"Adulthood is trying each of the same six passwords that you use for everything": The Scarcity and Ambiguity of Security Advice on Social Media

SRUTI BHAGAVATULA, Carnegie Mellon University, USA
LUJO BAUER, Carnegie Mellon University, USA
APU KAPADIA, Indiana University Bloomington, USA

In order to keep one's computing systems and data secure, it is critical to be aware of how to effectively maintain security and privacy online. Prior experimental work has shown that social media are effective platforms for encouraging security-enhancing behavior. Through an analysis of historical social media logs of 38 participants containing almost 200,000 social media posts, we study the extent to which participants talked about security and privacy on social media platforms, specifically Facebook and Twitter. We found that interactions with posts that feature content relevant to security and privacy made up less than 0.09% of all interactions we observed. A thematic analysis of the security- and privacy-related posts that participants interacted with revealed that such posts very rarely discussed security and privacy constructively, instead often joking about security practices or encouraging undesirable behavior. Based on the overall findings from this thematic analysis, we develop and present a taxonomy of how security and privacy may be typically discussed on social networks, which is useful for constructing helpful security and privacy advice or for identifying advice that may have an undesirable impact. Our findings, though based on a fraction of the population of social media users, suggest that while social networks may be effective in influencing security behavior, there may not be enough substantial or useful discussions of security and privacy to encourage better security behaviors in practice and on a larger scale. Our findings highlight the importance of increasing the prevalence of constructive security and privacy advice on online social media in order to encourage widespread adoption of healthy security practices.

CCS Concepts: • **Security and privacy** → **Social aspects of security and privacy**.

Additional Key Words and Phrases: security, privacy, security information, social media

ACM Reference Format:
Sruti Bhagavatula, Lujo Bauer, and Apu Kapadia. 2022. "Adulthood is trying each of the same six passwords that you use for everything": The Scarcity and Ambiguity of Security Advice on Social Media. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 264 (November 2022), 27 pages. https://doi.org/10.1145/3555154

1 INTRODUCTION

Computers and online accounts have become ubiquitous for the majority of the population [70]. With the growth of new technology, the number of dimensions across which security and privacy need to be maintained is also growing. Ideally, computer systems would be designed to shield users from the complexity of implementing and maintaining security and privacy [2, 72]. However, the current state of the digital world requires users to be aware of how to protect the security of their online accounts and the privacy of their data [6]. For example, defending against social engineering

Authors' addresses: Sruti Bhagavatula, Carnegie Mellon University, Pittsburgh, USA, sbhagava@alumni.cmu.edu; Lujo Bauer, Carnegie Mellon University, Pittsburgh, USA, lbauer@cmu.edu; Apu Kapadia, Indiana University Bloomington, Bloomington, USA, kapadia@indiana.edu.

2573-0142/2022/11-ART264
https://doi.org/10.1145/3555154

attacks like phishing continues to rely on some degree of security awareness [15, 47], as does reacting to account breaches by resetting passwords on potentially many online accounts [10], and recognizing that sensitive data should not be shared in online conversations [69]. As a result, despite the increased user burden, security and privacy awareness in general remains essential for people to have the tools and knowledge to keep their systems and data secure and private [6, 39, 49].

Prior work has shown that social media are some of the most prevalent environments for discussing security and privacy [26] and that, in experimental settings, they are effective platforms for encouraging friends to adopt security-enhancing behavior [21, 22, 24]. Inspired by these findings and the prevalent use of social media among adults [62], we study what and how much content about security and privacy is shared on social networks. Using a dataset of real-world Facebook and Twitter behavior, we examine in particular how prevalent the sharing of information about security and privacy may be in practice and what such information may typically convey, e.g., personal anecdotes and educational discussions. More specifically, we focus on two research questions: (1) how common is sharing and consuming security and privacy content on social media?; and (2) what are the different ways in which security and privacy are mentioned in social media posts? These "mentions" can refer to individual posts containing security- or privacy-related terms or back-and-forth dialogues about security and privacy. To answer these questions, we study the Facebook and Twitter logs of 38 participants who were enrolled in the Security Behavior Observatory (SBO) [32, 33], a longitudinal study of computer security behaviors collected from participants' home computers (see Sec. 3). We chose to collect and analyze data of SBO participants because the SBO allows us to also analyze the security behaviors of participants by collecting browser and system usage logs. We use these security behaviors as part of an exploratory analysis of the relationship between interactions with security and privacy content on social media and users' security behavior. We discuss the limitations of the SBO dataset in Sec. 6.

Perhaps surprisingly, we found that interactions with security- and privacy-related posts were scarce. When counting such interactions we considered both the consuming of content (i.e., liking, saving, or commenting on posts) and the sharing of content (i.e., (re-)sharing or retweeting posts). Only 131 of the 194,081 Facebook posts that participants interacted with, and 44 of the 6,883 Twitter posts, were related to security and privacy (0.09% in total). In fact, of the 38 participants almost half (47%) did not interact with any security and privacy content at all. We further found that the amount of interactions with security- and privacy-related posts was not correlated with any demographic characteristics or with technical savviness. To validate the above findings, we constructed a dataset of 15,040 tweets made by 1,000 random Twitter users and identified security and privacy posts within this set. We found that only 0.08% of the posts were related to security and privacy—slightly less than in our main dataset.

We examined the few security and privacy posts further to understand what they were trying to convey, to understand, for example, whether the posts were educational in nature and could help encourage healthy security behavior. We examined these posts through a thematic analysis [13] and identified five themes in what the posts about security and privacy conveyed. One of these themes described posts that presented obviously constructive security and advice. Posts with this theme may be the most clearly helpful in encouraging desirable security practices. However, only 10 participants interacted with such posts and not as often as with posts that demonstrated the other less constructive themes. A different, more prominent theme described posts that mentioned security and privacy but were either ambiguous about whether they were promoting desirable or undesirable practices or recounted unconstructive anecdotes or jokes about security and privacy. The other three themes included brief mentions of security and privacy topics without any substance (which 16 participants interacted with), encouragement or demonstrations of detrimental security

and privacy practices (seven participants), and information about public policy topics related to security and privacy (six participants).

Our results showed that, for participants in our sample, security and privacy were not topics that people frequently interacted with or were exposed to on social media. Given that our sample was slightly skewed towards technically-savvy and educated people (see Sec. 3.2), we suspect that the general population may have even less exposure than we observed, as hinted at by the even smaller amount of security- and privacy-related posts we found in the random Twitter sample. Even when posts referenced security and privacy, rarely was this to deliver constructive security advice or recount secure behavior, with posts more often providing little to no useful commentary on security practices. As a result, participants who interacted with the security and privacy posts we analyzed may have obtained higher security awareness, but are not likely to have taken away actionable advice or adopted security-enhancing behaviors. The latter is supported by an exploratory investigation we performed on the relationship between users' security behaviors and the posts they interacted with. While prior work showed through interventions that demonstrating healthy security practices on social media can correlate with increased adoption of healthy security behavior, our findings suggest that without interventions, the amount of constructive security advice being shared could be too low to effectively encourage healthier security behavior. Achieving more widespread adoption of security best practices could potentially be helped by two broad steps: increasing the spread of security and privacy advice using wide-reaching channels such as social media; and presenting the advice in such a way that user burden is reduced and such that it incites changes in their security behavior. To achieve this first step, content creators and distributors could structure the information such that it discusses security constructively. These content creators can take lessons from other domains about how posts' popularity can increase on social media. We conclude with a discussion about how our findings inform future work in these directions, as well as a discussion on what could constitute effective security education as part of the disseminated advice.

2 RELATED WORK

We discuss related work on the role of social influence in security and privacy and the applications of social media as an educational tool.

2.1 Social influence in security and privacy

Prior work has repeatedly shown that social influence plays a role in the adoption of security-enhancing habits.

Informal stories told by friends and family members were found to be significant sources of information and encouragement to advise others about security [65]. A training program that designated individuals in their community to educate others in the community about security was found to be a feasible approach to security education both through in-person and social-media interactions [58]. Researchers also found that users were inclined to change their security behavior as a result of indirect group or peer pressure [20]. When making security decisions, participants in a study were more influenced to take defensive measures if they were told that several of their friends made that decision as well [29]. Similarly, users of a proposed filesystem implemented security features when they were shown that other users had also implemented them [24]. Social networks have also been used to directly implement social influence to encourage different kinds of behavior. Studies that examine social influence within a social network introduced the concept of "social announcements," which tell users how many other users in their network are using certain security features. Researchers found that these social announcements were correlated with increased adoption of security features and awareness [21]. Disseminating information about

security features through a social network increased the adoption of these features, but this adoption also depended on other factors, such as the perceived visibility of the feature and the number of distinct social groups within the network surrounding the users that shared the information.

Perceived benefit or visibility has been found to be an important factor in people changing their security behavior despite social pressure [6, 41, 84]. Social networks have also served as effective platforms for people to engage in discussions about security and privacy experiences, advice, and complaints [26]. Overall, prior work has shown that social influence is an effective tool for encouraging adoption of certain security-enhancing behaviors and decisions. However, such social influence has also been seen to have a negative effect on computer users' security practices. Researchers found that users viewed security-conscious people as paranoid or as exhibiting undesirable behavior [35]. Therefore, they were less inclined to implement practices taken up by security-aware people, e.g., using encrypted mail.

Inspired by existing work, our paper measures—using empirical, real-world data—how social media may be being used as a platform to talk about security and privacy.

2.2 Social media as an educational tool

In the work described above, social network interventions were used to influence other users to adopt security features. While this could be considered educational about other people's behaviors, social media has also been used as a purely educational tool for spreading awareness and education in areas not related to security and privacy.

