Inf Syst Front (2007) 9:85–101
DOI 10.1007/s10796-006-9019-y

End-to-end privacy control in service outsourcing of human intensive processes: A multi-layered Web service integration approach

Patrick C. K. Hung · Dickson K. W. Chiu · W. W. Fung · William K. Cheung · Raymond Wong · Samuel P. M. Choi · Eleanna Kafeza · James Kwok · Joshua C. C. Pun · Vivying S. Y. Cheng

Published online: 22 December 2006

Abstract With the recent adoption of service outsourcing, there have been increasing general demands and concerns for privacy control, in addition to basic requirement of integration. The traditional practice of a bulk transmission of the customers' information to an external service provider is no longer adequate, especially in the finance and healthcare sectors. From our consultancy experience, application-to-application privacy protection technologies at the middleware layer alone are also inadequate to solve this problem, particularly when human service providers are heavily involved in the outsourced process. Therefore, we propose a layered architecture and a development methodology for enforcing end-to-end privacy control policies of enterprises over the export of personal information. We illustrate how Web services, augmented with updated privacy facilities such as Service Level Agreement (SLA), Platform for Privacy Preferences Project (P3P), and the P3P Preference Exchange Language (APPEL), can provide a suitable interoperation platform for service outsourcing. We further develop a conceptual model and an interaction protocol to send only the required part of a customer's record at a time. We illustrate our approach for end-to-end privacy control in service outsourcing with a

A preliminary version of this paper appears in the 7th International Conference of Electronic Commerce (ICEC2005). We have generalized and extended our approach to service outsourcing of human intensive processes. We also found P3P more appropriate in general.

P. C. K. Hung (✉)
Faculty of Business and Information Technology,
University of Ontario Institute of Technology, Osahwa, Canada
e-mail: patrick.hung@uoit.ca

D. K. W. Chiu
Dickson Computer Systems, Kowloon, Hong Kong
e-mail: dicksonchiu@ieee.org

W. W. Fung · J. Kwok · J. C. C. Pun · V. S. Y. Cheng
Department of Computer Science, The Hong Kong University of Science and Technology,
Kowloon, Hong Kong

W. W. Fung
e-mail: wwfung@cs.ust.hk

J. Kwok
e-mail: jamesk@cs.ust.hk

J. C. C. Pun
e-mail: punjcc@cs.ust.hk

V. S. Y. Cheng
e-mail: vivying@cs.ust.hk

W. K. Cheung
Department of Computing,
Hong Kong Baptist University,
Kowloon, Hong Kong
e-mail: william@comp.hkbu.edu.hk

R. Wong
School of Computer Science and Engineering,
University of New South Wales,
Sydney, Australia
e-mail: wong@cse.unsw.edu.au

S. P. M. Choi
School of Business and Administration,
The Open University of Hong Kong,
Kowloon, Hong Kong
e-mail: schoi@ouhk.edu.hk

E. Kafeza
Department of Marketing and Communications,
Athens University of Economics and Business,
Athens, Greece
e-mail: kafeza@aueb.gr

tele-marketing case study and show how the software of the outsourced call center can be integrated effectively with the Web services of a bank to protect privacy.

Keywords Web service integration · Privacy policies · Need-to-know principle · Layered architecture · SLA · P3P · APPEL

1 Introduction

Service computing has become a cross-discipline that covers the science and technology of bridging the gap between business services and IT services. Its goal is to enable IT and computing technologies to perform business more effectively and efficiently. One notable essence of service computing is that enterprises acquire functionalities from external partners as needed (Zhang, Li, & Lam, 2004); that means services are outsourced in this emerging information system frontier. However, service outsourcing activities have been widely adopted long before the current sophistication of service-oriented architectures. Power and Trope (2005) has recently reported this trend, which has raised a general concern of privacy issues.

Marketing is a strategy for selling products more effectively and efficiently. This includes sales promotion strategies for making consumers recognize a product's existence and persuading them to take purchase actions, circulation strategies for efficiently delivering the desired product, and continuation strategies such as after-sales service and claim processing (Diamond, 2000). Under the current complex commercial world, enterprise cooperation that involves the transfer of customer data from one company to a partner service company is common, especially during the outsourcing of services. The traditional approach of a bulk transmission of customers' information to a marketing company cannot meet the privacy demands, especially in the finance and healthcare businesses. To illustrate the problem and solution, we present a case study in the outsourcing of tele-marketing activities in a bank based on our consultation experience. A bank often passes its customer's data onto a marketing partner in a traditional batch mode. Only can the security of data transport be guaranteed, with say the Public Key Infrastructure (PKI). This can be achieved in a number of ways. One common way is to use email with Pretty Good Privacy (PGP). Another possibility is to establish a secure networked connection (such as an extranet) from the bank to the marketing company, where the data is encrypted before transmission. The information will be under the custody of the marketing partner after the data transmission. The bank can only rely indirectly on the contract with the marketing partner for further privacy protection. So, there is an obvious gap between common practices and the current privacy requirements.

As businesses often overlook online privacy and protection of personal data, Constantinides (2002) reports the concerns of millions of consumers about the way their personal information is being used. For example, credit card users' information has to be passed to the marketing company for the ease of the calling case confirmation. Further customer behavior is usually impulsive and the desire to buy will be cool down soon afterwards. Therefore, the management would like the purchase deal confirmed during the marketing process, rather than passing the list of interested customers back to the bank staff to call again to complete the deal. With the current approach, the credit card information for those customers who are not interested in the product would have unnecessary information passed onto the marketing company. No one knows who would possibly misuse this information and therefore the current approach needs to be improved.

On the other hand, contemporary privacy technologies for service-oriented architectures, such as the Platform for Privacy Preferences (P3P) Project, focus mainly on program-to-program interactions, particularly on the middleware layer. These are inadequate to address the problem because human service personnel are often heavily involved in outsourced service processes, such as tele-marketing, which involves various call center personnel. To the best of our knowledge, methodologies for the design of applications in conformance to such privacy requirements have not been reported.

In order to address this problem, we present in this paper a layered architecture and a development methodology for end-to-end privacy control over the export of each individual customer's record and fields through a Web services platform, according to the corresponding enterprise's privacy control policies for service outsourcing. We further develop a conceptual model and an interaction protocol to send only partial customer records at a given time over the Internet. We illustrate our approach for end-to-end privacy control with a tele-marketing case study from our consultation experience and show how the Web services of an enterprise, augmented with updated privacy facilities, can be integrated effectively with the software of an outsourced partner. The remainder of this paper is organized as follows: Section 2 introduces the current problem in end-to-end privacy control in service outsourcing. Section 3 compares related work. In Section 4, we present an overview of our development methodology. Section 5 details a technical approach to enforce the end-to-end privacy control based on our case study. Section 6 discusses about the technical challenges and infrastructure for the implementation, whereas Section 7 discusses the summary and future research directions.

2 Privacy issues in service outsourcing

Security enables privacy, but privacy is a much broader concept than security. Privacy is a state or condition of limited access to a person (Schoeman, 1984). In particular, information privacy relates to an individual's right to determine how, when, and to what extent one's personal information may be released to another person or to an organization (Leino-Kilpi et al., 2001). One can imagine that information privacy is usually concerned with the confidentiality of sensitive information. One of the most significant objectives of enforcing privacy policies is to protect personal identifiable information (PII). Fischer-Hubner (2001) further points out that threat to information privacy may come from insiders and from the outsiders in each organization. Similarly an article "Web Firms Choose Profit Over Privacy" from Washington Post (published on July 1, 2003) states that almost all companies promise not to sell consumer data, but many do not mention that such information is often *rented*. This means that the list owner will not release the data to an outside marketer, but it will send messages to the list on the outsider's behalf. Consumers or even organizations often indicate that the privacy of their sensitive information is their foremost concern regarding the e-commerce activities on the Internet. Privacy control is usually not concerned with individual subjects. Subjects release their data to the custody of an enterprise while consenting to a set of purposes for which the data may be used.

In general, privacy policies describe an organization's data practices, what information they collect from individuals (e.g., consumers), and what (e.g., purposes) they do with it. A privacy policy is a set of rules and practices that specify or regulate how a system or organization provides user control for the personal information collected, managed, and used by an organization. A privacy policy rests on a formalized or semi-formalized security model that is necessary but not sufficient to provide privacy protections. In principle, a privacy policy is based on applicable legislated requirements that are interpreted through a privacy impact assessment, some form of privacy risk assessment and analysis, and a set of privacy assertions that refer to rules and practices to regulate how personally identifiable or sensitive information is managed and protected. This generally includes the full spectrum of organizational controls of information under security but also covers privacy issues such as consent management, unlinkability, unobservability, pseudonymity, and anonymity of data.

In the US, the Privacy Act of 1974 requires federal agencies to grant individuals access to their identifiable records maintained by the agency, ensure the accuracy and timeliness of existing information, and limit the collection of unnecessary information and the disclosure of identifiable information to third parties (Davis, 2000). A recent survey (Hinde, 2002) reveals ongoing concerns of bank officers, mostly procedural, about how to handle the anticipated privacy regulations of the US Gramm-Leach-Bliley (GLB) Act, which requires financial institutions to regularly communicate privacy policies to customers and provide adequate opportunities for "opting-out" of personal information disclosure to non-affiliated third parties. For example, based on the GLB Act, when a bank selects information from customers, it has to provide the following:

- Notice: The data collector has to disclose the informational practices before collecting the information. This means the bank is obliged to notify the customers about its intention to use this information for marketing purposes.
- Choice: The customers have a choice on how the information is used. For example, the customers can choose to participate or not in a certain marketing campaign.
- Access: the customers should have the access to their information given and could write to the bank for update and/or deletion, as well as query on who is using the data and for what purposes.
- Security: Data collectors (that is the bank) should make sure that the data are secure from unauthorized use, by taking all security precautions when the data is transmitted, stored, and copied.

