Chapter I

Employing Graph Network Analysis for Web Service Composition

John Gekas
University of Essex, UK

Maria Fasli
University of Essex, UK

ABSTRACT

The Web services paradigm has enabled an increasing number of providers to host remotely accessible services. However, the true potential of such a distributed infrastructure can only be reached when such autonomic services can be combined together as parts of a workflow, in order to collectively achieve combined functionality. In this paper we present our work in the area of automatic workflow composition among Web services with semantically described functionality capabilities. For this purpose, we use a set of heuristics derived from the connectivity structure of the service repository in order to effectively guide the composition process. The methodologies presented in this paper have been inspired by research in areas such as graph network analysis, social network analysis and bibliometrics. In addition, we present comparative experimentation results in order to evaluate the presented techniques.

INTRODUCTION

The increasing popularity the Web Service paradigm and the Semantic Web have gained recently shows clearly the overall need for unified access to semantically meaningful Web-based resources, whether these resources are data sources (such as Web sites) or functionality providers (in the form
Employing Graph Network Analysis for Web Service Composition

of Web applications and Web services). Numerous and valuable efforts have been presented in these research areas, coming both from industrial and academic colleagues. Furthermore, the notion of Service-Oriented Architecture (SOA) has arisen, referring to any decentralised software platform that allows the development, deployment, and integrated access to Web service applications. It has also been strongly argued that the full potential of such a service-oriented infrastructure can only be fulfilled if effective mechanisms for resource discovery and service composition exist as well (Berners-Lee, 2001).

It seems to be common practice that the process of service discovery and composition usually takes place within a predefined search space (often called service repository or service network), as opposed to the discovery of services over the Web on the fly. In industrial environments, many real world hosts (commercial or not) that offer service brokerage, query, discovery, and/or composition mechanisms seem to be following this approach as well: remote methods (http://www.remotemethods.com/), WSIndex (http://www.wsindex.org/), and Strike Iron Web Service Marketplace (http://www.strikeiron.com/default.aspx) all operate on predefined (but extensible) Web service networks. Noncommercial service discovery providers operate on given service networks as well, such as XMethods and SalCentral.

In this article, we examine the applicability of graph network analysis as a potential solution to the Web service composition problem domain. The presented approach involves the representation of the problem domain search space (Web service directory/composition network) as a graph network, and the use of specific graph network analysis metrics in order to “guide” the composition algorithm to successful solutions. The purpose of the network analysis metrics is to examine the link structure of the composition network (partly or as a whole) and use this information in order to assess which Web services are most likely to be useful with regards to a particular composition request.

We believe that the problem domain of Web service composition poses a number of restrictions, inherent to the nature of the research problem itself:

- The size of the Web service composition network may be extremely large. Most Web service directories are open, public communities where every developer is able to publish their Web services, thus network size may be increasing rapidly.
- Web service directories are dynamic and flexible communities by nature: new Web services are constantly added to existing Web service networks, whereas other services might be made obsolete. Programmable Web (http://www.programmableWeb.com/), a Web service listing directory that also produces daily estimates regarding the size of the Web service landscape, had observed a growth rate of 2.81 new composed services per day as of June 2006.

Due to the above restrictions, it may be difficult for a composition mechanism to possess complete knowledge of the search space it operates on. This could be either due to the prohibiting size of the composition network, which would make the information collection process time-consuming and computationally-intensive, or to the fact that constant updates would need to be performed so as the information remains up to date. Thus, we believe that composition mechanisms should be able to operate under conditions of incomplete knowledge on the Web service network they operate. Furthermore, even when complete information on the search space is readily available, the service composer should be able to provide successful solutions with only partial evaluation of the search space, so as not to be affected by the potentially very large size of the composition network.