Journal of the American Medical Informatics Association, 26(6), 2019, 547–552
doi: 10.1093/jamia/ocz005
Advance Access Publication Date: 12 March 2019
Research and Applications

Research and Applications

Evaluation of a mandatory phishing training program for high-risk employees at a US healthcare system

William J Gordon,[1,2,3] Adam Wright,[1,2,3] Robert J Glynn,[2,4,5] Jigar Kadakia,[3] Christina Mazzone,[3] Elizabeth Leinbach,[3] and Adam Landman[2,3,6]

[1]Division of General Internal Medicine and Primary Care, Brigham and Women's Hospital, Boston, Massachusetts, USA, [2]Harvard Medical School, Boston, Massachusetts, USA, [3]Partners HealthCare, Boston, Massachusetts, USA, [4]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA, [5]Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA, and [6]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA

Corresponding Author: William J. Gordon, MD, MBI, 75 Francis St, Boston, MA 02115, USA (wjgordon@partners.org)

Received 8 November 2018; Revised 21 December 2018; Editorial Decision 27 December 2018; Accepted 21 January 2019

ABSTRACT

Objective: The study sought to understand the impact of a phishing training program on phishing click rates for employees at a single, anonymous US healthcare institution.

Materials and Methods: We stratified our population into 2 groups: offenders and nonoffenders. Offenders were defined as those that had clicked on at least 5 simulated phishing emails and nonoffenders were those that had not. We calculated click rates for offenders and nonoffenders, before and after a mandatory training program for offenders was implemented.

Results: A total of 5416 unique employees received all 20 campaigns during the intervention period; 772 clicked on at least 5 emails and were labeled offenders. Only 975 (17.9%) of our set clicked on 0 phishing emails over the course of the 20 campaigns; 3565 (65.3%) clicked on at least 2 emails. There was a decrease in click rates for each group over the 20 campaigns. The mandatory training program, initiated after campaign 15, did not have a substantial impact on click rates, and the offenders remained more likely to click on a phishing simulation.

Discussion: Phishing is a common threat vector against hospital employees and an important cybersecurity risk to healthcare systems. Our work suggests that, under simulation, employee click rates decrease with repeated simulation, but a mandatory training program targeted at high-risk employees did not meaningfully decrease the click rates of this population.

Conclusions: Employee phishing click rates decrease over time, but a mandatory training program for the highest-risk employees did not decrease click rates when compared with lower-risk employees.

Key words: information security, phishing, health information technology

BACKGROUND AND SIGNIFICANCE

Cybersecurity is an increasingly important component of the infrastructure of healthcare delivery organizations, and recent attacks on healthcare information systems have negatively impacted hospital operations, resulting in canceled clinical appointments and procedures, financial cost, and negative press.[1–8] Digital information systems are crucial for many aspects of clinical work in the United States and worldwide, and providing safe and effective care depends on our ability to secure these systems to the maximum extent possible.

Phishing, which the National Institute of Standards and Technology defines as the act of tricking individuals into disclosing personal information through deceptive means,[9] is a common threat vector targeting hospital employees. A successful phishing attack

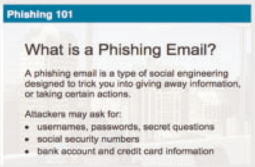
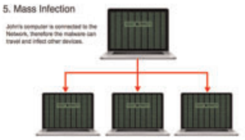
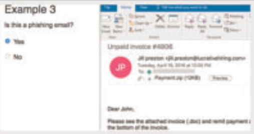
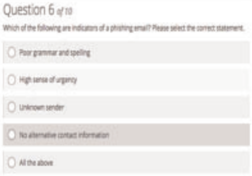
	Description	Example
Section 1: Overview of Phishing	• Definition of social engineering • Anatomy of a phishing email • Demonstration of a phishing attack • Impact of phishing • How to report a phishing email	 Phishing 101 What is a Phishing Email? A phishing email is a type of social engineering designed to trick you into giving away information, or taking certain actions. Attackers may ask for: • usernames, passwords, secret questions • social security numbers • bank account and credit card information
Section 2: Phishing scenario	• Identification of the target • Setting the bait • Hooking the target • Mass infection • Data compromise • Consequence of a Phishing attack	 5. Mass Infection John's computer is connected to the Network, therefore the malware can travel and infect other devices.
Section 3: Identifying a Phish	• Interactive examples of phishing emails	 Example 3 Is this a phishing email? ● Yes ○ No JP Dear John, Please see the attached invoice (.doc) and remit payment the bottom of the invoice.
Examination	• 10-question summary examination	 Question 6 of 10 Which of the following are indicators of a phishing email? Please select the correct statement. ○ Poor grammar and spelling ○ High sense of urgency ○ Unknown sender ○ No alternative contact information ○ All the above

