New Paradigms for Access Control in Open Environments

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

***Abstract*— Many access control models and policies have been proposed in recent years for different purposes. Access control is now evolving with the complex environments that it supports. In open environments such as the Internet, the decision to grant access to a resource is often based on the characteristics of the requestor rather than its identity. Also, people have often little control over their personal information once it has been disclosed to third parties. Privacy and secondary usage regulations are increasingly demanding attention.**

In this paper, we present the emerging trends in the access control field to address the new needs and desiderata of today's systems. In particular, we discuss two new access control paradigms and outline some R&D challenges that should be addressed.

***Index Terms*— Access control, privacy, semantics**

I. Introduction

Everyday business operations of companies, as well as the activities of individual citizens and the services offered by the public sector are becoming increasingly centered around the communication infrastructure. Thanks to availability of post-third generation mobile networks, user transactions are no longer bound to the traditional office-centered environment, but can be started virtually anywhere and at any hour [8]. Therefore, resources may be accessed in a variety of contexts, and users requesting access may be required to disclose a rich set of distributed information about themselves, including dynamic properties such as their location or communication device as well as conventional, identity-related user attributes. Experience has shown that some users may choose to abort a transaction rather than disclosing what they consider private information, while others wish to retain a degree of control over its secondary use, in order to protect their privacy. To build a pervasive, seamless computing and telecommunication infrastructure environment, some novel access control models and languages need to be developed. To this purpose, the main functional requirements that must be taken into account are the following.

- *Privacy.* Access control needs to guarantee the enforcement of the privacy requirements. The problems that need to be considered are principally two: first, the definition of privacy access control policies requires considering, expressing, and combining protection requirements taking in account both direct and indirect release of information. Second, information may not be under the control of a single authority; privacy policies related to information may take in consideration the privacy requirements of the owner, but also the privacy requirements of the collector and possible privacy law. Note that, all privacy requirements may be associated with the data during their movement among different parties in the system and the parties that receive the information must follow the privacy rules when managing them. These multiple authorities scenario should be supported from the administration point of view providing solutions for modular, large-scale, scalable policy composition and interaction [4], [27].
- *Anonymity.* Many services do not need to know the real identity of a user (e.g., a digital library could be accessible by a user that presents a certificate issued by a given association and stating the user's membership in the association). Pseudonyms, multiple digital identities, and even anonymous accesses must be adopted when possible.
- *Expressiveness.* Simplicity and expressiveness of the access control system is another key aspect. The access control system should be simple to make easy the management task of specifying and maintaining the security specifications. The access control system should be expressive to make it possible to specify in a flexible way different protection requirements that may need to be imposed on different objects.
- *Client-side restrictions.* In addition to traditional server-side access control rules, users should be able to specify restrictions about the usage of their information once released to a third party.
- *Semantics-aware restrictions.* Advanced metadata pave the way for the "semantics-aware", including, infrastructure self-awareness (semantic grid), awareness of application concepts (e.g., via domain ontologies), and awareness of physical locations (based on mobile 3G technologies).
- *Context-aware restrictions.* Protection requirements may need to depend on the evaluation of some conditions (e.g., system's predicates or conditions that make access dependent on the information being accessed). An access control system should then allow the specification of

generic constraints on subjects, objects, and on contextual information.

In this paper, we present the emerging trends in the access control field to address the new needs and desiderata of today's systems. The remainder of this paper is organized as follows. Section II describes the main characteristics of two proposals for addressing access control in open environments. Section III presents new research challenges in the access control area. Finally, Section IV concludes the paper.

II. Extending traditional access control

We now present the main features of the attribute-based and semantics-aware access control paradigms [16]. These two new paradigms address some of the functional requirements above-mentioned. As we will see, they are expressive and can support anonymity, context-aware restrictions, and semantics-aware restrictions. The support of privacy permissions and rights requires extended access control solutions. Some relevant work in this direction is described in [2], [3]. However, the problem of enforcing and implementing a privacy-aware access control solution is still an open issue.

