Secondary Education: Measuring Secondary Uses of 2FA Phone Numbers

Min Hee Kim, Christina Yeung, Daniel Salsburg, Joseph A. Calandrino*
Office of Technology Research and Investigation
Federal Trade Commission

Abstract

If a security feature requires user data, concerns over secondary uses of that data may influence user adoption of the feature. We explore secondary uses of phone numbers that users share for two-factor authentication. Some companies have reused these numbers for purposes unrelated to security, such as targeted advertising. Focusing on top sites, we assessed user-observable secondary uses of phone numbers in two ways. First, we examined web traffic for evidence that sites share numbers with third parties when the user enrolls in two-factor authentication. Second, we monitored calls, voicemail, and text messages to the phone numbers over a two-month period after enrollment. We observed neither form of secondary use in our analysis. Our results suggest a consistent norm against these secondary uses, with potential implications for companies considering practices that deviate from these norms.

1 Introduction

In 2018, Facebook users reported receiving text messages about friends' posts at mobile phone numbers that the users had provided to Facebook only for two-factor authentication (2FA) [16]. Later research suggested that Facebook also used 2FA phone numbers for targeted advertising [18]. As part of the company's $5 billion 2019 settlement with the US Federal Trade Commission, it agreed to restrictions on secondary uses of these phone numbers [5]. In 2019, Twitter also reported accidentally using 2FA phone numbers for advertising purposes [19]. In this paper, we examine user-observable secondary uses of phone numbers that users provide for 2FA.

We focus on phone-based forms of 2FA in which the party authenticating a user—such as a website—typically provides a single-use code (a one-time passcode, or OTP) to the user via a phone call or, more often, an SMS text message. After receiving the code, the user provides it back to the authenticating party. The logic is that this process demonstrates access to calls and text messages to the user's phone number. Knowledge of the code therefore implies possession of the user's phone, yielding a "something you have" second authentication factor that often complements a traditional password. While security concerns have led some to recommend other forms of 2FA over these phone-based methods (e.g., [10, 11]), phone-based 2FA remains common [4].

A user's phone number can be valuable for purposes well beyond security, such as targeted advertising [17]. If a security measure requires users to provide information, possible unwanted secondary uses of that information could discourage users from adopting the measure. Furthermore, unexpected uses of the information also could erode user trust, possibly causing users to avoid similar measures even on sites that do not engage in those practices. We explore the extent to which websites take phone numbers that users provide for 2FA and use those numbers for other purposes.

Limited prior work considers secondary uses of these phone numbers, focusing primarily on Facebook [18]. We took a broader approach. We examined the top 500 sites on Tranco [9], a website popularity-ranking list. We found 32 sites for which we could establish free accounts with phone-based 2FA. For each site, we set up multiple accounts with this form of 2FA and checked for secondary uses of the provided phone numbers in two ways. First, we analyzed web traffic when enrolling in 2FA, searching for direct evidence that the sites share phone numbers with third parties. Second, we examined calls, voicemail, and text messages to the 2FA phone numbers for a two-month period after enrollment, monitoring for activity indicative of secondary uses. While this broader focus may have missed site-specific or other difficult-to-detect secondary uses, it offers larger insights regarding industry norms.

*This report was prepared by staff of the Federal Trade Commission's Office of Technology Research and Investigation and does not necessarily reflect the views of the Commission or any individual Commissioner.

Who Are You?! Adventures in Authentication (WAY) 2020.
August 7, 2020, Virtual Conference.

We present the results of our traffic analysis as well as our call, voicemail, and text message monitoring. Our experiments revealed no evidence of either secondary use. This absence of evidence is itself significant and revealing, suggesting an industry norm against at least these particular secondary uses of information that users provide for security. Because we examined more observable secondary uses, our findings also indicate that users might reasonably be surprised not only by these specific uses but also by less observable secondary uses. As a result, companies should carefully consider factors like what disclosures might be appropriate if making secondary use of data that users provide to facilitate security features.

The remainder of this paper is organized as follows. In Section 2, we discuss related work. Section 3 presents our data collection and analysis approach, and Section 4 provides our findings with regards to secondary uses. In Section 5, we discuss the implications of those findings. Section 6 concludes and explores possible future work.

