Measuring the Human Factor of Cyber Security

Brian M. Bowen, Ramaswamy Devarajan, Salvatore Stolfo
Department of Computer Science
Columbia University
{bb2281, rd2446, sjs11}cs.columbia.edu

Abstract—**This paper investigates new methods to measure, quantify and evaluate the security posture of human organizations especially within large corporations and government agencies. Computer security is not just about technology and systems. It is also about the people that use those systems and how their vulnerable behaviors can lead to exploitation. We focus on measuring enterprise-level susceptibility to phishing attacks. Results of experiments conducted at Columbia University and the system used to conduct the experiments are presented that show how the system can also be effective for training users. We include a description of follow-on work that has been proposed to DHS that aims to measure and improve the security posture of government departments and agencies, as well as for comparing security postures of individual agencies against one another.**

I. Introduction

Lord Kelvin taught us that without numbers there is no science. Indeed, quantification lies at the very heart of scientific progress. Without measurement, one cannot know what has been learned or achieved and whether our knowledge has been advanced or progress has been made. The field of computer and information security requires the foundational science that provides the means for assessing the strength of organizational security postures. For the Department of Homeland Security needs, solid metrics may be applied as a means of assessing the strength of one organization relative to others, and to help identify vulnerabilities.

There are a few notable successes in the field of computer security where metrics have been well established and profitably applied to practical technologies with mathematically provable security properties. Cryptography has provided tools for researchers and developers to devise many practical and widely deployed technologies that provide for confidentiality. The formal analysis of the hardness of certain computational problems establishes a metric to judge the strength of an encryption scheme based upon key length, for example. However, cryptography alone does not provide all of the security guarantees we may want. Other areas of computer security have provided metrics to evaluate the relative merits of specific technologies, such as detection accuracy rates of competing intrusion detection systems. Computer security is not just about technology and systems, but must also take into consideration the people and processes that rely on the systems. In this work, we propose an approach to measuring organizational security and educating users that relies on mimicking attackers' actions in social attacks.

Social attacks include those that occur when an attacker uses any of a variety social attack vectors that may range from email and telephone to in-person encounters. According to the 2010 Verizon Data Breach Investigations Report [2], social attacks were used in 28% of the breaches for 2009 and nearly a quarter of these attacks occurred due to phishing. In these types of attacks, victims are sent spoofed emails that appear to be benign notifications from a bank, a social networking site, or a software upgrade. When victims take the bait, they are often greeted with some form of malicious software that attempts to install itself on victim's machine. Although there have been many technological advances that seem to hold promise in stopping these attacks, so far, none of them have proven 100% effective allowing the problem to continue. In fact, the vulnerability posed by phishing is often used effectively in the largest, most costly attacks happening today, like the recent attack on RSA [3].

Although the primary focus of the work is on measuring organizational security, the results suggest that the system also provides utility in training users. The defense approach we are advocating in this paper involves better educating users to be cautious of suspicious emails. Traditional training techniques can be beneficial, they are often not enough. Our technique involves testing users' vulnerability using a variety of decoy emails; those that fall victim to our phony phishing attacks are informed so that they may learn and change their behavior later. Subsequent tests of the same users show that this method works, although sometimes it takes several iterations of testing and teaching.

How a user responds to security significant events is also an important consideration for organizational security. For example, if someone receives a spear phishing email, how do they handle it? They could simply delete it, but optimally, they may report it so that appropriate actions are put in place to protect other users. The processes that an organization has in place and how they are socialized are essential for strong defense. The proposed system has been designed to automatically monitor emails sent to determine when a user falls for them, but it can also be used to test user responses to security events, an area that often gets little attention. These metrics may be used to evaluate and quantitatively determine the effectiveness of organizational policies and reporting processes.

A. Summary of Results

The preliminary results obtained from experiments conducted at Columbia University were conducted over the course of a year with a randomly selected group of 4000 students, staff, and faculty. Each of these participants was sent one of four types of phony phishing emails manually modeled after real phishing emails. The experiments began with 500 emails being sent for each of the four types. Users that fell for the fake phishing emails were presented with messages indicating so. Only users that fell for the bogus emails were selected for the next round. In summary, it took a total of four rounds before all users were able to identify the emails as being bogus. The results suggest that users can be trained using decoy technology to be cognizant of potential threats and provide a useful metric for assessing organizational security. Applying the same set measurements laterally across multiple organizations can be useful in measuring one organization's security posture relative to another's.

