2023

Invoking Suspicion for Improved Accuracy in Phishing Email Identification

Meg Harris

Deanna House

Invoking Suspicion for Improved Accuracy in Phishing Email Identification

Meg Harris
University of Nebraska at Omaha
megharris@unomaha.edu

Deanna House
University of Nebraska at Omaha
deannahouse@unomaha.edu

ABSTRACT

Threats to private and public sector environments are increasing in sophistication and becoming increasingly difficult to identify proactively. While there is urgency to stay abreast of critical patches and protective measures to ensure that the technological risks are mitigated, the human element is a consistent group that attackers focus on from a social engineering standpoint. Not only are humans unpredictable, but they have varying perspectives, backgrounds, and levels of experience that influence their behaviors and motivations. While training is a recommended best practice for improving individuals' capabilities for identifying illegitimate messages, there are still many phishing attacks that successful even with training in place. This research explores the element of suspicion and how suspicion could be invoked to influence behavior related to phishing message response.

Keywords

Suspicion, phishing, human element, deception

INTRODUCTION & LITERATURE REVIEW

Phishing attacks and areas that can influence the outcomes of attacks, such as email habits (Vishwanath, 2015), have been studied in prior research. While protective controls in securing both private and public sector organizations do have an effect on mitigating the risk of cyber-attacks, the human element is still an extremely vulnerable attack vector. According to the most recent Verizon Data Breach Investigations Report, 82% of breaches involved a human element (Verizon, 2022). Earlier studies on security countermeasures, such as D'arcy, Hovav, and Galletta (2009) have discussed the importance of security, education, training and awareness (SETA) programs. However, the recommendations for security training can be difficult to justify, particularly when research has shown that click rates can sometimes be higher when the detection of a phishing email is difficult (Steves, Greene, & Theofanos, 2020). Organizations across multiple industries frequently employ such programs, but it is concerning that approximately 2.9% of employees click on phishing emails (Verizon, 2022). When the magnitude of compromised records, is taken into consideration, this could result in millions of clicks. Many of these aspects of phishing attacks have been studied in prior research including types of phishing attacks (Tandale & Pawar, 2020); (Jakobsson & Young, 2005); (Aonzo, Merlo, Tavella, & Fratantonio, 2018), countermeasures to mitigate risk related to phishing attacks (Tandale & Pawar, 2020); (Vayansky & Kumar, 2018); (Azeez, Misra, Margaret, Fernandez-Sanz, & Abdulhamid, 2021); (Wu, Miller, & Garfinkel, 2006), the weakness represented by the human element (Thompson, 2013); (Anthony, 2019) (Alghenaim, Bakar, & Rahim, 2023); (Ferreira & Teles, 2019). However, our research is aimed at filling the gaps in the extant literature addressing concerns of why people click on such a high percentage of links and possibly more importantly, why this percentage is rising.

Suspicion

Suspicion is a detection method that any and all email users can employ and thus it is a central concept in identifying phishing emails. The American Psychological Association (*Suspiciousness – APA Dictionary of Psychology*, n.d.) defines suspiciousness as "mistrust of the motives or sincerity of others," while Bobko et al. (2014) define suspicion as "a person's *simultaneous* state of cognitive activity, uncertainty, and perceived malintent about underlying information that is being electronically generated, collated, sent, analyzed, or implemented by an external agent" (p. 293). However, Coppola and House (2019) bring to light the lack of an accepted definition of suspicion in the literature though it may be clear that the consistent features in definitions of suspicion lie in the intentions of others.

Many have affirmed that suspicion is impacted by situational elements and therefore is an emotional state, to an extent. In keeping with this classification of suspicion as an emotional state, prior research has shown that a variety of factors can cultivate suspicion. It can be developed or enhanced via cultural differences, values, and communication styles (Luu, 2017), an innate bias to assume honesty and truthfulness (Van Swol et al., 2013), conspiratorial thinking caused by personality traits such as paranoia, cynicism, alienation, and anxiety (Radnitz & Underwood, 2017). However, external factors can also influence suspicion. These include the fluctuating situational opinions of the public, repeated exposure to insinuating information, (Radnitz & Underwood, 2017), the demeanor of the provider of information, (Van Swol et al., 2013), and deception (Epley & Huff, 1998).

