A Pervasive P3P Negotiation Mechanism for Robust Ubiquitous services

Ohbyung Kwon

School of International Management, Kyunghee University
1, Seochun-Dong, Ghyheung-Gu, Yongin, 446-701, Korea
Tel:+ 82-31-201-2306, Fax: + 82-31-204-8113, E-mail: obkwon@khu.ac.kr

Abstract

Only a few P3P-based privacy aware systems address the discrepancy between a service provider's privacy policy and the user's typical concerns—hence, putting service usage at risk. Moreover, since users are typically nomadic in pervasive computing services, their specific privacy concerns would dynamically change according to the surrounding context. This leads us to develop a dynamically adjusting P3P-based policy for a personalized, privacy-aware service as a core element of secure pervasive computing. Hence, the purpose of this paper is to propose a pervasive P3P-based negotiation mechanism for privacy control which functions in a dynamic and flexible way.

Keywords:

P3P, Pervasive Computing, Ubiquitous Shopping

Introduction

To fully realize ubiquitous mobile computing's potential, applications designers must incorporate users' personal privacy preferences [4]. To protect users' privacy in ubiquitous or pervasive computing settings, the Privacy Profile Negotiation Protocol (PPNP) has been proposed in the literature. PPNP, initiated by Tohda Lab at Keio University in Japan, manages users' privacy profiles to prohibit transferring profiles to any but the trusted services. In a similar effort, ETH Zurich has been developing a privacy aware system (pawS) which applies P3P.

When users register a service, they must typically submit personal information, which relates directly to their privacy concerns. Because of these concerns, users reluctantly or warily provide their personal information, or may choose to deny getting service. On the other side, service providers need users' personal information to be as rich as possible with less personal information to help differentiate different types of customers. Figure 1 shows the discrepancy in available information between service provider and the user. Ideally, it would be possible to increase the allowable personal information—complying with the service provider's request—while at the same time increasing the users' comfort both by minimizing service providers requiring excessive personal information, and users having to submit other personal information which is of no value to the service provider.

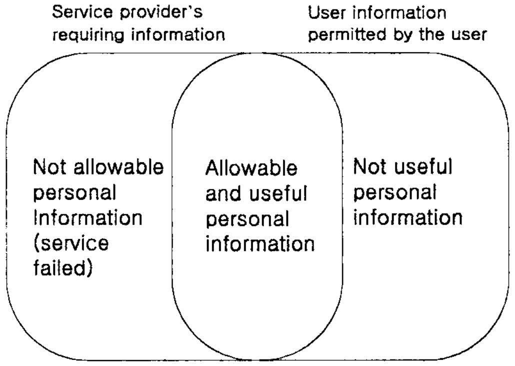

Service provider's requiring information
User information permitted by the user
Not allowable personal Information (service failed)
Allowable and useful personal information
Not useful personal information

Figure 1 - Personal Information Discrepancy

To automatically share which is useful personal information or not with service providers and users, the Platform for Privacy Preferences (P3P) is known as one of the most significant efforts currently underway to enable web-based service users to gain control over their private information. P3P provides a way for a Web site to encode its data-collection and data-use practices in a machine-readable XML format known as a P3P policy [5]. However, so far efforts for balancing privacy protection and service quality have been proven to be ineffective in current techniques including P3P [8].

Hence, this paper aims to propose the concept of P4P (Pervasive Platform for Privacy Preferences) over P3P, and to develop a P4P-based negotiation methodology for privacy-aware pervasive computing services. The methodology includes mutual privacy policy, context-aware policy design, and personalization. Especially, the methodology considers user's current location and personal profile with preferences as external context and internal context, respectively.

Platform for Privacy Preferences

In recent years, privacy protection has been one of the most active and spotlighted issues in electronic business. Accordingly, developing solutions to these privacy concerns that are both technically and socially secured becomes important. For example, a great number of commercial websites provide a privacy policy, because these sites do require personal information such as name, e-mail address, certain preferences, and even a social security number (SSN). So far, specifying a site's privacy policy to inform the users before they register for services has been known as sound way to decrease users' privacy concerns.

The design of P3P is partially originated from the code of ethics for user agents. Based on the CMA code of ethics, a user agent should follow a sort of the code of ethics as listed in Table 1 [10].

