

Approximate Service Retrieval

Eran Toch

Abstract

The rigidity of enterprise software systems is a fundamental problem in information systems engineering, demonstrated by the high expenses involved in developing new systems or enhancing existing ones. The problem is becoming acute as the pace of changes in information systems becomes faster.

The field of Semantic Web addresses this problem by formally specifying Web services and extending the current World-Wide-Web with formal languages and ontologies that describe domain knowledge. However, most works in the field deal with automatically inferring Web service compositions using logic-based methods. This approach has severe limitations when dealing with incomplete and uncertain information, which is the wide spread case of Web services. Furthermore, experiments we conducted indicate that the notion of similarity used by logic-based service discovery approaches is not compatible with the perception of human subjects. Rather, humans employ a much broader concept of similarity, which includes relations between services and similarities between instance sets.

This research advocates for approximate service retrieval in which the retrieval process returns approximate results. The basic question of this research is, therefore: what is considered valid approximation in service retrieval? For answering this question, we defined and evaluated a set of similarity (affinity) patterns for semantic Web services, i.e., semantically annotated Web services. The patterns are based on both semantic approximation and functional approximation: semantic approximation is done using the structure of the ontology, while the functional approximation uses the composition structure. Similarity is calculated by measuring the inherent information value lost in bridging the differences between two services. In comparison to the state-of-the-art works in the field of semantic Web service retrieval, our approach produced results with better recall and with equal precision.

As the new notion of approximate similarity imposes new challenges on the characteristics and complexity of service retrieval algorithms, we present in this research an

efficient method for approximate retrieval of Web services, which is based on indexing of semantic and functional service properties. Our results show that this index allows us to present an algorithm for query processing which is sub-linear in time complexity on the average case.

Finally, we demonstrated a proof of concept of our approach by developing OPOS-SUM, a search engine for Web services. The results of this search engine are automatically transferred into a Web-based user interface, called Liquid-Interface, allowing the user to immediately use prototypes that simulate the approximate compositions functionality.