Designing Effective Phishing Interventions for the Elderly

Nicholas Doorlay
Department of Computer Science, California Polytechnic State University
San Luis Obispo, California

Abstract

Phishing, the most common type of internet scam, is becoming more and more prevalent each year. While there have been many attempts at training users in phishing prevention, these often occur in the context of universities or workplaces; this leaves particularly vulnerable populations, such as the elderly, without the training needed to properly protect themselves. This paper suggests a novel solution to phishing prevention designed to address the specific challenges that older adults face; this solution combines tooling and training in the form of a browser extension that nudges users toward more secure online behavior. Combined with optional supplemental training, we believe this solution has the potential to measurably reduce phishing susceptibility among the most vulnerable.

Introduction

Phishing, often attributed as the most common type of scam, is the fraudulent act of sending messages, often over email or text, pretending to be a reputable source in order to obtain personal information, such as credit card numbers [4]. In 2022 alone, Americans lost $39.5 billion to phishing scams; a staggering number, and one that's only expected to go up [16]. While phishing is a relatively simple type of scam, it poses one of the greatest threats to internet security due to its prevalence and efficacy; it only takes one lapse of concentration to fall for a phishing attack. It is commonly believed that older members of the population have an inherently higher susceptibility to Internet scams, particularly phishing, due to their age [3]. However, this

susceptibility seems to be due to different, mutable factors, rather than age. In their 2017 study, Gavett *et al.* compared susceptibility to phishing emails across different age groups; while older members of the population proved to be the most vulnerable, this study concluded that the largest predictor of susceptibility was not age, but rather lack of education or awareness of phishing, which is often correlated with increased age [9]. These findings suggest that simple educational interventions can be incredibly effective in reducing phishing vulnerability, particularly among older members of the population who may otherwise suffer from higher susceptibility.

While Gavett *et al.* found that older adults' susceptibility is due to their lack of awareness and education, other studies have shown that there are particular aspects of mitigating phishing attacks that seem to be particularly challenging for this population [9]. Sannd *et al.* found that the main vector for concern is the ability of older adults to differentiate between authentic emails and deceptive emails [13]. This study revealed that older adults place greater trust in an email displaying a well-known brand, than in their own reading of the same email to check for context-based grammar and syntax errors. Additionally, the study reached the conclusion that while poor or unusual grammar, syntax, spelling and any associated context within a message can be a useful indicator of likely illegal activity, older people remain less likely to recognize the indicators with enough consistency to expect a mitigation of ransomware and phishing attacks.

Similar to Australian *et al.*, Nino *et al.* found that within a phishing awareness survey given to older adults, the most common factor that the respondents failed to identify were links in the email that looked untrustworthy [4, 11]. This study compared respondents' reliance on technical factors (sender's email address, links present, URLs), design factors (grammar, professional design, spelling), and pretexting (email subject, emotional factors) for determining the legitimacy of emails. The findings suggest that, for the older participants, design factors were

relied on the most, while technical factors the least; this suggests that older adults are most vulnerable to attacks where the email includes URLs to malicious sites or an untrustworthy sender address, but the rest of the email appears normal.

As is evident by the above research on the topic, older members of the population have an alarmingly high level of susceptibility to phishing scams; contrary to popular belief, however, this susceptibility is due to factors that can be addressed, such as lack of awareness or education, rather than age. Research suggests that the best way to address this vulnerability could be through a combination of tooling and training [7]. Tooling could provide information about whether or not a particular website is potentially malicious, while training could provide end-users with the means to understand the tooling and know what to look for when navigating the internet in order to avoid falling victim to phishing scams. Tooling without training, or vice versa, leaves an incomplete solution that is often more harmful than it is good [7]. This paper outlines a potential solution that combines tooling and training to provide a holistic solution that can measurably reduce phishing scam susceptibility for elderly end-users.

The goal of our potential solution is to measurably reduce phishing scam susceptibility for elderly (defined as ages 65 and older) end-users. As such, our solution is informed by research in numerous areas, such as effective nudge design, effective phishing training design, state of the art phishing detection, and effective messaging and wording for discussing secure online behavior. The main contribution of this paper is a novel approach to phishing prevention for the elderly in the form of a browser extension that performs security nudges and provides the user simple, research-backed training. We believe this approach to be an effective strategy for preventing phishing as it lies at the intersection of tooling and training, an area of research that is promising yet unexplored.

