P3P Privacy Enhancing Agent

Hsu-Hui Lee
Computer Science
San José State University
One Washington Square
San José, CA 95192 USA
kenny_lee@sjsualum.org

Mark Stamp
Computer Science
San José State University
One Washington Square
San José, CA 95192 USA
stamp@cs.sjsu.edu

ABSTRACT
Protecting personal privacy information is an inherently difficult problem. Privacy enhancing agents are software agents that help web users to protect their private information by collecting web site P3P [1] information and exchanging knowledge of web site privacy practices with other agents. Based on the information collected, the software agents begin the decision-making and negotiation process. A suggestion or warning is presented to the user if any discrepancy is detected [4]. In this paper, we outline our design for a P3P privacy enhancing agent (PEA).

Categories and Subject Descriptors
H.3.5 [**Software**]: Online Information Services – W*eb-based services, Data Sharing.*

General Terms
Algorithms, Management, Security, Verification.

Keywords
Privacy, Personal Identifiable Information (PII), Platform for Privacy Preference (P3P), Software Agent, Ontology.

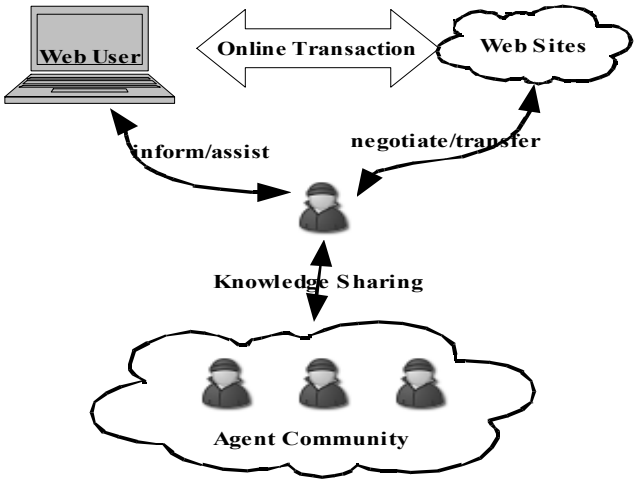
1. INTRODUCTION
The Web has become a major marketplace where people buy and sell goods and information. According to a recent report, "More than one billion people in the world have access to the Internet, a quarter of them with broadband or high-speed connections". However, despite the continued rapid growth of the Internet, e-commerce growth actually appears to be leveling off. From a recent study on U.S. retail e-commerce sales growth, "eMarketer estimates that retail e-commerce sales will increase an average 18.6% per year between 2005 and 2009—that's strong growth, but still a downturn from the 26% annual rate seen between 2001 and 2005" [2]. Of course, many factors might contribute to this relative downturn in the sales growth rate. However, concerns over the collection and use of personal identity information (PII) and security and privacy on the Web in general, are certainly a negative factor.

We have developed a Privacy Enhancing Agent (PEA) to address the privacy concerns from the client side by helping users analyzing web site P3P policies. The primary tasks of our PEA agents are (1) P3P policy retrieval (2) User privacy preference enforcement (i.e., matching web site P3P policies with user preferences) (3) Decision making and negotiation processes based on user preferences, transaction history with the particular web site and related privacy practice knowledge from other agents.

PEA is implemented as a software agent built on top of the Java Agent DEvelopment Framework (JADE) [3]. PEA is able to take appropriate actions to assist users and to enhance the user experience by protecting privacy in a transparent way, which should increase user confidence that PII is being handled properly. An additional benefit is that PEA enables users to navigate efficiently since it handles all routine PII requests. PEA should increase consumer confidence in online transactions, at least with respect to PII issues.

SWS'06, November 3, 2006, Alexandria, Virginia, USA.
ACM 1-59593-546-0/06/0011.

2. ARCHITECTURE OVERVIEW



The PEA agent has access to its owner's personal privacy preference and to the virtual agent community to which the agent is affiliated. Armed with this information and the user's transaction history, the PEA agent monitors the user's Web transactions and follows the user's privacy preferences if a transaction involves any privacy information exchange.

2.1 SYSTEM COMPONENTS
The main components of our PEA system are: the JADE framework, a traffic monitor agent, a P3P knowledge module, a negotiation module, an archive module, a user interface module and a mobility module. Each component has the following function within the PEA system: (1) The JADE framework is the foundation of the PEA agent system. Detailed information about JADE can be found at http://jade.tilab.com. (2) The traffic monitor agent monitors Web traffic to and from a user's Web browser and notifies the system when any privacy information transaction is detected. (3) The P3P knowledge module has knowledge about P3P policy specifications and it has the responsibility to translate any well-formed P3P privacy information into the system ontology [5], which can be understood by any PEA agent. (4) The negotiation module is responsible for analyzing and matching a web site's privacy policy with the user's privacy policy rules and constraints. (5) The archive module is a book-keeping agent which manages all

transaction history and relevant privacy policy files. (6) The user interface module is responsible for informing and alerting the user regarding privacy related issues, and for collecting user input. (7) The mobility module acts as a liaison for PEA system. Its purpose is to migrate to another PEA system to retrieve knowledge of privacy policy practice relevant of a specific Web site.