Table 1 - Code of ethics for user agents

Category	Code of ethics
Notice and communication	Provide mechanisms for displaying a service's information practices to users. Provide users an option that allows them to easily preview, and agree to or reject each transfer of personal information that the user agent facilitates. Not be configured by default to transfer personal information to a service provider without the user's consent. Inform users about the privacy-related options offered by the user agent.
Choice and control	Include configuration tools that allow users to customize their preferences. Allow users to import and customize P3P preferences from trusted parties. Present configuration options to users in a way that is neutral or biased towards privacy.
Fairness and integrity	Be usable without requiring the user to store user personal information as part of the installation or configuration process. Act only on behalf of the user according to the preferences specified by the user. Accurately represent the practices of the service provider.
Security	Provide mechanisms for protecting the personal information that users store in any data repositories maintained by the agent. Use appropriate trusted protocols for the secure transmission of data. Warn users when an insecure transport mechanism is being used.

To specify a privacy policy in a complete and standardized manner, a sort of policy specification language has emerged such as EPAL and P3P. Among these, the P3P specification defines the syntax and semantics of P3P-based privacy policies. Since the specifications are in a machine-readable format, user agents can understand P3P specifications and then automate decision-making on behalf of their users based on the specifications when appropriate, so that users need not read the privacy policies at every site they visit. The user agent is a kind of program whose purpose is to mediate interactions with services on behalf of the user under the user's preferences. The user agents can be built as an e-wallet or any user data management tools.

P3P provides a way for a Web site to encode its data-collection and data-use practices in a machine-readable XML format known as a P3P policy [5]. With the help of these P3P specifications, Web users can more easily understand what data will be transferred to the web site when they visit. To do so, the P3P specifications include:

- A standard schema for data that a Web site may wish to collect, known as the "P3P base data schema."
- A standard set of uses, recipients, data categories, and other privacy disclosures.
- An XML format for expressing a privacy policy.
- A means of associating privacy policies with Web pages or sites, and cookies.
- A mechanism for transporting P3P policies over http.

Various privacy-aware ubiquitous or pervasive systems based on P3P have been proposed over the last decade. Ackeman proposed a technical mechanism to inform users regarding data requests and their consequences [1]. A privacy control module could be embedded in the privacy-aware systems as middleware [6]. The Personal Context Agent Networking (PeCAN) knowledge architecture consists of both client-side and web-side architectural data components and services, which inform the user of online privacy and trust within e-commerce tasks [7]. Representative commercial examples that use a P3P policy include Microsoft's Explorer Ver6.0 and AT&T's Privacy Bird.

However, current research studies seldom address methods to resolve privacy concerns between service providers and users in a pervasive computing environment. Actually current P3P is being discussed for better and complete specification and use [2]. Hence, a privacy-aware service in a pervasive computing environment for nomadic users would require an amended and tailored P3P specifications.

Pervasive Platform for Privacy Preferences (P4P)

P4P is an extension of a conventional P3P that additionally considers specifications useful for context-related privacy control. In representing data elements in P3P specifications, a tree structure is applied. For example, a data element, vehicle.model, is a child of a data element vehicle. Even though the current P3P specification does not consider contextual data, we can extend the P3P specifications to dynamically changing data elements as the same notation system. To represent these dynamically changing data elements, we suggest the following method:

P3P data schema.**CONTEXT**.*context field.*

For instance, if one needs to discern GPS position data with both public activity and private activity, then the data element in P4P is represented as:
user.current.GPS_position. **Activity**.public and
user.current.GPS_position. **Activity**.private.

To deliver the context data element to the service providers, the context information could be acquired from either a personal context ontology or a user agent. The context model with category and property applied in P4P is listed in Table 2.

Table 2 - Proposed context model

Context category	Property	Examples
Activity	Public Moderate Private	Current schedule
Location	Public Moderate Private	Current location
Social	Public Moderate Private	Nearby person
ComputationalEntity	Public Moderate Private	Device, network, TV channel, etc.
Physical environment	Public Moderate Private	Temperature, Climate, etc.

Meahwhile, the service provider's privacy preferences are necessary, as well as the user's privacy preferences, for the service agent to be able to negotiate with user agents—aiming for a consensus as to what extent information should be provided to the other side. In comparison with the users, service providers also prefer the users of high reputations and hence might have more value. Hence, we can say that the service providers certainly have their own privacy concerns to protect information embedded in the services: *e.g.,* company profiles, product information, service profiles, and even employee private data. These mainly make privacy preferences of the service provider slightly difference from those of the clients, as already considered in P3P. Consequently, the privacy preference purposes and categories of the service provider are shown in Table 3.

