Web Service Discovery and Client Goals

Article *in* Computer · February 2009

DOI: 10.1109/MC.2009.31 · Source: IEEE Xplore

CITATIONS
10

READS
94

2 authors, including:

Qusay H. Mahmoud
Ontario Tech University
265 PUBLICATIONS **4,589** CITATIONS

SEE PROFILE

Web Service Discovery and Client Goals

Eyhab Al-Masri and Qusay H. Mahmoud
University of Guelph

Adding Web service quality attributes to queries can help define client goals and improve search results.

Search engines greatly facilitate the finding of information on the Web, but how well they retrieve relevant results and accurately rank them can limit Web searching ability. Various studies looking at ways to improve Web search results have determined that identifying the goals of users performing queries is a key quality factor.

The underlying goals of Web service discovery are quite different from those of Web search. For example, users searching on the Web often seek information about a particular topic by viewing or obtaining a resource available on webpages. However, clients searching for Web services often seek offerings about a particular domain of interest to integrate into their systems.

Further, clients look for Web services that not only can meet their requirements but are also capable of performing the required functionalities with an acceptable degree of *quality of Web service* (QWS).

A Web service standard such as UDDI does not guarantee the validity or quality of information it contains, making it incapable of providing QWS measurements for registered Web services. In addition, search engines such as Google, Yahoo, AlltheWeb, and Baidu are not well suited for discovering services because their retrieval techniques were specifically developed to obtain pertinent data on webpages and not Web services.

Consequently, clients may not be able to effectively assess the overall quality of one or more Web services in accomplishing the required functionality. Further, the nature of how Web services are developed and deployed makes them highly volatile and subject to performance changes over time. Moreover, as Web services proliferate, a larger number of them will compete in providing similar functionalities—for example, real-time stock quotes.

The nature of Web service discovery imposes additional requirements to the most common—keyword-based—information retrieval methods. Because clients are more concerned with the degree to which Web services can achieve the required functionality, discovering Web services using quality attributes combined with keyword-based methods becomes the natural solution.

Associating quality of service with Web services might therefore provide clients with ways to improve the discovery process, help identify clients goals when performing queries, and properly articulate queries tailored to their needs.

QUALITY OF WEB SERVICE

Quality of service is a broad term often associated with network resources and widely used in the context of networking and multimedia systems. QoS is also applicable to other computing resources such as Web servers. However, Web services are not monolithic—they are small segments of larger applications. Thus, numerous Web services will eventually offer the same functionality.

QWS will become a significant differentiator among the competing implementations. To that end, QoS for Web services must transcend system-centric quality measures to encapsulate not only implementation details that may influence performance metrics but also deployment and user-experience issues.

Although the stringency of QWS client requirements can vary, a set of parameters that provide an overall assessment of the Web service's behavior in delivering the required functionality is critical. As Figure 1 shows, we distinguish two basic types of QWS parameters. *Objective* parameters correspond to measurements that are concrete or quantitative,

 Published by the IEEE Computer Society

while *subjective* parameters correspond to qualitative measures either based on the client's perception or regulated by the service provider.

Performing Web service queries that are based on objective and subjective QWS parameters can help identify client goals more accurately and thereby significantly improve the service discovery process.

WEB SERVICE QUERIES

Queries are the primary means for determining user information needs in a format that data-retrieval systems understand. In this context, terms that comprise queries are the fundamental building blocks for performing search and lookup tasks. Unfortunately, Web queries are considerably different from queries in traditional information-retrieval systems.

Several studies analyzing Web query length, structure, and reformation indicate that such queries are short and often contain one or two terms. Other studies show that a small percentage of Web users exploit advanced search features when performing queries.

Because Web service interfaces are syntactically described using XML, a large portion of Web Services Description Language documents consist of WSDL syntax—tags, rules, schemas, and so on—with only a small portion of text.

To determine the actual composition, we used our Web Service Crawler Engine (WSCE) to retrieve 3,184 WSDL documents and then analyzed the amount of textual versus syntactic information (E. Al-Masri and Q.H. Mahmoud, "Investigating Web Services on the World Wide Web," *Proc. 17th Int'l World Wide Web Conf.*, ACM Press, 2008, pp. 795-804). We found that the former comprised 18 percent of the documents and the latter the remaining 82 percent.