A. Attribute-based access control

Whereas traditional access control systems were based on the identity of the party requesting a resource, in open environments such as the Internet, this approach is not effective because often the requestor and the resource belong to different domains. Therefore, a more appropriate approach it seems one where the access decision is based on properties (attributes) of the requestor and of the resource. Basing authorization on attributes of the resource/service requester provides flexibility and scalability that is essential in the context of large distributed open systems, where subjects are identified by their characteristics. As an example, suppose that an online bookstore offers discounts to students at accredited universities. In this case, a user has to submit a student ID to receive a student discount. The online bookstore is therefore not interested in the identity of the user; it only needs a proof that the user is a student.

Attribute-based access control differs from the traditional discretionary access control model by replacing both the *subject* by a set of attributes and *objects* by descriptions in terms of available properties associated with them. The meaning of a stated attribute may be a granted capability for a service, an identity, or a non-identifying characteristic of a user (e.g., a skill). Here, the basic idea is that not all access control decisions are identity-based. Therefore, new approaches based on *digital certificate* are becoming widespread, being much more suitable for the open communication infrastructure [26]. Digital credentials are digitally signed assertions about the credential owner by a credential issuer. Recent research and development efforts are based on the *attribute certificate* [17] that can be used for supporting attribute-based access control systems. An attribute certificate contains attributes that specify access control information associated with the certificate holder (e.g., age, citizenship, credit status, group membership, role, security clearance). The decision to access a resource is therefore based on the attributes in the requestor's credentials. In an attribute certificate, attributes need to be protected in a similar way to any other certificate: they are therefore digitally signed sets of attributes created by *attribute authorities*. Attribute authorities are responsible for their certificates during their whole lifetime, as well as issuing them.

A first attempt to provide a uniform framework for attribute-based access control specification and enforcement was presented by Bonatti and Samarati in [5]. They propose a uniform framework for regulating service access and information disclosure in an open, distributed network system like the Web. Like in previous proposals, access regulations are specified as logical rules, where some predicates are explicitly identified. Attribute certificates are modeled as *credential expressions* of the form credential_name(attribute_list), where credential_name is the attribute credential name and attribute_list is a possibly empty list of elements of the form "attribute_name=value_term", where value_term is either a ground value or a variable. Besides credentials, the proposal also allows to reason about declarations (i.e., unsigned statements) and user-profiles that the server can maintain and exploit for taking the access decision. Yu et al. [31], [32], [30] developed a service negotiation framework for requesters and providers to gradually expose their attributes.

These approaches, however, are not designed for enforcing privacy policies. For instance, privacy issues that are not addressed by traditional approaches include protecting user identities by providing *anonymity*, *pseudonymity*, *unlinkability*, and *unobservability* of users at communication level, system level, or application level. Therefore, the consideration of privacy issues introduces the need for rethinking authorization policies and models and the development of new paradigms for access control and, in particular, authorization specification and enforcement. In [2] we have introduced different types of *privacy policies* (i.e., traditional access control policies, release policies, data handling policies, and sanitized policies) and presented a privacy-enhanced authorization model and language. The model allows the definition and enforcement of powerful and flexible access restrictions based on generic properties associated with subjects and objects. In short, the main elements of the authorization rules are the following.

- *Subject expression.* Each expression identifies a set of subjects having specific properties. Each user is then associated with a *profile* that defines names and values of some properties that characterize the user.
- *Object expression.* The characterization of the entities to be protected should be specified through expressions. As for subjects, each object is associated with a profile which defines the names and values of their properties.
- *Actions.* Policies must be able to make distinctions based on the type of actions being requested on objects.
- *Purposes.* Data access requests are made for a specific purpose, which represents how the data is going to be used by the recipient.
- *Conditions.* Additional conditions such as conditions dic-

tated by legislation, location-based conditions, and trust conditions.
- *Obligations.* To improve privacy, users can define some obligations attached to the data. Therefore, when a certain access is allowed, the parties involved must take some additional steps, following the defined obligations.

