Fighting Fraud without Fingerprinting

Fighting Fraud without Fingerprinting slides
Presenter: Philipp Pfeiffenberger <philippp@google.com>
Scribes: Nick Gaw,

Queue - add yourself at the bottom (include topic, if you like our scribes write out your question/comment)

- QUEUE CLOSED

Attendees (sign yourself in):

1. Panos Astithas (Google Chrome)
2. Michael Ficarra (F5)
3. Brian May (dstillery)
4. Kevin Gibbons (F5)
5. Cornelius Witt (eyeo)
6. Steven Valdez (Google)
7. Ryan Kalla (Google Chrome)
8. Charlie Harrison (Google Chrome)
9. Joey Knightbrook (Snap)
10. Aykut Bulut (Google Chrome)
11. Jean Luc DI Manno (Fime)
12. Martin Thomson (moz://a)
13. Sameer Tare (Mastercard)
14. Johann Hofmann (Google Chrome)
15. Ben Kelly (Google Chrome)
16. Ben Wiser (Google Android)
17. Erik Anderson (Microsoft Edge)
18. Tara Whalen (Cloudflare)
19. Mihai Cirlanaru (Google Android)
20. Brianna Goldstein (Google Chrome)
21. Aram Zucker-Scharff (The Washington Post)
22. Matthew Finkel (Apple)
23. Per Bjorke (Google, Ad Fraud team)
24. Sam Macbeth (DuckDuckGo)
25. Edwin Daniel(Google Chrome)
26. Benjamin VanderSloot (Mozilla)
27. Tim Huang (Mozilla)
28. Filipa Senra (Google Chrome)

29. Satoru Takagi (KDDI)
30. Zainab Rizvi (Google Chrome)
31. Borbala Benko (Google Counter-Abuse)
32. Tim Huber (Google Chrome)
33. Shubhie Panicker (Google Chrome)
34. Vinod Panicker (Amazon Ads)
35. Chris Fredrickson (Google Chrome)
36. Nick Doty (CDT)
37. Mark Nottingham (Cloudflare)
38. Christian Dullweber (Google Chrome)
39. Shunya Shishido (Google Chrome)
40. Antonio Sartori (Google Chrome)
41. Nicolas Pena Moreno (Google Chrome)
42. Gerhard Oosthuizen (Entersekt)
43. Tony England (Visa)
44. Chaals nevile (Consensys)
45. Ane Fabo Aranzabal (Google Chrome)
46. Scott Hendrickson (Google)
47. Kaustubha Govind (Google Chrome)
48. Alex Christensen (Apple)
49. Jonathan Foote (Fastly)
50. Sergey Kataev (Google)
51. Nicola Tommasi (Google Chrome)
52. Christos Bacharakis (eyeo)
53. Alex Rudenko (Google Chrome)
54. Ben Savage (Meta)
55. Erik Taubeneck (Meta)
56. Benjamin Case (Meta)
57. Alex Koshelev (Meta)
58. John Wilander (Apple WebKit)
59. Sami Tikkala (Visa)
60. Wojciech Filipek (eyeo)
61. Yoav Weiss (Google Chrome)
62. Kapil (reCAPTCHA)

Minutes

Philipp: Want to clarify the problem we’re exploring. Fraud - use cases defined by the W3C anti-fraud CG
… crosses privacy, accessibility and security
… Need to minimize ???

Tension arises bc users need to establish trust and the methods for establishing trust are often behind the scenes now.

Sites can require phone num, additional info or inspect client state.
Hard to distinguish between signal collection for fingerprinting from signal collection for fraud detection
Websites can also be malicious. Shouldn't give those new capabilities.
Want to discourage cross-site tracking without making abuse worse. Need to preserve open-ness.

Let's talk about WEI

Google considered exposing platform details, but that created openness concerns around blocking of users
We had proposed a holdback to try to address open-ness concerns, but was clear from the feedback that didn't go far enough.
Neither the browser nor the user is incentivized to maintain the holdback. Users could override the holdback.
Valid concerns that we want to address. Need to consider ecosystem openness front and center.

