Meet your Online Tracking Profiles with TrickTrack

Martin Degeling
Ruhr-University Bochum
martin.degeling@rub.de

1. INTRODUCTION

We would like to present TrickTrack, a work-in-progress prototype for transparency about and obfuscation of profiles based on browsing histories. TrickTrack is a browser addOn for Firefox that provides insights about information online tracking services can conclude from web browsing histories and offers means to obfuscate these profiles based on users' own preferences.

2. THE PROBLEM WITH ONLINE TRACKING AND PROFILING

Profile based advertising, recommendation and tailoring of information has become a widespread functionality to make services more effective and efficient. They are adopted by a broad range of online services from social networks to shopping websites. However, the underlying automatic personalization of information requires the processing of more data about users than they are possibly aware of. In most cases personalization takes place without explicit consent and in ways that are not transparent to users and is therefore limiting informational self-determination. Personalization is accompanied with a loss of control. It has long been an issue in system design [1] that users are hardly aware of what information is disclosed to which service and when, especially when data is collected without their explicit consent.. Therefore they do not have a sufficient understanding of what the assumed characteristics of their digital identities – also known as *data doubles* [2] - are, nor are there tools that effectively support this understanding [3]. While some see benefits in personalized recommendation most users just dislike and ignore common forms of *targeted advertisements* also refered to as *online behavioural advertisement (OBA)* [4],[5]. Negative effects based on this technology like discrimination [6], [7] or filter bubbles [8] are less recognized and users are often not even aware of these possibilities. On a more general level the effects of tracking can be described as a threat to the autonomy of internet users, while privacy can be regarded as a mechanism to protect individuals' autonomy and self-determination [9]. Tracking technology is not only used for delivering personalized ads but also to alter the content of websites. Some even assume that it is used to alter prices to motivate customers to buy a product [11]. Although it can be argued that it is in the interests of users the real drivers for this development are the needs of service providers to increase sales or reduce risks. However this can, in the end, result in an exclusion of specific users without their beeing transparent about it e.g. when products are simply not offered to those that live in a specific region [11]. Although a growing number of tracking blocker plugins try to mitigate tracking and therefore these effects, the number of users of ad and tracking blocker is currently only about 5% of those that surf the web [12]. In addition blockers often reduce the functionality of a website and do not offer any means to understand the profiles that are built and the effects they have.

3. THE NEED FOR TRANSPARENCY

Especially in the European Union researchers and activists argue for more regulation of profiling, especially with regard to transparency [13], but resistance from the industry against new regulation ist strong. We do support the demand for more transparency and privacy literacy [14] and therefore built TrickTrack as a tool to empower users to gain knowledge about and control over (at least some of their) profiles. TrickTrack is a browser addOn for Firefox that visualizes profiles which can be created based on the websites a user has visited. It combines the browser history with aggregated information about tracking services and data from audience measurement sites. At the moment the information provided by tracking services is limited. Only a few offer individual access to stored information[1] while the majority only refer to standardized privacy policies. And even those, that offer transparency and means of control try to turn the users rights into their benefit by encouraging them to increase the accuracy of the profiles created without consent. Those pages are dependent on services goodwill and might as well disappear sooner or later. In addition we know from [14] that transparency pages do not represent the profiles used by the services. TrickTrack therefore comes with it's own data set of characteristic connected with webpages and combines them based on the browser history available.

TrickTrack offers insights on three levels: statistics about the amount of data available in the users browser history, an estimation of an interest profile and a summary of the most likely socio-demographic characteristics of the user.

3.1 Tracking Information

Figure 1 show the starting page of TrickTrack with information extracted from the browsing history to make users aware of the amount of data available. It shows the number of websites visited, general cookies and tracking cookies.

3.2 Interest Profile

On a second page an estimated interest profile is shown (see Fig. 2). The interest categories are based on google interest categorization used for their ad services. It was extracted by simulating users and observing which interests are related with which web page on Googles Ad-Settings page (see fn 2). Additional information is provided to explain how these interests

Figure 1: Tracking Information

1 See https://www.google.com/settings/ads or http://bluekai.org/registry

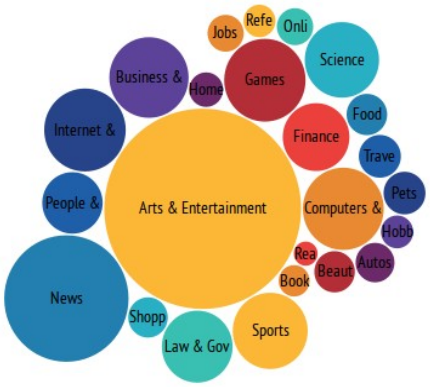

Figure 2: Interest Profile

are determined and how they are used by google and other advertisers.

