PRIME

Privacy-aware Access Control Policies

Introduction

PRIME project is a large-scale research effort aimed at developing an identity management system able to protect users personal information and to provide a framework that can be smoothly integrated with current architectures and online services. In this context an important service for helping users to keep the control over their personal information is represented by access control solutions enriched with the ability of supporting privacy requirements. To fully address the requirements posed by a privacy-aware access control system, the following different types of privacy policies have been defined in the context of PRIME Project.

1. *Access control policies.* They govern access/release of data/services managed by the party (as in traditional access control). Access control policies define authorization rules concerning access to data/services. Authorizations correspond to traditional (positive) rules usually enforced in access control systems. An access control rule is an expression of the form:

   ```
   <subject> with [<subject_expression>] can <actions> on
    <object> with [<object_expression>] for <purposes> if
                        [<conditions>]
   ```

2. *Release policies.* They govern release of properties/credentials/personal identifiable information (PII) of the party and specify under which conditions they can be released. Release policies define the party's preferences regarding the release of its PII by specifying to which party, for which purpose/action, and under which conditions a particular set of PII can be released. Although different in semantic access control and release policies share the same syntax.

3. *Data handling policies.* They define how personal information will be (or should be) dealt with at the receiving parties. Data handling policies regulate how PII will be handled at the receiving parties (e.g., information collected through an online service may be combined with information gathered by other services for commercial purposes). Users exploit these policies to define restrictions on secondary use of their personal information. In this way, users can manage the information also after its release. Data handling policies will be attached to the PII or data they protect, and transferred as sticky policies to the counterparts. A DHP rule is an expression of the form:

   ```
      <recipients> can <actions> for <purposes> if
   [<gen_conditions>] provided [<provisions>] follow
                       [<obligations>]
   ```

A prototype providing functionalities for integrating access control, release and data handling policies evaluation and enforcement has been developed in the context of PRIME project.

Rough use cases

The reference scenario is a distributed infrastructure that includes three parties: *i)* users are human entities that request on-line services; *ii)* service provider is the entity that provides on-line services to the users and collects personal information before granting an access to its services; *iii)* external parties are entities (e.g., business partners) to which the service provider may want to share or trade personal information of users. The functionalities offered by a service provider are defined by a set of objects/services. This scenario considers a user that needs to access a service. The user can be registered and characterized by a unique user identifier (user id, for short) or, when registration is not mandatory, characterized by a persistent user identifier (pseudonym). Three major use cases are listed in the following.

- *E-commerce.* A major factor in the evolution of the Web has been the widespread diffusion of e-commerce, that is, the ability of purchase, sell, and distribute goods and services to customers. A primary concern in the development of e-commerce was to provide a secure Global infrastructure through solutions for secure data exchange and systems for protecting e-services from unauthorized accesses. However, in the last years, the focus is shifted from the protection of server-side resources to the protection of users privacy. If users do not have confidence that their private data are managed in a privacy-oriented way by the server, they will refuse participation in e-commerce. In this scenario, it is mandatory to provide to users the possibility of protecting their privacy and their sensitive data, still accessing the on-line services.
- *Online healthcare system.* Healthcare systems support interactions among patients, medical and emergency personnel, insurance companies, and pharmacies. These systems allow for anonymous access to general information and advice, and enforces access control to individual patient records according to general rules, context (e.g., treatment, emergency), and the patient's specific choices (e.g., primary care physician, health insurance). In this context, it is important to ensure to the patients enhanced privacy functionalities to define restrictions regulating access and management of their data.
- *Location-Based Service (LBS).* Technical improvements of location technologies permit to gather location information with high accuracy and reliability. Physical location of individuals is then rapidly becoming easily available as a class of personal information that can be processed for providing a new wave of online and mobile services, such as location-based access control (LBAC) services. In addition to LBAC services, many mobile network providers offer a variety of location-based services such as point of interests proximity, friend-finder, or location information transfer in case of an accident (e.g., 911 emergency service). Such services naturally raise privacy concerns. Users consider their physical location and movements as highly privacy sensitive, and demand for solutions able to protect such an information in a variety of environments.