P4P Negotiations

Pervasive P3P (P4P) and negotiation mechanisms based on P4P are developed and evaluated, with the goal to increase the possibility of service match by decreasing the gap between the information that service providers require, and personal information which the user approves sharing. Since P4P basically assumes a nomadic user and a pervasive service running in any space, contextual data elements are considered in the specification so that the service may keep track of the user's data that may be dynamically changing. Based on these dynamically changing privacy preferences, negotiation about what information should be passed and to what extent opened to the other side is a core mechanism of the ideas addressed in this paper.

Areas of future research include aspects as follows:

- Full-fledged development of P4P-based privacy aware pervasive systems
- Implementation of the negotiation mechanism
- Provision of complete P4P categories, purposes and retentions
- Evaluation of the negotiation mechanism to reach at optimal performance

At the moment we plan to extend the presented mechanism so that it may be available in legacy privacy-aware pervasive systems.

Privacy concerns arise when there is tension for an individual between the gains earned by being accessed and the needs to hide the personal information [9]. In this paper, the economics-based model of user privacy is adopted for the negotiation method. The utility is basically computed by the trade-offs between the cost of lack of services and the cost of surveillance: delivering one's private information to the other side. It has been often observed in the context of e-commerce that privacy preferences and actual behavior have a trade-off relationship, and these cause complications to the users [3].

As seen in Figure 2, for a data element *i*, the costs of surveillance, $y = \alpha_{2,i} e^{\beta_{2,i} c_i}$, and of lack of service, $y = \alpha_{1,i} e^{-\beta_{1,i} c_i}$, depend on the level of surveillance, c_i.

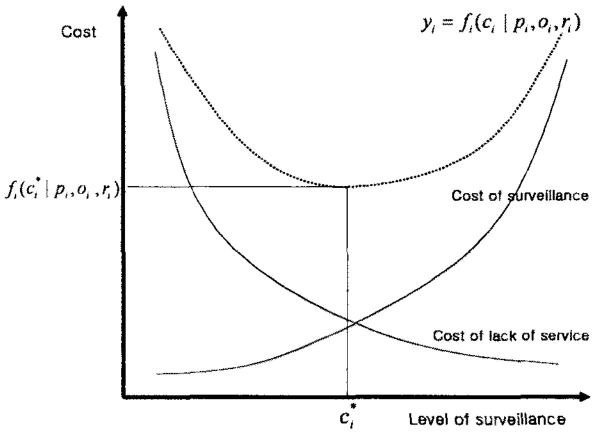


Figure 2 - Trade-off Relationship for Negotiation

The negotiating method is represented as follows:

STEP 1: The user agent gets the service provider's P3P policy. The P3P policy *P* is represented as:

$$P = \{(c_1, p_1, o_1, r_1), \ldots, (c_i, p_i, o_i, r_i), \ldots, (c_N, p_N, o_N, r_N)\} \text{ --- } \mathbf{(1)}$$

where N is the number of data elements included in the P3P policy, $c_i \in C$, $p_i \in P, o_i \in O = \{always, opt-in, opt-out\}, r_i \in R$, where C, P, O and R indicates a set of categories, purposes, options and retentions, respectively.

STEP 2: The user agent produces an optimal solution by minimizing the total cost. The total cost of the user's side is represented as (2):

$$TC_U = Max\{f_1(c_1^* \mid p_1,o_1,r_1), f_2(c_2^* \mid p_2,o_2,r_2),\ldots, f_N(c_N^* \mid p_N,o_N,r_N)\} \quad \text{------------ (2)}$$

where c_i indicates the *i*th category.

For all i, $f_i(c_i^* \mid p_i,o_i,r_i) = \alpha_{1,i}e^{-\beta_{1,i}c_i} + \alpha_{2,i}e^{\beta_{2,i}c_i}$, since the cost of lack of service is $y = \alpha_{1,i}e^{-\beta_{1,i}c_i}$ and the cost of surveillance is $y = \alpha_{2,i}e^{\beta_{2,i}c_i}$. Hence, the optimal level of surveillance of *i*th data element is derived as (3):

$$c_i^* = \frac{log\dfrac{\alpha_2\beta_2}{\alpha_1\beta_1}}{\beta_1 - \beta_2}, \text{ if } \beta_1 \neq \beta_2 \quad \text{---------- (3)}$$

STEP 3: Suppose that $TC_U = f_M(c_M^* \mid p_M,o_M,r_M), 1 \leq M \leq N$. Then the optimal set of user preferences, $\{c_1^*, c_2^*, \ldots c_i^*, \ldots c_N^* \mid p_1,\ldots,p_N,o_1,\ldots,o_N,r_1,\ldots,r_N\}$, is passed to the Negotiator, so that the Negotiator may compare the user's preferences to the service agent's preferences.