Have to consider ecosystem open-ness. Introduce several takeaways:

1) Critical that we don't ship things that require specific browsers or configs. No DRM for the web
2) Wanted to have a coupled approach for CHrome and Android. No longer proposing that, stopping exploration of WEI on the web.
3) Any rate limiting or throttling depends on a scarce resource, but don't want this to become a barrier for access. E.g PAT are a scarce resource. But need this to not become a barrier for access
4) In a Privacy Pass scheme, the website talks to an issuer, issuer talks to attesters.
5) How can we enable new attesters to join. How can the issuer make sure that attesters remain reliable over time?
6) More comprehensive forum to discuss open-ness. We need people who care about openness, privacy and anti-fraud in the same conversation

Don't think the status quo here is great, want to explore joint proposals.

Want to consider a reduced use-case - request is associated with a single user within a per-site limit.

Make progress on problems, not solve them all.

How could this be done? 1P cookie or partitioned storage?

But there is zero cost for attackers to pretend to be a new user.

Better solutions - PrivacyPass & rate-limiting extension

However, we need to increase clarity on how new attesters and issuers can join the ecosystem.

Have to identify what value remains after we narrow the use case - might result in high latency of attestation, for example.

How scarce does something need to be in order to be used for rate limiting? Rate limiting against a 1000$ device is more expensive than against a $1 device.
Identity is also a source of scarcity - e.g. library vs. utility bill

Where is the line of utility for anti-fraud that also allows for latency in this kind of system?

In a world with multiple attesters, the issuer has to be able to evaluate the quality of attestations over time.

Will websites actually provide feedback on utility/missed abuse?

Would love a way for us to have conversations with people who care about anti-abuse, open-ness and anti-fraud together.

Questions:

- Gerhard: Could you rephrase the question? Are you asking for a new group? More ppl to join the AFCG?
    - PP: Difficult to fold in all aspects of this discussion, especially openness. Question is there interest from folks from all these viewpoints
- Mark Nottingham: you want to design something that by its properties gives you what you want. Properties may need to be applied at runtime, using some governance mechanism. Has that been on your mind?
    - PP: I’m wondering how that agreement would be verifiable from the outside?
    - Mnot: similar to how the CAB forum operates, added technical means to enforce, but at the end of the day browsers need to trust CAs
    - PP: That kind of assessment and evaluation would be to allow certain collection processes in certain contexts?
    - Mnot: More of a concern on whether you’re arbitrarily allowing some people as attesters vs not. That could have an effect on openness of the web. Could we design a way that doesn’t discriminate?
    - PP: Having an onboarding practice is part of that design. How does it work over time though? Need qualitative feedback over time.
    - Privacy Pass like scheme with issuers and attesters, with governance and rules for what it means to be either

- Nick Doty: I don't think it's the case that the concerns about openness hadn't been heard or raised before – I feel confident that those concerns were being raised. We had discussion. However, I do think it's going to be the case that it's going to be hard to get ppl interested in anti-fraud if they aren't seeing industrial use. But do think that if we were to narrow down to a specific issue, I could help bridge in others.
    - There's interest in governance about new issuers/attesters, I'm more interested in what is attested. Need some accountability mechanism for that. Just agreement to start with would be good, but would probably need some out of band enforcement mechanism.
    - PP: Thanks for all the input - getting more ppl involved would be helpful. Agreed that it'd be helpful to find a more atomic property that we want to attest - would help focus conversation.
- Ben VanderSloot: Part of the challenge is that it's all so abstract. I understand that the scarce resource might vary from attester to attester, but one ecosystem concern is exactly that resource and how scarce it is. Might be more scarce on a platform that has extensions installed. DRM on the web idea starts there. People in different countries could have different scarce resources. Set of eyeballs from one country could have different value to set of eyeballs from another country. Don't want to see that.
    - PP: And there are a number of different use cases that have come up in conversation with different companies and anti-fraud providers. Both things like geo attestation and rate limiting – seems like there is a low bar for brute force rate limiting. There are risks in all of this – need to make sure that we're not creating a social credit score, shared across identity providers. That's definitely not an outcome we'd support.
    - Ben: Which is where a larger ecosystem of attesters helps solve some of these issues, but there are separate concerns to that.
- Aram: Who is an auditor, what do they audit, what qualifies them? There's 0 cost to be a new user, but what attackers desire is to be an existing user. Where do you depart from PrivacyPass original features? Why does it need auditing?
    - PP: I don't think we'd make definitive statements about auditors being required. For PP, we only know who the issuer is - and if a new platform shows up and says that they want to join the ecosystem, how do they do that?
    - Aram: Why, if they have a problem with the issuer, wouldn't the person working with them withdraw?
    - PP: That's the risk - causing the attester to be cautious about their issuers. Have an existing issuer vs. a new startup, how can they prove that they're just as reliable? I think this would lead to new attesters having a hard time getting to adoption.
    - Aram: Wouldn't new issuers arise to address this market gap?
    - <discussion about indie attesters/issuers>

- Charlie: general question. When talking about issuers and attesters, what are we talking about? Are they sites? Governments? They could be anything?

