<<LOGO>>

CompanyN

Prepared for:
President
IT Manager
Information Security Analyst

Prepared by:
Andrew Bates, UIS graduate student
Thomas Caldera, UIS graduate student
Kari Grafton, UIS graduate student
Shanita Hunt, UIS graduate student
Pedro Valencia, UIS graduate student

Faculty Research Supervisor:
Sahar Farshadkhah, Assistant Professor

College of Business and Management
Management Information Systems Department
Capstone course project, Spring 2021

<<This is an anonymized copy of the follow-up report provided to the client after the completion of the phishing test.>>

PHISHING TEST FOLLOW-UP REPORT

TEST DATE: APRIL 2021

Background

Phishing attacks are on the rise, and we should remember that the attackers are often several steps ahead of the technologies and personnel intended to defend against them. Even the best anti-phishing tools are merely reactive and can only prevent against known attacks. To help guard against new attacks we need to be proactive; users need to be trained to be the front line against these attackers. A survey of 524 companies who experienced a data breach between August 2019 and April 2020 found that 19% of the breaches were a result of compromised credentials. The average cost of the studied data breaches was $3.86 million, and it took an average of 280 days to contain the breach.[1]

CompanyN identified a need to assess the employees' security awareness and readiness of responding to phishing attacks. Prior tests have been performed and education has been provided to the employees on how to guard themselves from falling for phishing e-mails. The goal of this test was to gauge the training's effectiveness and provide ongoing phishing training for the employees.

Test Design

While previous phishing tests have attempted to acquire user credentials, this test uses just an e-mail link for three reasons. First, research has shown that users who log in before seeing the education page are confused as to why they are seeing information about not clicking on links, likely because the login created a gap between clicking the link and seeing the education.[2] Second, the same research also suggests that the majority of users who will click on a phishing site will also provide their credentials, so it seemed unnecessary. And finally, because collecting credentials introduces extra security complications that would be preferable to avoid in this particular case.

Care was taken in designing an e-mail with the appropriate level of persuasiveness. Phishing scams as common and obvious as "Nigerian prince" scams should be well-known enough at this point that it would limit the usefulness of any training provided. The most complex phishing e-mails tend to include elements that are beyond the average computer user's understanding, and this level of attention to detail is less commonly seen in real-world phishing attacks. While these are important to train, without enough details of previous tests to create a baseline to compare against, the decision was made to design a more average-difficulty test to both give the most accurate overall metrics and to create a baseline that a training program of repeated phishing tests can be built off of.

Figure 1 shows a screenshot of the e-mail for this phishing test from within CompanyN's environment. The primary strategy chosen for this phish was to acquire a domain one letter off from the company's real domain in the hope that the users would miss the subtle difference. The

[1] IBM. (2020). *Cost of a Data Breach Report 2020.* Armonk, NY: IBM Corporation.

[2] Kumaraguru, P., Rhee, Y., Acquisti, A., Cranor, L. F., Hong, J., & Nunge, E. (2007). Protecting People from Phishing: The Design and Evaluation of an Embedded Training Email System. *CHI 2007*, (pp. 905-914). San Jose, CA.

domain companym.com was available and was acquired for the test. There are two places that the user could have noticed this discrepancy: in the from address or by hovering over the "sign in" link to view the actual URL. The e-mail system was set up to flag external senders as shown in the screenshot of the e-mail, but experience says that many users ignore this warning. The secondary strategy chosen for this phish was to make the e-mail look like something that could legitimately come up in the course of the user's work. It is not unreasonable to think a user would receive a message from a coworker or client that required extra security, and this e-mail was designed based on some simple examples of legitimate e-mails of this sort.