Each access request results in an *access decision* that can take three different forms: *yes* (i.e., the access request is granted), *no* (i.e., the access request is denied), and *undefined*. An undefined response means that current information is insufficient to determine whether the request can be granted or denied. For instance, suppose that a user can access a service if she is at least eighteen and can provide a credit card number. Two cases can occur: *i)* the system knows that the user is not yet eighteen and therefore returns a negative response; *ii)* the user has proved that she is eighteen and the system returns an undefined response together with the request to provide the number of a credit card.

B. Semantics-aware access control

Security and privacy concerns are increasingly important in open environments, where controlling the release, retention, and secondary use of personal data have become key issues. While encryption-based technologies such as the Public Key Infrastructure [23] guarantee credentials' unforgeability, a framework for empowering the user with full control over information release during the exchange of certificates on the Web is still missing [14], [22]. Key requirements for this framework include the following two aspects.

- A data model for representing credential information and a language enabling: *i)* end users to state policies expressing their preferences on the disclosure and acceptable secondary use of personal data; *ii)* service providers to dynamically define the requirements to be met by clients.
- A decision mechanism enabling uniform evaluation and enforcement of policies.

Advanced modeling of *Personally Identifiable Information* (PII) (i.e., any kind of information that can be linked to a specified individual) allows for controlling data release according to users' privacy requirements. On the other hand such a model can assist system administrators in the specification of the information required for a resource or service to be granted. The structure referenced by policy requirements is rooted on a set of application-dependent elements referencing the formal definitions of credentials to be stored and requested by the system. This part of the knowledge base is aggregated from a pool of heterogeneous normative sources and constitutes the domain knowledge the negotiating parties are required to agree upon. Privacy policies can then constrain the disclosure of PII and are associated with either instances of credentials (e.g., the VISA card) or abstractions defining them (e.g., credit card information). Abstractions allow the sharing of the same preferences among multiple instances of the same credential. Moreover, data items within an instance of credential (e.g., the name on the credit card) or the corresponding fields in the underlying definition can be singled out to reach the finest granularity. Furthermore, correspondences can be drawn between the attributes of distinct credentials (e.g., the user's name) by mapping their definitions to a common structure. This way, privacy preferences can span along multiple credentials. To refer and reason about credentials, we developed an approach that exploits the base data schema of the Platform for Privacy Preferences (P3P) for expressing data-collection and data-use practice in a standard format [1]. Our approach relies on an ontology-based representation of the standard P3P base data schema, showing the internal structure of complex credentials, like a driving license, in term of fine-grained items such as a surname or a birth-date. Then, the P3P ontology is applied in order to augment access control policies (e.g., ones written in a standard language like XACML [1]), obtaining enhanced rule set including the reasoning patterns required by the application. For instance, if the original access control policy required the surname on a driving license to be `Smith`, its augmented version will accept as a valid alternative that the same value appears on a passport. Augmented access control policies can be used as a replacement of the original ones, automatically incorporating ontology knowledge; perhaps more promisingly, they can also be evaluated side by side with the original ones, allow for flagging cases of semantics inconsistency and assessing the scope and quality of the original policies.

Although P3P base data schema provides us with a well understood type-space for the definition of cross-cutting properties linking semantically equivalent data items, P3P data schemata still lack the expressive power and the clearly defined semantics required for the definition of complex user credentials [29]. Semantic Web languages like OWL [28] and RDFS [24] lend themselves very well to advanced representation of personal information inasmuch they allow for integrating credentials' structural definitions with a data schema expressing the meaning of the information to be exchanged, thus defining cross-cutting relationships linking semantically equivalent data items (e.g., birth dates) appearing in multiple credentials (e.g., a passport and a driver license). In our previous work [12] we addressed the problem of using ontology to model the *portfolio*, that is the entity enclosing all the sensitive data stored by the system at both sides of the transaction [7]. Specifically, we presented some techniques allowing for the informative content of a user credential to be decomposed into atomic components, so that users can non-ambiguously single out items to be released. Obviously, this work represents only a first step towards a semantics-aware access control, and much work is still to be done before this encoding can be used in practice. The first required extension is related to the data type awareness of OWL: current reasoners are only required to support the `xsd:integer` and `xsd:string` datatypes, while our model requires the full expressiveness of XML Schema in the definition of data type properties representing the portfolio items. We also need to constrain values allowed by such properties so that it is possible to specify, among the possible instances of a given credential, those satisfying the

requirements imposed by a policy. Other than specifying single data items, policy rules will also address whole credentials provided with zero-knowledge proof technologies such as Idemix [6], [21]. The second improvement is mapping the policy preference language to the OWL syntax so that policies and requirements associated with them can be exchanged as triples: this way it is not necessary to translate policies from the original format to the corresponding OWL representation.