Prior work has used social media to increase awareness in various scientific domains. For instance, numerous prior works have discussed and shown the effectiveness of spreading awareness about environmental conservation through social media posts [38]. One of these studies has shown that not only is social media (i.e., Facebook, Twitter, YouTube) effective for this, but that people often use social media for the purpose of absorbing new information, making social media as an educational tool particularly promising [45]. In a similar vein, researchers have also found that social networks used for messaging such as WeChat can be used to increase awareness of wildlife conservation measures and to reduce misunderstandings of policies implemented by policymakers and recommendations by scientific experts [82]. They found that particular characteristics of the content shared such as more images and fewer words were correlated with higher engagement and popularity. Other work has highlighted the issue that dissemination of scientific issues and awareness should not solely rely on traditional scientific journals or legacy media to reach beyond the scientific community; they reported that in addition to simply using social media to spread awareness, the specific way that people use social media is also a determining factor in how issue awareness is spread [54]. Social media has also been extensively been used in other areas for example, to launch mental health awareness campaigns [71] and to increase geographic awareness across the world [83].

Existing work in this space motivates our goal of studying how often and how effectively security and privacy are discussed on social media in the wild. In particular, this will help us understand how to improve the dissemination and presentation of security and privacy information on social media such that it can be used as an educational tool to reach more people.

3 DATA COLLECTION AND DATASET

3.1 Data collection

We used data collected by the Security Behavior Observatory (SBO) project. The SBO is a longitudinal study of the usage patterns and security behaviors of Windows computer users [32, 33]. The SBO started recruiting participants and collecting data in October 2014 and ceased data collection in

July 2019; participants were enrolled in the study at different times and for different durations. We used data collected from January 2015 upto April 2018. SBO participants' home computers were instrumented with software that periodically and automatically collected data. The data collection software collected data via system-level components and browser extensions in Chrome, Mozilla Firefox, and Internet Explorer. System details such as system configuration, installed software, operating system updates, updates to the filesystem, and other system events were collected via system-level processes. Browser events including browsing history, content permissions granted to websites, and passwords entered were collected by the browser extensions.

The SBO project was approved by the ethics review board at its home institution (Carnegie Mellon University); we obtained additional approval for its use. Participants were compensated by the SBO project with $30 for enrolling and an additional $10 each month they were enrolled in the study.

Collecting Facebook and Twitter data. We collected additional social media data from a subset of consenting SBO participants through another study approved by the relevant institutions' ethics review boards. All Facebook posts are not visible or searchable to the public while Twitter posts created by non-private accounts are publicly available online. Specifically, we collected Facebook and Twitter data between January 2015 and April 2018. For Facebook, we collected the content on participants' "Activity Log" pages (a log of all activity of a Facebook user). For Twitter, we collected participants' tweets (including retweets and replies to tweets), the tweets they've favorited, the Twitter accounts they follow, and the Twitter accounts that follow each participant. The logs we collected covered *all* historical social media activity (not including private messages) irrespective of the device on which participants may have used the application. In particular, the logs for a given Facebook or Twitter user contained posts made by the user and posts the user interacted with (e.g., liked, commented on, saved, or re-shared). The posts a user interacted with were visible to that specific user, for example, they could have been public to the internet, made by friends within the network, or made in groups to which the user belonged. When displaying the text of posts in this paper (see Sec. 5), to protect participant anonymity, we only report the text of Facebook posts that are private to users or are made by public Facebook pages[1]. Participants who consented to this additional study received $15 compensation in the form of an Amazon gift card.

We created developer applications for both Facebook and Twitter [30, 79] and asked participants to log into our applications through a webpage presented after the consent page to our additional study. Participants had the option to provide data for one or both platforms. We collected a participant's Facebook data as follows: using the participant's login credentials, we fetched their name using the Facebook API [30]. We then exclusively stored the MD5 hash of the name for subsequent use, ensuring that we were never directly working with the participant's identifying information. Since the API does not provide functionality to retrieve the activity log of a user, we instrumented the SBO browser extension to trigger data collection when the participant logged into Facebook and visited the Facebook homepage from their SBO computer. Before starting to collect data, the extension ensured that the hash of the name on the visited Facebook homepage matched the hash of the name fetched by the API. When the above criteria were met, the extension loaded the participant's activity log or page likes webpage in an invisible browser tab, scrolled through the page, and downloaded the contents. If data collection was halted due to logging out, it was resumed the next time the participant logged in to Facebook. For Twitter, we used the credential tokens provided when participants logged into our Twitter app and used the Twitter API [79] to fetch the above-mentioned Twitter data for each participant.

[1]We do not display Twitter posts because most posts are public and they don't differentiate between regular users and pages.

3.2 Dataset

Our study is based on the longitudinal data collected by the browser extensions and the additional collected social media data for 38 participants. Specifically, 34 participants provided us with Facebook data, 16 with Twitter data, and 12 with both Facebook and Twitter data. The participants who provided Facebook data interacted with 5708 Facebook posts (i.e., they liked, commented on, saved, created their own, or shared others' posts) on average over the four-year period. Participants who provided Twitter data interacted with 601 tweets (i.e., they favorited, commented on, created their own, or retweeted others' posts) on average. Tables 5 and 6 in App. A.2 describe the number of posts for each type of interaction. App. A.1 describes the type of data associated with each Facebook and Twitter post. Other datasets we use in our analyses are also summarized in App. A.1.

Participants' ages ranged from 21 to 81 years with a mean age of 33. Participants were female-skewed (74%). A little more than half (53%) knew at least one programming language and 18% had programming as their primary profession. 61% were students and 53% had a bachelor's degree or higher.

We discuss limitations of our dataset (such as the small sample size) in Sec. 6.

4 HOW COMMON WERE SECURITY AND PRIVACY DISCUSSIONS ON SOCIAL MEDIA?

Our first research question examines how often participants interacted with content that mentioned topics explicitly related to computer or online security and privacy on Facebook and Twitter.

We compiled a set of all the Facebook and Twitter posts the participants interacted with during the timeframe discussed in Sec. 3 (see Sec. 3.2 for the definition of "interacted"); this included all the Facebook posts for the 34 Facebook users and Twitter posts for the 16 Twitter users. For a given interaction with a post on Facebook or Twitter, we then extracted and concatenated all pieces of text corresponding to the components of the post. For example, if a post was being re-shared, we considered both the original text and the shared text. Similarly, if a post was commented on, we considered the original post and the comment text. We identified posts related to computer or online security and privacy by examining this text for each post.

We iteratively categorized posts as related to digital security and privacy as follows. Two researchers who are domain experts in security and privacy initially created a list of regular expressions (regexes) by reviewing all posts and by brainstorming for regexes related to digital security and privacy (see App. A.3 for this initial list). We then systematically built a list of regexes, starting with the aforementioned initial list, by iterating through the following steps:

(1) Match the regexes in the list to the set of all collected posts.
(2) Look through each post matched by any of the regexes. Identify strings within the matched posts that are relevant to security and privacy, but which do not yet have corresponding regexes in our list.
(3) Manually examine each post in a random sample of one hundred of the unmatched posts. Identify posts that should be classified as security- or privacy-related and strings relevant to security and privacy within each post.
(4) Construct regexes for each of the new strings from the previous two steps and add them to the list of regexes.
(5) Repeat steps 1–4 until steps 2 and 3 yield no new strings from which regexes could be created.

After constructing the final list of regexes (see Table 1), we flagged all the posts that matched any of the regexes. To ensure the absence of false positives, we manually examined all the flagged posts and unflagged them if they appeared not to be related to security and privacy. We verified

we did not incur false negatives by inspecting random samples of 100 unmatched posts at each iteration, and once at the end before stopping, which showed no false negatives.

We verified that sampling 100 posts at the end of each iteration was sufficient to ensure no false negatives. We did this by sampling and examining 1000 posts that were not flagged as related to security and privacy after executing the above procedure. We found that none of these 1000 posts were related to security and privacy, implying that our approach unlikely missed relevant posts.

We used the above approach after exploring a few possibilities. Originally, one approach was to construct a topic model based on the text of individual, uncategorized posts. However, the resulting topic models did not produce coherent topics. Another approach we considered was to manually divide a set of posts into posts related and unrelated to security and privacy according to some criteria; then build a topic model over posts in these two categories; and then use the topic model to categorize additional posts. However, it was difficult to obtain a sound initial categorization without an approach like the iterative approach described above.

We found 131 of the 194,081 Facebook posts (0.07%) and 44 of the 6,883 Twitter posts (0.6%) in our dataset to be related to security and privacy. To confirm that this proportion wasn't an artifact of our dataset, we compared it to the proportion of security and privacy content in a dataset of 100 tweets made by each of 1000 randomly selected Twitter users from a public dataset [14]; the resulting dataset contained 15,040 tweets, as not all Twitter users had 100 tweets to collect. Matching our final list of regexes against the set of random tweets, 0.08% of posts were flagged as related to security and privacy. To ensure the absence of false negatives as a result of using the list of regular expressions for our collected data on the separate random Tweets dataset, we sampled 100 random unflagged tweets of the 15,040 posts and found that they were correctly categorized as not related to security and privacy.

Based on the results for two different datasets, we find that security and privacy content was scarcely interacted with on social media by the participants.

Security and privacy keywords		
password	cybersecurity	bitcoin
social security number	cyber.security	net.*neutral
security camera	de-verify	secure
key.*security	security.*protocol	jailbreak
security.*account	security	jail.*break
security question	privacy	comp.*virus
phishing	hack	

Table 1. Lists of regexes used to flag Facebook or Twitter posts related to security and privacy. Initial regexes are in black while the regexes in green were added via the iterative process. Although "bitcoin" may not seem directly related to security and privacy, we considered it to be related because it is a digital currency that provides privacy protection for its users.