Figure 1. Overview of phishing training program.

might allow an attacker to steal someone's username and password, and with that information, log in to clinical systems, write prescriptions, or steal data. Phishing emails can be indiscriminate and sent to large groups of employees or they can target specific individuals (eg, senior management), a technique commonly called "spear phishing."[10] Though there are numerous strategies to mitigate phishing risk—for example, training employees, automatically detecting phishing emails before they are delivered, or blocking commonly used phishing accounts, to name only a few[11–17]—limited evaluations of effectiveness have occurred and phishing remains an important cybersecurity risk for healthcare organizations.[5]

To address this risk of email as a phishing vector, numerous organizations have begun deploying phishing simulation programs, in which employees are sent fake phishing emails (often based on legitimate phishing emails). Employees that click then receive real-time, brief phishing education, usually through a website that comes up after clicking on the phishing email.[18,19] Many institutions have implemented such programs, and we have been working with multiple institutions across the country to better understand phishing programs and strategies to reduce risk. One of these institutions was willing to share employee-level data. This institution had started a program in 2015, and through this program, found that many employees continued to click on phishing emails, despite repeated simulated failures with educational content provided after clicking. In 2017, the information security officers at this institution, in conjunction with hospital leadership, decided to target a mandatory, detailed phishing awareness and training program for all employees that had clicked on at least 5 phishing simulations during the course of their employment. We sought to evaluate the impact of this mandatory training program and to understand whether such a program would improve simulated phishing click rates for these "high-risk" employees.

MATERIALS AND METHODS

Study design and setting

We partnered with a tertiary care academic medical center that was willing to share employee-level data. This institution has opted to remain anonymous due to the sensitive nature of this work, though the institution's identity was available to study authors and journal editors.

This institution had partnered with a third-party vendor (Cofense [formerly PhishMe]) since 2015 to manage the email delivery and tracking of their phishing simulation campaigns (in which a campaign is generally defined as 1 email sent over a fixed period of time to a large group of employees).[20] This vendor also provides the initial training material that is viewed upon clicking on a phishing email. All employees that clicked on at least 5 emails before campaign 16 were required to undergo phishing awareness training; this population was labeled "offenders."

After completion of the 15th campaign, all offenders were notified via email that they were being targeted for intensive information security training given their prior susceptibility to phishing simulation. They were enrolled in an online course (offered through a third-party online learning management system [HealthStream]).[21] This course was created by the information security and privacy team at their institution. The training program consisted of 3 main sections: (1) an overview of phishing, (2) a phishing scenario, and (3) how to identify a phishing email (Figure 1). Finally, to complete the training program, each employee had to pass a 10-question test on the material presented in the online video (Figure 1). They could retake this test as many times as needed to pass. Network access was suspended until employees successfully completed the training and passed the exam.

Measurements

We collected data for all employees that had received simulated phishing emails since 2015. For each employee and each phishing "campaign," we collected the date of that campaign and whether or not the employee clicked on the email.

To ensure a consistent email recipient population, we limited our analysis to those employees that had completed all 20 campaigns. All employee emails were anonymized before analysis.

Statistical analysis

We calculated descriptive statistics including overall click rates and click rates per campaign. We calculated the total number of clicks by employee across all 20 campaigns and created a frequency histogram with these results. We segmented our population into 2 groups, offenders and nonoffenders, in which offenders had clicked on at least 5 phishing campaigns at any point during their employment, before campaign 16. All offenders were offered a training intervention, whereas nonoffenders were not offered training. We calculated click rates for each group, before and after training was offered to the offenders, to assess trends in overall click rates in these 2 populations.

All analyses were conducted using R (version 3.5.1 for Statistical Computing, Vienna, Austria). The Partners HealthCare Institutional Review Board approved this study.

RESULTS

Since the launch of the phishing training program in July 2015, 20 campaigns were sent, every 2-3 months, until May 2018. A total of 35 580 unique employees received 390 908 emails over the course of the intervention period; 5416 of these employees received all 20 campaigns, and 772 clicked on at least 5 emails and were classified as "offenders." A total of 740 employees completed the mandatory phishing training program (Figure 2). Supplementary Table S1 has the number of emails sent, number of emails clicked, and calculated click rates for each campaign. None of the simulated phishing emails sent were "spear phishing" emails.