III. Research challenges

We now outline some of the key aspects and challenges to be addressed for managing and protecting information in the service-centric information society.

a) Contextual information: Context information is used by policy infrastructures to allow environment factors to influence how and when policy is enforced. Generally speaking, context defined the conditions that must be verified for the policy to be applied. Therefore, context information should be made available to any authorized service/application at any time. Still unauthorized information leaks should be prevented, also to avoid loss of privacy, for example, on the user's whereabout. This requirement suggests a globally accessible, secure infrastructure for distributing context metadata, involving a variety of devices from portable computers to mobile phones and seamlessly dealing with their different standard formats. Also, information generated from different applications should not remain restricted to the local context; integrating context information with user profiles paves the way to advanced applications where user context can be exploited for service discovery and composition. In order to achieve these goals, context representation must be semantically unambiguous, interoperable, human readable and processable by machines.

b) Ontologies: Due to the openness of the scenario and the richness and variety of security requirements and attributes that may need to be considered, it is important to provide parties with a means to understand each other with respect to the properties they enjoy (or request the counterpart to enjoy). Therefore, common languages, dictionaries, and *ontologies* must be developed. In interoperable e-business architectures based on the semantic web vision, *ontology-based domain models* are used as controlled vocabularies for resources description, allowing users to obtain the right resources at the right time [15]. While research on developing standards and tools that ultimately will lead to the existence of the semantic web is increasing [25], many issues still need to be solved to enable integrating the result of this research into access control languages. For instance, the high expressive power of semantic web metadata allows for using multiple different syntaxes to carry the same semantics. While no constraints can be posed a priori on the content of resources' descriptors, a standard syntax must be adopted for metadata used to describe subjects and objects within access control policies. Also, a standard syntax should be used for subjects' descriptions. In our view, metadata underlying access control, reputation, and trust must come together with those aimed at reputation management as the cornerstone of the new generation secure information infrastructure.

c) Filtering and renaming of policies: As discussed is Section II, since access control can return the information about which conditions need to be satisfied for the access to be granted ("undefined" decision), the problem of communicating such conditions to the counterpart arises. To fix the ideas, let us see the problem from the point of view of the server (the client's point of view is symmetrical). The naive way to formulate a credential request, that is, giving the client a list with all the possible sets of credentials that would enable the service, is not feasible, due to the large number of possible alternatives. Also, the communication process should not disclose "too much" of the underlying security policy, which might also be regarded as sensitive information. As an example, consider a medical database where the access to patient records is granted if the requester is a senior researcher and is a member of a particular association. Consider now user `Alice` that wants to access the patient records in the medical database. `Alice` needs to show or prove to the system that she satisfies the policy. However, neither `Alice` nor the system want to disclose their private information. More precisely, `Alice` does not want to reveal her credentials, as her credentials contain sensitive information about her (e.g., health, date of birth, marital status, and so on). Analogously, the system does not want to reveal the policy, even to those who satisfy the policy, so as to make it harder for a malicious user to know which credentials she should illicitly obtain.

d) Outsourcing: A recent trend in the information technology area is represented by *database outsourcing*. Companies shifted therefore from fully local management to outsourcing the administration of their databases by using externally service providers [9], [13], [18], [19], [20]. The main problem in outsourcing data to external service providers is that sensitive data become stored on a site that is not under the data owner's direct control. This problem is solved by encrypting the data and by adopting techniques that enable external service providers to execute queries on encrypted data, otherwise all the relations involved in a query would have to be sent to the data owner for query execution. Even if database outsourcing has been studied in-depth in the last few years, there are new interesting research challenges that have to be investigated. In particular, the problem of guaranteeing an efficient mechanism for implementing selective access to the remote database is an open issue. As a matter of fact, all the existing proposals for designing and querying encrypted/indexing outsourced databases assume the client has complete access to the query result. However, this assumption does not fit real world applications, where different users may have different access privileges. A first attempt to address this issue is presented in [10], [11], where the authors propose an approach for enforcing access control policies based on key-derivation mechanisms.