5 HOW DID POSTS ACTUALLY TALK ABOUT SECURITY AND PRIVACY?

In Section 4 we reported on the low occurrence of interactions with security- and privacy-related content in our set of Facebook and Twitter posts. In this section, we dive deeper into the posts themselves by conducting a thematic analysis of the Facebook and Twitter posts that the social network users interacted with. The purpose of this analysis is to characterize the types of security-related content people may come across as they're browsing social media. While prior work has studied how helpful security and privacy information is presented across the web [64, 68], in this section we study what information people may actually be exposed to which may go beyond only

helpful content. In particular, we identify themes of how the posts discuss security and privacy and what they convey.

5.1 Methodology

As the first step of the thematic analysis, two researchers familiarized themselves with every security- and privacy-related Facebook and Twitter post in our dataset. If a link was present in the post, we considered the link's content as part of the post. Using an inductive approach, one of the researchers conducted a round of open coding of each post [8]. Axial coding was then used to derive higher-level codes or categories related to what the post appeared to convey [46], e.g., "sarcasm about security advice," "demonstrating a constructive action," "story about bad experience." They then coded each post according to these derived codes. Following this, the researcher identified a set of overarching themes across the identified higher-level categories. Through frequent discussions among the research group, the researchers iterated on the categories and the resulting themes to ensure that every post could be described by at least one of the themes (i.e., posts could be described by more than one theme). Table 2 shows an example of codes and themes in this process. Because we conducted a thematic analysis, we did not compute the inter-rater reliability between two coders, a decision that is supported by prior work [3, 5, 53].

5.2 Results

We identified five major themes in how the Facebook and Twitter posts mentioned security and privacy.

5.2.1 Quick mentions of security or privacy. Some posts mentioned a security or privacy topic (e.g., passwords, net neutrality, and Bitcoin) but did not elaborate on that topic or that topic was not the focus of the post. Such posts also included announcements about something related to security and privacy but without details of the actual phenomenon, e.g., a post that announces a cybersecurity article or paper but doesn't describe the topic. The following are examples of posts that exhibit this theme.

> *Overseen: Girl from The Ring supports data privacy*
>
> *I told y'all Bitcoin would crash lol*
>
> *Atoms, stars, the solar system, cyber security, creativity - just a few of the amazing things you can learn about in your spare time.*

In all three posts, terms related to security and privacy are mentioned (i.e., "data privacy", "bitcoin", and "cyber security"). However, these terms do not describe the main topic of the post.

Of the 34 Facebook users and 16 Twitter users, 14 and five respectively interacted with posts corresponding to this theme; a total of 16 participants interacted with posts in this theme across both Facebook and Twitter. When reporting on the "total" for the subsequent themes, we mean the total over the union of the Facebook and Twitter users.

5.2.2 Ambiguity in the intended message. Some posts discussed security and privacy in more detail but did not convey a clear message of encouraging or discouraging certain security behaviors. These posts were sometimes sarcastic, included anecdotes about security experiences, or joked about security topics.

> *Right to Privacy. LOL.*

This post talks about the right to privacy but follows it up with "LOL" (the acronym for "laugh out loud") which indicates that the poster thinks the idea of the right to privacy is funny. It is unclear whether they are trying to convey something constructive or otherwise with this post.

Passwords: it's just too much. (Based on a true story)

Here, the poster talks about passwords being too much but without explaining why they are "too much" or the security advantages and disadvantages of passwords.

> *The mandatory computer security training I have to do for work just advised me to "only install mobile apps that are absolutely necessary". Also apparently hackers can kidnap your child, steal your laptop to sell on the black market (specifically the black market), and also take your job. Those tricksy hackers.*

This post appears to express sarcasm about recommended security behavior. While the poster does speak about the advice to only install necessary apps, the second hyperbolic sentence suggests that they may be ridiculing the advice.

A total of nine participants interacted with posts in this theme, encompassing eight Facebook users and three Twitter users.

5.2.3 Constructive security or privacy advice. A less frequent type of post encourages constructive security practices directly or indirectly; for example, by sharing an anecdote about how certain security practices helped them, criticizing poor security practices, or explicitly delivering constructive advice.

> *Hi everyone. If you received an email from me. Don't open it, good ole gmail was hacked. # thanksgmail*

In this post, the poster tells their friends to not trust and open emails sent by their hacked account. This advice is constructive to making sure their friends are not victims of the hacker given the specific situation.

> *If it said you posted it, you may have gotten "phished" and need to change your password.*

In the above post, the poster suggests how to tell if someone's account was compromised as a result of phishing and describes the recommended remediation to protect their account.

The following post again explicitly provides constructive advice for healthy security behavior.

> *MINER ALERT: Should anyone see a AD here on FB for free downloadable photo enhancer called InPixio, do not...I repeat.. DO NOT download as it is harmful to your computer, my Internet Security scanner raised red flag that it was automatically removed on account of being not just virus related but hacks your computer as well*

In total, 10 participants interacted with constructive posts in this theme, including nine Facebook users and two Twitter users.

5.2.4 Demonstration of detrimental security practices. Posts can often demonstrate undesirable security practices, e.g., by asking for advice to help execute these practices, sharing passwords inside posts, or speaking positively about unsafe practices.

> *The wifi password is probably puravida Thank you Costa Rica...*

> *I want WomanlyAlways1895 on my gravestone Lol does anyone know what the for sisters password on the chi o wix page is*

In the above two posts, the posters mention or ask for a password in a post. Whether the password information was intended to be public or whether it was posted to a private group, the post demonstrates a bad practice of posting or asking for passwords online which may encourage the same behavior in others.

A total of seven participants interacted with these negatively constructive posts, all of whom were Facebook users.

Open codes	*Axial codes*	*Themes*
announcement of security engineer job	announcement of security-related jobs/events without discussing security concepts	Quick mentions of security and privacy
announcement of security project with no project description		
link to security/privacy podcast with no description of the content		
mentions Bitcoin without security implications	mentions concept related to security but security is not the focus	
opinion about net neutrality without discussing security/privacy implications		
announcement of a security incident	sharing articles discussing incidents/breaches with helpful information	Constructive security or privacy advice
link to article about breached passwords		
opinions on why a security feature would be fine and not bad	commentary on bad/good security practices	
comment on bad security practices by online banking sites		

Table 2. Sample codes and their groupings into axial codes and subsequently, broader themes.

5.2.5 Information or advice about public policy topics. The final theme describes a different flavor of posts, which are less about the technical details of security and privacy but instead (or additionally) provide information on security and privacy topics in the public policy sphere. For example, posts that provide information on net neutrality and its implications or links to events or talks about the intersection of public policy and security or privacy would be described by this theme. The following are two examples of this kind of post:

> *Do you enjoy Netflix? Do you find yourself spending too much time on FB? If net neutrality goes away, our Internet bills go up and we give power to companies like Comcast and Spectrum to control what information we can access. Here's what you can do...*

> *The FCC just announced its plan to slash net neutrality rules, allowing ISPs to block apps, slow websites, and charge fees to control what you see and do online. They vote December 14th. Call your representatives today to tell them to fight for net neutrality! Learn how to do that at http://battleforthenet.com*

Posts in this theme occurred the least frequently among participants with only four Facebook users, four Twitter users, and six participants in total interacting with such posts.

For each of the five themes, Tables 3 and 4 show the number of posts in each theme that correspond to each interaction type.

Interaction/Theme	*1*	*2*	*3*	*4*	*5*
liked	32	13	17	4	4
commented	17	20	14	5	3
created/shared	3	1	1	0	0

Table 3. Number of Facebook posts in each theme that correspond to each interaction type.

Interaction/Theme	*1*	*2*	*3*	*4*	*5*
liked	21	3	3	0	6
commented	1	0	0	0	0
created/shared	7	1	0	0	2

Table 4. Number of Twitter posts in each theme that correspond to each interaction type.

5.3 Exploring the relationships between social media interactions and security behavior

Out of the posts related to security and privacy we analyzed, we could consider that the kinds of posts that may actually trigger changes in security behavior or habits are the posts in the third theme (constructive security or privacy advice). However, only 10 participants interacted with those posts and such posts occurred infrequently (20% of the 175 security and privacy posts and 0.02% of all social media posts in our dataset). With the low number of security and privacy posts we found overall (Sec. 4) and the even scarcer amount of constructive advice in these posts, we hypothesized that though participants' security behavior may be affected by social influence in a social network in experimental settings [21], the discussions we observed about security and privacy on Facebook and Twitter may not have been prevalent enough to trigger such an effect in the wild.

We tested this hypothesis through an exploratory statistical analysis. In this analysis, we analyzed features related to participants' social media interactions and outputs related to their security behavior.

We computed 44 features representing social media interactions with posts related to technology, security and privacy, and data breaches. We identified posts in the latter two categories in a similar manner as described in Sec. 4. These features described, for example, the number of interactions a participant had with posts in each category, the number of such interactions that involved consuming content or sharing content, and the popularities and familiarities of the authors of posts participants interacted with. A list of all features and their descriptions can be found in Table 8 in App. A.4.

We computed 25 outcomes representing security behavior. These outcomes spanned behaviors concerning a variety of system and browsing events. Among the system events, the outcomes described, for example, the number of computer updates on participants' computers, the frequency of software updates, the number of antivirus software programs installed, and the number of times participants connected to open and insecure Wi-Fi networks. Browsing-related outcomes included the number of times participants' changed passwords, the number of malicious websites they visited and files they downloaded, and the number of websites to which they granted microphone, camera, or location permissions. A list of all security behavior indicators along with their descriptions can be found in Table 9 in App. A.4.

We built quantile regression models [48] to model the relationships between the above social media interactions and security behavior outcomes. We found no demographics or social media factors to be consistently statistically significant. Therefore, we conclude that interactions with security and privacy content on these platforms, whether as a result of sharing or consuming, were not associated with improved security behavior.

