IDEA GROUP PUBLISHING

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITP4987

Platform for Privacy Preferences: An Innovative Technology and Standard in E-Commerce

Noushin Ashrafi
University of Massachusetts, 100 Morrissey Blvd. MA 02125, noushin.ashrafi@umb.edu

Jean-Piere Kuilboer
University of Massachusetts, 100 Morrissey Blvd. MA 02125, jeanpierre.juilboer@umb.edu

ABSTRACT

Since the mid 1990's the advancement in enabling technology has made the use of web ubiquitous. Specifically, with the e-commerce on the rise the issues of privacy protection and system security have come into focus. E-business is permeating everyday life of more and more citizens around the world while privacy issues remain unresolved. This research is aimed at evaluating the state of a new technical approach to privacy protection. We are using a sample of 500 commercial enterprises to evaluate the adoption of the Platform for Privacy Preferences (P3P), one year after it was recommended by the World Wide Web Consortium (W3C).

INTRODUCTION

The computing field today is far from what the environment was twenty years ago. In the industrialized countries and increasingly around the world the Internet is used as a common tool for communication, information gathering, and trade medium. Along with this wide adoption some underlying weaknesses such as lack of security and disregard for personal privacy have emerged (Reagle & Cranor 1999)

For the end-user consumers and IT professionals, security issue is often regarded as a technical issue to be solved by off-the-shelf solutions such as virus checkers or firewalls. Privacy problem, however, has received little attention and is likely to get worse without proactive measures. According to a recent report by Zona Research, a low percentage (16%) of managers and IT staffers surveyed said that their company addressed privacy issues (Surmacz, 2001).

Why security and privacy remain an unresolved issue after so many years is not really a mystery. Both are more of a process than a static problem and the environment keeps changing while little resources have been applied to the matter. Budget and mentality have evolved slowly even if the new distributed architecture is far from the vault-setting environment characterizing computing systems of twenty years ago.

On the other hand the e-business economy is growing (Rohde, 2002) and there is a sense of emergency to some notion of trust and respect for privacy before another wave of issues brings instability to this emerging infrastructure. Research groups, both from academia and industry, have embarked in the development of Privacy Enhancement Technologies. Particularly active in this area are AT&T with demonstrated commitment to new initiatives such as P3P and IBM with its Privacy Research Institute.

P3P is an emerging standard developed by the World Wide Web Consortium. It should provide a simple, automated way for users to gain more control over the use of personal information on the web sites they visit, and for the web site developer a simple means to express their web site information collection/distribution/retention policies.

From a technical perspective, P3P reference files are written in the Extensible Markup Language (XML) and there is still a dearth of tools to help companies automate the implementation process. Among the proponents of P3P, IBM is distinguishing itself and experimenting with the development of a P3P policy editor (IBM, 2000). AT&T has a research project dedicated to developing a P3P add-on to Microsoft Internet Explorer (AT&T, 2003), which matches personal preferences of the visitors on privacy issues with the XML published privacy policies of the companies.

Despite these efforts, there is little published information on what actually P3P is, what its capabilities are, and to what extent it is being adopted by e-commerce organizations. This study aims at providing a brief explanation of P3P both as a new technology and as a standard. We use our empirical data to assess its adoption in the e-commerce environment

PLATFORM FOR PRIVACY PREFERENCES: THE EXPLANATION

Before entering into an argument about the respective advantages and disadvantages of P3P adoption the concept should be defined. P3P is a simple Internet standard aimed at facilitating the exchange of information about web site privacy policies. It includes a set of XML documents illustrating organizational priorities regarding customer information. It aims at keeping the information secure and restrains from sharing it with outside parties and provides a way to control the use of personal data collected through web site visit. In its most primitive form it consists of a set of rules published on the web server of the organization providing the user with a machine readable snapshot of the organization privacy position. On the client side, the latest browsers (e.g. Microsoft IE 6) contain new privacy features based on the P3P project. P3P was designed from the start for a machine to machine dialogue allowing negotiation and matching of stated policies to consumer preferences. An example of such negotiation would involve the visit of a transaction web site where the potential buyer will be asked to provide a number of personal information such as address, telephone, credit-card number, and through his/her purchase will disclose his/her various interests.

