

MozPETs - a Privacy enhanced Web Browser

Lars Brückner and Marco Voss
Darmstadt University of Technology
{brueckner,voss}@ito.tu-darmstadt.de

Abstract

Usability problems are a major obstacle for the further deployment of privacy enhancement technologies. Most tools are highly specialized and require profound knowledge to be used properly. We propose a set of privacy protection tools for technically unskilled users. The tools were integrated within the open-source browser Mozilla.

Keywords: privacy enhancement, identity management, web applications, end users, open source

1. Introduction

Privacy and security problems and the resultant lack of trust in the internet concern users and service providers. Despite numerous results from the research community, the actual deployment of security and privacy enhancement technologies has not changed significantly in the recent years. Although many tools are available free of charge, only few users have the necessary knowledge and are willing to spend the effort to use these tools. Our work is an attempt to improve the usability of privacy enhancement technologies. We focus on the typical home user, who accesses the internet to shop, read news and e-mail. The main threats to privacy are disclosure of personally identifiable information to unknown parties, spam, phishing, user tracking through clickstream analysis, and the sharing of data between multiple sites and third parties.

2. Privacy Enhancing Technologies

Privacy research has been done on a lot of different technologies [16, 15, 21]. Most research focused on anonymity for network access, publishing, authorization, and payment. E-mail fraud and phishing has lead to increased research on technical countermeasures [11]. Data licenses were proposed to guarantee access rights to personal information [10]. An identity management system combines privacy enhancement and security technologies to minimize

the disclosure of personal information, and provides the user with means to make informed decisions whether certain data is disclosed or not [17]. The current identity management tools differ a lot in features and scope [18].

However, the web looks different. Privacy invasive technologies, such as tracking users with web bugs and third party cookies across multiple sessions and different sites, are common practice among site operators. Data mining is a key element of many commercial sites' business plans. In general, the privacy and security features of today's web browsers are the same as of the Netscape Navigator 6 released in late 2000. The user can modify the cookie policy, wallet components store logins, passwords and other personal information. Many security tutorials advice users to disable cookies and active content, including JavaScript. Applying these settings makes the web nearly unusable, as most sites will not work correctly. After this experience the disappointed user will restore the old settings.

Anonymous web access or effective anti-profiling measures, such as blocking web bugs or suppressing of referer headers is not supported by the current browsers. The user has to install and maintain additional tools, such as JAP [6], CookieCooker [5], Junkbuster [19], and Bugnosis [3] to achieve this. The process to find the right tools, configure them correctly, and deal with possible errors is too complex for most users.

Also, self-regulatory systems like P3P [22] and privacy seals were build on the false assumptions that companies would try to market fair information practices as an advantage and that users would actually spend the time to read the privacy policies and change their actions [4, 2]. P3P was deployed widely when the Microsoft Internet Explorer started to check for P3P policies as a prerequisite to accept cookies. Unfortunately, this reduced P3P to an option of the cookie menu. Only few P3P agents like Privacy Bird [12] actually used the P3P policy to provide better information for users. Privacy Seals such as TRUSTe have been introduced to convince consumers that a web site follows appropriate privacy practices. However, seal issuers have little powers to influence their clients [13, 7], and [20] shows that most consumers neither understand nor care about privacy seals.

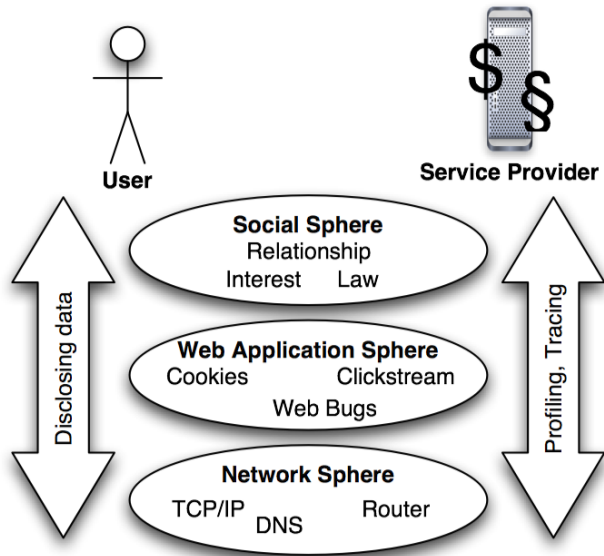


Figure 1. Spheres Overview

3. Web application spheres

To find the right tools for our scenario, we have to consider the achieved gain of privacy, possible compatibility problems, and the additional effort for the user. Another point are possible countermeasures by the service providers. If a certain privacy enhancing technology gains wide acceptance, providers might look for even more invasive alternatives or ban users from the service.

