Privacy Enhanced Authorizations and Data Handling

Ernesto Damiani, **Sabrina De Capitani di Vimercati**, **Pierangela Samarati**
Università degli Studi di Milano
Dipartimento di Tecnologie dell'Informazione
26013 Crema - Italy
{damiani,decapita,samarati}@dti.unimi.it

Abstract

The protection of privacy is an increasing concern in today's global infrastructure. Many research efforts have been therefore devoted to the development of privacy protecting technology. In particular, as an important step for helping users to maintain control over the use of their personal information, current access control solutions have to be enriched with the ability of supporting privacy requirements. In this position paper, we outline some emerging requirements and R&D challenges of privacy protection related to access control and discuss how they should be addressed in the framework of a comprehensive approach to *privacy-aware access control*.

1 Introduction

Access control is important to prevent users from accessing confidential information. Traditional access control (AC) systems are based on regulations (*policies*) that establish who can, or cannot, execute which actions on which resources. Available AC languages allow the specifications of policies with reference to generic attributes/properties of the parties and the resources involved [9]. Such languages have reached a high level of standardization. For instance, basic security protocols such as SAML, and policy languages like XACML have been developed for some time, as well as security standards for Web Services, such as the WS-* series of proposal. While these building blocks are now firmly in place, they are only starting to take possible privacy constraints on these properties into account. *Personal information privacy* is today an issue that most people are concerned about, particularly with the advent of the Internet, which has increased the possibilities of information distribution, combination, and reuse [13]. Personal information privacy is about collection, management, use, and protection of *personal identifiable information* (PII). It is widely recognized that a standardized format for privacy policies would allow consumers to quickly assess whether a particular site's privacy policy satisfies their privacy goals. This idea was behind the development of proposals like P3P and EPAL (see Section 2). However, the notions of privacy and access control policies need more work to be fully integrated in a standard fashion within a common framework of *Privacy-aware access control* [3]. Privacy-aware access control encompasses two notions: *i)* guaranteeing the desired level of privacy of information exchanged between different parties and controlling access to services/resources based on this information; and *ii)* controlling secondary use of information disclosed for the purpose of access control enforcement. A standard privacy-aware authorization language should combine these two notions and should be simple and expressive enough to support the following privacy requirements [11].

- *Openness.* Policies and privacy practices should be transparent.
- *Individual control.* Users should be able to specify who can see what information about them and when.
- *Collection limitation.* Parties who collecting personal data for the purposes of a transaction must gather no more data than what are strictly necessary for carrying out the transaction itself.
- *Purpose specification.* Those who collect and disseminate personal data must specify the purpose for which they need these data. The collected data must therefore be used only for the purposes specified.
- *Consent.* Users should be able to give their explicit and informed consent on how to use their personal data.
- *Data quality.* Those who collect and disseminate personal data must maintain accurate information. Users therefore should be able to access their personal information to modify it when needed.
- *Security.* Adequate security mechanisms for data protection have to be applied, according to the sensitivity of the personal data collected.

As a standard solution, the development of a privacy-aware authorization language specification must be based on a through understanding of information privacy as well as information security and IT systems, networks, and applications.

2 Relations with existing standards

The Internet community is experiencing a proliferation of access control (e.g., XACML, WS-*) and privacy (e.g., EPAL, P3P) technologies. The eXtensible Access Control Markup Language (XACML) with a privacy policy profile [10, 12] is an XML-based languages designed to express and interchange access control policies against objects that are themselves identified in XML. In addition to the language, XACML defines both an architecture for the evaluation of policies and a communication protocol for messages interchange. P3P [14] is a project widely acknowledged that addresses the need of a user to assess that the privacy practices adopted by a server provider comply with her privacy requirements. P3P permits the definition of server privacy practices in a standard format, allowing users to automatically understand and match these practices against their privacy preferences. Thus, users need not read the privacy policies at every site they interact with but they are always aware of the server practices in data handling. Some drawbacks of P3P are the lacking of a formal and unambiguous language to define user privacy preferences, of a technical mechanism to verify that Web sites respect users policies and of a process to negotiate the privacy practices between the interacting parties. In addition, P3P scope is restricted to Web sites only. EPAL [5] is an XML-based markup language that formalizes enterprise-internal privacy policies. It approaches the problem on the server side and addresses the need of a company to specify access control policies, with reference to attributes/properties of the requestor, to protect private information of its users. EPAL is designed to enable organizations to translate their privacy policies into IT control statements and to enforce policies that may be declared and communicated in P3P. XACML, however, provides most (if not all) of the expressive power of EPAL.