PLATFORM FOR PRIVACY PREFERENCES: THE TECHNOLOGY

P3P is currently implemented using the XML. While the bulk of the implementation reside on a visited web site and is managed by the web server, P3P will not be effective without a client counterpart. In the initial connection protocol the negotiation will involve exchange of TCP/IP address and will validate connection parameters. Once connected, P3P can enter into action by providing the P3P from the company's point of view (e.g. privacy policies elicited in XML addressing the questions of what information will be collected, what will be done with it, how long it will be kept, how it will be shared, the measures taken to protect it, etc..). The client through a web browser or a customized application will automatically match these policies with his/her own preferences and accept or deny the connection and ensuing transactions.

This will be possible only if the client is sophisticated enough to be able to express the preferences and configures adequately a P3P browser to match them. This obviously depends on a chain of events and resources on both sides:

1. The organization expresses its policies through valid P3P (i.e. XML) documents
2. The organization makes the policies available to web visitors in a standardized way (i.e. location, format, access)
3. The browser on the client side supports the P3P negotiation
4. The browser know where to find the policies
5. The client is able to express the privacy policies preferences
6. The negotiation of attributes is successful
7. The client is able to interpret the results of the interaction offered by the browser
8. The client has preferences applicable to generic web sites other wise the process will be a resource burden.

PLATFORM FOR PRIVACY PREFERENCES: THE STANDARD

P3P 1.0 is the result of a project started about five years ago. Since the beginning of Internet commercialization stakeholders have been aware that the rules of information transfer had changed from a consumption of information to a full exchange. Following pressures regarding a lack of privacy protection in Internet data exchange, organizations hosting web sites started to include privacy policies on the sites without standard context, location, or defined format. This made these human readable documents virtually unusable. Visitors had difficulty understanding, matching, and comparing differing policies. This lead to equally difficult avenues for data collection for purpose of customer relationship management or mass customization as there were increasing risks of litigation or negative publicity about abuse of personal data in an equivocal non regulated environment. This lead to the design of an agent-based system based on XML aimed at standardizing and automating the privacy preferences negotiation process.

The P3P standard is endorsed by the W3C and has a number of prominent stakeholders such as IBM, Microsoft, and AT&T. They work concurrently on next release of P3P and associated standard effort on a P3P preference exchange language (APPEL) aimed at providing the users with a mechanism to encode their preferences about privacy. The P3P standard is currently in its first inception. With version 1.0 in effect since April 2002, currently the W3C has an active group working on the next implementation including features left out of the first release. These new features include mechanisms to allow sites to offer a choice of P3P policies to visitors; visitors (through their user agents) to explicitly agree to a P3P policy; for non-repudiation of agreements between visitors and web sites; and user agents to transfer user data to services.

PLATFORM FOR PRIVACY PREFERENCES: THE ADOPTION

The Platform for Privacy Preferences (P3P) is emerging as a standard way for web sites to encode their privacy policies (Cranor et al. 2002). By mid March 2002, just before the publication of P3P first recommendation by the W3C, Jupiter Media Metrix reported that from the top ten web sites with the highest traffic, six web sites had adopted P3P. This led to the erroneous assumption that widespread adoption would follow after the passage of the recommendation by W3C. A study published by Ernst and Young in September 2002, six months after the recommendation by W3C, found only 24% of the top 100 sites posting a P3P policy. A follow up study of 500 sites by Ernst and Young in January 2003 revealed that 18% had adopted P3P.

Our study conducted in June of 2003 involves 500 interactive companies and analyses the rate of the diffusion of P3P practices one year after the initial standard was published. We checked the web sites of 500 companies that were listed as top 500 interactive companies by Interactive Week Magazine published in June 30, 2000. The listing included public and private companies and provided us with the background information about these 500 companies including their total revenues, online revenues, type of business, and most importantly their web site URL.

Table 1: comparison of P3P adoption

Study by	Date	Sample Size	% that adopted P3P
Jupiter Media Metrix	March 2002	10	60
Ernst & Young	September 2002	100	24
Ernst &Young	January 2003	500	18
Our Study	June 2003	500	13

We visited these 500 web sites and conducted our research to find out what percentages of the top 500 companies have implemented P3P. We were also interested to find out if company size (based on total revenue) was a factor influencing the implementation of P3P. We found that only 13% of top 500 interactive companies had adopted P3P as of June 2003. Table 1 compares the results of our study with previous ones:

These numbers show a decline in the rate of adoption by time and as the sample size increases. A cursory analysis of study conducted by Jupiter Media Metrix points out that one possible reason for the high rate of P3P adoptions could have been be due to the fact that their small sample included large companies such as Yahoo, MSN, Lycos and some industry P3P sponsors such as IBM, ATT, and Microsoft. These may have biased the results to a high adoption rate of 60 %.