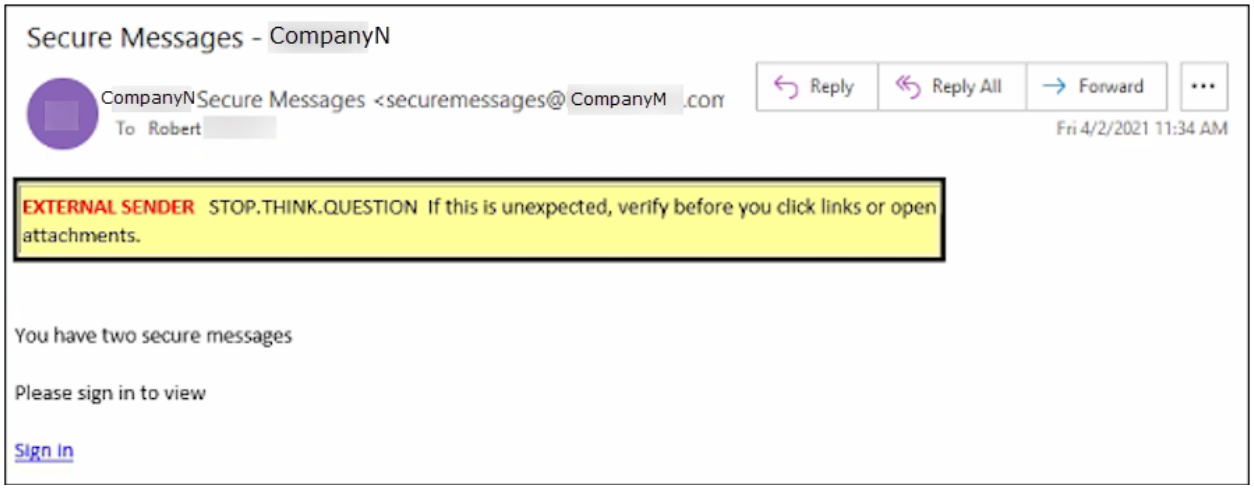

Figure 1 - Phishing e-mail for this test, shown inside the client's environment

The day of the week and time of day selected to send the e-mail was strategic as well. One of the strategies of social engineering is to get users to act without thinking, so a time was selected when users were likely to be busy and/or preoccupied. Mondays are a good day, because people are trying to get caught up from the weekend and may not be giving things the attention they should be. The two times of day considered were shortly before lunchtime and shortly before the end of the day. Shortly before lunchtime was chosen because sending at the end of the day leaves it more likely that the user will ignore it until the next day, whereas if the user leaves the e-mail until after lunch, a much smaller time gap is created.

The client provided a list of 521 e-mail addresses for the test, which were broken up into smaller groups, so the e-mail server did not see such a large number of identical e-mails coming in at the same time and block it as suspicious. The list was divided up by time zones so all e-mails would be sent around the recommended time frame for the time zone, a random number was assigned to each e-mail address, and these random numbers were used to parse the list into groups of 25-26 addresses, with one group of 6 addresses for the Pacific time zone. Groups were scheduled to be sent with gaps of 4 minutes in between.

Before the actual phishing test, a trial run of the test was done to make sure the e-mail would successfully make it to a user's inbox. At this point, the server logs were also checked to see if the 1-pixel white image planted in the body of the e-mail would show as accessed, which would

indicate that the user opened the e-mail. This trial run did show the image accessed, but the final test results show some users who clicked the link without showing as having accessed the e-mail, so this suggests that only some users' e-mail clients are loading the images without first asking for an ok from the user. For this reason, estimated active users will be calculated and used in the metrics.

Sending of the e-mails ran smoothly and the schedule was adhered to ± 2 minutes. A few e-mails quickly bounced back as undeliverable, and an attempt was made to correct the e-mail address and add them back into a later list. The test was considered "open" until Friday of the same week, at 5 p.m. at the office in the latest time zone. At this time, the test was considered "closed" and the server logs for the duration of the test were pulled to be used to calculate the metrics.