One can imagine that this activity is very important to the bank because there is no tolerance in breach of customer information privacy. For example, an incident happened in late 2004 (CIBC, 2005). A Canadian bank called CIBC misdirected facsimiles that contained the personal information of its customers to a company in the US and another in Dorval, Quebec over a number of years. However, CIBC did not appropriately recover customer personal information and this was deeply disturbing to the Office of the Privacy Commissioner of Canada.

In summary, a customer should be fully aware of how one's personal information is being used. So, the bank should also provide its customers with a list of the partner companies to which marketing activities may be outsourced and the customer should have the option to trust these companies or not. Alternatively, the bank could also ask the customer whether one trusts the companies that the bank trusts, i.e., delegate one's trust to the bank. In this case, the bank is acting like a certificate authority in the sense that it has investigated the outsource partner's policies and guarantees that it is trusted. Therefore, the bank should have checked at least the following: (a) the outsourcing company has the appropriate policies and technical background to verify different appropriate roles which have access to a different set of identifiable information, and (b) the data sent will be secured with respect to inside personnel as well as outsiders. This guarantees that no other

personnel besides the one specified by the role can access the data.

3 Technical background and related work

To introduce the technical background, we first present a centralized privacy access model here. According to the NIST draft standard (NIST, 2005), RBAC naturally fits with many service outsourcing applications. In RBAC, permissions are associated with roles, and users classified into appropriate roles thereby acquiring the roles' permissions. In addition, roles can be granted new permissions, and permissions can be revoked from roles as needed. The significant benefit of deploying RBAC is its flexibility to meet the changing needs of an organization. Cheng and Hung (2005) propose a RBAC model with privacy-based extension as shown in Fig. 1. When a request is subjected to access control, the core RBAC enhanced with a privacy-based extension (purpose, recipient, obligation, and retention) either grant or deny the permission according to the content of the request, a set of obligations, and a set of retention policies. Figure 1 presents an access control model of core RBAC with privacy-based extension. When a request arrives at the access control mechanism, the core RBAC enhanced with privacy-based extensions (purpose, recipient, obligation, and retention) either grant or deny it and returns a set of obligations and a set of retention policies. Referring to Ferraiolo, Kuhn, and Chandramouli (2003), the core RBAC model mainly includes the following entities:

- SUBJECTS, the set of subjects in the system
- USERS $\subseteq$ SUBJECTS, the set of human users in the system
- ROLES, the set of roles that describes the authority and responsibility on a member of a role
- OBJECTS, the set of objects in the system
- OPS, the set of operations that can be executed in the system
- PRMS $= 2^{\{\text{OPS}\times\text{OBJECTS}\}}$, the set of permissions that approve a particular operations to one or more objects in the system
- UA $\subseteq$ USERS $\times$ ROLES, a many-to-many mapping between users and roles (user-to-role assignment relation)
- PA $\subseteq$ PRMS $\times$ ROLES, a many-to-many mapping between permissions and roles (role-permission assignment relation)
- Assigned_users: (r:ROLES)$\rightarrow$2USERS, the mapping of role r onto a set of users. Formally: assigned_users$(r) \subseteq \{\forall_{i=1,2,\ldots,n} u_i \in \text{USERS} | (u_i, r) \in \text{UA}\}$
- Assigned_permissions: (r:ROLES)$\rightarrow 2^{\text{PRMS}}$, the mapping of role r onto a set of permissions. Formally: assigned_permissions$(r) \subseteq \{\forall_{i=1,2,\ldots,n} p_i \in \text{PRMS} | (p_i, r) \in \text{PA}\}$.
- Subject_user: (s:SUBJECTS)$\rightarrow$USERS, the mapping of subject s onto the subject's associated user.
- Subject_role: (s:SUBJECTS)$\rightarrow 2^{\text{ROLES}}$, the mapping of subject s onto a set of roles. Formally : subject_role$(s) \subseteq \{\forall_{i=1,2,\ldots,n} r_i \in \text{ROLES} | (\text{subject_user}(s), r_i) \in \text{UA}\}$

The following set of privacy-based entities (purpose, recipient, obligation, and retention) is added to the proposed core RBAC as shown in Fig. 1.

- PURPOSES, is the set of purposes that describes the user's purpose(s) of a request submission.
- RECIPIENTS $\subseteq$ SUBJECTS, is the set of recipients of the result generated by the set of collected object(s) such as analysis reports.
- OBLIGATIONS, is the set of obligations that may be taken after the decision of permission is made. In general, an obligation is opaque and is returned after the permission is granted. The obligations describe what promises a subject that must be made after gaining the permission.
- RETENTIONS, is the set of retention policies that are to be enforced in the object(s) in effect. Each data custodian may have its own retention policy to enforce the usage of datasets.

In RBAC, a subject can never have an active role that is not authorized for its user (Ferraiolo et al., 2003). With all the privacy-based extension (purposes, recipients, obligations and retentions), the role authorization in the core RBAC model (Ferraiolo et al., 2003) is revised as follows:

- DECISION:{ALLOW,DENY}, a decision for describing whether the access is granted or denied.
- OWNER $\subseteq$ SUBJECTS $\times$ OBJECT, a one-to-many mapping between subjects and objects. It is a set of tuples (s, o) where $s\in$SUBJECTS and $o\in$OBJECTS.

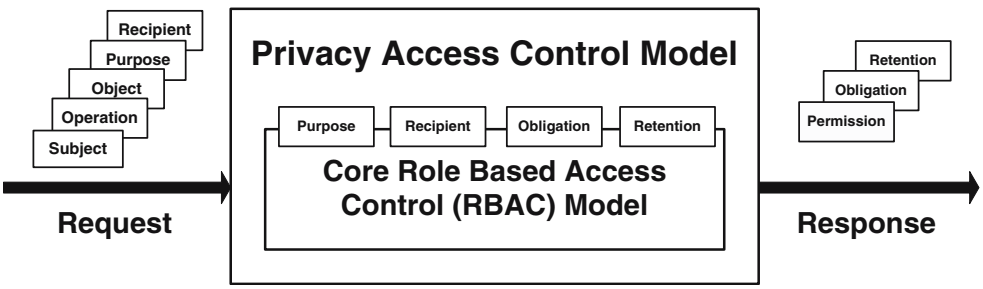


Fig. 1 An overview of privacy access control

- Access: SUBJECTS × OPS × OBJECTS × PURPOSES× RECIPIENTS → DECISION × OBLIGATIONS × RETENTIONS, the core part in determining whether a access is granted or denied with any obligations and retention policies according to the subject that invoke the access, operation of the access, the objects that the subject requesting, the purposes for this request, the recipients of the result generated by the set of collected objects. Formally: access(s, op, $\{o_1, o_2,\ldots, o_i\}$, $\{pp_1, pp_2,\ldots, pp_j\}$, $\{rp_1, rp_2,\ldots, rp_k\}$)=(ALLOW, $\{obl_1, obl_2,\ldots, obl_m\}$, $\{rt_1, rt_2,\ldots, rt_i\}$) if subject s can access any object in $\{o_1, o_2,\ldots, o_i\}$ using operation op for any purpose in $\{pp_1, pp_2,\ldots, pp_j\}$ with any recipient in $\{rp_1, rp_2,\ldots, rp_k\}$; (DENY, ∅, ∅) otherwise. If the access is granted, a set of obligations $\{obl_1, obl_2,\ldots, obl_m\}$ and also a set of retention policies $\{rt_1, rt_2,\ldots, rt_i\}$ for corresponding set of objects $\{o_1, o_2,\ldots, o_i\}$ are returned to subject s.
- Object_owner: OBJECTS→SUBJECTS, the mapping of an object to its owner. Formally: object_owner(o)= $\{s\in$ SUBJECTS$|(s,o)\in$OWNER$\}$
- Owner_object: SUBJECTS→2^{OBJECTS}, the mapping of a subject to all the objects the subject owns. Formally : $\text{owner\_object}(s) \subseteq \{\forall_{i=1,2,\ldots,n} o_i \subseteq \text{OBJECTS}|(s,o_i) \in \text{OWNER}.\}$.
- To restrict the one-to-many mapping between subjects and objects, we add the following rules: s_i, $s_j\in$SUBJECTS, $s_i\neq s_j$, obj∈OBJECTS, object_owner(obj)=$s_i\oplus$object_owner (obj)=s_i, where ⊕ means logical exclusive-or.

This approach assumes that a data provider (object owner) has the absolute right to determine whether a data requestor (subject) can access a certain data object or not. Referring to the tele-marketing example, the data provider is the bank and the data requestor is the marketing company. In this scenario, the marketing company requests the access of the credit card customer's information (object) from the bank in order to accomplish the outsourced activities.

