Architecture and Implementation of CC/PP Harmonization with P3P based on Web Privacy

Fumihiro Kato
Faculty of Environmental Information, Keio University
5322 Endoh, Fujisawa, Kanagawa 252-8520, Japan
torry@tom.sfc.keio.ac.jp

Wataru Okada
Faculty of Environmental Information, Keio University
5322 Endoh, Fujisawa, Kanagawa 252-8520, Japan
wataru@slab.sfc.keio.ac.jp

Kazuhiro Kitagawa
W3C Keio
5322 Endoh, Fujisawa, Kanagawa 252-8520, Japan
kaz@w3.org

ABSTRACT
We propose the architecture that harmonizes CC/PP with P3P to protect user's privacy on Web Sites. CC/PP is a proposed specification of a framework to adapt Web contents with a variety of device capabilities and user preferences. CC/PP does not handle privacy issues by itself in spite that CC/PP manages personally identifiable information. We have designed the architecture that CC/PP can protect privacy of personally identifiable information by harmonizing with P3P. P3P is a proposed specification for users to gain more control over personal information on Web sites. To demonstrate and test our idea, we have implemented a browser and a server which are compatible with CC/PP and P3P. We believe that P3P can complement the security aspect of CC/PP.

Keywords
CC/PP, P3P, HTTP Extension Framework, Privacy

1. INTRODUCTION
CC/PP (Composite Capabilities/Preference Profiles) [1, 2, 3, 4] and CC/PP Exchange Protocol [5, 6] are proposed specifications of content negotiation framework to customize and adapt Web contents with a variety of device capabilities and user preferences by W3C and IETF. CC/PP describes a device capabilities profile of Web client using RDF [7]. A client has already gotten the default device information beforehand, and a client exchanges the only modified part with a server or a proxy on HTTP Extension Framework [8]. However, CC/PP does not describe the trust model for exchanging privacy sensitive profile information between a client and a server in spite that CC/PP manages personally identifiable information. In short, it is expected to protect privacy on CC/PP.

It is very important to protect privacy of personally identifiable information. The Web community also has recognized that privacy issues are a fundamental part of the Web. We believe that P3P (The Platform for Privacy Preferences) [9] takes in charge of this part.

P3P is a proposed specification for users to gain more control over their personal information on Web sites by W3C. P3P describes policy files of Web sites using XML, and a client automatically retrieves and interprets the policy files from the Web site when users visit there. Then, it makes a comparison between the policy of the Web site and the user. In addition, P3P policy file defines only the information which Web sites need. Therefore, the information to be exchanged between a client and a server is only the minimum requirements. We believe that P3P can complement the security aspect of CC/PP.

2. ARCHITECTURE
In this section, we describe the trust model using CC/PP and P3P. Although CC/PP is protocol independent, we assume using HTTP protocol.

2.1 Use Case
Here, we explain a simple scenario of our harmonization.

1. When a user visits the Web site, user's client retrieves a P3P policy reference file from well known location of P3P.
2. If there is a P3P policy reference file in location other than well known location, the Web server sends a P3P header to the client. Then, the client interprets the P3P header, and gets a P3P policy reference file from that location written in the P3P header.
3. The client interprets the P3P policy reference file, retrieves and interprets any P3P policy files.
4. The client makes a comparison between the Web site's policy and the user's one, recognizes what kind of information the server wants to get and decides which information the client may send to the server.
5. The client collects all the information which should be sent to the server and converts them to CC/PP-Diff Header.
6. The client conveys CC/PP-Diff Header to the server.
7. The server resolves and retrieves default profiles from CC/PP repository, and merges received CC/PP-Diff Headers and them.
8. The server adapts contents or services to the client and transfers it to the client.

2.2 Trust Model
With current CC/PP, a user cannot specify the information which should be sent or not. Due to lack of privacy protection mechanism of CC/PP, a client conveys all information to a server.

Our architecture allows that a client conveys the only necessary information which are approved by a user to protect

user's privacy. It also ensures that a client does not send their information if a user do not approve sending them. Furthermore, it assures that a client can understand what purpose a server needs the information for.

It guarantees that a server receives its only necessary information. The contents and services that each server provides are usually different, and the information which each server needs is naturally different.

We do not describe the security issues of the communication channel, such as encryption. It may be covered by SSL or TLS, though it is beyond the scope of this paper.

2.3 Requirements

A client must understand P3P, make CC/PP-Diff Header and be compatible with HTTP Extension Framework.

