Usage Profiles:
A Process for Discovering Usage Patterns over Web Services and its Application to Service Evolution

Bruno Vollino, Instituto de Informática, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil
Karin Becker, Instituto de Informática, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil

ABSTRACT

As part of web services life-cycle, providers frequently face decision about changes without a clear understanding of the impact on their clients. The identification of clients’ consumption patterns constitutes invaluable information to support more effective decisions. In this paper, the authors present a framework that supports the discovery of service usage profiles, to bring awareness on the distinct groups of consumers, and their usage characterization in terms of detailed service functionality. The framework encompasses a process to cluster client applications and derive usage profiles. The paper also discusses how usage profiles can help to access the real impact on clients of incompatible changes performed over service descriptions, and presents a usage-oriented compatibility assessment algorithm. Experimental results are presented for both the profile discovery process and profile-based compatibility analysis.

Keywords: Compatibility, Data Mining, Usage Patterns, Usage Profiles, Web Service

INTRODUCTION

Web services became vital for the business of many companies in the software industry, especially with the advent of the software on demand paradigm, such as SaaS (Software as a Service). As in any business, providers have interest in understanding the needs of their clients to avoid customer attrition, and to attract new clients. Many providers focus on large scale service provision, and have very little knowledge about their clients. At the same time, they face hard decisions related to the maintenance of deployed services, service versioning to avoid breaking clients, and service redesign evolution to keep up with clients expectations. Typically, these decisions are made without a clear understanding of the possible outcomes, frequently based on worst-case scenarios. Understanding the usage clients make of services is thus invaluable to support web service life-cycle (Papazoglou, Andrikopoulos et al., 2011).

DOI: 10.4018/jwsr.2013010101
Data mining techniques have been applied in many business segments to discover knowledge about clients, which is hidden in large volumes of data (Tan, Steinbach et al., 2006). Web service mining (Liang, Chung, et al., 2006) aims at discovering patterns of service usage, i.e. specific ways in which web services (or their operations) are used repeatedly by a group of users with similar properties, as well as are correlated to each other. Usage analysis have been used to support the recommendation of services (Yu, 2012; Zhang, Ding et al., 2011; Kang, Liu et al., 2012; Rong, Liu et al., 2009), the discovery of service composition communities (Zhang, Yin et al., 2009; Wang, Wang et al., 2012), or process discovery for applications such as process documentation, conformance checking or process optimization (Motahari-Nezhad, Saint-Paul et al., 2011; Musaraj, Yoshida et al., 2010; Tang & Zou 2010; van der Aalst, 2012). van der Aalst (2012) highlights that, even when predefined interaction models are available, very often the reality differs of the expected behavior, justifying the deployment of sophisticated techniques to capture the actual usage patterns of services by their client applications.

Our work is focused on the usage analysis as a support for the service evolution lifecycle (Yamashita, Vollino et al., 2012; Silva, Vollino et al., 2012; Yamashita, Becker et al., 2012). Our approach is to empower providers with an understanding of the overall impact of changes in the whole set of client applications, enabling sound decisions in terms of evolution strategies. Providers can leverage usage impact information to make decisions about the creation, maintenance and decommissioning of versions. For that purpose, they must have a clear understanding of the patterns involved in the overall requests clients make (the operations they request, the structure of the messages exchanged, co-occurrence of operations, among others), and leverage these patterns to group clients with a similar service usage behavior, which we refer to as usage profiles.

We have explored usage profiles for the quantification of change impact in terms of affected clients (Yamashita, Vollino et al., 2012) or financial metrics (Silva, Vollino et al., 2012). Another possible application is compatibility assessment. Compatibility has been traditionally addressed in terms of a worst-case scenario, i.e. based on the possibility of breaking existing clients (Andrikopoulos, Benbernou et al., 2012; Becker, Lopes et al., 2008; Fang, Lam et al., 2007). However, clients are bound to specific functionality, rather than the entire service interface, and therefore, incompatible changes may have different effects on clients (Yamashita, Becker et al., 2011; Zou, Fang et al., 2008; Ponnekanti and Fox, 2004). Usage-oriented compatibility assessment can support service evolution management by providing relevant information about the change impact on client applications. For instance, providers can evaluate the trade-offs between the costs of provisioning multiple versions of a service, and the benefits of not breaking clients. Service designers can also proceed with certain incompatible changes they would otherwise hesitate to perform due to the possibility of breaking clients, in case the impact is not considered significant to the business.

The contributions of this paper are twofold: (a) a framework that guides the discovery of usage profiles over monitored clients requests, through a knowledge discovery process (KDD) (Tan, Steinbach et al., 2006), and (b) a profile-based compatibility assessment algorithm, which identifies the changes that are incompatible with regard to the current usage of a specific group of clients at a fine-grain.

The usage discovery framework encompasses components for: (a) monitoring and logging of clients requests, (b) inputting this data in a general purpose Usage Database, and (c) applying a knowledge discovery process to derive usage profiles. The framework predefines tasks that require minimum user intervention for the selection and transformation of relevant data, data mining using clustering techniques, and summarization of clusters as profiles. We present experiments based on synthetic data, simulating requests to a real service. The paper extensively details the ideas
Integrating Intelligent Mobile Agents with Web Services  
[www.igi-global.com/article/integrating-intelligent-mobile-agents-web/3061?camid=4v1a](www.igi-global.com/article/integrating-intelligent-mobile-agents-web/3061?camid=4v1a)

Big Data Virtualization and Visualization on Cloud  
[www.igi-global.com/chapter/big-data-virtualization-and-visualization-on-cloud/217870?camid=4v1a](www.igi-global.com/chapter/big-data-virtualization-and-visualization-on-cloud/217870?camid=4v1a)