Traditional business-to-business applications connect trading partners through a centralized architecture. A major drawback is that setting up an additional connection with another trading partner is costly and time consuming. Web services are now becoming an increasingly popular technology for supporting business-to-business (B2B) applications. Web services technology is based on a set of eXtended Markup Language (XML) standards such as Universal Description, Discovery and Integration (UDDI), Web Services Description Language (WSDL), and Simple Object Access Protocol (SOAP). A WSDL document describes the Web service interface such as what operations the Web service supports, what protocols to use, and how the exchanged data should be packed. From another viewpoint, the WSDL document can be deemed as a contract between the Web services requestor and provider. Then, the Web services provider may publish the WSDL document to the Web services broker, via UDDI registries. In fact, UDDI is a "yellow pages" of WSDL documents. UDDI provides a standard means for describing businesses and their services and allowing the online discovery. By contrast, a requestor entity can be a person or organization that expects to make use of a provider entity's Web service for achieving its business requirements. The Web services requestor may find a Web service that matches with certain specific requirements by using UDDI registries. In this scenario, the Web services broker would act as a matchmaker between the Web services provider and requestor. From another point of view, the Web service broker is working as a discovery agency, like Web search engine such as Google and Yahoo. Once a Web service is found in the UDDI registries, the Web services requestor gets the correspondent WSDL document and tries to bind with the Web service via a SOAP message. SOAP is an XML-based messaging protocol that is independent of the underlying transport protocol (e.g., HTTP, SMTP, and FTP). SOAP messages are used both by services requestors to invoke Web services, and by Web services to answer the requests. Therefore, the Web service receives the input SOAP message from the Web services requestor and generates an output SOAP message to the Web services requestor. This model is called the publish-find-bind model. Web services are nowadays widely supported by all the major vendors such as IBM, Microsoft, Sun, HP, and BEA. In contrast, the benefits of adopting Web services include faster time to production, convergence of disparate business functionalities, a significant reduction in total cost of development, and easy to deploy business applications for trading partners (Ratnasingam, 2002). Another difference between traditional business-to-business applications and Web Services is a secure environment versus an exposed environment. Ratnasingam (2002) expects financial services to be early adopters of Web services that may involve a set of diverse trading partners working closely together in a highly competitive market.

With the recent boom of B2B applications, there are increasing demands and discussions about privacy technologies for supporting different business applications in the industry and research community. At this moment, there is still no standardized Web services privacy control technology. To enable privacy protection for Web service consumers across multiple domains and services, the World Wide Web Consortium (W3C) published a document called "Web Services Architecture (WSA) Requirements" that defines some specific privacy requirements for Web services as a future research topic. There

are five specific privacy requirements (Ref: AC020) for enabling privacy protection for the consumer (user) of a Web service across multiple domains and services as follows (W3C, 2002):

AR020.1: The WSA must enable privacy policy statements to be expressed about Web services.
AR020.2: Advertised Web service privacy policies must be expressed in P3P (W3C, 2005).
AR020.3: The WSA must enable a consumer to access a Web service's advertised privacy policy statement.
AR020.5: The WSA must enable delegation and propagation of privacy policy.
AR020.6: Web Services must not be precluded from supporting interactions where one or more parties of the interaction are anonymous.

Privacy technologies have also been investigated in the Web business environment for a period of time (Senicar, Jerman-Blazic, & Klobucar, 2003). Although many other companies have already been providing various privacy tools in the past few years, very few privacy *standards* exist beyond a principal statement made by IBM and Microsoft: "organizations creating, managing, and using Web services will often need to state their privacy policies and require that incoming requests make claims about the senders' adherence to these policies" (IBM & Microsoft, 2002). WS-Privacy has been mentioned in industry for a period of time for defining subject privacy preferences and enterprise privacy practice statements for Web services (IBM & Microsoft, 2002). At the time of writing, the WS-Privacy specification has not been released to the public yet.

IBM has recently proposed the Enterprise Privacy Authorization Language (EPAL) (Leino-Kilpi et al., 2001) to address the growing need for privacy authorization languages. EPAL is used to specify the privacy authorizations for the actual enforcement of intra- or inter-enterprise privacy control, by abstracting the data models and user-authentications from all the deployment details. In addition, EPAL is an interoperability language for defining the enterprise privacy policies on the data handling practices in the context of fine-grained positive and negative authorization rights. The goal behind EPAL is to enable an enterprise to encode its privacy-related data-handling policies and practices in XML for facilitating privacy-enforcement in enterprise information systems. Its recent emergence as a fine-grained privacy-related language standard is in response to the irreversible trend of having more and more dynamically formed and evolving federations of organizations in this e-business era. The importance of managing personal identifiable information (PII) for such dynamic "coalitions" has been echoed by a recent project of Network Associates called AMBer (AMBer, 2003). EPAL works together with the eXtended Access Control Markup Language (XACML) (OASIS, 2003) as well as the recent development of access control in Semantic Web (Agarwal, Spick, & Wortmann, 2004) to achieve privacy control. However, this centralization approach is often inappropriate for tele-marketing and many other service outsourcing scenarios because the service provider is not playing an active role to ask for data access to the enterprise. In fact, the service provider is playing a passive role to receive the information from the enterprise after making the contract.

The P3P working group at the W3C develops the P3P specification for enabling Web sites to express their privacy practices (W3C, 2005). On the other hand, P3P user agents allow users to automatically be informed of site practices and to facilitate decision-making based on the Web sites' privacy practices. Thus, P3P also provides a language called P3P Preference Exchange Language (APPEL) version 1.0 for expressing users' preferences of making automated or semi-automated decisions regarding the acceptability of machine-readable privacy policies from P3P enabled Web sites (W3C, 2005). Further, Lategan and Olivier (2002) propose a conceptual model for enhancing the decision-making at the user agents by using the Chinese Wall security policy based on the P3P framework. Though the P3P framework is not mainly designed for supporting Web services privacy policies, the P3P working group is currently studying the feasibility of applying a revised version of P3P for this purpose. We show in our case study that P3P and APPEL can be used for Web services privacy as well as an illustrative privacy meta-language that is also based on the basic principles of P3P.

Replicated information under different formats and classifications within an enterprise can cause difficulties in information security management and auditing. Stoica and Farkas (2004) propose the use of ontology as a concept unifier for alleviating any possible security threats. We believe that the need of concept unification will be even more important for privacy enforcement in inter-organizational system integration, like that between a bank and a tele-marketing company. Apart from the handling of heterogeneities, different entities may also need to resolve conflicts or come up with consensus when the privacy rules do not accommodate each other. In this case, human intervention may be needed in order to resolve the privacy conflicts. Lupu and Sloman (1999) as well as Yee and Korba (2004) provide further details on the research issues of conflict resolution, which are out of scope of this paper. However, we believe the concept unification process can at least automate the integration of most common privacy concerns. As a bank is a much larger organization with much better technology support, a outsource partner

(usually serving one bank because of possible conflict of interest) can simply follow the ontology of the bank.

In particular, the OASIS ebXML Collaboration-Protocol Profile and Agreement Technical Committee (OASIS, 2002) has formed an auto-negotiation sub-team since August 2001. Their primitive goal is to automate the negotiation process between two negotiation parties for bargaining different technical issues in the context of Collaboration-Protocol Profile (CPP). As a result, an agreement between two negotiation parties is expressed in the format of Collaboration-Protocol Agreement (CPA). In addition, they are planning to support the negotiation of higher-level issues such as business parameters and legal matters. However, their work mainly focuses on the CPP and CPA templates. Furthermore, WS-Privacy has been mentioned in industry for a period of time for defining subject privacy preferences and enterprise privacy practice statements for Web services (IBM & Microsoft, 2002). At this minute, the WS-Privacy specification has not been released to the public yet.

Recently, some Web services research papers have been published. For example, Hong, Ng, Lederer, and Landay (2004) discuss a privacy risk study model that contains two parts for defining privacy issues. The first part is a privacy risk analysis and the second part looks at privacy risk management. Next, Hung, Ferrari, and Carminati (2004) introduce a vocabulary-based Web services privacy framework with a vocabulary independent privacy meta-language to describe all the WSA privacy requirements properly (W3C, 2002). In addition, Carminati, Ferrari, and Hung (2005) discuss the privacy issues related to discovery agencies in the context of publish-and-find Web services with different technical approaches to tackle the privacy issues in UDDI registries. Cattaneo, Faruolo, Petrillo, and Persiano (2004) present a SOAP extension for protecting the privacy of users of a Web Service. Yee and Korba (2004) present a Privacy Policy Compliance System (PPCS) for handling security requirements and implementations. However, none of these recent works present an integration model of enforcing end-to-end privacy control like this paper.

By combining the contemporary technologies as discussed above, this paper arrives at a solution to replace the traditional uncontrolled transfer of customers' data in bulk during the outsourcing of marketing activities. We demonstrate that Web services can adequately facilitate the operation and management such outsourcing activities by employing a layered architecture for end-to-end privacy control on the export of individual customer's record and fields according to the corresponding enterprise's privacy control policies through a Web services platform augmented with technologies like the P3P and the APPEL. As far as we know, there have not been any published papers taking an approach similar to ours. Nor are there methodologies addressing the application design for the end-to-end privacy control of service outsourcing that extensively involves human service personnel in the processes.