App. A.4 describes these analyses and results in greater detail.

6 LIMITATIONS

We analyzed the data of a (relatively) small subset of participants in the SBO due to the difficulty of recruiting participants to agree to additional data collection. Popular social networks such as Facebook do not provide access to activity logs through their API, and obtaining access to such data is otherwise difficult [30]. Our dataset is a tradeoff between a large number of participants (which it doesn't have) and extensive, hard-to-obtain, longitudinal Facebook data (which it does have).

As part of our approach for detecting security and privacy posts in the Facebook and Twitter posts, we ensured the lack of false positives by randomly sampling 100 posts in each iteration and sampling 1000 unmatched posts at the end. Even though we found no false negatives in these random samples, there is a chance that these samples still missed relevant posts. Additionally, the regular expressions corresponding to search terms we constructed ended up appearing to correspond to very specific phrases. However, since these keywords and expressions were constructed based on the above process whose effectiveness we validated, we believe that the posts that discussed security and privacy discussed them using these specific terms.

Although 53% of the 38 participants interacted with security and privacy posts, the majority of them only interacted with two or fewer posts. Only 29% of the 38 participants interacted with more than two security- and privacy-related posts; therefore, those 29% accounted for the majority (93%) of the security and privacy posts we analyzed. Although this corresponds to a low number of participants whose behavior we analyzed, this also substantiates our findings that participants did not often interact with security and privacy posts.

The data we collected contains information only about posts that participants explicitly interacted with on Facebook or Twitter. Therefore, we could analyze only those posts, and not, for example, posts that participants may have seen but with which they did not interact. While such posts may further contribute to our understanding of how security and privacy is discussed on social media and how often, we were particularly interested in studying active engagement with content as opposed to passive engagement (i.e., seeing posts without interaction), including because evidence of active engagement implies that the participant has actively comprehended a post, as opposed to idly scrolling past it without reading or understanding it.

The participants in our study were female-skewed with a relatively high percentage of people who knew at least one programming language. Such technically savvy people have been found to be more likely to read about security information [10]. Even though our results are not based on a representative population, we believe the frequency of security and privacy posts will likely be even lower with a less technically savvy population. Furthermore, our thematic analysis reveals important patterns for how security and privacy may be discussed on social media that are likely independent of demographics (mentioned in Sec. 1).

The SBO collected data exclusively from Windows computer users. People who use other operating systems may exhibit different social media and security behaviors. However, as Windows is the most commonly used OS for personal computers [11], we do not believe our findings are fundamentally affected by this.

Finally, due to the nature of the SBO data collection infrastructure, participants may be skewed towards people with lower concerns about security and privacy. Despite this, participants tended to self-report high on the SeBIS intentions scale [27], answers to which they were optionally asked to provide at the time of enrollment to the SBO. On average, 64% of the participants whose data we studied indicated a frequency of "sometimes" or higher for the extent that they: a) secure their devices, b) generate strong and varied passwords, c) demonstrate proactive awareness of security issues, and d) update their computer software. Finally, several studies have successfully used the SBO for exploring various questions related to security and privacy [9, 16, 33, 37, 60].

7 DISCUSSION

Our analysis of the social media interactions of 38 participants and almost 200,000 posts revealed that security and privacy were not frequently discussed. Although most of the 38 participants interacted with *some* security and privacy posts, about a third of the participants accounted for a majority of the interactions with the security and privacy posts. Through our thematic analyses, we found that, more often than not, discussions about security and privacy revolved around jokes or sarcasm or merely mentioned a security- or privacy-related buzzword without discussing it further. Posts sometimes described security experiences with clear lessons but very rarely did posts about security and privacy include constructive, actionable advice. Similarly to how prior work has characterized the different ways in which constructive security advice is available in articles across the web [64, 68], our analysis characterized the security- and privacy-related content people may encounter and actively engage with on social media.

We next discuss the implications of our findings on increasing the dissemination of security and privacy information through social networks and what may constitute effective security and privacy education for the consumers of this information.

7.1 Disseminating security and privacy advice in social networks

We were surprised to find only 175 posts related to security and privacy out of the 194,081 social media posts that we examined. It is possible that security and privacy posts' popularity and spread is overshadowed by other topics, that users are not explicitly interacting with security- and privacy-related posts even if they encounter them on social media, or that there is not enough security content to spread. Prior work has also shown that security and privacy content does not reach all computer users equally [66, 67], which could also be reflected in social networks and responsible for the scarcity of security and privacy posts that we observed. Researchers have found that social influence—particularly within social media—can in experimental settings encourage people to take security-enhancing actions [20–22, 24, 29]. However, through measurement of historical social-media interactions and behavior, our findings and exploratory statistical analyses suggest that the positive influence of real social media posts and discussions on security behavior may not be as high or as likely as experimental results suggested.

An important step toward security and privacy content on social media impacting security behavior is to increase the number of interactions with security and privacy content. Future work could first evaluate whether increasing the volume of security and privacy content on social media would result in more interactions with such content. In general, social media has been highlighted as having an important role in information diffusion [7]. Future work could explore how to launch large-scale security and privacy publicity campaigns on social media, using lessons learned from previous cybersecurity awareness campaigns [1, 73]. These campaigns have typically been conducted through mass media such as radio and television. However, organizations that typically rely on these mass media outlets have the potential to reach a wider group of people through social media than just through more traditional outlets. For example, most news organizations maintain a Facebook page or Twitter account with large follower bases through which they share news and updates. Therefore, news media outlets may be able to play a significant role in furthering security and privacy campaigns on social media. Social media influencers have also had success with reaching many social media users with the goals of providing advice or influencing the purchase of products [40, 77]. They too could potentially help broaden the reach of security and privacy campaigns.

In addition to increasing the volume of security and privacy content on social media, encouraging users to interact with such content remains a separate problem [42]. Future work could take lessons

from previous work examining what makes posts in other topic domains popular [43, 50, 52]. The effectiveness of approaches to increasing interactions could be measured by the popularity of posts related to security and privacy, using metrics such as the number of likes and comments a post receives on Facebook and the equivalent metrics on other social networks.

Finally, posts about security and privacy also need to present advice in a way that incites the adoption of healthy security practices. In Sec. 5, we found that some posts constructively discussed security and privacy either by providing explicit advice or by conveying an anecdote in which certain specific security practices helped the post author. Future work could evaluate the effectiveness of these constructive posts on a large scale to identify recommendations for how security and privacy advice can be presented so as to encourage adoption of good practices.

7.2 Security education

Although our work suggests that people engage with security and privacy information on social media relatively infrequently compared to how often they engage with posts on other topics, the degree to which users *should* be educated about security and privacy remains an open question. Implementing security in systems has often been an afterthought in system design [36, 76] and making security features usable has been even less of a priority [34]. As a result, systems often place the burden of deciding to implement security on its users, requiring them to make complex decisions in the process [18, 23, 44]. Much of the research in the security and privacy community advocates for factoring security into system design from the start and relieving people of the burden of security decision-making as much as possible [2, 72].

Although reducing user burden is the ideal scenario, many current systems, tools, and services still require their users to be partly responsible for maintaining security, e.g., by creating and remembering strong passwords. For example, social engineering attacks (e.g., shoulder-surfing, phishing) are multi-faceted with no fix-all technical solution; avoiding them may require some savviness on the part of the user [15, 47] as well as awareness of the level of risk in various situations they encounter [19]. Therefore, security education is still important to inform people of the decisions they may need to make and steps to take to achieve the level of digital security they expect [6, 39, 49]. Amidst the complex security requirements of systems, several tools have been created to assist users. For example, password managers have been shown to be an effective way to create unique, strong passwords [31]. However, the use of password managers is still uncommon [59]. Our study suggests that computer users are unlikely to learn to improve their security behaviors through their general interactions with online content, at least in the context of social media. This suggests that more intentional security and privacy education and designing systems that further remove the burden of making security and privacy decisions are both necessary.

8 CONCLUSIONS

We analyzed real Facebook and Twitter logs and security behavior data from 38 participants over almost four years to empirically study *how often* and *how* security and privacy is discussed on social media, with the goal of understanding whether social media posts could be helpful in encouraging healthy security behavior. We identified a surprisingly low number (0.09%) of posts out of over $200k$ posts in the social media logs we analyzed to be about security and privacy. As our participant sample was slightly skewed toward technically savvy participants, it is likely that the fraction of security and privacy posts interacted with by the general population is even lower, which is supported by our analysis of the frequency of security and privacy posts on Twitter. We gained insight into the underlying nature of the security- and privacy-related posts on Facebook and Twitter by conducting a thematic analysis. We uncovered five major themes in how the social media posts we analyzed talked about security and privacy. Only one theme described posts that

spoke constructively about security and privacy, and these constructive posts made up only 20% of the already small pool of security and privacy posts we identified. Posts often spoke about security and privacy in passing, through jokes or sarcasm, or through anecdotes without conveying information about healthy or unhealthy security behavior. Though prior work suggests social media can be a highly effective platform for influencing healthy security habits, our findings suggest that there may not be enough constructive discussions on social media in practice such that people are educated and incentivized enough to make changes in their security habits, which exploratory statistical findings substantiated. Based on our findings, we discussed directions for future work toward achieving widespread constructive security awareness and behavior—both in terms of how to increase the dissemination of advice using social networks and how to improve the efficacy of security and privacy posts in encouraging healthier, more effective security behavior.

ACKNOWLEDGMENTS

This work was supported in part by the Carnegie Mellon University CyLab Security and Privacy Institute. Parts of the dataset we used were created through work supported by the National Security Agency under Award No. H9823018D0008. We would also like to thank Sarah Pearman and Jeremy Thomas for help with working with the SBO data.