Clicking the link took the user to an education page designed specifically for this test, and structured into 5 parts (Figure 2, Figure 3):

1. A brief, light-hearted attention-getter including an image that represents phishing in a slightly amusing way, and large text telling the user they have clicked on a simulated phishing test authorized by CompanyN.
2. A very brief description of what phishing is.
3. An interactive piece telling the user that two things in this particular e-mail could have indicated to them that it was a potential phishing e-mail and asking the user to see if they could identify them. Clicking the image revealed the indicators to the user with callouts giving more information. This section uses rule-based training to teach users specific things to watch out for.
4. Three ways to think critically about an e-mail before acting. This encourages the user to be mindful in what they are doing, in order to counteract the attempt of social engineering methods to get the user to act without thinking.
5. The steps the user should take if they suspect an e-mail may be a phish. These are kept as succinct as possible, so the user does not have too much to remember.

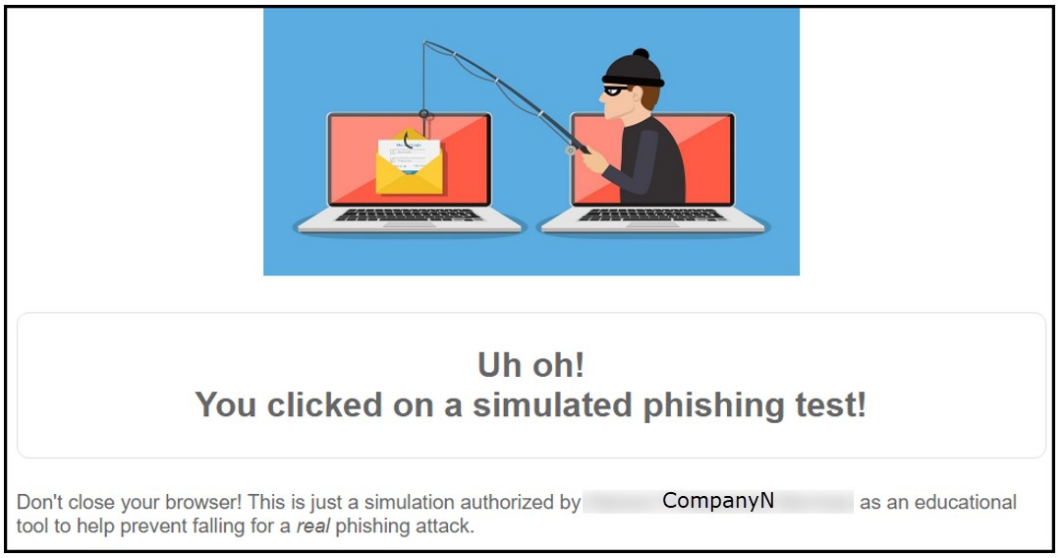

Figure 2 - Phishing education page, part 1

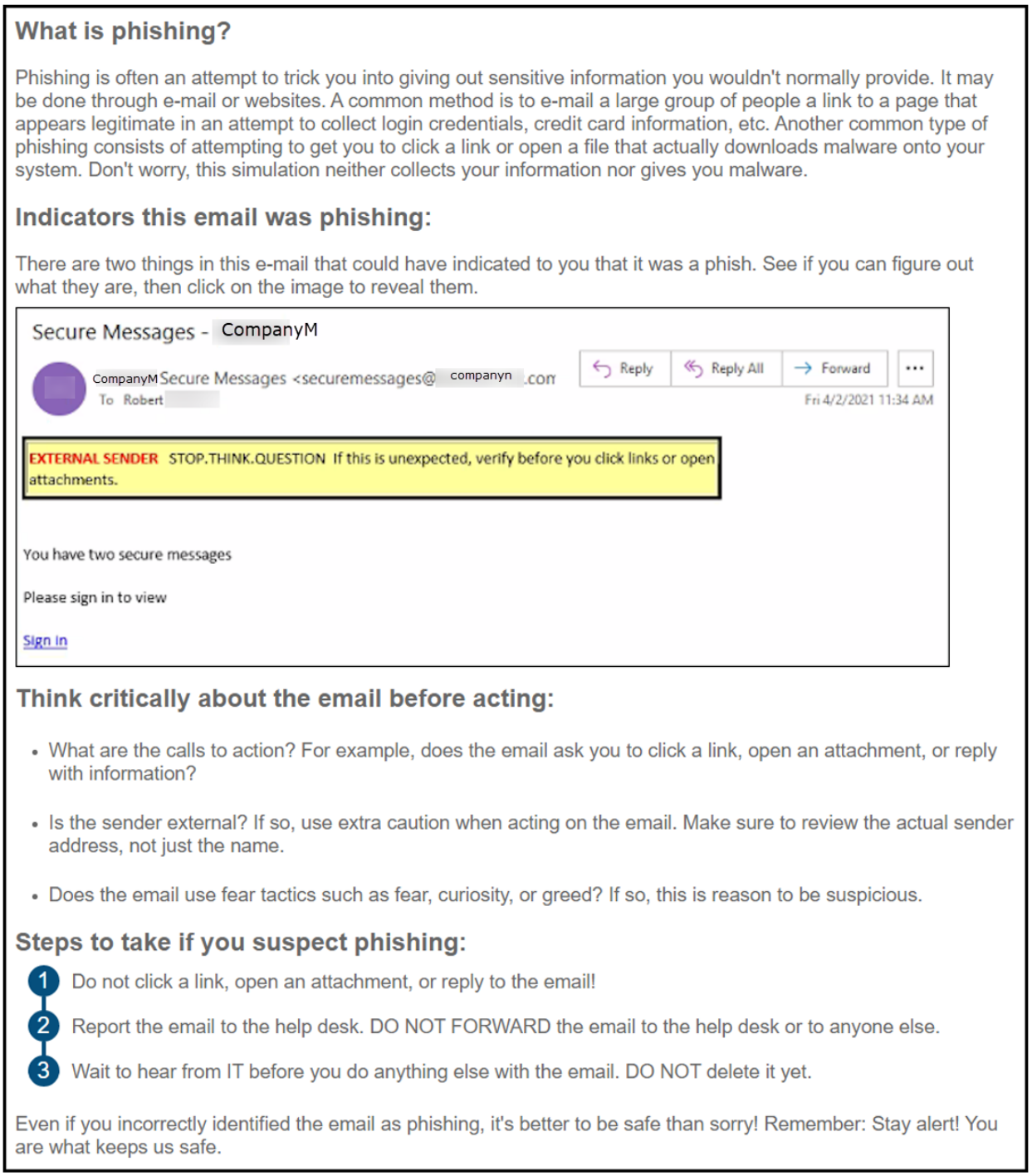

Figure 3 - Phishing education page, part 2

Results

The majority of the data needed for the metrics were gathered from the server logs, with the remaining data received from the client and the test e-mail inbox. Table 1 lists each value used in the metrics, how the value was obtained, and the value resulting from this test. All values and metrics exclude the people at the company who knew the test was taking place.

It would have been possible to include demographic attributes of the users in the metrics to see if and how this affected things, but the literature has very mixed opinions on whether demographic attributes actually contribute. Even if there was any significant contribution found, this information only becomes actionable if it can be used to tailor the training in some way. Given the literature's disagreement on this matter and the extra amount of work that would go into trying to tailor the training to different groups of demographics with potentially very little benefit, demographic attributes have not been included in these metrics.

