Computers in Industry 63 (2012) 193–209

Contents lists available at SciVerse ScienceDirect

Computers in Industry

journal homepage: www.elsevier.com/locate/compind

A fuzzy matchmaking approach for Semantic Web Services with application to collaborative material selection

Min Liu [a,b], Weiming Shen [a,b,*], Qi Hao [b], Junwei Yan [a], Li Bai [c]

[a] The Key Laboratory of Embedded System and Service Computing, Ministry of Education, School of Electronic and Information Engineering, Tongji University, Shanghai 200092, PR China
[b] National Research Council Canada, London, Ontario N6G 4X8, Canada
[c] Shanghai Lixin University of Commerce, Shanghai 201620, PR China

ARTICLE INFO

Article history:
Received 21 May 2010
Received in revised form 19 April 2011
Accepted 7 October 2011
Available online 10 November 2011

Keywords:
Semantic Web Services
Similarity
Matchmaking
Fuzzy-set
Material selection

ABSTRACT

A critical step in the process of reusing existing WSDL-specified services for building web-based applications is the discovery of potentially relevant services. However, the category-based service discovery, such as UDDI, is clearly insufficient. Semantic Web Services, augmenting Web service descriptions using Semantic Web technology, were introduced to facilitate the publication, discovery, and execution of Web services at the semantic level. Semantic matchmaker enhances the capability of UDDI service registries in the Semantic Web Services architecture and it is able to recognize various degrees of matching for Web services. Based on the Semantic Web Service framework, semantic matchmaker, specification matching and probabilistic matching, this paper proposes a fuzzy matchmaking approach for Semantic Web Services to support a more automated and veracious service discovery process in collaborative manufacturing environments. In the proposed approach, the membership function is set up on the semantic information annotated on WSDL specification of a web service, and the similarity of service objects is evaluated by the similarity degree of service attributes. A collaborative material selection case study in the die casting process for thermoelectric fan housing is used to illustrate the proposed approach.

1. Introduction

The service-oriented computing paradigm and its realization through web service technologies provide a promising solution for the seamless integration of enterprise applications to create new value-added services [1]. Industrial practices witness a growing interest in the service-based system integration and interoperability for supply chains, dynamic alliances, e-business, extended/virtual enterprises, and virtual organizations [2–4].

A critical step in the process of reusing the existing WSDL-specified services for building large scale distributed and heterogeneous applications is the discovery of potentially relevant services as shown in Fig. 1. UDDI (Universal Description, Discovery, and Integration) [5] servers are essentially catalogs of published WSDL specifications of available services. These catalogs are organized according to categories of business activities. Providers advertise web services by adding their WSDL specifications to the appropriate UDDI directory category. Through a well-defined API, service requesters can browse the UDDI directory by category to discover existing services potentially relevant to what they want to query.

However, this category-based service discovery is clearly insufficient, because it relies on the shared common-sense understanding of an application domain among the developers who publish and request services. The assumption is that the provider knows the appropriate categories to publish its services; on the other hand, the requester must browse the "right" categories to search for relevant services and discover which candidate services are more likely to be useful in the context.

Semantic Web Services (SWS) [6,7], augmenting Web service descriptions using Semantic Web technology, were introduced to address the above problem and to facilitate the autonomous publication, discovery, and execution of Web services at the semantic level. Moreover, semantic Web service description languages, such as Ontology Web Language for Services (OWL-S) [8] and Web Service Modeling Ontology (WSMO) [9], were proposed as abstractions of syntactic Web service description languages such as WSDL. OWL-S describes the categories, the inputs, the outputs and the consequences of Web services in terms of concepts defined in the OWL ontology, and it also provides the grounding constructs for specialization into WSDL constructs for

* Corresponding author at: The Key Laboratory of Embedded System and Service Computing, Ministry of Education, School of Electronic and Information Engineering, Tongji University, Shanghai 200092, PR China.
E-mail addresses: lmin@tongji.edu.cn (M. Liu), weiming.shen@nrc.gc.ca (W. Shen), qi.hao@nrc.gc.ca (Q. Hao), jwyan@tongji.edu.cn (J. Yan), baili@lixin.edu.cn (L. Bai).

0166-3615/$ – see front matter. Crown Copyright © 2011 Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.compind.2011.10.001

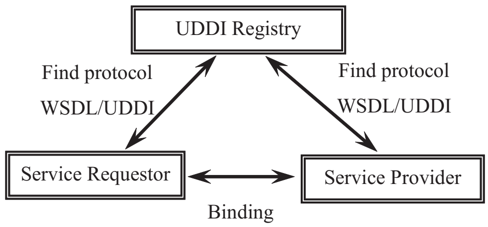

Fig. 1. Web service architecture framework.

compatibility with existing Web services. To support programmatic service discovery, Semantic matchmakers [10,11], software agents which accept and keep track of the descriptions of available services from providers and match them against the requirements from service requesters [6,12], significantly improve the capability of UDDI service registries and web services location in the Semantic Web Services architecture. Intelligent matching algorithms between advertisements and requests described in OWL-S have been implemented in these matchmaking agents to recognize various degrees of matching for web services. Furthermore, a fuzzy matchmaking framework for web services uses fuzzy logic to abstract and classify the underlying data of web services as fuzzy terms and rules [13,14]. Another fuzzy-valued similarity measure applies a fuzzy-valued description of QoS parameters to select the most appropriate services [15]. Some other fuzzy matchmaking frameworks of Semantic Web Services are proposed to abstract and classify the underlying data of Web services by fuzzy terms and rules [16–18].