Prior research has also shown the difficulty in measuring suspicion as a person's current psychological state at the time of measurement can be a confounding factor (Radnitz & Underwood, 2017), impacting the way they perceive and internalize information. It can also be difficult to measure suspicion as individuals can have concurrent inconsistent inferences of received information meaning that they could be simultaneously suspicious of motives while also taking in the information at face value (Ham & Vonk, 2011).

Deception, Trust, and Distrust

Deception, by contrast, is "distortion of or withholding of fact with the purpose of misleading others" (*Deception– APA Dictionary of Psychology*, n.d.) and as can be seen in this definition is often perpetrated in one of two methods: commission (a lie) or omission (withholding information). Prior work has found that deceivers can justify their claims to the information-receiver by providing plausible information, hoping to boost their credibility, along with other tactics like guiding the communication toward their desired end goal (Paik & Van Swol, 2017).

Deception has been linked with distrusting beliefs in prior work (Jarvenpaa & Majchrzak, 2010) which often times lead to suspicion. It is worth distinguishing trust and distrust, which are said to be decisions, from suspicion which is centered in uncertainty (Luu, 2017). Suspicion, along with situational abnormalities and a general distrusting disposition have been identified as antecedents to distrusting beliefs while trusting beliefs, distrusting beliefs, and ambivalence have been noted in prior work as antecedents to trusting intentions (Moody et al., 2014).

Phishing

Phishing is defined as "a technique for attempting to acquire sensitive data…through a fraudulent solicitation in email or on a web site, in which the perpetrator masquerades as a legitimate business or reputable person," according to NIST (n.d.). Spear-phishing is known as a more targeted approach to phishing in which the perpetrator purports to be someone known and trusted by their attack target. Both are forms of social engineering in which the target of the attack is manipulated or deceived so that the attacker achieves their goals of system access or theft, among others. The Anti-Phishing Working Group (APWG) is an organization whose goal is to eradicate theft and fraud perpetrated primarily via phishing methods (*APWG | About the APWG*, n.d.). Their data shows 1,270,883 unique phishing attacks in just the third quarter of 2022 alone (*APWG | Phishing Activity Trends Reports*, 2022). Unfortunately, that is the record-setting quarter ever observed by the APWG since its inception in 2003. August of 2022 had the highest monthly number of phishing attacks at 430,141. The APWG 2022 third quarter report goes on to say that the increase in attacks is a result of some targets being attacked multiple times by "persistent phishers." Suffice it to say, phishing is a real and significant issue for individuals and organizations.

Research Questions

Due to the above identified gaps in the research, we pose the following research questions (RQ):

> **RQ1:** How does the level of suspicion change (in relation to baseline state suspicion) after receiving legitimate emails in the medium or high suspicion categories? How does behavioral motivation change as a result of changes in level of suspicion?
>
> **RQ2:** How does the level of suspicion change (in relation to baseline state suspicion) after receiving illegitimate emails in the low or medium suspicion categories? How does behavioral motivation change as a result of changes in level of suspicion?
>
> **RQ3:** If there is a change in suspicion level related to RQ1, what is the impact of contextual suspicion?

RQ4: If there is a change in suspicion level related to RQ2, what is the impact of contextual suspicion?

PROPOSED METHODOLOGY

The researchers seek to understand how suspicion can affect an individual's motivation to click on a link in an email. This research will utilize an experiment to measure changes in suspicion levels, resulting changes in behavioral motivation in alignment with behavioral motivation theory (Fowles, 1987), and the impact, if any, of contextual suspicion. Contextual suspicion is suspicion that is impacted by the environment and state of mind of the email receiver at the time of reading. Distler (2023), discusses context in terms of physical, social, internal, technical, task, and temporal context to determine its influence on reactions to spear-phishing emails stating that the influence of context on responses to phishing emails is not well understood yet. He goes on to state that the "context a person is currently in seems to be a highly relevant factor to investigate," (Distler, 2023, p. 4) then providing an example of a tired or stressed person being influenced differently based on this internal context. Other authors have called contextual suspicion by other names such as state suspicion. Bobko et al (2014, p. 3) describe state suspicion as "a person's simultaneous state of cognitive activity, uncertainty, and perceived malintent about underlying information." The researchers begin by building a three by two model depicting suspicion (three levels: low, medium, and high) and email legitimacy (two states, legitimate and illegitimate). The researchers will scrape an existing email dataset for keywords, which we will call suspicion markers, and classify them into the three suspicion level categories. These keyword markers could consist of language not commonly used any more, for example "Dear X". The suspicion levels indicate how suspicious the receiver of the email is upon reading it. If the reader trusts the email, this would indicate a low suspicion level. If they feel that something is wrong and would not want to click on any links, this would indicate a high suspicion level.