We then calculated the total number of clicks per employee over the 20 campaigns (Figure 3). In our set, only 975 (17.9%) employees clicked on 0 phishing emails over the course of the 20 campaigns; 3565 (65.3%) clicked on at least 2 emails during this time period and 90 (1.6%) clicked on at least 10.

We then looked at per-campaign click rates, split by offenders and nonoffenders (Figure 4). We included the small set of offenders who did not complete the training (n = 32) in the offender group, as they still received the emails informing them they were offenders, as well as numerous follow-up emails. Across the 20 campaigns, there was marked variation in click rates, with concordant spikes in more challenging campaigns (6, 8, and 17), and nadirs in less challenging campaigns (5, 7, 16, and 19), with click rate as a proxy for how challenging a campaign was. Overall, there was a clear decrease in click rate for each group over the 20 campaigns, particularly for the offenders. Despite the overall decrease, there was marked variability in click rates per campaign, which suggests that there are campaign-specific factors that strongly influence click rates. The mandatory training program, initiated after campaign 15, did not have a substantial impact on click rates, and the offenders remained more likely to click on a phishing simulation on campaigns 16-20 despite the notification that they were being targeted for mandatory training and undergoing the mandatory online training. Campaign 17, for example, still had an offender click rate of almost 25%, despite the mandatory training program. Examination of all employees, including those that did not participate in all 20 campaigns (eg, they were not employed for the entire intervention period) showed similar results (Supplementary Figure S1).

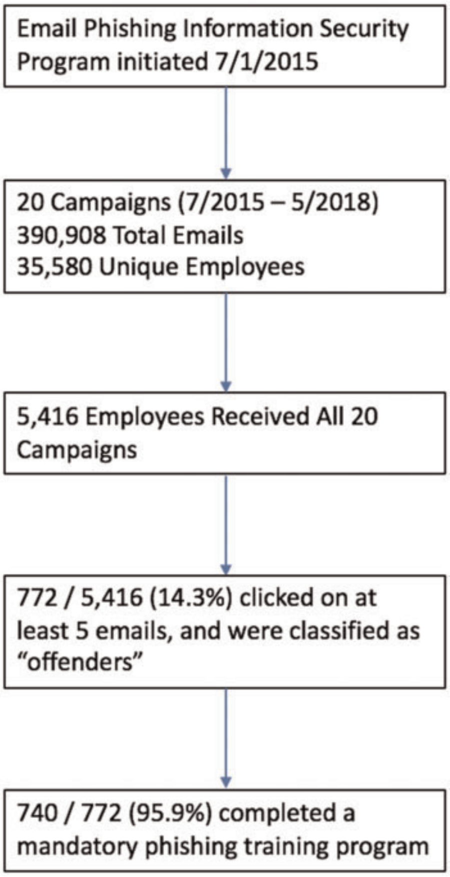

Email Phishing Information Security Program initiated 7/1/2015
20 Campaigns (7/2015 – 5/2018)
390,908 Total Emails
35,580 Unique Employees
5,416 Employees Received All 20 Campaigns
772 / 5,416 (14.3%) clicked on at least 5 emails, and were classified as "offenders"
740 / 772 (95.9%) completed a mandatory phishing training program

Figure 2. Study design and data acquisition.

DISCUSSION

In this study, we examine a phishing simulation information security program at 1 institution and the effect of a mandatory training program for employees that repeatedly clicked on simulated phishing emails. We find that phishing click rates are alarmingly high, but generally improve with repeated simulated phishing campaigns. Importantly, the mandatory training program for employees who clicked on 5 or more simulated phishing campaigns itself did not have a meaningful impact on click rates—the "offenders" remained more likely to click on phishing emails than nonoffenders did, with click rates between 10% and 25% post-training. The real-time training (provided after an employee clicked on any phishing simulation) may have been more effective at narrowing the gap between offenders and nonoffenders given the rapid decrease in click rates after the first several campaigns, particularly for the offenders. Because it only takes 1 successful phishing attack to cause substantial damage to a healthcare organization, these click rates (with or without training) are highly concerning.

