AN ANALYSIS OF ANONYMITY ON THE INTERNET

FRANCOIS SEARLE AND MAREE PATHER
Port Elizabeth Technikon, South Africa, francoissearle@balltron.co.za maree@petech.ac.za

Key words: Anonymity, Privacy, Internet

Abstract: Internet technologies bring with them a host of ramifications for the personal user, including a threat to personal privacy and anonymity. Anonymity and privacy are very closely integrated. Anonymity can be regarded as a way to protect one's privacy. Third parties, providing Internet services, can record and collect user data in order to create user profiles. This is done in numerous ways, such as using cookies. The available anonymizing services have various deficiencies. The objective of this paper is to analyze current threats to anonymity, as well as popular countermeasures being used, with a view to providing an improved, systematic model integrated with operational controls to improve anonymity.

1. INTRODUCTION

Companies are becoming increasingly interested in knowing their client market to more effectively direct their marketing efforts. Health and insurance institutes want client information to improve their decision-making operations. This encourages the data "mining" companies to come up with mechanisms for gathering more detailed information about clients and prospective clients. In the heat of terrorist attacks, federal investigators are also using more invasive "surveillance" methods. Internet technologies thus bring with them a range of implications for the individual user, including a threat to personal privacy and anonymity. (Junkbusters, 2002)

The objective of this paper is to analyze current threats to anonymity on the Internet, as well as popular countermeasures being used. Deriving from these, possible improvements - and the integration of different features from existing models to enhance their efficacy – will be suggested.

2. BACKGROUND

Anonymity can be defined as the ability to use the Internet in such a manner that one's identity is hidden. Privacy refers to the ability to store information or to transmit information, in such a manner that only the intended party is able to use the information (Internet Security and Privacy for Activists and Citizens, 2002). Anonymity and privacy are very closely integrated. Anonymity can be regarded as a means for protecting one's identity, thereby allowing one a degree of privacy. The difference between individually identifiable information and personally identifiable information should first be clarified. Individually identifiable information contains "demographic" information e.g. interests, hobbies, etc. unique to a user. This is regarded as allowing anonymity, since the individual's identity is still concealed. Personally identifiable information contains personal information, such as one's name, address, etc., which uniquely identifies an individual.

According to the Federal Trade Commission (FTC) that surveyed a comprehensive list of websites, 92% of the websites collected personal information. 30% of users feel comfortable giving personal information over the Internet and 42% of Internet users are concerned about their personal information being sold. (Federal Trade Commission, 1998)

Legislation regarding privacy concerning the Internet is very slow moving, with a large amount of legislation pending. The FTC together with the Network Advertising Initiative (NAI) came up with a set of principles, the NAI Principles, to be enforced by self-regulatory committee. The NAI principles, as well as the South African Electronic Communications and Transaction Bill (SA Electronic Communications and Transactions Bill, 2002), are aligned with the FTC's framework for fair information practice principles. They recommend that websites adhere to the following:

1. Provide notice and raise awareness of their site's information practices

2. Provide the user the ability to choose and give consent for which data is collected.
3. Provide the user access to the data about him, thereby being able to ensure that the data collected is accurate and complete.
4. Ensure the integrity of data collected; cross-referencing it against various sources. (Federal Trade Commission, 1998)

Unfortunately, since websites are not abiding by the NAI principles and the self-regulatory committee has failed to ensure compliance, these principles seem relatively ineffective.

Online Profiling

Advertising companies place advertisements, called "banner ads" on web pages. In return, the website may receive benefits such as increased traffic due to the advertising. A company's advertisements will almost certainly be seen more than once during a session, thereby enabling companies to compile a "demographic" user profile. Although this information is unique, it is not personally identifiable. Unfortunately, this changes when the user fills in a form on a web page such as a survey, or registration forms. Personally identifiable information can then be captured. All of the above is done without any consensual participation of the user. By matching different sets of information gathered, the advertising companies are able to create fairly accurate personal profiles. (Roha, 2000)

The biggest cause for concern is that highly confidential information, such as financial and medical histories, could be derived from a user's browsing habits. A survey by the California HealthCare Foundation confirmed that some health care sites are allowing advertisers to collect personally identifiable information. This appears to violate even their own privacy policies. Some privacy-advocates are concerned that online profiles, especially information about one's financial or medical status, could be influence decision-making, such as denying health insurance. Another cause for concern regarding online profiling is the inaccuracy of the information. When a single user's login credentials are used by multiple users, it is difficult for the advertising companies to distinguish between the different users. (Roha, 2000)

3. CURRENT THREATS AND SOLUTIONS

The need for accumulating information fuels numerous intrusive technologies that threaten one's anonymity. They in turn encourage a range of countermeasures.