II. Related Work

The computer security field demand for techniques to evaluate and compare security designs and organizations. Many techniques have been proposed and explored [1], but these typically focus on systems and technologies rather than people. Our work aims to demonstrate techniques aimed at measuring organizational security through its people rather than just with the technology on which they rely.

The proposed system is designed for educating users and measuring organizational security using decoy emails. Traditional security training classes can be beneficial for organizations, but they are not enough and there are more effective methods [4]. Our technique involves testing users' vulnerability using a variety of decoy emails; those that fall victim to our phony phishing attacks are informed so that they may learn and change their behavior. Traditional approaches for training users about the threat posed by phishing rely on classes and informational warnings. Efforts to raise user awareness have focused on testing users to demonstrate their vulnerability [5]. Some tools have been created to support the sending of fake phishing emails for purposes of pen testing and training [6], [7], but these rely on an administrator to manually construct and send the emails to targeted individuals. None of these tools focus on the development of formal metrics for measuring organizational security such that they can be used for relative comparisons for comparing one organization against another.

A similar study was conducted at Indiana University, which involved social phishing and spoofing [8]. As part of the study, the researchers launched harmless phishing attacks on the students, specifically targeting students aged 18-24. The experiment was performed with intent to show that social context can be used in effective phishing. Unlike our efforts, they did not focus on how useful metrics could be obtained from the experiments or on techniques for improving an organizational security posture.

The Honeynet project titled Know your enemy provides practical information on the practice of phishing and draws on data collected by the German Honeynet Project and UK Honeynet Project [9]. This paper discusses on the various techniques and tools used by the phishers, providing three examples of empirical research where real-world phishing attacks were captured using Honeynet. It also specifies the variety of malware that are used by spammers in automating the email address for generating genuine looking emails in tricking the users.

III. Phony Phish System

The goal of the Phony Phish System is to provide an automatic means to generate and send benign phishing emails that can be used to measure an organization's security and educate users. The system consists of several components as shown in Figure 1.

a) Crawler Module:: This component was designed to crawl a directory and obtain a list of target identities to perform the experiments on. For the experiments described in IV, this module was used to search the Columbia University directory, select users, their role within the university, and which department they belong to.

b) Email Generator:: The email generator integrated all of the components and was used to deliver emails for the experiments. This component takes real email as input and performs processing on them to change names and using the Stanford Named Entity Recognition [1] engine. It also functions to anonymize user identify information through the use of unique hashes. For the generation of beacon'ed documents, the email generator relies on the Decoy Document Distributor introduced in [10].

c) Web Application:: A web application is used to collect user responses when they click on links and submit forms containing credentials. It tracks responses using a base 64 encoded query string that is attached to user requests. The system does not the store the identify of users, only the time at which the link was clicked, the department that the user belongs to, and the role of a user within the organization.

IV. Experimental Analysis and Results

Experiments were conducted by sending 500 emails for each of four different types of decoy emails. Using standard statistical techniques [11], this sample size was determined to be significant for measuring a single population parameter (*i.e.,* will a user open an email) with a 5% margin of error and 95% confidence for the approximately 70,000 IDs in the Columbia University directory. A second consideration for the choice of using 500 was for practical reasons. Our intent was to have a sample size large enough to draw scientifically significant conclusions without burdening an unnecessary number of subjects. Although we had permission from the university, and the approval of the Columbia University Institutional Review Board, the subjects were unwitting participants. The nature of this kind of experiment has the potential to cost users in both time and aggravation. Given that this was our first attempt at such an experiment, we decided we would start with 500 emails for each of the four types and adjust as necessary.

