Online Tracking and Web Advertising

a survey and discussion by

Josh Berlin and David Wetterau

May 1st, 2014

Abstract

In this paper we discuss the issues surrounding online tracking and the use of targeted advertisements. We present the history of advertising and then discuss the tracking mechanisms that are used on the web today. We then analyze several factors that influence online tracking from a corporate perspective and highlight the impacts that such practices have on users and specifically user privacy. Finally, we summarize relevant regulation in the United States and Europe and the potential impacts that such regulation could have on the industry in the future.

Introduction

The roots of modern advertising date back to the Egyptians, who used papyrus to make posters and primitive flyers, and to Indian rock paintings from as early as 4000 BCE ("Advertising"). In ancient societies, where most people were unable to read, it was common for tradesmen to advertise their products using large billboards with pictures of their trade or product. As society evolved and people became more literate, with the introduction of industrial printing presses and mass market newspapers, publishers started selling advertising space to subsidize the price of publications. By the beginning of the 20th century, advertising agencies that designed full blown campaigns including billboards and posters started popping up.

The rapid technological innovations in media including the introduction and popularization of radio and television provided advertising agencies and companies with opportunities to place advertisements on a wider scale than ever before. These technologies enabled companies to reach customer bases across the country and build images that incentivized loyalty and increased market share. Further, as governments recognized the power of such evolving mediums, they began using advertising as a way to reach their citizens. In 1942, United States President Roosevelt created the War Advertising Council, now known more commonly as the Ad Council, to encourage national support for the United States' involvement in World War II. After the war, the Ad Council cemented its role as most know it today by aggregating donated ad space in media sources around the country and publishing public service announcements ("Ad Council"). The growth of national and multinational chains by the 1970s was arguably the result of their abilities to advertise at scale.

Although the technologies that existed at the end of the 20th century allowed companies to advertise across broad geographic regions, there was a limited extent to which a company could gauge who saw their ads or how relevant an ad was to viewers. Companies could only target demographics and age groups, not individual users. For example, children's breakfast cereal ads often aired with cartoons. Nonetheless these technologies still play an important role in marketing today, as evidenced by the millions of dollars that companies pay each year for coveted super bowl spots.

The introduction of computers, and most importantly the Internet, changed the granularity of advertising arguably more than any innovation had in the past. The first banner ad was displayed by HotWire to make a bit extra cash from its website in 1994 ("Web Banner"). Leading up to the dot-com boom, new Internet technologies allowed search engines

and websites across the Internet to offer “relevant” advertising like never before. Small businesses could now advertise without incurring major costs. Even the neighborhood plumber could be listed, and show up as a targeted advertisement to a potential customer in their area. Large companies also benefited from targeted advertising, leading to overall market expansion for producers. More recently, the increased mining of social network data and browser cookies have allowed more fine-tuned advertising over time.

Companies are willing to pay such a premium for web advertising because of the ability to individually target users and market segments using Internet advertisements compared to their abilities to do so in newspapers, radio, or television. Although many users are aware that they are being tracked to some extent online, they are likely unaware of how detailed these digital profiles really are. Such practices reasonably evoke worry among consumers, and according to a Pew Research Center study, “68% of Internet users believe current laws are not good enough in protecting people’s privacy online” (Rainie et. al).

Modern online tracking evolved from a simple mechanism designed to make the web stateful. Online advertisers have realized that they can leverage this originally benign technology to identify and track Internet users. Once obtained, this data can be used to provide consumers with targeted ads. Although this technology benefits advertisers, it raises serious privacy concerns for consumers. Historically, with the introduction of new technologies, governments have created policies to protect both businesses and consumers; however, there are no regulations currently in place that completely address online tracking technologies. There have been efforts among governments, consumer protection groups and concerned businesses to develop regulatory standards for online tracking, but none of the proposed standards have been widely accepted.

The rest of this paper is organized as follows. We provide a brief overview of web tracking techniques and discuss why companies want to track users. Subsequently, we consider how online tracking can be harmful and contrast it with possible benefits. Lastly, we talk about current and past efforts to regulate online tracking and their effectiveness as well as what users can do to stop themselves from being tracked online.

Online Tracking Techniques

Contemporary online advertisers use a variety of different mechanisms to track users across many different websites and build a digital profile about individuals. Although some modern tracking techniques are becoming increasingly complex, many of the original, more basic mechanisms are still used in practice today.

One of the first tools that companies leveraged to track users is the browser cookie. The browser cookie initially evolved as an innocent mechanism to store and access information about user sessions across visits to the same website. Many websites use browser cookies to save information about shopping carts or login information so that users don’t need to login each time they visit a website from the same computer. The use of browser cookies is restricted to the same origin, which means that cookies can only be accessed by the domains that created them. For example, Amazon can access all the cookies on a user’s browser that have been created by Amazon.com, but it can’t read or write those for Google.com. Because cookies are restricted to the domains (companies) that created them, they were not seen as a privacy concern when they were first created. The ability to

return to Amazon later in the day and see the same shopping cart that had been created earlier was appreciated as a time saver to most users.