e) Negotiation strategy: Credentials grant parties different choices with respect to what release (or ask) the counterpart and when to do it, thus allowing for multiple trust

negotiation strategies [32]. For instance, an *eager* strategy, requires parties to turn over all their credentials if the release policy for them is satisfied, without waiting for the credentials to be requested. By contrast, a *parsimonious* strategy requires that parties only release credentials upon explicit request by the server (avoiding unnecessary releases).

f) Composite services: In case of a composite service (i.e., a service that is decomposable into other services called *component services*) there must be some semi-automatic mechanisms to calculate the access control policy of a composite service from the access control policies of its component services.

IV. Conclusions

In this paper, we have surveyed the current state and future trends in the access control area and we have seen that they are a crucial part of tomorrow's communication infrastructure supporting mobile computing systems. We highlighted the critical necessity for privacy protection and described how advanced metadata can be used to address the new challenges posed by access control in an open, service-oriented environment. 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.

Acknowledgments

This work was supported in part by the European Union within the PRIME Project in the FP6/IST Programme under contract IST-2002-507591 and by the Italian MIUR within the KIWI and MAPS projects.

References

[1] C. Ardagna, 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.

[2] C. Ardagna, E. Damiani, S. De Capitani di Vimercati, and P. Samarati, "Towards privacy-enhanced authorization policies and languages," in *Proc. of the 19th Annual IFIP WG 11.3 Working Conference on Data and Applications Security (IFIP)*, Nathan Hale Inn, University of Connecticut, Storrs, USA, August 2005.

[3] P. Ashley, S. Hada, G. Karjoth, C. Powers, and M. Schunter, *Enterprise Privacy Authorization Language (EPAL 1.1)*, 2003, http://www.zurich.ibm.com/security/enterprise-privacy/epal.

[4] P. Bonatti, S. De Capitani di Vimercati, and P. Samarati, "An algebra for composing access control policies," *ACM Transactions on Information and System Security*, vol. 5, no. 1, pp. 1–35, February 2002.

[5] P. Bonatti and P. Samarati, "A unified framework for regulating access and information release on the web," *Journal of Computer Security*, vol. 10, no. 3, pp. 241–272, 2002.

[6] J. Camenisch and E. V. Herreweghen, *Design and Implementation of the idemix Anonymous Credential System*, IBM Zurich Research Laboratory, http://www.zurich.ibm.com/ jca/papers/camvan02.pdf.

[7] P. Ceravolo, E. Damiani, S. De Capitani di Vimercati, C. Fugazza, and P. Samarati, "Advanced metadata for privacy-aware representation of credentials," in *Proc. of the International Workshop on Privacy Data Management*, Tokyo, Japan, 2005.

[8] A. Corallo, M. Cremonini, E. Damiani, S. De Capitani di Vimercati, G. Elia, and P. Samarati, "Security, privacy, and trust in mobile systems," in *Mobile and Wireless Systems beyond 3G: managing new business opportunities*. Idea Group Inc., 2004.

[9] E. Damiani, S. De Capitani di Vimercati, M. Finetti, S. Paraboschi, P. Samarati, and S. Jajodia, "Implementation of a storage mechanism for untrusted DBMSs," in *Proc. of the Second International IEEE Security in Storage Workshop*, Washington DC, USA, May 2003.

[10] E. Damiani, S. De Capitani di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, and P. Samarati, "Metadata management in outsourced encrypted databases," in *Proc. of the 2nd VLDB Workshop on Secure Data Management (SDM'05)*, Trondheim, Norway, September 2005.

[11] 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.

