

An Architecture for Identity Management

Brian R. Richardson and Jim Greer

Abstract— Personalization of on-line content by on-line businesses can improve a user’s experience and increase a business’s chance of making a sale, but with stricter privacy legislation and Internet users’ increasing concerns about privacy, businesses need to ensure they do not violate laws or frighten away potential customers. This paper describes the design of the proposed Identity Management Architecture (IMA). The IMA system allows users to decide on a per business basis what personal information is provided, gives users greater access to their personal information held by on-line businesses, and does not rely on a trusted third-party for management of personal information.

Index Terms— Identity, Privacy, .NET, PIPEDA.

I. INTRODUCTION

IN order to complete any commercial transaction on-line, people must provide personal information such as their name, address, phone number, email, credit card number, etc. On-line businesses often record more information than is actually needed to process a transaction. Some businesses monitor and record what types of products are bought or even the customers browsing patterns. This is done to form a detailed profile that will allow the business to target a customer with future advertising of products more closely related to his or her interests. As a result, businesses may be inadvertently in violation of privacy law and customers may be unaware of the extent to which their personal data is being stored or used. Individuals often try to counter such actions by supplying false or misleading data in an attempt to conceal their identities. Thus, the following three factors formed the basis of our approach:

- Legislation: Canada’s Personal Information Protection and Electronic Documents Act.
- Personal Concerns: The increasing concerns of Internet users about what information on-line businesses record.
- Tool Support: The lack of an available privacy tool that allows for management of multiple identities.

Even though there are a wide range of privacy tools available, there is currently no tool for users to manage what information they have provided to an on-line business they regularly visit. This paper presents the design of the Identity Management Architecture (IMA) that allows for several

privacy issues to be addressed. This is accomplished by providing Internet users with more control over their personal information while interacting with on-line businesses. This is achieved with four main features:

- Providing for the creation and management of multiple discrete personal identities.
- Allowing users to restrict and manage the access that businesses have to identifying information.
- Providing users with the ability to request from a business what personal information is stored.
- Providing businesses with a simple means of answering such requests.

The IMA system is designed around two main components: the IMA Manager, which is the client application, and the IMA Web Service, which is the web service deployed by participating on-line businesses. Each business that wishes to participate in the IMA approach must follow the standard defined for the IMA Web Service, must implement this service, and deploy it on its web site. Through this service, all interactions with the IMA Manager application are handled.

One of the key features of the IMA system, which is not offered by other personal information management tools, is the ability to create and manage multiple identities from within a single user account. The use of multiple identities is more than just partitioning a user’s personal information a business will see, but also allowing people to interact with a business for more than one purpose. For example, if someone shops at an on-line computer parts store, sometimes for work purposes and other times for personal purposes, he or she may want to create a “Work” identity and a “Personal” identity. Creating separate identities allows someone to more easily manage these two separate relationships with a business. This may be beneficial to a business, especially one that personalizes web site content based on the interests of the user. If someone is using his or her “Personal” identity, the business may use the browsed products and recently purchased products to make suggestions to the user about other products that may be of interest. If this same person visits the business’s web site at another time using the “Work” identity, the business will be better able to tailor content more towards the interests associated with this identity.

II. BACKGROUND

This section discusses the current state of Internet privacy in relation to the use and management of personal information. This section also provides a brief background on current Canadian privacy legislation and discusses how this

Manuscript received August 20, 2004.

B. R. Richardson is with the ARIES Laboratory, Department of Computer Science, University of Saskatchewan, Canada (e-mail: Brian.Richardson@usask.ca).

J. Greer is with the ARIES Laboratory, Department of Computer Science, University of Saskatchewan, Saskatoon, Canada (e-mail: greer@cs.usask.ca).

legislation restricts how a business may use a customer's personal information.

A. Privacy Legislation

The Canadian Personal Information Protection and Electronic Documents Act (PIPEDA) defines what personal information a business can track about a customer and how the business may use that information [8]. Information such as name, age, weight, height, income, purchases, spending habits, education, home address, and phone number are just a few examples of the types of information covered. This act only applies to information that is collected about an identifiable individual [9].

In January of 2004 PIPEDA was extended to apply to the handling of personal information gathered in any commercial transaction [9]. It is the third phase of this act that worries many businesses. The act puts restrictions on how information is obtained (i.e., consent must be obtained before information is gathered), and what types of information businesses are allowed to record.

