A Semi-Automated Security Advisory System to Resist Cyber-attack in Social Networks

Samar Muslah Albladi and George R S Weir

University of Strathclyde, Glasgow G1 1XH, UK
{samar.albladi; george.weir}@strath.ac.uk

Abstract. Social networking sites often witness various types of social engineering (SE) attacks. Yet, limited research has addressed the most severe types of social engineering in social networks (SNs). The present study investigates the extent to which people respond differently to different types of attack in a social network context and how we can segment users based on their vulnerability. In turn, this leads to the prospect of a personalised security advisory system. 316 participants have completed an online-questionnaire that includes a scenario-based experiment. The study result reveals that people respond to cyber-attacks differently based on their demographics. Furthermore, people's competence, social network experience, and their limited connections with strangers in social networks can decrease their likelihood of falling victim to some types of attacks more than others.

Keywords: Advisory System, Social Engineering, Social Networks.

1 Introduction

Individuals and organisations are becoming increasingly dependent on working with computers, accessing the World Wide Web and, more importantly, sharing data through virtual communication. This makes cyber-security one of today's greatest issues. Protecting people and organisations from being targeted by cybercriminals is becoming a priority for industry and academia [1]. This is due to the huge potential damage that could be associated with losing valuable data and documents in such attacks.

Previous research focuses on identifying factors that influence people's vulnerability to cyber-attack [2] as the human has been characterised as the weakest link in information security research. When investigating human behaviour toward online threats, it is important to focus on the interaction between the individual's attributes, their current context, and the message persuasion tactic [3]. Most previous studies that have considered persuasion tactics in social engineering exploits, have focused on phishing as the common type of cyber-attack while limited research has investigated other types, such as malware, or clickjacking. Figure 1 shows that 37% of participants in the current study fell victim to a phishing scam attack that asked them to validate their Facebook account using a phishing link, while only 28% fell victim to a phishing attack that asked them to register their information to enter a prize draw. Consequently, the present study

argues that people's vulnerabilities change depending upon the type of cyber-attack and our investigation addresses the human characteristics associated with victimisation for a range of cyber-attacks which, in turn, facilitates the design of a semi-automated security advisory system that relies on the idea of people segmentation and targeting. Segmentation, Targeting, and Positioning (STP) strategic approach is a well-known model that has been extensively applied to modern marketing research [4]. According to this model, there are three main processes to segment people in order to deliver them an effective and 'focused-to-need' messages. We have adopted this approach to design a security advisory system based on social networks users' characteristic and associated threat vulnerability.

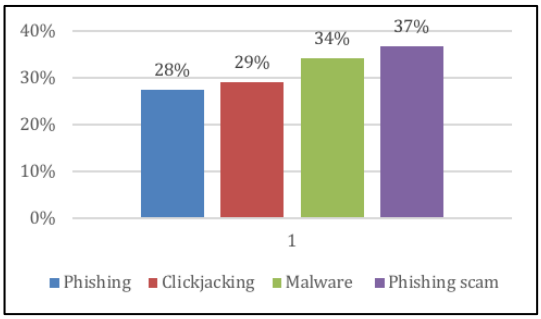

Fig. 1. Victim percentages

The material presented here is organised as follows. Section 2 provides a brief literature review. Section 3 describes the study methodology while Section 4 presents the results of the analysis. Discussion of the results is provided in Section 5. An outline approach to a semi-automated user advisory system is proposed in Section 6. Finally, Section 7 offers conclusions from the study.

2 Literature review

Criminals in social communication channels use advanced methods to access sensitive information to help them increase the success rate of their attacks and this so-called social engineering attacks. In this type of attack, the aim is often not to target systems, but rather individual's users or organizations. In order to protect against this type of cyber-attack, it is necessary to investigate and understand the reasons why people are not able to detect these attempts to penetrate their data and devices.

