12-12-2021

Bait the hook to suit the phish, not the phisherman: A field experiment on security networks of teams to withstand spear phishing attacks on online social networks

Robert Lamprecht
University of Innsbruck, robert.lamprecht@student.uibk.ac.at

Andreas Eckhart
University of Innsbruck

Ryan T. Wright
University of Virginia

Bait the Hook to Suit the Phish, not the Phisherman: A Field Experiment on Security Networks of Teams to Withstand Spear Phishing Attacks on Online Social Networks

Robert Lamprecht[1]
Faculty of Business and Management, University of Innsbruck,
Innsbruck, Austria

Andreas Eckhart
Faculty of Business and Management, University of Innsbruck,
Innsbruck, Austria

Ryan T Wright
McIntire School of Commerce, University of Virginia,
Charlottesville, USA

ABSTRACT

In this paper, we present our research in progress of a field experiment conducted to observe the impact of collective security behavior of teams when being targeted with a spear phishing attack on online social networks. To observe the shaping of security networks in teams, fifteen different honeypot profiles were created to send spear phishing messages after an initial bonding of eight weeks to the target group of 76 people. The experiment simulated a regular communication on online social networks of three teams of an international organization. The team members were entangled in personal and individual chats on an online social network to later react to an unexpected and unforeseen spear phishing message. As previous research has shown, various aspects influence the spear phishing susceptibility, but the collective security behavior has currently been neglected. This work plans to evaluate how security networks are being formed, the factors relevant to shape those networks and efforts to protect against spear phishing attacks.

Keywords: spear phishing, social engineering attacks, social network, group context, collective security

[1] Corresponding author. robert.lamprecht@student.uibk.ac.at

INTRODUCTION

Lacking cybersecurity is among the highest risks in the next ten years in terms of likelihood and impact. Failures from Cybersecurity remain one of the top ten global risks which impact our society and our global ecosystem (World Economic Forum 2021). But risks resulting from the cyber domain are an evolution, but not a revolution. With regards to social engineering in particular, phishing (Bose and Leung 2007; Wright et al. 2010) is still attack mode number one.

In phishing research, several studies with various research foci explored phishing and the way people react when receiving seemingly legitimate emails (Abbasi et al. 2021; Aleroud et al. 2020; Caputo et al. 2013; Jensen et al. 2017; Martin et al. 2019; Wright et al. 2014). Predominately these studies dealt with the individual aspects of phishing, using traditional means of electronic communication such as email messaging, neglecting the developments of new forms of digital communication. Along with this, also research dealing with developments of new technology platforms, such as social media, aside from traditional desktop or laptop devices and their influence on the success of phishing attacks is rare (Anti-Phishing Working Group 2021; Jensen et al. 2017). Within this research in progress, we present an approach to analyze the creation and formation of security networks and their effectiveness when resisting spear phishing attacks performed on social networks.

The following sections present the background of spear phishing and related social phishing research, followed by the methodology for our field experiment. The paper concludes with the current progress, the status and the potential implications form this work.

RESEARCH BACKGROUND

The social influence phenomenon is one of the key concepts to direct, instruct, coordinate and influence other members of a certain species (Cialdini et al. 1990, 1991). Several researchers address this phenomenon and the respective areas such as conformance (Nail and Ruch 1992), obedience, a form of social influence that requires performing an activity under the order of an authority (Milgram 1963), or persuasion or attitude change (change in response to a message) (Hovland et al. 1953), compliance (Asch 1956) (change in response to an explicit instruction) , social forces (change in response to the structure of the social situation) (Milgram 1967) and help (change in response to someone´s need) (Brehm 1966). The work of Asch and the subsequent research demonstrated that the judgements of one person can be influenced by the judgments of others (as social norms). Two motives were identified (1) informational social influence through the acceptance of the social norm as valid information and (2) normative social influence through reward for compliance with the social norm or penalty of noncompliance (Deutsch and Gerard 1955).

