Discovering Important Services Based on Weighted $k$-Core Decomposition

Muchou Wang, Wenzhou University, Wenzhou, China  
Yiming Li, Wenzhou University, Wenzhou, China  
Sheng Luo, Wenzhou University, Wenzhou, China  
Zhuxin Hu, Wenzhou University, Wenzhou, China

ABSTRACT

With the development of service-oriented architecture, the number of services is expanding rapidly. Important services usually have high quality, and they can be recommended to users if the users do not give any keyword. However, how to discover the important services is still a problem facing many people. In this article, the authors propose a novel approach to discover important services based on service networks. First, their approach uses service networks to abstract services and the relations between them. Second, the authors employ the weighted $k$-core decomposition approach in the field of complex networks to partition the service network into a layered structure and calculate the weighted coreness value of each service node. Finally, services will be ranked according to their weighted coreness values in a descending order. The top-ranked services are the important ones the authors’ approach recommends. Experimental results on a real-world data set crawled from ProgrammableWeb validate the effectiveness of their approach.

KEYWORDS

Complex Network, K-Core Decomposition, Service Computing, Service Importance, Service Network

INTRODUCTION

With the development of the World Wide Web and the web service technology, the form, application mode, and the design and development method of software have changed significantly (Zhang & Zhang, 2013). Web service technology has become the mainstream implementation of service-oriented architecture (SOA) (Papazoglou, 2003), which advocates to use web services to compose a software system. Generally, there exist two main types of services on the Internet, i.e., SOAP (Simple Object Access Protocol) services and lightweight RESTful services (Zhang & Zhang, 2013). With the development of SOA, the number of services is expanding rapidly, and a web of services is forming, making discovering suitable services a hard task to be resolved. However, discovering suitable services is the premise of service sharing and reuse, whose quality will affect the compatibility and substitution of service composition. So far, many service discovery approaches have been proposed, such as QoS-based service discovery approaches, ontology-based service discovery approaches, and situation-based service discovery approaches (Xu et al., 2015). Important services usually have high quality, and they can be recommended to users if the users do not give any key word. To the best of our knowledge, no research work has been reported on discovering the important services.
But how to define the concept of “service importance” and how to propose approaches to discover those important services?

Network science has provided us with powerful tools to explore the structure and behavior of complex software systems. In software engineering, Pan et al. (Pan et al., 2018a; Pan et al., 2018b) analyzed the structure of Java software by using techniques in complex networks and found many useful software evolution laws. In service computing, Pan et al. (Pan et al., 2018c) clustered services by using a complex network representation of services and their couplings. Oh et al. (Oh et al., 2008) analyzed the structural properties of WSDL (Web Services Description Language) at different levels and found that they all have “small world” and “scale free” properties. They believed that such properties can help improve the performance of service discovery and service composition. The works of Pan et al. and Oh (Pan et al., 2018c; Oh et al., 2008) inspire us to use techniques in complex networks to discover the important services.

In this work, our purpose is to discover the important services for service discovery. We propose a novel approach, named DISSN (Discovering Important Services based on Service Networks) to fulfill the objective of this work. DISSN takes a two-step way. First, DISSN uses two types of service networks to abstract services and the relations between them. Second, we employ the weighted k-core decomposition approach in the field of complex networks to partition the service network into a layered structure and calculate the weighted coreness value of each service node in the service network. The weighted coreness of each node can be seen as a measure of its importance. Finally, services are ranked according to their weighted coreness values in a descending order. The top-ranked services are the important ones our approach recommends. Our empirical experiments are performed on a real-world data set crawled from the famous Mashup and Open API directory, programmableWeb (PWeb). Comparisons are performed between our approach and other four different approaches, and the similarity and difference are illustrated.

The main contributions of this work are as follows:

- We propose a weighted service network from the usage history of services. It is very general and can be used in many other similar situations. Specifically, we use the composition history of services and employ service networks to formally abstract such a composition relation: First, we use a two-mode graph to abstract the usage relations between composite (i.e., a bigger service which is composed of many atomic services) and services (i.e., atomic services which is not composed of other smaller services). Second, by one-mode projection (i.e., one coupling exists between two services if the two services are listed in the profile of at least one composite), the service coupling network can be built.

- Weighted k-core decomposition is employed to quantify the service importance from a network perspective. By a weighted network representation, we can employ the weighted k-core decomposition in the field of complex network to partition the layered structure and calculate the weighted coreness value of each service node which can be used as an indicator of service importance.

- Our approach is validated using a real-world data set crawled from the famous Mashup and service directory PWeb. Our data set is published online, making our work can be replicated.

The rest of the paper is structured as follows. First, the authors review the related works. Then, the authors give the details of their DISSN approach. Afterwards, the empirical validation of DISSN using a real-world data set is given. Finally, the authors conclude the paper.
Managing Semantic Metadata for Web/Grid Services
[www.igi-global.com/article/managing-semantic-metadata-web-grid/3090?camid=4v1a](www.igi-global.com/article/managing-semantic-metadata-web-grid/3090?camid=4v1a)

Strategic Best-in-Class Performance for Voice to Customer: Is Big Data in Logistics a Perfect Match?
[www.igi-global.com/chapter/strategic-best-in-class-performance-for-voice-to-customer/217931?camid=4v1a](www.igi-global.com/chapter/strategic-best-in-class-performance-for-voice-to-customer/217931?camid=4v1a)