Identity Management and Citizen Privacy

James B. D. Joshi
University of Pittsburgh, USA

Saubhagya R. Joshi
University of Pittsburgh, USA

Suroop M. Chandran
University of Pittsburgh, USA

INTRODUCTION

E-government systems aim to transition traditional paper-based systems to "paperless" digital information systems to automate and streamline government operations and services. This transformation to digital form raises daunting challenges related to protecting identity and privacy of the citizens. Electronic fraud and identity theft are among the biggest risks to an e-government system that may potentially undermine its success. The CSI/FBI 2005 (Gordon, Loeb, Lucyshyn, & Richardson, 2005) Computer Crime and Security Survey reports more than $30 million in losses attributed to theft of proprietary information and more than a $31 million dollars loss related to unauthorized accesses. According to the data collected by the Consumer Sentinel and Identity Theft Data Clearing House, identity theft accounts for almost 40% of the fraud complaints (FTC, 2005). It is estimated that billions of records are available in both private and government databases describing each citizen's finances, interests, and demographics. For instance, personal healthcare information about the diseases and health cases inflicting the general population are available in different places including insurance companies and pharmacies. While accessing such data is important for detecting epidemics and bio-terrorism, such accesses can easily encroach into citizen privacy. This demands a balancing act in dealing with issues related to privacy, accountability, national security, and/or good governance. Because of the heterogeneity of an e-government system, the task of protecting identity information as citizens interact with different sub-systems becomes exacerbated. Users typically may need to maintain multiple identities or complete anonymity while interacting with multiple interoperating systems raising severe privacy and identity management problems. For an e-government system to be reliable, and hence successfully deployed, the privacy and identity management issues need to be properly addressed and incorporated in its infrastructure design.

Privacy may be defined as *"the right of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated"* (Jajodia, 1998). Thus, unlike security, which is *organization-centric,* privacy is a *person-centric* concept and refers to the control that an individual has over the use of his or her personal information. One aspect of the privacy problem is the anonymity of the users, which is aimed at protecting the *identity* of the users. The *identity* of an individual is a collection of personal data associated with the individual that uniquely identifies him or her. Associated with each identity is a set of attribute-value pairs, also known as *credentials*, typically representing a user's qualifications and personal attributes including sensitive personal information such as name, age, and social security number. The capability to identify entities (subjects, objects, and resources) is essential in order to know what to protect from whom. Depending upon the context, a subset of the identity may be used to signify an individual. Such a "*partial identity*" is typically bound to the individual with a *pseudonym* (Köhntopp & Berthold, 2000), and may or may not uniquely identify the individual. Typically, multiplicity of identities for individual entities becomes necessary because of the requirements of anonymity, personal data protection, and controlled access to resources in multidomain e-government systems. Moreover, the notion of privacy and identity is inherently complex and may often be contradictory; furthermore, each stakeholder could have a different perspective on them. In large multidomain e-government system, identity management would typically aim towards providing mechanisms that ensure identity dependability to build and maintain trust and confidence between the interacting entities.

E-GOVERNMENT PRIVACY AND IDENTITY MANAGEMENT ISSUES

An e-government provides citizens with services that involve processing their personal information. With more and more personal information being stored in multimedia format (text, image, audio, and video), much heightened socio-psychological concerns with regards to privacy can be seen. Privacy vulnerabilities arise even if data is available in statistical or aggregate forms, or that allow personal information to be inferred. Furthermore, the fact that the government can carefully monitor every transaction and resource accesses made by a citizen can discourage citizen participation, thus affecting successful deployment of e-government systems. Compared to the general e-commerce environments, the e-government systems have increased obligation/responsibility towards maintaining privacy of the citizens (Dempsey, 2003).

Challenges

Several privacy and identity management related challenges exist that need to be addressed to ensure the success of an e-government environment.

Privacy Policy Specification

Development of a comprehensive privacy policy specification framework is a major challenge. There is a need for an expressive privacy policy specification and enforcement framework that supports flexible, fine-grained policy specification, and facilitates auditing and monitoring while maintaining scalability as well as cost-effectiveness. In particular, users' privacy preference could be very diverse, and could depend on the context and purpose of use of personal information. In general, natural language based specification would be immensely desirable but it can introduce severe problems in machine-readability, introducing difficulty in removing ambiguity and consistency in the policies. Formal, expressive languages are necessary and should be augmented with scalable correctness evaluation tools to facilitate proper administration and enforcement of privacy policies.

Active Content

One privacy challenge is introduced by Internet technologies such as Web browsers, whose vulnerabilities can be exploited to compromise privacy. For instance, cookies, the data stored on the client's machine and exchanged between the clients and the server to maintain connection information, can be used for the purpose of gathering user information. Use of executable content such as Java applets and ActiveX controls is another source of security vulnerability, that could be used to obtain personal information. Tools and techniques to ensure that such active content does not violate privacy requirements of users are crucial.

Multidomain Environment

Two characteristics of e-government services exacerbate privacy and identity management problems: (1) sharing of citizens' information among different government agencies and (2) differing privacy preferences of the citizens, heterogeneous privacy requirements of different e-government sub-domains, and potential use of identities to access different e-government sub-systems. While facilitating citizens' services, e-government domains may need to exchange users' information, including their identities and credentials. Further, the domains' privacy policies themselves may need to be integrated to provide transparency of the underlying privacy-preserving information sharing activities. Diverse privacy requirements of the citizens need to be addressed by the e-government infrastructure to ensure that all citizens feel safe to interact with the e-government systems. Further challenge is to facilitate unknown users to interact with e-government systems raising trust issues coupled with privacy requirements. A facility to establish trust between e-government systems and the users without unnecessarily divulging sensitive information is essential. To benefit from different applications one often requires multiple identities, which introduce multiple risks of exposure and fraud. If a user has to authenticate each time he accesses a different e-government sub-systems, it will involve multiple risks of identity theft.

Anonymity

Users often prefer anonymity during online transactions particularly when his or her activities can reveal sensitive information about him or her. A partial identity that cannot be used to uniquely identify an individual provides a degree of anonymity. Anonymity does not imply that no information is released; instead, the released information should not reveal the actual identity of a user (Damiani, di Vimercati, & Samarati, 2003).For example, a patient may collect/order medicine from a pharmacist anonymously. Here, the pharmacist should be able to associate the patient's partial identity to the prescription. Some interactions, however, cannot be conducted anonymously, for example when a doctor diagnoses a patient, his or her identity must be uniquely verified and his or her health records must be accessed.