Distinctive features of this language

Access Control/Release Model and Language

1. *XML-based syntax*. The language provides a XML-based syntax for the definition of powerful and interoperable access control and release policies.
2. *Attribute-based restrictions*. The language supports the definition of powerful and expressive policies based on properties (attributes) associated with subjects and objects.
3. *Credential definition and integration*. The language supports requests for *certified data*, issued and signed by authorities trusted for making the statement, and *uncertified data*, signed by the owner itself.
4. *Anonymous credentials support*. The language supports definition of conditions that can be satisfied by means of zero-knowledge proof.
5. *Support for context-based conditions and metadata*. The language allows the definition of conditions based on physical position of the users and context information, and integration with metadata identifying and possibly describing entities of interest.
6. *Ontology integration*. Policy definition is fully integrated with subject and object ontology in defining access control restrictions. Also, the language takes advantages from the integration with credentials ontology that represents relationships among attributes and credentials.
7. *Interchangeable policy format*. Parties need to specify protection requirements on the data they make available using a format both human- and machine-readable, easy to inspect and interchange.
8. *Interactive enforcement*. Rather than providing a simple yes or no decision, policy evaluation provides a way of interactively applying criteria to retrieve the correct sensitive information, possibly managing complex user interactions such as the acceptance of written agreements and/or online payment.
9. *Variables support*. Currently, access control/release language supports two placeholders, one for the subject and one for the object. This solution represents a good trade-off between expressivity and simplicity but can be easily extended to support variables definition.

Data Handling Model and Language

1. *Attribute-based restrictions and XML-based syntax*. As for access control/release language, data handling language supports the definition of powerful and expressive XML-based policies based on properties associated with subjects and objects.
2. *Customized policies*. Data handling policies are defined through a negotiation between the user and the service provider. When a user requires a service, predefined policy templates are provided by the service provider as a starting point for creating data handling policies. The templates are then customized to meet different privacy requirements. A user can directly customize the templates or it can be supported by a customization process that automatically applies some user privacy preferences. If the customized data handling policies will be accepted by the service provider, the personal information provided by the user

will be labeled with the customized data handling policies. This represents the most flexible and balanced strategy for the definition of data handling policies.

3. *Stand-alone policies*. Data handling policies are defined as independent rules. Personal data are then tagged with such data handling policies, which physically follows the data when they are released to an external party, thus building a chain of control coming from the data owner.

Relation to standards

XACML v2.0

XACML version 2.0 was ratified by OASIS standards organization on 1 February 2005. Similarly to PRIME languages, XACML proposes a XML-based language allowing the specification of attribute-based restrictions. Main differences with PRIME languages are as follows.

- XACML does not explicitly support privacy features.
- Although XACML supports digital credentials exchange, it does not provide request for certified credentials.
- XACML does not support and integrate location-based conditions and ontology.

P3P/APPEL

P3P allows Web sites to declare their privacy practices in a standard and machine-readable XML format. Designed to address the need of the users to assess that the privacy practices adopted by a server provider comply with their privacy requirements, P3P has been developed by the World Wide Web Consortium (W3C). Users specify their privacy preferences through a policy language, called A P3P Preference Exchange Language (APPEL), and enforce privacy protection by means of an agent. Similarly to PRIME languages, P3P proposes a XML-based language for regulating secondary use of data disclosed for the purpose of access control enforcement. It provides restrictions on the recipients, on the data retention and on purposes. Main differences with PRIME languages are as follows.

- P3P does not support privacy practices negotiation. The users in fact can only accept the server privacy practices or stop the transaction. The opt-in/opt-out mechanisms result limiting.
- P3P does not support definition of policies based on attributes of the recipients.
- P3P does not provide protection against chains of releases (i.e., releases to third parties).