# Investigating Gaze Behavior in Phishing Email Identification

Francesco Pietrantonio, Alessio Botta, Giorgio Ventre
*Dep. Electrical Eng. and Information Technology*
*University of Naples Federico II Naples, Italy,*
f.pietrantonio@studenti.unina.it,
alessio.botta, giorgio.ventre@unina.it

Lugi Gallo
*Cyber Security Lab TIM S.p.A.,*
*Via Reiss Romoli 274, 10148 Turin, Italy*
luigi1.gallo@telecomitalia.it

Stefania Zinno
*Dep. Social Sciences*
*University of Naples Federico II, Naples, Italy,*
stefania.zinno@unina.it

Laura Mancuso, Roberta Presta
*Scienza Nuova Research Centre,*
*University Suor Orsola Benincasa, Naples, Italy,*
laura.mancuso@studenti.unisob.na.it,
roberta.presta@centroscienzanuova.it

***Abstract*—The pervasiveness of phishing signals the insufficiency of current measures. Through a multidisciplinary approach, we conducted an eye-tracking study on how and where users look when they have to classify an email as phishing or legitimate. Furthermore, we investigated whether there is a difference between expert and non-expert subjects. The study showed firstly, better performance in recognising phishing emails by experts. Secondly, eye movement data showed the use of different email inspection methods between experts and non-experts. This could open up scenarios in the area of the improvement of training courses and the development of more intuitive email client interfaces in the suggestion of important clues in the recognition of phishing emails.**

***Index Terms*—phishing, eye-tracking, training, interfaces**

## I. Introduction and Motivation

Phishing poses a pervasive and devastating threat in today's digital landscape, exploiting deceptive communications to prompt individuals to reveal sensitive data, engage in illicit activities, or spread malware. In Q3 2022, a record high of 1,270,883 phishing attacks (the worst quarter ever observed) were reported by the Anti Phishing Working Group [1]. Moreover, a report by IBM [2] found that the average cost of a corporate data breach in 2022 was $4.35 million, and that phishing was the second most common cause of these breaches. However, the consequences are reflected not only in economic terms, but also in environmental, reputational, and time and productivity losses [2]–[5].

Emails are the main vector for these attacks due to their affordability [3], widespread use, and security vulnerabilities stemming from the SMTP protocol [6], [7] and configuration errors in protocols such as SPF, DKIM, and DMARC [6]. Moreover, it is easy for attackers to collect a large number of email addresses [8].

Phishers use various techniques to increase the persuasiveness of their message, including visual manipulation (involving the alteration of visual elements and formatting), domain squatting (involving the registration of domain names similar to legitimate ones), and exploiting cognitive vulnerabilities and human psychology.

There are several categories of phishing, such as Spear Phishing (targeting specific individuals or groups), Whaling (targeting high-profile individuals or executives), Clone Phishing (employing minor modifications to legitimate emails), Calendar Phishing (employing invitations to fake events containing malicious links), and Lateral Phishing (exploiting trust between colleagues within a company).

Common countermeasures include antivirus and web/mail filtering to block malicious attachments, websites, and messages. However, since attackers exploit human weaknesses, efforts are also made to protect the user through guidelines, security policies, and staff training courses.

The persistence of phishing emails as a pervasive phenomenon highlights the insufficiency of current countermeasures, thereby motivating our research. Furthermore, the complex nature of this phenomenon has required a highly multidisciplinary approach and the search for diversified expertise. Potential factors contributing to this result could be email client interfaces not effectively directing users' attention to relevant cues, phishing training courses' dubious efficacy and knowledge retention, and a lack of consideration for human-related factors [9].

This research extends our earlier work [10], [11] on phishing at the University of Naples "Federico II". A key contribution is the Spamley Web App [1], developed to collect user characteristics through a survey and simulate an email client presenting users with emails, inspired by real ones, to categorize as phishing or legitimate. The survey, divided into two sections, covers demographics, education, work, email use context, computer skills, phishing detection abilities and awareness, and the second part delves into users' personality traits and

978-3-903176-58-4 © 2023 IFIP.

[1] https://spamley.comics.unina.it/

cognitive vulnerabilities. The gathered data was employed to study the relationship between user traits and phishing identification skills. In this work we complement the analysis by including data characterising the user's ocular behaviour.

## II. Related Works

The related works in phishing susceptibility and education present a complex picture with contrasting findings.