The research work on social influence by Robert Cialdini introduces six core principles of influence which are reciprocity, scarcity, consistency, authority, liking and social proof (Cialdini 2009). Current phishing research in the field of Information Systems (IS) addresses several of Cialdini's principles (Oliveira et al. 2017; Parsons et al. 2019; Taib et al. 2019; Vishwanath et al. 2011; Williams et al. 2018). According to researchers (Caputo et al. 2013), spear phishing is defined as type of attack in the domain of cyberthreats which attempts to get unauthorized access to users or organizations IT systems for the threat actors' benefit. Access is usually obtained by crafting seemingly legitimate email messages which impersonate trust relationships (e.g., using

the power of authority) and networks inside a company or a group of people, which actively work together for a dedicated context (Frauenstein and Flowerday 2020).

Phishing and spear phishing research aims to understand how the people's perception of susceptibility changes or adjusts over time, as the phishing attacks are evolving and becoming more sophisticates for the recipients (people) to detect (Chen et al. 2020). Although, current literature agrees upon the importance of fighting spear phishing, research synthesizing the current state of knowledge on spear phishing is scarce so far (Hong 2012). The existing approaches are fragmented, uncovering gaps in its definition and configuration (Parsons et al. 2019; Silic and Back 2016; Silic and Lowry 2020). The current definitions primarily focus on email communication as the sole means of electronic communication to interact with the victim.

While existing research very much focuses on electronic communication such as email, it somewhat neglects to consider literature on the social media channels. Literature acknowledges this gap calling for future research to explore security awareness trainings in other contexts aside from traditional phishing by adapting the training techniques for encouraging appropriate behavior in other IT security settings such as information disclosure on social media (Jensen et al. 2017; Parsons et al. 2019; Silic and Back 2016).

Many researchers have examined how individual users can be trained to identify and detect traits of phishing and successfully avoid being a victim (Albladi and Weir 2020; Caputo et al. 2013; Jampen et al. 2020; Jensen et al. 2017; Puhakainen and Siponen 2010), however less have examined the group aspect by means of human relation and how a group or collective of people (Mattke et al. 2020) could act against external phishing attacks in general, and against spear phishing attacks in the social world (Algarni et al. 2015; Benenson et al. 2017; Choi et al. 2015; Jiang et al. 2013; Krasnova et al. 2010; Oliveira et al. 2017; Yazdanmehr et al. 2020), in

particular (Jagatic et al. 2007; Wright et al. 2020). Therefore, we formulate the following research question: How do security networks evolve in an organization and what is their impact, effect and characteristic when withstanding spear phishing attacks performed on online social networks?

METHODOLOGY

Spear Phishing Field Experiment Design

For this work, we conducted a randomized mock spear-phishing field experiment in three different teams of an international organization. The field experiment used the online social network (OSN) Instagram and reflects the approach of Wright et al (2014) to answer our research question. The organization is keen to raise the awareness of the different teams and to prepare them to cope with the risks resulting from the use of OSN.

As part of this preparation, 15 honeypot Instagram profiles were prepared with different cover stories and were operated over a period of six months to create seemingly legitimate online content. Thereafter, the honeypot profiles tried to start interacting with the target profiles (list is provided by the organization). The aim of the honeypot profiles was to establish trust (Cialdini et al. 1991), to influence targets for non-conformance (Nail and Ruch 1992) and non-compliance with the organization´s policies (Asch 1956) and to provoke a behavioral change (Milgram 1967). One team received a dedicated security awareness training to address the social media risks and to measure the effectiveness and memory of this intervention.

Prior the start of the spear phishing field experiment, the social network profiles of the teams were analyzed to retrieve public available insights into the organization. Data of all public Instagram profiles was automatically collected, using a custom-developed scrape engine,

including, but not limited to posts, likes, comments and answers. This data provided insights into the team exposure (e.g., content shared, information disclosure) and the shaped online network of the teams which acted as a starting point to identify possible exposed profiles.

Interaction within the field experiment was solely done on Instagram like getting connected with the team members, messaging with them and to exfiltrate information from the team members. Before the end of the field experiment, which lasted eight weeks, all target profiles received an Instagram phishing message, either from a source they already know (i.e., one of the honeypot profiles) or from an unknown source (i.e., an Instagram page which publishes random pictures). Sources of phishing messages were assigned randomly but were equally distributed among the three teams.

To measure phishing susceptibility, we tracked the message seen notification for every target on Instagram. As this measure can be deactivated in the Instagram privacy settings, we coded a unique number for every target into the link prior sending it on Instagram. Every click on the link created a log entry, if the target clicked on the link and no entry, if the link was not clicked. Recent phishing research also used this objective measure (Wright et al. 2014).

