The Peculiar Case of Tailored Phishing against SMEs Detection and Collective Defense Mechanisms at a Small IT Company

DECISION: accept

AUTHORS: Pavlo Burda, Abdul Malek Altawekji, Nicola Zannone and Luca Allodi

Summary of Reviews

- Review 1: 2 (4)
- Review 2: 1 (3)
- Review 3: 1 (3)

Reviews

Review 1

TOTAL SCORE: 2 (accept)
Overall evaluation: 2 (accept)
Reviewer's confidence: 4 (high)

The paper presents a study on the effect, reaction and perception of spearphishing against a SME organisation. The study involved real phishing attacks to a real company, affecting 30 employees. The assessment was supplemented with semi-structured interviews of nine selected individuals after the phishing attacks were conducted. Results are promising, indicating SME reacted better and faster to spearphishing than larger enterprises.

Overall the paper is well written, with a sound structure, clear research questions and methodologies.

Comments and questions:
- In the introduction, the authors make reference to the importance of SME in developing countries and the lack of resources they have. However, the study focuses on a Dutch SME. The reference to developing countries seems exaggerated and introduces bias to the reader as SMEs from developing countries are not being studied.
- The oversight of using the appropriate language (english vs dutch) in the spearphishing campaigns impersonating the CEO can be a strong bias in the experiment and the results. The choice of language and the effect on the study should be discussed further.
- The experiment duration (lasted one day) seems short. Maybe 3 to 5 people receiving the same spearphishing in the same hour? Authors should discuss whether this timeframe seems real and whether expanding this timeframe would have produced less sharing among individuals.

Review 2

TOTAL SCORE: 1 (weak accept)
Overall evaluation: 1 (weak accept)
Reviewer's confidence: 3 (medium)

This paper sets out to evaluate the effectiveness of tailored phishing (also known as spear phishing) on small and medium enterprise. This is a gap in current literature as most research has focused on evaluating tailored phishing

against larger corporations. For this paper, the author(s) crafted an experiment with 30 employees from a SME IT company in the Netherlands. The 30 employees were randomly assigned into two conditions. Employees in the first condition (n = 15) received the tailored phishing email, which included personal and professional information (e.g., place of residence and years in current company) that were publicly available. Employees in the second condition received a generic phishing email. Emails in both conditions impersonated the CEO of the company and included a link to a well-known website for trips in the Netherlands. Overall, only eight out of the 30 employees interacted with the phishing emails. In addition, a warning email was sent out to all employees within 10 minutes of the phishing experiment. The author(s) concluded that results from the experiment was inconclusive. However, to further understand employees' interaction and reaction to the phishing emails, the author(s) conducted a semi-structured interviews with nine employees to addressed four topics: a) security awareness, b) rational and emotional response to the phishing emails, c) emotional drive that led to reporting behavior, and d) the behavior towards the tailored phishing emails. Of the nine employees, three of them received the tailored email and five received the generic phishing email. One interviewee did not participate in the experiment but is the system administrator for the company. Via thematic analysis, the author(s) identified the cognitive steps (e.g., detection, reaction, and action) taken by employees and factors affecting these steps (e.g., awareness, tailoring effects, and reporting). This paper contributes to the current literature via its innovative approach in methodology and sample selection.

1) Overall, the author(s) provided detailed description and explanation for the mixed-method approach of this paper. The author(s) described the steps and consideration taken in crafting the phishing emails, as well as the approach to coding the transcripts from the interviews. The author(s) also provided a great overview on how qualitative coding was carried out, which is not an easy task.

2) The author(s) included a section on ethical consideration within the paper. This is especially important since interviews were utilized for data collection.

3) In "Related work", here are some suggestions for the author(s):
• The significance of Vahdad's research needs to be elaborated. For example, which target-related variables did their research uncover that improved the effectiveness of phishing attacks. This will help provide some contexts in the target-related variables that were utilized in crafting the tailored phishing emails in this paper.
• The sentence "Extant research on phishing reporting highlighted the …" ended with "factors behind this [22]". It is unclear what "this" was referring to as several factors associated to reporting behaviors were mentioned prior to that.
• Since the author(s) relied on citation [15] for comparison, the author(s) should summarize the context and findings from the citation. It will help situate the current paper.

4) In "Research gap and contribution", the author(s) should use the word "including" instead of the abbreviation.