It is important to investigate the main entities that encapsulate and contribute to the success of social engineering-based attacks in order to understand why people get easily deceived by this kind of attack. Krombholz et al. [5] proposed a taxonomy of social engineering sophisticated attacks in the virtual communication networks. The taxonomy comprised three main entities that have been argued to form every social engineering attack, the operator of the attack, the type of the attack, and the attack channel. The attack can be originated by either a person which reflected a limited number of victims

such as spear phishing [6] or by a malicious software which usually targeted a considerable huge number of users such as the cross-site scripting attack in SN [7].

In the virtual environment of social networks, there has been limited research available to help explain why people are easily deceived by social engineering attacks. An investigation of people's social network habits and their relation to people's vulnerability to SN phishing attacks revealed that the willingness of raising the number of connected friends as well as maintaining frequent use of the network have a high impact on user behaviour [8]. Yet, another study [9] went further by investigating whether the impact of social network usage on people's likely victimisation differs among different social network platforms and found a statistically negative relationship between frequent usage of multipurpose dominant SN such as Facebook and victimization. This means that high frequency of using Facebook does not lead to an increase in the likelihood of victimisation. Furthermore, another study [10] found that connecting with a large number of profiles on Facebook would lead to a non-controllable online network which ultimately increases individuals' vulnerability. Perceptual-related factors have also been identified as affecting vulnerability to cyber-attacks, as a high-level of risk propensity could cause people to fall victim to cyber-attacks [9].

One of the proposed solutions to deal with enduring online threats is to understand the victim's background and examine their reaction by conducting a real attack, such as the case of sending phishing emails to a particular group of users [11], [12]. In contrast, due to ethical considerations, the majority of studies [13]–[15] used scenario-based experiments to examine people's vulnerabilities. Among many identified characteristics that are believed to predict potential victims [2], [16], demographics are the most controversial variables. Moreover, most of the earlier mentioned studies are focused on one type of attack although criminals have several ways to perform social engineering attacks. This has indicated the need for further investigation of people's vulnerability to different types of cyber-attack and the need to explore which groups of users are more vulnerable to specific types of cyber-attack in a social network context. Identifying the characteristics of most vulnerable individuals for a particular type of attack could help designing an advisory system to push awareness messages to vulnerable individuals. We expect that designing such security advisory system based upon observed user behaviour and characteristics could reduce people susceptibility to different types of cyber-attacks in social networks.

3 Methodology

In order to design our advisory system, we first collect the user data. An online questionnaire has been designed as an assessment tool to examine participants' perception and behaviour toward different threats in a social network context.

An invitation email was sent to faculty staff in two universities asking them to distribute the online-questionnaire among their students and staff. Participants were presented with the online-questionnaire which has 3 parts. The first part asked about demographics. The second part includes questions that measure study constructs such as the three scales used to measure the user's competence to deal with cyber-attacks [17].

The final part includes the scenario-based experiment as participants were presented with four images of Facebook posts, each post includes a type of cyber-attack such as phishing for sensitive information (Attack 1), clickjacking with an executable file (Attack 2), malware attack (Attack 3), and a phishing scam that impersonates a legitimate organization (Attack 4). These four cyber-attacks have been chosen from the most prominent cyber-attacks that occur in social networks [18]. Participants were asked to indicate their response to these attacks, as if they encountered them in their real accounts, by rating a number of statements such as "I would click on this button to read the file" using a 5-point Likert-scale from 1 "Strongly Disagree" to 5 "Strongly Agree".

After that, we examine whether the collected user data could help us designing a semi-automated advisory system that classifies participants into different vulnerability segments in order to target their needs by providing personalised awareness messages.

4 RESULTS

We have tested which group of people are vulnerable to each type of cyber-attack in the scenario-based experiment, based upon their rating response to the different statements. Table 1 describes the mean from the five-point Likert-scale and its corresponding vulnerability level.