Before the SETA measure, which uncovered the background of the field experiment to the target audience, the participants in the field experiment completed a post-experiment-survey that discusses the advice network (Cross et al. 2001), the professional network (Mertens et al. 2020), the access network (Cross et al. 2002), the advice confirmation network (De Lange et al. 2004) and the friend network (Criado and Such 2015) when needing help to deal with challenges in the online social network Instagram. The survey furthermore discusses spear phishing self-efficacy (Schuetz et al. 2020; Wang et al. 2016) on online social networks (Frauenstein and Flowerday 2020), common ways to detect attacks, and gathered control variables that were

included in the analysis. The invitation to participate in the SETA measure was masqueraded as a general training, to avoid any bias when answering the survey.

The results of the spear phishing field experiment were shared with the management of the organization for use and to understand the need for training and other interventions of exposed team members. To explore the effectiveness of the spear phishing attacks, we gathered the honeypot profiles used in the attacks, rated the manipulation techniques used in every profile and then examined the effectiveness of the profiles across the tree teams.

NEXT STEPS

Currently, the data from this case, including the spear phishing messages, the post-field experiment survey and the OSN data has been collected and is ready to be analyzed. The next stage of the research will deal with the data analysis using R, Gephi and a custom-developed Python analysis code.

For example, we suggest that the victim´s exposure depends on the position inside a network. A general analysis, using the phishing funnel (Abbasi et al. 2021), will compare how many individuals failed the spear phishing test who did not receive a treatment and compare it with those individuals who did receive a treatment.

For the spear phishing messages sent, data analysis will compare, if the time of the day has an impact on phishing susceptibility with the data received from the timestamps of the target´s activity on Instagram.

For each type of honeypot profile, the analysis will evaluate, if more targets were victims of the phishing attack. The analysis included a comparison of the target´s position (centrality) in the OSN Instagram with the position inside the team from the post-field experiment survey.

Finally, additional variables including demographics (e.g., age, position in the team, social media experience, education) and other details (e.g., time of day of phishing response) and their effect will be analyzed.

REFERENCES

Abbasi, A., Dobolyi, D., Vance, A., and Zahedi, F. M. 2021. "The Phishing Funnel Model: A Design Artifact to Predict User Susceptibility to Phishing Websites," *Information Systems Research* (32:2).

Albladi, S. M., and Weir, G. R. S. 2020. "Predicting Individuals' Vulnerability to Social Engineering in Social Networks," *Cybersecurity* (3:1), Cybersecurity.

Aleroud, A., Abu-Shanab, E., Al-Aiad, A., and Alshboul, Y. 2020. "An Examination of Susceptibility to Spear Phishing Cyber Attacks in Non-English Speaking Communities," *Journal of Information Security and Applications* (55:September), Elsevier Ltd, p. 102614.

Algarni, A., Xu, Y., and Chan, T. 2015. "An Empirical Study on the Susceptibility to Social Engineering in Social Net- Working Sites: The Case of Facebook.," *European Journal of Information Systems* (26:6), pp. 661–687.

Anti-Phishing Working Group. 2021. "Phishing Activity Trends Report 2nd Quarter 2021." (https://docs.apwg.org/reports/apwg_trends_report_q2_2021.pdf).

Asch, S. E. 1956. "Studies of Independence and Conformity: I. A Minority of One against a Unanimous Majority.," *Psychological Monographs: General and Applied* (70:9), US: American Psychological Association, pp. 1–70.

Benenson, Z., Gassmann, F., and Landwirth, R. 2017. "Unpacking Spear Phishing Susceptibility," *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (10323 LNCS), pp. 610–627.

Bose, I., and Leung, A. C. M. 2007. "Unveiling the Mask of Phishing: Threats, Preventive Measures, and Responsibilities," *Communications of the Association for Information Systems* (19:April).

Brehm, J. W. 1966. "A Theory of Psychological Reactance.," *New York, Academic Press,* Oxford, England: Academic Press.

Caputo, D. D., Pfleeger, S. L., Freeman, J. D., and Johnson, M. E. 2013. "Going Spear Phishing: Exploring Embedded Training and Awareness," *IEEE Security and Privacy* (12:1), pp. 28–38.