3.1 Threats

To fully understand the threat to one's anonymity, one needs to first understand the different ways that privacy and anonymity are threatened on the Internet. The following sections discuss typical examples of such threats.

3.1.1 Cookies

Cookies are small text files stored on the user's computer. There are a few different types of cookies, of which "Advertising" cookies are the most threatening. These cookies are placed on a user's computer by banner advertising companies. They are therefore not limited to one website. Cookies are generally only read by the company that places them, but could also be read by other companies that have an agreement to share user information stored in cookies. Synchronizing or exchanging their acquired user-profiles enables advertising companies to have a profile on a user before the user has visited their sites. (Wilsker, 2001)

3.1.2 Web bugs

A web bug is a 1x1 pixel graphic mostly served by advertising companies as an alternative to banner advertisements. Because the graphic is so small, and therefore loads fast, the user is unaware of it. This is a fairly new way used to collect user information. Web bugs have increased by 488% over the past few years (Nua Internet Surveys, 2002). A web bug is not much more than a picture and is therefore not detected by anti-virus software. Web bugs can place a cookie on the user's computer although disabling cookies does not disable all web bugs. The Privacy Council have demonstrated that a web bug can be used to steal a computer user's entire e-mail address book and monitor emails merely by clicking on a bugged page. Even more devious is the web bugs' ability to place tiny applications on the user's hard drive, which secretly collects user information by monitoring documents for specific words such as "financial". (Wilsker, 2001)

3.1.3 Exploit Software

So-called "exploit" software is typically software that comes bundled or even part of software downloaded from the Internet. Spyware is software that tracks almost all surfing activities, and secretly sends that information to a third party. Scumware is software that takes control of the user's machine for other malicious purposes. There were an estimated 48000 different Malware programs, also known as pests, purposefully damage the user's computer and captures information. (Wilsker, 24 May 2002, email)

3.1.4 Web Enabled Software

Microsoft and Netscape do not ensure secure browsing by default settings. Many users are ignorant of these settings and as a result are exposed to the threats on the Internet. Furthermore, Microsoft and Netscape still permit "third-party" cookies.

Microsoft's new XP range makes use of a ".NET Password" online identification system. Microsoft has admitted that this new system could allow hackers the capability to steal credit card numbers and personal information. (Junkbusters, 2002)

3.1.5 Exploits in Hypertext Transfer Protocol (HTTP)

HTTP is the predominant protocol on the Internet. Every HTTP message has a header that provides the server with information that may threaten the user's anonymity. The Hypertext Transfer Protocol (HTTP/1.1) RFC 2616 refers to a few precautions regarding the protocol, some of which, regarding anonymity, are listed below:

- Information, such as user's name, location, email address, passwords, encryption keys, etc., can easily be leaked.
- The website server is able to save information that is confidential in nature.
- Web servers and proxy servers should take particular care in securing the information gathered in their log files as it can be used maliciously. (HTTP RFC 2612, 1999)

It seems obvious that HTTP is not an entirely anonymous protocol and therefore one must look to other methods to ensure anonymity. (Eckert & Pircher, 2001)

3.1.6 Usenet (Newsgroups) Threats

User profiles, with real names, email addresses and demographic information can be complied by scanning through Usenet newsgroup postings (Goncalves, Donkers, Harkin, Hart, Niles, Toyer & Willis, 1997).

3.1.7 Internet Service Providers Threats

Internet Service Providers (ISPs) can monitor everything a user does online, such as: what one sends over the Internet, as well as the files one downloads. Most anonymity services address anonymity at third party level. (Privacy Rights Clearinghouse, 2000)

3.1.8 Email Threats

Email headers contain user and routing information that are transmitted in clear text. Most businesses provide their employees with email facilities and some randomly monitor their employees' email, which introduces a new threat to personal anonymity. (Privacy Rights Clearinghouse, 2000)

There are other additional indirect threats to anonymity, such as authentication, by which the user clearly identifies himself and Web spiders that browses through the Internet and gathers information from websites. Being aware of the threats to anonymity creates a positive cautiousness when using the Internet.