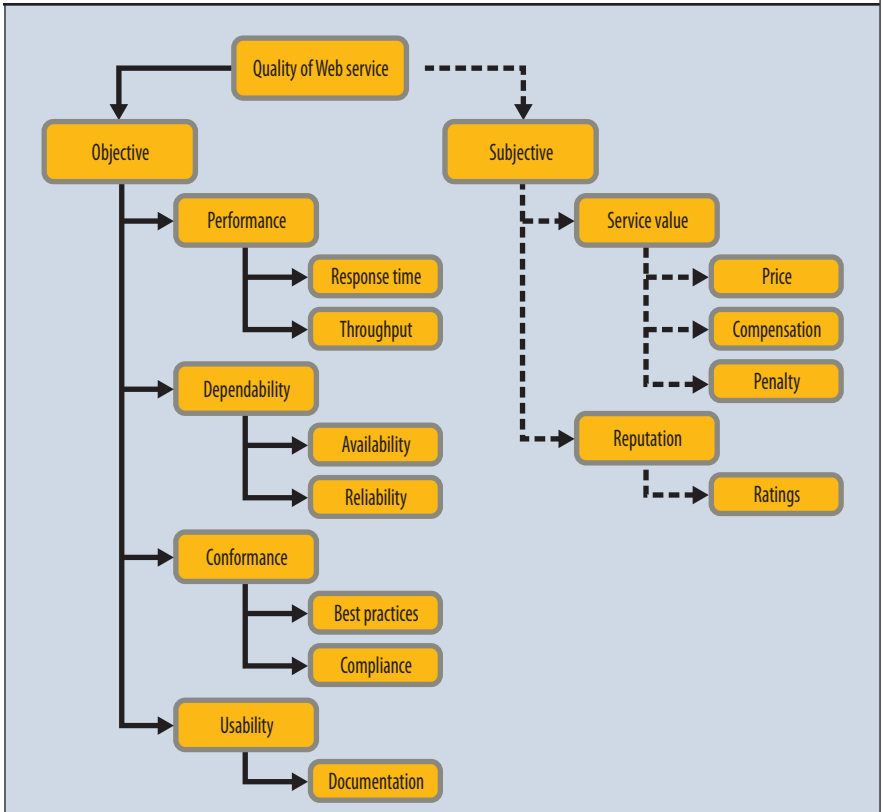

Figure 1. Basic quality of Web service parameters. Objective QWS parameters correspond to concrete or quantitative measurements, while subjective parameters correspond to qualitative measures either based on the client's perception or regulated by the service provider.

SERVICE CLIENT QUERY GOALS

Given the impact of query variations on Web service clients' experience, improving the quality of search results is critical.

In line with other research studies that have focused on Web users' search goals to improve page ranking and answer presentation, we conducted a survey to determine clients' goals when discovering Web services.

We first asked participants to become acquainted with the concept of QWS and then respond to questions relating to types of service discovery goals. Based on the 144 valid responses received, we determined that Web service client goals can be classified into two main categories with respect to QWS criteria.

A query is *exploratory* when the client wants to explore all Web services about a particular domain or interest without any QWS criteria in mind. A query is *informational* when a client wants to discover Web services about a particular domain or interest with associated QWS information.

With exploratory queries, clients presume any Web service to be the "best" answer and are willing to discover multiple results. With informational queries, however, clients want to discover only quality results. According to our survey, 16 percent of Web service client goals were exploratory and 84 percent informational.

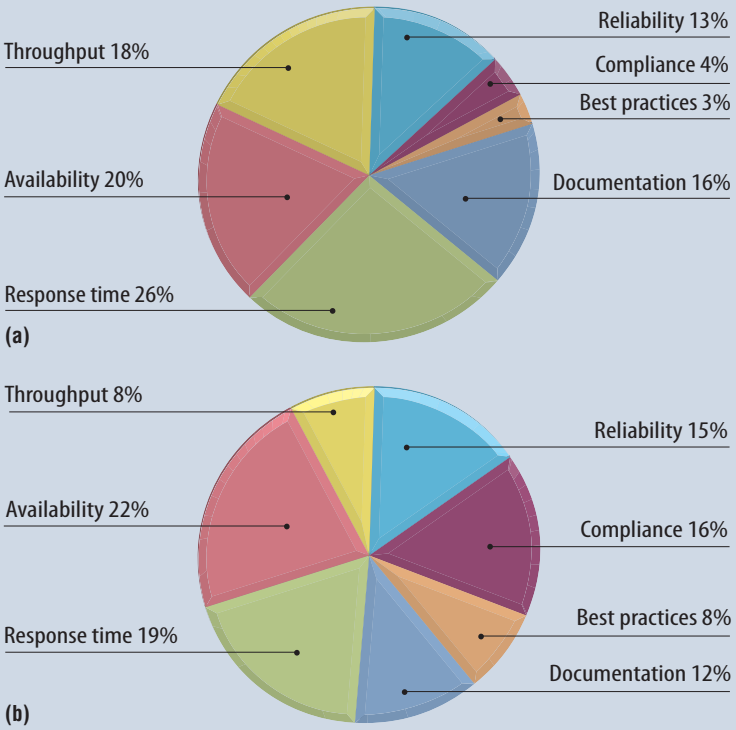

Figure 2. Percent distribution of QWS criteria for (a) directed and (b) semidirected informational queries.