This large number of security standards is causing some confusion and seems increasing the effort for developers to build on-line services. Moreover, while the main concepts underlying these standards are quite similar, policy enforcement algorithms differ greatly from one another, due to the fact that different techniques are used to refer to resource properties values state and evaluate policy conditions based on them. Some of the main features that the privacy-aware access control language specification should consider are briefly summarized. First, it must define an interchangeable, human- and machine- readable policy format. This format should be easy-to-check for being compliant with externally defined privacy regulations and, also, it should be simple enough to be readily understood by non-specialists. Second, it must support interactive enforcement, managing complex user interactions such as the acceptance of written agreements and/or on-line payments. Third, it must be non-proprietary, which means that no private ownership may be asserted over its whole or its parts and it must also little encumbered as possible by any pre-existing Intellectual Property Rights for its whole or any part of its specifications. Fourth, it should incorporate and/or provide for interoperability with the above-referenced most widely accepted standards. Finally, it must be vendor and/or platform neutral.

3 Privacy-aware authorization language: R&D challenges

We now outline some R&D challenges that should be addressed for developing a flexible and expressive privacy-aware authorization language.

- *Context (including location) information.* Context information is used by the policy infrastructure to allow environment factors to influence how and when policy is enforced. Generally speaking, context information is a set of metadata identifying and possibly describing entities of interest, such as *subjects* and *objects*, as well as any ambient parameters concerning the technological and cultural environment (including location) where a transaction takes place. As far as policy enforcement is concerned, context contains information enabling verification of policy conditions and, therefore, it should be made available to any authorized service/application at any time and in a standard format. Still unauthorized information leaks should be prevented, also to avoid loss of privacy, for example, on the user's whereabouts [1, 2]. This requirement suggests a globally accessible, secure infrastructure for distributing context metadata, involving a variety of devices (e.g., portable computers and mobile phones) and seamlessly dealing with their different standard formats. A major factor harnessing the potential of context representation is lack of a standard context representation metadata layer. Today, context information formats vary from one user to another and also over time, e.g. location privacy may not be uniformly protected when a roaming cell-phone user relies on diverse location services provided by different mobile network operators during a journey.

Privacy-aware access control needs a sound standard model for context representation, semantically rich but unambiguous (see below). From the technological point of view, context must be highly interoperable, human-readable and processable by machines.

- *Semantic-aware privacy policies.* Tomorrow's applications will provide users with highly customized environments (e.g., for mobile learning, work and entertainment) including a much richer representation of the user's operating context, for example, representing the task that she is currently engaged in, the work team she is part of, and so on. This perspective needs highly expressive privacy-aware authorization languages, fully supporting the new generation of context-aware mobile services. While preserving enforcement efficiency, these policy languages will support sophisticated distributed evaluation, including some simple reasoning based on context information. Here, early standardization will be in our opinion a key success factor to prevent heterogeneous incompatible solution that would make context-aware mobile applications extremely costly and their development too risky. A central element of semantic aware privacy policy is the use of semantic portfolio (as an evolution, say, of P3P credentials) supporting controlled access to contextual resources (e.g. personal, company and public services) subject to user-specified privacy constraints. This field needs to combine existing standards (e.g., OWL Semantic Web reasoning engine, location tracking functionality, OWL Rule Extension, and so on) with new enforcement techniques [7, 8].

- *Secondary use.* Although users provide personal information for use in one specific context, they often have no idea on how such a personal information may be used subsequently. In other words, users do not always realize that the information they disclose for one purpose (e.g., name, date of birth, and address within an on-line transaction) may also have secondary uses (e.g., access to existing data for purposes of grouping together users on the basis of common characteristics such as age or geographic location). Therefore, even if users consent to the initial collection of their personal information, they must also be given a mechanism to specify whether or not to consent to the future use of that information in secondary applications. P3P is a good starting point but it is not widely adopted by the service providers and presents some major limitations on the user side. The main limitation is that the users have a passive role: a service provider defines a privacy policy which users can only accept or reject. We believe that a better way to enforce the informed consent principle is to offer users a richer, more active role in establishing how their personal data should be used. This implies that a new type of privacy policy, which we call *data handling policy*, regulating the secondary use of a user's personal data needs to be developed [4]. Users should therefore use these policies to define how their information will be used and processed by the counterpart. To this purpose, an automatic negotiation of preferences between users and servers needs to be supported.