3.2 Current Popular Anonymizing Models

There are currently several ways available to protect one's anonymity. One of the most well known ones is the *Anonymizer* website that allows the user to browse the Internet anonymously. *Anonymizer* removes privacy and security threats by, for example, rewriting the web pages and hiding the user's IP address. Unfortunately, the user has to pay for the service. Such services usually result in a reduction of available bandwidth. *Anonymizer's* president warned that today's complex software makes it extremely difficult to find every possible vulnerability. (Anonymizer, 2002)

Crowds, a proxy server protocol written in perl, is based on the concept of a single person "blending into a crowd". A request is sent from the user's browser to an affiliated member of a "crowd". That member then either sends the request to the destination or passes it to another member, until it is finally submitted. A user's computer might be used to submit other

member's requests. The origin of the message is obfuscated. A disturbing weakness of *Crowds* is that the user's request is routed through other machines where sensitive data can be captured. (Reiter & Rubin, 2002)

Another way to provide anonymity is through *onion-routing*. It uses a layered approach to hiding the user's identity until it reaches the destination. An anonymous route is determined through numerous "onion routers". Each request is layered with the next router's address (Figure1), which could optionally be encrypted. When the "onion" arrives at a router it removes its layer revealing the next address and passes it on. The connection is then destroyed after the data is sent. (Goldschlag, Reed & Syverson, 2002)

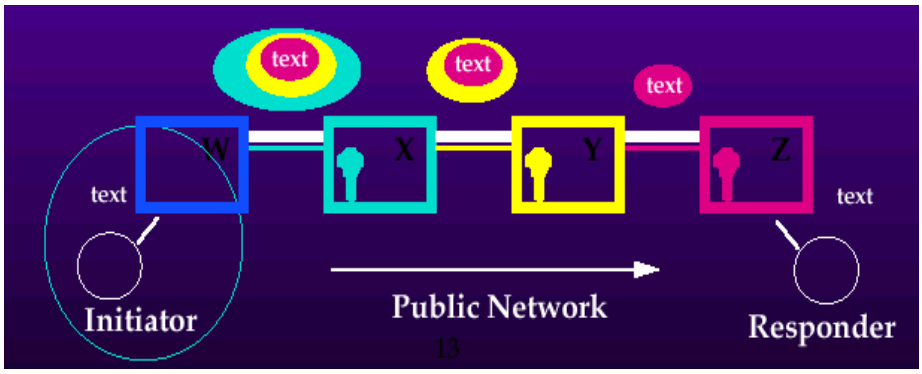

Figure 1. A message passing through the onion-routers. (Goldschlag, et al, 2002)

Infomediaries are a relatively new technology that allows the user to choose what personal information is shared with each third party. The third party is then provided with the chosen level of information (Goncalves, et al, 1997).

There are several other anonymizing services available. Most are based on the three models described above. All have unique advantages and disadvantages. A better model could possibly derived by integrating different aspects from different models.

4. A SYNTHESIS OF ANONYMITY SOLUTIONS

The objective of this section is not necessarily to create untraceable connections but to keep the user anonymous; revealing as little as possible about the user while browsing the Internet. It seems that an anonymity solution should, at a minimum, accomplish the following:

- There should be reasonable resistance against attacks.
- The system must be trusted and easy to use by its users.
- The system should be as local as possible.
- There should minimal reduction in speed and efficiency during browsing.
- The system should be modular; this will allow uncomplicated replacements, and the failure in a component will not affect the entire system.

Most anonymizing services seem to lack a systematic approach. The proposed approach will try systematically to prevent threats in the connection from the user's computer to the web server(s). (Rennhard, Rafaeli, Mathy, Plattner, & Hutchison, 2002)

4.1 Operational Controls

To ensure the success of the technical controls, operational controls are essential. These controls encourage and guide the users to use the Internet anonymously. The subsequent subsections will discuss some suggested operational control guidelines.

4.1.1 Create a "side" profile

Create a "side pseudonymous" profile that does not contain any personal information about the user. Also, create a side email account or alias, for the side profile. When the user needs to provide personal details on the Internet this profile will allow the user to stay anonymous. Write the side profile details down to maintain consistency and prevent uncertainty. (McCandlish, 2001)

4.1.2 Do not reveal unnecessary personal information.

The OECD Guidelines on Protecting of Privacy and Transborder Flows of Personal Data specifies that collection of data should be limited and should be relevant to the purpose for which it is gathered. Therefore, the user should not feel obligated to provide any information that is not needed. (Organisation for Economic Co-operation and Development, 1999)

4.1.3 Keep personal email address private

When subscribing to mailing lists, chat rooms and any other place where the user's email address is required use the "side" email address. Be cautious to unsubscribe from a mailing list and replying to any spam email, as it will

only notify the sender that the user's actual email account is active. (McCandlish, 2001)

