A Trust-Powered Technique to Facilitate Scientific Tool Discovery and Recommendation

Jia Zhang, Carnegie Mellon University, Silicon Valley, CA, USA
Chris Lee, Carnegie Mellon University, Silicon Valley, CA, USA
Petr Votava, NASA Ames Research Center, Silicon Valley, USA & Science Mission Directorate, NASA Headquarters, Washington, D.C., USA
Tsengdar J. Lee, Science Mission Directorate, NASA Headquarters, Washington, D.C., USA
Shuai Wang, Carnegie Mellon University, Silicon Valley, CA, USA
Venkatesh Sriram, Carnegie Mellon University, Silicon Valley, CA, USA
Neeraj Saini, Carnegie Mellon University, Silicon Valley, CA, USA
Pujita Rao, Carnegie Mellon University, Silicon Valley, CA, USA
Ramakrishna Nemani, NASA Ames Research Center, Silicon Valley, CA, USA

ABSTRACT

While the open science community engenders many similar scientific tools as services, how to differentiate them and help scientists select and reuse existing software services developed by peers remains a challenge. Most of the existing service discovery approaches focus on finding candidate services based on functional and non-functional requirements as well as historical usage analysis. Complementary to the existing methods, this paper proposes to leverage human trust to facilitate software service selection and recommendation. A trust model is presented that leverages the implicit human factor to help quantify the trustworthiness of candidate services. A hierarchical Knowledge-Social-Trust (KST) network model is established to extract hidden knowledge from various publication repositories (e.g., DBLP) and social networks (e.g., Twitter and DBLP). As a proof of concept, a prototyping service has been developed to help scientists evaluate and visualize trust of services. The performance factor is studied and experience is reported.

Keywords: Knowledge-Social-Trust (KST) Network Model, Scientific Workflows, Social Networks, Trust-based Recommendation, Workflow Recommendation

DOI: 10.4018/IJWSR.2015070102
1. INTRODUCTION

Software as a Service (SaaS) has been widely considered as a promising new software delivery and provisioning technique to support modern software engineering. While software is published as programmable services (components) on the Internet, software developers can leverage appropriate services as components to build new value-added software, faster than before. In recent years, SaaS has been successfully applied to the scientific world and has greatly facilitated scientific application and workflow design and development (Roure, Goble et al. 2009). However, the open feature of the science community has led to many scientific services published on the Internet with similar functionalities. Due to the unique feature of scientific applications (services) where many parameters have to be properly tuned to ensure their appropriate use, how to help a scientist select appropriate services remains a challenge.

The Services Computing community has been working on the topic of service discovery for over a decade, when a holistic set of solutions have been proposed. Most of the methods analyze either syntactic or semantic meanings of candidate services, and conduct matchmaking processes between the candidate services and the desired requirements before making a selection. In addition, not only functional requirements but also non-functional requirements (QoS features) are used to help select proper services. However, similar to deciding business partners, the human factor may also play an important role. For example, a service published by a reputable research group may be more likely to be considered. For another example, if a scientist knows a collaborator has been using a specific service for several years, the scientist may trust the service more. Such scenarios show that human trust over a candidate service is important for effective service selection.

In this project, we have conducted a comprehensive study of the human factor in scientific service selection. The technical issues on which we targeted are two-fold:

- Where can we draw hidden knowledge to help decide the trustworthiness of a candidate service?
- How to quantify the trustworthiness of a service?

As illustrated in Figure 1, our core idea is to turn the problem of measuring trust on services into measuring trust on human (e.g., corresponding service developers). We have established a Knowledge-Social-Trust (KST) network model to compute the trustworthiness of a service producer regarding specific user requirements and context. Human trust is calculated as the summation of two main components which are “knowledge factor” and “social factor.” The knowledge factor provides a score of human’s reputation that is evaluated from users’ expertise, work experience, educational background, and so on. The social factor is evaluated from social interaction and relationship which could cover co-authorship, working relationship, friendship, and online interactions via Facebook or Twitter (called tweet). As shown in Figure 1, this project has demonstrated that a significant amount of such knowledge can be extracted from various data sources including publication repositories and social networks.

As a newly emerged type of social network, Twitter (http://twitter.com/) is increasingly extending its role from a communication channel into an important platform for seeking and sharing real-time information – a social sensor network. In this project, we have demonstrated that Twitter may contribute to service selection beyond social relationships, including knowledge reputations and service usage history.

Our previous work examined the workflow repository myExperiment (http://www.myexperiment.org) and extracted service usage history information (Zhang, Tan et al. 2011), e.g., what types of workflows in which a service is usually used; and how different services are usually used together. In this work, we have further demonstrated that hidden knowledge related to human factor can also be extracted from software usage history to support service discovery and selection.
Adaptation Space: A Design Framework for Adaptive Web Services
[www.igi-global.com/article/adaptation-space-design-framework-adaptive/3042?camid=4v1a](www.igi-global.com/article/adaptation-space-design-framework-adaptive/3042?camid=4v1a)

The Full Knowledge of Big Data in the Integration of Inter-Organizational Information: An Approach Focused on Decision Making
[www.igi-global.com/chapter/the-full-knowledge-of-big-data-in-the-integration-of-inter-organizational-information/217864?camid=4v1a](www.igi-global.com/chapter/the-full-knowledge-of-big-data-in-the-integration-of-inter-organizational-information/217864?camid=4v1a)