B. Legislation Compliance

With the introduction of PIPEDA more restrictions have been placed on how businesses obtain consent when asking a user for personal information. Consent can no longer be obtained by just a general statement asking for a user's personal information, but must be more specific [11]. Businesses can no longer use difficult to understand legal statements when asking for consent, but must explain clearly why the information is requested and for what purpose it will be used [11].

Under PIPEDA a business must disclose to users, upon request, what personal information the business has about them. In order to be able to process these requests a business must ensure it has appointed someone within the organization as a privacy officer who is in charge of fielding requests from customers and also ensuring that the business's handling of personal information is in compliance with PIPEDA [11]. Some companies will need to hire additional employees to deal strictly with customer requests for personal information.

C. Privacy of Internet Users

On-line businesses already understand the benefits of gathering information about visitors to their web sites since it allows them to get a better understanding of visitors' interests, and thus improving the chances of turning visitors into customers. This has resulted in growing concern among Internet users about what information a business is tracking about them. For example, a survey conducted on Internet users showed several concerns that would cause a visitor to a web site to be less likely to divulge personal information [2].

Rated as highly important by 69% of respondents:

- Is the information used in an identifiable way
- Type of information collected
- The purpose for which this information is being collected

Rated as very important by 62% of respondents:

- Whether a site is run by a trusted company or organization
- Whether a site will allow a visitor to access information recorded about them
- Whether a site will remove someone from a mailing list upon request

For on-line businesses, getting Internet users to visit their web site and be willing to disclose personal information is crucial to survival. After all, if a visitor to a business's web site does not trust how that business handles personal information and is therefore not willing to disclose personal information, then a transaction such as purchasing a product or subscribing to a service will be much less likely to occur. Since most users are not aware of how businesses gather and use personal information, some businesses believe they don't need to respect the privacy preferences of users. This is a dangerous attitude to have towards users' privacy. It is also illegal in Canada.

D. Privacy Technologies

In recent years privacy technologies have been developed to help users to manage and protect their personal information. In this section several technologies that have been designed to help protect users' privacy will be discussed and compared.

1) Platform for Privacy Preferences

The Platform for Privacy Preferences (P3P) standard was developed by the World Wide Web Consortium (W3C) to allow Internet users to define privacy preferences and have these preferences compared to a business's privacy policy automatically when visiting the web site [13]. Since privacy policies are often long and difficult to understand, P3P offers a simple interface that allows Internet users to define their preferences once, and then have these preferences automatically checked at each participating P3P web site they visit.

P3P relies on informing visitors to a web site about the contents of a business's privacy policy. The problem with relying on this as the sole form of privacy protection for Internet users is that they may not always change their actions based on the privacy differences shown between their preferences and the business's privacy policy [10]. Although a user may be informed about the ways a business uses personal information, the user may not alter his or her behaviour in terms of browsing activity or when asked questions while shopping, especially when interested in a product or service [10]. The IMA system is not meant as a replacement for P3P, but rather as a tool for Internet users who decide to establish a relationship with an online business by providing an identity, to stay informed about which online businesses have their personal information, and to know what information is stored about them at each business.

2) TRUSTe

TRUSTe is a non-profit organization that provides privacy seals to web sites that follow a set of privacy specifications [12]. When a user sees a TRUSTe seal on a business's web

site it provides assurances that the privacy policy posted accurately represents what personal information the business gathers and how it is handled.

In the last few years TRUSTe has come under fire for problems that have occurred. One such problem was Yahoo's decision to change the way it would use personal information, while including information collected previously under a different privacy policy bearing the TRUSTe seal [4]. Similar incidents have also occurred with Real Networks and eBay. This type of misuse of the TRUSTe seal, and the organization's unwillingness or inability to do something about it has brought into question the validity of this TRUSTe seal program. Seth Ross from PC Guardian had this to say:

"A trustmark does more harm than good by creating an illusion of privacy where none exists. A meaningless logo may induce people to make information disclosures that they would otherwise avoid" [4].

3) *Privacy Critics*

A privacy critic is an agent that provides the user with warnings of potential privacy risks to the actions the user is attempting to take [1]. It is then up to the user to decide either to take the suggestion, or ignore it. Unlike other privacy technologies that block certain information, a privacy critic takes no action on behalf of the user, but rather just provides the user with warnings without stopping the user from making a bad decision [1]. If a user has decided not to listen to the privacy critic and releases information, there is no way for the user to get that information back. The IMA system does not provide any of the features of a privacy critic, but it does allow a user to view information given to a participating business and update that information.

