Personal Information Circulation Technology Based on Policy Control

Minoru Sakuma[†], *Daisuke Hamuro, Akinori Shiraga, and Masayuki Kobayashi*

Abstract

With the boom in online shopping, personal information is now frequently circulated on networks, which raises the importance of protecting the privacy of individuals. One well-known privacy protection architecture on the Web is P3P (Platform for Privacy Preferences). It enables privacy information control by comparing the policies of a user and a Web site. But there is a problem: once personal information has been sent to the Web site, the user has lost control of it. Moreover, P3P is focused on the Web and does not support sensor networks such as ones using ID tags, which have become popular.

This paper discusses an architecture for the secure circulation of personal information that can solve these problems. It also discusses results for a prototype individual location information circulating system for a wireless LAN. This trial showed that personal information could be circulated in real time by this architecture.

1. Introduction

The rise in online activities such as online shopping is leading to more and more personal information circulating on the network such as names, addresses, telephone numbers, and credit card numbers. Data security has become an important issue. According to a questionnaire from the Japanese Ministry of Public Management, Home Affairs, Posts and Telecommunications [1], 77.7% of respondents answered that their biggest anxiety concerning online shopping was personal information being passed on to third parties.

On the other hand, the user's needs for service personalization are rising, so new services using personal information will be beneficial, such as service customization using personal information (address, preferences, etc.), a recommendation service using purchasing history at online shops, or an advertisement service using the user's location. In fact, amazon.com [2] provides a recommendation service using the purchasing history of users.

† NTT Information Sharing Platform Laboratories
Musashino-shi, 180-8585 Japan
E-mail: sakuma.minoru@lab.ntt.co.jp

2. Existing problems about personal information circulation

2.1 Problems from the viewpoints of users and service providers

Problems and requirements can be considered from the viewpoints of users and service providers (e.g., online shops).

- From the viewpoint of users

In Web shopping, if a user is asked to input information, he has no choice because he cannot buy goods unless he fills in his personal information such as mail address or telephone number. But once the information has been entered on the Web page, the user has lost control of it and may consequently receive spams (unsolicited email) or direct telephone advertising. Most people want to avoid this, so they do not want to provide their personal information. With the existing technology, the privacy policy on a Web page is usually the only means for users to check

how personal information will be handled.

• From the viewpoint of service providers

Personal information about customers is an important business asset and service providers want to utilize it positively for marketing and other purposes. However, the damage caused by the leakage of personal information is becoming increasingly serious, so service providers want to manage personal information securely.

2.2 Support for personal information acquired in a sensor network

Personal information has conventionally been acquired through intentional actions of users (e.g., they enter it on a Web page). Recently, however, it has also been acquired without their conscious awareness from sensors such as ID tags, which are often part of a sensor network, so sensor networks must be taken into consideration. Acquiring personal information in a sensor network has the following problems.

- With Web page input, the user can to some extent control information that he intentionally input: he can choose not to input any more information than necessary. However, in a sensor network, he cannot control the circulation of personal information because it was acquired unbeknownst to him.
- Service providers that acquire personal information can show how they treat it in a privacy policy, which conventionally many companies do. But sensor network service providers have no suitable channel for informing users how the personal information will be used, because a sensor network does not have any mechanism equivalent to the privacy policy of a Web site.

2.3 Technology trends about handling personal information

As a standard for the handling of personal information, P3P (Platform for Privacy Preferences) was standardized by the World Wide Web Consortium in 2002 [3]. Its features are described below.

- It offers a standardized form for handling personal information acquired in a Web site.
- It uses XML descriptions to enable user agents to distinguish among different handling conditions.

P3P is currently implemented in some Web browsers and has recently attracted attention as a standard specification for privacy policies. A personal information management architecture based on P3P policy has also proposal [4].

Now, we shall concentrate on the effectiveness of using P3P to solve the problems and meet the requirements mentioned in sections 2.1 and 2.2. In handling personal information on a network, there are three problems.

Problem 1: The privacy policy on a Web browser window is complicated, so it is hard to check.

Problem 2: A user cannot control how his information is treated after he has provided it.

Problem 3: A user cannot control the information acquired passively from sensors, which do not provide a privacy policy to a user.