Abbr	Name	Metric	How Gathered	value
P	population	users the email was successfully sent to	received list from the client	513
C	clicked	users who clicked the link in the email	count of IDs in the server logs who have a GET request for index.php	66
T	trained	users who clicked the image to see the indicators of potential phishing	count of IDs in the server logs who have a GET request for img2.png	4
R	reported	users who reported the email as suspicious	received list from the client	95
A	active participants	users who opened the email	count of IDs in the server log who have a GET request for the hidden image in the email	counted: 212 estimated: 288
N	non-active participants	users who did not open the email	population minus active participants (P-A)	counted:301 estimated: 225
LI	location: in office	users who clicked the link from one of the client's office locations	count of IDs whose GET request for index.php shows as from an IP in the list of IPs the client provides as their NAT addresses	22
LR	location: remote	users who clicked the link from a remote location	count of IDs whose GET request for index.php shows as from an IP not in the list of IPs the client provides as their NAT addresses	44

Table 1 – Values for the phishing test metrics

The most obvious metric to report is click rate, the number of users who clicked the link (C) divided by the number of people in the test population (P). Multiplying this by 100 gives us a percentage:

$$click\ rate = C/P * 100 = 66/513 * 100 = 12.87\%$$

One interesting extra piece of information is that the first user to click the phishing link did so 4 seconds after the e-mail was sent. This expresses just how important it is for the help desk team to be made aware of any real phishing attacks as quickly as possible.

Also relevant is how many of the people who went to the phishing page clicked on the image to reveal what the phishing indicators were. This divides the number of people who clicked over to the second image (T) by the same C used above to give a rate of phished users who were also trained by reading at least a portion of the education page:

$$trained\ rate = T/C * 100 = 4/66 * 100 = 6.06\%$$

Figure 4 shows the above two metrics graphically. The whole bar represents the population the test was sent to, with grey being people who did not click the link and the blues being the people who did click. The small sliver of dark blue on the right is the people who we considered trained.

Figure 4 - Users who did and did not click the phishing link

But just looking at the click rate and trained rate fails to show the whole picture. An essential piece is to know how many people reported the e-mail as suspicious. Matching up whether each person clicked the link or not and whether each person reported the e-mail as suspicious or not gives the following matrix:

	Reported	*Did not report*
Clicked link	*Q1*	*Q2*
Did not click link	*Q3*	*Q4*

Dividing each quadrant by P and multiplying by 100 gives a percentage for each value. This gives a better breakdown by quadrants, as follows:

Quadrant 1 – These users are halfway there. Despite clicking on the phishing link, they then followed up by reporting it to the help desk.

Quadrant 2 – These users are the most dangerous. Not only did they click the link, they did not report it to the help desk so that mitigating action could be taken as soon as possible.

Quadrant 3 – This is ideally where you want your users, not clicking the link, but still reporting the e-mail as suspicious.

Quadrant 4 – This is a mixed bag. It contains some people who are halfway there, in that they did not click the link, but they also did not report the e-mail to the help desk as suspicious (a second best to being in quadrant 3). But it also contains the people who never even opened the e-mail and so never would have had the chance to click the link.

This test grouped users in the following way:

	Reported	*Did not report*
Clicked link	*0.39%*	*12.48%*
Did not click link	*18.13%*	*69.01%*

Users have been specifically told not to forward a phishing e-mail to the help desk to report it, so if only the users who reported it to the help desk properly are counted, the numbers are more concerning:

	Reported properly	*Did not report properly*
Clicked link	*0.19%*	*12.67%*
Did not click link	*7.60%*	*79.53%*

Representing these two matrices visually more easily gives us another insight. (see Figure 5) While the percentage of users in the ideal group (Q3) is higher than the users in the most dangerous group (Q2), when accounting only for users who reported through the proper channels, this reverses.