4) *Privacy Incorporated Software Agent*

With the Privacy Incorporated Software Agent (PISA), proposed by Dr. John J. Borking, the vice-president of the Dutch Data Protection Authority, the goal is to address areas of potential privacy risks by building a Privacy Enhancing Technology Agent [3]. The PISA agent allows other agents to perform tasks requested by the user yet ensures only the minimal personal information required for a task is provided and nothing more. The PISA agent would be placed either between the user and other agents to prevent those agents from possessing any more information than necessary, or it would be placed between the agents and all outside systems [3]. The downside of the PISA agent is that implementing the agent so that it would not accidentally divulge information to the wrong source and at the same time be compliant with privacy legislation is difficult.

5) *Privacy Policy Compliance System*

The Privacy Policy Compliance System (PPCS) proposed by G. Yee and L. Korba is based on the idea of allowing a web user to check how a business will use his or her personal information [14]. The proposed PPCS system allows consumers to see if their privacy preferences match those of the provider before their personal information is provided. At the same time PPCS allows businesses to be in compliance

with privacy legislation by allowing them to obtain consent by disclosing to a consumer how they plan to use personal information. The IMA system is different from PPCS in that IMA does not involve privacy policies that explain users' preferences. Instead it deals strictly with the personal information and allows users to manage their personal information across multiple businesses, without the use of a central source (i.e., third party service).

6) *Enterprise Privacy Architecture*

The Enterprise Privacy Architecture (EPA) offered by IBM provides a system for businesses to evaluate their existing uses of customer information and locate potential privacy risks that may exist [6]. By evaluating a business's business processes in terms of uses of personal information, EPA allows a business to see how a customer's information is being used and shows how to place protection on that information to ensure any inappropriate disclosure does not occur [6]. While the services offered by EPA allow a business to evaluate how information is managed in their business processes, the IMA system looks at how information is managed between the customer and the business.

7) *Liberty Alliance Project*

The Liberty Alliance Project is based on the idea of allowing users to connect multiple sets of personal information, that exist across several on-line businesses, into one easy to manage identity. This allows for the convenience of a single sign-on service, as well as easier management of personal information across multiple businesses [5]. The Liberty Alliance architecture allows an Internet user to store his or her personal information with a trusted business. When the user needs to access a service provided by another business, which is part of the same group of associated businesses, the user's chosen trusted business provides authentication of the user, as well as the user's identity information [5].

What makes this system architecture unique, is that rather than relying on a trusted third party, such as .NET Passport, to provide a user's identity to each business the user accesses, it allows the user to have a business he or she trusts store and pass identity information from one business to another, which is part of the same group of associated businesses [5]. One downside to this design is that identity management across multiple businesses is restricted to the set of businesses that have formed associations with each other. Even if two businesses are participating in the Liberty Alliance system, if these two businesses do not have an identity sharing relationship between them, the user will not be able to use the same identity at both businesses.

8) *Microsoft .NET Passport*

The Passport system created and managed by Microsoft provides a single sign-in service that allows Internet users to have one account for access to all Passport participating web sites [7]. The Passport system handles authentication of a user by having the sign-in page on each participating web site authenticate the user by contacting the Passport system [7].

No personal information about a user is ever provided to a

business unless the user gives consent. Consent is obtained when a user signs into a business's web site using Passport [7]. The Passport system stores only the information a user provides on sign-up. This is the only information that is provided to a business when a user accesses the business's web site. All information gathered about the user while at a business's web site is stored only by the business. It is left up to the business to determine what it will do with the information gathered (i.e., share the user's personal information with other businesses).

A user may update his or her email address, phone number, or postal code through the Passport home web site and have that information propagated to Passport-approved businesses. This does not give the user access, however, to the personal information a business participating in the Passport system may have gathered. Instead, it is up to the user to read the privacy policy of each partner business's web site to determine what information the business might collect and maintain. If the user had access to the information a business has stored (i.e., profile), this would give the user much more control over personal information.

One of the main features of the Passport system is the single sign-in service. Although this is convenient, it does not allow for any sort of management of multiple identities. Passport does not allow the creation of multiple identities (i.e., more than one set of personal information) to be associated with a single account to allow a user to choose on a per business basis what personal information a business receives. Although it is true that this could be accomplished by creating multiple Passport accounts, this defeats the purpose of a single sign-in service, since this approach would require a user to manage several Passport accounts, to remember multiple usernames and passwords, and to remember which account had been used at each business.