Participants will then be gathered and a baseline suspicion level will be measured prior to assigning them the task of reading emails from the selected dataset and denoting the emails as either legitimate or illegitimate. The researchers will remeasure the suspicion level after they read the emails with the classified suspicion indicators to observe the change. The change in suspicion level will be compared to the accuracy of email legitimacy notation by participants.

EXPECTED CONTRIBUTIONS TO RESEARCH AND PRACTICE

It is commonly known that humans are the weakest component of a security plan (Sasse et al., 2001); (West et al., 2009). This is especially true with regards to phishing attacks. Unlike other attacks, like ransomware for example, where the perpetrator can achieve the goal of the attack without the target taking action, phishing is only successful if the target buys into the perpetrator's ruse. This is a key reason for continual and consistent training. However, despite regular education and preparation, phishing attacks, which are persistently perpetrated, are continually successful which leads us to believe that current training programs have room for improvement. Refining education starts with gaining knowledge about the sources of the problem. We posit that the source of the problem is primarily due to the attack target's lack of suspicion. We theorize that if targets maintained higher levels of suspicion, they would lessen their likelihood of falling victim to phishing attacks.

Therefore, this study contributes to research and practice by adding to the body of knowledge surrounding suspicion. This study identifies markers that elicit high, medium, and low levels of suspicion in attack targets and identifies which suspicion-level markers cause an individual to label an email as either legitimate or illegitimate. With the knowledge of which types of markers educe suspicion, we can improve phishing education, along with other types of cyber-attacks and social engineering, mitigating the likelihood of attack success. This work is conducted with the ultimate goal of protecting organizations and individuals from phishing, and other, attacks.

CONCLUSION AND LIMITATIONS

While this work aims to improve education with the larger goal of protecting individuals and organizations from phishing and other types of attacks, it may also have unintended consequences of aiding attackers. The knowledge of suspicion markers and the ability of attack targets to identify illegitimate emails may help attackers to improve their phishing emails to the point that the markers identified in our study are no longer used in phishing emails. Further research could utilize the results of our study to further refine abilities to maintain high levels of suspicion without the use of specific markers, further diminishing the likelihood of successful phishing attacks.