Problem 1 can be solved because privacy policy judgment can be automated by a user agent. However, problem 2 cannot be solved because P3P targets policy judgment, and personal information acquired as a result of the judgment cannot be controlled. Problem 3 cannot be solved either because P3P targets the Web and does not support sensor networks. Therefore, problems 2 and 3 are remaining issues that need to be solved.

3. Study of policy-based personal information circulation architecture

3.1 Requirements and solutions

In general, personal information circulation involves the following entities.

- Users: they are the subjects described by the circulating information.
- Information providers: they own user information and provide the information to the information requesters. Examples include Web site providers and sensor network providers.
- Information requesters: they request user information and provide personalized services by the user information. Examples include contents or advertisement providers and marketing companies.

Analyzing the existing problems, we extracted two requirements and studied ways to meet them.

First, the privacy policy is conventionally determined between the personal information sender and the receiver. When personal information flows from the user to the first provider, the policy is determined between the user and the provider. However, when the first provider passes the information on to a second provider, the policy is determined between the two providers: the user cannot indicate his policy. Therefore, the first requirement is that users should be able to indicate their policies wherever their information is. To meet this requirement, we make the policy judgment mechanism independent of the sender or receiver and apply it to all information flows.

Second, policy judgment is conventionally carried out between the user and the information requesters because there is no information provider (e.g., a sensor network provider). Therefore, the second requirement is that information providers can indicate their intentions. To meet this requirement, we add the policies of information providers and policy judgment is carried out using the policies of the user, the information requesters, and the information providers.

3.2 Overview of proposed architecture

The architecture we derived from the above study is shown in Fig. 1. There are four entities: the users, information providers, information requesters, and the policy judgment mechanism. The first three have their own policies, and the policy judgment mechanism has a policy which is generated in the policy judgment. An outline of the process is shown below.

(1) The three entities register their policies in the policy judgment mechanism. Specifically, it is as follows.
- A user registers a "user policy", which describes his individual privacy conditions.
- Information providers register "provider policies", which describe the kind of information that can be offered and its handling conditions.
- Information requesters register "requester policies", which describe the kind of information they want and how they will handle it.

(2) The policy judgment mechanism matches these three policies and generates the "information circulation policy", and then selects users that match the circulation conditions.

(3) The policy judgment mechanism transfers the information circulation policy to the information provider that owns the users' personal information.

(4) The information provider processes the personal information according to the information circulation policy specified by the policy judgment mechanism and transfers the information to the information requester. Simultaneously, the provider appends information treatment conditions and indicates that the information must be treated according to the specified treatment conditions.

3.3 Policy architecture and judgment method

Examples of the policies of the three entities are shown in Fig. 2. The various conditions are basically based on the P3P standard, but are extended in some parts. The provider and requester policies consist of the following three conditions, and the user policy contains only information treatment conditions and does not contain any personal information itself.
- Conditions for items that can be provided or which are being requested (kind of personal information)

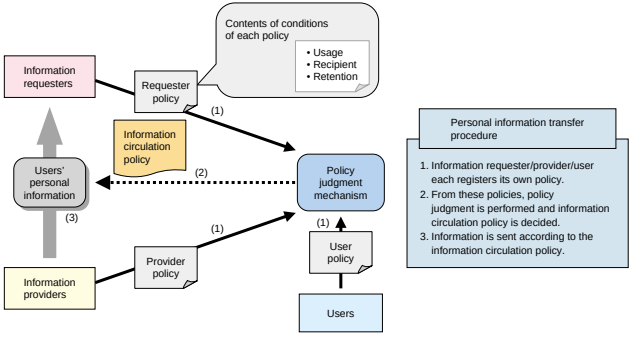

Fig. 1. Outline of proposed architecture.

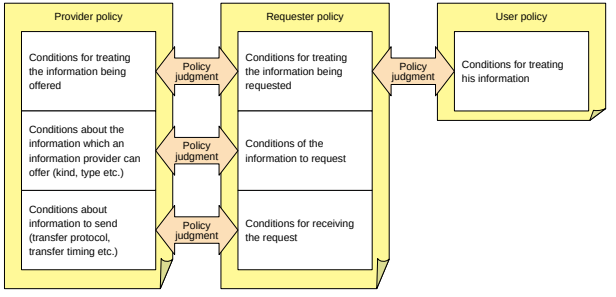

Fig. 2. Contents of the three policies.