STEP 4: For any category *i*, service agent sets $c_i = c_i^*$. Then the total cost of the service provider is represented as (4):

$$TC_S = Max\{g_1(p_1^*,o_1^*,r_1^* \mid c_1),\ldots,g_j(p_j^*,o_j^*,r_j^* \mid c_j),\ldots,g_N(p_N^*,o_N^*,r_N^* \mid c_N)\} \quad \text{--------------(4)}$$

STEP 5: The optimal set of the service provider's optimal preference is represented as (5):

$$\{p_1^*,\ldots,p_j^*,\ldots,p_N^*,o_1^*,\ldots,o_j^*,\ldots,o_N^*,r_1^*,\ldots,r_j^*,\ldots,r_N^* \mid c_1,\ldots,c_i,\ldots,c_N\} \quad \text{--- (5)}$$

Then (5) is passed to the Negotiator.

Implementation

P4P is actually considered in COEX (Convention & Exhibition), Korea's largest shopping mall, and the largest retail and entertainment complex in Asia, to examine how P4P works. The COEX mall, opened in 1979 ISO 9001 certified in 2004, has a vision as a representative space of experiencing Korea in the 21st century. COEX aims to provide young people with entertainment and a culture space, and to provide domestic or foreign visitors with convention and exhibition spaces. It is the largest underground shopping space in Asia, with over 200 stores and a total facilities area of up to 118,000 square meters, about 14.5 times of the space of standard Olympic main stadium.

To explain how our P4P-based negotiation system works, let's suppose a proactive pervasive computing service called, MyEntrance. MyEntrance is tested in the shopping mall as follows:

"Joseph has an RF-tag which contains data about his ID. He enters into a shopping mall service zone though a revolving door. When he passes through the revolving door, an RF-reader attached to the door attempts to read the data of Joseph's RF-tag. The MyEntrance service shows personalized events and ads in the shopping mall to the users. The service requires the user's ID, and optionally his phone number and home address. Among those, the user ID will be used for tailored service and user preference. A phone number and home address will be inputted for pseudo-decision. The user ID will not be retained and the other data will be stored for a certain period. These privacy policies are already predefined in the service provider's P4P file. Then the user agent on Joseph's behalf accesses the P4P file and finds that Joseph does not want to share his phone number and home address. To resolve these conflicts, a negotiator in the zone communicates with both the user and service agents to dynamically formulate a P4P file only for Joseph and specifically applied at that time. Consequently, the user and service agents reach an accord to provide user preference, rather than phone number and home address, so that Joseph may get served. As Joseph approaches the kiosk that announces events and ads, the service agent is ready to provide Joseph with some detailed product information. However, since the service agent wonders if Joseph may reuse the content in an unwanted way, the service agent queries the user's privacy policy."

Using the P3P files from both sides, the MyEntrance service as a P4P-aware pervasive service will be performed as shown in Figure 3.

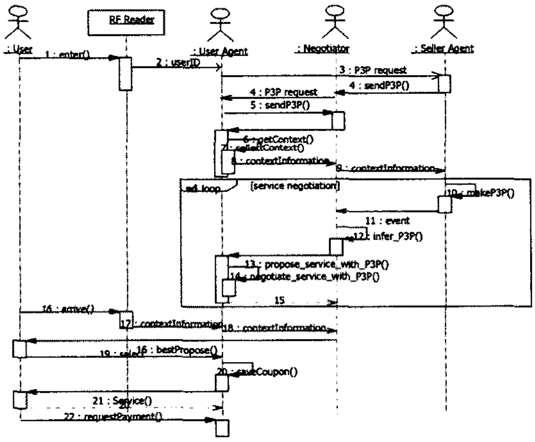
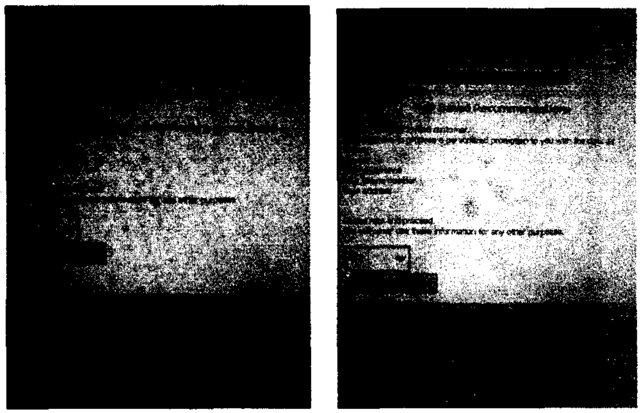