A server must understand CC/PP, adapt contents, send P3P header and be compatible with HTTP Extension Framework. Nevertheless, sending P3P header is not mandatory in our architecture. Because a server does not have to send P3P header if there is a P3P policy reference file in well known location of P3P. Please note that it is not P3P1.0 specification compliant.

3. IMPLEMENTATION

In this section, we describe implementation of our client and server based on this architecture to demonstrate briefly.

3.1 Client

We have developed a client which is a revised version of "PANDA" [10], written in C with GTK and Libwww[13]. This is a experimental system and it doesn't support all of CC/PP and P3P functionality. We can currently use only optional method on HTTP Extension Framework, and we can not use mandatory method because HTTP Extension Framework has not been in Libwww. However, it is sufficient for CC/PP to use only optional method.

3.2 Server

The server was developed based on Apache server. CC/PP and HTTP Extension Framework functionality were added to Apache as its module written in Perl. We did not implement mandatory method of HTTP Extension Framework as well as client side.

4. DISCUSSION

CC/PP is an important framework to change how to provide the information of each device and to adapt content in appropriate format. Moreover, by harmonizing CC/PP with P3P, our architecture can automatically manage the information of the privacy.

As a result of our architecture, it was easy to develop our client and server.

To demonstrate our idea, we are also developing Location Base Service with GPS and IEEE802.11b using our proposed harmonization to protect privacy of personal information. In addition, we will implement our architecture on various devices, for example PDA and cellular phone with i-appli and BREW.

We noticed one problem when we implemented our architecture. The format of the description and the transfer method are different between CC/PP and P3P.

- CC/PP: RDF + HTTP Extension
- P3P : XML + HTTP header

It seems that P3P will prepare RDF data models representing P3P policies and policy reference files. However, P3P header is not sent via HTTP Extension yet. We believe that it should be integrated with HTTP Extension.

5. ACKNOWLEDGEMENTS

We thank Kinuko Yasuda for developing the first version of PANDA. We also thank Nobuo Saito, Tatsuya Hagino, and Norio Toyama for their help in improving our idea. We would like to thank Hidetaka Ohto, W3C CC/PP WG member who inspired us with this harmonization. Special thanks to NTT Docomo for providing financial support.

6. REFERENCES

[1] Mikael Nilsson, Johan Hjelm, Hidetaka Ohto. Composite Capabilities/Preference Profiles: Requirements and Architecture. W3C Working Draft 21 July 2000. http://www.w3.org/TR/CCPP-ra/

[2] Mikael Nilsson. Composite Capabilities/Preference Profiles: Terminology and Abbreviations. W3C Working Draft 21 July 2000. http://www.w3.org/TR/CCPP-ta/

[3] Franklin Renolds, Chris Woodrow, Hidetaka Ohto. Composite Capability/Preference Profiles (CC/PP): Structure. W3C Working Draft 21 July 2000. http://www.w3.org/TR/CCPP-struct/

[4] Graham Klyne. CC/PP Attribute Vocabularies. W3C Working Draft 21 July 2000. http://www.w3.org/TR/CCPP-vocab/

[5] Hidetaka Ohto, Johan Hjelm. CC/PP Exchange Protocol based on HTTP Extension Framework. W3C Note 24 Jun 1999. http://www.w3.org/1999/06/NOTE-CCPPexchange-19990624

[6] Johan Hjelm. Content Negotiation Header in HTTP Scenarios. IETF Internet draft 31 October 2000.

[7] Ora Lassila, Ralph R. Swick. Resource Description Framework, (RDF) Model and Syntax Specification. W3C Recommendation 22 February 1999. http://www.w3.org/TR/REC-rdf-syntax

[8] H. Nielsen, S. Lawrence. An HTTP Extension Framework. IETF RFC 2774 February 2000.

[9] Lorrie Cranor, Marc Langheinrich, Massimo Marchiori, Martin Presler-Marshall, Joseph Reagle. Platform for Privacy Preferences(P3P1.0) Syntax Specification. W3C Candidate Recommendation 15 December 2000. http://www.w3.org/TR/P3P

[10] Kinuko Yasuda, Takuya Asada, and Tatsuya Hagino. Effects and Performance of Content Negotiation Based on CC/PP. In Proceedings of MDM'2001 Jan 2001.

[11] Henrik Frystyk Nielsen. Libwww - the W3C Protocol Library. http://www.w3.org/Library/