3.3 Socio-Demographic Profile

A third page shows a number of socio-demographic characteristics. Based on multiple sources for audience measurement TrickTrack estimates the most likely characteristic for the attributes gender, age, income, education, ethnicity and wether or not the user has kids (see Fig. 3).

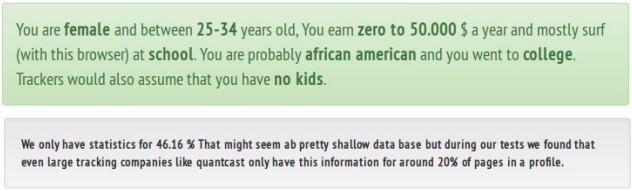

Figure 3: Profile

3.4 Anti-Profiles and Obfuscation

TrickTrack offers basic means for obfuscating and influencing profiles. On the users request the addOn opens a tab and automatically surfs to pages (a) found be inverse to the interest profile or (b) support profile characteristics choosable by the user. Although tools for obfuscating web-navigation trails like TrackMeNot have been criticized for theoretical attack vectors we see obfuscation as a way to empower users to interact with their profiles [17].

4. Evaluation and Plans

The first prototype was evaluated with ten users to test utility and usability. The reactions were positive in general but several problems related to methodological impreciseness were found.

First, to not be dependent on the data provided by tracking services, we simulated users to build up a database of website/interest relations. For many pages this data is unstable, meaning visiting website A does not always lead to the same interest B. Therefore the profiles shown can not be considered correct in a way that they reflect the same interest profile presented by Google. Although it is stated in the text, and users reported that TrickTrack as well as Googles interest profile do not fully represent them, they get confused by the difference, which can also lead to "tracking is ok, because it's wrong"-attitude. Second, we have not formally proven that obfuscation works effectively. Since the obfuscation is designed in a way that it may only prove itself users may feel less tracked, although they are not.

Nevertheless users reported that they appreciate to have more transparency about their profiles and that it is fun to influence them. By letting TrickTrack surf to "cliché" sites that were identified to support a diverse profile the recognized that they were confronted with pages outside of their "filter bubble". In the ongoing development the problems are addressed and a final version could be released at HOTPETS.

4. REFERENCES

[1] J. P. Hourcade, A. Cavoukian, R. Deibert, L. F. Cranor, and I. Goldberg, "Electronic Privacy and Surveillance," in *ACM Conference on Human Factors in Computing Systems*, New York, NY, USA, 2014, pp. 1075–1080.

[2] K. D. Haggerty and R. V. Ericson, "The surveillant assemblage," *Br. J. Sociol.*, vol. 51, no. 4, pp. 605–622, 2000.

[3] P. Leon, B. Ur, et.al., "Why Johnny Can't Opt out: A Usability Evaluation of Tools to Limit Online Behavioral Advertising," in *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, New York, 2012, pp. 589–598.

[4] M. Malheiros et. al., "Too Close for Comfort: A Study of the Effectiveness and Acceptability of Rich-media Personalized Advertising," in *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, New York, 2012, pp. 579–588.

[5] L. Agarwal, N. Shrivastava, S. Jaiswal, and S. Panjwani, "Do Not Embarrass: Re-examining User Concerns for Online Tracking and Advertising," in *Proc. of the 9th SOUP*, New York, 2013, pp. 8:1–8:13.

[6] O. H. Gandy, *The Panoptic Sort - A Political Economy of Personal Information*. Boulder u.a.: Westview Press, 1993.

[7] L. Sweeney, "Discrimination in Online Ad Delivery," Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 2208240, Jan. 2013.

[8] E. Pariser, *The Filter Bubble: What The Internet Is Hiding From You*. Penguin UK, 2011.

[9] B. Rössler, *The value of privacy*; Polity, 2005.

[10] T. Vissers, N. Nikiforakis, N. Bielova, and W. Joosen, "Crying Wolf? On the Price Discrimination of Online Airline Tickets," *HotPET Symp.*, 2014.

[11] A. Danna and O. H. Gandy, "All That Glitters is Not Gold: Digging Beneath the Surface of Data Mining," *J. Bus. Ethics*, vol. 40, no. 4, pp. 373–386, 2002.

[12] Pagefair, "Adblocking goes mainstream," 2014.

[13] M. Hildebrandt, "The Dawn of a Critical Transparency Right for the Profiling Era," *Stand Alone*, pp. 41–56, 2012.

[14] A. Datta, M. C. Tschantz, and A. Datta, "Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination," *arXiv:1408.6491*, 2014.

[15] F. Brunton and H. Nissenbaum, "Vernacular resistance to data collection and analysis: A political theory of obfuscation," *First Monday*, vol. 16, no. 5, Apr. 2011.