Spy Watch: A Tool for Transparency in Web Tracking

Neel Guha
neelguha@gmail.com

Gunn High School,
Palo Alto, CA

Abstract. This is in part a position paper and in part a description of a Google Chrome privacy extension we built. We make the case that our first goal when tackling the issue of privacy and third party tracking of user's browsing should be to increase transparency. We describe a browser extension we have built, that enables users to see which third party has observed what about them. We discuss the results of a survey we conducted which suggests that users are divided on the issue of online web tracking. While some users expressed strong uncomfort at being tracked, others expressed tolerance. User opinion was also divided on how to block third party sites.

1 Introduction

Third party tracking on the web, through cookies and other means, is one of the most widespread forms of user data collection. Since much can be inferred about a user by mining the pages he/she has visited (i.e. interests, gender, purchases, trips, location, etc), third parties can construct profiles of users and gather sensitive personal information. Astonishingly, third party tracking occurs on the most well known and respected sites. The top ten (by traffic) online newspapers in the US have an average of 13 third party tracking sites on their homepages! The surreptitious nature of this tracking, often occurring without the user's consent or knowledge, is alarming.

In recent months, the debate over online third party tracking between privacy advocates and advertisers has become polemic. When Mozilla announced plans for its "Do Not Track Browser", the International Advertising Bureau compared the move to a "nuclear first strike" and claimed that Mozilla was disrupting a business model under the "cloak of privacy" (`http://goo.gl/5bNS7`). Conversely, some privacy advocates have adopted the extreme position that all third party tracking should be banned. A more nuanced understanding that users have very different attitudes towards online tracking is missing.

2   User's Rights

Our position is that transparency is the first step required to tackle third party tracking. We believe that every user has the right to know:

1. Which sites are currently tracking them;
2. What data each of these sites has gathered about them.

Such knowledge will empower users to better understand who is collecting their data and how it is being used. Equipped with such knowledge, users can make more judicious decisions regarding tracking. We believe a population of informed users can contribute on the policy level and impact relevant legislation.

3   Spy Watch

In accordance with our position, we have created Spy Watch (`http://goo.gl/wd7Lx`), a Chrome extension and Firefox addon that seeks to increase transparency in online tracking. Through a clear, easy-to-use interface, Spy Watch provides users with answers to the following questions:

1. For a given page visit, which are the other sites (watchers) that know of this visit (Figure 4)?
2. Which are the sites that are watching the user (Figure 5)?
3. What are the URLs a third party site has watched the user visit (Figure 6)? In the example shown in the figure, we can see that doubleclick.net knows that a user has visited nytimes.com and espnfc.com.

While there are other privacy extensions (such as Ghostery and DoNot-TrackMe), Spy Watch differs in several critical aspects. Firstly, Spy Watch allows users to view on which pages each third party has watched them (number 3 above). We believe this information enables users to determine how invasive a certain third party may be. Secondly, Spy Watch does not block cookies or tracking sites. As a result, Spy Watch does not break or interfere with the functionality of certain sites relying on cookies.

4   Survey

User feedback on Spy Watch led us to the realization that different users can have very different views on online third party tracking. We investigated further by conducting an online survey (found at `http://goo.gl/WId8n`) whose aggregate results are shown in Figures 1, 2, and 3. Participants in the survey came from a wide range of demographics that included high school students, retirees, social science faculty, computer science faculty and IT professionals. As the results show, reactions to the level of online tracking varied significantly (Figure 1). While some participants had little to no concern about being tracked, others

expressed concern at the level of tracking. Further, users had different criteria for blocking sites (Figure 3). Some participants wanted to block all sites, whereas others prefered being able to specify which sites to block. Still other participants supported blocking sites based on the fraction of the browsing history known by that site. It is clear that any uniform policy applied to all users regarding which cookies should be blocked and which should be allowed will leave a large fraction of the user population dissatisfied.

