Background and specification progress report.

Daniyar Irishev
K1763572

Contents

Background and context.

Online behavioural advertising (OBA) is one of the biggest phenomena on the world-wide web that has seen much more use in the recent years. It is a big part of the experience for the daily internet user whether they are using social media, news websites or entertainment websites. It is also crucial for the entities that allow their platform to be used for online advertising as it is usually a major part of their income.

However, targeted advertising brought attention of the public to issues regarding data privacy and morality. Even with attempts to solve it such as the GDPR, users do not feel that their data is safe. Currently advertisers cannot explain the reason behind showing a particular ad to a particular user as most parts of the process are automated. Use of provenance has potential to track users' data in a way that the ads are explainable in the future. Specifically, automatically generating provenance records within advertising cookies. Considering that there exists PROX-XML – an XML standard for PROV, it can be implemented with relevant ease.

Currently there are several methods of data gathering and tracking used for OBA. This project will focus on one particular technique that uses cookies. Cookies themselves have different variations depending on implementation and purposes. The main goal of this project is to utilise provenance in order to provide more background to targeted ads that are based on cookies, such as the origin, who/what handled them and other.

Requirements

The system will be required to automatically generate provenance for every tracking cookie created for a specific user. During the lifetime of this cookie, the provenance needs to be

updated automatically in events of the cookie being sent to some party and it being used for OBA.

Several cookies can be used to give a user a specific "tag" that lets advertisers show relevant ads based on them. Provenance should hold record of these events as well.

Relevant literature.

- Cranor, L. (2012). Can Users Control Online Behavioral Advertising Effectively?. *IEEE Security & Privacy Magazine*, 10(2).
- Herosmyth. (2018). *Who Took the Cookie? The Science Behind Targeted Ads.* [online] Available at: https://www.herosmyth.com/article/who-took-cookie-science-behind-targeted-advertising [Accessed 13 Nov. 2019].
- Luc Moreau and Paul Groth (2013). Provenance: An Introduction to PROV. Synthesis Lectures on the Semantic Web: Theory and Technology. Morgan & Claypool.
- PROV Model Primer. (2013). [Blog] *W3C Working Group Note*. Available at: https://www.w3.org/TR/prov-primer/ [Accessed 17 Dec. 2019].
- Hardcastle, F. (2019). *A Sociotechnical Exploration of Online Behavioural Tracking and Advertising Technologies and Practices*. Ph.D. University of Southampton.
- Soltani, A., Canty, S., Mayo, Q., Thomas, L. and Hoofnagle, C. (2009). *Flash Cookies and Privacy*. Berkeley: UC Berkeley School of Law.