The decoy emails were modeled after various types of phishing attacks that occur in the wild. All of the emails were

[1] http://nlp.stanford.edu/ner/index.shtml

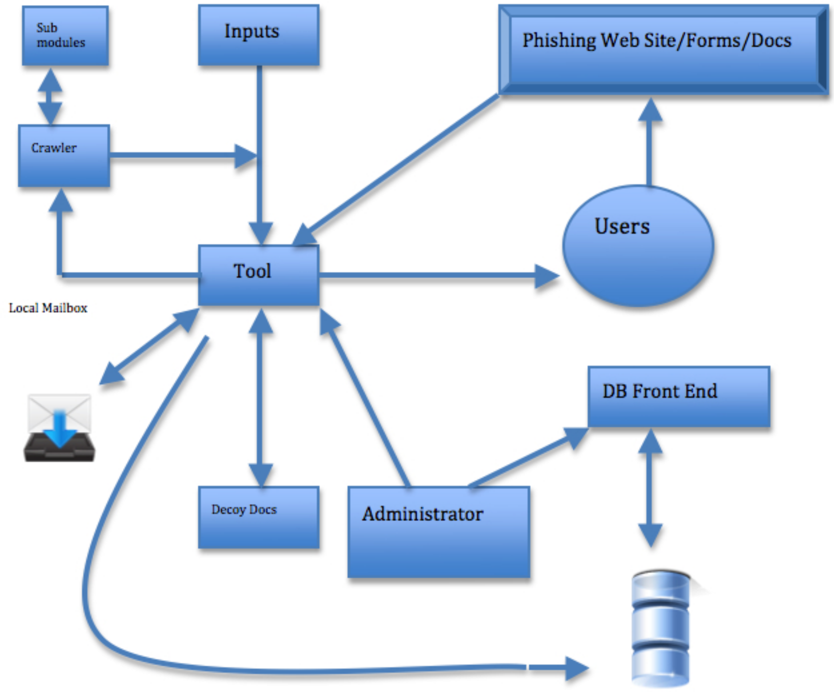

Fig. 1. Components of the Phony Phish System.

sent using an external email account from a popular webmail provider. Users that fell victim to the phony phishing emails were presented with the following message:

The Columbia University IDS Lab is conducting experiments designed to measure the security posture of large organizations and to educate users about safe practices so that they avoid falling prey to malicious emails. The emails automatically generated and sent to users of Columbias network and email system are designed to test whether users violate basic security policies. Although our emails are completely benign, please be aware that many emails are sent that are designed to trick unsuspecting users into giving up identity information.

The four different types of emails and their results are summarized below:

- **Email with internal URLs:** The content of these emails were from email received with an external source, but the URLs were changed to point to our IDS severs. The goal of these emails were to see how many users bothered to look at the address of the recipient before opening the email.
- **Email with external URLs:** The content of these emails was modified from emails received with an external source. The emails were designed to lure those interested in obtaining the Apple iPad. The URLs were changed to point to our external servers in the .info domain.
- **Forms to obtain credentials:** The content of these emails contained links to forms asking users for credentials to see how many users were willing to expose their credentials. Credentials were not stored.
- **Beacon Documents:** These emails contained PDF attachments that emitted a beacon to our servers when opened. The beacons were designed so that every user emitted a unique response enabling us to track them. An evaluation of the beacons is provided in [10].

A. Result Summary

Table I and Table II provide an overview of all of the results obtained from two rounds of experiments. Over the course of several weeks, offenders were repeatedly targeted until they stopped falling victim. The results between the two rounds of experiments were fairly consistent. The most important point that can be gleaned from the data:

In all cases, users can be trained to be cautious of suspicious looking emails, but sometimes it takes several iterations of testing. In our experiments, the slowest learners took at most four iterations as shown in II.

Table III presents a role-wise breakdown of the results. It can be seen that students are more susceptible to the phishing emails than university staff. This may be attributed to the fact that student population flux is high and the tendency to adapt to Columbia policies may not be uniform. Staff awareness of, and adherence to security policies is greater than that of the students, who may simply have no regard for them. A similar observation can be made in case of decoy

TABLE I
THE NUMBER OF RESPONSES FOR EACH ROUND FOR THE FIRST EXPERIMENT TO MEASURE THE USER RESPONSE TO PHONY PHISH.

Decoy Type	1^{st} **Round**	2^{nd} **Round**	3^{rd} **Round**	4^{th} **Round**
Email with internal URLs	52	2	0	NA
Email with external URLs	177	15	1	0
Forms to obtain credentials[2]	39/20	4/1	0	NA
Beacon Documents	45	0	NA	NA

TABLE II
THE NUMBER OF RESPONSES FOR EACH ROUND FOR THE SECOND EXPERIMENT TO MEASURE THE USER RESPONSE TO PHONY PHISH.

