PRINCETON UNIVERSITY

CENTER FOR INFORMATION TECHNOLOGY POLICY
PRINCETON UNIVERSITY

Enhancing the Security of Data Breach Notifications and Settlement Notices

Ryan B. Amos[1], Mihir Kshirsagar[1], Edward W. Felten[1], Arvind Narayanan[1]

Executive Summary

Recent data breach settlement notifications use insecure methods to communicate with the public that expose consumers to significant risk. In particular, the email notifications are frequently written in ways that make it difficult for users to distinguish real settlement notifications from scams. This lowers the bar for scammers to create fake, phishing emails, potentially victimizing users twice. In this document, we provide recommendations for stakeholders to mitigate such risks.

At a high level, we recommend the creation of a centralized database of settlements and breaches, so that users have a way to verify the notices distributed. We further recommend that no settlement or breach notice include a URL to a new domain; instead, notices should include a URL to a page on a trusted domain, such as a government-run domain or the breached party's domain. If desired, that page can then redirect users to a dedicated domain. This allows users to safely ignore links to unrecognized domains, helping avoid phishing. Such a process also has value outside the data breach context as courts and government agencies increasingly turn to electronic notices to inform the public.

Current Notification Practices Create Significant Risk of Phishing Attacks

Phishing is a well-established social engineering attack over electronic communication that is responsible for tens of millions to billions of dollars in damages each year.[2] In 2019, the FTC issued a warning about phishing sites impersonating the Equifax settlement website to scam users[3], and later that year KnowBe4, a security awareness training company, issued a warning that Yahoo settlement phishing scams are likely.[4] One potential reason phishing is prevalent after the Equifax breach is that the company used a new domain to inform the settlement class: equifaxbreachsettlement.com.

As we mentioned previously, users had no simple mechanism to distinguish the real domain from fake domains, so the situation was ripe for phishing exploits. The

---

[1] We are researchers associated with the Center for Information Technology Policy (CITP) at Princeton University. CITP has launched a new Technology Policy Clinic that provides nonpartisan research, analysis, and commentary to policy makers, industry participants, journalists, and the public. This paper is a product of that Clinic and represents the independent views of the authors.

[2] Hong, Jason. "The State of Phishing Attacks." *Communications of the ACM* 55.1: 74-81.

[3] https://www.consumer.ftc.gov/blog/2019/07/equifax-data-breach-beware-fake-settlement-websites, https://beta.washingtonpost.com/business/get-there/the-equifax-settlement-has-already-spawned-deceptive-websites/2019/08/01/bba971d2-b493-11e9-8949-5f36ff92706e_story.html

[4] https://blog.knowbe4.com/scam-of-the-week-yahoo-massive-data-breach-settlement-phishing-attacks

FTC's own consumer education advice on phishing would suggest that these notifications should be treated as suspicious emails.[5] But the FTC's suggestion that consumers verify the authenticity of the URL is difficult to implement because these are new domains that are specifically created for each settlement. Indeed, even if an enterprising consumer went online to verify the authenticity of the URL they may still be vulnerable because any news sources face the same challenge of verifying the URL's authenticity. For example, Equifax's twitter pointed consumers to an incorrect URL.[6]

Another recent data breach settlement concerning Yahoo has similarly listed links to their settlement distribution information on new domains created specifically for that settlement. Since these domains are newly established, users have no way to know if these domains are truly connected to the settlement in question or if they are malicious domains created to steal personal information, passwords, or money. These same concerns apply equally to settlement notices as well as the initial breach notices.[7]

We note that these issues are not limited to electronic notices. We found the Premera settlement notice to also use the same domain pattern we observed in the Yahoo and Equifax settlement notices. The notice was distributed by physical mail, and included no reliable mechanism to verify its veracity.[8] Phishing attacks and other scams can be equally effective when conducted over mail, and physical mail can lend undue credibility to potentially malicious notices, as noted in previous FTC actions.[9,10]

How Viable is an Attack?

It is in fact quite simple to perform a phishing attack. Consider the recent Yahoo Data Breach. At the time of writing, we found the following domains to be available within a few minutes of searching, keeping in mind the real domain is yahoodatabreachsettlement.com:

- yahoodatabreachsettlement.net
- yahoooobreachsettlement.com
- yahoodatabreachsettlemant.com
- yahoodatabreachpayment.com

A would-be scammer could easily purchase any of these domains for around US$20.[11]

To test the viability of sending a phishing email, we repurposed a legitimate informational email from Yahoo and changed the links to point to a fake domain (from

[5] https://www.consumer.ftc.gov/articles/how-recognize-and-avoid-phishing-scams
[6] https://www.theverge.com/2017/9/20/16339612/equifax-tweet-wrong-website-phishing-identity-monitoring
[7] It's not unreasonable consumers might expect to give away personal information on a breach notice. As an example, Equifax asked consumers for their social security number to check if they had been breached. An archived version of this page is available here: https://web.archive.org/web/20171228005926/https://www.equifaxsecurity2017.com/
[8] https://komonews.com/news/consumer/that-post-card-in-the-mail-about-money-from-a-premera-blue-cross- breach-settlement-is-real
[9] https://www.consumer.ftc.gov/blog/2016/09/prize-scams-your-mailbox
[10] https://www.consumer.ftc.gov/blog/2015/04/sweepstakes-scam-uses-ftc-commissioners-name-vain
[11] As an illustration, we bought one of these domains for less than US$15.

