Engineering Applications of Artificial Intelligence 26 (2013) 96–109

Contents lists available at SciVerse ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Magentix2: A privacy-enhancing Agent Platform

Jose M. Such*, Ana García-Fornes, Agustín Espinosa, Joan Bellver

Departament de Sistemes Informàtics i Computació, Universitat Politècnica de València, Camí de Vera s/n, València, Spain

ARTICLE INFO

Article history:
Received 1 December 2011
Received in revised form
11 June 2012
Accepted 18 June 2012
Available online 10 July 2012

Keywords:
Privacy
Agent Platforms
Multi-agent Systems
Security
Trust
Reputation

ABSTRACT

Agent Platforms are the software that supports the development and execution of Multi-agent Systems. There are many Agent Platforms developed by the agent community, but they hardly consider privacy. This leads to agent-based applications that invade users' privacy. Privacy can be threatened by two main information activities: information collection and information processing. Information collection can be prevented using traditional security mechanisms. Information processing can be prevented by minimizing data identifiability, i.e., the degree by which personal information can be directly attributed to a particular individual. However, minimizing data identifiability may directly affect other crucial issues in Multi-agent Systems, such as accountability, trust, and reputation. In this paper, we present the support that the Magentix2 Agent Platform provides for preserving privacy. Specifically, it provides mechanisms to avoid information collection and information processing when they are not desired. Moreover, Magentix2 provides these mechanisms without compromising accountability, trust, and reputation. We also provide in this paper an application built on top of Magentix2 that exploits its support for preserving privacy. Finally, we provide an extensive evaluation of the support that Magentix2 provides for preserving privacy based on that application. We specifically test whether or not privacy loss can be minimized by using the support that Magentix2 provides, whether or not this support introduces a bearable performance overhead, and whether or not existing trust and reputation models can be implemented on top of Magentix2.

1. Introduction

A Multi-agent System (MAS) consists of a number of agents that interact with one-another (Wooldridge, 2002). MAS represents a key issue, especially from the development point of view in Distributed Artificial Intelligence (DAI). This is because the MAS community has produced both methodologies and actual frameworks to make the implementation of agent-based applications possible. In particular, Agent Platforms (APs) are the software that supports the development and execution of MAS. APs provide all the basic infrastructure (for message handling, tracing and monitoring, run-time management, and so on) required to create MAS (Wooldridge, 2002).

There are many APs developed by the MAS community—for an overview of current APs and the features they provide refer to Alberola et al. (2010). However, privacy is seldom considered (Piolle et al., 2007; Such et al., in press). This leads to agent-based applications that invade individuals' privacy. This is due to the fact that an agent usually encapsulates personal information describing its principal[1] Fasli (2007a), such as preferences, names, and other information. Moreover, agents carry out interactions on behalf of their principals so that they exchange personal information. For instance, agents act on behalf of their principals in agent-mediated e-commerce (Sierra, 2004), as personal assistants (Mitchell et al., 1994), in virtual worlds like Second Life[2] (Weitnauer et al., 2008), as recommenders (Montaner et al., 2003), and so on.

The modern conception of privacy started more than a hundred years ago, with the seminal work of Warren and Brandeis (1890) *The right of privacy*. These two lawyers defined privacy as "the right to be let alone". They were pioneers in considering the implications of technology in privacy. Specifically, they were very concerned about the implications of instantaneous photographs and portraits in injuring the feelings of the people in those photographs and portraits. Privacy was later recognized as a fundamental human right by the United Nations Declaration of Human Rights, the International Covenant on Civil and Political Rights, the Charter of Fundamental Rights of the European Union, and many other international treaties (Acquisti et al., 2008).

* Corresponding author.
E-mail addresses: jsuch@dsic.upv.es (J.M. Such), agarcia@dsic.upv.es (A. García-Fornes), aespinos@dsic.upv.es (A. Espinosa), jbellver@dsic.upv.es (J. Bellver).

[1] In this paper, we use the terms principal and user indistinctly to refer to the user that the agent is acting on behalf of. Principals are also called agent owners, or simply users in the related literature.

[2] http://secondlife.com/

0952-1976/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.engappai.2012.06.009

In the second part of the 20th century, Alan Westin defined privacy as "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated" (Westin, 1967). This is what is currently known as the informational self-determination right (Rannenberg et al., 2009). The concept of informational self-determination changed the right to privacy from the right to be let alone to its current incarnation as a means to limit the abuse of personal data (Schermer, 2007). Informational self-determination represents today's European understanding and regulation of privacy in the context of information and communication technology (EU Directives 95/46/EC, 45/2001/EC, and 2002/58/EC).

