Modeling and Respecting Privacy Specification when Composing DaaS Services*

Salah-Eddine Tbahriti, Department of Computer Science, Claude Bernard Lyon1 University, Villeurbanne, France

Brahim Medjahed, Department of Computer and Information Science, University of Michigan-Dearborn, Dearborn, MI, USA

Chirine Ghedira, Magellan Research Laboratory, IAE-Jean Moulin Lyon3 University, Lyon, France

ABSTRACT

The concept of Web service composition has undergone many evolutions and improvements including especially the apparition of new category of services, on which the composition process is made, called “Data-As-A-Service (DaaS). However, privacy is still among the key challenges that keep hampering DaaS service composition solution. Indeed services may follow different, conflicting privacy specifications with respect to the data they use and provide within a composition. In this paper, the authors propose an approach for privacy-aware composition of DaaS services. The authors’ approach allows verifying the compatibility of privacy specifications of services involved in a composition. In the case when any composition will be incompatible in terms of privacy, the authors introduce a novel approach based on negotiation to reach compatibility of concerned services. The negotiation approach is cautiously operated with without any privacy damaging of services. The authors validate the applicability of their proposal through a set of experiments.

Keywords: DaaS Service Composition, Data-As-A-Service (Daas), Negotiation, Privacy, Specifications

1. INTRODUCTION

The concept of DaaS Service Composition becomes an excellent tool to elicit integration of a wealth of information sources and to response complex and personalized queries (Roman Vaculin 2008). It allows organizations to share costs, skills, and resources by enabling remote interactions between their applications and information systems. Simply put, this concept refers to the aggregation and combination of data and/or functionalities to create value-added business processes (i.e., composite applications). While the ability of the Web service composition to integrate heterogeneous data sources and to provide customized information

DOI: 10.4018/jwsr.2012100102

Copyright © 2012, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
has boosted personal and business productivity, it has raised significant concerns regarding individuals’ and organizations’ privacy. Indeed, despite important efforts aimed at preserving privacy (Fung 2010), privacy leakage incidents on the Web continue to make the headlines. Besides, the emergence of analysis tools makes it easier to analyze and synthesize huge volumes of information, hence increasing the risk of privacy violation. Ensuring the privacy within service composition is more challenging task. Let us illustrate some privacy challenges through the following scenario.

1.1. Privacy Challenges in Service Composition

We consider the following epidemiologist’s query Q: “What are the ages, genders, zip, DNA and salaries of patients infected with H1N1; and what are the global weather conditions of the areas where these patients reside?” and a subset of services shown in Table 1. To answer this kind of queries, we use a mediator-based approach to compose DaaS services (Mahmoud Barhamgi 2010) in which, a mediator component selects, combines and orchestrates (i.e., gets output data from a service and uses it as input data to call another service) services to answer queries. It also carries out all the interactions between composed services (i.e., relays exchanged data among interconnected services in the composition). The result of the composition process is a composition plan (CP), which consists of a set of services that must be executed in a particular order depending on their input and output parameters. Input parameters are identified with a first “$” character and output parameters with a “?”. Hence, service $S(a, b)$ requires an input value $a$ and provides an output value $b$. Thus, the following CP can answer the query $Q$: $CP = \{S_{1,1}, S_{4,1}, S_{2,2}, S_{3,3}, S_{3,1}\}$. $S_{1,1}$ is invoked with $H1NI$, then for each obtained value of $SSN$, the services $S_{1,1}, S_{2,2}$ and $S_{3,3}$ are invoked to obtain the corresponding values of DNA, DoB (i.e., date-of-birth), Zip and salary. Finally, $S_{3,1}$ is invoked with the patients’ Zip to get information about the weather-conditions (note that other CP can be found with the services of Table 1).

1.2. Privacy Challenges in Service Composition

Services involved in a CP may have conflicting privacy concerns with respect to their exchanged data. Some services may require some input data that other participating services cannot disclose because of their privacy specifications. For instance, let us assume that $S_{1,1}$ discloses its data (i.e., $SSN$ values) to a third-party service for use with a “limited time” restriction. $S_{1}$ meanwhile attests that it keeps collected data (i.e., $SSN$ values) for an “unlimited time”. $S_{1,1}$ and $S_{3}$ are incompatible in terms of privacy regarding $SSN$ values. $S_{1,1}$ (which provides $SSN$ values) judges that a long retention of $SSN$ by a third-party is a risk for privacy, while $S_{1}$ would use that data as long as possible to perform several tasks that are not considered as a privacy risk. Such a conflict invalidates the CP in terms of

Table 1. Web services subset

<table>
<thead>
<tr>
<th>DaaS Services</th>
<th>Semantics Services Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{1,1}$ ($x, ?s$)</td>
<td>Returns “$SSN$” of patient infected with a disease= “$x$”</td>
</tr>
<tr>
<td>$S_{2,1}$ ($x, ?s$)</td>
<td></td>
</tr>
<tr>
<td>$S_{2,2}$ ($s, ?d, ?g$)</td>
<td>Returns $d= “DoB”, and $g= “gender” of patient identified by $s= “SSN”$</td>
</tr>
<tr>
<td>$S_{3,1}$ ($s, ?z, ?r$)</td>
<td>Returns $z= “zip”, and $r= “salary” of patient identified by $s= “SSN”$</td>
</tr>
<tr>
<td>$S_{3,1}$ ($s, ?n$)</td>
<td></td>
</tr>
<tr>
<td>$S_{3,2}$ ($s, ?n$)</td>
<td>Returns $n= “DNA” of patient identified by $s= “SSN”$</td>
</tr>
<tr>
<td>$S_{3,1}$ ($z, ?w$)</td>
<td>Returns $w= “Weather-condition” of address $z= “zip”$</td>
</tr>
</tbody>
</table>
A Metamorphic Testing Methodology for Online SOA Application Testing
[www.igi-global.com/chapter/metamorphic-testing-methodology-online-soa/41517?camid=4v1a](www.igi-global.com/chapter/metamorphic-testing-methodology-online-soa/41517?camid=4v1a)

Extensible Architecture for High-Performance, Scalable, Reliable Publish-Subscribe Eventing and Notification
[www.igi-global.com/article/extensible-architecture-high-performance-scalable/3108?camid=4v1a](www.igi-global.com/article/extensible-architecture-high-performance-scalable/3108?camid=4v1a)