A real-world study with 515 participants demonstrated that the embedded training system PhishGuru led to users retaining knowledge after 28 days and that a second message reduced the likelihood of people giving away personal information [12]. This aligns with the results of a roleplay survey with 1001 participants from Sheng et al. [13], which also showed reduced susceptibility with prior exposure to education. Moreover, a study analysing a dataset of phishing emails sent to employees of client companies during their routine work over 1.5 years by PhishCo, a company hired to raise security awareness, further solidified this premise by reporting a decrease in employee click rates following long-term awareness training [14]. At last, a study conducted with university members using an actual phishing email as stimulus [15], underscored the importance of prior knowledge of phishing tactics and suggested that familiarity mitigates visceral triggers' influence and expands the use of deception indicators in the decision-making process.

However, these results should be interpreted with caution. It was reported that participants still fell for 28% of phishing messages post-training, thus illustrating the incomplete nature of the solution [13]. Moreover, a study by Caputo et al. [16] investigated its effectiveness in reducing employees' susceptibility to spear phishing in a medium-sized DC-based company, and found that training effects were lost between 28 and 90 days, and many participants were either non-clickers or all-clickers regardless of training type, casting further doubt on the effectiveness and longevity of training.

These contrasting findings culminate in the work by Lain et al. [17], which presented findings from a large-scale, 15 month phishing experiment involving over 14,000 company employees, that received simulated phishing emails in their normal working context. This study contests the effectiveness of embedded training during simulated phishing exercises, therefore exemplyifying the ongoing debate in the literature about the efficacy and longevity of the impact of phishing training.

Adding another dimension to this discourse, the study by Wang et al. [15] emphasized the influence of "visual triggers" and "phishing deception indicators" on user decision-making: results showed that focusing on visceral triggers (e.g. urgency) increased response likelihood, while focusing on phishing deception indicators (e.g. grammatical errors) decreased it. This suggests that user behavior towards phishing might not be solely dependent on training but also on the specific characteristics of the phishing email. Moreover, Lain et al. [17] presented findings from a large-scale, 15 month phishing experiment involving over 14,000 company employees, that received simulated phishing emails in their normal working context. They confirmed the effectiveness of email warnings, suggesting the importance of taking into consideration the interaction between the user and the interface when designing phishing countermeasures.

However, the characteristics of the human interacting with the interface should also be taken into consideration. It was reported that gender has no significant effect on phishing susceptibility [12] and that different educational materials reduced information entry into phishing sites by 40%, regardless of demographics [13]. On the other hand age was found to affect phishing susceptibility, with participants aged 18-25 being more vulnerable than older participants [12], [13]. Also, the authors in [14] found that some psychological vectors were more successful in tricking users into falling for phishing scams.

The contrasting results and the incomplete nature of the solutions underscore the need of our research, incorporating human-machine interaction and diverse expertise, an angle not yet fully explored in the current body of literature.

## III. Methodology and Expected Outcomes

Our goal is to address a fundamental question: What factors contribute to individuals' success to identify phishing attempts?

We consider two potential determinants:

- The role of experience (i.e. the participants' general knowledge in computer science and/or their professional engagement with the topic, meaning that experts possess at least a foundational or intermediate level of proficiency in one or both aspects) in detecting phishing attempts.
- The importance of attention (i.e. the focus on specific aspects of an email) in this process.

Two hypotheses related to these determinants can be formalized.

| | |
|---|---|
| H1 | Experts demonstrate better phishing recognition than non-experts |
| Verified if: | Performance(E) > Performance(NE) |
| If true: | Experience positively affects phishing recognition |
| If false: | Experience does not significantly influence phishing recognition |

| | |
|---|---|
| H2 | Experts and non-experts differ in their email inspection behavior |
| Verified if: | Eyetracking(E) $\neq$ Eyetracking(NE) |
| If true: | Distinct inspection techniques exist between experts and non-experts |
| If false: | Inspection techniques do not vary between the two groups |

Our expectations are twofold: first, that experience or familiarity with computer science will have a significant impact on the recognition of phishing. Second, that experts employ different strategies when inspecting emails than non-experts. These different inspection methods may explain the better performance of experts in detecting important cues in the recognition of phishing emails.

To test these hypotheses, we conducted an experiment on how and what people look at emails to determine whether they are phishing or not. Specifically, the eye-tracking wearable device Pro Glasses 2, developed by Tobii, was used. The device consists of a pair of glasses that use an infrared light to illuminate the subject's eyes and detect the centre of the pupil by reading corneal reflections. This data is used to track eye movements during the inspection of an email. We then used the Tobii Pro Lab software to process the collected data.