Background & Related Work

A nudge, a term borrowed from economics, is an "intervention that maintains freedom of choice but steers people in a particular direction" [8]. In the context of online security, a nudge is a small intervention that pushes, but doesn't force, a user toward better security practices. When designed appropriately, nudges can be incredibly effective; appropriate design, in this context, includes two main criteria: 1) the user's choices should be respected, and 2) the direction, salience, and firmness of a nudge should be proportional to the user's benefit from the suggested course of action [1]. For example, a nudge that results in very little security benefit should be much less firm than a nudge that results in significantly more security benefit.

Phishing awareness training, generally speaking, is very effective at helping users identify phishing scams online [9]. Not all training, however, is created equal. In their 2020 study, Jampen *et al.* suggests that training tools should have progressive difficulty; if a user displays the correct anti-phishing behavior when tested, they should be moved onto more difficult tests of phishing awareness [10]. This progressive difficulty phishing training was found to be incredibly effective when compared to standard, constant difficultly training. Jampen *et al.* also warns of "security fatigue" in response to training, where a trainee receives so many warnings and popups that the training has the unintended effect of making them more susceptible to phishing scams [10].

Anti-phishing Phil, a product of Wombat Security, is a popular phishing awareness training game; as described on the Carnegie Mellon University website, Anti-phishing Phil is, "an interactive game that teaches users how to identify phishing URLs, where to look for cues in web browsers, and how to use search engines to find legitimate sites" [6]. In their 2007 study,

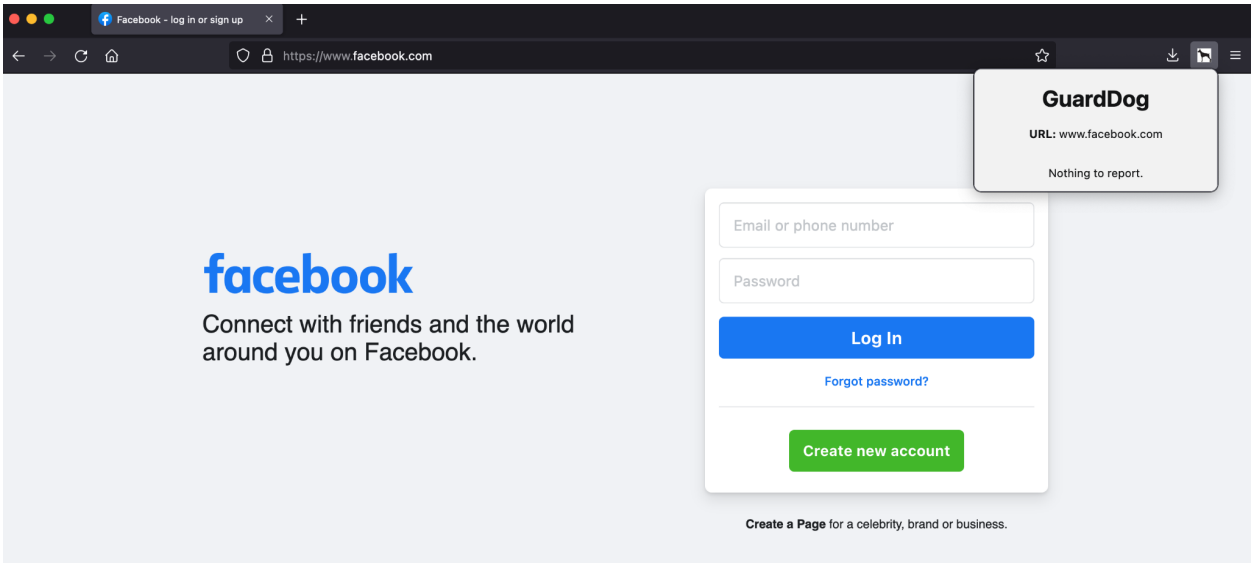
Sheng *et al.* performed an analysis of the efficacy of Anti-phishing Phil [14]. The study found Anti-Phishing Phil to be incredibly effective at reducing susceptibility to phishing attacks, attributing this efficacy to the fact that the game doesn't just bring awareness to phishing, but makes users more knowledgeable about techniques they can use to identify phishing web sites.