    - PP: Sorry for not adding more context. Reference to the Privacy Pass protocol. Tries to separate into 2 entities: the issuer, which talks to the website; and the attester, which knows about the device
- Ben Savage: I live in Singapore. Have a gov id system. Can only register once with a real gov id, address, etc. Have an app on the phone that you can use to give to apps, websites etc. Very expensive limited resource, difficult to spin up. Have you considered gov issued identity resources?
    - PP: In the mix of how we can compose different types of attesters, we could consider things like government ids. But the centralization of control is something we'd need to consider. And what other things do you have to fall back on? What else can you fall back on in order to provide attestations?
    - Ben: I guess it depends on the use case you're trying to use it for. What's the domain you had in mind? What type of fraud are we focused on?
    - PP: High volume of use, either bulk account creation. For user consent, if users consent to wanting to use that opens another possibility. Account creation, ad fraud, maybe. Other things. Optimization problem between openness, privacy, security is most relevant here.
    - In government ID case, you may have something that is tied to something with your citizenship (your retirement account/pension system). May be a place for further exploration
- Jonathan: There's some analog in the US with a few states starting to provide virtual IDs. I like Ben's idea about making a targeted use case around bootstrapping trust using that
- Brian May: This problem is more easily solved if we divide the domains into high-value where we need to identify the user (bank account) and others, such as browsing the web with ads shown, where it's less reasonable to identify. Shouldn't come up with something that could be applied to different domains
    - PP: Ads use case is super tricky and we could separate that. Advertisers care about activity across different publishers. State that has to be persisted across sites, not great.
    - Brian: If 1P can assert to the browser I know the user over time would be nice
    - PP: Is very limited to cases of being logged in
    - Brian: Was thinking of 1P cookies
- Matt: Seems like we're rehashing a bunch of topics that have come up over the years- but think it'd be helpful to refocus on the specific question about if this specific use case could help. The new attester problem is something that has come up a bunch of times, and is probably a topic for a different forum.
- Per: I'd like to learn more about the open-ness concerns and takeaways. What are the principles required here?
    - PP: The openness concerns that were surfaced were at a lower level than the browser. The browser is a pass-through. Concerns are around platform choice, what happens when a new platform wants to join that, reduce friction for their users? Rooted, unlocked devices, etc. also a concern. Prevents the user from modifying their devices, reduces choice was some feedback.

- John Wilander: (on the threat model. Are we considering a malicious user but otherwise regular technology? Or are we considering a compromised/rigged browser? Or even a compromised/rigged operating system? If we trust the tech, and this is truly just about rate limiting, then a prescriptive request from the site could work. I have not thought through all the implications of that, but it is a tool in making things more private – instead of pulling a bunch of information from the device and making a decision server-side, have the server prescribe its desired behavior and the client to make it so without data leakage.)
    - PP: On the threat model, the idea is to figure out something scarce that can be counted against, regardless of user intention, ideally not limiting the user's ability to choose different platform
    - John: You're restricting yourself here. The server could tell the browser "only allow the user to send the request 10 times per day"
    - PP: The browser is acting on behalf of the user, and the user can modify their browser
    - John: So, considering a modified/compromised browser in your threat model?
    - PP: Yes, this should work if you unlock bootloader, hack your motherboard, whatever.
    - John: Then of course the attacker can change the fingerprint over and over in the browser and fool you.
    - PP: cat and mouse game, yeah. And you can look for unexpected FPs. That's the way it works today. Moving that to the browser requires attestation
    - John: Aren't we comparing this with FP?
    - PP: For the sake of rate-limiting, yes.
    - John: But if you're assuming a compromised browser that doesn't work.
    - Per: You're right that we assume the browser is compromised in many of these cases. The challenge is that if you have a rich set of attributes, it gets hard for bad actors to figure out all the attributes and the organic distributions of them. Bad actors haven't been able to reverse engineer the expected distribution.
    - John: Got it - sounds like a meaningful difference.
    - Per: Yes, we assume that both the browser and OS could be compromised.
- Gerhard Oosthuizen: There's a lot of interest in this. The use case and type of attack vector is irrelevant in my scenario. We do gov ids and selfies etc. Might be useful to write down why people use fingerprinting and craft a path towards how do we replace that fingerprinting?
    - PP: I think that's a good tack to take - we had a breakout this morning about fingerprinting, the valuable use cases, and ones that are invasive. We need to have further community conversations about thi.
- Kapil : Would like to share that there are other actors in the flow: Site owners and simple end users. Site owners are looking for affordable ways to let all "good" users in. Least friction possible. Find other methods (SMS, MFA) to be too expensive to implement. User is sitting on a big stack, multiple layers, many places where we could implement solutions to get the best experience. IDs are scarce but would require a lot of plumbing