4.1.4 Be cautious when revealing personal information

Only provide trusted people and sites with personal information, and verify privacy policies and seals before signing up. Some websites provide rewards or simply require one to sign-up; in such cases provide the third party with one's "side" profile. (McCandlish, 2001)

4.1.5 Be conscious of Internet security

When dealing with financial matters make sure it is done through a secure connection. Browsers normally indicate this with a lock in the status bar. This will decrease the chances of malicious scripts capturing sensitive information or redirecting one to an unsecured site. The Internet is not secure by design, therefore: unless the connection is encrypted, everything is sent in clear text. When downloading free software one should by wary of exploit software that is secretly embedded within some downloaded applications. (McCandlish, 2001)

4.1.6 Beware of creating a personal web page

Personal web pages can reveal a great deal; an online biography or Curriculum Vita wipes out one's anonymity. Web spiders or robots capture these types of information. Instead, direct queries to a "side" email address. (McCandlish, 2001)

4.2 Technical Controls

Suggested technical controls would include the following:

4.2.1 User's Computer

In today's connected, email-reliant world, anti-virus software is imperative. This should be installed on every user's computer and updated regularly. This will ensure protection against Trojan horses, embedded scripts, etc. The operating system security is an important component of ensuring privacy and anonymity. For instance, configuring user profiles, rights and permissions, as well as implementing encryption, would be essential.

4.2.2 Internet firewall

A firewall is as important as anti-virus software. There is a wide variety of personal firewalls available for a single user environment (e.g. Outpost firewall). The firewall should allow policy based access control. These access controls should be the implementation of the corporate security policy and reinforce the operational controls. The firewall should also perform network address translation (NAT) and filter Internet requests. NAT allows an entire network to seem like one computer from the outside as the firewall translates all the IP addresses from the LAN to the single IP address of the firewall on the Internet (Figure 2) and visa versa. NAT will therefore prevent third parties from identifying a specific user within the LAN. (Tiny Software, 2002)

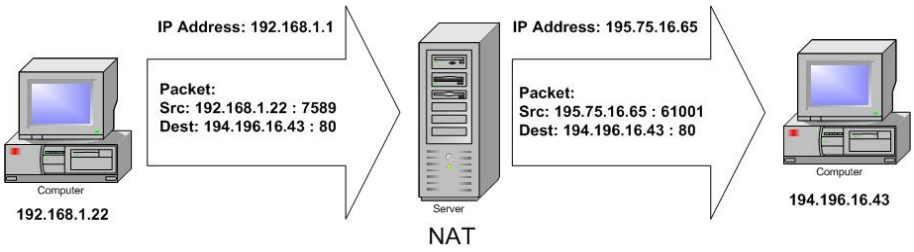

Figure 2. Network Address Translation (Tiny Software, 8 March 2002)

The Internet service provider should preferably provide this service, as NAT would be ineffective when implemented by a single computer connected to the Internet.

4.2.2.1 The gateway should enforce the following:

- It must "silently" block unknown traffic from the outside. The term silent refers to the way the firewall reacts to packets that arrive; when a packet arrives at the firewall that has no known connection from inside, it is denied access without sending a "port closed" message. This will give the requester the perception that the computer is off or not connected.
- It must only allow particular applications to connect from inside and only at certain ports. This will firstly ensure that the application only functions at the predefined ports. Therefore even if a known application has been modified, e.g. by a malicious plug-in, it will only be allowed to connect at the predefined port. Secondly, any malicious software on a user's computer trying to connect to the outside will be blocked, as its port is not predefined.

- It should block or filter threats such as cookies, active content, etc. This will significantly decrease threats related to web browsing as the banner advertisements, web bugs and active content will not be loaded by the user's browser. Some firewalls close the HTML-enabled email security hole by blocking cookies and active content.

(Goncalves, et al, 1997)

A number of firewall applications need to be updated regularly. This ensures that the applications' monitoring capability include protection for latest security exploits. Updates will ensure that the latest application security holes are patched. It is beneficial to use software companies that do regular and emergency updates. There are applications available that scan the Internet for the latest updates of all the software installed on one's computer.

4.2.3 The Internet

Headers of Internet protocols, such as HTML and SMTP, include fields that contain identifiable information. Filtering these headers by encrypting or removing such fields will increase the user's anonymity. Anonymizing services, such as Anonymizer, use this means to achieve anonymity. Applications are slowly becoming available to perform such tasks. Such applications will eliminate the disadvantages of online filters, for example, reduced bandwidth, cost, etc. (Eckert, et al, 2001)

To further increase one's anonymity on the Internet, the gateway should operate as an onion router. Ideally, one's Internet Service Provider should provide one with an onion-routing proxy server. As discussed onion-routing hides the user's request from all parties involved.

