Web Service Ranking in Service Networks

John Gekas
University of Essex
Wivenhoe Park, Colchester
Essex, UK.
0044 1206 873805
jgekas@essex.ac.uk

ABSTRACT

In this paper, we present the concept of *web service ranking*: a service rank is a quantitative metric that in some way shows the "importance" of a service within a web service network. The ranks we briefly introduce here are based on a variety of metrics, borrowed from graph network and social network analysis, and thus the "importance" of a web service is defined differently in the context of each ranking method. We also attempt to explain how web service ranking can be used in the context of web service discovery and composition, so that successful solutions can be found with traversing as little of the web service network as possible.

Categories and Subject Descriptors

D.2.11 [**Software Engineering**]: Software Architectures – *Domain-specific architectures.* G.2.2 [**Discrete Mathematics**]: Graph Theory – *Graph algorithms, Network problems*.

General Terms

Algorithms, Measurement.

Keywords

Web service networks, web service ranking, web service discovery-composition, graph networks, semantic web.

1. INTRODUCTION – RELATED WORK

In this paper we present the idea of applying network analysis mechanisms in networks of semantic web services: we believe that information derived from service network link analysis can prove highly useful in order to provide effective service discovery and composition mechanisms. In this context we present the notion of *web service ranking*: ranking measures given to web services belonging to a particular service network, measuring the "importance" of the service within the network. A number of ranking mechanisms can be employed, depending on the analysis criteria used in order to calculate the ranks. The ranking systems we briefly present in this paper are based on social network and graph network analysis methods, and are mainly aimed towards connectivity measures and link analysis within the web service network. We assume that web service ranking takes place within a directory of web services, for which compatible, semantic descriptions are available; available semantic information about each web service can allow us to analyse whether different services are "compatible" and can be "linked" – an important concept for performing network link analysis.

The motivation/vision behind this work is that such ranking systems could be used by service discovery and composition systems that operate as service directory search mechanisms, attempting to "extract" successful solutions to a request by traversing as little of the overall search network as possible. Thus, the ranking systems should serve as a "heuristic" guide towards successful solutions. In this context, we believe that the use of social network analysis methodologies can provide invaluable help in facilitating resource-efficient querying mechanisms for service discovery and composition. Experimental evaluation of which particular ranking methods are the most effective towards the highest service discovery performance is out of the scope of this paper. However, a detailed performance evaluation of some of the proposed ranking systems can be found at previous work done by the authors [2], [3].

Even though a number of research approaches have addressed the problems of web service discovery and composition, the areas of web service network analysis and web service ranking have not been directly approached. However, [4] propose a service composition approach based on a best-first graph search algorithm, where the services that are evaluated first by the algorithm are the ones that "can lead to the largest number of data types" – which vaguely encapsulates the notion of a service rank as a connectivity measure. Furthermore, [1] propose an architecture supporting similar service composition approaches, where the best-first composer can be led by "service ranking mechanisms specified by the user". The composer makes use of a "priority queue" (the *heap*), where services are added in descending order, depending on their rank. The particular approach mainly focuses on the architectural design of such a service directory and the specification of a query language in which the ranking systems can be specified, rather than the service ranking mechanisms themselves. Finally [5] is a search engine project specialising in searching within directories of java applications/classes. The search algorithm is based on the *rank* of a java component, where the rank is calculated as the part of the directory the component can be linked to.

We believe that network analysis can prove extremely useful when applied in the context of semantic web service networks: using the appropriate analysis tools, web services can be ranked in terms of "importance" or "usefulness" within a directory, so that the discovery of successful solutions can be performed in a resource-effective way. Finally, to our knowledge, no other analytical approach has applied concepts borrowed from network analysis to web service directories.

2. WEB SERVICE RANKING

The ranking metrics we describe can be categorised in two different ways: (a) *local* and *global*, depending on whether local or global network knowledge is needed for their estimation, and (b) *absolute* and *relative*, depending on whether the measurement

Demos and Posters of the 3rd European Semantic Web Conference (ESWC 2006), Budva, Montenegro, 11th - 14th June, 2006

is of absolute scope or refers to a particular client request. Based on this categorisation, we can see that there are four different combinations a service rank can fall under. Our service metrics are based on graph network analysis metrics that have been used is other research areas, such as social network analysis and bibliometrics. Furthermore, we have to note that these ranks rely on the analysis of the link structure of the service network (there is only one exception – *data type semantic similarity*). In the context of a web service network, a *forward link* is defined as when the output data provided by a particular service is sufficient in order to call another service. The following table shows the network metrics used, along with the category combination they fall into, based on the above categorisation:

Table 1: Categorisation of Web Service Ranks.