Online advertising companies quickly found a way to take advantage of browser cookies to give them more information on a user's web browsing habits across different websites. Common websites on the Internet today contain content from many different domains and companies: advertisements from an ad network, 'Like' buttons from Facebook, and embedded videos from YouTube among others. On a website where content originates from multiple domains, each domain may create and access its own cookies. For example, if a website contains an ad from DoubleClick, DoubleClick might first read and send the DoubleClick.com cookie to the DoubleClick server. The DoubleClick server might recognize the user, log this request as well as the website the user is visiting, and serve up an ad customized to that user. Cookies that are created by domains and companies displaying content on a webpage other than the website a user is visiting directly, such as those created by an ad served by DoubleClick, are referred to as "third party cookies". As a result of the prominence of several major advertising networks across the Internet, these advertising networks are able to access their third party cookies on many pages as users browse websites that use these ad providers. These advertising networks can subsequently use this information to build profiles for individual users. Many unsuccessful lawsuits have been placed to try to protect users from companies that use such third party cookies to generate and then sell profiles of users as they browse the Internet (Tene et. al 291). Luckily, browser designers came to the rescue and have made it easier for users to delete or block these sorts of cookies on a site by site basis. In fact, some modern browsers block third party cookies by default.

Although online advertising companies may have been temporarily dismayed by the creation of mechanisms in modern browser to delete user cookies, they soon discovered a means to regenerate cookies deleted by a user. A common technique that companies have devised to do so is with so called "super cookies". Super cookies take advantage of the fact that although the user might clear some forms of local session storage such as cookies, they might not empty all the other resources that web content can access. Historically, one of the most prominent uses of supercookies involved storage associated with Flash plugins. Flash is a plugin used commonly across the Internet, frequently for animating and displaying videos and advertisements. Content using the Flash plugin in browsers is allowed to create and access "local shared objects". Companies figured out that they could take advantage of this functionality to simulate cookies, but with 25 times the capacity of normal cookies, no expiration date, and no user accessible browser restrictions (Soltani et. al). For a long time, clearing the information stored in user browsers did not clear the storage associated with the Flash plugin, something likely a result of the fact that the Flash plugin is managed and installed separately from most browsers. Companies soon figured out they could use this loophole to store copies of cookies in Flash storage and use this storage to regenerate normal browser cookies when users visit a website after clearing their normal cookies (Soltani et. al 2). Adobe, the creator of Flash, eventually took note of this hole and integrated Flash plugin management with browsers so that clearing browser storage and cookies also clears Flash storage. Although Adobe fixed this issue with Flash, it wouldn't be surprising if similar holes exist in many of the other plugins used in modern browsers.

A concerned user might be able change their cookie settings and disable all their browser plugins to avoid being tracked online via the aforementioned techniques. Unfortunately, there are other techniques such as browser fingerprinting that aren't so easy to avoid. The concept behind browser fingerprinting is that, like a human fingerprint, browser

installations and the systems on which they run are subtly different and unique. As discussed in “Cookieless Monster: Exploring the Ecosystem of Web-based Device Fingerprinting”, many commercial browser fingerprinting tools used on websites can tell a remote server information about the user’s system including what operating system, web browser, web browser plugins, and even which fonts are installed on the system (Nikiforakis et. al 3). This information can be combined at a server to form a unique fingerprint that can be used to identify the user without the need for cookies or any storage on the user’s system. This fingerprinting technology was originally invented as a way for banks to combat fraud: if the user could be matched to their own unique “online fingerprint”, then the bank was more assured that the user was actually the owner of the bank account. However, this technology can also be used for the arguably less noble purpose of user tracking. In efforts to educate people about browser fingerprinting and also conduct research on fingerprinting effectiveness, the Electronic Frontier Foundation set up a website, www.panopticlick.eff.org, where users can go and see if their browser can be uniquely identified using common fingerprinting techniques. Both authors of this fingerprinting paper, using different computers, had unique browsers with more than 21 bits of identifying information. Assuming numbers from Cisco that estimate 8.7 billion devices on the Internet as of 2012 (“Connections”), if an ad company wanted to uniquely identify each device on the Internet they would only need around 33 bits of uniquely identifiable information. Nikiforakis et. al additionally show in their paper that “over 800,000 users, who are currently utilizing user-agent spoofing extensions, are more fingerprintable than users who do not attempt to hide their browser’s identity” (Nikiforakis et. al 2). Fingerprinting is potentially more disturbing than other tracking techniques because of its ability to easily identify users who try to avoid tracking, theoretically making such users more vulnerable to enhanced tracking or even easier identification as a result of their attempts to stay anonymous.