Decoy Type	1^{st} **Round**	2^{nd} **Round**	3^{rd} **Round**	4^{th} **Round**
Email with internal URLs	69	7	1	0
Email with external URLs	176	10	3	0
Forms to obtain credentials	69/50	10/9	0	NA
Beacon Documents	71	2	0	NA

documents. There exists a Columbia University policy that prohibits columbia from downloading documents originating from unknown email addresses. The results suggest that the policy may be regarded more highly by the staff population than the student population. Table IV provides the school-wise breakdown of the results. It is evident from the table that there is almost uniform distribution of vulnerable users across the schools affiliated with the university.

Table V provides staff-wise breakdown of the results obtained. One observation that can be made from these results are that non-academic staff are more vulnerable to phishing attacks than the academic staff.

Another observation was that it appeared users are less likely to respond to emails that appear to be from internal sources, but have an external sender address. These emails were indeed suspicious, but we do not have a good way to account for the differences in the content. For example, the external emails (row 2) appear pertain to the Apple iPad. At the time the emails were sent out, the emails would have been appealing to the masses. On the other hand, the internal emails (row 1) resembled those distributed by the university and are likely less appealing to the masses. Hence, there is insufficient data to make any conclusion concerning these differences.

The number of users that actually entered their credentials to the bogus forms seemed alarmingly high. We did not record the credentials and we did not validate them to ensure they were valid. However, we believe it is likely that at least some of the users entered valid credentials.

B. Challenges

One of the challenges in conducting these studies lied in managing the user discontent that was generated as a result of being an unwitting participant in the studies. Despite the messages that were generated to describe the nature of the experiments and the benefit they may bring, some users still brought issue. The largest challenge for us was in accommodating the users that were not so easily fooled by the fake phishing emails, or who correctly identified them as being suspicious. Unlike the users that did fall for the phony phishing emails and who were presented with messages describing the experiments, these users were not made aware of the study. Consequently, some of them notified the university, which required us to make contact with them and alleviate their concerns on an individual basis. In future studies, we may notify study participants by email after the study is complete to mitigate this challenge.

Another challenged we faced was the confusion that arose due to the generated emails and their similarity to real emails. Since the generated emails were mined from Columbia servers, there was a high likelihood that the content and themes might be among the those currently circulating. To overcome the issue we applied natural language processing techniques to modify themes.

V. UTILITY TO THE DEPARTMENT OF HOMELAND SECURITY

The Department of Homeland Security has the lead responsibility for securing the nations information technology infrastructure for public, private, and international entities. This goal cannot be achieved solely through technological improvements, but must also take into consideration the people and processes that rely on the technology. The proposed system focuses on measuring and improving the security posture of organizations through their people. A fully developed production system that extends upon the one proposed within this work could be used to support the DHS mission of securing government departments and agencies. The proposed system could provide DHS with a means of assessing the security posture of individual departments and agencies as well as for comparing security postures of individual agencies against one another.

VI. OPEN PROBLEMS: BELIEVABILITY OF BOGUS PHISHING EMAILS

The success of the system relies on its ability to synthesize emails that are semantically equivalent to real phishing emails. Depending on the goals, resources, and sophistication of attackers, phishing emails may vary in their level of believability to a user and ability to bypass traditional security means. The least sophisticated phishing attacks are those aimed at tricking the largest sets of recipients and are sent out across the net broadly. They often contain a common invariant or signature

TABLE III
ROLE WISE SPLIT UP PERCENTAGE.

Role	1^{st} **Experiment**	2^{nd} **Experiment**	3^{rd} **Experiment**	4^{th} **Experiment**
Students	11.6	48.8	14.4	36.5
Staffs	19.6	21.6	12.8	4

TABLE IV
SCHOOL WISE SPLIT UP PERCENTAGE.