Chen, R., Gaia, J., and Rao, H. R. 2020. "An Examination of the Effect of Recent Phishing Encounters on Phishing Susceptibility," *Decision Support Systems* (133:September 2019), Elsevier, p. 113287.

Choi, B. C. F., Jiang, Z. J., Xiao, B., and Kim, S. S. 2015. "Embarrassing Exposures in Online Social Networks: An Integrated Perspective of Privacy Invasion and Relationship Bonding," *Information Systems Research* (26:4), pp. 675–694.

Cialdini, R. B. 2009. *Influence: Science and Practice (5th Edition)*, p. 272.

Cialdini, R. B., Reno, R. R., and Kallgren, C. A. 1990. "A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places.," *Journal of Personality and Social Psychology* (58:6), US: American Psychological Association, pp. 1015–1026.

Cialdini, R. B., Reno, R. R., and Kallgren, C. A. 1991. "A Focus Theory of Normative Conduct.Pdf," *Journal of Personality and Social Psychology* (58:6), pp. 1015–1026. (http://media.cbsm.com/uploads/1/AFocusTheoryofNormativeConduct.pdf).

Criado, N., and Such, J. M. 2015. "Implicit Contextual Integrity in Online Social Networks," *Information Sciences* (325), Elsevier Ltd., pp. 48–69.

Cross, R., Borgatti, S. P., and Parker, A. 2001. "Beyond Answers: Dimensions of the Advice

Network.," *Social Networks* (23:3), Netherlands: Elsevier Science, pp. 215–235.
Cross, R., Borgatti, S. P., and Parker, A. 2002. "Making Invisible Work Visible: Using Social Network Analysis to Support Strategic Collaboration," *California Management Review* (44:2), pp. 25–46.
Deutsch, M., and Gerard, H. B. 1955. "A Study of Normative and Informational Social Influences upon Individual Judgment.," *The Journal of Abnormal and Social Psychology* (51:3), US: American Psychological Association, pp. 629–636.
Frauenstein, E. D., and Flowerday, S. 2020. "Susceptibility to Phishing on Social Network Sites: A Personality Information Processing Model," *Computers and Security* (94), Elsevier Ltd, p. 101862.
Hong, J. 2012. "The Current State of Phishing Attacks," *Communications of the ACM* (55:1), pp. 74–81.
Hovland, C. I., Janis, I. L., and Kelley, H. H. 1953. "Communication and Persuasion; Psychological Studies of Opinion Change.," *Communication and Persuasion; Psychological Studies of Opinion Change.*, New Haven, CT, US: Yale University Press.
Jagatic, B. T. N., Johnson, N. A., and Jakobsson, M. 2007. "Social Phishing," *Communications of the ACM* (50:10), pp. 95–100.
Jampen, D., Gür, G., Sutter, T., and Tellenbach, B. 2020. "Don't Click: Towards an Effective Anti-Phishing Training. A Comparative Literature Review," *Human-Centric Computing and Information Sciences* (Vol. 10), Springer Berlin Heidelberg.
Jensen, M. L., Dinger, M., Wright, R. T., and Thatcher, J. B. 2017. "Training to Mitigate Phishing Attacks Using Mindfulness Techniques," *Journal of Management Information Systems* (34:2), pp. 597–626.
Jiang, Z., Heng, C. S., and Choi, B. C. F. 2013. "Privacy Concerns and Privacy-Protective Behavior in Synchronous Online Social Interactions," *Information Systems Research* (24:3), pp. 579–595.
Krasnova, H., Spiekermann, S., Koroleva, K., and Hildebrand, T. 2010. "Online Social Networks: Why We Disclose," *Journal of Information Technology* (25:2), pp. 109–125.
De Lange, D., Agneessens, F., and Waege, H. 2004. "Asking Social Network Questions: A Quality Assessment of Different Measures," *Metodološki Zvezki* (1:2), pp. 351–378.
Martin, S. R., Lee, J. J., and Parmar, B. L. 2019. "Social Distance, Trust and Getting 'Hooked': A Phishing Expedition," *Organizational Behavior and Human Decision Processes* (May 2018), Elsevier Inc.
Mattke, J., Maier, C., Reis, L., and Weitzel, T. 2020. "Herd Behavior in Social Media: The Role of Facebook Likes, Strength of Ties, and Expertise," *Information and Management* (57:8), Elsevier B.V., p. 103370.
Mertens, N., Boen, F., Steffens, N. K., Haslam, S. A., and Fransen, K. 2020. "Will the Real Leaders Please Stand up? The Emergence of Shared Leadership in Semi-Professional Soccer Teams," *Journal of Science and Medicine in Sport*, Sports Medicine Australia.
Milgram, S. 1963. "Behavioral Study of Obedience.," *Journal of Abnormal Psychology* (67:4), pp. 371–378.
Milgram, S. 1967. "The Small World Problem," *Psychology Today* (1:1), pp. 60–67.
Nail, P. R., and Ruch, G. L. 1992. "Social Influence and the Diamond Model of Social Response: Toward an Extended Theory of Informational Influence," *British Journal of Social Psychology* (31:3), pp. 171–187.
Oliveira, D., Rocha, H., Yang, H., Ellis, D., Dommaraju, S., Muradoglu, M., Weir, D., Soliman,