4 Methodology and model

Regarding the drawback of current practices, we adapted the layer architecture of Cheng and Hung (2005) to provide a full solution containing the elements in the layered architecture as depicted in Fig. 2:

- *Contracts and Laws*: provide a base line level for regulations, codes of practice, and acceptable policies. In addition, the contract specifies the requirements of marketing activities for both parties as well as the penalties for the violation of other clauses and remunerations.
- *Audit Application and Provided Seal Service A*pplication: help ensure that the service company complies with the policies and improves service reliability. In particular, the service company must use the provided seal programs for the outsourced activities to further ensure that data is only used for the desired purposes. Typical enterprise software infrastructures such as Java 2 Enterprise Edition (J2EE) and Microsoft .NET are best suited for this purpose.
- *Privacy Access Control Rules*: ensure only the necessary PII is revealed to the appropriate personnel, with further help of a*nonymity tools*. This layer is responsible for enforcing user privacy. Based on the privacy model discussed above, users can have their own privacy profiles and can be created, viewed, modified and deleted at any time. This layer aims to handle and store different privacy profiles of individual users for which can be presented in P3P and APPEL.
- *Ontology*: specifies the role of the personnel and software as well as meaning and structures of the data in both parties in particular for the purpose of privacy control. We propose to use the Web Ontology Language (OWL) or the DAML language to create ontologies and to markup information so that it is machine readable and understandable. DAML-S defines an upper ontology for describing the semantics of Web services (DAML, 2003).
- *Web Services*: provide controlled access points in the form of a programmatic interface. Usually a Web services provider may have a different privacy profiles to different requestors. To resolve the conflicts of privacy profiles, this Web services application layer can provide controlled access points in the form of a programmatic interface. To illustrate the negotiation of a privacy policy profile, Yee and Korba (2004) propose the Privacy Policy Compliance System (PPCS) to advocate privacy policy negotiation to protect personal privacy. The compliance system can be

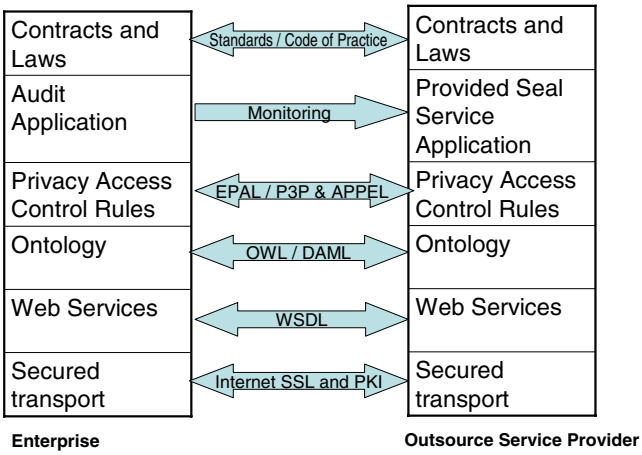


Fig. 2 A layered model for tackling privacy protection in service outsourcing

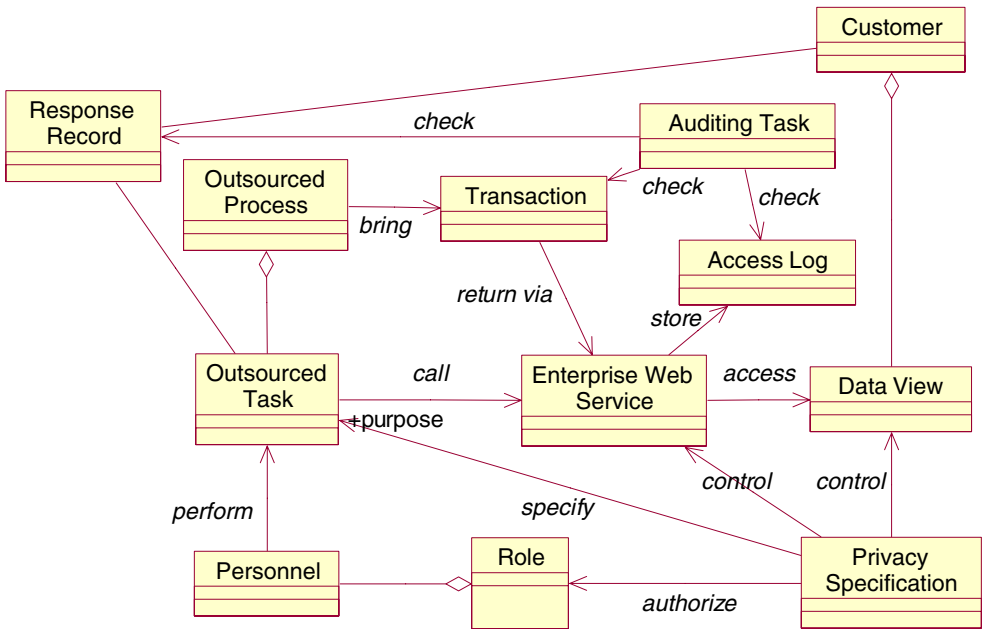
used for negotiating a user's individual privacy. As we only allow access to an individual customer's record one at a time for processing, the use of semi-automatic control such as email or file transfer is not viable. We also observe that a Web-based program hosted at the enterprise is inadequate for integration with the partners' systems and it could eventually have scalability problems.

- *Secured Transport*: support secure data in transit such as the Public Key Infrastructure (PKI) and Secure Socket Layer (SSL). Data confidentiality and integrity is provided in this layer. The existing encryption tools such as SSL and PKI can be used to support secure data in transmission and data exchange. Storage of personal information at the outsource service provider should be avoided as far as possible. Both J2EE and .NET have good library support for secured transport implementation.

Our methodology relies on the need-to-know basis to release just the *sufficient* amount customer information to a partner (Cheng & Hung, 2005). Figure 3 depicts a conceptual model of our Web services-based approach based on the requirements outlined in our layered model. We highlight some of the important steps and the main strategies in the system development methodology as follows, before explaining some details in the next session:

- Outsource contracts should comply with the requirements in laws and code of practices of the industry. In particular, privacy requirements should be specified explicitly and in details because outsource partners are often unfamiliar with requirements of individual industries and impossible to know the other additional requirements of the enterprise. These are translated into the privacy specifications.
- As most enterprises may already have some privacy policy guarding the operations of their employees in their own existing marketing activities, this may serve as some basic hint. However, the rules must be stricter for outsourcing.
- Study carefully the outsourced process. According to the principles of RBAC, identify the different employees' roles of both parties required for the facilitation of each task of the process.
- Based on the "need-to-know" principle, partition the data into *views* according to the need of different stages of the outsourced process, in consideration of the sensitivity of each data field (Cheng & Hung, 2005). Associate the access control based on the employees' roles of both parties as well as individual task requirements. Determine

Fig. 3 A conceptual model of web-service-based privacy access control



contingencies and necessary override mechanisms in order to avoid ad hoc decisions.
- Express these rules in a high level language (such as P3P and APPEL).
- Make sure that outsource partners understand not only the privacy rules but also the ontology based on which these rules are defined.
- In constructing the seal marketing application for the outsource partners, the in-house version may serve as a blue-print but definitely not those of the outsource partners. This is because the in-house program normally already has some tailored privacy protection features. In particular, if such in-house program has been written with enterprise support features (such as with J2EE), the adaptation is relatively easy and reliable as mainly the access control policy and data access path (currently via Web Services) have to be modified.
- Analyze the data requirement at each stage of the outsourced service activity based on the need-to-know principle and formulate the minimal data views on the customer data for each stage for such privacy restrictions.
- Ensure that the outsource partners access the data normally one at a time via only pre-defined Web service calls and from the authenticated seal programs (especially for the call center software) provided by the enterprise. Utmost care should be taken for batch processing of customer data, such as report generation.
- If software other than the seal application is allowed to access the Web services, Service Level Agreement (SLA) must be constructed accordingly and compliance must be validated. A SLA is a formal contract between a Web services requestor and provider guaranteeing quantifiable issues at defined levels only through mutual concessions (Sahai, Durante, & Machiraju, 2002).
- All the data access, the customers' response, and transactions should be recorded for auditing purposes.
- The enterprise should construct additional auditing programs to monitor the execution of the external seal programs at the outsource partners to ensure (a) privacy constraints are met, and (b) the integrity and correctness of transactions. Otherwise, alerts should be sent to the management. For example, we may want to ensure that the access of customers' data from a outsource partner should not be higher than a certain rate otherwise we may suspect a leakage of information.

For example, an enterprise can pre-filter the data to block all customers who refuse to participate in a particular activity. A minimum set of information (e.g., the name of the customer, salutation, and the phone number) according to the data requirement of the outsource process (for example with the approach of Cheung, Chiu, & Till, 2003) is used for the service personnel to contact the enterprise's customers. The service personnel have a restricted view of the data and only one at a time. The information fed into this step is "adequate, relevant, and not excessive in relation to the purpose for which they are processed." At the end of the processing, the service company should destroy the data supplied by the enterprise. With this approach, the customer's data are also ensured up-to-date as the information is only transmitted to the service partner upon usage.

5 Towards end-to-end privacy control

In this section, we illustrate some of the technical details of our approach towards end-to-end privacy control based on our case study. The personnel of external marketing partners execute marketing processes, each composed of different marketing tasks, for the bank. Privacy access control specifications govern the access of the customers' data views (instead of the whole record) according to the role of the service personnel as well as the marketing task which can adequately reflect the purpose for the access. Web services enable the seal marketing application to access the data in such a prescribed way and all the access are stored in an access log. Successful transactions are sent back to the bank also via Web services. All the customers' responses are recorded for auditing against the access log and transaction records. We now proceed to discuss further details of our approach.