In their 2020 study, Adebowale *et al.* discusses development of an intelligent phishing detection scheme, using deep learning algorithms, that classifies websites as malicious with 93.28% accuracy [2]. The model, which was trained on over one million URLs and ten thousand images, analyzes urls, images, text, and frames on websites in order to classify them as either malicious or benign.

Protection Motivation Theory (PMT) is a framework for understanding responses to triggers in response to potential threats; it is commonly referenced when discussing effective messaging that encourages individuals to protect themselves from actions that might be harmful [15]. In their 2019 study, Bavel *et al.* showed that PMT can and should be used when designing nudges to improve online security behavior [5]. The study identifies two different kinds of messages – coping messages, which inform users that it is easy to protect themselves and indicate which steps to take, and fear appeal messages, which warn users of the potential harm that their behavior could leave them vulnerable to. Bavel *et al.* conducted an experiment comparing online behavior for a group exposed to only coping messages, a group exposed to only fear appeal messages, and a group exposed to a combination of the two [5]. Of these three, the study found the group exposed to only coping messages to demonstrate the most secure online behavior.

Guard Dog

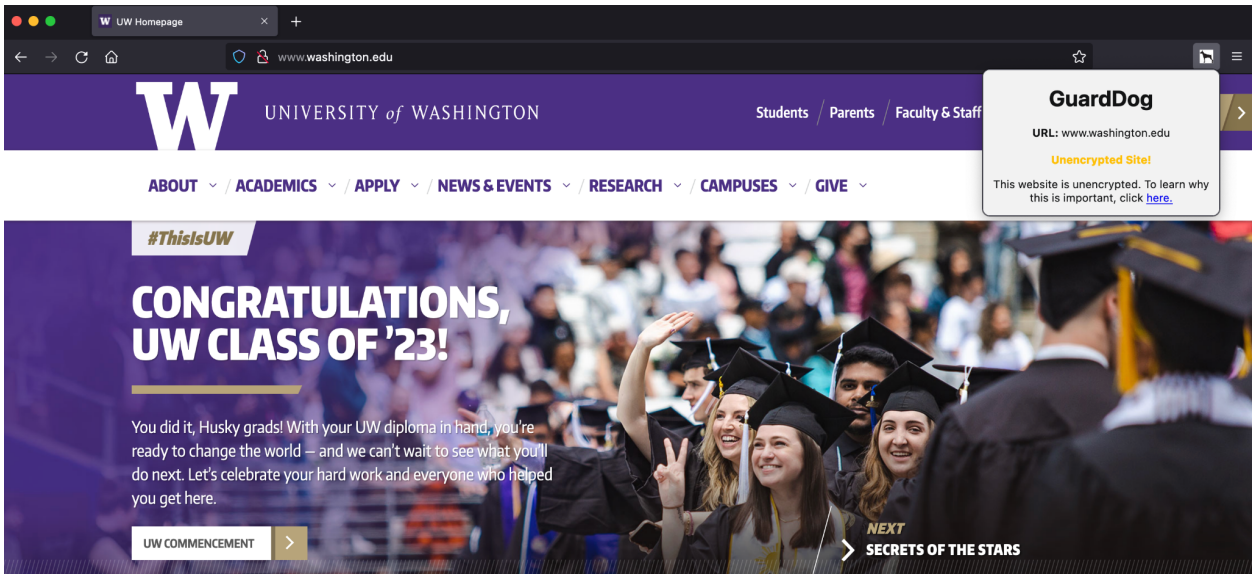
Inspired by Carnegie Mellon University's Privacy Bird, the working title of this paper's solution is Guard Dog [12]. Guard Dog is an open-source browser extension currently designed only for Mozilla Firefox. The code package for this browser extension includes HTML and CSS to represent the the browser extension's popup (figure 1), as well as JavaScript to define different behavior for different websites. For example, our implementation involves functions to parse the current tab's URL and extrapolate security-related data, such as whether or not the URL contains the substring 'https'. Based on the results of this parsing, Guard Dog will display different popup messages in different security scenarios, reducing the end-user's phishing susceptibility.