The techniques discussed thus far have mostly tracked a user’s online behavior by taking advantage of browser mechanisms, but there are certainly other tools advertising companies can use to learn about a user’s online behavior. The form of tracking users outside of browser technologies that is arguably most prominent in the news currently is the use of deep packet inspection by Internet service providers (ISPs). This technique has been likened to “postal employees opening envelopes and reading the letters inside” (Tene et. al 298), where a user’s “letters” are data packets and the “postal employees” are the ISPs. Deep packet inspection allows ISPs to build extensive and complete profiles about a user’s online behavior and requires significantly less effort for the party doing the tracking since the ISPs do not need to worry about the user switching devices or erasing a tracking cookie. Privacy groups have been successful at combating this intrusive technique when it is used without the user’s knowledge, but that doesn’t mean uneducated consumers can’t still opt in to these practices (Diaz). Companies such as AT&T currently offer 30% off the price of their new high speed fiber service in exchange for permission to perform this data collection. Once they have this information, AT&T claims that “they may use your personal information to direct another advertiser’s ad to you, but that advertiser would never have access to your Personal Information” (Brodkin). Advertisers are likely willing to pay AT&T at least as much as AT&T discounts the customer for these profiles, so it seems probable that AT&T is providing these companies with information they find valuable and difficult to obtain on their own.

Ultimately, there are many different tools that advertising companies use to try to track user behavior online, and even as browser and plugin developers fix certain vulnerabilities that enable such tracking, new and improved technologies are developed all the time to provide companies with more information that can be used to identify and profile users.

Why Companies Want to Track Users

As discussed in the previous section, companies go to great lengths to track and identify users across the web. In this section we will discuss some of the motivations for companies to engage in user tracking.

A recent study sponsored by the Digital Advertising Alliance “found that advertisers pay three times more per impression for cookie driven ads, and seven times more if the cookie is 90-days old” (Bachman, “DAA Study…”). According to the same study, “Small websites [are] dependent on independent interest-based ads for 60 percent of their ad revenue”. In response to critiques of targeted advertising, companies that are strong proponents and users of targeted advertising often cite these technologies as the economic backbones of the Internet as we know it today. Ultimately companies are willing to pay significantly more for targeted advertising than traditional advertising or even sponsored clicks because they believe that targeted advertising has significantly more impact on the users they reach.

In 2012, Orbitz published results from their sales that showed that “people who use Apple Inc.’s Mac computers spend as much as 30% more a night on hotels” (Mattioli). As a result of these statistics, Orbitz decided to use a user’s operating system, part of the information that is sent along with the user agent to each website a user visits, as a factor in the ordering of hotels that are displayed by default. Further, studies by market research firms have shown that the “average household income for adult owners of Mac computers is $98,560, compared with $74,452 for a PC owner” (Mattioli) and that “over 41% of Mac owners were 34 years old or younger” (Mattioli). Another result of a similar study suggests that “users of tablets -- who by a large majority are iPad owners -- tend to place bigger online orders on average than users of laptops and desktops”. Just knowing a small bit of information, which didn’t involve any tracking cookies or other technology, but utilizes a very basic form of targeting, helps Orbitz and other retailers sell more expensive goods to consumers, and ultimately increases their bottom line.

“How much can Behavioral Targeting Help Online Advertising”, a paper published by Microsoft Research in 2009, attempted to quantify a few statistical tenets of behavioral ad targeting. In the paper, the researchers discuss several types of statistical analysis they performed on the search results and behaviors of about 6 million users from a commercial search engine. Based on this dataset and the experiments they conducted, the researchers came to several conclusions about the potential impacts of behavioral targeting in advertising. First, they found that “users who clicked the same ad can be over 90 times more similar than users who clicked different ads” (Yan 262). Next, they found that by modeling the behavior of users in these groups and by targeting ads they could possibly enhance the click through rates of ads (rates of users who visit the link from an ad) by “as much as 670%” (262). Lastly, in contrast to the statistic from the earlier article about advertisers paying more for older cookies, the authors suggest that short term user behavior may be a better metric for segmenting users than long term behaviors.

In a study sponsored by the Network Advertising Initiative, Howard Beales outlines some of the economics behind online targeted advertising. Beales points out that if “advertising better matches consumer interests, consumers are more likely to respond to the

message, and advertisers will be willing to pay more for ads delivered to such an audience" (Beales 1). Beales' research, which is based partly on statistics collected from 9 of the top 15 of online advertising networks circa 2009, compared advertising revenues and conversion rates of targeted advertising compared to conventional 'run of network' banner style ads that are not targeted. Beales' data shows that conversion rates, which indicate the percent of users who make a purchase after clicking on an ad, along with price, more than double with the use of behavioral targeting.