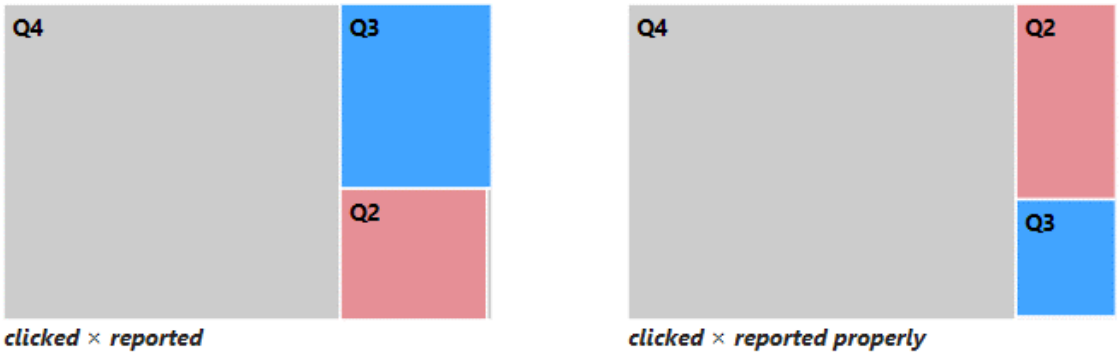

Figure 5 - Clicked × reported matrixes

34 users also directly contacted one of the two IT employees involved in running the test either by phone or through Microsoft Teams, but these reports are not counted in the metrics as they did not go through the proper channels.

The trial run with the phishing test e-mail showed that the e-mail client is configured to display remote images for at least some users, so a social engineering resistance (SER) metric can be calculated. SER corresponds to the probability that an adversary, selecting one user at random, will obtain secret information.[3] It is important to note that since the metrics measure how likely

[3] Hasle, H., Kristiansen, Y., Kintel, K., & Snekkenes, E. (2005). Measuring Resistance to Social Engineering. *International Conference on Information Security Practice and Experience* (pp. 132-143). Springer-Verlag Berlin Heidelberg.

the hackers are to *succeed*, lower numbers are better.[4] This set of metrics uses the same P and C values from previous metrics, but also takes into account the users who did and did not open the e-mail.

This is where the estimated numbers of active users come into play. It is unknown how many of the users who did not click the link opened the e-mail, but the number of users who showed as active (opened the e-mail) and non-active (did not open the e-mail) within the set of users who clicked the link is known. Using the assumption that users who did not click the link accessed the e-mail in the same ratio as users who did click the link, estimated numbers of active and non-active users for the entire population can be calculated. Those estimated numbers will be used for the social engineering resistance metric as they are likely closer to the actual value.

Three metrics are computed here, each based on a different assumption of how non-active participants would have acted if they had opened the e-mail. The three parts of the social engineering resistance metric are:

Optimistic assumption – This value assumes that none of the non-active participants would have clicked the link, so divides P by C. (Note that this calculation is the same as the click rate above.) This gives the low bound for the range.

$$SER_{LOW} = (C/P) * 100 = (66/513) * 100 = 12.87\%$$

Average assumption[5] – This value assumes that the non-active participants would have clicked the link at the same rate as the active participants, so C is divided by A to find what percentage of the active participants clicked the link.

$$SER_{AVG} = (C/A) * 100 = (66/288) * 100 = 22.92\%$$

Pessimistic assumption – This value assumes that all of the non-active participants would have clicked on the link, so adds together the number of clicks and the number of non-active participants, then divides by P. This gives the high bound.

$$SER_{HIGH} = (C+N)/P * 100 = (66 + 225)/513 * 100 = 56.73\%$$

Together these three numbers give a social engineering resistance of:

$$SER = <SER_{LOW}, SER_{AVG}, SER_{HIGH}> = < 12.87\%, 22.92\%, 56.73\%>$$

[4] Note that the original article inverted these numbers to show a preference for higher values, so the numbers look better, but skipping the step makes it slightly easier to understand.

[5] The average assumption was not included in the original paper that designed the SER, but it is a natural extension of the method.

which gives a fuller picture of the state of things. This suggests that an attacker targeting any one random employee has somewhere between a 13-57% chance of success.

