EKATA

Analysis: Simulating Obfuscated Emails with Ekata Data

Apple's obfuscated email capability is not new. It has existed since iOS 13 but in a different form. Before iOS 15, Apple offered users obfuscated email at signup when using Apple single sign-on (SSO). It required SSO integration support and all created email addresses came from "@privaterelay.appleid.com." This allowed businesses to opt-out of using any email with that domain name during account creation.

However, with iOS 15 and iCloud+, obfuscated emails no longer require SSO integration, and they no longer distinguish between normal and obfuscated emails. To make the challenge even more difficult, a user can generate up to 100 emails per iCloud account.

Simulating Obfuscated Emails

We wanted to provide evidence to understand how Ekata Identity Risk Score and Identity Network Score would be affected by an obfuscated email. We used a planned historical data test with labels and ran it through our point in time process in two ways:

- First, we sent the data through as normal.
- Then we sent the same data, but replaced all the emails with a unique (but random) email identity, such as adhcai7efaheafhe@icloud.com.

What might change when an email is obfuscated?

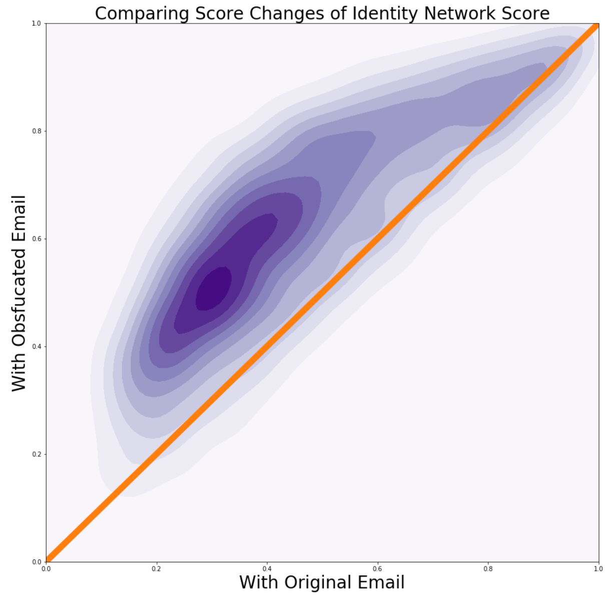
In the obfuscated set, I can expect that we have no information on the email in the Ekata Identity Graph and we have never seen the email before in the Ekata Identity Network. Thus:

- Email to name will always result in "no name found".
- Email first seen days will always equal 0.
- All other network signals, such as phone email first seen days, will also be 0.

All of the above are interpreted by our models for Identity Risk Score and Identity Network Score as higher risk, and thus will return higher scores for the same information than in the non-obfuscated set. Thus, we decided to focus on our two scores, which are consistently most predictive.

What did the data show?

As expected, we found that scores generally increased arbitrarily for both Identity Risk Score and Identity Network Score. The graph to the right shows this shift for Network Score. The x-axis shows the original score. The y-axis shows what the score would be with the same identity information but using an obfuscated email address. Anything above the line had a higher score and anything below the line had a lower score.



Even though low scores moved higher, high scores also increased. This was expected, as signals like "email first seen” and “email to name” indicated risky responses. However, what was interesting is that the predictiveness (measured by AUC) changed by less than 1%. In a rules scenario, we achieved similar results using a score threshold of 300 with non-obfuscated emails and a score threshold of 415 when the score threshold with obfuscated emails.

Conclusion

The predictability of Ekata’s scores is minimally affected by email obfuscation. While score thresholds may change, equivalent results are possible with new thresholds. Current customers concerned about the new iCloud+ technology should consider modifying their rule sets, but be confident that Ekata’s 5-1 identity solution will still deliver value.

For those currently using rule based systems with email only features for mitigating fraud risk, without making changes you can expect to see a lot more false positives as technology enabling privacy in the digital world continues to evolve. Adding scores to your risk assessment, such as Identity Risk Score and Identity Network score which consider more identity elements including phone, address, email, IP, and name, can simplify the challenges of risk assessment as consumer behavior changes.

Learn how Ekata can help identify your good customers and stop fraud. Contact us today.