REFERENCES

[1] 2020. Get Safe Online. (2020). https://www.getsafeonline.org/
[2] Anne Adams and Angela Sasse. 1999. Users are not the enemy. *Commun. ACM* (1999).
[3] Imtiaz Ahmad, Rosta Farzan, Apu Kapadia, and Adam J Lee. 2020. Tangible privacy: Towards user-centric sensor designs for bystander privacy. In *Computer Supported Cooperative Work and Social Computing (CSCW)*.
[4] Elham Al Qahtani, Mohamed Shehab, and Abrar Aljohani. 2018. The effectiveness of fear appeals in increasing smartphone locking behavior among Saudi Arabians. In *Symposium On Usable Privacy and Security (SOUPS)*.
[5] David Armstrong, Ann Gosling, John Weinman, and Theresa Marteau. 1997. The place of inter-rater reliability in qualitative research: An empirical study. *Sociology* (1997).
[6] Maria Bada, Angela Sasse, and Jason Nurse. 2015. Cyber Security Awareness Campaigns: Why do they fail to change behaviour?. In *International Conference on Cyber Security for Sustainable Society*.
[7] Eytan Bakshy, Itamar Rosenn, Cameron Marlow, and Lada Adamic. 2012. The role of social networks in information diffusion. In *International Conference on World Wide Web (WWW)*.
[8] Lucia Benaquisto and Lisa Given. 2008. *The SAGE encyclopedia of qualitative research methods*.
[9] Sruti Bhagavatula, Lujo Bauer, and Apu Kapadia. 2020. (How) Do people change their passwords after a breach?. In *IEEE Workshop on Technology and Consumer Protection*.
[10] Sruti Bhagavatula, Lujo Bauer, and Apu Kapadia. 2021. What breach? Measuring online awareness of security incidents by studying real-world browsing behavior. In *European Symposium on Usable Security (EuroUSEC)*.
[11] Ed Bott. 2013. Latest OS share data shows Windows still dominating in PCs. *ZDNet* (Apr 2013). https://www.zdnet.com/article/latest-os-share-data-shows-windows-still-dominating-in-pcs/
[12] Kevin J Boudreau, Nicola Lacetera, and Karim R Lakhani. 2011. Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Science* (2011).
[13] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* (2006).
[14] Jon Bruner. 2013. Tweets loud and quiet. (2013). https://www.oreilly.com/content/tweets-loud-and-quiet/
[15] Jan-Willem H Bullée, Lorena Montoya, Wolter Pieters, Marianne Junger, and Pieter H Hartel. 2015. The persuasion and security awareness experiment: reducing the success of social engineering attacks. *Journal of Experimental Criminology* (2015).
[16] Casey Canfield, Alex Davis, Baruch Fischhoff, Alain Forget, Sarah Pearman, and Jeremy Thomas. 2017. Replication: Challenges in using data logs to validate phishing detection ability metrics. In *Symposium On Usable Privacy and Security (SOUPS)*.
[17] Stephane Champely. 2018. *pwr: Basic Functions for Power Analysis*. https://CRAN.R-project.org/package=pwr R package version 1.2-2.
[18] Lorrie Faith Cranor and Simson Garfinkel. 2005. *Security and usability: designing secure systems that people can use*. O'Reilly Media, Inc.

[19] Sanchari Das, Jacob Abbott, Shakthidhar Gopavaram, Jim Blythe, and L Jean Camp. 2020. User-centered risk communication for safer browsing. In *Financial Cryptography and Data Security*.
[20] Sauvik Das, Tiffany Hyun-Jin Kim, Laura Dabbish, and Jason I Hong. 2014. The effect of social influence on security sensitivity. In *Symposium On Usable Privacy and Security (SOUPS)*.
[21] Sauvik Das, Adam DI Kramer, Laura Dabbish, and Jason I Hong. 2014. Increasing security sensitivity with social proof: A large-scale experimental confirmation. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*.
[22] Sauvik Das, Adam DI Kramer, Laura Dabbish, and Jason I Hong. 2015. The role of social influence in security feature adoption. In *Computer Supported Cooperative Work and Social Computing (CSCW)*.
[23] Rachna Dhamija and Lisa Dusseault. 2008. The seven flaws of identity management: Usability and security challenges. *IEEE Security & Privacy* (2008).
[24] Paul DiGioia and Paul Dourish. 2005. Social navigation as a model for usable security. In *Symposium On Usable Privacy and Security (SOUPS)*.
[25] William J Doll, TS Raghunathan, Jeen-Su Lim, and Yash P Gupta. 1995. A confirmatory factor analysis of the user information satisfaction instrument. *Information Systems Research* (1995).
[26] Paul Dunphy, Vasilis Vlachokyriakos, Anja Thieme, James Nicholson, John C McCarthy, and Patrick Olivier. 2015. Social media as a resource for understanding security experiences: A qualitative analysis of #password tweets. In *Symposium On Usable Privacy and Security (SOUPS)*.
[27] Serge Egelman and Eyal Peer. 2015. Scaling the security wall: Developing a security behavior intentions scale (SeBIS). In *ACM Conference on Human Factors in Computing Systems (CHI)*.
[28] Paul D Ellis. 2010. *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press.
[29] Pardis Emami Naeini, Martin Degeling, Lujo Bauer, Richard Chow, Lorrie Faith Cranor, Mohammad Reza Haghighat, and Heather Patterson. 2018. The influence of friends and experts on privacy decision making in iot scenarios. In *Computer Supported Cooperative Work and Social Computing (CSCW)*.
[30] Facebook. 2018. Facebook for Developers. (2018). https://developers.facebook.com/
[31] Michael Fagan, Yusuf Albayram, Mohammad Maifi Hasan Khan, and Ross Buck. 2017. An investigation into users' considerations towards using password managers. *Human-centric Computing and Information Sciences* (2017).
[32] Alain Forget, Saranga Komanduri, Alessandro Acquisti, Nicolas Christin, Lorrie Faith Cranor, and Rahul Telang. 2014. Building the security behavior observatory: An infrastructure for long-term monitoring of client machines. In *Symposium and Bootcamp on the Science of Security, HotSoS*.
[33] Alain Forget, Sarah Pearman, Jeremy Thomas, Alessandro Acquisti, Nicolas Christin, Lorrie Faith Cranor, Serge Egelman, Marian Harbach, and Rahul Telang. 2016. Do or do not, there is no try: User engagement may not improve security outcomes. In *Symposium On Usable Privacy and Security (SOUPS)*.
[34] Simson Garfinkel and Heather Richter Lipford. 2014. Usable security: History, themes, and challenges. *Synthesis Lectures on Information Security, Privacy, and Trust* (2014).
[35] Shirley Gaw, Edward W Felten, and Patricia Fernandez-Kelly. 2006. Secrecy, flagging, and paranoia: adoption criteria in encrypted email. In *Proceedings of the CHI Conference on Human Factors in Computing Systems, CHI*.
[36] Peter Gutmann and Ian Grigg. 2005. Security usability. *IEEE Security & Privacy* (2005).
[37] Hana Habib, Jessica Colnago, Vidya Gopalakrishnan, Sarah Pearman, Jeremy Thomas, Alessandro Acquisti, Nicolas Christin, and Lorrie Faith Cranor. 2018. Away from prying eyes: Analyzing usage and understanding of private browsing. In *Symposium On Usable Privacy and Security (SOUPS)*.
[38] Suraya Hamid, Mohamad Taha Ijab, Hidayah Sulaiman, Rina Md Anwar, and Azah Anir Norman. 2017. Social media for environmental sustainability awareness in higher education. *International Journal of Sustainability in Higher Education* (2017).
[39] Bartlomiej Hanus and Yu Andy Wu. 2016. Impact of users' security awareness on desktop security behavior: A protection motivation theory perspective. *Information Systems Management* (2016).
[40] Paul Harrigan, Timothy M Daly, Kristof Coussement, Julie A Lee, Geoffrey N Soutar, and Uwana Evers. 2021. Identifying influencers on social media. *International Journal of Information Management* (2021).
[41] Cormac Herley. 2009. So long, and no thanks for the externalities: The rational rejection of security advice by users. In *Workshop on New Security Paradigms*.
[42] Cormac Herley. 2014. More is not the answer. *IEEE Security & Privacy* (2014).
[43] Liangjie Hong, Ovidiu Dan, and Brian D Davison. 2011. Predicting popular messages in Twitter. In *International conference on World Wide Web (WWW)*.
[44] Johannes Kaiser and Martin Reichenbach. 2002. Evaluating security tools towards usable security: A usability taxonomy for the evaluation of security tools based on a categorization of user errors. In *IFIP World Computer Congress*.
[45] Amandeep Kaur and HS Chahal. 2018. Role of social media in increasing environmental issue awareness. *Researchers World* (2018).