3. PRIVACY POLICY TRANSACTION

The PEA agents manage and protect a user's privacy information by supporting the HTTP and P3P protocols. To illustrate the process, consider a simple P3P transaction flow. The transaction starts with a user requesting for a web page. The traffic monitor agent intercepts the response from the web server and looks for P3P policy information. If a P3P policy is found, the monitor agent retrieves the policy file and passes it to the knowledge engine, which is composed of P3P knowledge agents, negotiation agents and other support modules. The knowledge engine processes the P3P policy based on privacy rules (see Section 4) and either completes the transaction—transparently from the user's perspective—or requests user intervention.

4. PRIVACY POLICY NEGOTIATION

In PEA, negotiations are based on the idea of defining constraints on the information sets. The constraints are used to specify the conditions under which personal information can be accessed. A P3P privacy preference policy contains a request to gain access to a particular set of information. PEA checks the request and verifies whether the request satisfies all of the user-defined constraints. If the constraints are satisfied, access to the information is granted. Otherwise, the request is rejected or a negotiation process can be used to try to find an alternative solution based on additional historical information or user input.

4.1 INFORMATION, RULES and FACTS

Now we define the basic terms, information, rules and facts used by the PEA system in the negotiation process.

4.1.1 Personal Information

Personal information consists of a set $\mathbf{I} = \{d_1,d_2,d_3,...,d_n\}$ where **I** is a finite set of personal information and each d_i, for $1 \le i \le n$, is a data element.

4.1.2 Rules

A rule r is used to define the specific circumstances under which an information set can be accessed. A rule is of the form

$r = (D_r, C_r)$ where $D_r \subseteq I$ and $C_r = \{c_1,c_2,.....,c_n\}$

where r is a pair (D_r,C_r) and D_r denotes a set of personal information from **I**, and C_r denotes a set of constraints on D_r. Each constraint $\mathbf{c_i}$, for $1 \le i \le n$, represents a particular condition which must be met in order to satisfy rule r.

4.1.3 Facts

Facts f are used to describe a request for a set of data and the conditions under which the request is made. Then

$f = (D_f, V_f)$ where $D_f \neq \varnothing$ and $V_f = \{p_1,p_2,.....,p_n\}$

where a fact f is defined by a pair (D_f, V_f) where D_f denotes the requested data and V_f denotes the conditions under which the data D_f is requested. The conditions specified by V_f contain name-value pairs of the form $p_i = (n_i,v_i)$, for $1 \le i \le n$, where n_i denotes the name of an item and v_i denotes the value of the item.

4.1.4 Collection of Facts

The collection of facts relevant to a particular request is denoted

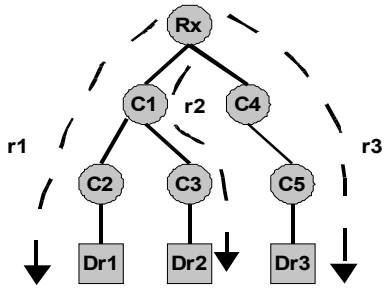
$F = \{f_1,f_2,...,f_n\}$ where each f_i, for $1 \le i \le n$, is a fact, as defined in previous section.

4.1.5 Rule Set

A rule set is denoted $R = \{r_1,r_2,...,r_n\}$ where each r_i, for $1 \le i \le n$, is a rule, as defined above.

4.2 TREE REPRESTATION of A RULESET

According to the definition in Section 4.1.2, a rule is a pair that defines constraints on a set of data. To represent a rule set, as per the definition in Section 4.1.5, we will use a tree structure, which will be very useful when we evaluate rules and match facts with rules. We begin with a simple rule set, say, R_x. Suppose the access to three different sets of information is controlled by three rules of the form $r_1 = (D_{r1},\{c_1,c_2\})$, $r_2 = (D_{r2},\{c_1,c_3\})$, $r_3 = (D_{r3},\{c_4,c_5\})$, where $R_x = \{ r_1 , r_2 , r_3 \}$. Note that both r_1 and r_2 contain the same constraint c_1. For this example, the rule set tree for R_x is given in the figure below:



Each rule is represented by a path from the root node R_x to the leaf node representing the data element D_{ri}, for i = 1,2,3. With this tree representation, we can easily see the constraints that lie between the root node of the rule set and the data elements, which are the leaf nodes.

Based on such a tree structure, rule evaluation is simply a depth-first search on the ruleset tree.

5. CONCLUSION

We believe that PEA can be a valuable asset to users. PEA will enable a user to manage, negotiate and analyze personal privacy information on the World Wide Web. In this way, PEA can make the Web safer and more secure with respect to the management of personal private information.

6. REFERENCES

[1] World Wide Web Consortium (W3C), Platform for Privacy Preferences Project – P3P, W3C, Cambridge, MA, http://www.w3.org/p3p

[2] eMarketer (2006 June), "US Retail E-Commerce", http://www.emarketer.com/Report.aspx?ecom_us_jun06

[3] JADE (Java Agent Development Framework) , http://Worldjade.tilab.com/

[4] Mark S. Ackerman, Lorrie Cranor, "*Privacy Critics: UI Components to Safeguard Users' Privacy*" , ACM Press, May 1999

[5] Giovanni Caire, David Cabanillas, "*Application-Defined Content Languages and Ontologies*", TILab S.p.A., Nov. 2004.