for simple sites. Something simpler might be a good enough proxy, and then next layers can increase requirements.

    - PP: Thank you - that's a good question to consider.
- Brian: should we consider proof of work? - For the rate limiting problem, could we do something like proof of work to prove that it's a browser and not a browser farm? I think there also needs to be some communication to users – for example, if we implemented proof of work, and a user had a low-powered machine, we'd need to be clear to the user about what is happening.
    - PP: POW is challenging given mobile devices, if you're a cybercrime group less of a problem. User being aware of whats happening: How is the user aware? FP is invisible to the user, have been using this for a while and it's not great but it works. The balance of having an attester knowing something about you, having that user value needs to be made more explainable.
- Vinod: We should be thoughtful about how much burden of proof / work we put on legitimate parties, because there can be backlash from users for using their resources for solving problems that are not apparent to them. They don't directly benefit from using their resources for this, communication problem.
- Jacynda: how would that impact assistive technology?
    - PP: It's critical, but it's a shame that AT is also the first way in which clients can be automated. Noted
- Ben VanderSloot, Mozilla - PoW leads to boiling oceans, Government ID is scary because of scope creep, could become more generic fraud mitigation and could e.g. force users to be US/EU citizen
    - PP: In your mind, are there any identity providers that could augment that (for example, a reddit account with a bunch of karma)
    - Ben: I don't know anything off the top of my head
- Brian May: This topic is in desperate need of attention, user data is getting more scarce, which is good, but we need to protect the integrity of the internet.
- Sami T (Visa): For your point about payments. We have the web payments WG. We don't like FP, but depending on who you ask. In payments, the IDs don't work everywhere, but there are local regulation requirements around it. But it exists in some countries. In the rest of the world, it's about "is this that person". That's where fingerprinting becomes really interesting. Let's say that a purchase starts and then a government ID is issues - many steps along the way - is this a real person, is this the same person ,etc. Different kinds of attack vectors in that process. Interested in limiting those.
    - PP: Thanks! Agreeing with Gerhard that we need to split out these use cases.
- MT: I'm not convinced that we can get people to stop using FPs. Looking for supplementary ways of more information would just provide better fingerprints. So don't like the framing
    - PP: Don't want to take an absolutist stance, just hope that there will be less FP over time. Browsers could use blocklists. Push people to better ways of doing this.

    - MT: Starting to reach the conclusion that we’re not going to stop people from FPing, can increase the anonymity set size(?) for these folks. Might need a different framing for this.
- Kevin Gibbons: a good motivating use case is "reddit spam": you want to let humans sign up for new accounts, but humans don't want to use their government ids.

