) \cdot [\boldsymbol{TM}(i,h) / \sum_{j=1}^{N} \boldsymbol{TM}(i,j)]$). β is the dampening coefficient introduced to follow principle 1, $\boldsymbol{F}_0(i) \cdot [\boldsymbol{TM}(i,h) / \sum_{j=1}^{N} \boldsymbol{TM}(i,j)]$ is used to follow principle 2 and 3, where $\boldsymbol{TM}(i,h)$ is the weight of edge *<i, h>*, $\sum_{j=1}^{N} \boldsymbol{TM}(i,j)$ is the out-degree of website *i*.

3 Services Selection Using Trustworthiness

3.1 Assessing Trustworthiness for Web Services

In the section above, we get the fidelity of supporting websites. In this section, we will make use of the fidelity of websites to assess the trustworthiness of web services.

Given a service *s*, we use *T(s)* to represent *s*'s trustworthiness. *T(s)* can be measured in three dimensions, i.e. the dimensions of service's *SP*, *WH*, and *SR*. Thus, *T(s)* can be represented as a triple: $T(s)=<T^P(s), T^H(s), T^R(s)>$, where $T^P(s)$, $T^H(s)$, and $T^R(s)$ is the trustworthiness measured in the dimension of *s*'s *SP*, *WH*, and *SR*, respectively. The computation is as follows:

$$T^P(s) = \frac{\sum_{p \in \boldsymbol{P}_s} F(p)}{|\boldsymbol{P}_s|} \bullet \sqrt{|\boldsymbol{P}_s|}, T^H(s) = \frac{\sum_{h \in \boldsymbol{H}_s} F(h)}{|\boldsymbol{H}_s|} \bullet \sqrt{|\boldsymbol{H}_s|}, T^R(s) = \frac{\sum_{b \in \boldsymbol{R}_s} F(b)}{|\boldsymbol{R}_s|} \bullet \sqrt{|\boldsymbol{R}_s|}$$

$\boldsymbol{P_s}$, $\boldsymbol{H_s}$, and $\boldsymbol{R_s}$ is the set of service *s*'s *SP* websites, *WH* websites, and *SR* websites, respectively. In addition, we give rewards to the services which are more popular by multiplying the number of their supporting websites.

Then we convert the multi-dimensional representation into a single-dimensional representation in the following way:

$$T(s) = \chi \bullet T^P(s) + \delta \bullet T^H(s) + \gamma \bullet T^R(s), \quad \text{where } \chi + \delta + \gamma = 1 \qquad (*)$$

3.2 Services Selection Using Trustworthiness of Web Services

The essential purpose of services selection is to provide service consumers with a novel approach to choose the appropriate service more easily. The common practice is to provide a mechanism for ranking the functionally similar service candidates. Thus, in this section, by using services' trustworthiness, we propose an approach for ranking services candidates.

Assume the criteria used for services ranking is ***K***. Given a service *s*, we use ***K(s)*** to represent such criteria of *s*. ***K(s)*** could be represented as a two-tuples: ***K(s)*** = *<M(s), T(s)>*, where *T(s)* is the trustworthiness of *s* assessed using supporting websites' fidelity, and *M(s)* is the evaluation of *s* with respect to some other metrics, e.g. availability[1]. An approach for comparing services according to ***K*** is needed.

We argue that the assessed trustworthiness of services reflects the quality of services in some degree; it plays the role as an assistant for services selection instead of a 'decision maker'. Thus, if the gap between *M(s)* is larger than a given threshold, the comparing is determined on the aspect of *M(s)*, otherwise determined by *T(s)*.

4 Case Study

The scenario of the case is: A service consumer needs a service providing weather forecast. But there are a bunch of services providing this function, and the performance of them are roughly the same. Which service should the consumer select?

There are 16 web services providing weather forecast in our collected dataset[2]. We monitored the availability of these 16 services for two weeks[3]. Then we rank these 16 services in decreasing order of availability (the availability is used as *M(s)* here). The top 6 services[4] in the ranked list are shown in the left part of Table 1. The first column from left is the ranking order; the second column from left is the ranked list in decreasing order of availability; the third column from left is the corresponding availability. The availability of the top 6 services ranges from 100% to 98%; it would be hard for users to choose a service from them just according to their availability.