(Figure 1, the Guard Dog popup when performing a secure action)

As discussed in Fusaro *et al.*, nudges are an incredibly effective way to influence users toward better security practices [8]. In the context of our proposed solution, Guard Dog nudges a user when certain unsafe actions are performed, such as visiting a website that does not use the HTTPS internet protocol (figure 2). In order to adhere to the design standards outlined in Fusaro

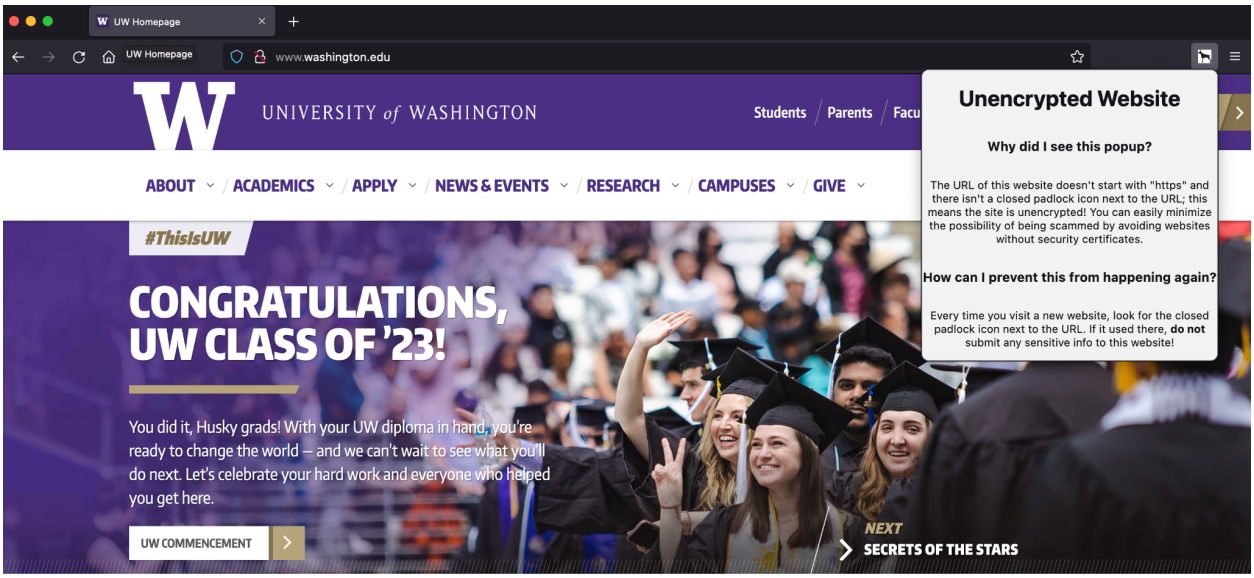
et al., these nudges must respect the user's choice if they decide to visit these websites despite the outlined risk; as such, Guard Dog's implementation allows the user to ignore the nudge if they choose to do so [8].



(Figure 2, example nudge for unsecured website)

By analyzing phishing awareness training like the popular Anti-Phishing Phil, Sheng *et al.* concluded that including short, optional trainings that teaches users how to identify phishing URLs on their own can be incredibly effective [14]. Drawing inspiration from this, the Guard Dog popups also include a link to learn more information about why the user received a warning; if clicked, the user receives an additional popup that provides supplemental training to prevent this unsecure action in the future (figure 3). The wording of this supplement training is incredibly important; in their 2019 study, Bavel *et al.* concluded that the terminology and tone used when discussing secure online behavior makes a significant impact on nudge efficacy [5]. These results influenced the messaging of this study's proposed solution to include only coping messages (instead of fear appeal messages), opting for phrases such as, "The URL of this website doesn't