Although more broad than targeted advertising alone, "big data" is the driving force behind determining how to target ads to individuals. In an interview with CNN, Scott Howe, the CEO of Acxiom, which is one of the largest data aggregators in the market, suggests that "companies like Acxiom are trying to get intelligent about what you might be interested in and who you are… [and that] data generates tremendous values for both people and businesses" (Morris). Indeed, in an entry on "Why Big Data is the new competitive advantage", the Ivey Business Journal discusses how big data and the insight it gives corporations will play a key role in future enterprise. In a scenario posited by the journal, a "next-generation retailer will be able to track the behavior of individual customers from Internet click stream, update their preferences, and model their likely behavior in real time. They then will be able to recognize when customers are nearing a purchase decision and nudge the transaction to completion" (McGuire).

The use of targeted advertising helps companies present ads that are more relevant to users, and research has shown that the use of such strategies can have a significant impact on user interaction with online advertising. Companies take advantage of these technologies to grab the attention of users and, in the end, try to increase their bottom lines.

Online Tracking Considered Harmful

Although there are certainly many benefits that companies, advertisers, and even governments might experience from tracking users across the Internet, these benefits do not necessarily align for users in the same way that they do for these organizations. In this section we discuss some of the downsides of online tracking from the user's perspective.

Many studies have been performed to analyze what users think of online tracking. Unsurprisingly, "surveys have consistently shown opposition to third parties collecting and using browsing activity" (Mayer et. al 4). When a company uses a method of online tracking, it's arguable that they intrude on what most users believe to be their rights to privacy. The information companies gather through online tracking can be used irresponsibly in a multitude of ways to inflict harm upon users.

One of the ways online tracking can harm users is economically. As mentioned earlier, research has shown that in general, Apple device owners have a higher median income and are generally younger than the users of other devices. With simple techniques, websites can tell whether or not a user is using an Apple device through browser fingerprinting. Online retailers such as Orbitz admit to "showing Mac users pricier hotel options in some cases after finding they tend to spend as much as 30% more a night on hotels on the site" (Mattioli). By prioritizing the display of more expensive hotels to Mac users, Orbitz is likely making more money and also satisfying some set of Mac users by targeting them. At the same time, a Mac user might miss out on low price deals if they aren't shown them, resulting in an economic harm to the user because of user tracking and targeting.

Although users in the previous example may lose money, there are certainly other more serious scenarios that might result from online tracking. Many people are probably somewhat aware they are being tracked online, but most users likely consider their online behavior to be anonymous unless they explicitly identify themselves. Significant research has been done by privacy researchers to show that users can be easily identified even from “completely anonymous” datasets (Narayanan). Once a user on the Internet has been successfully identified, irresponsible companies or other third parties can sell this information to organizations the user has never interacted with. An organization that obtains such data and has no responsibility to the user can leak this information which could obviously cause some embarrassing truth about the user to become known and cause them personal harm. Possible nightmare scenarios could include the identification of users who visit adult websites.

In extreme cases, online tracking can be used by a government to censor and suppress the government’s subjects. The Wall Street Journal recently published an article about how the Libyan government purchased software from the French company “Amesys” that “was a major part of a broad surveillance apparatus built by Col. Gadhafi to keep tabs on his enemies” (Sonne et. al). This software performed deep packet inspection to look at the Internet traffic of Libyan citizens and to monitor their online activities. In Libya, the use of this tracking technology led to nationwide government censorship. “Amesys says its ‘strategic nationwide interception’ system can detect email from Hotmail, Yahoo and Gmail and see chat conversations on MSN instant messaging and AIM”, which the Wall Street Journal claims “helped Libya sow fear as the country erupted in civil war” (Sonne et. al).

To summarize, online tracking that was originally intended to be used beneficially for more relevant advertising for users and safer online banking can be misused to harm Internet users as well. Simple demographic research combined with online tracking can lead to online marketplace injustice. De-anonymized traffic records of Internet users can be used for blackmail, extortion, or public ridicule, and deep packet inspection has already been used to secretly censor and suppress entire nations.

Why Tracking is a Good Thing

Although there are certainly many possible downsides to web based tracking, at the end of the day, tracking users to provide targeted advertisements can also provide a better, more personalized Internet browsing experience.

Many of the currently published academic studies on web tracking focus on the possible technological implications of web tracking mechanisms, and relatively few attempt to evaluate the reactions of users to targeted or relevant ads versus their conventional, run of network counterparts. As pointed out by Sableman et. al in a legal brief on consumer attitudes towards behavioral advertising, some of the designs of prominently cited studies that do involve interviewing users about their opinions “raise concerns about the validity of the results obtained” (Sableman et. al 101). Of the two frequently cited studies that Sableman et. al discuss, they point out that one was conducted exclusively offline using “questions that .. did not present real-life situations” (101), and the other presented many questions heavily focused on consumer privacy rather than the perceived utility of targeted ads. In the list of responses to the posed questions, the second study even provided three negative choices alongside two positive choices, a practice that likely subtly leads users to choose a more

negative option. Resultingly, as an author interested in writing a paper on the negative aspects of targeted advertising, it's easy to design a user study where the results inevitably present common perceptions of relevant advertising and related technologies in the desired way.

