Poster: Phishing on Facebook

Sovantharith Seng
Rochester Institute of Technology
Rochester, New York
ss2816@rit.edu

Mahdi Nasrullah Al-Ameen
Clemson University
Clemson, South Carolina
malamee@clemson.edu

Matthew Wright
Rochester Institute of Technology
Rochester, New York
matthew.wright@rit.edu

***Abstract*—Facebook, the largest social networking site (SNS) with over one billion users, has been woven into the everyday life of many people. Online attackers are increasingly employing phishing attacks on Facebook due to its wealth of personal information and users' lack of security knowledge, seeking to fool their victims by using fake or compromised accounts. These attacks are hard to recognize by the Facebook defensive system and users alike, and few studies have been done on how users interact with such attacks. This study aims to take the initial step in understanding the thought process of users and influences of different variables when they interact with their Facebook newsfeed.**

I. Introduction

Social networking sites (SNSes), especially Facebook, have become an integral part of life for many people. Despite the complex use and plethora of information on these networks, a large fraction of users are lacking in security knowledge and awareness about how to navigate SNSes securely [1]–[3]. On top of that, Some SNSes, Facebook in particular, have complicated systems of security and privacy settings due to their complex structure.

Phishing attacks exploit human errors in online navigation [1]. The attacker's goal is to either collect login credentials from the victim in order to gain access to their online accounts or have the victim visit a crafted malicious site with a drive-by download. According to the Phishing Activity Trends Report[1] by the Anti-Phishing Working Group, phishing attacks has increased to over 1.2 million attacks in 2016. That is a 65% increase compared to 2015. Recently, there is also an increase in phishing attacks on SNSes using fake or compromised accounts. In SNSes, attackers can improve their chance of being clicked by creating targeted attack using information shared on the platform or using a link shortener (e.g. bitly.com) or specialized obfuscation services to disguised their malicious destinations.[2] Although this requires more time and effort, it is generally more successful and harder to be detected by current defense systems [4] and users. With over one billion active users [5], a successful attack in Facebook is worth the additional effort.

To date, there has been little research into understanding the efficacy of attackers' strategies in carrying out phishing attacks over SNS, which is important for understanding how to improve SNS defense mechanisms and user awareness. Our proposed study aims to fill this gap. In particular, we will investigate the importance of different aspects of a post (phishing or otherwise) that influence the user's decision of whether or not to click the link. This paper briefly describes the design of the study in relation to prior work.

[1]https://docs.apwg.org/reports/apwg_trends_report_q4_2016.pdf
[2]https://apps.lazza.dk/facebook

II. Related Work

We now discuss prior studies on phishing on SNSes and users' vulnerabilities.

Phishing on Social Networking Sites: Dhamija et al. [1] showed a correlation between the success of the attacks and the low knowledge level of the users of the users as well as with the level of authenticity in the look and feel of the spoofed email and website. Chhabra et al. [6] discussed the rise of attacks using shortened URLs on Twitter. Shortened URL through third-party services such as bitly.com and owl.ly are widely used to reserve character space and provide memorable links for advertisement or personal use. However, attackers can use this service to misdirect their victims, fooling them by redirecting to a phishing website instead of the real one. According to the study, 89% of references on Twitter were reported to be inorganic (i.e. shared by automated bots instead of real users). Vishwanath [4] pointed out the lack of statistics provided by Facebook regarding phishing attacks and fake account statistics, which makes it hard for researchers to conduct formal studies on the platform. They reported that approximately 1 in 10 Facebook accounts is a fake or a duplicate account. He stated that SNS has become a very attractive attack vector because of its continuing success. According to their study, attacks on Facebook have an approximately 40% success rate, compared to a success rate of just 1% for traditional email phishing. His findings indicate that attackers typically either post malicious links on a newsfeed, mimicking something of interest to the victims, or personally contact the victims through a private message. Alam et al. [7] noted that the success of targeted phishing is correlated with the amount of information the attacker has. Therefore, if an attacker is a friend with the victim or uses a compromised account of a friend of the victim, they will have little difficulty in fooling the victims without getting noticed. Since SNS users expose a lot of personal information through the site, particularly to their connections, the high success rates reported by Vishwanath may be considered unsurprising.