Table 1. Description of the scale mean

Mean	Likert scale	Vulnerability Level
1.00-1.79	Strongly Disagree	Low vulnerable
1.80-2.59	Disagree	
2.60-3.39	Neither Agree nor Disagree	Moderately vulnerable
3.40-4.19	Agree	High Vulnerable
4.20-5.00	Strongly Agree	

Demographic differences

To examine whether user demographics have an impact on user susceptibility to social engineering victimization, every demographic variable has been tested individually to identify which group of people are more vulnerable to a certain type of attack. Female participants are found to be more vulnerable than male participants to all considered cyber-attacks. Figure 2 shows that among the four selected types of attack, the phishing scam that impersonates a Facebook technical support message is most successful among male and female participants with a mean of 1.92, 2.32 respectively.

Generally, younger adults are less vulnerable to cyber-attacks than older adults (as appears in Figure 3). Surprisingly, in the phishing that offers a prize as well as in the malware attack, the oldest group (45-55) was most vulnerable (phishing=2.60, malware=2.80) while the mid-aged group (35-44) was the least likely to respond to these kinds of attack (phishing=1.71, malware=1.64).

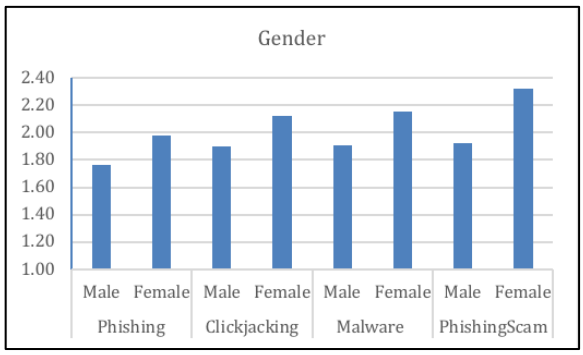

Fig. 2. Gender Comparisons of Vulnerability to SE

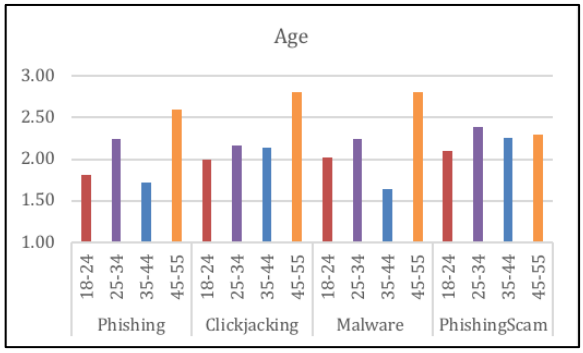

Fig. 3. Age Comparisons of Vulnerability to SE

The analysis of different groups with various education levels and their response to the four types of SE attacks revealed that master's degree holders are more vulnerable to clickjacking than to other types of cyber-attack (m=2.14). While high school and bachelor's degree holders are more vulnerable to the phishing scam that impersonates a legitimate SN provider (with a mean of 2.10, and 2.31 respectively).

Users with a technical education background were shown to be less vulnerable to cyber-attacks. In contrast, Business School students are more vulnerable to the phishing attack that offers a prize than other attacks, while Humanities and Arts students are more vulnerable to the malware attack. Medical and Science students are more vulnerable to the phishing scam that impersonates a Facebook technical support message.

Prevention factors

In order to investigate if user characteristics can prevent user's vulnerability to specific types of attacks, three factors have been chosen (user's competence, Social network experience, low connections with strangers) to consider whether their prevention effect is similar across the four types of attacks. Multiple regression tests have been conducted to test the impact of these three variables on preventing users from falling victim to cyber-attacks. These factors are proved to decrease people's vulnerability to the four considered cyber-attacks, when combined together in our study. This section will present the result of their impact on the four types of attacks as shown in Figure 4.