	Local	**Global**
Absolute	Absolute Degree(ADR), Hubs-Authorities(HAR), QoS Rank(QR).	PageRank(PR), Closeness(CR), Betweenness(BR).
Relative	Relative Degree(RDR), HITS, data type semantic similarity(SSR).	Depth-Limited Walks(WR), Flow(FR).

We will not go into a detailed description of each of the above ranks, due to space restrictions. However, it would be useful to briefly describe some of them and some of the cases they can be useful in:

Degree-based Ranks: The ADR, RDR and PR ranks are based on the degree of a web service - the number of services a web service can "feed", as a normalised percentage. RDR shows the part of the ADR that belongs in the semantic data type category specified in the request. ADR and RDR are both simple and "light" estimates of how important a service is, since they can be calculated directly, without requiring global knowledge of the service network. PR (similar to the PageRank algorithm used by the Google search engine) of course is global by nature, making it "heavier" to estimate, but also richer in informational value.

Hubs-Authorities – based Ranks: HAR and HITS both examine the relation between the number of services that link to a specific service (in-degree) and the number of services that service links to (out-degree). This is important, since there are web services and semantic data types whose one degree type tends to be much higher than the other – e.g. web services that operate on *non-functional parameters* (like e.g. *Transaction Confirmation ID*). Such high degrees would be able to attract the composer even though they could potentially lead to dead-ends in the service network, and thus should be identified.

Non-Functional Ranks: A number of the presented service ranks focus on the non-functional aspects of service composition. For instance, QR (which is calculated with regard to a specific QoS attribute) has the form of a percentage ranging from 0 to 1: this rank examines the specified QoS value of the services a service links to, estimates an average, and places it on a normalized [0,…,1] scale compared to the range of the QoS values found within the service network. Such a rank can be extremely useful when a service composition request explicitly declares QoS restrictions. Also, the FR and WR ranks examine how many alternative routes exist between two web services, which can be useful when we are interested in the reliability of a workflow solution (i.e. if a part of the solution is unavailable, will the solution fail?).

Non-Connectivity Ranks: Even though most ranks presented here are related on connectivity aspects of the web service networks, this does not always have to be the case. A useful rank that is not related to connectivity degrees is the SSR (semantic similarity rank): this rank evaluates how semantically related the data provided by two web services are. For this purpose, we assume semantic data items are defined in some form of classification/ontology – SSR is estimated as the graph distance in the ontology graph, between the data types in question.

3. DISCUSSION

The above list of web service ranks is by no means complete: similar ranks can be defined with regards to any property/attribute inherent to web service networks, depending of course on the way the rank is intended to be used. In our work, the particular ranks were chosen because they give an idea of the connectivity structure of the network.

In our work, the above ranks are used for web service discovery and composition: the service composition mechanism is defined as a graph search algorithm, that traverses the service network (search space) with the purpose of extracting successful solutions by searching as little of the search space as possible. In this sense, the composer makes use of a *Priority Queue*, where web services are added in series of how "important" and "useful" they are considered to be: this assessment is made based on the rank used at each particular case.

A detailed presentation of which ranks seem to perform better in the context of service composition is out of the scope of this paper – however, we can claim that web service ranks that measure attributes relatively to the request seem to perform higher than the absolute ones. A detailed experimental setting and performance analysis can be found at [2] and [3].

4. REFERENCES

[1] Walter Binder, Ion Constantinescu, Boi Faltings: Directory Support for Large-Scale, Automated Service Composition. Software Composition 2005: 57-66.

[2] John Gekas, Maria Fasli: Automatic Web Service Composition Based on Graph Network Analysis Metrics. OTM Conferences (2) 2005: 1571-1587.

[3] John Gekas, Maria Fasli: Service Network Structure Analysis for Web Service Discovery and Composition. In proc. BASEWEB 2006.

[4] Seog-Chan Oh, Byung-Won On, Eric J. Larson, Dongwon Lee: BF*: Web Services Discovery and Composition as Graph Search Problem. EEE 2005: 784-786.

[5] Silvestri F., Puppin D., Laforenza D., Orlando S.: Toward a Search Engine for Software Components. IEEE Web Intelligence, Beijing, China, September 20-24, 2004.

[6] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. In WWW Conference, volume 7, 1998.