Looking at the components and parties involved with web access, we identified three different spheres: the network, the application, and the social sphere. The spheres have different privacy properties; more important, the user and the provider have very different influences on the processes that occur within the spheres.

The network sphere contains the TCP/IP stack, all intermediate routers, required third-party servers (e.g. DNS) and their operators. Within an e-commerce scenario, neither the user nor a service provider have any real influence on the network. The user and the provider have the common interest that the contents of their communication remains private and is not tampered with. Users have the interest to hide their real IP address to avoid being tracked.

The application sphere contains the user's web browser, the service's web server and all application logic (HTML, JavaScript, CGI-scripts) that is stored on the web server and its backends. The application sphere is dominated by the provider who has the full control on the application logic. The logic defines the transactions and what data is collected during the application. The provider can implement cookies

or other technologies to analyze the user's clickstream. The user can influence the application only by customizing his browser and installing other privacy-enhancement tools. If an application does not work because of such customizing, the user has often no choice but to deactivate the protection.

The social sphere contains the real world: the user, the business which owns the web server, and their real-life interests and social contexts. In the end, every application is build to fulfill a human demand that originates in the social sphere. A big difference to the network sphere is that users are aware that they disclose social information (e.g. their name) when they enter it with the keyboard. People are aware that this information is often shared with other parties. However, most users are not aware that cooperating services can correlate disjunct sets of social data with tracking information gained in the network and application spheres.

4. The MozPETs prototype

Our analysis of the problems, their relation to the spheres and existing tools brought us to the conclusion that a set different protection technologies is required. We also concluded that the protections should be integrated within the browser to reach our usability goals.

Working with an open source browser gave us the flexibility to change any aspect of the browser. While Mozilla's core libraries for network and rendering are implemented in C++, the GUI and most of the application logic is implemented in XUL, an XML based interface description language, and JavaScript. A powerful extension mechanism allows the user to add modifications written in XUL and JavaScript to the browser without recompiling. We used the following methods to implement our prototype: changing the browser's defaults, integrating of existing tools and extensions, and development of new extensions for areas we could not find a suitable tool for. Our prototype has the following features:

To protect communication on the network sphere, we implemented a wrapper that integrates anonymizer JAP.

To improve privacy within the web application sphere, our prototype aims to protect the user against embedded tracking methods without losing compatibility. A first step was to force all cookies to be treated as session cookies, and deny all third-party cookies to hamper profiling across browser sessions. Then we modified the extension Ad-Block [1] to block the loading of arbitrary third party content. We found that this setting would break many sites and would make our prototype unusable. Therefore we have chosen to block only a few types of third-party content by default, and to provide the user with a GUI to add more new rules on his own.

We developed two new components, MozPAw and

Clickstream Analyzer, to educate the user about the remaining third party content and other potential harmful technologies. MozPAW inspects the current web page to identify potential threats to the user's privacy. Each web site is tested against a set of filters that check for certain properties of the page, e.g. number and desired validity of cookies, or the amount and type of third-party content. The result of all tests is aggregated and displayed as an emoticon next to the URL address bar (see figure 2). The user can open a special sidebar to examine the detailed results of the tests. The



Figure 2. The MozPAW privacy icon rates the current page

Clickstream Analyzer records the domains of all web sites a user visits, and all domains that have provided third-party content (e.g. banners or web-bugs). The result is presented as graph (see figure 3) that shows all domains and how they are connected with each other. The user can now identify clusters of co-operating service providers who could combine their access logs to profile him. He can try to break the tracking by adding new rules to the blocking component described above.