[46] Judy Kendall. 1999. Axial coding and the grounded theory controversy. *Western Journal of Nursing Research* (1999).
[47] Iacovos Kirlappos and Angela Sasse. 2011. Security education against phishing: A modest proposal for a major rethink. *IEEE Security & Privacy* (2011).
[48] Roger Koenker and Gilbert Bassett Jr. 1978. Regression quantiles. *Econometrica: Journal of the Econometric Society* (1978).
[49] Elmarie Kritzinger and SH Solms. 2010. Cyber security for home users: A new way of protection through awareness enforcement. *Computers & Security* (2010).
[50] Tal Laor. 2019. "Hello, is There Anybody Who Reads Me?" Radio programs and popular Facebook posts. *International Journal of Interactive Multimedia & Artificial Intelligence* (2019).
[51] Arunesh Mathur and Marshini Chetty. 2017. Impact of user characteristics on attitudes towards automatic mobile application updates. In *Symposium On Usable Privacy and Security (SOUPS)*.
[52] Masoud Mazloom, Robert Rietveld, Stevan Rudinac, Marcel Worring, and Willemijn Van Dolen. 2016. Multimodal popularity prediction of brand-related social media posts. In *ACM International Conference on Multimedia*.
[53] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice. In *Computer Supported Cooperative Work and Social Computing (CSCW)*.
[54] Julian M Mueller-Herbst, Michael A Xenos, Dietram A Scheufele, and Dominique Brossard. 2020. Saw it on Facebook: The role of social media in facilitating science issue awareness. *Social Media + Society* (2020).
[55] Stanley A Mulaik. 2009. *Foundations of factor analysis*. CRC Press.
[56] David Nettleton. 2014. *Commercial data mining: processing, analysis and modeling for predictive analytics projects*. Elsevier.
[57] James Nicholson, Lynne Coventry, and Pam Briggs. 2017. Can we fight social engineering attacks by social means? Assessing social salience as a means to improve phish detection. In *Symposium On Usable Privacy and Security (SOUPS)*.
[58] James Nicholson, Ben Morrison, Matt Dixon, Jack Holt, Lynne Coventry, and Jill McGlasson. 2021. Training and embedding cybersecurity guardians in older communities.. In *ACM Conference on Human Factors in Computing Systems (CHI)*.
[59] PasswordManager.com. 2020. 65% of people don't trust password managers despite 60% experiencing a data breach. (2020). https://www.passwordmanager.com/password-manager-trust-survey/
[60] Sarah Pearman, Jeremy Thomas, Pardis Emami Naeini, Hana Habib, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, Serge Egelman, and Alain Forget. 2017. Let's go in for a closer look: Observing passwords in their natural habitat. In *ACM SIGSAC Conference on Computer and Communications Security, CCS*.
[61] Karl Pearson. 1909. Determination of the coefficient of correlation. *Science* (1909).
[62] Andrew Perrin and Monica Anderson. 2019. Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. *Pew Research Center* (2019). https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/
[63] Kristopher J Preacher and Robert C MacCallum. 2003. Repairing Tom Swift's electric factor analysis machine. *Understanding statistics: Statistical issues in psychology, education, and the social sciences* (2003).
[64] Emilee Rader and Rick Wash. 2015. Identifying patterns in informal sources of security information. *Journal of Cybersecurity* (2015).
[65] Emilee J Rader, Rick Wash, and Brandon Brooks. 2012. Stories as informal lessons about security. In *Symposium On Usable Privacy and Security (SOUPS)*.
[66] Elissa M. Redmiles, Sean Kross, and Michelle L. Mazurek. 2016. How I learned to be secure: A census-representative survey of security advice sources and behavior. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*.
[67] Elissa M Redmiles, Amelia R Malone, and Michelle L Mazurek. 2016. I think they're trying to tell me something: Advice sources and selection for digital security. In *IEEE Symposium on Security and Privacy*.
[68] Elissa M Redmiles, Noel Warford, Amritha Jayanti, Aravind Koneru, Sean Kross, Miraida Morales, Rock Stevens, and Michelle L Mazurek. 2020. A comprehensive quality evaluation of security and privacy advice on the web. In *USENIX Security Symposium*.
[69] Chris Rose. 2011. The security implications of ubiquitous social media. *International Journal of Management & Information Systems* (2011).
[70] Camille Ryan. 2018. Computer and Internet Use in the United States: 2016. (2018). https://www.census.gov/library/publications/2018/acs/acs-39.html#:~:text=In2016,theAmericanCommunity,socializing,andaccessinghealthcare.
[71] Koustuv Saha, John Torous, Sindhu Kiranmai Ernala, Conor Rizuto, Amanda Stafford, and Munmun De Choudhury. 2019. A computational study of mental health awareness campaigns on social media. *Translational Behavioral Medicine* (2019).

[72] Angela Sasse and Ivan Flechais. 2005. Usable security: Why do we need it? How do we get it? *Security and usability: Designing secure systems that people can use* (2005).
[73] Matthew W Savage, Sarah E Jones, Jenna E Reno, and Shari Veil. 2017. A case study: Targeting the Stop. Think. Connect. cybersecurity campaign to university campuses. In *Oxford Research Encyclopedia of Communication.*
[74] Johanna Schönrock-Adema, Marjolein Heijne-Penninga, Elisabeth A van Hell, and Janke Cohen-Schotanus. 2009. Necessary steps in factor analysis: enhancing validation studies of educational instruments. The PHEEM applied to clerks as an example. *Medical Teacher* (2009).
[75] Mark Shevlin and Jeremy NV Miles. 1998. Effects of sample size, model specification and factor loadings on the GFI in confirmatory factor analysis. *Personality and Individual Differences* (1998).
[76] Curtis Steward Jr, Luay A Wahsheh, Aftab Ahmad, Jonathan M Graham, Cheryl V Hinds, Aurelia T Williams, and Sandra J DeLoatch. 2012. Software security: The dangerous afterthought. In *International Conference on Information Technology-New Generations.*
[77] Karthik Subbian, Dhruv Sharma, Zhen Wen, and Jaideep Srivastava. 2013. Social capital: The power of influencers in networks. In *International Conference on Autonomous Agents and Multi-Agent Systems.*
[78] Lynn Tsai, Primal Wijesekera, Joel Reardon, Irwin Reyes, Serge Egelman, David Wagner, Nathan Good, and Jung-Wei Chen. 2017. Turtle Guard: Helping Android users apply contextual privacy preferences. In *Symposium On Usable Privacy and Security (SOUPS).*
[79] Twitter. 2018. Use Cases, Tutorials, & Documentation | Twitter Developer. (2018). https://developer.twitter.com/en
[80] VirusTotal. 2018. VirusTotal Developer Hub. *VirusTotal* (2018). https://www.virustotal.com/
[81] Rick Wash, Emilee Rader, Kami Vaniea, and Michelle Rizor. 2014. Out of the loop: How automated software updates cause unintended security consequences. In *Symposium On Usable Privacy and Security (SOUPS).*
[82] Yinglin Wu, Ling Xie, Shiang-Lin Huang, Ping Li, Zengwei Yuan, and Wenhua Liu. 2018. Using social media to strengthen public awareness of wildlife conservation. *Ocean & Coastal Management* (2018).
[83] Xinyue Ye, Bo Zhao, Thien Huu Nguyen, and Shaohua Wang. 2019. Social media and social awareness. *Manual of Digital Earth* (2019).
[84] Yixin Zou, Abraham H. Mhaidli, Austin McCall, and Florian Schaub. 2018. "I've got nothing to lose": Consumers' risk perceptions and protective actions after the Equifax data breach. In *Symposium On Usable Privacy and Security (SOUPS).*

A APPENDIX

A.1 Datasets we study

We use the following datasets in the paper.

A.1.1 Facebook and Twitter data: This dataset contains information about every post or tweet participants interacted with and the public pages or Twitter users that the participants follow. Each Facebook interaction contains the timestamp of the interaction, the type of interaction (i.e., like, comment, save, or create/share), information about the posters (i.e., a post has multiple posters if it was re-shared), and the post text. Each Twitter interaction contains the timestamp of the original post interacted with, the type of interaction (i.e., like, reply, or create/share), information about the posters (i.e., a post has multiple posters if it was re-shared), and the post text.

Browsing history: This dataset contains participants' browsing history over Google Chrome, Mozilla Firefox, and Internet Explorer. The dataset contains information about every URL visited in the web browser, along with page titles and timestamps.

Password data: This dataset includes information about every entry made into a password field in a web page, as determined by a browser extension, including a salted one-way hash of the password and the URL of the form in which the password was submitted. We filtered this dataset to exclude passwords used during failed login attempts or entered by a user other than the main computer user by replicating the filtering process used by prior work that examined passwords collected through the SBO [9].

Installed software update history: This dataset contains information about all events related to installed software and software updates on participants' computers along with Windows updates.

interaction_type	*# of posts*
liked	118,343
commented	47,739
saved	1,098
created/shared	26,901

Table 5. Number of Facebook posts in our dataset that correspond to each type of interaction described in Sec. 3.2.

interaction_type	*# of posts*
liked	3,498
commented	389
shared/retweeted	2,996

Table 6. Number of Twitter posts in our dataset that correspond to each type of interaction described in Sec. 3.2.

Data about one event includes information about whether it corresponds to an update or new software installed, the version of the software, and the timestamp for when the software was installed or update was executed.

File system data: This dataset contains information about all the files present on participants' filesystems. Specifically, data about one file includes the file name and path, the hash of the file in both MD5 and SHA1, and the timestamp for when the file was created.

Operating system history: This dataset contains the history of the different Windows operating system versions (e.g., "Vista", "7", "8", "10") participants have used throughout the duration of the SBO. Data about one version includes the operating system version and the timestamp at which that specific version was installed.

Software update settings history: This dataset consists of the Windows software update settings participants set for their computers at various points in time. The settings are a combination of four different preferences: (1) whether Microsoft updates are enabled; (2) whether recommended updates are enabled; (3) whether the update service is enabled; and (4) when notifications are desired. Data about one settings combination at a point in time contains the answers to the above four questions as well as the timestamp this particular settings combination was set.

WiFi connection history: This dataset contains information about all the WiFi networks participants connected to from their SBO-instrumented computer. Data about one instance of connecting to a WiFi network includes the name of the WiFi profile, whether the network required authentication, the type of encryption used by that network, and the type of shared key authentication used.

Browser content settings: This dataset contains information about what content participants allow various websites to use within their browser, i.e., microphone, location, or camera permissions. These permissions may be set or modified for a website at various points in time. Data about one event corresponding to setting or modifying permissions for a website includes which of the above permissions the participants granted to that website as well as the timestamp for when those particular permissions were set.