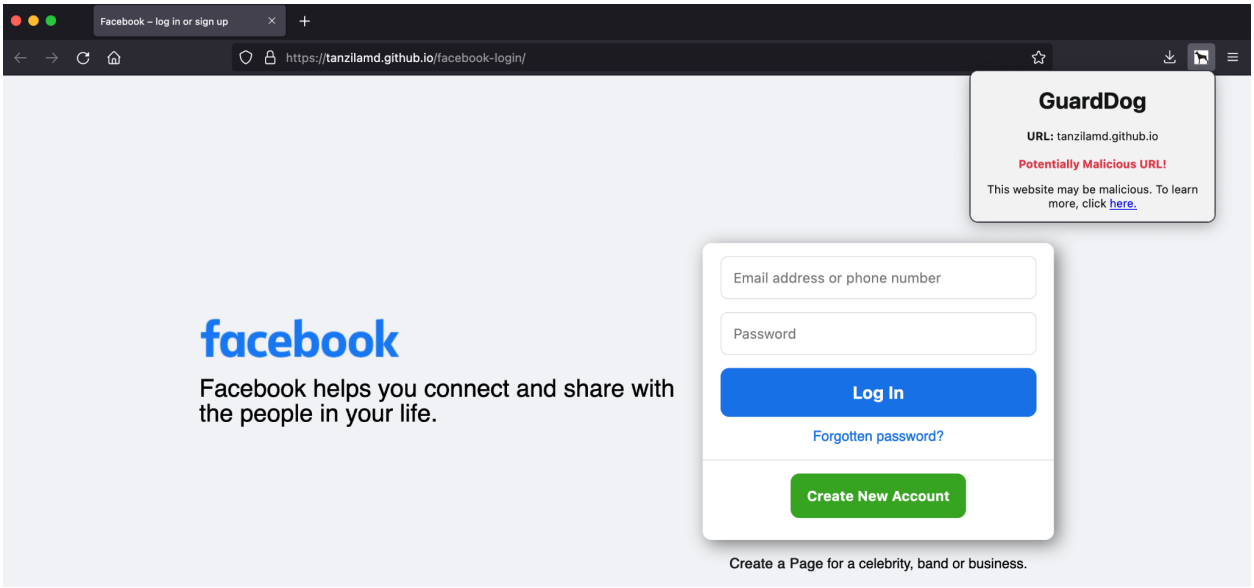
start with ‘https’ and there isn't a closed padlock icon next to the URL; this means the site is unencrypted! You can easily minimize the possibility of being scammed by avoiding websites without security certificates.”.



(Figure 3, supplemental training for unsecured website)

While much of this study is focused on effective security training, our goal was to create a holistic end product that includes effective tooling; as such, we attempted to incorporate state of the art phishing detection into our solution. Dhamija *et al.* discusses development of an intelligent phishing detection scheme which analyzes URLs and images in order to classify websites as malicious or benign with over 93% accuracy [7]. Inspired by this efficacy, this study intended to include a similar intelligent phishing detection scheme, implemented via a logistic regression classifier, trained on a dataset of both malicious and benign URLs, to provide users with nudges upon visiting a potentially malicious website; this classifier would be accessed via an API that is called upon each new website visit. This specific implementation proved too

cumbersome given this study's time constraint, so the implementation was changed to nudge based on a static denylist of known malicious domains gathered from reputable sources such as RiskAnalytics, the Canadian Institute for Cybersecurity, and the Cisco Talos Intelligence Group (figure 4). Similar to the nudges for unencrypted websites, these malicious websites provide the option to click through for supplemental training on why the user received the nudge.



(Figure 4, example nudge for website on malicious denylist)

Methodology

While the time constraints of this project did not allow for evaluation of our proposed solution's efficacy, we did design an approach that we would have taken given the time. Drawing on the phishing susceptible characteristics mentioned in Gavett *et al.*, we would recruit participants from a wide range of ages and familiarities with technology for our evaluation [9]. Users would be asked to classify websites as either malicious or benign, with one group having the Guard Dog browser extension installed and one control group without it installed.

With our current solution's implementation, the only two characteristics of a website that result in a nudge are 1) if a website does not implement HTTPS, and 2) if a website appears on our denylist. We would provide the user with websites that matched one of these characteristics but were benign, and websites that matched one of these characteristics and were malicious. We would then measure the percentage of websites that were correctly identified as malicious from both the control group and the group with Guard Dog installed to determine the efficacy of this study's proposed solution. This percentage of websites correctly identified, particularly among elderly users, would be our primary metric for measuring efficacy, similar to related studies such as Sheng *et al.* and Adebowale *et al.*[14, 2].

Conclusion

There were a number of features that we would have liked to implement given more time; the architecture of the browser extension was intentionally modular to provide easy future development. One feature that would potentially provide significant value is automatic malicious URL detection using machine learning. This would involve a machine learning model, accessible via an API, that is trained on malicious URLs and could analyze the current URL to determine if it is likely malicious. This is a feature we attempted to implement but proved to be too large an undertaking given the timeframe.