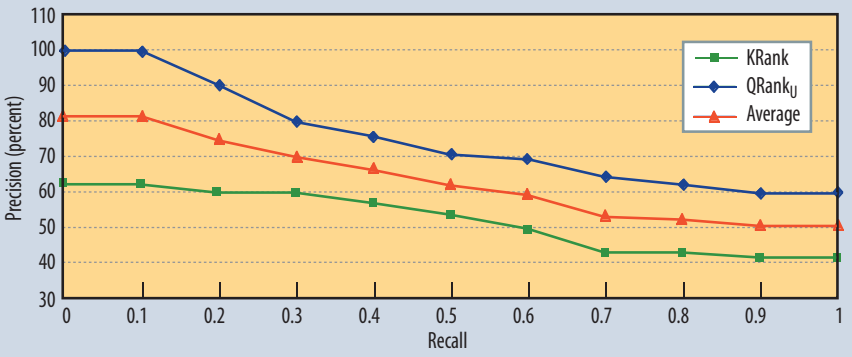

Figure 3. KRank versus QRank$_U$ performance. Average recall-precision curves for 10 service search queries at 11 standard cut-off recall levels. Using undirected informational queries improved average precision by 24 percent over using exploratory queries.

Our survey further identified three types of informational queries.

Directed queries focus on particular QWS criteria and consider their varying degrees of importance essential, while *semidirected* queries consider variations of QWS parameters but have less specific quality informational needs. Figure 2 shows the percent distribution of QWS criteria for both types of queries. *Undirected* queries look for high quality but do not rank QWS parameters in any order or according to any preference.

QWS AND SERVICE DISCOVERY

Because clients can be significantly influenced by service query variations, we used a ranking mechanism that we developed as part of our Web Service Broker Framework (E. Al-Masri and Q.H. Mahmoud, "QoS-Based Discovery and Ranking of Web Services," *Proc. 16th Int'l Conf. Computer Communications and Networks,* IEEE CS Press, 2007, pp. 529-534) to accommodate our survey's findings. Here we only consider objective QWS parameters as they are measurable.

The mechanism implements two main algorithms. KRank processes exploratory queries and uses keyword-matching techniques to rank returned Web services without considering their behavior. QRank$_U$ processes undirected informational queries and enables clients to specify those QWS parameters that they consider important. Clients can also identify a query's level of importance by assigning weights ranging from 0 to 1.

For our evaluation, we selected 10 service search queries from a subset of publicly available test suites based on actual user search studies. We then presented the search results to five different judges carefully selected from a pool of faculty members and graduate students from various institutions who are proficient in the use of service discovery and ranking algorithms. For each query, the judges evaluated the top 15 relevant search results—a total of 15 × 10 = 150 items to be judged for each result set. We obtained relevance judgments for both KRank and QRank$_U$ search results.

The relevance of returned results in a data-retrieval system indicates the appropriateness of the search results in satisfying users' information needs, and the quality of retrieved documents is a key measure of that evaluation.

Recall and precision are metrics commonly used to indicate the qual-

ity of retrieved items. Recall measures a system's ability to present all relevant items, while precision measures the system's ability to present only relevant items (R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*, Addison Wesley, 1999).

We applied the Text Retrieval Conference's methodology for calculating precision and recall at 11 standard cut-off levels to compare the performance of our two ranking algorithms. A cut-off level is simply a measure that defines the retrieved set of items.

As Figure 3 shows, we observed an average improvement in precision of 24 percent using undirected informational over exploratory queries. Based on these results, we conclude that quality-driven service discovery improves search results and provides a better means for clients to identify their service information needs.

To help the Web services community incorporate QWS parameters into the service discovery process, we have created the QWS Dataset (www.uoguelph.ca/~qmahmoud/qws). This public dataset currently consists of 2,507 Web services with their corresponding QWS metrics that we measured using commercial benchmark tools. It also includes 2,507 Web service interfaces collected using WSCE. Researchers can use the QWS Dataset for many applications including semantic analysis of WSDL documents and Web service classification, modeling, and composition.

In the future we plan to develop additional service ranking strategies and investigate the use of other QWS parameters such as cost and reputation. We also hope to extend our research in this area to mobile devices.

Eyhab Al-Masri *is a postdoctorial Fellow in the Centre for Mobile Education and Research at the University of Guelph, Ontario, Canada. Contact him at ealmasri@uoguelph.ca.*

Qusay H. Mahmoud *is an associate professor in the Department of Computing and Information Science at the University of Guelph. Contact him at qmahmoud@uoguelph.ca.*

Editor: Richard G. Mathieu, Dept. of Computer Information Systems and Management Science, College of Business, James Madison Univ., Harrisonburg, VA; mathierg@jmu.edu