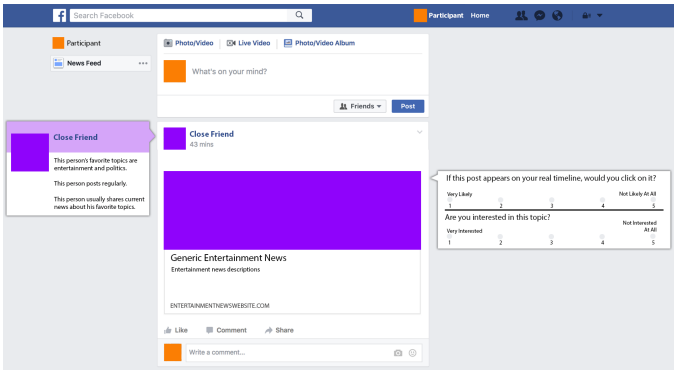

Fig. 1. Mock-up of the simulated interface

Users' Vulnerabilities: Dhamijia et al [1] pointed out to the users' lack of security knowledge. Second, the attacks are successful due to how Facebook is used. Joinson [8] found that some Facebook users gain gratification from the site by either social surfing, finding more information about other people, or expanding their social network. To find others and be found, users may fill out information on their profile and tailor their privacy settings to reach a wider audience. By doing so, these users are providing more information to the phishing attackers and exposing themselves as vulnerable targets. Lampe et al. [9] noted that there is a correlation between the completeness of profile details and the amount of online friends. Additionally, users who are receptive to new connections may also be vulnerable to accepting friend requests and messages from attackers posing as legitimate users. Furthermore, users with large amount of friends may be more vulnerable to interacting with unknown strangers or unaware that their friends' accounts have been compromised. Patil [10] conducted a study of fake accounts in SNS in 2012, in which they find that up to 40% of users would accept a fake account request. Boshmaf et al. [11] developed the Socialbot Network, a group of adaptive social bots that tricked up to 80% of Facebook users into accepting their friendship requests. I have not found any study, however, on whether users treat links from these fake accounts the same as those from accounts connected to them based on relationships that extend beyond Facebook. Furthermore, no studies I have seen examine whether and how users are looking for indicators of compromised accounts or fake posts.

III. Methodology

In this section, we briefly describe our planned method for this work in terms of study and user interface design.

Study Design: Our main goal of the proposed study is to identify the extent to which content type and relationship affect a user's decisions to click a post with link on Facebook. Since there are many possible variables that can affect the user's decision, we want to explore a small number of variables in this pilot study by using a simplified Facebook newsfeed simulation to determine if there is any indications that they have significant impacts. We intend to later conduct a larger study on MTurk with a more realistic Facebook newsfeed simulation.

In this pilot study, we will conduct a lab study with a think-aloud protocol. We have designed a simulated interface that resembles the Facebook newsfeed, shown in Fig. 1. Participants will be asked to interact with the simulated SNS interface, where they will respond to Likert-scale questions regarding their likelihood of clicking on a post shown on the newsfeed. In this case, a post may vary in terms of the type of contents (e.g., news, entertainment, and travel) and the relationship with the person who shared that post. In this way, the study will reveal how an attacker could exploit the content-type and relationship with users to make them clicking on a malicious like over social networking sites. Once the participants finish their interaction with the simulated SNS interface, they will complete another survey regarding their real-life behavior on social networking sites, like Facebook.