IRC Transcript

<npdoty> I hope it's clear that the openness/freedom concerns weren't out of the blue or unheard of before the WEI feedback
<AramZS> I am DEEPLY uncomfortable with the idea that countries might have different attestation resources yes.
<mnot> What if countries _did_ attestation (a la fedid) but the specific country issuing the attestation was hidden from the site?
<npdoty> some countries may have different expectations/conditions on freedom or anonymity
<mnot> Yes. But there might be other avenues.
<AramZS> I guess I don't understand why I wouldn't just go to an indie issuer who would then become a more significant player to deal with the overflow of people who are refused by the larger issuer. Why wouldn't an attester trust new issuers who start up if they have a reason to trust them?
<mnot> I suspect vetting of attesters needs a governance layer — it needs to be consistent between browsers, and guided by principles.
<mnot> (if it's tied to real-world things like government ids)
<npdoty> not everyone agrees that it's a good idea to have centralized government control of identity
<mnot> Nick: I'm thinking more of a one-bit assertion that this is a unique person. Might expand to do age attestation down the road.
<AramZS> But it doesn't *have* to be in order to accomplish these particular outcomes right? If I'm a publisher in Nowhere Indiana and I need to secure the impressions of Nowhere News dot com, why wouldn't I be just as likely to trust ... say the New York Times as I would the Drivers License App of Indiana ?
<npdoty> not everyone has access to a government id, or wants to, or feels safe disclosing their government id in the context of their online activity
<mnot> npdoty: not talking about disclosing.
<AramZS> In fact, in a healthy ecosystem presumably attesters would be more useful if they are interest-domain specific rather than globally or state general.
<AramZS> mnot: but wouldn't they have to disclose their gov't ID to *someone* at some point first, even if not to the site?
<michaelficarra> there's already the Verifiable Credentials standard (https://w3c.github.io/vc-data-model/) for cases when you want to fully identify yourself, but we care about fraud cases where the end user wants to remain at least pseudonymous
<mnot> depends on how the browser does it; see eg fedid
<michaelficarra> mnot: fedid consults a central authority, which is also undesirable for a number of reasons

<AramZS> Right but to npdoty's point, even with any of these other systems, *some* system has to be able to ID the user's ID as real?
<AramZS> Right, at some point the ID has to be *verified* right? And if it is gov't ID then it has to be verified by the gov't?
<mnot> There will be some form of centralization here; the question is how it is managed.
<michaelficarra> mnot: there need not be some for of centralization in the redemption, just the issuance
<mnot> yes
<michaelficarra> *form
<AramZS> Right, I guess I'm unclear on why the centralization needs to be *more centralized* than what inevitably will occur as a small number of attesters who operate at scale make a biz out of attestation? If we have to verify the attester beyond the signal of entities trusting them
<AramZS> It may just be getting late but I feel there's some assumed threat here that I'm not seeing.
<AramZS> oh lord let's not define fingerprinting again. Don't we have a document for this?
<AramZS> I thought we wrote a bunch of use cases down somewhere?
<kaustubhag> Aram: "If we have to verify the attester beyond the signal of entities trusting them" => are you referring to the question on the slide "Would websites provide feedback on utility / missed abuse?" or something else?
<AramZS> Yes, I would assume that if websites act on the signal that is a usable signal kaustubhag
<AramZS> If we're talking about a browser api, the information about that should easily be legible. Maybe you have to rank domains as more trustworthy or something like that, but if we are trying to determine if the entity is accomplishing an outcome on behalf of a site successfully, presumably the site knows?
<kaustubhag> My understanding of the privacypass formulation is that multiple "attesters" are hidden behind a single "issuer". The website knows who the "issuer" is, but not who the "attester" is. The question is why would any issuer risk their own reputation by onboarding a fledgling (potentially unreliable) attester
<kaustubhag> Apple's Private Access Tokens deployment currently own has one attester (Apple/iCloud) behind a CDN issuer; but privacypass folks have mentioned on some forums that the vision is for the issuer to issue a token based on some combination (string of OR conditions is the simple formulation) of the attestations available on the client
<AramZS> Presumably it is valuable to have more than one attester? If it is then it is valuable to have a new attester? If an attester has value and the issuer does not act on it, why would it not partner with a different issuer. If either has value I don't see either side getting blocked or captured.
<kaustubhag> only has*
<kaustubhag> PrivacyPass currently requires the website to choose the issuer (https://developer.apple.com/news/?id=huqjyh7k).The openness concern is that most websites will likely only choose the larger/established issuer(s)
<AramZS> kaustubhag - is there a cost to compete on?
<kaustubhag> Aram: Great question. I'm not aware of any discussions regarding what the privacypass issuer/attester ecosystem will look like. PATs is the first shipping deployment; and

the only issuers supported so far are CDNs (I'm assuming the website would choose their own CDN, but not sure)