Infotrust
eXchange

Discussion of a user/publisher optimized web advertising system

This file: Mozilla-User-publisher-optimized-ad-system-04-29-16 is found at: https://docs.google.com/document/d/1gmLktP9rDYEwYekn0RL7R8twLKiv0465p7StpA3Vpz4/edit

Author: Don Marti

Senior Advisor, Mozilla

From the Mozilla point of view we have two important principles that need to be reconciled.

“Individuals’ security and privacy on the Internet are fundamental and must not be treated as optional.”

When users share information about themselves they need to do it voluntarily and know what they are sharing and who they are sharing it to.

“Commercial involvement in the development of the Internet brings many benefits; a balance between commercial profit and public benefit is critical.”

From the user point of view:

- Users want free ad-supported sites
- Users express dislike for some common tracking practices (66% of adult Americans said they do not want marketers to “tailor advertisements to their interests”)

News and cultural works built at companies are an important part of the web's value, and we really don't want to see that type of site move off of the open web and into some kind of silo that is controlled by a single company.

Advertising right now is how a lot of news organizations hope to be able to pay the bills, as print advertising is going away. Advertising is certainly not there yet. The same content on the web brings in roughly 10% of the value of that content in print. Yet advertising is something that people are relying on being able to do on the web, and make it work somehow. But web advertising as it is practiced today is not consistent with the kind of user control over their own information that is important for a sustainable web. So information about users can effectively follow them from site to site resulting in unpredictable and unwanted privacy and security concerns.

Web advertising as it works today facilitates putting quality sites into direct competition with low-quality sites and with outright fraud. If you wanted to get an ad in front of people in Alameda, California, you could buy advertising in the Alameda Sun or you could go to some geo-targeted adtech intermediary that claims to reach the 94501 zip code and buy advertising there.

Today, much online ad money goes to intermediaries, some goes to low-quality sites, and a lot of the money vanishes into fraud. Unlike in print, a high-quality news site is in direct competition with low quality sites and fraud sites for ad money.

We have the opportunity to build a new advertising system that

1. works with Mozilla/EFF/privacy principles
2. doesn't break economic signaling, as targeted ads do
3. has immediate effects on current buzzword problems such as #adfraud
4. is small enough to implement in a reasonable amount of time.

Publisher strong points are domain expertise and reputation. A new system must use what publishers are good at, instead of trying to out-Facebook Facebook. (Don't hunt unicorns—shoot rats. Rats in grain elevators spoil much more grain than they actually eat. And companies that practice problematic forms of user tracking tend to reduce the total value of web advertising much more than they actually profit.)

Simple example of workflow

Advertiser: Super Duper Burgers has an internal customer database. Many of their existing burger-eating customers are JavaScript programmers. They want to reach more JavaScript programmers because Big Data says they're likely to be burger eaters.

User data source: JavaScript Developer Network has information on user interests and skills, and has issued attributes to JavaScript programmers.

Publisher: Web Developer News has many JavaScript-using readers, and wants to show advertisers that the site is a good way to reach JavaScript programmers.

The system needs to facilitate sharing the JavaScript attributes with trusted publishers.

The sharing mechanism does not need to know about JavaScript or burgers in advance. Just that some sites tell users: "here is an identifier or some other piece of data that applies to you" and other sites want to learn those identifiers if the user is willing to share them.

1. JavaScript Developer Network assigns an attribute to the user: "JavaScript developer."
2. User can choose to reject or accept the attribute. (Users may also visit a self-serve site, or install a browser extension that lets them pick arbitrary attributes not signed by an authoritative issuer.)
3. User visits Web Developer News, a trusted publisher. The publisher presents a user survey and the user agrees to respond.
4. The publisher explicitly asks the user to share the attribute as part of the survey, and the user agrees. The site can choose when to request an attribute, and the user can say yes, no, or never share attributes with this site. Sort of like the way that "share your location" works today.
   http://www.w3.org/TR/geolocation-API/#privacy_for_uas Sites will have to balance user data collection with the risk of driving users away. The decision on when to request attributes is up to each site. Some sites may offer an incentive to assign or request an attribute. Testing of alternate offers and workflows is important.

5. The publisher aggregates attributes from many users into an Audience Profile Book.
6. The advertiser parses multiple Audience Profile Books to assemble the desired audience, and buys advertising space on one or more publisher sites.

All of this can be implemented by Software as a Service providers working for the publisher and advertiser. All of these providers can be "Do Not Track" compliant.

Notes on system requirements

Can start with a suggested list of attributes, but niche publishers will want weird user attributes. Free-form user info must be possible.

Every user attribute can have its own home base (The Alameda Sun can give me an attribute that says I probably live nearby, and the Linux Foundation can issue me an attribute about what I do for a living.) May be storable encrypted, unreadable on server side, in a cloud service.

A single publisher might have multiple section audiences in its Audience Profile Book (travel section readers, live music news readers). The audience profile book will need to let publishers break out readers by section. ("arts and music" might be a section at one publication, but another might have want to have sections for "music" and "arts and literature") Publishers should be able to choose how to split it up. Advertiser doesn't get to reach a cohort of readers, though, just placement on that section's pages. A new reader will see the same ad on a given arts story that a frequent arts reader and survey taker does.

Compete where local newspapers' strong point faces creepy Internet companies' weak point

The big platforms have lots of user data. Not just a bigger head start, but more experts and more money, and the power to get more data faster. You can't win a Big Data race with the people who invented Big Data and can do it more efficiently, with higher scale, than anybody.