III. AN IDENTITY MANAGEMENT ARCHITECTURE

Personal information management systems such as .NET Passport or Liberty Alliance rely on either a third-party or another business as a form of storing and transferring someone's personal information. One goal of the Identity Management Architecture (IMA) is to avoid any use of a third party system and to not require businesses to communicate with each other for the purpose of providing a customer's information.

A. System Architecture

The IMA system has two main components:

- 1) IMA Manager (Client): An application that attaches to the user's web browser and handles the management of all user identities and web browsing history.
- 2) IMA Web Service (Business): A web service that each participating business provides to allow users of the IMA Manager to send and receive identity information.

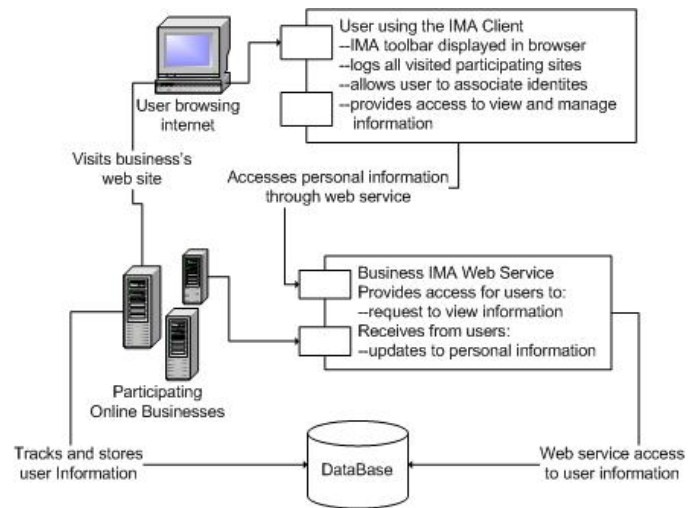


Fig. 1. General overview of the architecture of the IMA System. Shown here is an Internet user using the IMA Client to connect to a business that is participating in the IMA system by implementing the IMA Web Service.

The IMA Manager allows a user to contact a business's IMA Web Service to make a request to see what information the business currently has stored in the user's profile. A user may correct or remove information. If the information to be changed is contained in an identity, a user may modify the identity information stored in the IMA Manager and the application will automatically forward updates to all businesses associated with this identity. The user may associate another identity with a business at any time; this will be used for future visits to that business.

As shown in Fig. 1, the IMA Manager runs on the user's local web browser. The Manager receives from the web browser the URL of each site the user visits while on-line. It then checks to see if the business is participating in IMA by attempting to contact the IMA web service that all participating sites are required to make available. If this business is not participating, then this is shown in the IMA Manager's display. However, if the web service is available, this URL is stored and the service is contacted. The first communication with a business is the transmission of the user's preferred on-line identity. From this point on, each time the user returns to this web site the user will be identified by this identity. This will allow the business to associate information such as the products browsed with this identity. The business will then use this information to determine the user's interests.

B. Passport vs. IMA

What makes the IMA System's architecture different from Passport is the lack of a third party participant. As a result there are several differences in how personal information is handled in the absence of a third party. In Passport, a user's personal information is stored in two locations: in the Passport system and at each participating on-line business. Fig. 2 shows the typical flow of personal information in Passport.

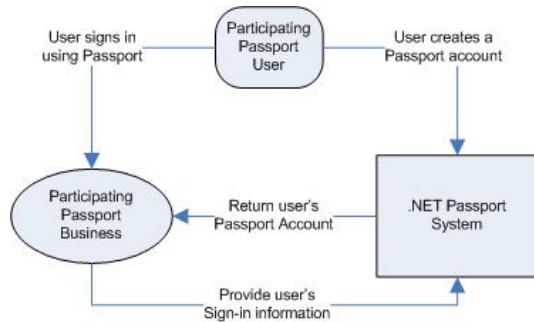


Fig. 2. Information Flow in .NET Passport. The .NET Passport third party system is involved in both sides of the relationship. The user must provide information to Passport by creating an account, and the business must retrieve the user's information from the Passport system.

In the IMA system a user's personal information is stored in two places: on the user's local computer and at each on-line business the user has provided with an identity. This is different from Passport's approach since the IMA system does not require a user to provide personal information to a third party in order to participate. Fig. 3 shows the typical flow of personal information in the IMA system.