[12] E. Damiani, S. De Capitani di Vimercati, C. Fugazza, and P. Samarati, "Extending policy languages to the semantic web," in *Proc. of the International Conference on Web Engineering*, Munich, Germany, July 2004.

[13] E. Damiani, S. De Capitani di Vimercati, S. Jajodia, S. Paraboschi, and P. Samarati, "Balancing confidentiality and efficiency in untrusted relational DBMSs," in *Proc. of the 10th ACM Conference on Computer and Communications Security*, Washington, DC, USA, October 27-31 2003.

[14] E. Damiani, S. D. C. di Vimercati, and P. Samarati, "Managing multiple and dependable identities," *IEEE Internet Computing*, vol. 7, no. 6, pp. 29–37, November/December 2003.

[15] J. Davies, D. Fensel, and F. van Harmelen, *Towards the Semantic Web: Ontology-Driven Knowledge Management*. John Wiley & Sons, Ltd, 2002.

[16] S. De Capitani di Vimercati, P. Samarati, and S. Jajodia, "Policies, models, and languages for access control," in *Proc. of the Workshop on Databases in Networked Information Systems*, University of Aizu, Japan, March 2005.

[17] S. Farrell and R. Housley, "An internet attribute certificate profile for authorization," RFC 3281, April 2002.

[18] H. Hacigümüs, B. Iyer, and S. Mehrotra, "Providing database as a service," in *Proc. of 18th International Conference on Data Engineering*, San Jose, California, USA, February 2002.

[19] H. Hacigümüs, B. Iyer, and S. Mehrotra, "Ensuring integrity of encrypted databases in database as a service model," in *Proc. of the IFIP Conference on Data and Applications Security*, Estes Park Colorado, August 2003.

[20] H. Hacigümüs, B. Iyer, S. Mehrotra, and C. Li, "Executing SQL over encrypted data in the database-service-provider model." in *Proc. of the ACM SIGMOD'2002*, Madison, Wisconsin, USA, June 2002.

[21] *Idemix anonymous credential system*, IBM Zurich Research Laboratory, http://www.zurich.ibm.com/security/idemix.

[22] "Privacy and identity management for europe (PRIME)," European RTD Integrated Project, http://www.prime-project.eu.org/.

[23] "Public key infrastructure (PKI)," National Institute of Standards and Technology, http://csrc.nist.gov/pki/.

[24] *RDF Vocabulary Description Language (RDFS)*, World Wide Web Consortium, http://www.w3.org/TR/rdf-schema/.

[25] *Semantic Web*, World Wide Web Consortium, http://www.w3.org/2001/sw/.

[26] L. Wang, D. Wijesekera, and S. Jajodia, "A logic-based framework for attribute based access control," in *Proc. of the 2004 ACM Workshop on Formal Methods in Security Engineering*, Washington DC, USA, October 2004.

[27] D. Wijesekera and S. Jajodia, "A propositional policy algebra for access control," *ACM Transactions on Information and System Security*, vol. 6, no. 2, pp. 286–325, May 2003.

[28] *Web Ontology Language (OWL)*, World Wide Web Consortium, http://www.w3.org/2004/OWL/.

[29] T. Yu, N. Li, and A. Anton, "A formal semantics for p3p," in *Proc. of the ACM Workshop on Secure Web Services*, Washington, DC, USA, October 2004.

[30] T. Yu, M. Winslett, and K.E. Seamons, "Supporting structured credentials and sensitive policies trough interoperable strategies for automated trust," *ACM Transactions on Information and System Security (TISSEC)*, vol. 6, no. 1, pp. 1–42, February 2003.

[31] T. Yu, M. Winslett and K.E. Seamons, "Prunes: Am efficient and complete strategy for automated trust negotiation over the internet," in *Proc. of the 7th ACM Conference on Computer and Communications Security*, Athens, Greece, November 2000.

[32] T. Yu, M. Winslett and K.E. Seamons, "Interoperable strategies in automated trust negotiation," in *Proc. of the 8th ACM Conference on Computer and Communications Security*, Philadelphia, PA, USA, November 2001.