There are several ways organizations can reduce phishing risk. First, there are technological solutions. Numerous vendor products exist to detect or reject suspicious inbound phishing emails, for example, based on machine learning rules of suspicious content or a database of flagged IP addresses. Additional email-based technical

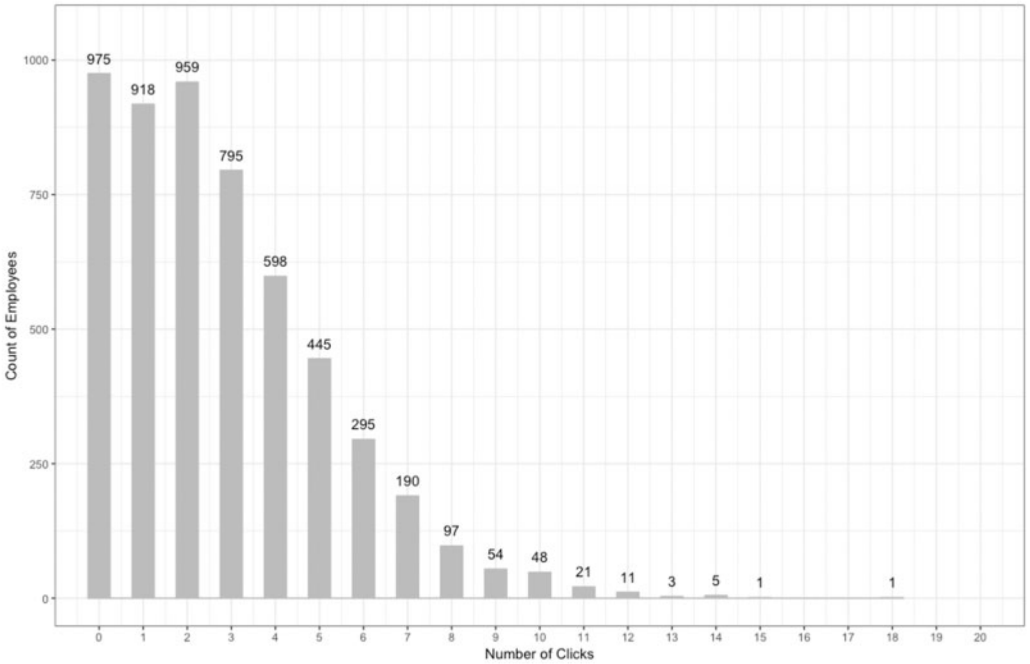


Figure 3. Distribution of number of clicks on simulated phishing emails, per employee for employees that received all 20 simulated phishing campaigns (5416 total employees).