5) In Section 3.4.3, the author(s) should consider including information on the number of coders that were involved in the process and whether any software/tools were used to assist with the coding process. In addition, the citation after the first sentence in this section should be citation [31] and not citation [30].

6) In Section 4.2.1 and the discussion on identified themes in Table 3, the author(s) should consider addressing the collapse of some topic into categories. For example, under "Tailoring effects", there are two topic that was mentioned only once but the significance of these instances was not illustrated (either via the inclusion of the quote or summary of the quote). Another suggestion is the transition between "Awareness" and "Detection". Specifically, the last quote in the sub-section "Awareness" was a quote that fall under the theme of "Detection. It is unclear why the quote was mentioned under the theme "Awareness". Another confusion is the categorization of "Link investigation". The author(s) might wish to explain why it was collapsed into the theme "Action" instead of the category "Technical" within the theme "Detection", especially since hovering over a link provide technical clues.

7) In Section 4.2.2, there was an over-reliance on long quotes. One suggestion would be utilizing shorter excerpts of data and engaging in deeper analysis that connects the contributing factors to the cognitive steps. For example, with the contributing factor "Reporting (expectations & reasons), the author(s). provided a long quote but did not elaborate the significance of the quote, especially given the number of categories for this theme.

Review 3

TOTAL SCORE: 1 (weak accept)
Overall evaluation: 1 (weak accept)
Reviewer's confidence: 3 (medium)

The paper looks into targeted phishing attacks on small-medium enterprises. Specifically, the authors launched a phishing campaign on a smaller company and then further investigated the success of the campaign using semi-structured interviews with the employees. The phishing campaign was not very successful - only a handful of employees interacted with the phishing link, mostly to investigate. The main reasons for the failure of the campaign, as identified through the interviews, were the expectation mismatch (the email didn't sound right) and the network effect (people warning each other).

I like that the paper focuses on smaller companies which are often overlooked in similar research. The related works section as well as the findings clearly showcase that smaller companies are different than big enterprises and why they should be investigated separately. I could see how the main findings of the study could be useful in practice and help inform the appropriate actions, both for small companies and policymakers. I believe the authors of the study sufficiently addressed ethical considerations. I also appreciate that the authors provided the full email text and interview questions in the appendix. Overall, I think the paper is interesting, clearly written and easy to read.

The following section contains some of my concerns in order of importance.

Firstly, I am concerned about the external validity of the findings. The paper partly addresses the small number of employees included in the phishing campaign, however, it also mentions that the company mostly consists of highly educated and technical staff. I wonder if the same experiment performed on a number of different small companies would lead to the same results. Further, among the threads to validity, two warnings that invalidated the results are mentioned. It is not completely clear if this means that people were warned that an experiment is happening or that a potential phishing is happening (either by google or by other employees). If it's the former, I agree with the validity claim. However, if it's the latter, wouldn't that actually be a part of the results (potential victims of a real phishing attack might be warned in the same way), so not really invalidating the results?

In terms of presentation and analysis, I would have wanted to see more distinction between the two treatments (tailored vs. non-tailored). I realise that the number of people that interacted with the phishing link is too small for any further analysis, but it would be nice to at least see how many of those belonged to each of the treatments. Without this, RQ1 is not fully answered. Similar holds for the interviews, it would be nice to see how the identified categories/themes correlate with clicking on the link or the treatment.

The study recorded 8 interactions with the links. I wonder to what extent this is the result of a very short campaign (1 day). It is not clear what happened after 1 day - I assume the links went down and interactions were not recorded anymore. I wonder how many people didn't even read the email in the time window and would have clicked on the link if the campaign continued longer. Additionally, the "victims" were informed about the experiment in the debriefing page. The page explicitly asks the victims not to talk to others about the experiment, although I imagine that not everyone respected the request. I wonder if that also partially led to the low success of the campaign.

Finally, a minor clarification request: In the beginning of second paragraph of section 4.1 you say: “We can observe that eight out of 30 employees who received the phishing email interacted with it. Specifically, five employees clicked on the link in the phishing email. Among these employees, two provided fake login details ”
The phrasing here suggests that out of 5 employees that clicked on the link, 2 provided fake details. However Table 1 and the fact that 8 people interacted with the link suggest that 5 people only clicked on the link, and two clicked on the link and then filled in random address. Which one is correct?

Overall, I believe the work addresses and interesting questions and provides some valuable insights. It could be improved in terms of clarify, analysis and presentation. Thus I recommend a weak accept.