In addition, the Spamley Web App was used and adapted to meet the specific requirements of this study. It allows for the collection of participant characteristics through a survey and the presentation of stimulus emails to participants. After viewing each email, participants are asked to classify it as phishing or legitimate using two buttons. At the end of the test, participants can find out the percentage of correct answers.

The study considers participants recruited from the campus of the University of Naples 'Federico II' and the campus of the University of Naples 'Suor Orsola Benincasa'. The sample is divided into two groups according to experience, as defined above. All participants are exposed to a consistent set of email stimuli (inspired by real emails) both phishing and legitimate. The order in which the stimuli are presented is randomised, respecting two constraints: emails of the same type cannot appear more than three times consecutively (e.g. Phishing-Phishing-Phishing is not allowed), and a constant alternation between phishing and legitimate emails is not allowed (e.g. Phishing-Legitimate-Phishing-Legitimate is not allowed). This approach helps to control order and sequence effects. Each stimulus email is displayed for 20 seconds, but participants can move on by clicking on the 'next' button even before time runs out. This option mitigates potential noise resulting from unintentional or disinterested examination of stimuli.

The visual stimulus is divided into regions called Areas of Interest (AoIs): header, body, and URL. Some of the following metrics are calculated for each area: Total Duration of Visit, which is the overall duration of visits to an AoI measured in milliseconds; Total Duration of Fixations (TDoFs), which is the overall duration of fixations to an AoI measured in milliseconds; Number of Fixations, which is the overall number of fixations to an AoI; Fixation Duration, which is the duration of a single fixation measured in milliseconds; Fixation rate, which is the number of fixations divided by a period such as the duration of a trial in seconds, giving the number of fixations per second; Coverage Ratio, which indicates the percentage of the area covered by the participants' gaze within a specific AoI; Number of Visits, which is the overall number of visits to an AoI; the order of the visits to AoIs in a participant's eye gaze path through AoIs (also known as scanpath).

## IV. Preliminary Results and Future Analyses

We conducted a preliminary study recruiting participants from the campuses of the two universities. Both the number of final participants and the number of visual stimuli (emails) were filtered. In particular, some participants were excluded due to low values in the Gaze Samples metric, calculated by dividing the number of correctly identified eye-tracking samples by the theoretical maximum given by the sensor sampling frequency. Spectacle wearers shows a low percentage of correctly identified samples, indicating the need for future studies to using eye-tracking devices with integrated prescription lenses. Additionally, emails with minimal visual content were found to be a challenge, as the software was unable to map gaze points to the stimulus image coordinate system, resulting in the inability to extract the data for these emails. Future studies should select visual stimuli appropriately to avoid this problem.

As a result, the sample consists of 28 participants, of which 15 identified as non-experts and 13 as experts. These participants were exposed to 12 emails, of which 5 were legitimate and 7 were phishing.

This study confirms the better phishing classification performance in participants identified as experts. Furthermore, this study also confirms the use of different email inspection methods among experts and non-experts. In particular, we found that experts focused more on the header than non-experts, in terms of total duration of fixations, indicating increased attention and awareness towards phishing indicators in this area (see Fig. 1).

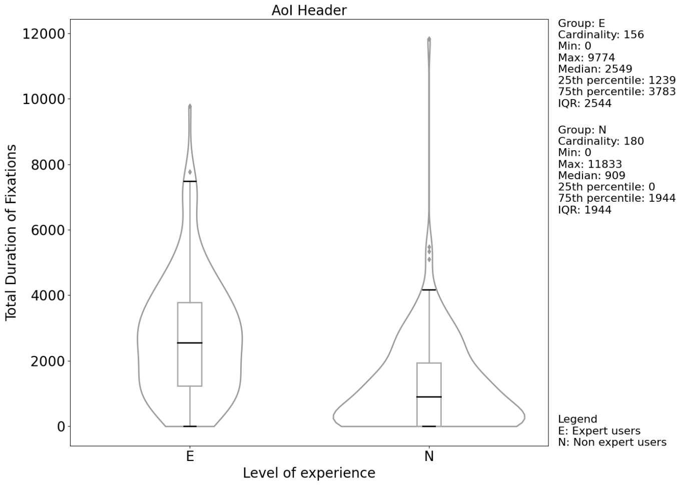

Fig. 1. TDoFs on the header between the two group of participants

This preliminary study showed the potential of eye-tracking technology in phishing research and the need to use a multidisciplinary approach. Futhermore, this study highlighted the challenges and issues that can be encountered during the research process, such as data quality issues, the choice of appropriate metrics and stimulus. The accumulated knowledge from this study offers a solid foundation for future research, which should involve more participants and a more diverse range of email scenarios, as well as consider smaller AoIs that capture individual elements in the header and body areas. In the future, a better understanding of the phishing phenomenon from a cognitive point of view could lead to the necessary insights to improve training courses and develop more intuitive email client interfaces for spotting the phishing.

