Online Profiling Project - Comment, P994809 / Docket No. 990811219-9219-01

Pursuant to the policies announced by the National Telecommunications and Information Administration (NTIA) of the United States Department of Commerce and the Federal Trade Commission (FTC) , Andrew Shen, on behalf of the Electronic Privacy Information Center (EPIC), formally submits the following reply comments in response to the statements made at the Public Workshop on Online Profiling that took place on November 8, 1999 and as a suggestion for further action on the part of the NTIA and FTC.

Andrew Shen
Electronic Privacy Information Center (EPIC)
Tel: (202) 544-9240
Fax: (202) 547-5482
Email: shen@epic.org

Executive Summary

Online profiling -- the recording of detailed online behavior of an individual -- is fast becoming the preferred business model for online advertisers. In many cases, these profiles can be connected to personally identifiable information and as such should be governed by Fair Information Practices. Based on recent experiences in related areas, there is little reason to believe that self-regulation will provide sufficient protections from the potential privacy risks of online profiling. Strong privacy protections enforced by independent agencies are necessary.

Statement

What is online profiling?

Online profiling is the practice of online advertisers to record online behavior for the purpose of producing targeted advertising. Online behavior encompasses every sort of interaction that can take place on the World Wide Web, including the pages viewed, searches conducted, and products or services purchased. Targeted advertising takes into account past online behavior in order to present Internet users what they think will be the most likely to buy. While advertisers have always tried to target audiences based on market demographics, the detailed knowledge about an individual and the ability to tailor and present advertisements for a single person is unprecedented. These online profiles can be connected to personal information -- such as a name or address -- if such information is provided directly to the advertiser or to affiliated websites.

Many online companies try to gather information from their customers, but online advertisers utilize online profiling differently than the majority of online companies in two respects. First, online advertisers operate in the background of web surfing. As most people do not go online to see banner advertisements, they are often unaware that the advertisements -- even if never clicked on -- can place cookies on their computer. Second, online advertisers operate on multiple sites so that profiles can be constructed based on consumer behavior on completely separate sites.

The information contained within these online profiles is not insignificant. Through the observation of sites visited, products viewed, and purchases made, one can infer important personal attributes, such as medical conditions, sexual preferences, and political or religious beliefs. As more people interact on the Internet in more significant ways, their actions and movements will reveal more about them.

Many people go online to find information about sensitive topics that are of great importance, such as medical conditions. Part of the reason for going online to search is so that they can anonymously find out such information. Probably much to their surprise, the presence of online advertising networks threatens that anonymity. Suppose someone wants to find out about breast cancer either because he or she is worried about the condition or knows someone who might be at risk. The individual might start such a search at a popular portal such as http://www.altavista.com. In addition to the search engine, altavista.com offers a directory of sites such as a category devoted to Health & Fitness. Under the Health & Fitness section, there is a heading for conditions. Clicking

on conditions takes you to another page where one can select the link for cancer. On the page listing sites that offer information about cancer, there is a heading for breast cancer. Upon selecting breast cancer, there is a compilation of different sites that provide information. One of these is http://www.allhealth.com, an affiliated website of http://www.ivillage.com, which is targeted towards female Internet users. The allhealth.com page focused on breast cancer offers tips on risks and symptoms associated with breast cancer, message boards for people with the disease, and chats with health experts. The person searching information has found all the information he or she may have been looking for after traveling over five different pages run by two different companies. What the individual may not know is that the advertisements on all these pages were provided by a single company, DoubleClick Inc. The DoubleClick banner advertisements on all these pages can track the individual through his or her search. While neither altavista.com nor allhealth.com may not know the beginning or the end of the search, DoubleClick has the ability to produce a complete record of the entire process. To be fair, DoubleClick publicly says that they do not track such information, but the fact is that they could.

Besides the invasion of privacy that online profiling presents, there may be far-reaching consequences for the online consumer due to such massive collection of information. A recent report by the Consumer Federation of America[1] points out that targeted advertising in an electronic medium may unfairly impact consumer choice by allowing the vendor to take advantage of detailed information about consumer interests and purchasing history.

[1] http://www.consumerfed.org/digitaltv.pdf

What are the privacy concerns about online profiling?

Online profiles can be connected to personally identifiable information, creating a detailed history of online behavior that can attributed to an offline person.