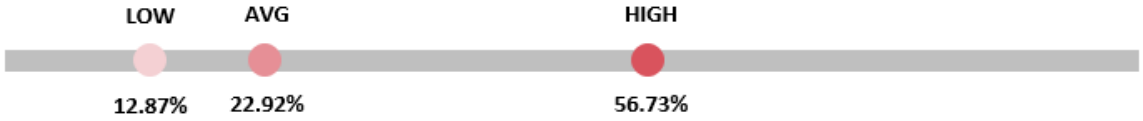

Figure 6 - Social engineering resistance (SER) metric

Because of the current pandemic and drastic increase in the number of employees working from home, it is worth looking into if there are behavioral differences between users working in-office (LI) and remote (LR). Each user's location was determined by comparing the IP address in the server logs with the list of internal NAT addresses provided by the client. Each set of metrics could be split by in-office and remote, but the most useful information comes from a straight comparison of the ratio of clicks from in-office users to remote users:

Ratio of in-office : remote clicks = LI:LR = 22:44 = 1:2

This does not necessarily mean that a user working remotely is twice as likely to click the phishing link as a user who is working in the office. Since this is only one data point, it is not enough to make that much of a prediction. It also does not take into account the ratio of employees working in-office to working remotely. What it does mean is that *for this particular test*, remote users clicked the phishing link at twice the rate of in-office users, so it is a metric worth keeping an eye on and possibly digging further into.

Performing WhoIs lookups of the users' IP addresses could also provide a potential hacker a plethora of information, including approximate user locations, information about the network, and so on. Most of the internal NAT addresses on the list provided by the client were also identified as belonging to the company this way. A determined hacker could also go after very specific information found this way. The WhoIs information on one user's IP address placed them at a specialized medical facility. If the assumption is made that the user is there as a patient (as opposed to doing work for the medical facility), this could give a hacker a very emotionally-charged approach to spear-phish the specific employee.

Recommendations

This test only included one phishing e-mail, but periodic repetition of a phishing test combined with the embedded training will keep the information fresh in the minds of long-time users and help train up new users. A review of the literature recommends phishing e-mails for training purposes should be done on a monthly basis.[6] Each iteration of the test serves both as an opportunity to teach and an opportunity to test the education that has been given in the past. These future tests are the place to introduce the more complex elements that show up in the more carefully-designed phishing attacks, to further the users' knowledge. This test can serve as both a design model and a baseline for future testing.

[6] Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. F., & Hong, J. (2008). Lessons from a Real World Evaluation of Anti-Phishing Training. *2008 eCrime Researchers Summit* (pp. 1-12). IEEE.

The major weakness of this test is the same issue that plagues so much of computer security: how do you get the users to care about the education? While this is essential to examine and would be expected of a company contracted to do regular testing, that is outside the scope of this particular project for this particular group of graduate students.

We recommend selecting one provider to do the testing, whether it is done in-house or by a penetration testing company, and gathering the same metrics from each test so they can be compared across time. Using the same provider across tests would also let the provider track individuals across tests, at a level of granularity which would not generally be provided to the client, to identify users who are a particular danger to the company and should be provided additional training. A dashboard giving a visual representation of the aggregate metrics from the series of tests would also be beneficial.

If monthly testing is found to be too much and the company is seeing consistently positive results, an alternate suggestion is to pick an acceptable threshold for the metrics (e.g., the SER is below a certain percentage and at least a certain percentage of users are reporting the phish), and once the numbers have met this threshold for several tests in a row, decrease to quarterly testing so long as the metrics stay within the threshold. But the rate of employee turnover should also be taken into account – remember that you are not just continuing training for existing users, you are training your new users as well.

Phishing attacks are unlikely to ever go away, given how successful they can be and often are, and from very little work on the attacker's end. But that does not mean that defending against them is a fruitless task. Users should be seen not so much as the "weakest link"; a shift needs be made towards seeing them as an integral brick in the defense structure against phishing attacks. Periodic, concise training that is tailored to your company's and users' needs is how you will reinforce this brick and strengthen the mortar between it and the other bricks in your defense structure.