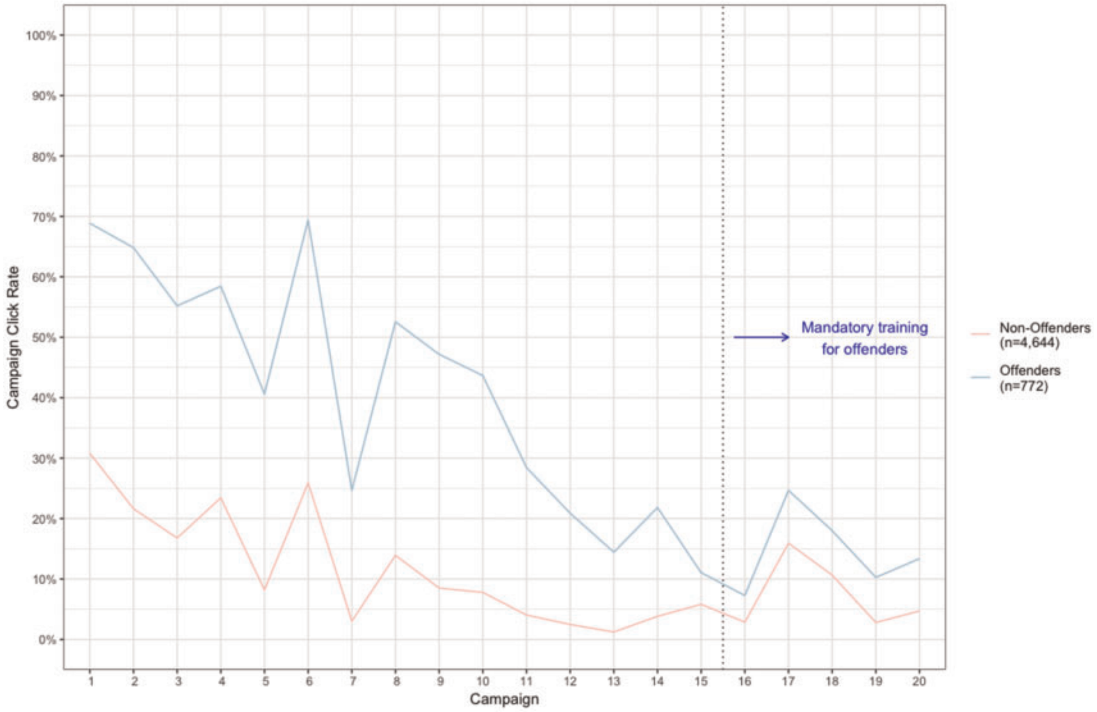


Figure 4. Click rates by campaign number, offenders vs nonoffenders, for employees that completed all campaigns. Shown are the click rates per campaign for employees that participated in all campaigns.

solutions include STARTTLS (which encourages encrypted emails), authentication techniques like Sender Policy Framework and Domain Keys Identified Mail, and flagging all noninternal emails with phrases like "[EXTERNAL]." The US Department of Homeland Security, for example, issued a Binding Operational Directive in October 2017 requiring many of these technologies to be in place across the federal branch, and the National Institute of Standards and Technology has also published technical guidance for improving email security.[22,23] Additional technical interventions include blocking actual phishing websites (preventing an employee from accessing a suspicious website) and 2-factor authentication, which requires employees to have a password and another piece of information (eg, a passcode from a mobile phone) to login, so that only knowing someone's password will not allow you to access systems under their credentials.

While technical solutions can mitigate risk, if a phishing email makes it through these barriers, it is up to the individual receiving the email to engage. User awareness, education, and training is crucial, and prior work has shown some benefit to phishing simulation and subsequent training.[24] The mandatory employee training program studied here was initiated after a multiyear simulation effort was already in place and click rates had already started to plateau. It is possible that up-front training (eg, after the first click or simply requiring it of all users) would reduce click rates more quickly. Additionally, it is possible that the format of the training program—online learning, a test that could be retaken as many times as needed—is not the most effective method of providing phishing training, and a different intervention, like an in-person training exercise, would have greater impact. Finally, because there is some educational content provided after an employee clicks on a phishing simulation email, it is likely that the offenders had already received some baseline education before the mandatory training program.

Our study has several limitations. First, this is a single-center study, and it is possible that other institutions, with different information security programs and different employee populations, would achieve different results. Second, we did not include employee attributes (eg, hospital role, education level, primary language, demographics) in our analysis. Such information, which we leave to a future study, would be helpful in understanding populations at higher or lower risk. Third, because of the nature of our data—retrospective, and gathered as part of an information security program (and not a research study)—our statistical options to evaluate the training program were limited. We considered quasi-experimental models (eg, difference in differences) and mixed-effect logistic regression models, but none could overcome the bias of our study group (the offenders) having been preselected by the outcome of interest (clicking on an email). Although not a randomized trial, we do feel that our results—that the training did not have a substantial impact—are valid, and would hold under an improved, prospective study design.

CONCLUSION

Despite significant international attention to cybersecurity in the past several years, we show that in a simulation environment, hospital employees remain remarkably susceptible to phishing emails. Though click rates decreased over time, a mandatory training program for the highest-risk employees in the context of a mature phishing simulation program did not substantially reduce click rates, and this population remained particularly susceptible to phishing compared with the general population. These click rates suggest that the US healthcare system is quite vulnerable to cybersecurity attacks, and significant work is needed to improve this vulnerability.

FUNDING

This work was supported by Harvard Catalyst (Harvard Clinical and Translational Science Center) (National Center for Advancing Translational Sciences, National Institutes of Health Award UL 1TR002541) and financial contributions from Harvard University and its affiliated academic healthcare centers. The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University, and its affiliated academic healthcare centers, or the National Institutes of Health.

AUTHOR CONTRIBUTORS

WJG, AW, and AL conceptualized the project. WJG conducted the analyses and wrote the manuscript. RG, CM, EL, and JK contributed to the analysis and writing of the manuscript.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

Conflict of interest statement. JK is on the Board of Advisors of Censinet and the Board of Directors of the Health Information Sharing and Analysis Center.