5.1 Protocols and architecture for tackling privacy issues

Figure 4 depicts our enhanced approach: "Information dissemination on demand." The idea is that the bank only passes the minimum necessary information onto the marketing company for doing the tele-marketing. This means only part of a customer record is sent at a time, as illustrated in the typical marketing process flow of the next sub-section. With this arrangement, the customer privacy is further protected and the bank has control over who can gain access and why the customer information is passed onto the marketing staff. By comparing the application list to the Web services log with the application list, we can identify whether an abuse of customer information has happened. Suppose a marketing campaign has a success rate of 20%. Using this approach, 80% of the customers who have rejected the product have their privacy protected, in contrast to the approach where a full set of customer information is emailed to the marketing company and left for them to control.

Both the bank and the marketing company are assumed to have a common ontology (vocabularies) to share. The credit card data ontology (Step 1 in Table 1) is generated

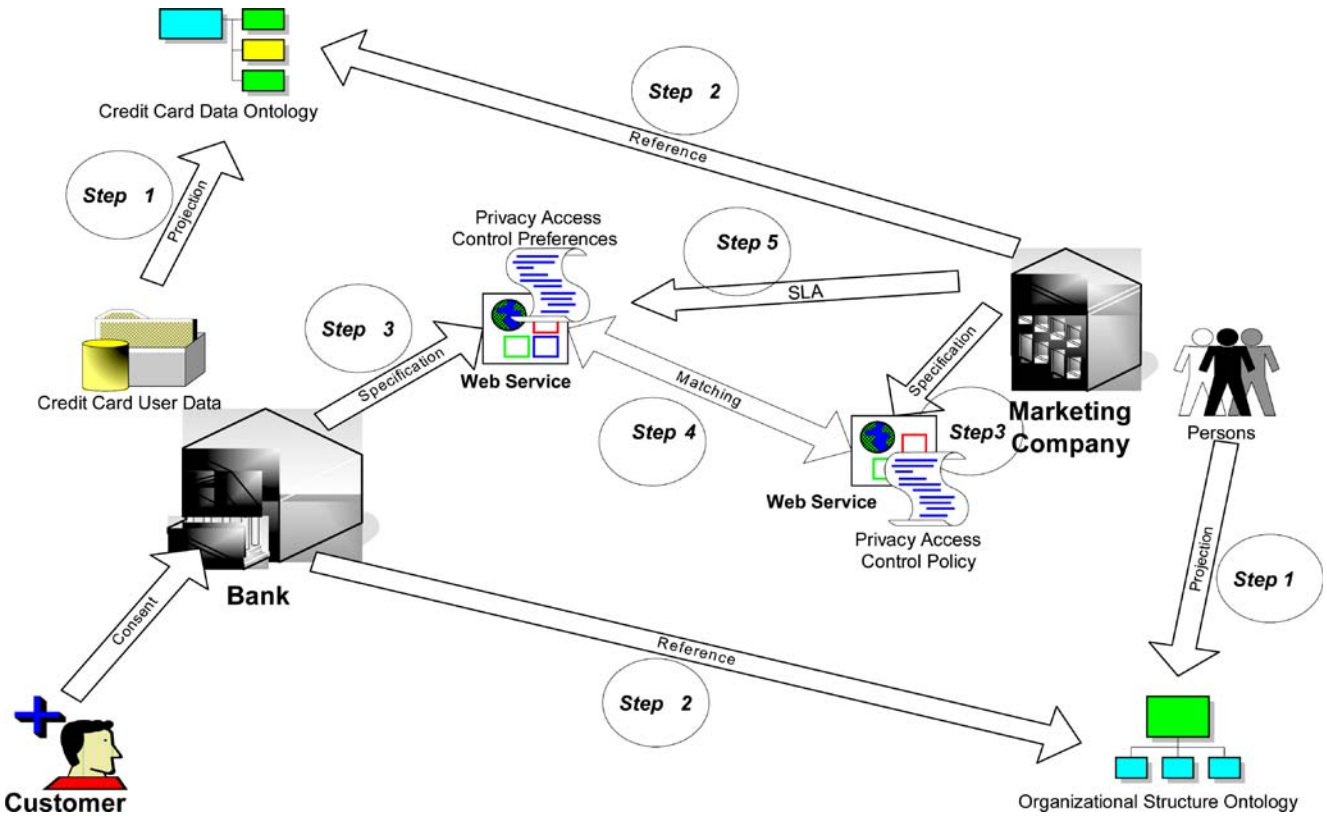


Fig. 4 A protocol and architecture for tackling privacy access control issues

from the credit card user data (maybe in some standardized schema) at the bank, while the organizational structure ontology (Step 1 in Table 1) is generated from the organizational structure in the market company. From the technical perspectives, the ontology is implemented by the OWL or DAML. Next, the bank has to specify its privacy access control preferences based on its own credit card user data ontology as well as the organization structure ontology (Step 2 in Table 1). Since the bank does not really know who at the marketing company will access the customers' data, the bank should gather the information of roles from the organization structure ontology. On the other hand, the marketing company also references its own ontology and the customers' data ontology developed in the OWL and DAML.

Then, the bank can specify the privacy access control preferences (Step 3 in Table 1) for explicitly declaring which role(s) can access (what type of privileges) which data subset(s) for which purpose, retention policy, obligations, and other privacy entities. For illustration, Fig. 5 describes an XML document of APPEL privacy preference to describe the bank's privacy control preferences. On the bank side, the privacy preference states that "the marketing company must be the only recipient to collect the data (credit card customer information) for tele-marketing purposes only." Referring to the Fig. 5, the assertion <STATEMENT></STATEMENT> describes the data practices as applied to data elements. The assertion <RECIPIENT></RECIPIENT> describes the legal entity, or domain, beyond the service provider as well as its agents where data may be distributed. In this case, the recipient is <ours/>, meaning that only the marketing company collects the data. The assertion <PURPOSE></PURPOSE> describes the reason(s) for data collection and use. In this case, the purpose is for tele-marketing only. Then, the assertion <DATA-GROUP></DATA-GROUP> describes the data to be transferred or inferred. Referring to the APPEL privacy preference shown in Fig. 5, there are two statements (rules) defined as follows: (a) only the marketing company <ours/>

Table 1 The steps shown in Fig. 4

Step	Bank	Marketing company
1	Generate the credit card data ontology	Generate the organizational structure ontology
2	Check the organizational structure ontology	Check the credit card data ontology
3	Specify the privacy access control preferences	Specify the privacy access control policy
4	Matchmaking process	
5	Request the Service Level Agreement (SLA)	Generate the Service Level Agreement (SLA)

```
<appel:RULE behavior="tele-marketing.request">
  ...
  <-- evidence (abbreviated) -->
  ...
  <POLICY>
    <STATEMENT>
      <RECIPIENT appel:connective="or-exact">
        <ours/>
      </RECIPIENT>
      <DATA-GROUP appel:connective="or-exact">
        <DATA ref="#credit-card-customer.*"/>
      </DATA-GROUP>
    </STATEMENT>
    <STATEMENT>
      <PURPOSE appel:connective="or-exact">
        <tele-marketing/>
      </PURPOSE>
      <DATA-GROUP>
        <DATA>
          <CATEGORIES appel:connective="or-exact">
            <DATA ref="#credit-card-customer.*"/>
          </CATEGORIES>
        </DATA>
      </DATA-GROUP>
    </STATEMENT>
  </POLICY>
  ...
  <-- evidence (abbreviated) -->
  ...
</appel:RULE>
```

Fig. 5 An illustrative APPEL privacy preference

can be the recipient of the data of credit card customer, and (b) the purpose must be only for <tele-marketing/>.

On the other hand, the marketing company does not know what are the credit card holders' names, addresses, and the like are, before the bank transfers the data to them. Then, the marketing company specifies the privacy access control policy (i.e., who will access what information for what purpose) based on the task (e.g., promotion campaigns) offered from the bank (bid a business contract). Referring to Fig. 6, the privacy policy at the marketing company states that "we are the only recipient collected the data (credit card customer information) for tele-marketing purposes" by using P3P.

There is a matching mechanism (Step 4 in Table 1) for both parties to check whether there is any conflict between the preferences (what the bank would like) and policy (what the marketing company is offering). While the bank is deciding whether to transfer the credit card holders' data to the marketing company (grant the project) or not, the comparison between these entities is one of the determinant factors. From the technical perspectives, middleware that provide an infrastructure to support the consistent enforcement of the privacy policies and preferences across the organization are required in this scenario. Based on this example, the P3P privacy policy at the marketing company matches with the rules set at the bank's APPEL privacy preference. As we have discussed before, the bank may also pass more customer information to the marketing company with the customer's consent. Before the bank sends the data to the marketing company via the Internet, the bank also have to concern about the communication security. To enforce end-to-end privacy, the marketing company adopts the PKI security mechanism to ensure the confidentiality of all transmitted data as well as to verify and authenticate the validity of each party in the context of public and private key. Then, Secure Socket Layer/Transport Layer Security (SSL/TLS) is used to authenticate peer clients with digitally signed certificates. These protocols provide endpoint authentication and communications privacy over the Internet using cryptography. The SSL is a commonly used protocol for managing the security of message transmissions over the Internet. SSL has recently been succeeded by TLS, which is based on SSL. SSL uses the public-and-private key encryption system from RSA, which also includes the use of a digital certificate. TLS and SSL are not interoperable. However, a message sent with TLS can be handled by a client that handles SSL but not TLS.