It is true that online profiles can remain anonymous, but many companies attempt to correlate personally identifiable information (PII), such as a name and address, with consumer profiles. This can be done in two ways -- either by giving an online advertiser PII or by combining information collected online with other sources of information. This submission will use one major online advertiser, DoubleClick Inc., as an example of how online advertisers make great effort to make online profiles personally identifiable.

One example of how DoubleClick Inc. attempts to obtain PII is through one of its websites, http://www.NetDeals.com. NetDeals.com is a sweepstakes site where one can win $1,000,000. To enter the sweepstakes, all one has to do is enter a name, mailing address, email address, and birthdate. The privacy policy which can be seen by scrolling through the text box at the bottom of the page, clearly states that the site is part of the DoubleClick network and that the personally identifiable information is compiled:

> "When you register on NetDeals you provide us with personally identifiable information such as your name, home address and e-mail. We combine that information with other information about you that is available to us. This includes other personally identifiable information and certain non-personally-identifiable information . . ."[2]

Another site part of the DoubleClick network that collects PII is http://www.iaf.net/. This site, the Internet Address Finder (IAF), does not offer anything quite as lucrative as

NetDeals. Instead it offers a way to look people up on the Internet and a way to "call" people over the Internet. Claiming close to 7 million listings, the form asks for a name, email address, organization, and domain (this is your organization's WWW address, or the part of your email address after the @ sign -- for example, epic.org). Two clicks away from the form collecting personal information is a vague statement on how the information entered into IAF can be connected to other data in the hands of DoubleClick. Under Frequently Asked Questions, there is the statement that "all information submitted by you for inclusion in IAF will become part of a database, the compilation copyright for which is owned exclusively by DoubleClick, Inc."[3] There is no mention that this information can and will be connected with non-PII obtained through cookies.

Besides the collection of personal information from other Internet sites, DoubleClick Inc. can obtain personal information by merging with other offline companies. On November 23, DoubleClick stockholders approved a merger with Abacus Direct Corporation. Abacus Direct is a company that does research on catalogue buying and has been doing so for many years. Abacus has within its possession 88 million 5-year buying profiles that contain such personal information such as name, addresses, and family makeup. DoubleClick Inc. is now able to combine any of the information it already has in its own possession with Abacus' database of personally identifiable information.

---

[2] http://win.netdeals.com/getaway/

Why self-regulation does not protect consumer privacy.

The self-regulatory plan presented by the Network Advertising Initiative (NAI) does not adequately protect consumer privacy. Providing notice and opt-out is insufficient due to the invisibility of online profiling and the fact that profiles are often associated with PII. Furthermore, previously established self-regulatory bodies have been reluctant to pursue privacy violations committed by companies and provide little reassurance that privacy will be protected.

Online advertising occurs in the background of most Internet activities, whether it is viewing a page or purchasing a product. Most consumers are not aware that banner advertisements are owned by third-party companies and are actively collecting information. Since the process occurs so invisibly, notice and opt-out are not enough. If most consumers are not aware of the existence of online advertisers, how can they be expected to read privacy policies and actively request to stop the collection of information? The collection of information must only take place after the consumer is made aware that it is taking place. Furthermore, notice and opt-out unfairly places a huge burden on the consumers to find out what companies are doing with their personal information and stop those actions if they disagree with them. Consumers should expect a baseline standard for their privacy from all companies.

Since the information contained within online profiles can be connected to offline individuals, the full array of Fair Information Practices should apply to the situation. As

---

[3] http://www.iaf.net/frames/faq.htm

stated in the 1980 Organization for Economic Cooperation and Development (OECD) Privacy Guidelines[4], "personal data" applies to "any information relating to an identified or identifiable individual (data subject)." In governing the collection and use of such personal data, the OECD defined guidelines, often known as Fair Information Practices, comprised of the following principles:

> Collection Limitation -- data collection occurs with the knowledge and consent of individual.
> Data Quality -- data collected is relevant and accurate.
> Purpose Specification -- use of information must be known at time of collection.
> Use Limitation -- data should not be used for purposes other than the ones specified at the time of collection.
> Security Safeguards -- reasonable measures are taken to protect data from unauthorized use.
> Openness -- individual should readily be able to avail themselves of data collection practices and be able to contact the entity collecting the data.
> Individual Participation -- ability of the individual to know if information has been collected and access the information if it does exist.
> Accountability -- data collector should be held accountable for failing to achieve any of the above principles.