"YahooDataBreachSettlement.com" to "TotallyNotYahooDataBreachSettlement.com"). We then used a GMail account to send a doctored message to a Yahoo mail account. Yahoo Mail failed to recognize the email as spam. The email looks identical, aside from the tweaked URLs, and even an experienced user may not be able to distinguish the two. The "Sender" field is not a reliable field, and can often be spoofed.[12]

After convincing a user to click on such a link, an attacker might ask the victim for their name, their social security number (for "tax reasons"), their Yahoo account details possibly including the password (to "prove ownership"), and more. This presents a grave risk to users' identities, potentially causing them to be twice compromised. This phenomenon is not limited to Yahoo. Other companies' data breach responses have followed similar practices.

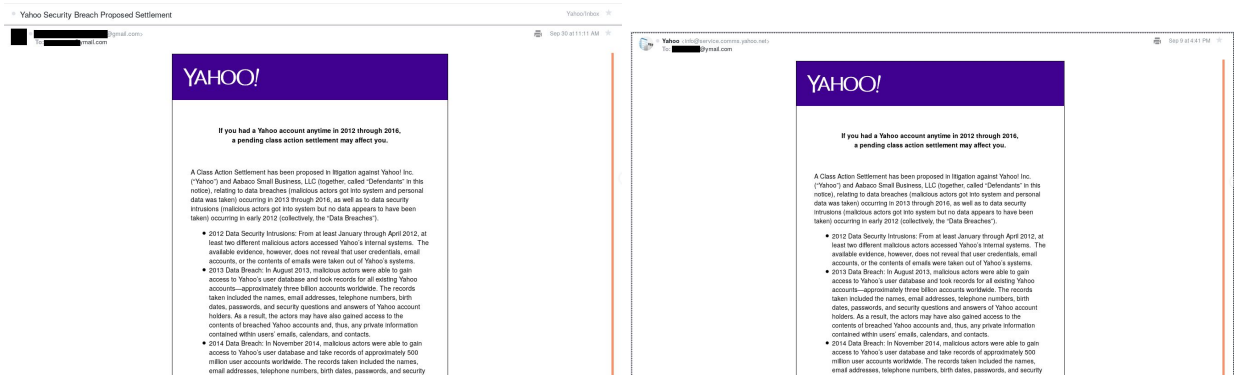

Figure 1: One of these is real, one is fake. Which is which?

Suggestions for Mitigating the Risk of Phishing Attacks

We recommend the creation of a centralized master list of settlements that consumers can use for reference and industry and courts can point to. The establishment of a centralized information resource to distribute consumer information has a precedent in the Consumer Product Safety Commission (CPSC) recall list.[13] Even for settlements made outside of the court system, information could be hosted on the centralized list. Indeed, the CPSC hosts both voluntary and court-ordered recalls.

Then there is the challenge of disseminating information about breaches. We propose relying on a well-established party to host a redirect[14] or landing page, and the settlement notifications link to the redirect, not directly to their unique website.

---

[12] Pandove, Kunal, Amandeep Jindal, and Rajinder Kumar. "Email spoofing." *International Journal of Computer Applications* 5.1 (2010): 27-30.

[13] https://www.cpsc.gov/Recalls/

[14] A redirection could be a response that serves an HTTP 301 response (See https://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html#sec10.3.2). In less technical terms, when the browser requests the redirecting page, the server sends a special response that tells the browser to automatically switch to a different URL. The browser then makes a new request to the new URL without requiring any user action.

The most obvious and preferred choice is for a government agency to host the redirects. For example, the Administrative Office of the United States Courts could host a redirect or landing page, e.g. settlements.uscourts.gov/yahoo2019, which leads users to the settlement redress page, e.g. yahoodatabreachsettlement.com. Courts could then direct litigants to use this method for disseminating notice.[15]

Since the previous option may require extra infrastructure, as an interim measure, companies can host redirects as a sub-domain on their primary domain. For example, Yahoo could host a redirect or landing page, e.g. settlement2019.yahoo.com, which leads users to the settlement redress page, e.g. yahoodatabreachsettlement.com.

Finally, if neither of the previous options are desirable, a third party (possibly non-profit) could host a redirect or landing page, e.g. yahoo2019.settlement-notices.org, which leads users to the settlement redress page, e.g. yahoodatabreachsettlement.com.

There are some challenges is deciding who should host the redirect. Does the government host the redirect or landing page, even in cases where the court or government agency is not directly involved in the settlement? Does the responsible party host it, even though that party may be tempted to direct traffic away from the settlement redress page? Can a third party be sufficiently trusted and have the resources to maintain the function over time? How will end users learn about this facility?

But broadly speaking, our core recommendation is that any notifications for a data breach settlement include links *only* to well-established domains. The page at the link could be a redirection, or it could lead to a landing page from which a user can find the settlement redress website. Both options are reasonable. The link provided to consumers in the notice should not point directly to a new settlement redress domain. Potential choices for well-established domains include a ".gov" domain, the responsible party's domain, or a widely trusted third party. Regardless of the chosen host, care must be taken to ensure that redirections to a domain are retired before that domain's registration expires. This prevents abandoned settlements domains from being purchased by would-be scammers and abusing the redirect to give authority to their scams.[16]

By using such an approach, stakeholders can encourage users not to put sensitive information into links from previously unseen domains, mitigating the chances of a phishing attack. Users will know not to trust links that don't come from the trusted pages.

[15] The Class Action Fairness Act requires approval of the federal courts for most large settlements.
[16] Another advantage of a redirect site is that this retirement process can be automated, so that domain registrations no longer open the door to phishing, as they currently would do.