In this paper, we propose a fuzzy logic based matchmaking approach to support discovery of potentially relevant Semantic Web Services for the collaborative material selection system. This approach expands the above-mentioned fuzzy matchmaking frameworks by defining the fuzzy set and similarity relation on web service attributes. It is composed of a theoretical framework for fuzzy matchmaking, and a semantic annotation specification of how the needed information of web service attributes can be captured as semantic annotation for WSDL interfaces, operations, faults, and XML Schema, and a fuzzy matchmaking algorithm for calculating the fuzzy similarity degree of web services providing the material selection services in the area of collaborative manufacturing.

The remainder of the paper is organized as follows: Section 2 reviews the related work; Section 3 presents the problem definition that describes a material selection problem in the die casting process for thermoelectric fan housing; Section 4 proposes a fuzzy matchmaking framework for semantic web service execution environment; Section 5 introduces a fuzzy definition for web services and designs the related fuzzy matchmaking algorithms for the proposed framework; Section 6 analyzes the fuzzy matchmaking results and evaluates different matchmaking approaches; Section 7 summarizes the contributions of this work and outlines some future considerations.

2. Related work

The problem of service matching and discovery is analogous to the problem of information retrieval and component retrieval [19]. Firstly, a WSDL specification declares a "software component" including a specification of its interface signature and a specification of where the actual implementation exists and how it can be used. Secondly, a WSDL specification usually includes a set of natural-language descriptions of the service and its elements. Therefore, a semantic information-retrieval method can be used to identify and order the most relevant WSDL specifications based on the similarity of their element descriptions with the query under question.

Information-retrieval is based on the probabilistic vector-space model [20,21]. In such a model, documents are represented as t-dimensional vectors, where each element of the vector corresponds to one of the distinct words contained in the document. The vectors are usually constructed after preprocessing that eliminates stop words (i.e., extremely common words that are unlikely to be searched) and stems the words in documents so that related words with a common stem are treated as a single word. Each term in the vector has a weight that reflects the importance of this term in the document. The weight value is proportional to the frequency with which the term appears in the document represented by the vector, and is inversely proportional to the number of documents that contain this term, thus, a common term importance indicator is the term-frequency inverse-document-frequency *tf-idf* ranking in probability [21], according to which, the importance of a word i in document j is $w_{ij} = tf_{ij} \times idf_i = tf_{ij} \times \log_2(N/df_i)$. Where, tf_{ij} is the frequency of term i in document j, idf_i is the inverse document frequency of term i, N is the total number of documents in the collection and df_i is the number of documents containing the term i.

Queries are also represented as vectors. The similarity between a document vector d and a query vector q can be computed as the inner product: $Sim(d,q) = d \cdot q = \sum_{i=1}^{t} W_{id} \times W_{iq}$. Where, W_{id} and W_{iq} are the weight values of term i. A higher similarity score indicates a closer similarity between the query and the retrieved documents. The major shortcoming of the vector-space model arises from the fact that each word does not have a unique, unambiguous meaning – homonyms have the same spelling but different meanings – and that many words have synonyms.

Signature matching [22] and specification matching [23] can address the component discovery. Signature matching considers only function types and ignores their behaviors so two functions with the same signature might have completely opposite behaviors. Although specification matching aims at addressing this problem by comparing software components based on formal descriptions of the semantics of the components' behaviors in terms of invariants and pre- and post-conditions of their methods, it requires the development of formal specifications and as these specifications have to be independently from the code, there is no guarantee that they correctly and completely reflect its behavior.

In service matching, automatic discovery of services relies on the ability to establish the most relevant services. In the literature [19,24–26], this problem was called "matchmaking". The following processes are needed in the matchmaking problem: (1) semantic matching – that is, identify services that deal with the same input/output (I/O) as requested; (2) capability/functional matching – that is, identify services that transform the same I/O in the desired manner; (3) ranking of services to select the best – that is, with extended discovery, rank the goal service as a whole, other than rank the individual services that participate in the goal service.

Semantic Web Services (SWS) empower Web services with semantics. There are two major initiatives aiming to realize Semantic Web Services by providing appropriate description means that enable the effective exploitation of semantic annotations with respect to discovery, composition, execution and interoperability of distributed Web Services, namely: OWL-S and WSMO. OWL-S defines an ontology for semantic markup of Web Services and is intended to enable the automation of Web Service discovery, invocation, composition, interoperation and execution monitoring by providing appropriate semantic descriptions of services. Web Service Modeling Ontology, led by the Semantic Web Services working group which includes more than 50 academic and industrial partners, aims to create an ontology for describing various aspects related to Semantic Web Services and its main objective is to solve the integration problem. Both initiatives have the common goal of providing a world-wide