Figure 3 - Sequence Diagram of MyEntrance Service

As shown in Figure 4(a) and Figure 4(b), the same buyer could get different suggestion according to the sellers. Woori Bank suggests using user preference and phone number to go to P4P-aware recommendation, while F. Café requires all data elements. These happen mainly due to the different recognition of the user's reputation value and the negotiation strategy. Or source credibility theory can support reputation mechanisms could be applied.



(a) Sample result by shop A (b) Sample result by shop B

Figure 4 – P4P Negotiation Results

Experiments

To show the technical and operational viability of the negotiation mechanism proposed in this paper, an experimental evaluation was carefully conducted. The experiments were conducted in an IBM Pentium PC with one 1.8G processor, 1Gbytes of RAM, and 120Gbytes hard disk.

The Effect of Negotiation Mechanism on Service Utilization

One of the major issues regarding privacy is to increase the rate of service utilization while addressing users' privacy concerns. It was mainly caused by the current limitation of the web-based systems in that the users cannot selectively submit the data items to get served. Even in cases of P3P-based privacy aware systems, only the service provider determines the content of the P3P. Hence, we need to examine if our negotiation mechanism is significantly useful to increase the rate of service utilization. To do so, a simulation approach was used for the experiment. According to the case service, MyEntrance, addressed in the previous section, UserID, user preference, phone number, home address, product information and service profile are considered for the negotiation:

To acquire the optimal level of surveillance for each data element, the value of $\alpha_1, \beta_1, \alpha_2$, and β_2 are set by random number generation. Then the cost of surveillance and the cost of lack of service are derived to automatically get the total cost and the optimal level of surveillance for each data element. If all of the optimal solutions of the data elements are collected, then final cost is derived by the selection of maximum total cost among the data elements. Then using a threshold value, a final decision whether the privacy policy could be accepted or not is made. If the user does not accept the policy, then the negotiation system informs the service provider to ask the possibility of the rule relaxation. For this decision, the user's reputation value is adopted. For our simulation approach, the reputation value is computed randomly using the random number generation function in Microsoft Excel 2003. According to the user's reputation value, to what extent the relaxation should be performed is determined: relaxing the purpose, retention, or even dropping the data element from the list. Then the temporally updated P3P is resent to the negotiator for the next round. To simplify our experiment, only two rounds were allowed for the relaxation. If more relaxations are allowed, the rate of service utilization will be increased than the outcome of this experiment. We performed 900 runs to get the results.

Table 3 is given to test the hypothesis that the negotiation mechanism is significant to increase the rate of service utilization as a main performance. Hence, the null hypothesis provides that the performance will yield equal results. If the *p-value* is greater than 0.05 or 0.01, then the null hypothesis cannot be rejected statistically. Based on such a principle, we can conclude that the statistical test results for the hypothesis indicates that the null hypothesis is strongly rejected statistically: less than one percent significance levels. We deduce that the negotiation mechanism works better in terms of rate of service utilization. Hence, the negotiation mechanism positively affects the usage of privacy-aware services.

Table 3 - Results of Statistical Test – One-way ANOVA (1)

Performance Measures	Mean rate of service utilization		Difference	*p-value*
	Without negotiation	With negotiation		
Rate of service utilization	25.63%	53.82%	28.2%	0.0000 ***

*** p<0.01

The Effect of Being Aware of Reputation on Service Utilization

We also conducted another experiment to examine to what extent considering the user characteristics in terms of reputation influences the performance of the negotiation mechanism. First, to identify how the reputation affects in service utilization, the mean rate of service utilization was compared between negotiation both with and without considering reputation. Table 4 shows the result: the rate of service utilization of negotiation mechanism without considering reputation is significantly higher.

