Chapter 17

Selective Service Provenance in the VRESCo Runtime

Anton Michlmayr  
Vienna University of Technology, Austria

Florian Rosenberg  
CSIRO ICT Centre, Australia

Philipp Leitner  
Vienna University of Technology, Austria

Schahram Dustdar  
Vienna University of Technology, Austria

ABSTRACT

In general, provenance describes the origin and well-documented history of a given object. This notion has been applied in information systems, mainly to provide data provenance of scientific workflows. Similar to this, provenance in Service-oriented Computing has also focused on data provenance. However, the authors argue that in service-centric systems the origin and history of services is equally important. This paper presents an approach that addresses service provenance. The authors show how service provenance information can be collected and retrieved, and how security mechanisms guarantee integrity and access to this information, while also providing user-specific views on provenance. Finally, the paper gives a performance evaluation of the authors' approach, which has been integrated into the VRESCo Web service runtime environment.

INTRODUCTION

The term ‘provenance’ is commonly used to describe the origin and well-documented history of some object and exists in various areas such as fine arts, archeology or wines. Provenance information can be used to prove the authenticity and estimate the value of objects. For instance, the price of wine depends on origin, vintage, and how the wine was stored. The notion of provenance was adopted in information systems to refer to the origin of some piece of electronic data (Moreau et al., 2008). Various research efforts have addressed...
Selective Service Provenance in the VRESCo Runtime

data provenance in different domains such as e-Science (Simmhan et al., 2005).

Service-oriented architecture (SOA) (Papazoglou et al., 2007) and Web services (Weerawarana et al., 2005) represent well-known paradigms for developing flexible and cross-organizational enterprise applications. Data provenance in such applications and the provenance of business processes as realized in Business Activity Monitoring (BAM) are important issues that have been addressed by several research projects (Curbera et al., 2008), (Rajbhandari & Walker, 2006), (Tsai et al., 2007). These approaches mainly focus on the provenance of the data produced, transformed or routed through an SOA system. In contrast to that, we argue that service provenance also plays a central role, for instance during service selection. If there are multiple alternative services available, service consumers might be interested in the history of the candidates. This includes creation date, ownership and modification information, as well as Quality of Service (QoS) attributes such as failure rate or response time. Additionally, service providers are also interested in service provenance, for instance, to identify services that do not perform as expected.

In this paper, we introduce a novel service provenance approach that has been integrated into the VRESCo runtime environment (Michlmayr et al., May 2009). In most current approaches, provenance information is captured at runtime and usually managed in a dedicated provenance store. In our approach, we have enhanced the existing VRESCo event processing mechanism (Michlmayr et al., 2008) in order to capture and maintain provenance information. Events are thereby published and correlated when certain situations occur (e.g., new service is created, service revision is added, QoS changes, service operation is invoked, etc.).

Security issues such as data integrity and access control represent a central problem, which is often neglected in provenance approaches (Tan et al., 2006). On the one hand, provenance information must be accurate while on the other hand, appropriate access control mechanisms should provide access to provenance information only to authorized parties. Moreover, service owners should define who is able to access which provenance information. For instance, while employees are able to access all information, sensitive in-house information might be hidden from business partners. Such security mechanisms are also discussed in this work.

The contribution of this paper is threefold: Firstly, we present a brief summary of related work in the field and position our work among these approaches. Secondly, we present the VRESCo service provenance approach including how provenance information is collected and retrieved at runtime. Furthermore, we give examples of its usage and applicability. Thirdly, we present access control mechanisms for Web service runtimes including authentication and authorization features. This also includes various types of visibility for events that are published in the runtime. It should be noted that the present paper represents an extended version of our work published in (Michlmayr et al., July 2009).

The remainder of this paper is organized as follows. Section 2 starts with the motivation of our work, while Section 3 presents related work regarding provenance. Section 4 then describes our provenance approach in detail, by showing how provenance information is collected, retrieved and visualized. Section 5 addresses the security mechanism of our provenance system, while Section 6 presents an evaluation of our work. Finally, Section 7 concludes the paper.

MOTIVATION

In this section, we want to present the motivation of our work. As already stated above, existing work on provenance mainly focuses on data provenance, meaning how and when data was created, transformed or accessed in processes (such as