Despite all these regulations, as the Internet has no governing or regulating body, privacy breaches are still possible. Nowadays, in the era of global connectivity (everything is inter-connected anytime and everywhere) with more than 2 billion world-wide users with connection to the Internet as of 2011,[3] privacy is of great concern. In the real world, everyone decides (at least implicitly) what to tell other people about themselves. In the digital world, users have more or less lost effective control over their personal data. Users are therefore exposed to constant personal data collection and processing without even being aware of it Fischer-Hübner and Hedbom (2008). Garfinkel (2001) suggests that nowadays users have only one option to preserve their privacy: becoming hermits and not using online social networks, e-commerce sites, etc. Considering the increasing power and sophistication of computer applications that offer many advantages to individuals, becoming a hermit may not really be an option. However, all of these advantages come at a significant loss of privacy (Borking et al., 1999). Recent studies show that 90% of users are concerned or very concerned about privacy (Taylor, 2003). Moreover, almost 95% of web users admitted they have declined to provide personal information to web sites at one time or another when asked (Hoffman et al., 1999).

In this paper, we describe the support that the Magentix2[4] AP provides for preserving privacy. The remainder of this paper is organized as follows. Section 2 introduces the main concepts treated in this paper. Section 3 gives a brief overview of the Magentix2 AP. Section 4 presents the support that Magentix2 provides for avoiding information processing. Section 5 presents the support that Magentix2 provides for avoiding information collection. Section 6 presents an application that takes advantage of the support for preserving privacy that Magentix2 provides. Section 7 presents the evaluation we carried out. Section 8 presents related relevant works. Finally, Section 9 presents some concluding remarks and future work.

2. Background

In this paper we consider two information-related activities that can represent a major threat for privacy: information collection and information processing (Rannenberg et al., 2009; Such et al., in press). These activities can lead to many privacy breaches (Solove, 2006). We now introduce both activities and outline how these activities can be prevented when they are not desired. We also detail the implications that preventing these activities may have in accountability, trust, and reputation.

2.1. Information collection

Information collection refers to the process of gathering and storing data about an individual. Personal data is transferred online even across the Internet. Without appropriate protection mechanisms a potential attacker could easily obtain information about principals without their consent. For instance, an attacker can be listening to transferred information over the network (files, messages, e-mails, etc.) and simply gather the information flowing in the network (Stallings, 2010). Moreover, the attacker could even use the information it gathers about an individual to impersonate her/his, which is known as *identity theft* (Koops and Leenes, 2006). For instance, in Bilge et al. (2009) the authors present how to clone an existing account in an online social network and to establish a friendship connection with the victim in order to obtain information about her/him.

In order to avoid undesired information collection, sensitive personal information must be protected from access by any other third party that is different from the agent to which the information is directed to. Therefore, avoiding information collection requires security to control the access to personal information (Petkovic and Jonker, 2007). In particular, confidentiality is a security property of a system that ensures the prevention of unauthorized reading of information (Stamp, 2006). In distributed environments, confidentiality usually means that sensitive information is encrypted into a piece of data so that only parties that can decrypt that piece of data can access the sensitive information.

Confidentiality can be achieved by using existing secure data transfer technologies such as Kerberos (Neuman et al., 2005), SSL (Frier et al., 1996), and TLS (Dierks and Allen, 1999). These technologies allow the encryption of messages before transferring them and the decryption of messages once they are received. As a result, if an agent A sends a message to an agent B using these technologies, A is sure that B will be the only one able to read this message.

Confidentiality is a necessary condition to preserve privacy, but it is not sufficient. It prevents undesired information collection from unauthorized third parties. If an agent A sends personal information to an agent B in a confidential fashion, external third parties will not be able to access it. However, agent B will obviously receive this personal information. The point is that agent B can then process the received personal information, unless specific measures for preventing information processing are adopted before sending this information.

2.2. Information processing

Information processing refers to the use or transformation of data that has already been collected (Spiekermann and Cranor, 2009), even though this information has been collected by mutual consent between two parties. An example of information processing is profiling (Hildebrandt and Gutwirth, 2008): "the process of 'discovering' patterns in data that can be used to identify or represent a human or nonhuman subject (individual or group) and/or the application of profiles (sets of correlated data) to individuate and represent an individual subject or to identify a subject as a member of a group (which can be an existing community or a discovered category) and/or the application of profiles to individuate and represent individuals or groups".

One of the most common types of profiling is called buyer profiling in e-commerce environments, in which vendors obtain detailed profiles of their customers and tailor their offers regarding customers' tastes. These profiles can represent a serious threat to privacy. For instance, these profiles can be used to perform *price discrimination* (Odlyzko, 2003). Vendors could charge customers different prices for the same good according to the customers' profiles, i.e., if a vendor knows that some good is of great interest to one customer, the vendor could charge this customer more money for this good than other customers for the same good. For instance, in 2000, Amazon started to charge customers

[3] http://www.internetworldstats.com/stats.htm to consult updated statistics on world Internet users and population.

[4] http://magentix2.gti-ia.upv.es