After discussing the flaws with many existing online advertising studies, Sableman et. al present the results of a study they designed to answer the question of "What do Internet users feel about online behavioral advertising?" (103). Participants were asked to "rate their feelings" (104) after visiting different categories of websites where they were presented with conventional ads or ads targeted to their prior behavior on the Internet. In order to attempt to compare the results to prior studies, they also asked some of the same questions that had been asked in the previous ones. Based on the results from these experiments, "relevant ad conditions were significantly more desired than irrelevant ads .. [and] study results strongly supported [their] hypothesis" (106). Although there is certainly room to question any study, Sableman et. al provide a convincing experiment that shows, at minimum, that consumers to do not unilaterally oppose receiving relevant ads.

Other studies have recently backed up claims that many people prefer advertising targeted to their online behavior. In a study conducted by the Digital Advertising Alliance, "consumers actually prefer targeted ads with nearly 70 percent responding that they'd like at least some ads tailored directly to their interests" (Bachman, "Targeted"). In another study where users were presented with video ads for travel website Kayak.com that were either tailored to their local airport or generic, users were significantly more receptive to the locally relevant advertisement ("Personalized").

Why did companies start using targeted advertising in the first place? In a previous section, we discussed the benefits of targeted advertising over conventional advertising from the advertisers perspective, but it is also important to discuss why ads are such a prominent part of the Internet in the first place. Ads make services like Google, Facebook and Twitter free for the end user while still allowing such companies to pay their employees well, provide value, and innovate.

The Internet has matured into a model of websites being subsidized by ads, and it is somewhat difficult to compare ad supported models to paid subscriptions because there exist relatively few paid consumer services in the current Internet marketplace. However, an interesting parallel to the development of the Internet that can be used as a proxy is the development of the mobile app market over the last five years. In a blog post discussing the evolution of mobile app market, Mary Gordon from Flurry analytics, a company that collects data on mobile app usage, points out the trend of consumers to choose free, ad-subsidized apps over their paid counterparts. She highlights that "as the outcome of consumer choice: people want free content more than they want to avoid ads or to have the absolute highest quality content possible" (Gordon). Gordon ultimately suggests that the compounding of such consumer choices has resulted in more mobile app developers deciding to make their applications ad subsidized rather than paid.

Another place in which consumers are starting to choose a subsidized model is in their relationship with their Internet service providers. As mentioned earlier, in December 2013 AT&T started advertising a service where customers can pay $30 less than the standard rate per month if they agree to allow AT&T to use deep packet inspection on their traffic and serve them super-targeted ads. AT&T is not the first Internet provider to offer such a service, but they clearly feel that providing their customers with more heavily targeted ads is a competitive substitution for $360 in income from the subscriber each year. As consumers are likely to give up some privacy to save $360, it seems likely that this may become common practice for

Internet service providers sooner rather than later.

As long as internet users repeatedly choose ad subsidized models over paid services, advertising agencies and companies that want to attract consumer attention will have a say in how software and services are designed. These companies prefer and are willing to pay more for ads that are displayed to targeted audiences in a relevant way, and the companies that provide the services are arguably just as responsible to the end users using their products as they are to the companies paying them to make them.

Government Regulation of Online Tracking

Irrelevant to one's opinion of online tracking and advertising, it has become a large industry: according to a Forbes article from 2013, digital ad spending in 2012 exceeded $100 billion. As online advertising grows as an industry, an important factor will be how governments decide to regulate or not regulate the use of online tracking.

In the United States, lawmakers have been actively working to create legislation to protect the privacy of Internet users. In fact, a subcommittee in the United States Senate has been explicitly created "and charged with '[o]versight of laws and policies governing the collection, protection, use, and dissemination of commercial information by the private sector, including online behavioral advertising" (Tene et. al 320). Additionally, the Federal Trade Commission (FTC) has proposed a three step plan called the "FTC Preliminary Report" to protect consumers by requiring companies to respect user privacy, give consumers options to control the sharing of their personal information, and be more transparent about how the companies collect data.

The first step in the proposed FTC Preliminary Report would require businesses to be more cautious with how they handle user data collected through online tracking. Included in this step is the requirement for businesses to secure the data they collect to cut down on accidental leaks of user data. Next, businesses are expected to only collect the data they need and only for as long as they need it in order to cut down on the amount of private information they are storing at any particular time. Lastly, this step calls on businesses to ensure that data is accurate. Although, interpreted in certain ways, these requirements could be a significant win for privacy advocates, their vagueness unfortunately makes this part of the proposal virtually toothless. Who decides what data is 'needed' for business? How could a company ensure that all their data is accurate, especially when even users make mistakes when entering their own information on websites?