2 Related Work

We focus on two primary areas of related work: secondary uses of 2FA user data and disclosure tradeoffs by users. Considerable prior work also explores 2FA more broadly, from usability factors [2, 13, 14] to security concerns [10, 11].

Secondary Uses of 2FA User Data. Limited prior work specifically investigates secondary uses of data that users provide for 2FA. Venkatadri et al. [18] examined data sources that Facebook uses for targeted advertising, observing that those sources included 2FA phone numbers. Rather than focusing on a single site, we looked more broadly at top sites, considering general, user-observable secondary uses. Busse et al. [1] conducted a user study of 2FA adoption incentives. Study participants expressed concerns about secondary uses of phone numbers provided for 2FA, such as sale of the numbers or their use for unwanted advertising. Due to privacy concerns, a participant in a study by Das et al. [3] declined to provide a phone number to a website for 2FA. These user concerns further motivate our study of real-world practices. In a user study by Redmiles et al. [12], the majority of participants understood the primary purpose of 2FA to be security.

User Disclosure Tradeoffs. Users may need to disclose personal information to receive a benefit under a wide variety of circumstances. For example, recent work explores user perspectives towards data collection to facilitate the response to COVID-19 [7, 15]. Hoofnagle and King [6] discuss a variety of circumstances in which tradeoffs may be required, such as product registration cards (disclosure of contact information versus benefits like recall notices), store loyalty cards (disclosure of sign-up and shopping information versus possible discounts), and even pizza delivery orders (disclosure of phone number and delivery location versus delicious pizza). A systematic survey of such tradeoffs is beyond the scope of this paper, but we build on the limited existing work studying this tradeoff in the context of 2FA.

3 Approach

To explore whether popular websites use 2FA phone numbers for additional purposes, we began by narrowing a list of popular sites to those that offer both free accounts and phone-based 2FA. We created accounts on those sites and enabled phone-based 2FA on the accounts. When enrolling in 2FA, we collected web traffic. We analyzed that traffic for secondary uses in the form of third-party sharing of phone numbers. For a two-month period, we also monitored calls, voicemail, and text messages to the provided phone numbers. We examined these communications for evidence suggesting use of the phone numbers for purposes beyond 2FA. All steps occurred on the desktop versions of websites.

3.1 Selecting Sites and Enabling 2FA

From April to August 2019, we compiled a list of popular sites that allow user accounts, narrowed the list to sites offering 2FA options, created accounts on those sites, and set up 2FA on the accounts.

Selecting Sites Offering User Accounts. On April 30, 2019, we collected the top 500 sites from Tranco [9], a website popularity-ranking list. We manually narrowed this list, removing sites that did not offer accounts, redirected to other sites for account creation (e.g., from YouTube to Google), required payment or a special relationship for an account (e.g., bank customer or employee), used only third-party login, provided only throwaway accounts, or forced users to install an application to create an account. We also excluded sites that failed to load as well as non-English sites and sites with practical constraints on account creation from the United States.[1] This resulted in a list of 177 sites.

Narrowing Based on 2FA Options. Given our preliminary list of sites offering user accounts, we sought to narrow the list to sites offering phone-based 2FA. Before including or ruling out a site, we created a user account on the site and manually verified available 2FA options. To reduce the possibility of missing difficult-to-find 2FA features on a site, we not only browsed the site looking for 2FA options but also reviewed support documentation on the site and searched the web and third-party sites for guidance on enabling 2FA on the site. We discovered that 57/177 sites offered 2FA, but 12 of these sites limited 2FA availability to particular categories of users or

[1] For example, Flipkart required an Indian phone number for account creation.

used 2FA only under circumstances that we could not reliably replicate (e.g., suspicious activity). Of the 45 remaining sites, 32 sites offered at least one option to enable 2FA with a single phone number.[2]

Setting up Accounts and 2FA. For each of the remaining 32 sites, we created user accounts and enrolled in phone-based 2FA. To help distinguish between activity resulting from 2FA enrollment and spurious calls, voicemail, and text messages, we created multiple accounts per site, each with a unique 2FA phone number. We provided VoIP numbers capable of receiving text messages where possible. If a website did not accept VoIP numbers, we used prepaid mobile phone numbers. We encountered three categories of sites:

- *Case 1 (24 sites):* The site did not require a phone number to create an account. In this case, we did not provide a phone number during account creation. We created two accounts on the site, enrolled both accounts in phone-based 2FA, and monitored the 2FA phone numbers.
- *Case 2 (4 sites):* The site required a phone number for account creation but allowed a different phone number for 2FA. In this case, we also created two accounts on the site. For each account, we used different phone numbers for account creation and 2FA. Each 2FA phone number was unique, and we monitored the 2FA numbers.
- *Case 3 (4 sites):* The site required or set the general phone number for the account to match the phone number for 2FA. One site automatically displayed the general account phone number during 2FA enrollment, preventing the user from changing it. Two sites prompted a user to set or change the general account phone number if the user elected to set a 2FA phone number. The final site allowed users to set a 2FA number, but doing so automatically set the general account phone number as well. We created four accounts per site. We enrolled two accounts in 2FA and left two accounts without 2FA. We set general account phone numbers for each account, even if we did not enroll the account in 2FA. We monitored phone numbers from all four accounts for patterns of differences based on whether accounts enrolled in 2FA.

Beyond the steps necessary to create accounts, enroll in 2FA, and confirm account enrollment, we did not generate any other activity for the study accounts. Once enrolled in 2FA, we simply logged back into the accounts to check that our phone numbers were receiving 2FA messages properly.

When enrolling in 2FA, we watched for statements indicating that the websites might use the phone numbers for purposes other than 2FA. We observed no such statements on tested sites, but we did not examine privacy policies or otherwise seek clarification.

[2]Two sites that we excluded offered more complex schemes mandating multiple phone numbers or phone numbers combined with other methods.

3.2 Monitoring and Analysis

Recall our goal to obtain a broader view of secondary uses than site-specific prior work. We considered practices that would be reasonably user-observable—perhaps with some level of technical expertise—on any site offering phone-based 2FA. We settled on two forms of secondary use.

Third-Party Sharing at 2FA Enrollment. When enrolling in 2FA, we used mitmproxy[3] to capture web traffic. We created all traffic captures in August 2019. We analyzed the traffic for transmission of the phone number to third parties. We considered this secondary use because third-party sharing is well known for other user data [8]. Beyond unencoded and Base64-encoded numbers, we searched for the numbers' MD5, SHA-1, and SHA-256 hashes. Given the phone number 202-555-0173, we would have searched for numbers of the following form: (202)555-0173, 202-555-0173, 2025550173, 555-0173, 5550173, and the first three cases preceded by 1. We examined headers, URLs, request content, and cookies.

Non-2FA Calls, Voicemail, and Text Messages. After we created all accounts and enrolled them in 2FA, we logged calls, voicemail, and text messages to the 2FA phone numbers during a two-month period: September and October 2019.[4] A previously reported secondary use of 2FA phone numbers involved text messages to the phone numbers [16].

We considered communications from any apparent source, not simply the site itself. We were concerned that chance activity unrelated to 2FA enrollment—such as a random robocall or a text message to a previous user of the phone number—could have affected our results. We took several steps to filter messages that were unlikely to represent secondary uses. If we provided a name during account creation, we removed text messages and calls associated with voicemail that addressed a different name. If the nature of a message seemed to be personal (e.g., a school party), we assumed the caller or sender of the text message entered the wrong number or intended to contact the previous phone subscriber.

In all other cases that we encountered (apparent spam messages, cut-off voicemail, silent voicemail, and calls that did not result in voicemail), we looked at patterns of activity across the two 2FA phone numbers on the site.[5] Unless we observed communication from similar phone numbers or the messages were similar across both 2FA numbers, we assumed random unsolicited communication rather than secondary uses. Because we looked for patterns across phone numbers, this approach could have excluded both targeted and unsystematic secondary uses, such as calls only to certain area codes.

[3]https://mitmproxy.org/

[4]If we created two accounts with 2FA and two accounts without 2FA (see Section 3.1), we logged this activity across phone numbers for all accounts.

[5]For sites that required the general phone number and 2FA phone number to match, we also considered the phone numbers for accounts without 2FA.

4 Results

For the 32 sites we analyzed, we uncovered no evidence of either form of user-observable secondary uses. We discuss each form in turn.