Another potential machine learning solution would be to analyze text and images on a webpage, not just URLs. Similar to Adebowale *et al.*, this solution would determine if a webpage or email is malicious even if the URL appears to be benign, providing even more protection for end-users who may not be able to accurately assess the safety of a webpage without the help of tooling [2].

There were additionally a number of challenges faced with the notification structure of the browser extension; it proved very difficult to generate popups that classified as nudges, visible to the user when necessary but not too intrusive. Given more time, we would fine-tune the notification structure in order to increase our solution's efficacy.

Finally, given more time, we would conduct an evaluation of our solution's efficacy in decreasing phishing susceptibility. As mentioned previously, this would be measured by the user's ability, with the help of the browser extension and training, to accurately classify a website as either malicious or benign. While phishing remains one of the most prominent internet scams, solutions such as our own have the potential to substantially decrease susceptibility, especially among our most vulnerable [4]. Our browser extension code can be viewed at https://github.com/doorlay/GuardDog.

References

[1] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, Yang Wang, and Shomir Wilson. Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online. In *ACM Computing Surveys, Volume 50 Issue 3*, 2017.

[2] Moruf Akin Adebowale, Khin T. Lwin, and M. A. Hossain. Intelligent Phishing Detection Scheme Using Deep Learning Algorithms. In *Journal of Enterprise Information Management*, 2020.

[3] Mohammed I Alwanain. Phishing Awareness and Elderly Users in Social Media. In *International Journal of Computer Science and Network Security*, 2020.

[4] Australian Competition and Consumer Commission. Scam Statistics. *https://www.scamwatch.gov.au/scam-statistics?scamid=all&date=2022*, 2022.

[5] René van Bavel, Nuria Rodríguez-Priego, José Vila, and Pam Briggs. Using protection motivation theory in the design of nudges to improve online security behavior. In *International Journal of Human-Computer Studies*, 2019.

[6] CyLab Usable Privacy and Security Laboratory. Anti-Phishing Phil. *https://cups.cs.cmu.edu/antiphishing_phil/*, 2008.

[7] Rachna Dhamija, J. D. Tygar, and Marti Hearst. Why Phishing Works. In *Conference on Human Factors in Computing Systems*, 2006.

[8] Roberta Fusaro and Julia Sperling-Magro. Much anew about 'nudging'. *https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/much-anew-about-nudging*, 2021.

[9] Brandon E. Gavett, Rui Zhao, Samantha E. John, Cara A. Bussell, Jennifer R. Roberts, and Chuan Yue. Phishing suspiciousness in older and younger adults: The role of executive functioning. *https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0171620*, 2017.

[10] Daniel Jampen, Gürkan Gür, Thomas Sutter and Bernhard Tellenbach. Don't click: towards an effective anti-phishing training. A comparative literature review. In *Human-centric Computing and Information Sciences*, 2020.

[11] Jean-Robert Nino, Gustav Enström, and Alan R. Davidson. Factors in Fraudulent Emails that Deceive Elderly People. In *International Conference on Human Aspects of IT for the Aged Population*, 2017.

[12] Privacy Bird. *http://www.privacybird.org/*.

[13] Premankit Sannd and David M. Cook. Older Adults and the Authenticity of Emails: Grammar, Syntax, and Compositional Indicators of Social Engineering in Ransomware and Phishing Attacks. In *International Conference on Information Processing (ICINPRO)*, 2018.

[14] Steve Sheng, Bryant Magnien, Ponnurangam Kumaraguru, Alessandro Acquisti, Lorrie Faith Cranor, Jason Hong, and Elizabeth Nunge. Anti-Phishing Phil: the design and evaluation of a game that teaches people not to fall for phish. In *Proceedings of the 3rd symposium on Usable privacy and security*, 2007.

[15] Ruth Shillair. Protection Motivation Theory. In *The International Encyclopedia of Media Psychology*, 2020.

[16] Truecaller Blog. Truecaller insights 2022 U.S. Spam & Scam Report. *https://www.truecaller.com/blog/insights/truecaller-insights-2022-us-spam-scam-report*, 2022.