CS133 Essay

Who is watching me online?

# AN ANALYSIS OF THE USE OF BEHAVIOURAL ADVERTISING METHODS

May 11, 2021

Abstract

This essay explores and evaluates the impact of online behavioural advertising (OBA), and how it is used by advertising networks to deliver suitable adverts using interest and content based tags. It discusses the ethics of building interest-based profiles, the legality and compliance surrounding the processing of required data, alternatives to behaviorally targeted advertising, and how these techniques can both enhance and damage people's online experience. It also discusses how current legislation (GDPR) tries to protect consumers from data harvesting, and concludes that future legislation should further protect consumers from privacy violations.

Felix Bowyer 2064363
University of Warwick

I declare that the work presented in this essay is my own, and that it has not been submitted for assessment on any other module.

Online behavioural advertising (OBA) is a targeted advertising technique which associates behavioural tags with internet users, in order to match users with the most relevant adverts online. It has seen increased usage, and whilst having been shown to be highly effective, it is not without controversy surrounding the ways that relevant data is collected, sorted, and used to build a profile of a user. Often, big data techniques are used to sort relevant information. Big data is characterised by large volumes and high velocity of data, as well as a large variety of type and source of data. In the case of OBA, this could involve analysing someone's entire online activity, to create a characterisation of their interests and generate relevant advertising tags. Current legislation, such as GDPR, attempts to regulate the handling of sensitive and personal data, but in many cases fails to deal with some relevant legal and ethical issues. There is an active and ongoing discussion around what the correct way to handle such data is, and around the justification of using such data at all.

There are a two main advertising methods in use online. The first is a contextual advertising system, which scans the web content currently being shown to somebody, and bases the served advertisement exclusively on that. The second is a behavioural advertising system, which has seen increased usage over the last 10 years. Google's "Inside AdSense" blog[1] describes it as recognising the types of web pages visited by a user, and using that to build a profile of a user. For example, if somebody were to visit a high number of sports pages, Google's AdSense network would label them as a sports enthusiast. As discussed by Brahim et al.[2], online behavioural advertising (OBA) has some clear advantages for advertisers over random (mass) and contextual advertising methods. OBA saves a lot of money for advertisers; companies don't waste money advertising towards people who will likely never engage with the product, and each served advert is much more likely to be interacted with, which means it is a lot cheaper to find people interested in the advert. Additionally, OBA may reduce competition between unrelated products - it allows the same advertising space to be used to advertise different products to different people. However, it also has limitations. OBA required to have some way to track internet users. This is often done directly with an account that you log in with to use a company's services, such as a Facebook account, however this tracking can also be achieved by correlating IP addresses, login locations and other factors. In their paper, Brahim et al. modelled behavioural advertising against a random model, and found that profit from advertisements can be sustained at a much higher level, and with a much lower budget, by using targeted advertising methods. Their methodologies only use a basic model, which may not perfectly model the real world - however, it should provide a reasonable estimate, or at least pick up on the trend that OBA has significant advantages; a fact that is reinforced by advertiser's increasing willingness to use it. What their research does not consider, however, is the impact that behavioural advertising has on the users being advertised to.

Behavioural advertising requires building up a profile on anybody being advertised to in order to be effective. This in itself comes with technical, legal and moral challenges.

It is quite difficult to harvest, process and use the mass amount of data needed to create a behavioural advertising system, and storing this data in a responsible way is just as difficult. Advertisers collect as much data as possible in order to profile users, and may, intentionally or not, collect extremely sensitive and personal data. Carrascosa et al[3]. discuss how advertisers often collect sensitive information such as sexual orientation, health and political beliefs. Whilst this information can greatly help to target advertisements, it is also information that people may wish not to be profiled on, and feel uncomfortable sharing with an advertising network. Possibly the largest impact of OBA is that it has become the de-facto expectation that everything you do online is being watched by an advertising algorithm. Large companies such as Google, Facebook and Microsoft often sell each other user data from their advertising networks, which helps them build their advertising profiles on people that don't even use their services. It has become an important question with the rise of targeted advertising as to whether it is right to have to share this degree of information in order to use any online service, particularly without explicit given consent, such as when IP addresses, advertising cookies and other factors are used to track someone.

Some effort has been made by lawmakers in recent years to limit how companies can collect, process and store personal data. Notably, in 2016 the European Union introduced the General Data Protection Regulation[4], which regulates data protection and online privacy in the EU and EEA. The UK's ICO provides a broad definition of what constitutes personal data[5], defined as any data which allows a person to be identified, either directly or indirectly through combination with other data. Under GDPR, unless a person has given informed consent, or there is a legal basis to do so, their personal data can not be processed. It also gives a person the right to access to the personal data being held about them by a company. Whilst GDPR makes a great effort to govern the processing of personal data, it is another matter to monitor and enforce it. In a study on 38 companies, Urban et al.[6] found that only 55% disclosed information within the legally required time of 30 days, and only 34% were able to send a copy of the data in time. Whilst this was a small scale study, it shows that many companies and advertising agencies are non-compliant. Sakamoto et al.[7] conducted a study into whether GDPR had a significant impact on the way that advertising agencies track users who have opted out from tracking. They found that whilst half of advertising agencies stop tracking immediately after enabling an OBA opt-out, many start tracking again once a user continues browsing, resulting in no statistically significant evidence that GDPR has changed the way users are tracked online. They also define two opt-out states, expected and compliant opt-out. An expected opt-out aligns with a user's expectations, and stops web tracking, whereas a compliant opt-out only minimally complies with guidelines, and continues to collect user data after an OBA opt-out; in effect, web tracking is not stopped. Both before and after GDPR was implemented, of the 133 studied advertising agencies, roughly half used a compliant opt-out mechanism only. This suggests that despite a slight decrease in tracking after implementation, GDPR has loopholes, and lacks the proper provisions to stop non-consensual behavioural advertising. WhoTracksMe, a privacy-focused tracker