Third-Party Sharing at 2FA Enrollment. Recall that we monitored web traffic when we enrolled in 2FA. We observed no evidence that tested sites transmitted the provided phone numbers to third parties. In all cases, observed transmission was solely to the first-party domain with Base64 encoding.

Because our goal was to monitor for observable secondary uses, this analysis may have missed obfuscated transmission, transmission at other times, or transmission of phone numbers directly between the site and third parties. If a site took such steps to transmit data, even a technically sophisticated user might be practically unable to discover the transmission.

Calls, Voicemail, and Text Messages. 2FA phone numbers received calls (900 total), voicemail (44 total), and text messages (58 total). No communication referenced the website associated with a 2FA phone number. Details of the communication and patterns across phone numbers suggest that no activity resulted from 2FA enrollment.[6]

We expected some amount of random communication unrelated to 2FA enrollment: one phone number received a text message before we shared it with any site. Other voicemail and text messages addressed recipients with names different from the name we used when creating the account on the website. For instance, we created two accounts on one site, enrolling both in 2FA. The first 2FA number received text messages about financial activity, but the second 2FA number received no similar messages. That second 2FA number instead received text messages about pharmaceuticals addressed to "Ronald," which was not the name associated with the user account on the site.

A variety of cases arose, and we cannot definitively rule out certain possibilities like targeted secondary uses (see Section 3.2). Nevertheless, the nature of the communication uniformly suggests coincidental activity, like unrelated spam messages, misdials, or communication with the previous user of the phone number.

5 Discussion

Our study revealed no evidence of secondary uses, but that absence is itself meaningful. We examined two observable forms of secondary uses. Neither form is obscure: third-party data sharing is a common practice, and past observed secondary uses of 2FA phone numbers involved sending text messages to those numbers. Among top sites, our findings suggest a consistent norm against these forms of secondary uses for 2FA phone numbers. Such norms would imply that these secondary uses are atypical and could surprise users. Although we focus on two secondary uses, other types also might surprise users, particularly if the result is less obvious than a call or text message to the user.

This suggests that companies should take care if considering secondary uses of information that users provide to facilitate security measures. Beyond considering disclosure of secondary uses, companies should also consider the broader implications of that use, including the possibility that it feeds user concerns about the measure [1]. Such an outcome could discourage adoption of the security measure in other contexts.

Companies should also consider whether they can take steps to address concerns regarding secondary uses. A company may be able to build trust and promote security by disclosing how it uses information that users share for security purposes, including any commitments to avoid secondary uses. Offering 2FA options beyond phone-based ones could allow users to reap the benefits of 2FA without sharing their phone numbers (and also could avoid security concerns associated with phone-based 2FA [10, 11]). The points in this paper remain applicable to any alternatives to phone-based 2FA. If a site instead requires users to install a mobile application for 2FA, unnecessary permission requests or data collection by the app could similarly discourage 2FA adoption.

6 Conclusions and Future Work

We examined two categories of user-observable secondary uses of phone numbers that users provide for phone-based 2FA. Our findings suggest a norm against such secondary uses by top sites. Companies should carefully consider whether, when, and how they use data that users provide to facilitate security measures.

Future work could address limitations in this study. For example, we excluded paid and non-English sites, but both are popular and important categories of sites. In addition, our two-to-four accounts per site had limited user activity, and the observation period was two months. Monitoring a larger number of accounts with greater user activity over a longer period of time could reveal cases in which secondary uses occur. Any work that scales up or automates this analysis could be valuable. Exploration of additional secondary uses—including site-specific or difficult-to-detect uses (e.g., targeted advertising)—could offer complementary findings as well.

Extensions to this work also include exploring additional cases in which users may need to disclose information for a security benefit. For example, mobile authentication applications provide an alternative to phone-based 2FA. Future research could examine the permissions those applications request and the data they collect.

[6] Anecdotally, many calls and text messages came from numbers associated with existing complaints, including consumer complaints to the Federal Trade Commission.

Finally, useful future projects could explore user understanding, attitudes, and behavior surrounding secondary uses of data for security measures. Such work could explore what uses are unexpected or unwanted. It could also examine the impact of secondary uses on the adoption of security measures, including whether secondary uses by one party influence user adoption of similar security measures elsewhere.