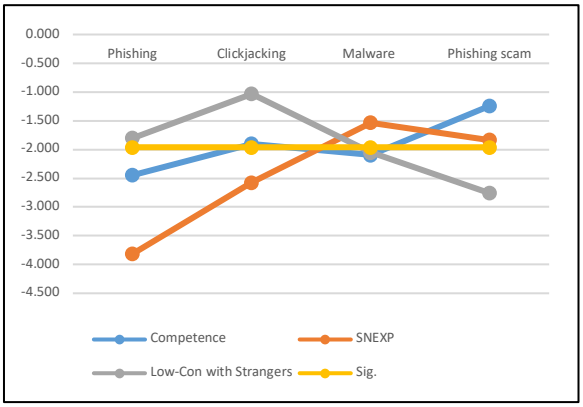

Fig. 4. Regression Analysis Results

User's competence: When analysing the impact of users' competence on decreasing users' susceptibility to social engineering victimization, the result in Figure 4 shows that measuring users' competence could identify less vulnerable individuals who can correctly detect phishing attack that offers a prize (t value=-2.447, P<0.05) and also detectors of malware attack (t value=-2.098, P<0.05). While competence could not prevent participants from falling victim to clickjacking and phishing scam attacks, as these relationships appear to be not significant (t value>-1.96).

Social Network experience: Regression analysis of the impact of social network experience on decreasing individuals' response to different kinds of cyber-attack indicated that among the four types of cyber-attacks, phishing attack that offers a prize (t value=-3.816, P<0.05) and clickjacking (t value=-2.573, P<0.05) are attacks that experienced social network users seem to have the ability to deal with and detect. It is also worth noting that there is a negative impact of social network experience on the other two cyber-attacks, however, this effect is still considered weak and not significant.

Low connections with strangers in SN: People with limited connections to strangers are less vulnerable to malware attack (t value=-2.049, P<0.05) as well as to the phishing scam that impersonates a legitimate organization (t value=-2.759, P<0.05). The result also shows that such low connections decrease users' vulnerability to phishing and clickjacking, although these relationships are not strong enough to be significant.

5 DISCUSSION

Some studies found no major variance between male and female in regards to response to email phishing [19]. Other studies have found women to be less susceptible to email phishing than men [20], [21]. Yet, female users have been repeatedly indicated as the weakest gender in detecting the risk in different cyber-attacks contexts such as email [14] or Social Network [15]. Our study also found that women are more vulnerable than men in all four types of attack. Furthermore, younger adults have been seen as a reckless group when dealing with risky emails, as stated by previous studies [14]. Yet, our study context is different and younger adults have shown their competency in

detecting social engineering attacks in social networks. This might be because their awareness and experience with social network settings and environment surpasses their knowledge of email environments and associated risks.

Benson et al. [22] state that students are less likely to fall for cybercrimes in social networks when compared with non-students. However, their study did not distinguish between the education levels of participants. Our study found that master's degree holders are more vulnerable to clickjacking attacks. This might be due to the fact that educated people usually seek new information even if there is risk associated with it. Further, a recent study [23] found that business students are more likely to open email phishing than humanities students. That study argues that this might be because business students are accustomed to a competitive environment and try to show their commitment by quick response to university emails.

The current study indicated three factors that can protect people from being deceived on social network websites. The individual competence level to deal with cyber-crime can be measured based upon three dimensions as proposed by [17], i.e., security awareness, privacy awareness, and self-efficacy. The result shows that this measure can significantly predict the individual's ability to detect phishing and malware attacks while decreasing the individual's vulnerability to clickjacking and scam attacks. Perception of self-ability to control the content shared on social network websites is considered a predictor of detection ability of social network threats [9].

A recent study investigating social engineering attacks in Facebook [15] found that the time elapsed since joining Facebook can be a significant predictor of susceptibility to victimisation. The more time has elapsed, the less vulnerable is the person. This accords with our findings as the years the individual spent using the network has been used as a measure of social network experience which also appeared to increase the individual awareness of the risk associated with using the network. Experienced users are more familiar with phishing and clickjacking, and thereby, easily detect them. Despite the fact that the relationship between SN experience and vulnerability to the other two attacks (malware, scam) are not considered statistically significant, experience with the network has decreased the likelihood of victimisation.