analytics group[8][9], recommends a proposed GDPR 2.0 should be machine-readable, and should require companies to publish in a standard location the following; a plain-text privacy policy, a list of third parties present on the site and their purpose, a standardised set of information regarding a company's data protection officer, and an open, transparent list of data incidents and court cases involving the mismanagement of personal data.

Legislation to date has focused on allowing users online the choice to be exempted from behavioural advertising. However, there are some advocates who think that OBE is overall a good thing, and others who want it to be banned entirely. Those who support the continued use of OBE claim that it can be beneficial to consumers, as well as companies. For example, there are certain online platforms, such as YouTube, or social media sites such as Instagram and Reddit, whose usage wholly depends on a form of behavioural advertising. YouTube uses a video recommendation algorithm, based on behavioural advertising, which decides which videos may appeal to a user, based on videos they have previously shown interest in. It is this very feature that has driven the popularity of YouTube; without personalised recommendations, it's many niche communities would be unable to grow, and it would be nowhere near as popular today. However, opposition claims that this algorithm is exploitative. A paper from Google employees, Zhao et al.[10], suggests that YouTube is actively working on making it's algorithm more addictive[11]. Bishop claims that the YouTube algorithm helps promote unrealistic body standards for women[12], by promoting and rewarding beauty vlogger content. Whilst her paper focuses partially on possibly accidental flaws in the YouTube algorithm, it demonstrates that behavioural advertising can have negative, reinforcing effects on people's mental health. Additionally, YouTube and other social media sites have been accused of radicalising people by recommending them conspiratorial or extremist political content[13], and during the recent pandemic, many social media platforms have been accused of spreading medical misinformation. Much of this evidence is speculative, or documented sensationalistically, though it is certainly undeniable that recommendation algorithms are having a polarising effect in online politics, through creation of echo-chambers. Whilst it is true that behavioural advertising has some benefits in recommendation, in both social media as well in as product recommendation on online stores, in both of these cases the main reason for using it is to boost profitability for the company running the service, which does not directly benefit the user; the biggest advocates of OBE are those who directly make a profit from it.

There are some alternatives to OBE that have been proposed, which avoid some of the more controversial ethical aspects, whilst keeping the personalised recommendation aspects which allow certain services to exist. Toubiana et al.[14] propose Adnostic, an advertising architecture which preserves user privacy, which is currently implemented in a Firefox extension. It allows advertisers to serve ads to relevant people based on behavioural tags, without violating the person's identity by exposing their advertising profile to the ad network. It does this by building an advertising profile in-browser, as opposed to one being built by the advertising network. Visited web pages are scanned for

content, and tags are extracted, building an interest profile based on Google's Adwords. When an advert needs to be shown to a user, the advertising network sends many adverts, and the most suitable advert is selected by the user's computer (not the ad network), using the local profile. This way, advertisers are reassured that their advert is being shown to relevant people, without a specific person's interests and behavioural tags being exposed to anyone else. However, this does come with some disadvantages - adverts can take up to 10 times longer to load, and it still doesn't solve the problem of exploitative advertising algorithms. It also creates a convoluted system for advertising payment. When an advert is shown, it is cryptographically verified in such a way that does not expose the user's identity, which could be computationally expensive over time.
Another way that OBE is avoided is by making it impossible to show relevant ads. AdNauseum[15] is a browser add-on created in protest of targeted advertising. Howe et al.[16] describe it's head on approach in their paper. Instead of trying to avoid targeted advertising altogether, AdNauseum quietly clicks random adverts in the background, and hides them from the user. This causes the advertising networks to build an incorrect profile about someone , by filling it with garbage behavioural tags, meaning it cannot tell what a user's real interests are. This makes it completely impossible to offer targeted advertising. However, this method is ethically questionable - advertisers often pay per click on adverts, meaning they still pay for adverts that are never seen, in a practice nicknamed "click fraud". Whether this is justified is a moral question up for debate.

Despite lacking legal legislation to stop behavioural advertising, some progress has been made by other companies prioritising customer privacy. When Apple released iOS 14 for their smartphones in April 2021, they made it a requirement for applications to ask for permission to track certain device data. In response, Facebook and Instagram started displaying pop-ups asking users to help keep the service free of charge, by allowing them to track their data[17]. Whilst it is true that restricting the tracking of data could cause service providers to make less profit, both Facebook and Instagram were still able to make money and provide a free service before they started tracking user's data for OBE, and there is no reason that they wouldn't be able to now.