Referring to Step 5 in Table 1, the bank is asking for the SLA at the marketing company. The SLA should specify what kind of security mechanism the marketing company adopts to ensure the communication security. Up to this moment, W3C does not yet have any specification for modeling Web services agreements. According to some existing SLA specifications, one potential solution that can be applied in this model is Web Service Level Agreement (WSLA). Further, WSLA even provides an extensible mechanism to include domain-specific vocabularies (IBM, 2003). WSLA can be used by both service provider and service customer to configure their respective systems to provide and supervise their services. An important aspect of WSLA is its capability to deal with the specifics of

```
<POLICY>
  ...
  <-- evidence (abbreviated) -->
  ...
  <STATEMENT>
      <RECIPIENT><ours/></RECIPIENT>
      <PURPOSE><tele-marketing/></PURPOSE>
      <DATA-GROUP>
        <DATA ref="# credit-card-customer.*"/>
      </DATA-GROUP>
  </STATEMENT>
  ...
  <-- evidence (abbreviated) -->
  ...
</POLICY>
```

Fig. 6 An illustrative P3P privacy policy

Fig. 7 An illustrative WSLA document for a tele-marketing web service

```
<SLA>
  ...
  <-- evidence (abbreviated) -->
  ...
  <ServiceDefinition name="tele-MarketingService">
    <Operation             xsi:type="wsla:WSDLSOAPOperationDescriptionType"
name="WSDLSOAPCollectData">
      ...
      <-- evidence (abbreviated) -->
      ...
      <SLAParameter name="communicationSecurity">
        <Metric>encryptionMechanism</Metric>
      </SLAParameter>
      <Metric name="encryptionMechanism">
        <Function xsi:type="wsla:publicKeyInfrastructure">
          <Metric>publicKey</Metric>
          <Value><LongScalar>
          mQGiBEJ60CIRBADCoT30IShJiLxaqU/4SPfkUXJDGwXl6vhFl/qCK9J1hHlVvG78
          fshuGUi63S6r5UiRqfYnaObRe0x+//k4NU4j2y7T07g5uTilbn6zr9D1N/hnwQJH
          kd2Jetz/xuw1i6ppFojI1v3FFigBv1v0IaOvzPQ56MwToy1TtDzuKHgIBQCg/8ef
          +Khs+6biExoJD+iCpy2WB+8EAKUVtmF1zfrT+FlJMfGlcznMYFncM1IfpOmWQjt7
          NxQy6iwc3bf5dCgDR/ppvnvn5/x9h6e7o+QqAccYBCqv33vOcBqgHyLOaKDPFUNM
          4UeODIWfaVqBJU2YtghHWATrkfQwZDvOwSMLqsPPKyx2wiV/0FL+Fav9ehW7Lqsb
          nOGlBAC5TJMKr/lzsfuxPsI5sLvnBAr/t16KwvhIaZULM0R/CyCoNJoBb05lwnEY
          oXoyCM/Y1ekL/OtJ2YRSybzu7HAaoG0i1LbkR/2AyQl1XKEhn5YUtXYUDpVT/JKV
          IfLhiiYWw1AzrRHzM+Q17JP8vzBSUAHf8c0o6kL4Myl6GYq9krQjUGF0cmljayBI
          dW5nIDxwYXRyaWNrLmh1bmdAdW9pdC5jYT6JAE4EEBECAA4FAkJ60CIECwMCAQIZ
          AQAKCRDMpAsFxAThDanjAJ9bvH57xL6VyDJqGXahtSL9Ah50TACgl/Arn1BioO9w
          uX2dvbzedrrWlcG5Ag0EQnrQIhAIAPZCV7cIfwgXcqK61qlC8wXo+VMROU+28W65
          Szgg2gGnVqMU6Y9AVfPQB8bLQ6mUrfdMZIZJ+AyDvWXpF9Sh01D49Vlf3HZSTz09
          jdvOmeFXklnN/biudE/F/Ha8g8VHMGHOfMlm/xX5u/2RXscBqtNbno2gpXI61Brw
          v0YAWCvl9Ij9WE5J280gtJ3kkQc2azNsOA1FHQ98iLMcfFstjvbzySPAQ/ClWxiN
          jrtVjLhdONM0/XwXV0OjHRhs3jMhLLUq/zzhsSlAGBGNfISnCnLWhsQDGcgHKXrK
          lQzZlp+r0ApQmwJG0wg9ZqRdQZ+cfL2JSyIZJrqrol7DVekyCzsAAgIH/3C1FU21
          e5cscG4w6dWifSbeOouO/Ck47ekXRh/EoG/2KNjDMlJZ8GSBoobf4GfA8/Yksr+6
          W5pBsbOa20gvLlUysoY09ljw1Ol5VsK0rwlNjbZBoH146U7D/+2DylJPLbWmuC/I
          5PIsIxPvhb741iccMfBB/81IybwXyJ6mwCT51cB4XQZN1dyZIcIVAp7FZ3iTRuNy
          /NIw46PCyvgn2HMgHiybrxnXKIBD7delDXAwcHGLEJ7iaLzgKyUXMSVEQn+ll4zo
          qKW8zKldrvGZ4mOyWpTejx5i5fTfzSBK0u+AAizBKykA5YFk2Oru5BOwngK9QCt9
          eLDkXZleL0O4r32JAEYEGBECAAYFAkJ60CIACgkQzKQLBcQE4Q0WNACfStFcAe0p
          So2i+xgL6sXLgpyctS8An0tb90bM9eu+LPhdZy7SnbmpuU3b
          =AnA4
          </LongScalar></Value>
        </Function>
      </Metric>
    </Operation>
    ...
    <-- evidence (abbreviated) -->
    ...
  </ServiceDefinition>
  ...
  <-- evidence (abbreviated) -->
  ...
</SLA>
```

particular domains and technologies. Figure 7 shows an illustrative example of WSLA at the marketing company.

5.2 Integrating typical call center software with Web services

To illustrate how privacy can be protected by sending a partial customer record at a time to a marketing partner, Fig. 8 depicts a typical marketing activity and the data requirement at each of its constituting tasks. The integration of the call center software with the bank's Web services for controlled access to the data through views at a different stage of a call is further explained below:

- The tele-salesperson authenticates his/her identity to the seal marketing application so that the role can also be determined. The login procedure is part of the role authorization procedure where the marketing partner specifies access control to the users of the data.
- The tele-salesperson can then work on an assigned or authorized campaign. Again we have a second level access control where each operator can view info only of this related campaign.
- The bank's host has already filtered a segment of customers for the campaign. So, the marketing seal program can only select a customer from such a list through Web services.
- At this point, the program can only obtain the customer's phone number, surname, and salutation. The application then dials the phone number, which is concealed from the screen and sent directly to a phone device.
- Only when the phone gets through will the customer's surname and salutation be displayed on the screen of the tele-salesperson.
- If the customer is interested in the product (or service), the tele-salesperson can now start to explain more details to the customer and therefore need more data to do personalized persuasion. Depending on the marketing campaign, such information could be the salary range, age range, etc. The seal program then obtains these data accordingly through the designated Web services and displays them.
- Only when the customer agrees to purchase the product, the program will display more sensitive information such as credit card number or addresses for confirming the customer's identity, payment, and/or product delivery.
- The details of the activities, particularly the customer's response, are recorded in the system. The bank staff will normally get the list of confirmed customers. For the marketing partner, they need to know how many customers each staff member has approached and among these, how many have been confirmed. This determines the marketing personnel's performance and thus their commission and salary. Therefore, the marketing partner could have a system for monitoring the callings of each staff.

The interface between the marketing company and the outsourcer would be the number of customers called and the number of customers who have confirmed the purchase. Sometimes, the outsourcer might perform sample check the recording tape. As confirmed customers would be involved in the process of delivering products and charging, the bank must make sure that the list of customers who have agreed to purchase the service/products are complete.

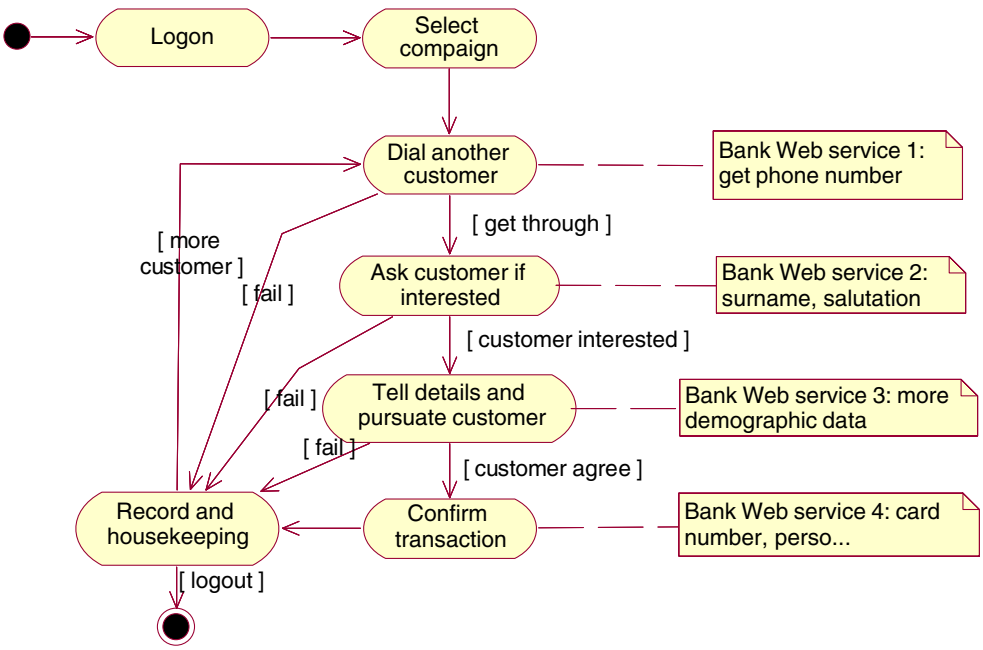


Fig. 8 An example marketing activity of an outsourced call center

The bank should construct additional auditing programs to monitor the execution of external seal programs at its marketing partners to ensure (a) privacy constraints are met, (b) the correctness of transactions. Otherwise, alerts should be sent to the management. For example, we may want to ensure that the access of customers' data from a service provider should not be higher than a certain rate otherwise we may suspect a leakage of information and then more sampling check should be carried out.