Previous studies claim that having a large network size is positively associated with online vulnerability [8], [10] as larger network's size included strangers as well as friends. Our study found that low connections with strangers can protect people from being deceived in social network specifically when encountering malware and phishing scam attacks. Limited connection with strangers is also decreasing the individual vulnerability to phishing and clickjacking attacks.

Our user study results provided an insight on the possibility to segment SN users based on their characteristics and vulnerabilities as a basis for a semi-automated security advisory system that responds to individual user vulnerabilities.

6 The architecture of a security advisory system

Our research on determining user vulnerabilities affords a basis for profiling users according to their weakness in respect of particular threats. In turn, this provides a

means to design a personalised advisory system that sends awareness posts to target individual user needs. For example, if the characteristics of the user are similar to those who are vulnerable to clickjacking, the advisory system might send awareness posts to the user and advise him/her on how to deal with this type of threat. The architecture of our proposed semi-automated advisory system is shown in Figure 5. A brief description of each component is given below.

Social networks users. A message must be sent to social networks users who want to register and benefit from the advisory system; ***Completing assessment survey.*** Any new user should start by completing a start-up survey that helps us assess participants' behaviour and perception in online social networks. This assessment survey result will profile the user in the most suitable segment later on to receive advice that suits the particular user needs. ***Pre-processing.*** The collected data will go through different screening and analysis tests such as construct reliability and validity tests. ***User classification.*** The segmentation process can be based on two different machine learning approaches: supervised or unsupervised [24]. Using unsupervised techniques such as clustering might be not suitable in our system as it requires no prior knowledge and clusters users based on patterns of unlabelled data. We aim to group users based on their vulnerability to different cyber-attacks. Therefore, supervised techniques such as classification are more appropriate to our goal, where the classes are predefined and the users grouped based on determined criteria. Thus, users will be classified into different groups based on the result of the scenario-based experiment in the assessment survey. Every segment should include users who shared similar characteristics that were found to increase vulnerability to a particular type of threat. For example, based on users' response to the phishing attack in the scenario-based experiment, users may be grouped into at least three segments: high, moderate, and low vulnerability. However, as we have considered user characteristics in the classification process, we might have more than one high vulnerable segment to a particular type of attack. For example, age and gender are among the factors that are included in the classification process, so it is possible to have two high vulnerable segments, e.g., one segment includes young-adult males and the other includes mid-aged females. This variation in the segmentation process can help us provide more individualised awareness messages.

Vulnerability Threshold. The local administrator can determine the threshold and the priority for each type of attack. For example, in our study we found that phishing scam is the most effective attack. Therefore, the threshold for this type of attack may be set to 3 which means high, moderate and low vulnerable segments will receive awareness advice on this type of threat. While the severity of malware attack is considered average in our study, we might set its threshold to 2, meaning that malware-related advice will be sent to the high and moderate rated vulnerability segments. Both phishing and clickjacking thresholds may be set to 1, meaning that only high vulnerable segments will receive advice for these two types of attack. Of course, a single user could be vulnerable to different types of attack and assigned to more than one segment. Therefore, the priority of the received type of advice is also determined by the attack's vulnerability threshold as assigned by the local administrator.

Segment filtering. In this step, segments are filtered based on threat thresholds. For each type of attack, only segments in the threshold vulnerability level will be addressed.

For instance, only segments with high vulnerability to phishing and clickjacking attacks may be considered. While according to the threshold of phishing scam, high, moderate, and low vulnerable segments may be taken into account.