5   User Directed Blocking

We support a policy that would enable the user to block and allow cookies of their choice in a simple and easy manner. Extensions like Ghostery and DoNot-TrackMe require significant user effort in selecting cookies to block/allow while providing little information regarding the invasiveness of each third party site. We are currently working on additions to Spy Tracker to enable this kind of functionality.

The blocking/allowing of third party tracking sites is comparable to the filtering of spam in email. Like spam, some tracking sites are harmful and should not be allowed to monitor the user. Other sites however, like regular email, can serve a beneficial purpose and help the user. The severity of a spam filter for a user is set by the user's behaviour and standards. Cookie spam should be viewed and tackled like email spam.

6   Conclusion

In conclusion, we believe that our primary goal should be to increase transparency so that a greater fraction of the user population is aware online tracking. When the privacy community understands these users' reactions, the course of action required will become clearer.

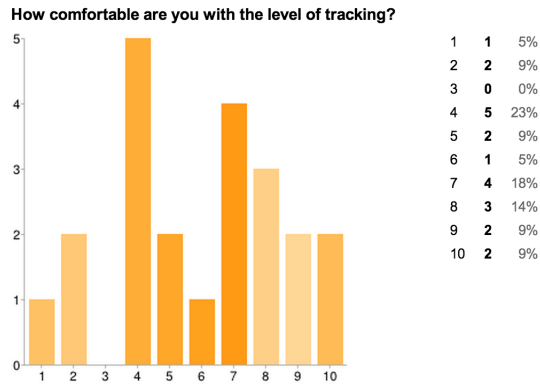**How comfortable are you with the level of tracking?**

1	1	5%
2	2	9%
3	0	0%
4	5	23%
5	2	9%
6	1	5%
7	4	18%
8	3	14%
9	2	9%
10	2	9%

Fig. 1: User reactions to the level of tracking on a scale from 1-10, with 1 being "absolutely fine" and 10 being "horrible"

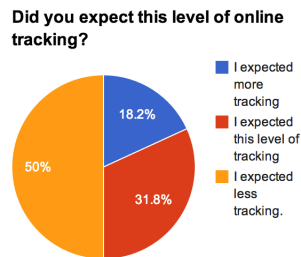**Did you expect this level of online tracking?**

- I expected more tracking — 18.2%
- I expected this level of tracking — 31.8%
- I expected less tracking. — 50%

Fig. 2: A comparison of user expectations of online tracking before and after using the extension.

Which of the following would you want to do?