REFERENCES

1. Clarke R, Youngstein T. Cyberattack on Britain's National Health Service—a wake-up call for modern medicine. *N Engl J Med* 2017; 377 (5): 409–11.
2. Nigrin DJ. When "hacktivists" target your hospital. *N Engl J Med* 2014; 371 (5): 393–5.
3. Gordon WJ, Fairhall A, Landman A. Threats to information security—public health implications. *N Engl J Med* 2017; 377 (8): 707–9. doi: 10.1056/NEJMp1707212.
4. Bai G, Jiang JX, Flasher R. Hospital risk of data breaches. *JAMA Intern Med* 2017; 177 (6): 878–80.
5. Ponemon Institute. Sixth Annual Benchmark Study on Privacy & Security of Healthcare Data. 2016. http://www.ponemon.org/blog/sixth-annual-benchmark-study-on-privacy-security-of-healthcare-data-1 Accessed October 26, 2018.
6. Ponemon Institute. 2016 Cost of Data Breach Study: Global Analysis. 2016. https://resources.idgenterprise.com/original/AST-0185855_SEL03094USEN.PDF Accessed February 25, 2019.
7. Spitzer J. Vanderbilt Warns Hospital Staff about Recent Phishing Attempts. *Beckers Hospital Review*. https://www.beckershospitalreview.com/cybersecurity/vanderbilt-warns-hospital-staff-about-recent-phishing-attempts.html Accessed October 28, 2018.
8. Davis J. 1.4 Million Patient Records Breached in UnityPoint Health Phishing Attack. *Healthcare IT News*. https://www.healthcareitnews.com/news/14-million-patient-records-breached-unitypoint-health-phishing-attack Accessed August 8, 2018.
9. National Institute of Standards and Technology. NISTIR 7298: Glossary of Key Information Security Terms. http://nvlpubs.nist.gov/nistpubs/ir/2013/NIST.IR.7298r2.pdf Accessed October 26, 2018.
10. Wright A, Aaron S, Bates DW. The big phish: cyberattacks against U.S. healthcare systems. *J Gen Intern Med* 2016; 31 (10): 1115–8.
11. Jansson K, von Solms R. Phishing for phishing awareness. *Behav Inf Technol* 2013; 32 (6): 584–93.
12. Dhamija R, Tygar JD, Hearst M. Why phishing works. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM; 2006: 581–90.
13. Dhamija R, Tygar JD. The battle against phishing: dynamic security skins. In: *Proceedings of the 2005 Symposium on Usable Privacy and Security*. New York: ACM; 2005: 77–88.
14. Parno B, Kuo C, Perrig A. Phoolproof phishing prevention. InInternational Conference on Financial Cryptography and Data Security 2006 Feb 27 (pp. 1–19). Springer, Berlin, Heidelberg.
15. Khan AA. Preventing phishing attacks using one time password and user machine identification. *Int J Comput Appl* 2013; 68: 7–11.
16. Miyamoto D, Hazeyama H, Kadobayashi Y. An evaluation of machine learning-based methods for detection of phishing sites. InInternational

Conference on Neural Information Processing 2008 Nov 25 (pp. 539–546). Springer, Berlin, Heidelberg.

17. Miyamoto D, Hazeyama H, Kadobayashi Y. SPS: a simple filtering algorithm to thwart phishing attacks. InAsian Internet Engineering Conference 2005 Dec 13 (pp. 195–209). Springer, Berlin, Heidelberg.
18. PhishMe. Enterprise Phishing Susceptibility Report. 2015. https://phishme.com/wp-content/uploads/2017/10/PhishMe_EnterprisePhishing-SusceptibilityReport_2015_Final.pdf Accessed October 26, 2018.
19. Ponemon Institute. The Cost of Phishing & Value of Employee Training. 2015. https://info.wombatsecurity.com/hubfs/Ponemon_Institute_Cost_of_Phishing.pdf Accessed October 26, 2018.
20. Cofense, Inc. https://cofense.com/. Accessed February 25, 2019.
21. HealthStream. https://www.healthstream.com/ Accessed October 26, 2018.
22. Chandramouli R, Garfinkel S, Nightingale S, *et al*. Trustworthy Email - NIST Special Publication 800-177. 2016. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-177.pdf. Accessed February 25, 2019.
23. Duke E. Binding Operational Directive BOD-18-01. 2017. https://cyber.dhs.gov/assets/report/bod-18-01.pdf Accessed October 28, 2018.
24. Kumaraguru P, Sheng S, Acquisti A, Cranor LF, Hong J. Lessons from a real world evaluation of anti-phishing training. In2008 eCrime Researchers Summit 2008 Oct 15 (pp. 1–12). IEEE.