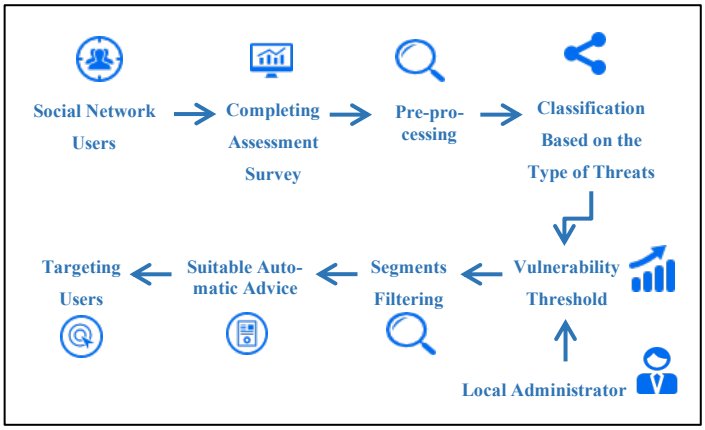

Fig. 5. Architecture of Semi-Automated Advisory System

Suitable automatic advice. Different user segments are vulnerable to different threats and require advice that is tuned to their needs. With this in mind, each of the identified threats has a set of recommendations that would help individuals to avoid falling victim to a particular threat. ***Targeting users.*** Each segment of users will receive automatic advice that aims to sensitise them to threats to which they are more vulnerable, while each single user can receive more than one package of advice, based on attack priorities that he/she is vulnerable to.

CONCLUSION

We are investigating why people easily fall victim to cyber-attacks in various online channels and whether vulnerabilities differ across cyber-attack categories in the context of social networks. The present study indicates that people respond differently to different types of cyber-attacks. A phishing attack that pretended to be from an authorized and legitimate organization (Facebook) is the most successful attack in our study with 37% of participants falling victim.

Female participants were found to be more vulnerable to social engineering victimisation than male participants. Younger and mid-aged adults show high detection ability compared to other age groups. Education is found to influence people's capability, as users with technical majors were found to be competent to detect cyber-attacks. Furthermore, the study result demonstrates that users' competence level, their experience with social networks, and low connections with strangers in the network play an important role in preventing people from falling victims to certain types of cyber-attacks.

The proposed semi-automated advisory system should help to address the problem of human vulnerabilities and weakness in detecting social engineering attacks. As-

sessing social network users and grouping them based on their behaviour and vulnerabilities is essential in order to focus relevant advice that meets users' needs. This is considered cost and time effective as users are only presented with insight on relevant threats. Furthermore, integrating individuals' needs as well as administrator's knowledge of existing threats, could avoid the overhead and inconvenience of sending blanket advice to all social network users.