Interface Design: Although we intend to design a simulated SNS interface that resembles a the real Facebook newsfeed in a later study (Fig. 1), we face some trade-offs to serve the purpose of this study. To limit the amount of biases from other variables and the keep the scope of this study manageable, our interface will differ from the real-life interface in the following ways: i) Elements on the right-side of the newsfeed, including chat bar, news updates, and the advertisements will be removed to make space for Likert-scale questions that the participants will require to answer for each of the shown posts, ii) The profile picture of each Facebook friend will be presented by a solid color box to control for various biases [10], iii) The name of each Facebook friend will be presented in a generic way, like 'Father' or 'Close Friend', providing a simplified presentation of the relationship with the participant and thereby controlling for variation in particular relationships, iv) There will be no 'preview details (e.g., images)', or 'like', 'reaction' or 'comment' shown with any post. The preview details will also be replaced with generic sentences.

In the simulated newsfeed, the participant's name will be just 'Participant' with having a solid color box as the profile picture. They are required to click on a friend's profile picture to view more details about them, where a pop-up on the left of a post will show the details about the person. The participants will also have the ability to hover over a post to see the link's destination (i.e., URL). The phishing and non-phishing posts will both have the same appearance. However, phishing post will have the link's destination that is different from what is shown in the preview. Furthermore, the topic of the post will be a mismatch from the listed interest of the profile.

References

[1] R. Dhamija, J. D. Tygar, and M. Hearst, "Why phishing works," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '06. New York, NY, USA: ACM, 2006, pp. 581–590. [Online]. Available: http://doi.acm.org/10.1145/1124772.1124861

[2] M. Tsikerdekis and S. Zeadally, "Online deception in social media," *Commun. ACM*, vol. 57, no. 9, pp. 72–80, Sep. 2014. [Online]. Available: http://doi.acm.org/10.1145/2629612

[3] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, "Social phishing," *Commun. ACM*, vol. 50, no. 10, pp. 94–100, Oct. 2007. [Online]. Available: http://doi.acm.org/10.1145/1290958.1290968

[4] A. Vishwanath, "Habitual facebook use and its impact on getting deceived on social media," *Journal of Computer-Mediated Communication*, vol. 20, no. 1, pp. 83–98, 2015. [Online]. Available: http://dx.doi.org/10.1111/jcc4.12100

[5] Facebook, "Company info — facebook newsroom," Dec 2016. [Online]. Available: http://newsroom.fb.com/company-info/

[6] S. Chhabra, A. Aggarwal, F. Benevenuto, and P. Kumaraguru, "Phi.sh/$ocial: The phishing landscape through short urls," in *Proceedings of the 8th Annual Collaboration, Electronic Messaging, Anti-Abuse and Spam Conference*, ser. CEAS '11. New York, NY, USA: ACM, 2011, pp. 92–101. [Online]. Available: http://doi.acm.org/10.1145/2030376.2030387

[7] S. Alam and K. El-Khatib, "Phishing susceptibility detection through social media analytics," in *Proceedings of the 9th International Conference on Security of Information and Networks*, ser. SIN '16. New York, NY, USA: ACM, 2016, pp. 61–64. [Online]. Available: http://doi.acm.org/10.1145/2947626.2947637

[8] A. N. Joinson, "Looking at, looking up or keeping up with people?: Motives and use of facebook," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '08. New York, NY, USA: ACM, 2008, pp. 1027–1036. [Online]. Available: http://doi.acm.org/10.1145/1357054.1357213

[9] C. A. Lampe, N. Ellison, and C. Steinfield, "A familiar face(book): Profile elements as signals in an online social network," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '07. New York, NY, USA: ACM, 2007, pp. 435–444. [Online]. Available: http://doi.acm.org/10.1145/1240624.1240695

[10] S. Patil, "Will you be my friend?: Responses to friendship requests from strangers," in *Proceedings of the 2012 iConference*, ser. iConference '12. New York, NY, USA: ACM, 2012, pp. 634–635. [Online]. Available: http://doi.acm.org.ezproxy.rit.edu/10.1145/2132176.2132318

[11] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, "The Socialbot network: When bots socialize for fame and money," in *Proceedings of the 27th annual computer security applications conference*. ACM, 2011, pp. 93–102.