A., Lin, T., and Ebner, N. 2017. *Dissecting Spear Phishing Emails for Older vs Young Adults*, pp. 6412–6424.

Parsons, K., Butavicius, M., Delfabbro, P., and Lillie, M. 2019. "Predicting Susceptibility to Social Influence in Phishing Emails," *International Journal of Human Computer Studies* (128:July 2018), Elsevier Ltd, pp. 17–26.

Puhakainen, P., and Siponen, M. 2010. "Improving Employees' Compliance through Information Systems Security Training: An Action Research Study," *MIS Quarterly: Management Information Systems* (34:4), pp. 757–778.

Schuetz, S. W., Benjamin Lowry, P., Pienta, D. A., and Bennett Thatcher, J. 2020. "The Effectiveness of Abstract Versus Concrete Fear Appeals in Information Security," *Journal of Management Information Systems* (37:3), Routledge, pp. 723–757.

Silic, M., and Back, A. 2016. "The Dark Side of Social Networking Sites:Understanding Phishing Risks," *Computers in Human Behavior* (60), Elsevier Ltd, pp. 35–43.

Silic, M., and Lowry, P. B. 2020. "Using Design-Science Based Gamification to Improve Organizational Security Training and Compliance," *Journal of Management Information Systems* (37:1), pp. 129–161.

Taib, R., Yu, K., Berkovsky, S., Wiggins, M., and Bayl-Smith, P. 2019. "Social Engineering and Organisational Dependencies in Phishing Attacks," *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 11746 LNCS), Springer International Publishing.

Vishwanath, A., Herath, T., Chen, R., Wang, J., and Rao, H. R. 2011. "Why Do People Get Phished? Testing Individual Differences in Phishing Vulnerability within an Integrated, Information Processing Model," *Decision Support Systems* (51:3), Elsevier B.V., pp. 576–586.

Wang, J., Li, Y., and Rao, H. R. 2016. "Overconfidence in Phishing Email Detection," *Journal of the Association for Information Systems* (17:11), pp. 759–783.

Williams, E. J., Hinds, J., and Joinson, A. N. 2018. "Exploring Susceptibility to Phishing in the Workplace," *International Journal of Human Computer Studies* (120:April), Elsevier Ltd, pp. 1–13.

World Economic Forum. 2021. *The Global Risks Report 2021: 16th Edition*, (16th ed.), World Economic Forum.

Wright, R., Chakraborty, S., Basoglu, A., and Marett, K. 2010. "Where Did They Go Right? Understanding the Deception in Phishing Communications," *Group Decision and Negotiation* (19:4), pp. 391–416.

Wright, R., Johnson, S., and Kitchens, B. 2020. *A Multi-Level Contextualized View of Phishing Susceptibility*, pp. 1–60.

Wright, R. T., Jensen, M. L., Thatcher, J. B., Dinger, M., and Marett, K. 2014. "Influence Techniques in Phishing Attacks: An Examination of Vulnerability and Resistance," *Information Systems Research* (25:2), pp. 385–400.

Yazdanmehr, A., Wang, J., and Yang, Z. 2020. "Peers Matter: The Moderating Role of Social Influence on Information Security Policy Compliance," *Information Systems Journal* (30:5), pp. 791–844.