Investigating phishing awareness using virtual agents and eye movements

Lynsay A. Shepherd
Andrea Szymkowiak

Poster: Investigating Phishing Awareness Using Virtual Agents and Eye Movements

Lynsay A. Shepherd
School of Design and Informatics
Abertay University
Dundee, United Kingdom
lynsay.shepherd@abertay.ac.uk

Andrea Szymkowiak*
School of Design and Informatics
Abertay University
Dundee, United Kingdom
a.szymkowiak@abertay.ac.uk

ABSTRACT

Phishing emails typically attempt to persuade recipients to reveal private or confidential information (e.g., passwords or bank details). Interaction with such emails places individuals at risk of financial loss and identity theft. We present an ongoing study using eye tracking metrics and varying interface components to assess users' ability to spot simulated phishing attempts. Findings seek to establish how users interact with email inbox interfaces and will inform future design of usable security tools.

CCS CONCEPTS

• **Security and privacy → Human and societal aspects of security and privacy**.

KEYWORDS

Human-centred security, phishing, eye tracking, virtual agents.

ACM Reference Format:

1 INTRODUCTION AND BACKGROUND

Citizens have become dependent on the Internet as a primary method of communication via the use of messaging and entertainment services, social media, and e-commerce, to name but a few. Users share vast amounts of personal data daily and may be unaware of the associated risks. Thus, the Internet provides an ideal setting for opportunistic fraudsters to target individuals. One method of deceiving users is via phishing emails, where the emails are carefully designed to coax users into revealing information such as bank details and passwords. Engaging with phishing emails can place users at risk of fraud and financial loss.

*Corresponding author.

The onset of the COVID-19 pandemic saw a significant rise in the number of cyberattacks and cybercrime. Phishing was a commonly used technique, with individuals and organisations targeted [9]. In 2022, IBM reported that data breaches caused by phishing cost organizations an average of $4.91 million [7]. Due to the personal and monetary damage phishing causes to individuals and businesses, research in this area is vital to help prevent such harm.

This paper outlines a novel method aimed to increase users' security awareness via virtual agents within an email inbox interface while classifying email exchanges. Eye tracking measures are used to investigate how users visually process aspects of the inbox.

1.1 Phishing Awareness

Despite the availability of anti-phishing educational materials, users continue to fall victim to this form of social engineering. Researchers have developed various techniques to alert users to potentially risky email content and links. Such work includes the use of machine learning-based browser extensions to detect phishing emails [2], security training involving cartoons and characters such as *"PhishGuru"* [8], or the development of serious games to educate users, e.g., *"PHISHY"* [3]. *"PHISHY"* [3] and *"PhishGuru"* [8] did not aim to investigate features of the characters in their relation to user behaviour. However, features of characters such as embodied or anthropomorphic virtual agents, i.e. human-like characters, may influence the user as a form of social presence.

1.2 Interface Design and Social Presence

Social presence was originally defined within telecommunication as the *"degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships"* [15]. The definition has also been extended to reflect technology mediated interaction [6]. Social presence in the form of an embodied agent or avatar has been beneficial in delivering persuasive messages to users [1], and can have a positive impact on student motivation and participation [14]. Whilst the effects of features of virtual agents have been extensively investigated in online and commerce environments [10] [11], their role in alerting users to potential phishing emails is less clear.

It is challenging to identify a possible expression of an agent that would alert users to potentially unsafe behaviour, as most research on emotional expression of agents has focused on the six universal (human) emotional expressions [5] (anger, surprise, disgust, enjoyment, fear, sadness), which may not necessarily relate a warning message to the user. Pelachaud and Poggi [12] researched the creation of subtle expressions, such as being 'sorry for' in virtual agents, and provided a framework for these expressions. However,

the association of expressions that could potentially alert a user to an unsafe situation with user behaviour, is under-researched.

1.3 Eye tracking and Phishing Awareness

Eye tracking measures can be used to investigate participants cognitive engagement with presented stimuli [4]. Work conducted by Yuan et al. [16] examined human cognitive models in the context of an authentication system. They suggested data gained from eye tracking devices, such as a viewer's scan paths, could provide a good understanding of how a viewer engages with presented material. Furthermore, previous work has utilised eye tracking measures to investigate how users look at phishing emails [13]. We build on such work by examining the visual processing of the emails in the presence of various interface agents in an effort to improve awareness in relation to secure behaviour.

2 METHODS AND PRELIMINARY RESULTS

We developed an interactive simulated email inbox environment, named *"Wee Inbox"* which gives researchers control over interactions with emails. *"Wee Inbox"* contains 12 messages which are a mix of legitimate and phishing emails. Each participant sees the same 12 messages in a randomised order. Participants were placed in one of three conditions: a control condition showing a mailbox, a "neutral" agent face, and a "warning" agent face presumed to be associated with an alerting expression. In line with Pelachaud and Poggi's [12] work, the agent expressions are subtle in nature rather than displaying an obvious emotion. Participants classified emails as legitimate or fraudulent. During the classification task, participants' eye movements were recorded (Figure 1).

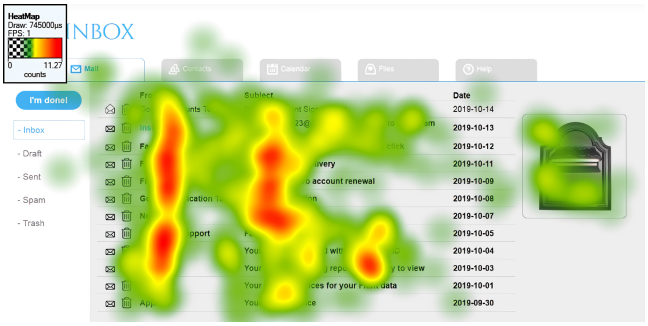

Figure 1: Example showing distribution of gaze (heatmap) (control condition). Red indicates a higher gaze count.