- Block all tracking sites — 22.7%
- Don't block anything
- Don't block sites that I have visited (eg: if I visited facebook, let facebook track me)
- Block sites which know more than a certain percentage (eg: 10%) of all the pages I have visited — 31.8%
- Only block sites that I specify — 36.4%
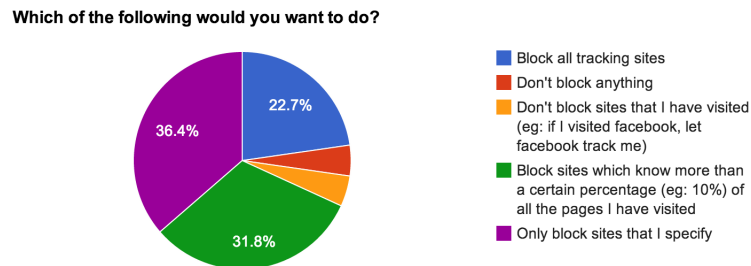
Fig. 3: User criteria for blocking/allowing sites to track
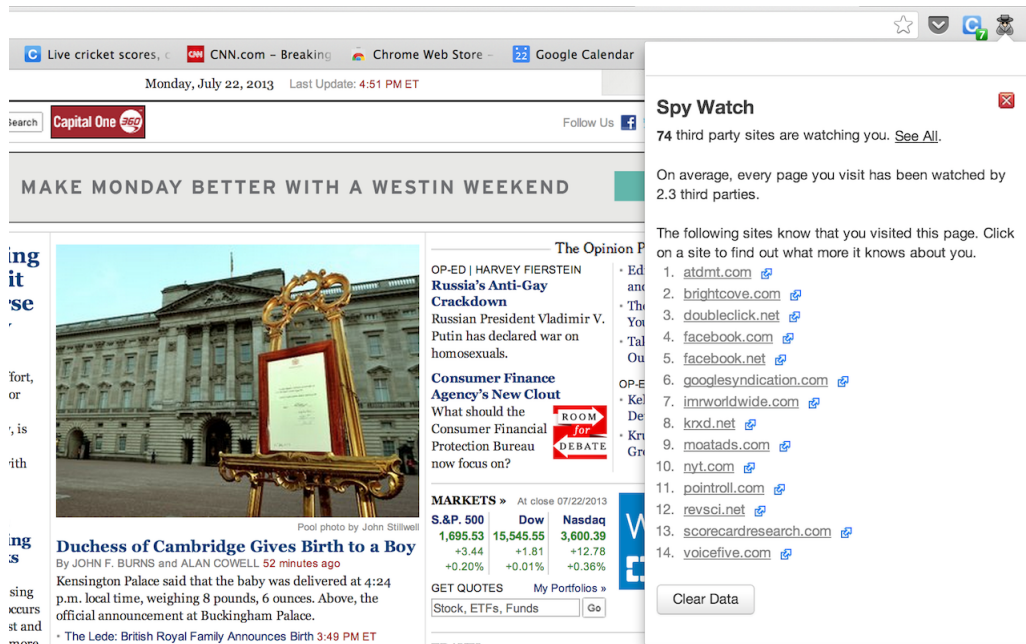
Fig. 4: A list of all the sites that could be watching a page (in this example, nytimes.com).
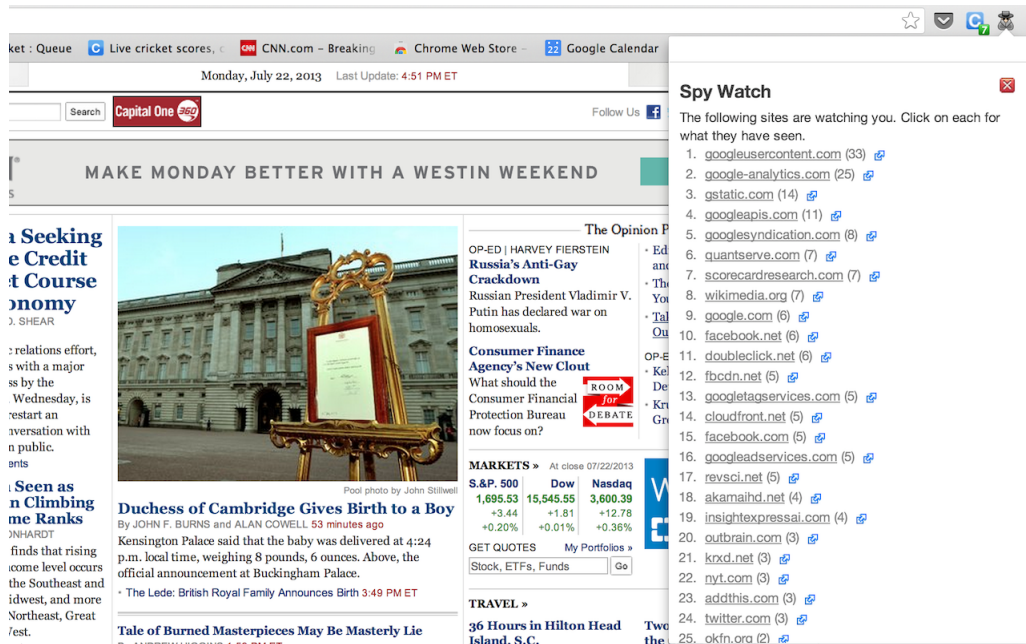
Fig. 5: A list of all the sites that are watching a user.

Fig. 6: A list of all the pages that were watched by doubleclick.net.