The second step of the FTC's Preliminary Report proposes a technology called Do Not Track that will act as "a user's centralized opt-out of online behavioral tracking" (Tene et. al 321). Although some tracking companies currently offer mechanisms for users to opt out of tracking, many of these opt out services store the user's preference as a cookie, which is erased whenever users clear their cookies (a behavior some do to erase tracking cookies in the first place). Do Not Track was originally a proposal for "a list-based registry of users, similar to the Do Not Call Registry" ("Do Not Track") but has since changed shape in the industry to the form of an HTTP header that is sent along with HTTP requests. If enabled in a user's browser, Do Not Track sends a simple flag on all of their web traffic to websites that indicates that the user does not wish to be tracked. It is entirely up to the receiving website to honor the request, and unfortunately there is no policy yet to require them to abide by the

user's request.

The last step in the FTC's proposal requests that companies increase transparency about how they handle collected user data. This step calls "for privacy notices to be clearer, shorter, and more standardized" (Tene et. al 321) as well as requiring companies to provide users with the data that has been collected about them. In addition, the proposal would require companies to obtain consent from Internet users to use data collected about them "in a manner materially different from that presented at the time of collection" (Tene et. al 321) which is intended to cut down on the abuse of user's data. It is not immediately clear what constitutes being "materially different" from a user's perspective.

Unfortunately, this proposal is by no means a requirement for companies in the United States yet. Proposed technologies like Do Not Track have been accepted by most of the modern browsers but are not required to be honored in any way by companies yet. Many companies see these proposals as a threat to their online advertising revenue as it could turn into "a government-sponsored, and possibly managed, ad-blocking program--something inimical to the First Amendment" (Tene et. al 324). At present, in the United States, the FTC investigates reported violations while debates on how to formally regulate online tracking continue in Washington.

Europe, on the other hand, already has a legal framework in place that applies to online behavioral tracking. The European Data Protection Directive and the European e-Privacy Directive regulate "the collection, processing, storage and transfer of personal data" (Tene et. al 307). The European Union has taken almost the opposite approach from what the FTC has proposed. Instead of requiring that websites provide users with an opt-out choice, Europe has decided instead to require users to provide explicit informed consent before their data is collected. In fact, the courts have even ruled that a browser setting to enable tracking does not show "the existence of valid informed consent" (Tene et. al 312). Historically, European nations have been aggressive in holding United States firms accountable to their privacy laws. Google and Facebook are frequently fined for changes to their privacy policies and for mishandling user data. Just last year, data agencies from six European governments launched investigations into Google's new privacy policy that they began using in 2012 (Leach) and an international group called "europe-v-facebook.org" has been created just to enforce these privacy laws.

Unfortunately, government regulation has not caught up globally with the pace of innovation in the technology and tracking techniques of online advertisers. Hopefully governments and companies can come to a compromise in the coming years about a set of internationally accepted regulations that protects user privacy while allowing companies to present users with relevant advertisements.

Stopping Companies from Tracking Me

There are certainly many different government efforts that are being discussed to regulate web tracking, but for right now the industry is mostly unregulated. In the meantime, security and privacy researchers as well as browser developers have proposed a few standards and tools that allow users to opt out of some forms of web tracking.

As mentioned in the previous section, Do Not Track is a header that users can set in their browsers to be sent to websites with each request. Websites and advertisers are by no

means legally required to pay attention to or honor these headers. DoNotTrack.us, a website maintained by Stanford and Princeton researchers, maintains a list of other sites that honor the Do Not Track headers, and, as of March 2014, Twitter, Pinterest and a few advertising data providers are the only companies that are on this list. Unfortunately, it doesn't appear many members of the industry will voluntary honor Do Not Track, and although the Digital Advertising Alliance, a representative of over 100 large advertising firms, was part of the working group developing the standard for more than two years, it formally stepped out of the Do Not Track working group as of September 2013.

As part of self regulation efforts, the Digital Advertising Alliance has created www.aboutads.info to allow users to opt out of certain third party cookies from members in its network. Both Safari and Firefox block third party cookies by default, which means that only websites that users directly visit will be able to store cookies on their computers. Options like these among advertising providers could potentially have an impact if sites don't collude with each other and share other identifying characteristics about users; however, all it takes is one company like Google to store, share and later monetize tracking data using a mechanism separate from cookies to subvert such protections.

Users who use any services that they don't pay for, such as Google and Facebook, are ultimately subjecting themselves to a certain amount of tracking across any pages that load plugins from these sites. The predominance of Facebook 'Like' buttons and Google '+1' buttons across the Internet makes it easy for Facebook and Google to track users who are logged in to their accounts and there might not be much stopping them from using this information later in ads.