References

1. B. B. Gupta, N. A. G. Arachchilage, and K. E. Psannis, "Defending against phishing attacks: taxonomy of methods, current issues and future directions," *Telecommun. Syst.*, vol. 67, no. 2, pp. 247–267, 2018.
2. S. M. Albladi and G. R. S. Weir, "User characteristics that influence judgment of social engineering attacks in social networks," *Human-centric Comput. Inf. Sci.*, vol. 8, no. 1, p. 5, Dec. 2018.
3. E. J. Williams, A. Beardmore, and A. N. Joinson, "Individual differences in susceptibility to online influence: A theoretical review," *Comput. Human Behav.*, vol. 72, pp. 412–421, 2017.
4. S. S. Andaleeb, "Market Segmentation, Targeting and Positioning," in *Strategic Marketing Management in Asia*, Emerald Group Publishing Limited, 2016, pp. 179–207.
5. K. Krombholz, H. Hobel, M. Huber, and E. Weippl, "Advanced social engineering attacks," *J. Inf. Secur. Appl.*, vol. 22, pp. 113–122, 2015.
6. J.-W. Bullee, L. Montoya, M. Junger, and P. Hartel, "Spear phishing in organisations explained," *Inf. Comput. Secur.*, vol. 25, no. 5, pp. 593–613, 2017.
7. S. Rathore, P. K. Sharma, and J. H. Park, "XSSClassifier : An Efficient XSS Attack Detection Approach Based on Machine Learning Classifier on SNSs," *J. Inf. Process. Syst.*, vol. 13, no. 4, pp. 1014–1028, 2017.
8. A. Vishwanath, "Habitual facebook use and its impact on getting deceived on social media," *J. Comput. Commun.*, vol. 20, no. 1, pp. 83–98, 2015.
9. G. Saridakis, V. Benson, J. N. Ezingeard, and H. Tennakoon, "Individual information security, user behaviour and cyber victimisation: An empirical study of social networking users," *Technol. Forecast. Soc. Change*, vol. 102, pp. 320–330, 2016.
10. S. L. Buglass, J. F. Binder, L. R. Betts, and J. D. M. Underwood, "When 'friends' collide: Social heterogeneity and user vulnerability on social network sites," *Comput. Human Behav.*, vol. 54, pp. 62–72, 2016.
11. I. Alseadoon, M. F. I. Othman, and T. Chan, "What Is the Influence of Users' Characteristics on Their Ability to Detect Phishing Emails?," in *Advanced Computer and Communication Engineering Technology*, Springer International Publishing, 2015, pp. 949–962.
12. A. Vishwanath, B. Harrison, and Y. J. Ng, "Suspicion, Cognition, and Automaticity Model of Phishing Susceptibility," *Communic. Res.*, 2016.
13. C. Iuga, J. R. C. Nurse, and A. Erola, "Baiting the hook: factors impacting susceptibility to phishing attacks," *Human-centric Comput. Inf. Sci.*, vol. 6, no. 1, p. 8, Dec. 2016.
14. S. Sheng, M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs, "Who falls for phish? A Demographic Analysis of Phishing Susceptability and Effectiveness of Interventions," in *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, 2010, pp. 373–382.

15. A. Algarni, Y. Xu, and T. Chan, “An empirical study on the susceptibility to social engineering in social networking sites: the case of Facebook,” *Eur. J. Inf. Syst.*, vol. 26, no. 6, pp. 661–687, 2017.
16. S. Albladi and G. R. S. Weir, “Vulnerability to social engineering in social networks: a proposed user-centric framework,” in *2016 IEEE International Conference on Cybercrime and Computer Forensic (ICCCF)*, 2016, pp. 1–6.
17. S. M. Albladi and G. R. S. Weir, “Competence measure in social networks,” in *IEEE International Carnahan Conference on Security Technology (ICCST)*, 2017, pp. 1–6.
18. H. Gao, J. Hu, T. Huang, J. Wang, and Y. Chen, “Security issues in online social networks,” *IEEE Internet Comput.*, vol. 15, no. 4, pp. 56–63, 2011.
19. P. Kumaraguru, J. Cranshaw, A. Acquisti, L. Cranor, J. Hong, M. A. Blair, and T. Pham, “School of phish,” *Proc. 5th Symp. Usable Priv. Secur. - SOUPS ’09*, p. 1, 2009.
20. W. Flores, H. Holm, G. Svensson, and G. Ericsson, “Using phishing experiments and scenario-based surveys to understand security behaviours in practice,” *Inf. Manag. Comput. Secur.*, vol. 22, no. 4, pp. 393–406, 2014.
21. J. Mohebzada, A. El Zarka, A. Bhojani, and A. Darwish, “Phishing in a University Community,” in *International Conference on Innovations in Information Technology (IIT)*, 2012, pp. 249–254.
22. V. Benson, G. Saridakis, and H. Tennakoon, “Purpose of social networking use and victimisation: Are there any differences between university students and those not in HE?,” *Comput. Human Behav.*, vol. 51, pp. 867–872, 2015.
23. S. Goel, K. Williams, and E. Dincelli, “Got Phished: Internet Security and Human Vulnerability,” *J. Assoc. Inf. Syst.*, vol. 18, no. 1, pp. 22–44, 2017.
24. M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” *Science (80-.).*, vol. 349, no. 6245, pp. 255–260, Jul. 2015.