REFERENCES

1. Alghenaim, M. F., Bakar, N. A. A., & Rahim, F. A. (2023). Awareness of Phishing Attacks in the Public Sector: Review Types and Technical Approaches. In M. A. Al-Sharafi, M. Al-Emran, M. N. Al-Kabi, & K. Shaalan (Eds.), *Proceedings of the 2nd International Conference on Emerging Technologies and Intelligent Systems* (pp. 616–629). Springer International Publishing.
2. Anthony, B. (2019). *Social Engineering: The Human Element of Cybersecurity* [M.S., Utica College].
3. Aonzo, S., Merlo, A., Tavella, G., & Fratantonio, Y. (2018, October). *Phishing Attacks on Modern Android | Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*.
4. *APWG | About the APWG*. (n.d.).
5. *APWG | Phishing Activity Trends Reports*. (2022).
6. Azeez, N. A., Misra, S., Margaret, I. A., Fernandez-Sanz, L., & Abdulhamid, S. M. (2021). Adopting automated whitelist approach for detecting phishing attacks. *Computers & Security*, *108*, 102328.
7. Bobko, P., Barelka, A. J., & Hirshfield, L. M. (2014). The construct of state-level suspicion: A model and research agenda for automated and information technology (IT) contexts. *Human Factors*, *56*(3), 489–508.
8. Bobko, P., Barelka, A. J., Hirshfield, L. M., & Lyons, J. B. (2014). Invited article: The construct of suspicion and how it can benefit theories and models in organizational science. *Journal of Business and Psychology*, *29*, 335-342.
9. Coppola, J., & House, D. (2019). Suspicion in Phishing and Organization Risk. *SAIS 2019 Proceedings*.
10. D'Arcy, J., Hovav, A., & Galletta, D. (2009). User Awareness of Security Countermeasures and Its Impact on Information Systems Misuse: A Deterrence Approach. *Information Systems Research*, *20*(1), 79–98.
11. *Deception– APA Dictionary of Psychology*. (n.d.).
12. Distler, V. (2023, April). The Influence of Context on Response to Spear-Phishing Attacks: an In-Situ Deception Study. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1-18).
13. Epley, N., & Huff, C. (1998). Suspicion, affective response, and educational benefit as a result of deception in psychology research. *Personality and Social Psychology Bulletin*, *24*(7), 759–768.
14. Ferreira, A., & Teles, S. (2019). Persuasion: How phishing emails can influence users and bypass security measures. *International Journal of Human-Computer Studies*, *125*, 19–31.
15. Fowles, D. C. (1987). Application of a behavioral theory of motivation to the concepts of anxiety and impulsivity. *Journal of Research in Personality*, *21*(4), 417–435.
16. Ham, J., & Vonk, R. (2011). Impressions of impression management: Evidence of spontaneous suspicion of ulterior motivation. *Journal of Experimental Social Psychology*, *47*(2), 466–471.
17. Jakobsson, M., & Young, A. (2005). *Distributed Phishing Attacks*.
18. Jarvenpaa, S. L., & Majchrzak, A. (2010). Research Commentary—Vigilant Interaction in Knowledge Collaboration: Challenges of Online User Participation Under Ambivalence. *Information Systems Research*, *21*(4), 773–784.
19. Luu, T. (2017). Cultural intelligence and state suspicion: Attachment styles as moderators. *Corporate Communications: An International Journal*, *22*(1), 113–132.
20. Moody, G. D., Galletta, D. F., & Lowry, P. B. (2014). When trust and distrust collide online: The engenderment and role of consumer ambivalence in online consumer behavior. *Electronic Commerce Research and Applications*, *13*(4), 266–282.
21. National Institute of Standards and Technology (NIST) Computer Security Resource Center (CSRC). (n.d.) phishing.
22. Paik, J. E., & Van Swol, L. M. (2017, May 12). *Justifications and Questions in Detecting Deception | SpringerLink*.
23. Radnitz, S., & Underwood, P. (2017). Is Belief in Conspiracy Theories Pathological? A Survey Experiment on the Cognitive Roots of Extreme Suspicion. *British Journal of Political Science*, *47*(1), 113–129.
24. Sasse, M. A., Brostoff, S., & Weirich, D. (2001). Transforming the 'Weakest Link'—A Human/Computer Interaction Approach to Usable and Effective Security. *BT Technology Journal*, *19*(3), 122–131.
25. Steves, M., Greene, K., & Theofanos, M. (2020). Categorizing human phishing difficulty: A Phish Scale. *Journal of Cybersecurity*, *6*(1), tyaa009.

26. *Suspiciousness – APA Dictionary of Psychology*. (n.d.).
27. Tandale, K. D., & Pawar, S. N. (2020). Different Types of Phishing Attacks and Detection Techniques: A Review. *2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC)*, 295–299.
28. Thompson, H. (2013). The Human Element of Information Security. *IEEE Security & Privacy*, *11*(1), 32–35.
29. Van Swol, L. M., Braun, M. T., & Kolb, M. R. (2013, April 25). *Deception, Detection, Demeanor, and Truth Bias in Face-to-Face and Computer-Mediated Communication—Lyn M. Van Swol, Michael T. Braun, Miranda R. Kolb, 2015*.
30. Vayansky, I., & Kumar, S. (2018). Phishing – challenges and solutions. *Computer Fraud & Security*, *2018*(1), 15–20.
31. Verizon. (2022). *2022 Data Breach Investigations Report*. Verizon Business.
32. Vishwanath, A. (2015, June 1). *Examining the Distinct Antecedents of E-Mail Habits and its Influence on the Outcomes of a Phishing Attack | Journal of Computer-Mediated Communication | Oxford Academic*.
33. West, R., Mayhorn, C., Hardee, J., & Mendel, J. (2009). *The Weakest Link: A Psychological Perspective on Why Users Make Poor Security Decisions* [Chapter]. Social and Human Elements of Information Security: Emerging Trends and Countermeasures; IGI Global.
34. Wu, M., Miller, R. C., & Garfinkel, S. L. (2006, April). *Do security toolbars actually prevent phishing attacks? | Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.