Table 4 - Results of Statistical Test – One-way ANOVA (2)

Performance Measures	Mean rate of service utilization		Difference	*p-value*
	Without considering reputation	Considering reputation		
Rate of service utilization	66.59%	53.82%	12.8%	0.0000 ***

*** p<0.01

However, in this case, a higher rate of service utilization definitely does not necessarily indicate excellence in performance simply because users with a lower reputation are less likely to provide value to the service provider. Such a user may even abuse or improperly use the service provider's information acquired in course of service use. Hence, the service provider may want to selectively allow the use of its service and information by considering user reputation as a factor. A user with a higher reputation should be more than welcomed by the service provider and *vice versa*. To discern with both simple and reputation-aware negotiations, we conducted a regression analysis to examine to what extent the reputation level could be a determinant. Figure 5 shows that a negotiation mechanism that considers the user's reputation looks more proportional to the level of reputation than the other mechanisms. To identify statistically, the F-values of each case are compared as shown in Table 5. According to the results of Table 5, we can conclude that a negotiation that takes reputation into consideration is smarter than other mechanisms: the mechanism can selectively attract the users while avoiding the users who have a lower reputation to decrease the service risk. The service provider could increase the rate of service utilization up to 82% for the very reputation users, while it decrease the rate of service utilization down to less than 30% for the very low reputation users.

Table 5 - Results of Statistical Test – Regression Analysis

Mechanism	d.o.f	MSR	MSE	*F-value*
No negotiation	899	0.0598	0,1910	0.3093
Negotiation without reputation	899	0.2072	0.2227	0.9306
Negotiation with reputation	899	18.5643	0.2284	81.2667***

* p<0.1, ** p<0.05, *** p<0.01

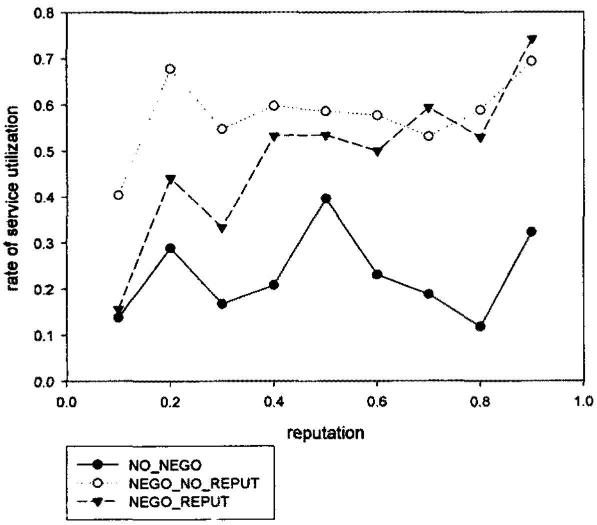


Figure 5 - Comparison of the Rate of Service Utilization with the Change of Reputation Level

Acknowledgements

This research is partially supported by u-market projects, Seoul and the ubiquitous Autonomic Computing and Network Project, the Ministry of Information and Communication (MIC) 21st Century Frontier R&D Program in Korea.

References

[1] Ackeman, M.S. (2004). "Privacy in pervasive environments: next generation labeling protocols," *Personal and Ubiquitous Computing*, Vol. 8 No. 6, pp. 430-439.

[2] Agrawal, R., Kiernan, J., Srikant, R., and Xu, Y.R. (2005). "XPref: A preference language for P3P," *Computer Networks*, Vol. 48, No. 5, pp.809-827.

[3] Berendt, B., Ganther, O., and Spiekermann, S. (2005). "Privacy in e-commerce: Stated preferences vs. actual behavior," *Communications of the ACM*, Vol.48, No.4, pp. 101-106.

[4] Chen, Y.R., and Petrie, C. (2003). "Ubiquitous mobile computing," *IEEE Internet Computing*, Vol. 7, No. 2 pp. 16-17.

[5] Cranor, L., Langheinrich, M., Marchiori, M., and Reagle, J. (2002). *The platform for privacy preferences 1.0 (P3P1.0) specification*, W3C Recommendation, HTML Version at www.w3.org/TR/P3P/.

[6] Hong, D., Yuan,, M., and Shen, V. Y. (2005). "Dynamic privacy management: a plug-in service for the middleware in pervasive computing," *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services*, Salzburg, Austria, pp. 1-8.

[7] Jutla, D.N., Bodorik, P., and Zhang, Y.J. (2006). "PeCAN: An architecture for users' privacy-aware electronic commerce contexts on the semantic web," *Information Systems*, Vol.31, No. 4-5, pp.295-320.

[8] Neustaedter, C., and Greenberg, S. (2003). "The design of a context-aware home media space for balancing privacy and awareness," *Lecture Notes in Computer Science*, Vol. 2864, pp.297-314.

[9] Redell, D. (1992). *Information Technology and the Privacy of the Individual*, Daft ACM Whitepaper on Computer and Privacy, September.

[10] W3C (2006)., http://www.w3.org/TR/P3P/.