5. CONCLUSION

The Internet is growing daily and creating a demand for new technologies. Consequently, threats will only increase if security, including anonymity, is not considered during the development of these technologies. Anonymity, similar to other security related matters, requires a technical fortification of the network connection to outside networks. Nevertheless, anonymity would be impossible without the assurance of operational controls enabling the user to maintain anonymity. These controls should be

integrated with any existing information security framework. There are many anonymizing services available, but, because they are not under one's control, one has to trust them to do as one expects. Solving the problem locally allows more control but requires more responsibility.

The prevention of anonymity *attacks* demands further in depth technical synthesis. Possibly improving the onion-routing model by integrating different other characteristics.

6. REFERENCES

Anonymizer. (2002). How Anonymizer Protects You. [Online]. Available: http://www.anonymizer.com/

Eckert, C. & Pircher, A. (2001) Internet Anonymity: Problems and Solutions. London : Kluwer Academic Publishers.

Federal Trade Commission. (2000). Online Profiling A Report to Congress Part 2 Recommendations. [Online]. Available: http://www.ftc.gov/os/2000/07/onlineprofiling.htm

Federal Trade Commission. (1998). Privacy Online: A Report to Congress. [Online]. Available: http://www.ftc.gov/reports/privacy3/priv-23a.pdf

Fielding, R., Gettys, J., Mogul, J.C., Frystyk, H., Masinter, L. & Leach, P. (1999). RFC2616: Hypertext Transfer Protocol – HTTP/1.1. [Online]. Available: http://www.w3.org/Protocols/rfc2616/rfc2616.html

Goldschlag, D., Reed, M. & Syverson, P. (Cited on 22 May 2002). Anonymous Connections and Onion Routing. [Online]. Available: http://www.onion-router.net/Publications/Briefing-1996.pdf

Goncalves, M., Donkers, A., Harkin, J., Hart, H., Niles, K., Toyer, K. & Willis M. (1997). Internet Privacy Kit. USA : Que Corporation

Internet Security and Privacy for Activists and Citizens. (Cited on 14 February 2002). [Online]. Available: http://www.astalavista.com/privacy/library/misc/security-privacy.shtml

Junkbusters. (Cited on 14 March 2002). What's News at Junkbusters. [Online]. Available: http://www.junkbusters.com/new.html

McCandlish, S. (2001). EFT Top 12 Ways to Protect Your Online Privacy. [Online]. Available: www.eff.org/Privacy/eff_privacy_top_12.html

Nua Internet Surveys. (2002). [Online]. Available: http://www.nua.ie/surveys/

Online Privacy Alliance. (2002). Rules and Tools for Protecting Personal Privacy Online. [Online]. Available: http://www.privacyalliance.org/resources/rulesntools.shtml

Organisation for Economic Co-Operation and Development. (1999). Guidelines on the Protection of Privacy and Transborder Flows of Personal Data [Online]. Available: www1.oecd.org/dsti/sti/it/secur/prod/PRIV-EN.HTM

Privacy Rights Clearinghouse. (2000). Fact Sheet 18: Privacy In Cyberspace. [Online]. Available: http://www.privacyrights.org/fs/fs18-cyb.htm

Reiter M. & Rubin A. (Cited on 8 May 2002). Anonymity Loves Company. [Online]. Available: http://www.research.att.com/projects/crowds/

Rennhard, M., Rafaeli, M., Mathy, L., Plattner, B. & Hutchison, D. (Cited on 22 May 2002). An Architecture for an Anonymity Network [Online]. Available: http://www.tik.ee.ethz.ch/~rennhard/publications/WetIce2001.pdf

Roha, R. R. (August, 2000). Prying Eyes. Kiplinger's Magazine. [Online]. Available: http://kipnew-live.worldweb.net/magazine/archives/2000/August/managing/e-privacy.htm

South Africa. Dept. of Communications. (2002). South African Electronic Communications and Transactions Bill [Online]. Available: www.gov.za/gazelle/bills/2002/b8-02.pdf

Tiny Software. (2001). Reference Guide: WinRoute Pro.

Wilsker, I. (June, 2001). Cookies, SpyWare, and other Privacy Threats on the Net. I/O Port Newsletter. [Online]. Available: http://www.tcs.org/ioport/jun01/cookies.htm