These eight principles that collectively form Fair Information Practices establish consumer control over personal information. Without such control, there is no real consumer privacy.

In addition, since information collection often occurs without the awareness of the individual Internet user, extra means should be taken to establish individual control over data collection. The collection of information over the Internet does not necessarily require any active participation of the individual. While many consumers are reasonably aware in the offline world when information is being collected via actively purchasing an

[4] http://www.oecd.org//dsti/sti/it/secur/prod/PRIV-EN.HTM

item or filling out a form, simply being online can now result in information collection. The more invisible the process, the more difficult it should be to obtain PII.

Business alliances that try to present uniform compliance with basic privacy principles fall far short of enforceable privacy protection. The principles that they propose to support do not assume that consumers have the right to control their information. For example, one such organization -- TRUSTe -- states that to obtain a "trustmark" (signifying a participant of TRUSTe) companies have to display "whether the user has an option to control its dissemination"[5].

TRUSTe has also proven to be incredibly reluctant to take action against privacy violations. Consider the recent controversy over the RealJukebox software, distributed by RealNetworks.[6] RealNetworks is a member company of TRUSTe. RealNetworks' privacy policy -- certified by TRUSTe -- did not disclose the existence of the Globally Unique Identifier (GUID) almost until the news about RealJukebox privacy invasions was released and still does not reveal that email addresses are transmitted along with GUIDs to RealNetworks[7]. Without tangible penalties, other than a wealth of bad publicity, companies will continue to maneuver without the knowledge of the consumers. This is not the first time TRUSTe has refused to launch an investigation into an apparent privacy violation. Just this past March, when Microsoft revealed that it used GUIDs to render documents created by Microsoft Office software and visitors to Microsoft

---

[5] http://www.truste.org/about/about_faqs.html
[6] http://www.wired.com/news/reuters/0,1349,32244,00.html
[7] http://www.junkbusters.com/ht/en/real.html

operated websites personally identifiable, TRUSTe refused to investigate or condemn Microsoft for invading consumer privacy.

Why legal protections can protect consumer privacy.

As the recent history of self-regulation demonstrates, companies that claim that they respect privacy will continue to abuse the privacy of the average consumer. Federal legislation establishing a baseline standard for online privacy is necessary because there are no other entities that can enforce privacy on such a wide scale. Also, the government has proven itself capable of legislating quickly and effectively on other Internet issues. In addition, independent polls of Internet users show widespread support for government action in online privacy.

While there exist many privacy advocacy groups like Electronic Privacy Information Center (EPIC), Junkbusters, and the Center for Media Education (CME), these groups have limited resources and authority. A privacy protection agency with greater resources for oversight and the authority to levy penalties is necessary. Online businesses, especially online advertisers that operate on the premise of collecting as much information as possible from the consumer, cannot be expected to police themselves on consumer privacy.

Moreover, the federal government has proven itself able to protect the consumer in the online environment. Roughly a year ago, Congress passed the Child Online Privacy Protection Act (COPPA). On October 20, the FTC unanimously approved rules on Child

Online Privacy which were widely applauded as a well-designed apparatus for acquiring parental consent before children under the age of 13 give out personal information. Although the Internet is a fast-moving, evolving medium, it is possible to legislate wisely and not stand in the way of future innovation. By establishing general procedures and principles such as those laid out in the 1980 OECD Privacy Guidelines, legislation can create a baseline standard for the privacy of all individuals online.

Evidence also shows that individuals online prefer government action, rather than industry self-policing, on online privacy. A poll commissioned by Business Week[8] and performed by Louis Harris & Associates and Dr. Alan Westin in March 1998 found that given three different possible approaches towards protecting Internet privacy (industry self-policing, government recommendations on privacy, or legislation), 50% of computer users chose legislation as opposed to only 21% in favor of self-regulation. Another poll conducted by Georgia Tech Research Corporation[9] found that in October 1998, 71.4% of survey respondents agreed strongly (40.6%) or somewhat (30.8%) with the statement "there should be new laws to protect privacy on the Internet."[10]

[8] http://@4WZJy4cASJ*2SwAA/1998/11/b3569104.htm
[9] Copyright 1994-1998 Georgia Tech Research Corporation. All rights Reserved. Source: GVU's WWW User Survey www.gvu.gatech.edu/user_surveys
[10] http://www.cc.gatech.edu/gvu/user_surveys/survey-1998-10/graphs/privacy/q59.htm