Chapter 4
Web Services Discovery with Rough Sets

Maozhen Li  
Brunel University, UK

Bin Yu  
Level E Limited, UK

Vijay Sahota  
Brunel University, UK

Man Qi  
Canterbury Christ Church University, UK

ABSTRACT

Web services are emerging as a major technology for building service-oriented distributed systems. Potentially, various resources on the Internet can be virtualized as Web services for a wider use by their communities. Service discovery becomes an issue of vital importance for Web services applications. This paper presents ROSSE, a Rough Sets based Search Engine for Web service discovery. One salient feature of ROSSE lies in its capability to deal with uncertainty of service properties when matching services. A use case is presented to demonstrate the use of ROSSE for discovery of car services. ROSSE is evaluated in terms of its accuracy and efficiency in service discovery.

INTRODUCTION

Web services are emerging as a major technology for developing service-oriented distributed systems. Potentially, many resources on the Internet or the World Wide Web can be virtualized as services for a wider use by their communities. Service discovery becomes an issue of vital importance for Web service applications. As shown in Figure 1, discovered services can either be used by Web service applications or they can be composed into composite services using workflow languages such as BPEL4WS (Andrews et al, 2003). UDDI (Universal Description, Discovery and Integration, http://www.uddi.org) has been proposed and used for Web service publication and discovery. However, the search mechanism supported by UDDI is limited to keyword matches. With the
Web Services Discovery with Rough Sets

development of the Semantic Web (Berners-Lee et al, 2001), services can be annotated with metadata for enhancement of service discovery. The complexity of this metadata can range from simple annotations, to the representation of more complex relationships between services based on first order logic.

One key technology to facilitate this semantic annotation of services is OWL-S (Martin et al, 2004), an OWL (Web Ontology Language, http://www.w3.org/TR/owl-features/Reference) based ontology for encoding properties of Web services. OWL-S ontology defines a service profile for encoding a service description, a service model for specifying the behavior of a service, and a service grounding for invoking the service. Typically, a service discovery process involves a matching between the profile of a service advertisement and the profile of a service request using domain ontologies described in OWL. The service profile not only describes the functional properties of a service such as its inputs, outputs, pre-conditions, and effects (IOPEs), but also non-functional features including service name, service category, and aspects related to the quality of a service. In addition to OWL-S, another prominent effort on Semantic Web services is WSMO (Roman et al, 2005) which is built on four key concepts – ontologies, standard WSDL based Web services, goals, and mediators. WSMO stresses the role of a mediator in order to support interoperability between Web services.

However, one challenging work in service discovery is that service matchmaking should be able to deal with uncertainty in service properties when matching service advertisements with service requests. This is because in a large-scale heterogeneous system, service publishers and requestors may use their pre-defined properties to describe services, e.g. in the form of OWL-S or WSMO. For a property explicitly used in one service advertisement, it may not be explicitly used by another service advertisement within the same service category. As can be seen from Table 1, the property $P_1$ used by the service advertisement $S_1$ does not appear in the service advertisement $S_2$. When services $S_1$ and $S_2$ are matched with a query using properties $P_1$, $P_2$ and $P_3$, the property $P_1$ becomes an uncertain property when matching $S_2$. Similarly, the property $P_3$ becomes an uncertain property when matching $S_1$. Consequently, both $S_1$ and $S_2$ may not be discovered because of the existence of uncertainty of properties even though the two services are relevant to the query.

It is worth noting that properties used in service advertisements may have dependencies, e.g. both $P_1$ and $P_3$ may be dependent properties of $P_2$ when describing services $S_1$ and $S_2$ respectively. Both $S_1$ and $S_2$ can be discovered if $P_1$ and $P_3$ (which are uncertain properties in terms of the user query) can be dynamically identified and reduced in the matching process. To increase the accuracy of service discovery, a search engine should be

![Figure 1. A layered structure for service-oriented systems](image-url)

Table 1. Two service advertisements with uncertain service properties

<table>
<thead>
<tr>
<th>Service advertisements</th>
<th>Property</th>
<th>Property</th>
<th>Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td>$P_1$</td>
<td>$P_2$</td>
<td></td>
</tr>
<tr>
<td>$S_2$</td>
<td></td>
<td>$P_2$</td>
<td>$P_3$</td>
</tr>
</tbody>
</table>