We have seen that there are many valid reasons concerning the morality and practicality of behavioural advertising. On the whole, OBE benefits large corporations and advertising networks, whilst harming internet users by exploiting their interests to make a profit or push certain content. There are cases where behavioural advertising can work in everybody's favour; in services that absolutely rely on behavioural recommendations, it can vastly improve a user's experience, and benefit the service provider. However, even in this scenario, it's usage should be limited, and designed carefully to prevent exploitative algorithms. It is absolutely imperative that stronger legislation is bought in to more effectively limit the data harvesting and tracking that OBE makes necessary, and to ensure everybody's right to privacy online. Behavioural advertising has a place online, but given there are more privacy-oriented alternatives, it's use should be limited.

## References


[1] Google. Driving monetization with ads that reach the right audience. https://adsense.googleblog.com/2009/03/driving-monetization-with-ads-that.html. Accessed: 2021-05-11.

[2] D. Laussel N. B. E. Brahim, R. Lahmandi-Ayed. *Is targeted advertising always beneficial?*, pages 551–585. 01 2014.

[3] R. Cuevas V. Erramilli N. Laoutaris J. M. Carrascosa, J. Mikians. I always feel like somebody's watching me - measuring online behavioural advertising. arXiv:1411.5281v3.

[4] European Union. Document 32016r0679 (General Data Protection Regulation). https://eur-lex.europa.eu/eli/reg/2016/679/oj. Accessed: 2021-05-11.

[5] ICO. What is personal data? https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/key-definitions/what-is-personal-data/. Accessed: 2021-05-11.

[6] Tobias Urban, Dennis Tatang, Martin Degeling, Thorsten Holz, and Norbert Pohlmann. A study on subject data access in online advertising after the gdpr. 07 2019.

[7] Takahito Sakamoto and Masahiro Matsunaga. After GDPR, still tracking or not? understanding opt-out states for online behavioral advertising. In *2019 IEEE Security and Privacy Workshops (SPW)*, pages 92–99, 2019.

[8] WhoTracksMe Privacy. GDPR - what happened? https://whotracks.me/blog/gdpr-what-happened.html. Accessed: 2021-05-11.

[9] Rémi Berson Josep M. Pujol Arjaldo Karaj, Sam Macbeth. Whotracks.me: Shedding light on the opaque world of online tracking. arXiv:1804.08959v2.

[10] Zhe Zhao, Lichan Hong, Li Wei, Jilin Chen, Aniruddh Nath, Shawn Andrews, Aditee Kumthekar, Maheswaran Sathiamoorthy, Xinyang Yi, and Ed Chi. Recommending what video to watch next: A multitask ranking system. In *Proceedings of the 13th ACM Conference on Recommender Systems*, RecSys '19, page 43–51, New York, NY, USA, 2019. Association for Computing Machinery.

[11] Karen Hao. Youtube is experimenting with ways to make its algorithm even more addictive. https://www.technologyreview.com/2019/09/27/132829/youtube-algorithm-gets-more-addictive/. Accessed: 2021-05-11.

[12] Sophie Bishop. Anxiety, panic and self-optimization: Inequalities and the youtube algorithm. *Convergence*, 24(1):69–84, 2018.

[13] Zeynep Tufekci. Youtube, the great radicalizer. https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html. Accessed: 2021-05-11.

[14] D. Boneh H. Nissenbaum S. Barocas V. Toubiana, A. Narayanan. Adnostic: Privacy preserving targeted advertising.

[15] https://adnauseam.io/. AdNauseum browser extension.

[16] Daniel C. Howe and Helen Nissenbaum. Engineering privacy and protest: a case study of adnauseam. Published by Cornell University.

[17] Chris Matyszczyk. Facebook threatens to make ios users pay. please do it, mr. zuckerberg. https://www.zdnet.com/article/facebook-threatens-to-make-ios-users-pay-please-do-it-mr-zuckerberg/. Accessed: 2021-05-11.

[18] Edward C. Malthouse and Hairong Li. Opportunities for and pitfalls of using big data in advertising research. *Journal of Advertising*, 46(2):227–235, 2017.

[19] David C. Dinielli Fiona M. Scott Morton. *Roadmap for a Digital Advertising Monopolization Case Against Google*. 05 2020.

[20] Rebecca Balebako, Pedro Leon, Richard Shay, Blase Ur, Yang Wang, and Lorrie Cranor. Measuring the effectiveness of privacy tools for limiting behavioral advertising. 05 2021.

[21] Dark patterns: How UX design tricks you into giving away your privacy. https://cliqz.com/en/magazine/dark-patterns-how-ux-design-tricks-you-into-giving-away-your-privacy. Accessed: 2021-05-11.

[22] Lauren Valentino Bryant. The youtube algorithm and the alt-right filter bubble. *Open Information Science*, 4(1):85–90, 2020.