Given the tendencies for these large conglomerates to adopt P3P, we decided to stratify our results along company size (overall revenues) and see if the results would be different. Our results suggested that the rate of adoption for the 100 largest companies was the highest (26%) and the rate declined as the company size decreased. For companies in 101-200 cluster adoption rate was 16% for those in 201-300, adoption rate was 13%, and for companies in 301-400 cluster, adoption rate was 9% and only 1% of the final 100 adopted P3P. This decline can be contributed to the organizational structure and resource levels that allow the larger organizations to experiment with new technology and be more responsive to consumer concerns where as smaller companies lag behind in adopting P3P and perhaps have privacy as a low item on their priority list.

PLATFORM FOR PRIVACY PREFERENCES: COMPETITIVE ADVANTAGE

Organizations are generally adopting technology to gain a competitive advantage. As Narayanan (2001) points out, "technology can often be a critically important element in the competitive battles between firms." Among the firms adopting P3P, a perceived advantage would be a factor fostering the frequent return of the users to the web site, hence furthering the business. As Culnan et al. (1999) suggest, "one goal of offering high quality service is to keep customer coming back and to attract new one through positive word of mouth." Privacy issues are becoming a major concern of the customers (Reagle and Cranor 1999) and an important aspect of quality service. Hence, an easy to use device intended to ensure the customer that the organization has a system to protect their privacy may characterize the quality service to which Culnan alludes.

AT&T Privacy Bird, a utility tool that evaluates the degree of equivalence between consumers' concerns and company's privacy policies, seems to be a move in the right direction. AT& T privacy bird provides the end-user with a tool warning of discrepancies between privacy issues important to them versus those addressed by the company. The biggest advantage of this tool is its ability to accommodate consumers' level of desire to protect their privacy.

Privacy preferences are divided into four categories: health or medical, financial or purchase, personally identified information, and non-personally identified information, each of which has a set of warning rules. A customer, using the privacy bird, has the option to choose low, medium, high, or customized levels of privacy protection. Table 2 (AT&T 2003) depicts the privacy preferences categories, rules within each category, and the four levels of protection. The table

Table 2. AT&T privacy categories and rules.

Privacy preferences categories and rules	low	medium	high	custom
Health or medical information				
Warn me at web sites that use my health or medical information for analysis, marketing, or to make decisions that may affect what content or ads I see, etc.	X	X	X	
Warn me at web sites that use my health or medical information to share with other companies (other than those helping the web site provide services to me)	X	X	X	
Financial or purchase information				
Warn me at web sites that use my financial information or information about my purchases for analysis, marketing, or to make decisions that may affect what content or ads I see, etc.			X	
Warn me at web sites that use my financial information or information about my purchases to share with other companies (other than those helping the web site provide services to me)		X	X	
Personally identified information				
Warn me at web sites that may contact me to interest me in other services or products via telephone			X	
Warn me at web sites that may contact me to interest me in other services or products via other means (email, postal mail, etc.)			X	
Warn me at web sites that may contact me to interest me in other services or products and do not allow me to remove myself from marketing/mailing lists	X	X	X	
Warn me at web sites that use information that personally identifies me to determine my habits, interests, or other characteristics			X	
Warn me at web sites that use information that personally identifies me to share with other companies (other than those helping the web site provide services to me)		X	X	
Warn me at web sites that do not allow me to find out what data they have about me		X	X	
Non-personally identified information				
Warn me at web sites that use my non-personally identifiable information to determine my habits, interests, or other characteristics			X	
Warn me at web sites that use my non-personally identifiable information to share with other companies (other than those helping the web site provide services to me)			X	

demonstrates that, for example, a consumer choosing the low level will be protected against three rules as marked under the column "low". If the web site does not accommodate three privacy issues check-marked in the table under column heading "low", a warning signal will be given. It is then up to the customer to decide whether to continue with his/her transaction on this web site of simply leave and navigate to another web site that will provide the same service or product with more a higher of privacy protection. A consumer with medium level desire to protect her/ his privacy would select medium level protection where six of privacy issues are check marked in table under column heading "medium". Consumers with high level desire for protection could choose high level protection, where twelve privacy issues are marked under 'high" column. The tool also offers the choice of customizing which issues are important to consumer and check marks them under "custom" column.