- Treatment conditions of personal information (e.g., purpose of user and retention)
- Communication conditions (e.g., transfer protocol and transfer timing)

An example of user policy setup is shown in Fig. 3. The Web-based graphical user interface lets a user set up a user policy simply by checking the various treatment conditions he wants. The specified policy is converted into an XML document and registered in the database of the policy judgment mechanism. The policy files for information providers and requesters are almost the same; they are also registered in a policy judgment mechanism database.

The policy judgment mechanism for determining information circulation conditions is explained below. This judgment mechanism is also an extension of the P3P mechanism, and it is fundamentally carried out by comparing XML tags in policy files.

(1) Conditions for items of circulated personal information

Judge whether items requested by the information requester are ones that an information provider can provide, and whether the circulation of this kind of information is allowed by the user and information provider.

(2) Conditions for treating personal information

Judge whether the treatment conditions requested by an information requester are permitted by the user and information provider.

(3) Conditions for communication

Judge whether the transfer protocol and transfer timing requested by the information requester are ones that an information provider can support.

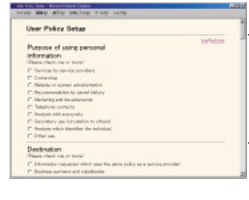

Fig. 3. Example of user policy setup screen.

Examples of information circulation policy and the transferred information (the personal information that is processed according to the information circulation policy) generated as a result of the policy judgment are shown in Fig. 4.

The policy file is also created in XML format. An example of the generated XML-format information circulation policy is shown in Fig. 5. In addition to user location information (e.g., address and latitude & longitude), it can describe environmental information (e.g., temperature and humidity). Finally, the transferred information is processed according to this information circulation policy and packed with it and sent to an information requester.

If a user sets privacy conditions that are unnecessarily restrictive, there is a risk that personal information will stop circulating, so he will be unable to

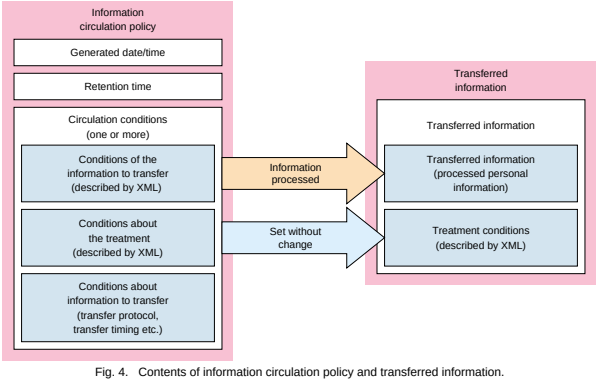

Fig. 4. Contents of information circulation policy and transferred information.

receive the services that he wants. To avoid that, a mediation function in the policy judgment mechanism notifies the user of this at the time of judgment, so he can decide whether to provide the required information. An example is shown in Fig. 6. In this case, the user does not permit his phone number and address to be circulated, but the information requester wants them. The policy judgment mechanism detects it and notifies the user. The mediation conditions are set up by the user.

3.4 Implementation of the prototype

We implemented a prototype based on the architecture described above. We assumed a service using location information of wireless LAN users and assumed that policy control is applied to the location of a wireless LAN access point and the authentication time acquired at the time of user terminal authentication. Specifically, the following were assumed.

- Users: People using the wireless LAN
- Information provider: The operator owning the wireless LAN access point and authentication server. It acquires the user's authentication information.
- Information requester: An operator asking the information provider for the user's policy-controlled authentication information
- Policy judgment mechanism: Mechanism executed by a neutral organization that holds and judges policies

We used an IEEE802.11b wireless LAN, which is popular now. The wireless LAN access protocol is IEEE802.1x [5] (EAP-TLS: Extensible Authentication Protocol-Transport Level Security [6]), which has high security. The protocol for cooperation between servers is SOAP (simple object access protocol), which has good affinity with XML. The system architecture is shown in Fig. 7. The system consists of the following components.