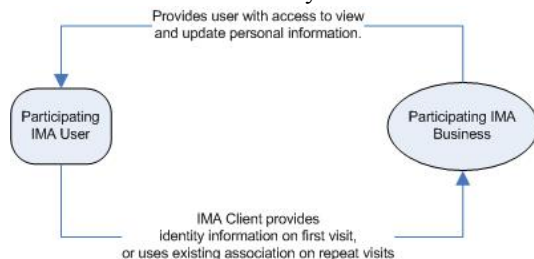


Fig. 3. Information Flow in the IMA System. Third party involvement is removed in the IMA system. As shown here, information is only passed between the participating user and the participating business.

Only that information contained in a single identity is sent to a business. This is unlike Passport, which uses only one set of personal information for a user's account and provides it to each partner business. If a user wishes to have more than one identity, multiple Passport accounts must be created which defeats the purpose of a single sign-in service. The IMA system is designed to provide management of multiple identities on behalf of the user.

One of the key features of the IMA system is that it provides users with the option to directly request what information a business has about them. The only information the user has access to view and update with Passport is the personal information that was entered when the account was created. This does not provide users with any access to additional information a business has about them and it does not provide any benefit to businesses in terms of compliance with information disclosure requirements placed on businesses by PIPEDA.

IV. IMPLEMENTATION

In order to demonstrate the design of this system, a small prototype implementation has been built. Each component of the system has been implemented using the .NET framework. There were four components that needed to be implemented for this project:

- IMA Toolbar
- IMA Manager
- IMA Web Service
- Example participating web site

The client side of the IMA system consists of the IMA Toolbar and IMA Manager. The IMA Toolbar was implemented as a toolbar for Internet Explorer. This toolbar handles basic functions such as showing the user if the current site they are visiting is participating, what identity is associated with that site, and allowing the user to login/logout. The IMA Manager application provides the user with management of identities, modifying settings, listing of current identity-business associations, and requesting of personal information.

The business side of the IMA system consists of the IMA Web Service and the participating business's web site. The IMA Web Service allows the IMA Manager application to communicate with a participating business. Identity information, as well as customer requests for information disclosure, is transferred to and from the IMA Web Service as an XML document. Each business that participates in the IMA system must implement the IMA Web Service. The business decides how to implement this web service and connect it to the user information stored in the business's database. All that is required is that a participating business follows the IMA Web Service interface and all required methods are made available to the IMA Manager. The base set of web service methods required by the IMA Manager include: authenticating a user, retrieving an identity, retrieving a profile report, updating an identity, updating a profile report, and adding of new identities for first time visitors.

V. IDENTITY MANAGEMENT

A. Identities

The IMA system allows users to set different identities, for example for personal, school, and work related activities. Having the ability to easily separate these identities allows web users to ensure that the information recorded about them at certain web sites is based on the identity they have assigned for that site. Users are able to select an identity to be used for each participating web site visited, but also have the ability to switch to another identity when their reason for using the Internet changes. The IMA tool also allows users to view a list of recently viewed web sites and change the identity associated with any site at a later time. This change results in the information for the alternate identity being immediately forwarded to the business. This alternate identity is then used for subsequent visits to this web site.

When shopping frequently at a number of different on-line businesses, it is convenient for customers to have each business store some of their personal "account" information such as name and mailing address and connect to that account with a username/password. However, when personal "account" information changes, the user will be required to update this information manually for each business. The IMA

tool allows a user to update an identity and have those changes automatically forwarded to all businesses with which this identity has been associated.

VI. LIMITATIONS

Personal information in the IMA system is always stored in two places; at the participating businesses the user has visited, and on the user's computer. Any responsible business will have security measures in place to protect stored personal information. However, the information stored on an individual's computer may be at risk of being compromised. In order to ensure no one other than the owner of the account has access to the information stored in it; the account is stored in an encrypted, password-protected file.

Even if an IMA participating business has excellent security measures in place to ensure each user's personal information is protected, there may still be increased security risks since this architecture promotes a more open exchange of personal information between users and businesses. If someone tried to make a request to a business for personal information while posing as another user, this could lead to the business disclosing a user's personal information to the wrong person. Since the only person with knowledge of the accounts username and password is the owner, unless this information is compromised, no one else will have access to it.