The use of eye tracking measures is of particular importance to the study. Eye movements flag whether participants have seen the stimuli, along with how much attention is given to aspects of the inbox. We utilise Areas of Interest (AOIs) to define the agent position on-screen. This allows us to investigate time to first fixation, fixation count, visit count, and visit duration in the AOI. We have currently analysed 39 participants using a Tobii TX-300 screen-based eye tracker (*control condition n=14, neutral agent n=12, warning agent n=13*). Preliminary results indicate participants looked at the neutral and warning agent at an earlier point in the experiment compared to the control. We have yet to investigate these results in relation to phishing classification performance.

3 CONCLUSIONS

Outcomes from our study seek to further research into usable security, identifying novel approaches when developing security tools. Owing to increasing financial losses incurred by businesses and individuals due to successful phishing attempts, research into how to tackle this form of crime continues to be ever more important. Through the implementation of varying interface designs and the use of eye tracking to investigate interactions, we suggest our work can provide a deeper understanding of how users engage with phishing emails with the ultimate aim to keep these users secure.

ACKNOWLEDGMENTS

The authors would like to thank The Carnegie Trust for the Universities of Scotland (Grant no. RIG007545), the Scottish Funding Council, and Abertay University's Emergent Technology Centre and Graduate School for supporting this work.

REFERENCES

[1] Amy L Baylor. 2009. Promoting motivation with virtual agents and avatars: role of visual presence and appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3559–3565.

[2] Paul Boyle and Lynsay A. Shepherd. 2021. Mailtrout: a machine learning browser extension for detecting phishing emails. In *34th British HCI Conference*. BCS, 104–115.

[3] Gokul CJ, Sankalp Pandit, Sukanya Vaddepalli, Harshal Tupsamudre, Vijayanand Banahatti, and Sachin Lodha. 2018. Phishy-a serious game to train enterprise users on phishing awareness. In *Proceedings of the 2018 annual symposium on computer-human interaction in play companion extended abstracts*. ACM, 169–181.

[4] Andrew T Duchowski. 2002. A breadth-first survey of eye-tracking applications. *Behavior Research Methods Instruments and Computers* 34, 4 (2002), 455–470.

[5] Paul Ekman. 1970. Universal facial expressions of emotion. *California Mental Health Research Digest* 8 (1970), 151–158. Issue 4.

[6] Charlotte N Gunawardena. 1995. Social presence theory and implications for interaction and collaborative learning in computer conferences. *International journal of educational telecommunications* 1, 2 (1995), 147–166.

[7] IBM. 2022. *IBM Report: Consumers Pay the Price as Data Breach Costs Reach All-Time High.* IBM. Retrieved February 3, 2023 from https://newsroom.ibm.com/2022-07-27-IBM-Report-Consumers-Pay-the-Price-as-Data-Breach-Costs-Reach-All-Time-High

[8] Ponnurangam Kumaraguru, Steve Sheng, Alessandro Acquisti, Lorrie Faith Cranor, and Jason Hong. 2010. Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology (TOIT)* 10, 2 (2010), 1–31.

[9] Harjinder Singh Lallie, Lynsay A Shepherd, Jason RC Nurse, Arnau Erola, Gregory Epiphaniou, Carsten Maple, and Xavier Bellekens. 2021. Cyber security in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic. *Computers & security* 105 (2021), 102248.

[10] Renaud Lunardo, Grégory Bressolles, and François Durrieu. 2016. The Interacting Effect of Virtual Agents' Gender and Dressing Style on Attractiveness and Subsequent Consumer Online Behavior. *Journal of Retailing and Consumer Services* 30 (05 2016), 59–66. https://doi.org/10.1016/j.jretconser.2016.01.006

[11] Jeunese Payne, Andrea Szymkowiak, Paul Robertson, and Graham Johnson. 2013. Gendering the machine: Preferred virtual assistant gender and realism in self-service. In *International Workshop on Intelligent Virtual Agents*. Springer, 106–115.

[12] Catherine Pelachaud and Isabella Poggi. 2002. Subtleties of facial expressions in embodied agents. *The Journal of Visualization and Computer Animation* 13, 5 (2002), 301–312.

[13] Kevin Pfeffel, Philipp Ulsamer, and Nicholas H Müller. 2019. Where the user does look when reading phishing mails–an eye-tracking study. In *Learning and Collaboration Technologies. Designing Learning Experiences: 6th International Conference, LCT 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, Proceedings, Part I 21*. Springer, 277–287.

[14] Jennifer C Richardson, Yukiko Maeda, Jing Lv, and Secil Caskurlu. 2017. Social presence in relation to students' satisfaction and learning in the online environment: A meta-analysis. *Computers in Human Behavior* 71 (2017), 402–417.

[15] John Short, Ederyn Williams, and Bruce Christie. 1976. *The Social Psychology of Telecommunications*. Wiley.

[16] Haiyue Yuan, Shujun Li, Patrice Rusconi, and Nouf Aljaffan. 2017. When eye-tracking meets cognitive modeling: applications to cyber security systems. In *International conference on human aspects of information security, privacy, and trust*. Springer, 251–264.