School	1^{st} **Experiment**	2^{nd} **Experiment**	3^{rd} **Experiment**	4^{th} **Experiment**
Fu Foundation	24.1	30.5	21.7	9.4
GSB	13.7	2.8	16.4	7.9
Columbia College	6.8	NA	8.2	34.5
School of social work	6.8	11.1	3.2	1.6
Grad-Arts and sciences	10.4	11.1	21.7	18.8
School of law	6.8	21.6	3.2	NA
Continuing Education	6.8	2.8	8.2	7.9
General Studies	3.4	11.1	3.2	4.7
School of public health	10.4	5.6	6.6	9.4
School of Nursing	3.4	NA	NA	NA
College of dental medicine	3.4	5.6	NA	NA
SIPA	NA	11.1	12.8	NA
College of Physicians	NA	NA	NA	1.6

TABLE V
STAFF ROLE WISE SPLIT UP PERCENTAGE.

Staff Role	1^{st} **Experiment**	2^{nd} **Experiment**	3^{rd} **Experiment**	4^{th} **Experiment**
Professor/Asst Prof	16.3	18.8	48.2	42.9
Postdoc	12.2	27.8	11.1	NA
Manager	NA	15.6	3.7	NA
Staffs	71.4	40.6	37	57.1

that allows them to be detected by email spam filters, which are commonly deployed at an organizations email gateway. They might also be obvious by direct observation of the individuals receiving, making them easily avoidable. As the specificity of targeting increases to specific companies, organizations within a company or at the extreme, an individual (i.e., spear phishing), the believability of these emails increases, making them more likely to bypass detection by an individual. Our initial approach has focused on constructing generative models for phishing emails using natural language processing and statistical techniques. So far, we have had success in modeling the least sophisticated attacks. Our future efforts will focus on generative models for the automated creation of more targeted spear phishing attacks. Developing a means to automatically create fake spear phishing emails remains an open problem. Successfully solving it requires addressing the fundamental properties for creating decoy-based systems that include variability, believability, enticingness, detectability, differentiability, and non-interference [10].

VII. CONCLUSION

The previous sections provided an overview of our system designed to create phony phishing emails. We presented the results of two rounds of experiments conducted at Columbia University in which approximately 4000 staff members and students were targeted for training using the bogus phishing emails. The results presented in the previous section suggest that users can be trained using decoy technology to be cognizant of potential threats. Applying the same set of measurements laterally across multiple organizations can be a useful in measuring one organization's security posture relative to another's.

ACKNOWLEDGMENT

This material is based on research partially sponsored by DARPA under the ADAMS program grant number W911-NF-11-1-0140 and from NSF under CCF grant 0950373.

REFERENCES

[1] S. Stolfo, S. Bellovin, and D. Evans, "Measuring security," *IEEE Security & Privacy Magazine*, pp. 72–77, 2011.

[2] W. Baker, M. Goudie, A. Hutton, C. D. Hylender, J. Niemantsverdriet, C. Novak, D. Ostertag, C. Porter, M. Rosen, B. Sartin, and P. Tippett, "2010 Data Breach Report," Verizon Risk Team and the United States Secret Service, Technical report, 2010.

[3] E. Mills. (2011, April) Attack on rsa used zero-day flash exploit in excel. CNET. [Online]. Available: http://news.cnet.com/8301-27080_3-20051071-245.html

[4] P. Kumaraguru, Y. Rhee, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, "Protecting people from phishing: The design and evaluation of an embedded training email system," in *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '07)*, San Jose, California, 2007.

[5] New York State Office of Cyber Security & Critical Infrastructure Coordination, "Gone phishing," A Briefing on the Anti-Phishing Exercise Initiative for New York State Government. Aggregate Exercise Results for public release., 2005.

[6] Core Security. (2010) Core impact pro. [Online]. Available: http://www.coresecurity.com/

[7] Phishme.com. (2011) Phishme.com. [Online]. Available: http://www.phishme.com/

[8] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, "Social phishing," *Commun. ACM*, vol. 50, pp. 94–100, October 2007. [Online]. Available: http://doi.acm.org/10.1145/1290958.1290968

[9] S. M. David Watson, Thorsten Holz. (2008) The honeynet project. [Online]. Available: http://www.honeynet.org/papers/phishing/

[10] B. M. Bowen, S. Hershkop, A. D. Keromytis, and S. J. Stolfo, "Baiting inside attackers using decoy documents," in *In Proceedings of the* 5^{th} *International ICST Conference on Security and Privacy in Communication Networks (SecureComm 2009)*, September 2009.

[11] R. V. Krejcie and D. W. Morgan, "Determining sample size for research activities," *Educational and psychological measurement*, vol. 30, pp. 607–610, 1970.