What do newspapers have that the giant Internet companies don't? Reputation.

But reputation for honestly reporting the news does not scale in the same way that Big Data does. The winnable long-term plan is to transform the web advertising medium. Right now we have a data-driven game based on following users around, delivering less and less valuable advertising, and making it up on volume. That's a game where large Internet companies engaged in adtech/adfraud have the advantage.

In a reputation-based game, creepy Internet billionaires are at a disadvantage, and newspapers can win.

Keep scope limited

Federated paywalls are a really good idea too. But the question is whether to build one incremental thing that works (keeping federated paywalls in mind so as not to break them in future) or to try to build one big project that includes both data sharing and federated paywalls—stuff that (1) directly competes with giant Internet companies and (2) is a prime attack point for griefers and fraud.

If we go incremental we can position it as an anti-fraud thing to start with. If you're buying ads on a site with suspiciously few entries in their Audience Profile Books, they're probably heavy on bots or light on users who trust them enough to take a survey. That way we're not eliminating adtech as we know it just to get a bigger piece for the publisher—adtech as we know it is just collateral damage of moving to a lower-fraud system. Current anti-adfraud measures are fatally limited by the perceived need to preserve existing adtech and track users across sites. If you move away from the idea of being able to target individuals, more and better anti-adfraud strategies are available.

Part of this complete breakfast

Any next-generation system needs to be developed in combination with advancing Tracking Protection technology, in applying Tracking Protection to more users, and with collecting data on the buying habits of Tracking Protection users. No advertiser or agency will bargain with users or publishers for fairly shared information if they can just take it. Any marketing spend has to be sold internally through a complex chain of agencies, intermediaries, marketing, and top management at an advertiser company.

Every decision has to be something that an individual marketer can sell to the next step in the chain. Current surveillance marketing is designed to be appealing to the current buying process. We have to cut off non-permissioned user data collection to get attention for data sharing.

High-value advertising must be an integrated part of a technical transition that also includes closing off the options for low-value advertising in the same medium. For example, publishers can warn users when they're vulnerable to third-party tracking, and encourage them to get protected. This will reduce the number of "bad" ads that those users see on any site.

Client side Server side

Make
low-value ads harder

Layer 1: block connections to untrustworthy trackers
Layer 2: don’t persist cookies and unsafe state
Layer 3: clean up problematic state (layers)

Safe ad blocker warnings: don’t block privacy tools as ad blockers. **Tracking warnings:** Inform users about alternatives so that they can configure in-browser tracking protection. **Reverse tracking walls:** offer bonus content to protected users.

Make
high-value ads easier

Attribute **assignment** with user control
Cross-site attribute **sharing**
with user control

Information Trust Exchange

attribute sharing Safe **web analytics**
Future: **federated paywalls**

Publishers can’t out-Facebook Facebook to offer creepier and more targeted ad placements. However, publishers can find common aspects of user-privacy-driven tracking protection improvements and publisher-business-driven tracking protection improvements and make them a priority. Realistically, we can assume that **advertisers and agencies will ignore the new system** until they see that it’s a way to reach a significant audience that they can’t reach in other ways.

The end of “right ad to the right user at the right time”

The current web advertising model depends on tracking the same user across multiple sites (either “anonymously” or using PII). Per-user targeting across the web is unsustainable and likely to become less available to advertisers in the future.

Targeted advertising fails to support news: News sites are unable to pay the bills with web ad revenue. If current trends continue, news ends when print does. Publishers are seeking higher-value models. Finding them is not optional.

Competition to reach users: Sites that produce content for a local audience or community of practice are now in direct competition with targeted ads, which can appear on low-value sites, to reach the same audience.

Fraud: Today’s adtech ecosystem makes fraud relatively easy and anti-fraud relatively difficult. Advertisers are effectively funding billions of dollars worth of artificial intelligence research, to create increasingly human-like adfraud bots. The presence of fraud in the system drives down the price of all web advertising, even ads on high-quality sites. Publishers and copyright holders, not adtech intermediaries, bear the costs of fraud.

User choices: Most users are willing to accept some advertising, but aversions to some targeting practices are widely held. Protection from strongly disliked practices, such as price discrimination and targeting by medical condition, is already a competitive advantage for the Apple Safari browser, and other browsers feel pressure to innovate.

Regulatory re-balancing of data risks: User identity theft and other security risks are negative externalities of data collection practices that enable, among other things, ad targeting. Regulators in some jurisdictions are likely to try to shift these costs to the firms that collect the data.

Ad blocking seems to be mainly a balance between the hassle of blocking (dealing with broken sites and anti-adblock warnings) and the hassle of not blocking (slow page load times, annoying ads). However, ad blocking did not go mainstream until retargeting showed users how ads were attempting to reach them, and were not just something on the site. The economic signal of non-targeted advertising is a way to shift the ad blocking balance back toward advertising.

A new advertising system will let the web lose the directly user-targeted ad while continuing to provide the information that advertisers need in order to measure ad performance and continue supporting high-quality sites.

Links

The Mozilla Manifesto
https://www.mozilla.org/en-US/about/manifesto/

List of protection layers
http://blog.aloodo.org/posts/protection-layers/

Tracking warnings for publishers
http://zgp.org/targeted-advertising-considered-harmful/#solution-tracking-protection-for-publishers

Ad blocking: why now?
https://digitalcontentnext.org/blog/2015/07/06/ad-blocking-why-now/