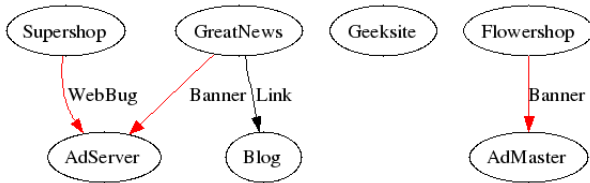


Figure 3. A Clickstream Graph reveals cross-site user tracking

The current web infrastructure ignores the social sphere. While service providers have CRM systems to manage their customer data, no equivalent tool exists on the user side. We developed the iJournal to support the user to keep track of personal information like real name and address after it was disclosed to websites. To use the iJournal, the user has to define once which information he wants to keep track of by entering the data in the browser's wallet component. Whenever the iJournal detects that personal information is about to be submitted to a website, the submission of the data is paused and the iJournal tries to profile the operator

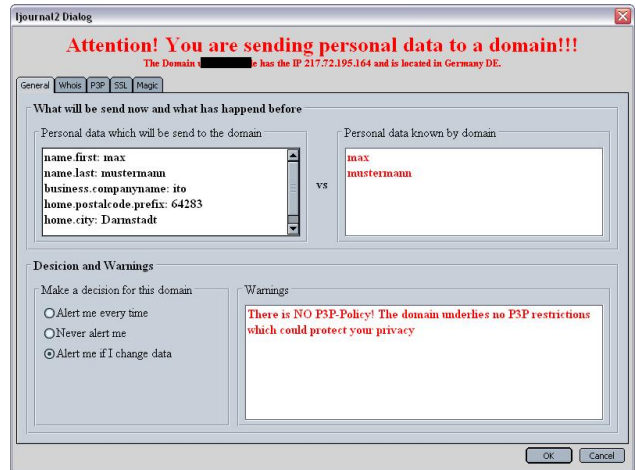


Figure 4. If the iJournal detects personal information in the user's input, it collects data about the provider and asks for confirmation

of the website. Therefore it analyzes the site's P3P policy, SSL certificate and Whois-record. The result is presented to the user (see figure 4). The user can choose to abort the transaction if he finds that the information about the service does not meet his current requirements. For example, the user might decide that he would rather not disclose certain data to a service as he finds out that it might be shared with others. In contrast to other P3P agents, the iJournal does not compare the service's policy to a user's preferences policy; it tries to provide valuable information about the current context, so that the user can make an informed decision. If the user authorizes the submission, the gathered data about the site and the type of the submitted data is stored in a journal for later use. The iJournal is a re-implementation of an earlier prototype [9].

Our prototype is a full browser which is compatible with existing web sites and provides a high level of privacy protection on the network and application sphere. Information of possible profiling on the network level is visualized graphically. The user can identify clusters of co-operating services providers. In conjunction with the information from the iJournal, the user himself can construct a worst-case profile.

5. Future Work

Mozilla provides a valuable platform to develop and test new privacy enhancement technologies. We plan to extend the work described here. We see the need to improve the interaction between the awareness components and the iJournal. Also, the browser and mail client should be enhanced to

fully support remailers and nym servers (like [14]). Another interesting area is the expansion to collaborative protection technologies by integrating a peer-to-peer network. Possible applications include the exchange of reputation values, a server-independent chat system, or spam detection. In the long term, even anonymous web access or remailer functionality could be implemented on top of a p2p network. We plan to integrate extensions developed by others too.

6. Conclusion

The majority of users has no knowledge how the internal mechanisms of the web work, and is thus not able to choose adequate privacy protection measures themselves. To empower these users to better protect themselves, software that covers the technical details and protects privacy by default is required. If this is not feasible, the user should be provided with related data that supports him to make an informed decision. We have presented a concept for extending a web browser with identity management capabilities that focuses on the technically unskilled user and his view on privacy and security problems. We have implemented our concept within the Mozilla Browser; the extension mechanism allowed us to create privacy protection components that integrate seamlessly with the user's browsing experience. The recent success of Mozilla's products gives the research community the opportunity to support the further adoption of privacy enhancement technologies by actively contributing their results to the open-source community. Our code is available as open source [8]. We encourage other researchers to base their future prototypes on Mozilla too.

Finally, we hope to motivate browser vendors to include more privacy protection features within their products.

Acknowledgments We thank our students Stéphane Alonias, Adalbert Biadatz, Robert da Campo, Sebastian Funk, Jens Hatlak, Torsten Hofacker, Pieter Hollants, Lars Kirsten, Alexander Kraus, Andreas Piesche, Christian Predikant, Björn Schneider, Benjamin Stritter, and Stephan Zimmer for their hard work on hacking Mozilla.