6 Discussion

In this section, we first compare the advantage of our approach to human-intensive service outsourcing practices over traditional ones such as sharing applications through Virtual Private Network (VPN) alone and Electronic Data Interchange. Then we discuss the technical challenges of our approach. To replace ad hoc email and bulk data exchange, one approach is to allow outsourced service providers to connect to the systems of the enterprise through VPN. This works may reasonably well for human intensive processes in some cases if the requirements are simple and the outsourcing activities are not much. In essence, the sealed application presented in the previous session is executed in the enterprise and the staff of the external service provider executes. However, there are many drawbacks. Opening up enterprise application to outsourced partners may create loopholes of malicious activities. This approach may not scale up when there are too many outsourced activities. Then, also the authorization management together with auditing will be a problem. In addition, the systems of the outsourced service provider cannot be integrated in a safe way. One implication is that the outsourced service provider can only trust in the enterprise for the job accounting, which may not be practical in many cases.

Another possible approach to support the service outsourcing activities is using Electronic Data Interchange (EDI). EDI is a computer-to-computer transfer of business information between two businesses that uses a standard format of some kind. Although Internet EDI is growing and offering new, flexible information interchange solutions for many trading partners, some elements of EDI remain difficult to transfer to the Internet due to their security concerns. To secure the communication channels on the Internet, the infrastructure of our architecture in Fig. 4 illustrates how to build on the platform of a Virtual Private Network (VPN). VPN is an extranet that uses public networks such as the Internet and their protocols to send sensitive data to partners, customers, suppliers, and employees using a system called IP tunneling or encapsulation. Further, the middleware for supporting our infrastructure in Fig. 4 is implemented by the IBM Tivoli Privacy Manager (IBM, 2005). The IBM Tivoli Privacy Manager builds P3P-based privacy policies and practices by providing an infrastructure across the organizations. With this approach, privacy-sensitive data is linked to policy at the point of collection, and subsequent requests to use the data are then filtered such as permitted or denied in according to policy and the data owner's preferences.

In fact, our architecture is not only limited to be applicable into marketing services. The methodology can also be adopted into other scenarios such as healthcare services, financial services, and insurance services. The main reason is that Web services technology consolidate all disparate data exchange and process interactions. This single border check enables all the governance policies and auditing procedures to be effectively executed and managed in a systematic manner.

The national and international legislative systems have sprung up to define and protect the privacy of personal data, and the legislations also apply to the borderless world of cyberspace. Fail to comply with the privacy legislations may lead to civil and/or criminal penalties and/or imprisonment, as well as the loss of reputation and goodwill when the non-compliance is publicized. As for legal compliance, effective privacy policy and system architecture is required in countries all over the world. However, because different countries have different requirements, mechanism for facilitating the customization of such policies is essential. To our best knowledge, there is no prior research on the privacy technologies for supporting service outsourcing and no solution to support Web services-based applications with privacy enforcement. This infrastructure perhaps was the first guideline setup for implementing multi-national privacy legislations compliance applications and the Web services-based applications. In fact, most of the technology components (for example SSL) inside the framework in Fig. 2 are not new. However, the idea of Web services integration and privacy framework are new. So one of the possible potential difficulties may turn up would be the interoperability between different layer. However, once a privacy policy description language has been standardized such as P3P and APPEL, we believed our architecture should be convincing even it is not fully implemented yet. It is therefore necessary to develop applications that are compliant with different legislations. There has been an increasing amount of discussion recently about a Web services privacy framework in both industry and the research community. However, there is still no comprehensive solution to tackle the various privacy issues in Web services have been defined and developed yet.

In this paper, we are proposing to tackle semantic issues in Fig. 2 by using the Web Ontology Language (OWL)

(W3C, 2003). OWL is a XML language proposed by the World Wide Web Consortium (W3C) for defining ontology. OWL ontology includes descriptions of classes, properties, and their instances, as well as formal semantics for deriving logical consequences in entailments. Referring to Fig. 2, we are currently working on the following technical research issues:

- *Ontology*: Adopt OWL with EPAL vocabularies
- *Privacy Access Control Policy*: Adopt EPAL with extended assertions;

From the practical and commercial perspective, we are also investigating research issues like:

- Critical success factors for the Web services-based end-to-end privacy control systems
- Cost and technical requirements for the involved parties
- The implementation issues of the system
- Extending the model to other applicable scenarios such as credit reference agencies

Knowledge about the corresponding equivalence and class hierarchies can effectively be described using Semantic Web ontology language like OWL. It could then further enable EPAL-based privacy policies to be properly transformed from those adopted by the bank to those used in the tele-marketing company for enforcement. To support more flexible and manageable transformation, inference engines like F-OWL (F-OWL, 2004) could be adopted. In a highly populated city such as Hong Kong where most businesses are mainly SME, the relationship between banks and telemarketing companies can be really complicated and sometimes aggressive, due to the competitive environment. Therefore, privacy enforcement is a critical issue that needs to be addressed in urgent.

7 Conclusion and future work

With the increase in the scale and complexity of the service outsourcing, neither traditional batch data transfer nor simple web-based software can meet the integration requirements. One can imagine that the demand for privacy-enhancing technologies is ever increasing. In this paper, we have introduced a layered architecture for the facilitation of privacy control based-on Web services, together with a methodology for this purpose. We have further formulated a conceptual model of Web-service-based privacy access control to facilitate the design of implementation architecture, from which outsourced service providers can be integrated with adequate control and auditing. We have also illustrated the practicability of our approach by demonstrating with a typical tele-marketing case study. We show how the software of a service partner can be integrated effectively with the enterprise's Web services, from which only the required part of a customer record is retrieved through the appropriate data views and sent one at a time to achieve strict end-to-end privacy. In this way, un-audited batch transfer of customers' data is no longer necessary and the protection of customer privacy is much enhanced.

From this case study, we illustrate that privacy technologies at the middleware layer (such as P3P) is inadequate to solve all the problems, particularly when human service providers are involved in the processes. Therefore we adopt a holistic approach to the problem in our methodology with a multiple-layer model. In particular, addressing the requirements from the application layer and the users' perspective is of vital importance, and has not been comprehensively discussed before. Our case study is also significant because not only can new systems for outsourcing be built in the way but also there exist many similar legacy systems and practices ignoring the importance of privacy protection.

We are finalizing our implementation plan with IBM WebSphere with J2EE and Web Services as the basic implementation infrastructure. Future work includes the exploration any potential usability and performance issues, particularly the mechanisms and tools for managing the interactions taking place between different layers in our model. We are currently looking into detailed issues of how ontology can further streamline this process in addition to role classifications. Ontology defines the terms used to present a domain of knowledge that is shared by people, databases, and applications. We are interested in the representation of the privacy access control policy in EPAL and the compliance of EPAL to the Web services. In particular, ontology encodes knowledge possibly spanning different domains as well as describes the relationships among them. While system interoperation with privacy enforcement using the ontology-based approach is promising in principle, more research efforts and larger scale deployments are required to meet this emerging information system frontier.

References

Agarwal, S., Sprick, B., & Wortmann, S. (2004). Credential based access control for semantic Web services. In *AAAI Spring Symposium—Semantic Web Services* (March).

AMBer (2003). *Network Associates' AMBer Project*. Online: http://www.networkassociates.com/us/nailabs/research_projects/security_infrastructure/amber.asp.

Carminati, B., Ferrari, E., & Hung, P. C. K. (2005). Exploring privacy issues in web services discovery agencies. *IEEE Security & Privacy Magazine* (September/October).