- *Protection of stored data.* None of the existing standards deals with how to protect data while it is being stored, either on the client side or, more important, on the server side. This point is however attracting increasing attention from regulatory bodies and final users, and should be addressed [6].

4 Conclusions

In this paper, we have surveyed the current state and future trends in the privacy-aware access control area. We highlighted the critical necessity for privacy protection and identified some R&D challenges to be looked at. In the future, we will continue contributing to research on privacy-aware data protection, while advocating and promoting standardization efforts on these leading-edge technologies.

References

[1] C.A. Ardagna, M. Cremonini, E. Damiani, S. De Capitani di Vimercati, and P. Samarati. Location-based metadata and negotiation protocols for LBAC in a one-to-many scenario. In *Proc. of the Workshop On Security and Privacy in Mobile and Wireless Networking (SecPri MobiWi 2006)*, Coimbra, Portogal, May 2006.

[2] C.A. Ardagna, M. Cremonini, E. Damiani, S. De Capitani di Vimercati, and P. Samarati. Supporting location-based conditions in access control policies. In *Proc. of the ACM Symposium on InformAtion, Computer and Communications Security (ASIACCS'06)*, Taipei, Taiwan, March 2006.

[3] C.A. Ardagna, E. Damiani, M. Cremonini, S. De Capitani di Vimercati, and P. Samarati. The architecture of a privacy-aware access control decision component. In *Proc. of the Construction and Analysis of Safe, Secure and Interoperable Smart devices (CASSIS'05)*, Nice, France, March 2005.

[4] C.A. Ardagna, S. De Capitani di Vimercati, and P. Samarati. Enhancing user privacy through data handling policies. In *Proc. of the 20th Annual IFIP WG 11.3 Working Conference on Data and Applications Security*, Sophia Antipolis, France, July 2006.

[5] P. Ashley, S. Hada, G. Karjoth, and M. Schunter. E-P3P privacy policies and privacy authorization. In *Proc. of the ACM workshop on Privacy in the Electronic Society (WPES 2002)*, Washington, DC, USA, November 2002.

[6] E. Damiani, S. De Capitani di Vimercati, S. Foresti, S. Jajodia, and P. Samarati. Key management for multiuser encrypted databases. In *Proc. of the International Workshop on Storage Security and Survivability*, Fairfax, Virginia, USA, November 2005.

[7] E. Damiani, S. De Capitani di Vimercati, C. Fugazza, and P. Samarati. Offline expansion of XACML policies based on P3P metadata. In *Proc. of the 5th International Conference on Web Engineering*, Sydney, Australia, July 2005.

[8] E. Damiani, S. De Capitani di Vimercati, C. Fugazza, and P. Samarati. Modality conflicts in semantics aware access control. In *Proc. of the 6th International Conference on Web Engineering*, Palo Alto, California, USA, July 2006.

[9] E. Damiani, S. De Capitani di Vimercati, and P. Samarati. New paradigms for access control in open environments. In *Proc. of the 5th IEEE International Symposium on Signal Processing and Information*, Athens, Greece, December 2005.

[10] *eXtensible Access Control Markup Language (XACML) Version 2.0*, February 2005. http://docs.oasis-open.org/xacml/2.0/access_control-xacml-2.0-core-spec-os.pdf.

[11] Organization for Economic Co-operation and Development. OECD guidelines on the protection of privacy and transborder flows of personal data, 1980. www.oecd.org/document/18/0,2340,en_2649_34255_1815186_119820_1_1_1,00.html.

[12] OASIS. *Privacy Policy Profile of XACML*, September 2004. http://docs.oasis-open.org/xacml/access_control-xacml-2_0-privacy_profile-spec-cd-01.pdf.

[13] Privacy and identity management for europe. http://www.prime-project.eu.org.

[14] World Wide Web Consortium. *The Platform for Privacy Preferences 1.1 (P3P1.1) Specification*, July 2005. http://www.w3.org/TR/2005/WD-P3P11-20050701.