- User terminals: owned by the user and connected to the wireless LAN. In the prototype, we used commercial PCs and commercial 802.11b wireless LAN cards.
- Wireless LAN access points: owned by the information provider and connected with user terminals. The prototype uses commercial IEEE802.11b wireless LAN access points, which comply with IEEE802.1x.
- Authentication server: owned by the information provider. It authenticates users and acquires and sends location information about access points according to the information circulation policy. The prototype uses a commercial authentication server (compliant with IEEE802.1x), database, and SOAP communication software on a commercial workstation.

```
<processingPolicy>
  <generationTime>20030403142204</generationTime>
  <limitation>20030503142204</limitation>
  <processingPolicyPart>
    <targetForm>
      <directDescription>
        <tgt:target xmlns:tgt="http://xxx.ntt.co.jp/target">
          <contents>
            <locationInfo>
              <globalPosition tgt:essential="true">
                ...
              </globalPosition>
            </locationInfo>
            <environmentalInfo>
              ...
            </environmentalInfo>
          </contents>
        </tgt:target>
      </directDescription>
    </targetForm>
    <trt:treatment xmlns:trt="http://xxx.ntt.co.jp/treatment">
      <treatmentPart>
        ...
        <retention>
          <time-limited>20030431235900</time-limited>
        </retention>
      </treatmentPart>
    </trt:treatment>
    <act:action xmlns:act="http://xxx.ntt.co.jp/action">
      <trigger>
        <eventOccurred>Accounting-start</eventOccurred>
      </trigger>
      <protocol>xxx:soap</protocol>
      <destination>
        <address>http://xxx.xx.xx.xxx:7778/soap/</address>
        <name>urn:InformationUpdateService</name>
      </destination>
    </act:action>
  </processingPolicyPart>
</processingPolicy>
```

Generated date/time
Retention term

Conditions about information to be transferred (kinds of information, etc.)

Treatment conditions (retention time of personal information, etc.)

Conditions for transferring information (transfer timing etc.)

Fig. 5. Example of information circulation policy.

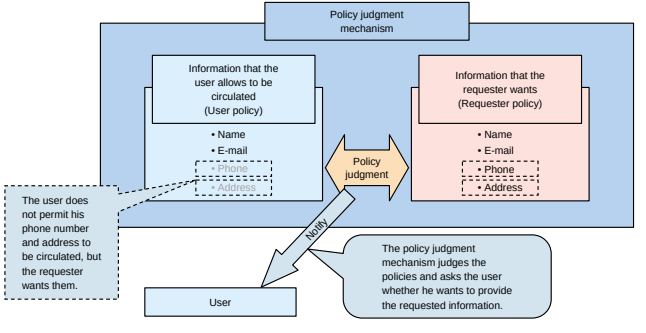

Fig. 6. Example of policy mediation.

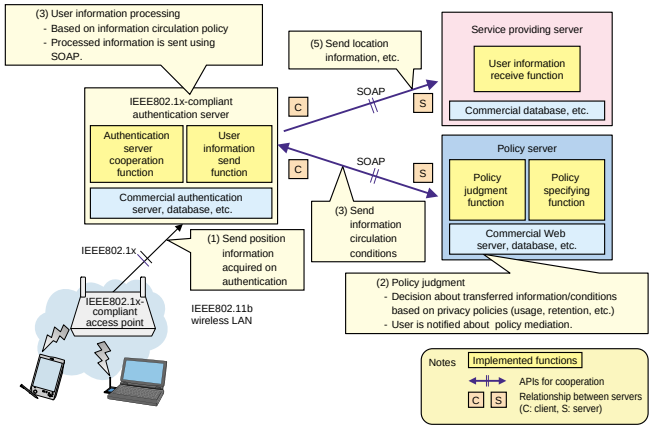

Fig. 7. Architecture of prototype system.

- Policy server: owned by the policy judgment organization. It performs policy judgment. The prototype uses a commercial Web server, database, and SOAP communication software on a commercial workstation.
- Service providing server: owned by the information requester. It receives personal information from the information provider (i.e., the authentication server) and provides services like a recommendation service or advertisement delivery service. The prototype uses a commercial database and SOAP communication software on a commercial workstation and provides a user information registration function. No service functions (recommending or advertising) are implemented in the prototype.