There is a trade off between the availability of the services or products that provide the convenience of on-line transactions and protecting consumer information. Consumers, who want to protect all health, financial, personal, and non-personal information, may not find many companies with a web site that promises such protection. Other consumers may ask for some protection against sensitive information such as those check marked under "medium" column. We chose the average consumer as the middle-of-the-road person that wants some privacy protection, but is not overly sensitive

To test what percentage of companies will accommodate such users, we visited those companies from the top 500 that have P3P policies. Four results were observed: sites with "positive image results" are those that accommodated six or more criteria for medium privacy protection; sites with "warning results" are those that accommodated some, but not all of six criteria; sites with negative P3P policies were those that had P3P policies but issue a negative report citing that their policies did not match any of the user's concern. Finally sites with "non-functioning" policies were those that have P3P policies, but did not respond to the request of matching privacy bird most likely due to syntax error or non-standard location. The results of the visits are outline in table 3 stratified by size.

Table 3. Privacy Preference match to P3P adopting sites (using ATT privacy Bird)

	First 100	101-200	201-300	301-400	401-500
Sites with positive image results	11%	8%	5%	3%	0%
Sites with warning results	11%	4%	2%	1%	0%
Sites with negative P3P policies	4%	4%	6%	5%	1%
Sites with non functioning policies	15.4%	15.0%	46.15%	55.56%	100%

Table 3 indicates that very small percentages of companies have P3P accommodating the privacy need of a pragmatic user and the proportion gets even less significant for smaller organizations. The companies in the lowest tier (401-500) are absolutely non compliant. Generally the organization could reap the benefits from positive positioning but surprisingly implementations are less than optimal.

CONCLUSION

Like many other innovations in E-commerce, P3P is in its infancy. Successful adoption of P3P and its likely descendants is highly dependent on the degree, stability and wisdom of administrative sponsorship and on the integration of technical, legal and market driven strategies. In the case of P3P the usefulness of the technology has thus far eluded both technical and top management. Early adopters and industry proponents have jumped onboard the P3P bandwagon, but as our study reveals, the majority and the mainstream of online stakeholders have yet to endorse the W3C recommendation either waiting for the next iteration, distrusting the proponents of the technology, or merely doubting the business value of the solution. To the well intentioned organization that embarked in the adoption of P3P without full understanding of the business and legal implications, a word of advice would be to get a book on the subject and experiment on a non production server. Validating the privacy policies through the W3C valuator and ATT Privacy Bird make a smooth transition to the technology.

REFERENCES

AT&T (2003). AT&T privacy bird. Retrieved October, 2003, from AT&T web site: http://privacybird.com

Culnan, M.J. and Amstrong, P.K. (1999) Information privacy concerns, procedural fairness, and impersonal trust. An empirical investigation. *Organization Science*, 10(1), 104-114.

Cranor L.F. and Wenning R. (2002). Why P3P is a good privacy tool for consumers and companies. Retrieved October 2003, from GigaLaw web site:http://www.gigalaw.com/articles/2002-all/cranor-2002-04-all.html

Cranor L. (2002). Web privacy with P3P. O'Reilly & Associates.

Ernst & Young P3P dashboard project. Retrieved October, 2003, from Ernst & Young web site: http://www.ey.com/GLOBAL/content.nsf/US/AABS_-_TSRS_-_Library

IBM (2000) P3P policy editor. Retrieved October, 2003, from IBM web site:http://www.alphaworks.ibm.com/tech/p3peditor

Narayanan, V. K. (2001) Managing Technology and Innovation for Competitive Advantage, Prentice–Hall.

Reagle J. and Cranor L.F. (1999) The Platform for Privacy Preferences. Communications of the ACM, 42(2), (Feb.), 48-55

Rohde, L. (2002) UN: Worldwide Use of Internet, E-commerce Still Growing. IDG News Service, Nov. 21, 2002.

Surmacz, J. (2001). Privacy a low priority. *CIO Magazine*, June 20, 2001.

W3C (2003). Retrieved June, 2003, from World Wide Web Consortium web site: http://www.w3.org/P3P/