References

[1] H. Aasted, W. Palant, Rue, S. Kinitz, and A. Spuler. Adblock project, 2005. <http://adblock.mozdev.org>.

[2] A. Acquisti and J. Grossklags. Privacy and Rationality in Individual Decision Making. *IEEE Security and Privacy*, 3(1):26–33, Jan 2005.

[3] A. Alsaïd and D. Martin. Detecting Web Bugs with Bugnosis: Privacy Advocacy through Education. In *Proceedings of the Second International Workshop on Privacy Enhancing Technologies (PET2002)*, April 2002.

[4] B. Berendt, O. Günther, and S. Spiekermann. Privacy in E-Commerce: stated Preferences vs. actual Behaviour. *Communications of the ACM*, 48(4):101–106, Apr 2005.

[5] O. Berthold. Cookiecooker home page, 2005. <http://www.cookiecooker.de>.

[6] O. Berthold, H. Federrath, and S. Köpsell. Web MIXes: A System for Anonymous and Unobservable Internet Access. In *Proceedings of the International Workshop on Design Issues in Anonymity and Unobservability*, Jul 2000.

[7] P. Boutin. Just How Trusty Is Truste?, April 2002. Wired News, <http://www.wired.com>.

[8] L. Brückner. Mozpets home page, 2005. <http://mozpets.sf.net>.

[9] L. Brückner, J. Steffan, W. Terpstra, and U. G. Wilhelm. Active Data Protection with Data Journals. In *Proceedings of the GI Jahrestagung (Schwerpunkt "Sicherheit - Schutz und Zuverlässigkeit")*, pages 269–280, 2003.

[10] S.-C. Cha and Y.-J. Joung. From P3P to Data Licenses. In *Proceedings of Third International Workshop on Privacy Enhancing Technologies (PET2003)*, 2003.

[11] N. Chou, R. Ledesma, Y. Teraguchi, D. Boneh, and J. C. Mitchell. Client-side Defense against web-based Identity Theft. In *Proceedings of the 11th Annual Network and Distributed Systems Symposium (NDSS'04)*, Feb 2004.

[12] L. F. Cranor, M. Arjula, and P. Guduru. Use of a P3P User Agent by early Adopters. In *WPES '02: Proceedings of the 2002 ACM workshop on Privacy in the Electronic Society*, pages 1–10, New York, NY, USA, 2002. ACM Press.

[13] M. Delio. Yahoo's 'Opt-Out' Angers Users, April 2002. Wired News, <http://www.wired.com>.

[14] E. Gabber, P. B. Gibbons, Y. Matias, and A. Mayer. How to Make Personalized Web Browsing Simple, Secure, and Anonymous. In *Proceedings of Financial Cryptography 97, LNCS 1318*. Springer-Verlag, 1997.

[15] I. Goldberg. Privacy-enhancing Technologies for the Internet, II: Five years later. In *Proceedings of the Second International Workshop on Privacy Enhancing Technologies (PET 2002)*, April 2002.

[16] I. Goldberg, D. Wagner, and E. Brewer. Privacy-enhancing Technologies for the Internet. In *Proceedings of the 42nd IEEE Spring COMPCON*. IEEE Computer Society Press, February 1997.

[17] M. Hansen and A. Pfitzmann. Anonymity, Unobservability, Pseudonymity, and Identity Management - A Proposal for Terminology, September 2004. Draft v0.21, available at http://dud.inf.tu-dresden.de/Literatur_V1.shtml.

[18] Independent Centre for Privacy Protection and Studio Notarile Genghini. Identity Management Systems (IMS); Identification and Comparison study, September 2003. <http://www.datenschutzzentrum.de/projekte/idmanage>.

[19] Junkbusters. The Internet Junkbuster, 2005. <http://www.junkbusters.com/ijb.html>.

[20] T. Moores. Do Consumers understand the Role of Privacy Seals in E-Commerce? *Communications of the ACM*, 48(3):86–91, Mar 2005.

[21] A. Rezugui, A. Bouguettaya, and M. Y. Eltoweissy. Privacy on the Web: Facts, Challenges, and Solutions. *IEEE Security and Privacy*, 1(6):40–49, Nov 2003.

[22] World Wide Web Consortium. P3P public overview, 2005. <http://www.w3.org/P3P>.