# References

[1] APWG, “Phishing activity trends report: 3rd Quarter 2022,” 2022-12. [Online]. Available: https://docs.apwg.org/reports/apwg_trends_report_q3_2022.pdf

[2] IBM Security, “Cost of a Data Breach Report 2022.” [Online]. Available: https://www.ibm.com/downloads/cas/3R8N1DZJ

[3] Internet Society, “Policy Brief: The Challenge of Spam,” 2015-10-30. [Online]. Available: https://www.internetsociety.org/policybriefs/spam/

[4] McAfee and ICF International, “The Carbon Footprint of Email: Spam Report 2009,” 2009. [Online]. Available: https://www.siskinds.com/wp-content/uploads/carbonfootprint_12pg_web_rev_na-1.pdf

[5] Mimecast, “The State of Brand Protection Report 2020.” [Online]. Available: https://www.mimecast.com/state-of-brand-protection/

[6] B. Holst-Christensen and E. Frøkjær, “Security issues in SMTP-based email systems,” in *2021 14th CMI International Conference - Critical ICT Infrastructures and Platforms (CMI)*, 2021, pp. 1–6.

[7] V. Riabov, “SMTP (simple mail transfer protocol),” in *Handbook of Computer Networks*, 2007-12, pp. 388–406.

[8] B. Stone-Gross, T. Holz, G. Stringhini, and G. Vigna, “The underground economy of spam: A botmaster’s perspective of coordinating large-scale spam campaigns,” in *Proceedings of the 4th USENIX Conference on Large-Scale Exploits and Emergent Threats*, ser. LEET’11. USENIX Association, 2011, p. 4.

[9] L. Allodi, T. Chotza, E. Panina, and N. Zannone, “On the need for new antiphishing measures against spear-phishing attacks,” *IEEE Security & Privacy*, vol. 18, no. 2, pp. 23–34, 2020.

[10] L. Gallo, A. Maiello, A. Botta, and G. Ventre, “2 Years in the anti-phishing group of a large company,” *Computers & Security*, vol. 105, p. 102259, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167404821000833

[11] L. Gallo, A. Botta, and G. Ventre, “Identifying threats in a large company’s inbox,” in *Proceedings of the 3rd ACM CoNEXT Workshop on Big DAta, Machine Learning and Artificial Intelligence for Data Communication Networks*, ser. Big-DAMA ’19. Association for Computing Machinery, 2019, pp. 1–7. [Online]. Available: https://doi.org/10.1145/3359992.3366637

[12] P. Kumaraguru, J. Cranshaw, A. Acquisti, L. Cranor, J. Hong, M. Blair, and T. Pham, “School of phish: A real-world evaluation of anti-phishing training,” in *Proceedings of the 5th Symposium on Usable Privacy and Security*, ser. SOUPS 2009 - Proceedings of the 5th Symposium On Usable Privacy and Security, 2009-07.

[13] S. Sheng, M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs, “Who falls for phish? a demographic analysis of phishing susceptibility and effectiveness of interventions,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’10. New York, NY, USA: Association for Computing Machinery, 2010, p. 373–382. [Online]. Available: https://doi.org/10.1145/1753326.1753383

[14] F. Quinkert, M. Degeling, and T. Holz, “Spotlight on phishing: A longitudinal study on phishing awareness trainings,” in *Proceedings of the 2021 International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, 2021.

[15] J. Wang, T. C. Herath, R. Chen, A. Vishwanath, and H. R. Rao, “Phishing susceptibility: An investigation into the processing of a targeted spear phishing email,” *IEEE Transactions on Professional Communication*, vol. 55, pp. 345–362, 2012.

[16] D. D. Caputo, S. L. Pfleeger, J. D. Freeman, and M. E. Johnson, “Going Spear Phishing: Exploring Embedded Training and Awareness,” *IEEE Security & Privacy*, vol. 12, no. 1, pp. 28–38, 2014-01. [Online]. Available: http://ieeexplore.ieee.org/document/6585241/

[17] D. Lain, K. Kostiainen, and S. Čapkun, “Phishing in organizations: Findings from a large-scale and long-term study,” in *2022 IEEE Symposium on Security and Privacy (SP)*, 2022, pp. 842–859.