A.2 Breakdown of interactions for Facebook and Twitter posts

Tables 5 and 6 show the numbers of posts that correspond to each type of interaction.

A.3 Regular expressions for identifying posts

For identifying posts related to security and privacy, technology, and data breaches as described in Sec. 4, we started with an initial list of regular expressions (regexes). After our iterative algorithm, we ended up with a final list of regexes for each category. Table 7 shows the regexes in both the initial and final lists.

Security and privacy	Technology		Data breaches
password	matlab	password	breach
social security number	airport.*security	software development	captcha
security camera	net.*neutrality	web design	data.*hack
key.*security	gprs	data.analy	data.*stolen
security.*account	[0-9]mb	data.economy	data.*compromise
security question	Java	coinbase	data.*breach
phishing	3g(\W|$)	JSFoo	
cybersecurity	computer	uptime	
cyber.security	Android [0-9]	plugin	
de-verify	internet.*service	motherboard	
security.*protocol	apple.*ios	autonomous	
security	ios.*apple	browser.*code	
privacy	ios.*android	frontend	
hack	android.*ios	webapp	
bitcoin	github	jschannel	
net.*neutral	wireless	\.js	
secure	browser extension	3d model	
jailbreak	api	apple.*invent	
jail.*break	bitcoin	chat.*dm	
comp.*virus	Nokia	Dell.*laptop	
	smart device	screens.*digital	
	hyperloop	game.*app	
	browser.*bug	(\W)Siri(\W)	
	prototype	programming	
	smartphone	matlab	
	jquery	programmer	
	callback	sensor	
	app.*service	robot	
	js.*app	algorithm	
	lg.*screen	python	
	c\+\+	arduino	
	ux.*ui	autolab	
	ui.*ux	package	
	js_channel	linux	
	node.*js		
	debug		

Table 7. Lists of regexes used to flag Facebook or Twitter posts within three categories. Initial regexes are in black while the regexes in green were added via the iterative process.

A.4 Statistical relationships between social media interactions and security behavior

We studied the relationship between people's interactions with security- and privacy-related posts on social media and their measured security behavior. To study this, we used behaviors describing interactions with social media content on Facebook or Twitter and participant demographics (age, gender, whether they're a student, and whether they know a programming language) as input variables to statistical analyses. The outcomes we studied were behaviors indicating better or worse security practices. For this analysis, we analyzed the data of the 38 participants who gave us social media data along with eight additional participants who we consider non-social media users for comparison. We identified these eight participants by examining all their webpage visits (as

recorded by the SBO) and determining that they did not visit webpages on the Facebook or Twitter domains since 2016.

The behaviors we were interested in modeling typically weren't directly part of the raw data we collected: to compute them, we processed the SBO system-level and social media data such that each participant was associated with all events pertaining to them from each dataset described in App. A.1. We distilled and augmented this data into features that captured participants' interactions with content on social media and measures of their security behavior.

For each participant, we created separate features for Facebook and Twitter since the way content is shared on each platform differs. These features spanned interactions with the following three topics: security and privacy, technology, and data breaches. We identified the security and privacy posts in Sec. 4 and posts in the second two topics following a similar process with different initial lists for each category (see Table 7). We designated three features to describe the total number of interactions with posts in each category on Facebook and three more to describe this number for Twitter. Prior work has found that the source of security and privacy advice plays a role in people's inclination to act upon it, often heeding the advice of experts or influential people, or friends and family [22, 29, 67, 84]. Therefore, we describe interactions with posts along dimensions related to the source of a social media post such as: the type of poster (friend or page), the familiarity of the poster (the number of times the user previously interacted with that poster), and the poster's popularity (e.g., number of followers). Common to both platforms, for posts in each category, other features describe: the number of posts that were consumed versus shared by the participant; the number of posts where the participant was the sole author of the post they were interacting with; and the average familiarity of the posters for each post participants interacted with. Features specific to Facebook included: the average number of followers of the public page posters; the number of posts either originally posted or re-shared by a regular Facebook user (likely a friend); and the number of posts either originally posted or re-shared by a public page. Features specific to Twitter included: the average number of followers of all the posters of posts interacted with; the average number of favorites of all posts interacted with; and the average number of retweets of all posts interacted with. A list of all features and their descriptions can be found in Table 8. If a participant did not have either a Facebook or Twitter account (including those who had no social media account or visits), their corresponding feature values were set to 0.

Each participant was also associated with outcomes representing their security behavior (as introduced in Sec. 3) which were inspired by or built upon previous work [9, 16, 33, 60, 81]. Features related to installed software and updates included: the number of times participants updated software; the number of times participants computers underwent security updates; and the number of antivirus software installed. Features related to operating system (OS) updates described the number of OS updates and the number of times the OS update settings were modified. Other features included the average strength across each domain's latest password, the number of open Wi-Fi networks participants connected to, the number of websites to which participants granted either microphone, camera, or location permissions, and the number of files in participants' filesystems flagged by VirusTotal [80] as malicious. A list of all security behavior indicators along with their descriptions can be found in Table 9.

In total, participants were represented by 44 features related to social media including interactions with security and privacy, technical, and breach content on Facebook and Twitter and 25 indicators related to security behavior. To reduce the number of features to analyze, we grouped highly correlated features together through factor analysis [55] on the social media features and the security behavior features separately. As a result, 44 social media features were reduced to six factors and 25 security behavior indicators were reduced to two factors as described in Sec. A.5.

feature_set	*feature_name*	*description*
Facebook	has_fb	Whether the participant had a Facebook account
	fb_num_{category}	Number of posts a participant interacted with that fell into {category}
	fb_num_{category}_consumed	Number of posts in {category} wherein the interaction involved consuming content
	fb_num_{category}_by_friend	Number of posts in {category} wherein at least one of the posters involved in the post was a friend
	fb_num_{category}_by_page	Number of posts in {category} wherein at least one of the posters involved in the post was a public page
	fb_num_{category}_was_poster	Number of posts in {category} wherein the Facebook user themself was the sole author of the post
	fb_{category}_pages_avg_popularity	Average number of followers of the public pages involved in posts in {category}
	fb_{category}_posters_avg_familiarity	Average familiarity of each poster involved in all posts in {category}
Twitter	has_twitter	Whether the participant had a Twitter account
	twitter_num_{category}	Number of posts a participant interacted with that fell into {category}
	twitter_num_{category}_consumed	Number of posts in {category} wherein the interaction involved consuming content
	twitter_num_{category}_was_poster	Number of posts in {category} wherein the Twitter user themself was the sole author of the post
	twitter_{category}_posters_avg_popularity	Average number of followers of the posters involved in posts in {category}
	twitter_{category}_posts_avg_favorites	Average number of favorites on all posts in {category}
	twitter_{category}_posts_avg_retweets	Average number of favorites on all posts in {category}
	twitter_{category}_posts_avg_familiarity	Average familiarity of each poster involved in all posts in {category}

Table 8. Features related to social media consumption describe posts in each of the following categories: sec_priv, tech, and breach. Features with "{category}" are repeated for each of the three categories.

We then computed statistical relationships between the factors describing social media consumption of content in the three categories and the factors describing the various security behaviors.

A.5 Results

A.5.1 Factor analysis. After collapsing highly correlated social media features into one feature and transforming the values for each feature into its Z-score, factor analysis yielded eight total factors based on a scree plot showing that eight factors had eigenvalues above one [74]. We considered six of those factors that had at least two features with a factor loading greater than 0.7 [25, 63, 75]. The remaining two factors did not have a sufficient number of factor loadings greater than 0.7 and were therefore, excluded from consideration. Based on the features with loadings above 0.7 in each factor, the factors described the number of: (1) interactions overall with technical posts on Twitter and those made specifically by familiar posters; (2) interactions with security- and privacy-related posts on Facebook made by familiar posters and the number of technical posts the user made on Twitter; (3) interactions with security- and privacy-related posts and breach-related posts on Facebook; (4) interactions overall with technical posts on Facebook and specifically, the number of technical posts the user themselves made; (5) security- and privacy-related posts and breach-related posts made by the user themselves; and (6) interactions with technical and security- and privacy-related posts made by familiar posters on Facebook.

feature_set	*feature_name*	*description*
Browsing data	browsing_num_sp_related_pages	Number of visits to webpages related to privacy policies or security settings
	browsing_num_ sp_errors	Number of visits to webpages that signaled a privacy or security error
	browsing_num_sp_ignored_errors	Number of times the above security or privacy errors were ignored by detecting if the participant continued to the page after seeing the error
	browsing_num_vt_domain_ip	Number of visits to webpages on domains flagged by VirusTotal's domain report and IP address report APIs
	browsing_num_vt_domain_as_url	Number of visits to webpages on domains flagged by VirusTotal's URL report API
	browsing_num_vt_url	Number of visits to webpages flagged by VirusTotal's URL report API
	browsing_num_gsb	Number of visits to webpages flagged by Google Safe Browsing
	browsing_num_uncommon_tlds	Number of visits to webpages under uncommon Top Level Domains (TLD) [a]
	browsing_num_private	Number of visits to webpages in private browsing mode
	browsing_avg_links_in_count	For the webpages visited, the average number of links to each page from other webpages as reported by AWIS
	browsing_avg_website_ranks	For the webpages visited, the average Alexa rank as reported by AWIS
	browsing_num_no_links_in_count	The number of webpages visited that did not have any links to it from other webpages as reported by AWIS
	browsing_no_rank	The number of webpages visited that did not have an Alexa rank as reported by AWIS
Updates	software_num_updates	The number of times installed software was updated
	software_num_antivirus	The number of antivirus programs installed on a participant's operating system [b]
	software_num_sec_updates	The number of Windows security updates executed
OS	os_num_updated	Number of times the operating system version was updated
	os_num_updatesettings_changed	Number of times the operating system update settings were modified
Filesystem	fs_num_vt	The number of files on the filesystem whose hash was flagged by VirusTotal's file report API
Passwords	pwds_avg_strength	The average strength of the latest passwords being used on each domain
WiFi profiles	wifi_num_open	The number of times a participant connected with an open WiFi network that did not require authentication
Browser content	content_num_camera	Number of URLs for which the participant granted camera access
	content_num_location	Number of URLs for which the participant granted location access
	content_num_microphone	Number of URLs for which the participant granted microphone access
	content_num_all_allowed	Number of URLs for which the participant granted camera, location, and microphone access

[a]Common TLDs were determined from https://www.lifewire.com/most-common-tlds-internet-domain-extensions-817511.
[b]The list of antivirus software we checked for can be found at https://support.microsoft.com/en-us/help/18900/consumer-antivirus-software-providers-for-windows#avtabs=win7.