Another potential threat to the IMA system is that a business may not publicly state that it is participating yet may secretly deploy an IMA Web Service. Through this service the business may attempt to receive personal information from visitors to the business's web site who are using the IMA Manager. The result is that a business may try to collect personal information from users, yet not provide the required access for users to personal information stored, or even notify users that the business is receiving information from the IMA Manager. The IMA Manager requires that the user must select and send an identity to any new business visited. The IMA Manager makes every effort to prevent a business from receiving identity information without the user having full knowledge of this action taking place.

VII. CONCLUSIONS AND FUTURE WORK

Even with the tools currently available for Internet users to protect their privacy on-line, there are few useful tools to protect privacy through personal information management. Single sign-in systems such as Passport provide some basic management of personal information, but lack flexibility and the ability for users to create multiple identities from within a single user account. These systems also require people to disclose their personal information to a third party and do not allow users to keep track of businesses that have been given their personal information. These areas of on-line privacy, where current technologies are unable to offer more control for users over their personal information, are where the IMA system attempts to fill this gap.

A larger more complete implementation of the IMA system will need to be built. The first issue that will need to be addressed is security. This will include determining the most secure way for a business to identify an IMA system user, and also ensuring that only businesses authorized to receive the user's information have access to it. The second issue that will need to be addressed is allowing users to access their accounts from multiple locations. Since the IMA system does not rely on a third party system for account access and storage, the IMA system will require a different approach. One option might be to store the encrypted account file with a single business the user trusts. The information can then be retrieved and used from other locations.

Once issues of security and usability have been addressed further, a larger scale study of the IMA system's design would need to take place. At the moment, the current design of the IMA system mainly looks at web sites that manage user profiles. Future work may include looking at ways to extend this system to such applications as education environments. Also a topic for a future paper may be to look at how existing systems, such as .NET Passport and Liberty Alliance, may be modified to incorporate features such as disclosure on demand and the use of multiple identities.

REFERENCES

- [1] M. Ackerman and L. Cranor, "Privacy critics: UI components to safeguard users' privacy" Conference on Human Factors in Computing Systems (CHI '99), ACM Press, 1999, pp. 258-259.
- [2] M. Ackerman, L. Cranor, and J. Reagle, "Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences" Proceeding of the ACM Conference on Electronic Commerce, 1999, pp. 1-8.
- [3] J. Borking, "Privacy Incorporated Software Agent (PISA): Proposal for building a privacy guardian for the electronic age" 2001; [1] http://www.pet-pisa.nl/dscgi/ds.py/Get/File-237/Report_PISA_Architecture_12122002-v7.pdf.
- [4] P. Boutin, "Just How Trusty Is Truste?" Apr. 2002; <http://www.wired.com/news/exec/0,1370,51624,00.html>.
- [5] S. Cantor et al., "Liberty ID-FF Architecture Overview" 2003; <http://www.projectliberty.org/specs/liberty-idff-arch-overview-v1.2.pdf>.
- [6] IBM, "Enterprise privacy architecture (EPA)" 2004; <http://www.zurich.ibm.com/pri/projects/epa.html>.
- [7] Microsoft, "Microsoft .NET Passport Privacy Statement" 2003; <http://www.projectliberty.org/specs/liberty-idff-arch-overview-v1.2.pdf>.
- [8] Privacy Commissioner of Canada, "A Guide for Canadians" 2001; http://www.privcom.gc.ca/information/02_05_d_08_e.asp.
- [9] Privacy Commissioner of Canada. The Personal Information and Electronic Documents Act: A Primer On its Privacy Provisions. Available online at <http://e-com.ic.gc.ca/english/privacy/632d30.html>.
- [10] S. Spiekermann, J. Grossklags, and B. Berendt, "E-privacy in 2nd Generation ECommerce: Privacy Preferences versus actual Behavior In Electronic Commerce (EC'01)" 3rd ACM Conference on Electronic Commerce, vol. EC '01, 2002, p. 38--47.
- [11] D. Ticoll, "Companies ignore privacy laws at their peril" Aug. 2003; <http://www.globetechnology.com/servlet/ArticleNews/TPStory/LAC/20030814/TWTICO14/TPTechInvestor/>.
- [12] Truste, "Frequently Asked Questions" 2001; http://www.truste.org/consumers/users_faqs.html.
- [13] World Wide Web Consortium (W3C), "P3P 1.0: A New Standard in Online Privacy" 2003; <http://www.w3.org/P3P/brochure.html>.
- [14] Yee, G. and Korba, L. Privacy Policy Compliance for Web Services. In Proceedings of the IEEE International Conference on Web Services. San Diego, California, USA, July 6-9, 2004.