In "Cookieless Monster: Exploring the Ecosystem of Web-based Device Fingerprinting", as published in a 2013 IEEE conference, researchers demonstrated that even without cookies, it's possible to track users based on the specific configurations of their browsers and machines. The researchers showed that even using a fake user agent string to avoid being "fingerprinted" leaves the user vulnerable because ultimately the properties of specific versions of browsers are easy to distinguish. One company offers a product that circumvents this by sending all user traffic through a remote web browser that is shared among all users of the service, but as such services are used by very few people, companies can easily still target ads to 'paranoid people who use anonymizing browser X'.

Ultimately the only way to completely prevent companies from tracking users across the Internet is to not use the Internet or technology at all. This is not a real possibility for most people, especially as society is becoming more dependent on connected systems, and hopefully government regulation improves in the future to provide more protection, visibility, and control to end users. The best way to work towards this goal for now is to talk about the issue of online tracking and approach government representatives to let them know this issue is important.

Conclusion

The Internet as we know it runs on online advertising. Websites are able to provide Internet users with free services because they are paid for displaying advertisements to their users. These advertising companies have realized that if advertisements are more relevant to the users that are shown them, they can be sold for more money and therefore generate more

revenue for the websites hosting the ads. In order to determine which advertisements are more relevant for each Internet user, these companies engage in online tracking to find out more about the user and create profiles of them.

Online tracking can be performed through a variety of different means. First party tracking cookies are simple mechanisms that a website can use to track a user's actions while they navigate and interact with the website. Unfortunately, these cookies can also be placed by third party websites while the user is unaware, which allows the third party to track a user's behavior across many other websites that all participate in the same ad group. A proactive user can clear their cookies from time to time, which can stop this method of online tracking, but other forms of tracking cannot be so easily counteracted. Super-cookies can be designed to stealthily regenerate tracking cookies even when a user deletes the cookies in their browser and methods like online fingerprinting can be used to identify users without using any cookies or client side web activity at all. Even worse, deep packet inspection can be performed by Internet service providers to examine all Internet traffic a user generates.

Companies engage in online tracking because targeted advertisements have a much higher chance of actually being useful for Internet users. Studies have shown that users are less annoyed by watching or seeing advertisements that are more relevant to them. The increased relevance also enables these advertisements to generate more revenue for all parties involved: the people paying to advertise their goods and services, the people paying websites to display the advertisements, and of course the websites displaying the ads.

Most people consider online tracking to be a major invasion of privacy. Companies that are smart about online tracking and targeting can change how they display offers to show users higher prices, possibly causing the users to miss out on good deals. If Internet user profiles are de-anonymized, users could find all of their online activities on display to the world which could lead to all sorts of embarrassing situations. In extreme situations, oppressive governments have used online tracking to monitor citizens' online activities in order to control the media and stay in power.

On the other hand, online tracking can also benefit Internet users. Targeted advertisements increase the likelihood that a user will see an advertisement for something that they actually end up buying and enjoying. The process of online tracking can also lead to personalized offers that can save users money. Without advertisements, the argument could be made that the free Internet as we know it wouldn't exist because websites would likely need to charge for their services. Even Internet service providers have taken advantage of online tracking to offer lower prices to subscribers in exchange for targeted advertisements.

The European Union has been aggressive in addressing online data privacy issues and has drafted legislation that companies oftentimes have a hard time complying with. These rules require careful collection and treatment of user data in Europe. In the United States, the Federal Trade Commission has proposed a three step plan to help regulate online tracking, but the major parts of the proposal are not yet legally enforceable. One of the steps involves the widespread adoption of a technology called Do Not Track that users could enable in their browser to prevent themselves from being tracked online. Currently browsers allow users to enable this feature, but it's up to the websites to honor the user's request to not be tracked. Some firms have argued back and forth with the FTC regarding this issue because they believe that this regulation could result in a government-run ad-blocking program that might infringe on their first amendment rights.

Today, users in the United States can take advantage of many privacy features available to them, such as activating Do Not Track in their browsers and clearing their cookies regularly. Unfortunately, nothing can guarantee protection from online tracking today. Users

concerned about their rights to online privacy should reach out to their local legislator to help promote knew rules and regulations on these online tracking companies.

It's arguable that most users are happier seeing relevant advertisements, but the tracking mechanisms that advertisers currently employ to reach this end result are much more controversial. From a user privacy perspective, the technologies that such companies currently use are very dangerous and the industry has failed to regulate itself. On the other hand, the results of these technologies pay for the services that we use on the Internet today. This debate between user privacy and the inherent benefits of targeted advertising will continue for the foreseeable future, and the results will have a significant impact on the future shape of the Internet. Hopefully, we as a society can come to a consensus on a system for tracking users that provides some degree of anonymity while still allowing advertisers to display more relevant advertisements.