Due to the small gap between the criteria of *M(s)*, we rank the 6 web services again using their assessed trustworthiness[5] (i.e. *T(s)*). The ranked list is shown in the right part of Table 1. The third column from right is the ranked list; the first column from right is the corresponding trustworthiness.

It is easy to notice the difference between the two ranked lists. For instance, in the left ranked list, the ranking order of service '*globalweather*' is 4; its ranking order is 1 in the right ranked list. Through analyzing the dataset, we find that the *SP* of service '*globalweather*' is '*webservicex.net*'. This website is a professional service provider which provides many fine web services. Moreover, service '*globalweather*' is also referenced by many websites among which there are several outstanding websites whose fidelity is high. In contrast, the fidelity of the supporting websites for service '*WeatherForecastService*' is relatively low; that is why its ranking order changes so much in the two ranked lists.

This case study indicates that the fidelity of web services' supporting websites does provide service consumers with a more comprehensive view of the services. With the

---

[1] We would not restrict the possible options for *M(s)*; we just use availability as an example here. Actually, the assessed trustworthiness of services could be used together with many metrics.

[2] The dataset was collected from the Internet according to the approach presented in [1] basically. We find these services in the dataset by searching with 'weather' 'forecast' as keywords firstly. Then we check the results manually.

[3] We check whether the service is available using the approach in section 2.2.1 once an hour. We use the ratio of successful check for each service as their availability.

[4] The reason for only listing the top 6 services is that their availability is high and very close with each other; it is hard for users to choose a service from them only using availability easily.

[5] α and β is set to 0.8 and 0.2; χ, δ and γ is set to 0.7, 0.2, and 0.1 respectively.

Table 1. Comparing of the two ranked lists

Rank	Rank according to *M(s)*		Rank according to *K(s)*		
	Service Name	***M(s)***	**Service Name**	***M(s)***	***T(s)***
1	WeatherWebService	100%	globalweather	99.2%	0.634
2	WeatherForecast Service	100%	WeatherWebService	100%	0.585
3	FastWeather	99.5%	FastWeather	99.5%	0.549
4	globalweather	99.2%	usweather	98.3%	0.549
5	usweather	98.3%	WeatherForecast	98%	0.527
6	WeatherForecast	98%	WeatherForecast Service	100%	0.43

comprehensive knowledge about web service candidates, service consumers can make a wiser decision in web services selection more easily.

5 Conclusion and Future Work

In this paper, we proposed a new feature for the trustworthiness of web service, i.e. the fidelity of supporting websites, to assist trustworthiness based services selection. Actually, the assessed trustworthiness of web service reflects the overall condition of the service's surrounding. However, we argue that the fidelity of supporting websites plays the role as an assistant instead of a 'decision maker'. The case study indicates the active effect of our approach. Some future work includes applying our approach on dynamic services selection, and proposing some other mechanisms to initialize the fidelity of supporting websites, etc.

References

1. Li, Y., Liu, Y., Zhang, L., Li, G., Xie, B., Sun, J.: An Exploratory Study of Web Services on the Internet. In: Proceedings of the IEEE International Conference on Web Services, ICWS, pp. 380–387 (2007)
2. Gyöngyi, Z., Garcia-Molina, H., Pedersen, J.: Combating Web Spam with TrustRank. In: Proceedings of the 13th International Conference on Very Large Data Bases, VLDB, pp. 576–587 (2004)
3. Liu, Y., Ngu, A., Zeng, L.: QoS Computation and Policing in Dynamic Web Service Selection. In: Proceedings of the 13th International Conference on World Wide Web, pp. 66–73. ACM Press, WWW (2004)
4. Zeng, L., Benatallah, B., Ngu, A., Dumas, M., Kalagnanam, J., Chang, H.: QoS-Aware Middleware for Web Services Composition. IEEE Transactions on Software Engineering 30(5), 311–327 (2004)
5. Maximilien, E.M., Singh, M.P.: A Framework and Ontology for Dynamic Web Services Selection. IEEE Internet Computing 8(5), 84–93 (2004)
6. Haddad, J.E., Manouvrier, M., Ramirez, G., Rukoz, M.: QoS-Driven Selection of Web Services for Transactional Composition. In: Proceedings of the IEEE International Conference on Web Services, ICWS, pp. 653–660 (2008)