The operating procedure is as follows.

(1) In advance, the user, information provider (authentication server), and information requester (service providing server) register their own policies in the policy judgment server. Then, at the wireless LAN spot, the user performs 802.1x (EAP-TLS) authentication from a user terminal and the location of the access point is registered in the authentication server at the same time.
(2) The authentication server asks the policy judgment server whether it may pass on the location information and the policy server judges this based on the policies.
(3) The policy server returns its judgment (i.e., the personal information circulation policy) to the authentication server.
(4) The authentication server processes location information according to the circulation conditions.
(5) The authentication server transfers processed location information to the service providing server.

3.5 Considerations

The prototype revealed that personal information circulation could be controlled according to a privacy policy of a user of a wireless LAN hotspot.

In the prototyping, we were concerned about performance because two policies are used in P3P, but in our architecture, three policies are used and each policy is extended, so policy judgment seemed to be

heavier than in P3P. Performance affects the usability of a service, so in this prototype, we tuned the database input/output and speeded up the policy judgment processing. As a result, we obtained real-time policy processing.

The prototype handled dynamic information about a user's location, but static information (such as the user's residential address) can also be handled in this architecture. As mentioned in the Introduction, there is a recommendation service using purchasing history. In the same way, various services using personal information about users and information about their environment (such as temperature) are being considered. This architecture can be used for such services.

4. Conclusion

Our personal information circulation architecture has two main features.

(1) XML-based policy control, which can apply not only to the Web but also to sensor networks (out of the scope of P3P) and other applications.
(2) A policy judgment mechanism, which can control all the personal information circulation flow (e.g., circulation to third parties).

We described a prototype that controls location information about wireless LAN users and showed that policy judgment can be processed in real time. In this architecture, if the personal information is removed from the network or improperly separated from treatment conditions, it cannot be protected (we cannot detect the misuse or trace the abuser), so we must solve these security issues.

Our architecture can be applied to personalized services using personal information, so we will continue to study the issues involved and work toward the goal of practical services.

References

[1] Ministry of Public Management, Home Affairs, Posts and Telecommunications, Japan, "Information and Communications in Japan White Paper 2002," 2002.
[2] http://www.amazon.com/
[3] W3C, "The Platform for Privacy Preferences 1.0 (P3P1.0) Specification," 2002.
[4] G. Karjoth and M. Schunter, "The Platform for Enterprise Privacy Practices—Privacy-enabled Management of Customer Data," 2nd Workshop on Privacy Enhancing Technologies (PET 2002), pp. 69-84, San Francisco, U.S.A., Apr. 2002.
[5] IEEE, "IEEE802.1x," 2001.
[6] IETF, "RFC2716 EAP-TLS," 1999.

Minoru Sakuma
Engineer, Software Architecture Project, NTT Information Sharing Platform Laboratories.
He received the B.A. degree in environmental information and the M.A. degree in media and governance from Keio University, Tokyo in 1994 and 1996, respectively. In 1996, he joined NTT Network Service Systems Laboratories. After working on the development of intelligent networks, his recent research area is privacy control technologies. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan.

Daisuke Hamuro
Research Engineer, Software Architecture Project, NTT Information Sharing Platform Laboratories.
He received the B.S. and M.S. degrees in physics from Tokyo Institute of Technology, Tokyo in 1992 and 1994, respectively. In 1994, he joined NTT Network Service Systems Laboratories. His recent research area is network security and privacy control technologies. He is a member of IEICE.

Akinori Shiraga
Engineer, Software Architecture Project, NTT Information Sharing Platform Laboratories.
He received the B.E. and M.E. degrees in applied physics from the University of Tokyo, Tokyo in 2000 and 2002, respectively. In 2002, he joined the NTT Information Sharing Platform Laboratories. His recent research area is privacy control technologies.

Masayuki Kobayashi
Senior Research Engineer, Communication Platform SE Project, NTT Information Sharing Platform Laboratories.
He received the B.S. degree in electrical engineering from Shinshu University, Nagano in 1981. In 1981, he joined NTT Electrical Communication Laboratories, Nippon Telegraph and Telephone Public Corporation (now NTT) and researched privacy control technologies. He is a member of IEICE.