Works Cited

“Advertising.” *Wikipedia, The Free Encyclopedia.* Wikimedia Foundation, inc. 4 Mar. 2014. Web. 8 Mar. 2014

“Ad Council.” *Wikipedia, The Free Encyclopedia.* Wikimedia Foundation, inc. 26 Feb. 2014. Web. 8 Mar. 2014

Bachman, Katy. "DAA Study: Targeted Advertising Is the Financial Engine of the Internet." *AdWeek*. AdWeek, 10 Feb. 2014. Web. 08 Mar. 2014.

Bachman, Katy. "Targeted Internet Advertising Isn't Feared by Consumers." *AdWeek*. AdWeek, 18 Apr. 2013. Web. 9 Mar. 2014.

Beales, Howard. "The value of behavioral targeting." *Network Advertising Initiative* (2010).

Brodkin, Jon. "AT&T offers gigabit Internet discount in exchange for your Web history." *Arstechnica.* 11 Dec 2013. Web. 16 Feb 2014.

"Connections Counter: The Internet of Everything in Motion." *Cisco*. Cisco, 29 July 2013. Web. 17 Mar. 2014.

Diaz, Sam. "ISPs Keep Their Distance from Deep Packet Inspection." *ZDNet*. ZDNet, 25 Sept. 2008. Web. 17 Mar. 2014.

"Do Not Track." Electronic Frontier Foundation. 16 Feb. 2014

Gordon, Mary E. "Flurry Blog." *The History of App Pricing, And Why Most Apps Are Free*. Flurry Analytics, 18 July 2013. Web. 09 Mar. 2014.

Leach, Anna. "Google Versus the EU." *Tech Europe*. Wall Street Journal, 4 Apr. 2013. Web. 17 Mar. 2014.

Mayer, Jonathan R., and John C. Mitchell. "Third-party web tracking: Policy and technology." *Security and Privacy (SP), 2012 IEEE Symposium on*. IEEE, 2012.

Morris, Jason. "Why Big Companies Buy, Sell Your Data." *CNN*. Cable News Network, 23 Aug. 2012. Web. 17 Mar. 2014.

Mattioli, Dana. "Why the Apple Demographic Is So Important to Orbitz and Retailers."*Digits Tech Blog*. Wall Street Journal, 26 June 2012. Web. 16 Mar. 2014.

McGuire, Tim, James Manyika, and Michael Chui. "Why Big Data Is the New Competitive Advantage." *Ivey Business Journal*. N.p., 01 Aug. 2012. Web. 17 Mar. 2014.

Narayanan, Arvind, and Vitaly Shmatikov. "Robust de-anonymization of large sparse datasets." *Security and Privacy, 2008. SP 2008. IEEE Symposium on*. IEEE, 2008.

Nikiforakis, Nick, et al. "Cookieless monster: Exploring the ecosystem of web-based device fingerprinting." *Security and Privacy (SP), 2013 IEEE Symposium on*. IEEE, 2013.

Ohm, Paul. "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization." *By Paul Ohm*. UCLA Law Review, 13 Aug. 2009. Web. 08 Mar. 2014.

"Personalized Online Video Ads Boost Branding." *EMarketer*. EMarketer, 24 Oct. 2011. Web. 09 Mar. 2014.

Rainie, Lee, Sara Kiesler, Ruogu Kang, and Mary Madden. "Anonymity, Privacy, and Security Online." *Internet American Life Project*. Pew Research Center, 5 Sept. 2013. Web. 16 Mar. 2014.

Sableman, Mark, Heather Shoenberger, and Esther Thornson. *CONSUMER ATTITUDES TOWARD RELEVANT ONLINE BEHAVIORAL ADVERTISING: CRUCIAL EVIDENCE IN THE DATA PRIVACY DEBATES*. Issue brief. N.p.: n.p., n.d. Print.

Soltani, Ashkan, Shannon Canty, Quentin Mayo, Lauren Thomas and Chris Jay Hoofnagle. "Flash Cookies and Privacy." (2009): Web.

Sonne, Paul, and Margaret Coker. "Firms Aided Libyan Spies." *The Wall Street Journal*. Dow

Jones & Company, 30 Aug. 2011. Web. 08 Mar. 2014.

Tene, Omer and Jules Polonetsky. "To Track or 'Do Not Track'." *Minnesota Journal of Law, Science & Technology.* 13.1 (2011): Web.

“Web Banner.” *Wikipedia, The Free Encyclopedia.* Wikimedia Foundation, inc. 3 Mar. 2014. Web. 16 Mar. 2014

Yan, Jun, et al. "How much can behavioral targeting help online advertising?."*Proceedings of the 18th international conference on World wide web*. ACM, 2009.