Chapter 76
Privacy–Aware Web Service Composition and Ranking

Elisa Costante
Eindhoven University of Technology, The Netherlands

Federica Paci
University of Trento, Italy

Nicola Zannone
Eindhoven University of Technology, The Netherlands

ABSTRACT
Service selection is a key issue in the Future Internet, where applications are built by composing services and content offered by different service providers. Most existing service selection schemas only focus on QoS properties of services such as throughput, latency and response time, or on their trust and reputation level. By contrast, the risk of privacy breaches arising from the selection of component services whose privacy policy is not compliant with customers’ privacy preferences is largely ignored. In this paper, the authors propose a novel privacy-preserving Web service composition and selection approach which (i) makes it possible to verify the compliance between users’ privacy requirements and providers’ privacy policies and (ii) ranks the composite Web services with respect to the privacy level they offer. The authors illustrate their approach using an eCommerce Web service as an example of service composition. Moreover, the authors present a possible Java-based implementation of the proposed approach and present an extension to WS-Policy standard to specify privacy related assertions.

INTRODUCTION
The Future Internet will be characterized by a new generation of applications built by composing services and data from different providers and organizations in order to provide users with added-value services tailored to their needs. Web services play a key role in realizing this vision because they can be advertised, located, and composed over the Internet using standards like WSDL, UDDI and BPEL, respectively. Typically, Web service composition is represented by a plan consisting of tasks that, at run-time, are instantiated to the actual services satisfying users’ requirements.

DOI: 10.4018/978-1-4666-8111-8.ch076
Due to the increasing number of services available offering similar functionalities, it is hard for users to select an optimal service composition among a list of candidate services that satisfy their needs. Therefore, service selection is a key challenge in the Future Internet.

The literature offers a large amount of work on Web service composition and selection. Most of the existing approaches focus on the identification of optimal Web services among a set of candidates based on constraints on the Quality of Service (QoS) performance of the candidates (Alrifai, Risse, & Nejdl, 2012; Chao & Younas, 2005; Hammond, Keeney, & Raiffa, 2002; Jeong, Cho, & Lee, 2009; Tran & Tsuji, 2008; Wang, Chao, Lo, Huang, & Li, 2006) or on their trust and reputation level (Maximilien & Singh, 2004; Paradesi, Doshi, & Swaika, 2009; Wang, Chao, Lo, Farmer, & Kuo, 2009; Z. Xu, Martin, Powley, & Zulkernine, 2007). To the best of our knowledge, only few works have investigated privacy issues in service selection (Massacci, Mylopoulos, & Zannone, 2006; Squicciarini, Carminati, & Karumanchi, 2011) and composition (Hewett & Kijsanayothin, 2010; Tbahriti et al., 2011; W. Xu, Venkatakrishnan, Sekar, & Ramakrishnan, 2006). Despite the limited effort, privacy plays a major role in Web service composition and selection. The orchestrator usually collects a large amount of personal data about their clients and eventually shares these data with the service providers providing the orchestrated services. This, however, may lead to risks of data misuse. For instance, a service provider may use client data for unlawful purposes. As a consequence, more and more users are considering privacy practices adopted by Web service providers as an important factor for service selection: users will more likely use Web services that customize the service provision based on users’ privacy preferences.

In this paper, we propose an approach to assist both users and Web service providers in composing and selecting optimal services with respect to their privacy preferences. We use AND/OR trees to represent the orchestration schema, component services and their privacy policies. Based on this representation, we present an algorithm that identifies the Web service compositions compliant with user privacy preferences. To help them to select the best Web service composition, our approach ranks admissible composite Web services (i.e., composite services whose privacy policy satisfy user preferences) with respect to their privacy level. The privacy level quantifies the risk of misuse of personal data based on three dimensions: sensitivity, visibility and retention period of information.

The contribution of this paper is three-fold. First, we propose a fine-grained model to express Web service providers’ privacy policies and users’ privacy preferences based on several privacy dimensions – sensitivity, purpose, retention period, visibility – while other approaches to privacy-aware service composition only consider one dimension, e.g. sensitivity or visibility. Second, we propose a Web service composition algorithm which merges into a single step the selection of services that satisfy users’ functional requirements and the selection of services compliant with users’ privacy requirements, while most existing approaches execute these two steps separately. Last but not least, we rank composite services with respect to the level of privacy they offer, while other approaches only focus on the generation of a privacy-preserving composition. We illustrate our privacy-aware composition and selection process using an eCommerce Web service as a running example. We also present a possible implementation of our approach as a Web service that does not require extending existing UDDI registries implementations. We also propose an extension to WS-Policy standard to allow service consumers and Web service providers to specify their privacy preferences and privacy policies by means of privacy assertions.

The remainder of the paper is structured as follows. In the next section we discuss related works. Then, we present a modeling framework for representing service orchestrations and users’