Acknowledgments

We thank Tyira Bunche, Emily Liu, and Phoebe Rouge for assisting with this study. We also thank Isabella Faccone for assisting with a preliminary study.

References

[1] K. Busse, S. Amft, D. Hecker, and E. von Zezschwitz. Get a free item pack with every activation! Do incentives increase the adoption rates of two-factor authentication? *Journal of Interactive Media (i-com)*, 18(3), Nov 2019.

[2] S. Das, A. Dingman, and L. J. Camp. Why Johnny doesn't use two factor: A two-phase usability study of the FIDO U2F security key. In *FC 2018*.

[3] S. Das, A. Kim, B. Jelen, J. Streiff, L. J. Camp, and L. Huber. Why don't older adults adopt two-factor authentication? In *2020 SIGCHI Workshop on Designing Interactions for the Ageing Populations - Addressing Global Challenges*.

[4] M. Engler. State of the auth: Experiences and perceptions of multi-factor authentication. *Duo Labs Report*, December 9 2019 (accessed May 20, 2020). https://duo.com/assets/ebooks/state-of-the-auth-2019.pdf.

[5] Federal Trade Commission. FTC imposes $5 billion penalty and sweeping new privacy restrictions on Facebook. July 24 2019 (accessed May 10, 2020). https://www.ftc.gov/news-events/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions.

[6] C. J. Hoofnagle and J. King. Research report: What Californians understand about privacy offline. May 15 2008 (accessed May 20, 2020). https://dx.doi.org/10.2139/ssrn.1133075.

[7] G. Kaptchuk, D. G. Goldstein, E. Hargittai, J. Hofman, and E. M. Redmiles. How good is good enough for COVID19 apps? The influence of benefits, accuracy, and privacy on willingness to adopt. May 20 2020 (accessed May 20, 2020). https://arxiv.org/abs/2005.04343v4.

[8] B. Krishnamurthy and C. E. Wills. On the leakage of personally identifiable information via online social networks. In *WOSN 2009*.

[9] V. Le Pochat, T. Van Goethem, S. Tajalizadehkhoob, M. Korczyński, and W. Joosen. Tranco: A research-oriented top sites ranking hardened against manipulation. In *NDSS 2019*.

[10] K. Lee, B. Kaiser, J. Mayer, and A. Narayanan. An empirical study of wireless carrier authentication for SIM swaps. In *ConPro 2020*.

[11] P. Markert, F. Farke, and M. Dürmuth. View the email to get hacked: Attacking SMS-based two-factor authentication. In *WAY 2019*.

[12] E. M. Redmiles, S. Kross, and M. L. Mazurek. How I learned to be secure: A census-representative survey of security advice sources and behavior. In *CCS 2016*.

[13] E. M. Redmiles, E. Liu, and M. L. Mazurek. You want me to do what? A design study of two-factor authentication messages. In *WAY 2017*.

[14] K. Reese, T. Smith, J. Dutson, J. Armknecht, J. Cameron, and K. Seamons. A usability study of five two-factor authentication methods. In *SOUPS 2019*.

[15] L. Simko, R. Calo, F. Roesner, and T. Kohno. COVID-19 contact tracing and privacy: Studying opinion and preferences. May 8 2020 (accessed May 20, 2020). https://arxiv.org/abs/2005.06056v1.

[16] N. Slatt. Facebook's two-factor authentication system auto-posts replies on your profile. *The Verge*, February 14 2018 (accessed May 10, 2020). https://www.theverge.com/2018/2/14/17014116/facebook-2fa-two-factor-authentication-auto-post-replies-status-updates-bug.

[17] G. P. Slefo. Google AdWords adds targeting by phone number, mailing address. *AdAge*, December 12 2017 (accessed May 10, 2020). https://adage.com/article/digital/google-adwords-adds-target/311609.

[18] G. Venkatadri, E. Lucherini, P. Sapiezynski, and A. Mislove. Investigating sources of PII used in Facebook's targeted advertising. In *PETS 2019*.

[19] Z. Whittaker. Twitter admits it used two-factor phone numbers and emails for serving targeted ads. *TechCrunch*, October 8 2019 (accessed May 10, 2020). https://techcrunch.com/2019/10/08/twitter-admits-it-used-two-factor-phone-numbers-and-emails-for-targeted-advertising/.