Table 9. Security behavior indicators across browsing and system-level datasets.

Applying the same criteria as above, factor analysis of security behavior indicators yielded eight factors of which we considered two. Again, based on the features in each factor with loadings above 0.7, the factors described: (1) the number of visits to websites flagged by VirusTotal or GSB and the number of links to them from other websites; and (2) the frequency of visits to URLs with uncommon TLDs and the number of antivirus software installed on participants' operating systems.

A.5.2 Analyzing relationships. We first conducted Pearson's pairwise correlation tests [61] between the six social media factors and each of the two security behavior factors and found that no pair exhibited a strong correlation (i.e., above 0.7 [56]). Considering the raw features before factor analysis, we also conducted pairwise correlations between each original social media and demographic feature against each security behavior and observed similar results.

Next, due to the size of our sample and the distribution of the data about participants, we constructed a non-parametric linear quantile regression model of the relationship between all six social media-related input factors and the four demographic features, and each of the two outcome factors [48]. We computed each of the two quantile regression models for the 25th, 50th (median), 75th, and 90th percentiles [12], resulting in a total of eight regression models. A power analysis with a desired power of 0.8, a p-value (α) of 0.05, and a medium effect size revealed that we required a sample size of 36, a criteria our dataset meets [4, 17, 28, 51, 57, 78]. We did not find any of the social media factors to be consistently correlated with either of the security behavior factors. In particular, computing the models for the 90th percentile revealed that being a student was correlated with a higher tendency to visit malicious URLs (first security behavior factor); and more interactions with security and privacy content (third social media factor) were correlated with having more antivirus software installed (second security behavior factor). Similarly, computing the model for the 50th percentile, more interactions with breach-related posts on Facebook (first social media factor) and technical posts on Twitter (second social media factor) were also correlated with having more antivirus software installed (second security behavior factor). However, no social media factor was correlated with any of the security behavior factors according to more than one model computed for one of the four percentiles, thus implying a lack of correlation [12].

feature_name	*baseline*	*coef.*	*std. err.*	*t*	*p*
(Intercept)		-0.979	0.524	-1.867	0.070
gender: male	female	-0.268	0.282	-0.950	0.349
knows_prog_lang: true	false	-0.175	0.281	-0.625	0.536
is_student: true	**false**	**0.672**	**0.321**	**2.091**	**0.044**
age		0.008	0.010	0.808	0.425
social_factor_1		-0.121	0.066	-1.841	0.074
social_factor_2		-0.098	0.069	-1.424	0.163
social_factor_3		<0.001	0.157	0.001	0.999
social_factor_4		-0.119	0.119	-0.997	0.326
social_factor_5		-0.060	0.085	-0.706	0.485
social_factor_6		0.030	0.085	0.353	0.726

Table 10. Quantile regression model for the 25th percentile studying the first security behavior factor.

feature_name	*baseline*	*coef.*	*std. err.*	*t*	*p*
(Intercept)		0.072	0.502	0.143	0.887
gender: male	female	0.049	0.236	0.210	0.835
knows_prog_lang: true	false	-0.098	0.327	-0.298	0.767
is_student: true	false	-0.064	0.374	-0.171	0.865
age		-0.011	0.010	-1.189	0.243
social_factor_1		0.109	0.065	1.677	0.102
social_factor_2		**0.258**	**0.064**	**4.059**	**<0.01**
social_factor_3		0.058	0.135	0.434	0.667
social_factor_4		0.007	0.108	0.064	0.950
social_factor_5		0.150	0.141	1.068	0.293
social_factor_6		-0.127	0.102	-1.239	0.223

Table 11. Quantile regression model for the 25th percentile studying the second security behavior factor.

feature_name	*baseline*	*coef.*	*std. err.*	*t*	*p*
(Intercept)		-0.7097	0.588	-1.207	0.235
gender: male	female	-0.1913	0.315	-0.608	0.547
knows_prog_lang: true	false	-0.7034	0.368	-1.911	0.064
is_student: true	**false**	**0.904**	**0.411**	**2.198**	**0.035**
age		0.012	0.011	1.039	0.306
social_factor_1		-0.100	0.106	-0.949	0.349
social_factor_2		-0.117	0.124	-0.942	0.353
social_factor_3		0.04‘	0.126	0.323	0.748
social_factor_4		-0.115	0.128	-0.901	0.374
social_factor_5		-0.250	0.126	-1.982	0.055
social_factor_6		0.070	0.119	0.592	0.557

Table 12. Quantile regression model for the 50th percentile studying the first security behavior factor.

Received April 2021; revised November 2021; accepted March 2022

feature_name	*baseline*	*coef.*	*std. err.*	*t*	*p*
(Intercept)		0.067	0.336	0.198	0.844
gender: male	female	0.185	0.180	1.026	0.312
knows_prog_lang: true	false	0.012	0.211	0.054	0.957
is_student: true	false	-0.365	0.235	-1.551	0.130
age		-0.001	0.006	-0.182	0.857
social_factor_1		**0.152**	**0.060**	**2.509**	**0.017**
social_factor_2		**0.251**	**0.071**	**3.523**	**0.001**
social_factor_3		0.038	0.072	0.524	0.604
social_factor_4		0.014	0.073	0.184	0.855
social_factor_5		0.067	0.072	0.925	0.361
social_factor_6		**-0.147**	**0.068**	**-2.160**	**0.038**

Table 13. Quantile regression model for the 50th percentile studying the second security behavior factor.

feature_name	*baseline*	*coef.*	*std. err.*	*t*	*p*
(Intercept)		0.374	0.539	0.695	0.491
gender: male	female	-0.162	0.337	-0.479	0.635
knows_prog_lang: true	false	-0.809	0.474	-1.708	0.097
is_student: true	false	0.368	0.492	0.748	0.460
age		0.008	0.010	0.730	0.470
social_factor_1		-0.029	0.122	-0.235	0.815
social_factor_2		-0.281	0.213	-1.322	0.195
social_factor_3		**0.518**	**0.114**	**4.551**	**<0.01**
social_factor_4		-0.103	0.125	-0.825	0.415
social_factor_5		-0.245	0.202	-1.217	0.232
social_factor_6		0.169	0.151	1.120	0.270

Table 14. Quantile regression model for the 75th percentile studying the first security behavior factor.

feature_name	*baseline*	*coef.*	*std. err.*	*t*	*p*
(Intercept)		0.005	0.431	0.012	0.991
gender: male	female	0.151	0.219	0.690	0.495
knows_prog_lang: true	false	0.235	0.280	0.837	0.408
is_student: true	false	-0.351	0.301	-1.165	0.252
age		0.005	0.008	0.639	0.527
social_factor_1		0.151	0.075	2.024	0.051
social_factor_2		0.191	0.131	1.451	0.156
social_factor_3		0.069	0.073	0.940	0.353
social_factor_4		-0.007	0.075	-0.091	0.928
social_factor_5		0.025	0.069	0.356	0.724
social_factor_6		-0.067	0.088	-0.763	0.451

Table 15. Quantile regression model for the 75th percentile studying the second security behavior factor.

feature_name	*baseline*	*coef.*	*std. err.*	*t*	*p*
(Intercept)		-0.262	0.885	-0.296	0.769
gender: male	female	0.071	0.474	0.150	0.881
knows_prog_lang: true	false	-1.227	0.865	-1.419	0.165
is_student: true	false	1.041	0.930	1.120	0.270
age		**0.036**	**0.018**	**2.053**	**0.048**
social_factor_1		0.176	0.188	0.938	0.354
social_factor_2		-0.534	0.461	-1.159	0.254
social_factor_3		**1.111**	**0.190**	**5.845**	**<0.01**
social_factor_4		0.260	0.187	1.390	0.173
social_factor_5		-0.422	0.440	-0.961	0.343
social_factor_6		0.207	0.254	0.815	0.420

Table 16. Quantile regression model for the 90th percentile studying the first security behavior factor.

feature_name	*baseline*	*coef.*	*std. err.*	*t*	*p*
(Intercept)		-1.509	1.721	-0.876	0.387
gender: male	female	0.301	0.643	0.469	0.642
knows_prog_lang: true	false	1.130	0.821	1.376	0.178
is_student: true	false	-0.234	0.922	-0.254	0.801
age		0.053	0.030	1.744	0.090
social_factor_1		-0.010	0.205	-0.050	0.960
social_factor_2		0.177	0.515	0.342	0.734
social_factor_3		-0.144	0.507	-0.284	0.778
social_factor_4		-0.262	0.199	-1.310	0.199
social_factor_5		0.265	0.209	1.269	0.213
social_factor_6		-0.231	0.303	-0.763	0.451

Table 17. Quantile regression model for the 90th percentile studying the second security behavior factor.