Cattaneo, G., Faruolo, P., Petrillo, U. F., & Persiano, G. (2004). Providing privacy for Web services by anonymous group

identification. In *Proceedings of the IEEE International Conference on Web Services (ICWS)* (pp. 166–173, 6–9 July).
Cheng, V. S. Y., & Hung, P. C. K. (2005). An integrated privacy framework for HIPAA-compliant web services. *International Journal of Health Information Systems and Informatics (IJHIS)*.
Cheung, S. C., Chiu, D. K. W., & Till, S. (2003). A data-driven methodology to extending workflows across organizations over the internet. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS)*, CDROM, 10 pages, Jan.
CIBC (2005). CIBC's privacy practices failed in cases of misdirected faxes. In *Office of the Privacy Commissioner of Canada, 2005*. Online: http://www.privcom.gc.ca/incidents/2005/050418_01_e.asp.
Constantinides, E. (2002). From physical marketing to web marketing: the web-marketing mix. *Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS)*, pp. 2628–2638, 7–10 Jan.
DAML (2003). DAML-S: Semantic markup for Web services. The DAML Services Coalition, Version 0.9. Online: http://www.daml.org/services/daml-s/0.9/daml-s.html.
Davis, J. C. (2000). Protecting privacy in the cyber era. *IEEE Technology and Society Magazine, 19*(2), 10–22 (Summer).
Diamond (2000). *Marketing, Diamond*.
Ferraiolo, D. F., Kuhn, D. R., & Chandramouli, R. (2003). Role-based access control. In *Computer Security Series*. Norwood, MA: Artech House.
Fischer-Hubner, S. (2001). IT-security and privacy. In *Lecture Notes on Computer Science* 1958.
F-OWL (2004). An OWL inference engine in Flora-2. Online: http://www.fowl.sourceforge.net/.
Hinde, S. (2002). *The perils of privacy* (pp. 424–432). IS Audit Editor. New York: Elsevier.
Hong, J. I., Ng, J. D., Lederer, S., & Landay, J. A. (2004). Privacy risk models for designing privacy-sensitive ubiquitous computing systems. In *Proceedings of the 2004 conference on Designing interactive systems: processes, practices, methods, and techniques* (August).
Hung, Patrick C. K., Ferrari, E., & Carminati, B. (2004). Towards standardized Web services privacy technologies. In *Proceedings of the 2004 IEEE International Conference on Web Services (ICWS)*, 6–9 July, pp. 174–181.
IBM (2003). Web Service Level Agreement (WSLA) Language Specification, Version 1.0.
IBM (2005). IBM Tivoli Privacy Manager for e-business. Online: http://www-306.ibm.com/software/tivoli/products/privacy-mgr-e-bus/.
IBM, & Microsoft (2002). Security in a web services world: A proposed architecture and roadmap. White Paper, Version 1.0.
Lategan, F. A., & Olivier, M. S. (2002). A chinese wall approach to privacy policies for the Web. In *Proceedings of the 26th Annual International Computer Software and Applications Conference (COMPSAC'02)*.
Leino-Kilpi, H., Valimaki, M., Dassen, T., Gasull, M., Lemonidou, C., Scott, A., & Arndt, M. (2001). Privacy: A review of the literature. *International Journal of Nursing Studies, 38*, 663–671.
Lupu, E. C., & Sloman, M. (1999). Conflicts in policy-based distributed systems management. *IEEE Transactions on Software Engineering, 25*(6), 852–869.
National Institute of Standard and Technology (NIST) (2005). *Role based access control standards roadmap*. 12 May 2005. Online: http://www.csrc.nist.gov/rbac/rbac-stds-roadmap.html.
OASIS (2002). *Automated Negotiation of Collaboration-Protocol Agreements Specification. ebXML Collaboration Protocol Profile and Agreement Technical Committee*, Version 0.04, 2002. Online: http://www.oasis-open.org/committees/ebxml-cppa/negotiation.
OASIS (2003). *eXtensible Access Control Markup Language (XACML)*. Online: http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml.
Power, E. M., & Trope, R. L. (2005). Averting security missteps in outsourcing. *IEEE Security & Privacy Magazine, 3*(2), 70–73 (March–April).
Ratnasingam, P. (2002). The importance of technology trust in web services security. *Information Management & Computer Security, 10*(5), 255–260.
Sahai, A., Durante, A., & Machiraju, V. (2002). Towards automated SLA management for web service. In *HP Technical Report*.
Schoeman, E. D. (1984). *Philosophical dimensions of privacy: An anthology.* New York, NY: Cambridge University Press.
Senicar, V., Jerman-Blazic, B., & Klobucar, T. (2003). Privacy-enhancing technologies—Approaches and development. *Computer Standards & Interfaces, 25*, 147–158.
Stoica, A., & Farkas, C. (2004). Ontology guided Security Engine. *Journal of Intelligent Information Systems, 23*(2), 209–223 (Special issue).
W3C (2002). Web services architecture requirements. *World Wide Web Consortium (W3C) Working Draft*, 14 November 2002. Online: http://www.w3.org/TR/2002/WD-wsa-reqs-20021114.
W3C (2003). OWL Web Ontology Language. Web-Ontology (WebOnt) Working Group, *World Wide Web Consortium (W3C)*, 2003. Online: http://www.w3.org/2001/sw/WebOnt.
W3C (2005). The platform for privacy preferences 1.1 (P3P1.1) specification. In World Wide Web Consortium (W3C) Recommendation, 1 July.
Yee, G., & Korba, L. (2004). Privacy policy compliance for Web services. In *Proceedings of the IEEE International Conference on Web Services (ICWS)*, 6–9 July, pp. 158–165.
Zhang, L.-J., Li, H., & Lam, H. (2004). Services computing: Grid applications for today. *IT Professional, 6*(4), 5–7 (July–Aug).

Patrick C. K. Hung is an Assistant Professor at the Faculty of Business and Information Technology since July 2004. He is currently collaborating with Boeing Phantom Works (Seattle, USA) and Bell Canada's Privacy Center of Excellence on security- and privacy-related research projects, and he has filed two US patent applications based on "Mobile Network Dynamic Workflow Exception Handling System" with Boeing. Before that, he was a Research Scientist with Commonwealth Scientific and Industrial Research Organization (Canberra, Australia) and a Visiting Assistant Professor at the Department of Computer Science in the Hong Kong University of Science and Technology. From 2000 to present, Patrick has been serving as a panelist of the Small Business Innovation Research and Small Business Technology Transfer programs of the National Science Foundation (NSF) in the States. He is an executive committee member of the IEEE Computer Society's Technical Steering Committee for Services Computing, a steering committee member of IEEE EDOC "The Enterprise Computing Conference," and an associate editor/editorial board member/guest editor in several international journals.

Dickson K. W. Chiu is the founder of Dickson Computer Systems. Besides, being an experienced consultant, he also teaches in Universities full-time and part-time. He is currently visiting the Hong Kong Polytechnic University as an Assistant Professor. He received the B.Sc. (Hon.) degree in Computer Studies from the University of Hong Kong in 1987. He received the M.Sc. (1994) and the Ph.D. (2000) degrees in Computer Science from the Hong Kong University of Science and Technology, where he worked as a Visiting Assistant Lecturer after graduation. From 2001 to 2003, he was Assistant Professor at the Department of Computer Science at The Chinese University of Hong Kong. His research results have been published in over 60 technical papers in international journals and conference proceedings, such as IEEE Transactions, Information Systems, and Decision Support Systems. He is a mini-track chair in the decision technologies track of the HICSS conference. Dr. Chiu is a Senior Member of the IEEE as well as a member of the ACM and the Hong Kong Computer Society.

W. W. Fung received his BEng (Hons) in Computer Systems Engineering from the University of Warwick (England), M.Phil and Ph.D in Computer Science from the Hong Kong University of Science and Technology. His Master research was on computer vision and robotics, while his Ph.D research was on cryptography for tamper-resistant devices. In addition to research activities in university, he has working experience in government and banking industry. Since 2001, he has been participating in digital forensics and has delivered various security and forensic training to people from various background, including officers from law enforcement.

William K. Cheung is an associate professor in the Department of Computer Science at Hong Kong Baptist University. His research interests include pattern recognition, machine learning, and artificial intelligence with applications to data mining, information extraction, recommender systems, and Web and Grid service management. He received his PhD in computer science from the Hong Kong University of Science and Technology.

Raymond K. Wong is a Senior Lecturer in Computer Science & Engineering at University of New South Wales, Australia. He is also a Project Leader at National ICT Australia. His research interests include XML data management, mobile technologies, and document processing. He received his PhD in Computer Science from Hong Kong University of Science and Technology in 1997.

Samuel Choi is currently a lecturer at Open University of Hong Kong. He received both his Bachelor and Master degrees in computer science from University of Manitoba (Canada), and his PhD degree in computer science from Hong Kong University of Science and Technology. Before joining OUHK, Dr. Choi taught at Hong Kong University of Science and Technology and was a post-doctoral teaching fellow in the Department of Computer Science at Hong Kong Baptist University. His primary research interests are electronic commerce, intelligent agents and artificial intelligence.

Eleanna Kafeza is a lecturer at Athens University of Economics and Business. She received her PhD from Hong Kong University of Science and Technology where she also held a visiting assistant lecturer position. Her research interests are in workflow management systems, legal issues in web contracting, web services and grid computing.

James Tin-Yau Kwok received the Ph.D. degree in computer science from the Hong Kong University of Science and Technology in 1996. He then joined the Department of Computer Science, Hong Kong Baptist University as an Assistant Professor. He returned to the Hong Kong University of Science and Technology in 2000 and is now an Assistant Professor in the Department of Computer Science. His research interests include kernel methods, artificial neural networks, pattern recognition and machine learning.

Joshua C. C. Pun received his M.Phil and Ph.D in Computer Science from the Hong Kong University of Science and Technology. His current research interests include web quality, web metrics and web technology.

Vivying S. Y. Cheung received her Master of Philosophy (MPhil) degree in Computer Science at the Hong Kong University of Science and Technology (HKUST) in 2006. She got her Bachelor of Engineering degree in Computer Science with a minor in Mathematics from HKUST in 2004. Her published work includes research papers in several international conferences and journals. She is now working at IBM Hong Kong as a consultant. Her research interests include security protocols, e-Healthcare security & privacy and Web services.