Proceedings of the 2005 IEEE
Workshop on Information Assurance and Security
United States Military Academy, West Point, NY

Community Security Awareness Training

Barbara Endicott-Popovsky, Ivan Orton, Kirk Bailey, Deb Frincke, *Member, IEEE*

NIST Special Publication 800-50 outlines standards for the development and implementation of security awareness training. [1] Recognizing that the "people factor" is the weakest link, NIST recommends that all users of any information system be made aware of their roles and responsibilities in maintaining security. [1] Further, to be effective, any awareness event should be designed for the intended audience, built around a message and desired outcomes and gain attention. [1]

Such a security awareness event was conducted for the business community leadership in Seattle, Washington. The purpose was to alert them to the risks of identity theft through misuse of online search engines. The means adopted for focusing attention, was a Google-hacking contest. It was anticipated that object lessons from this demonstration would (1) alert community leaders to take appropriate measures to ensure protection of personal and private information stored in their organizations' databases and (2) to open the way to influencing legislative change in the State, where the event sponsors contend the statutes are outpaced by technological advances.

The contest outcomes were significant. In a little over an hour, the winning team identified over one hundred million potential opportunities for identity theft. The results drew local and national attention. [2, 3] In addition, discussions have begun with the State's Attorney General's office regarding possible legislative reforms.

Based on observations of this trial, the authors suggest that a security awareness program, based on NIST standards, can be effective, not only for organizations, but for specifically defined communities, as well. This paper describes the event, the outcomes and the authors' conclusions. The approach presented in this paper could be repeatable in any community for a variety of purposes.

I. Background

On the morning of March 4, 2005, members of the Agora, a forum for airing current issues of concern among information assurance professionals, held a meeting in Seattle, Washington. The leadership had devised a security awareness event designed to raise the local business community's consciousness of the vulnerabilities of networked systems to Google-hacking[1]. [1]

Meeting quarterly for the last ten years, the group has been responsible for solving problems arising from the unintended consequences resulting from the proliferation of digital infrastructure through insecure public networks. [2, 4] Most recently, the Agora successfully tackled legislative change, at the State level, regarding cyber stalking, one of the fastest growing crimes on the Internet. [4]

Responding to a particularly egregious case involving an employee of the City of Seattle, members of the Agora undertook a two-year project of tracking down and assisting in the eventual prosecution of the stalker, but not before becoming the impetus behind some of the first cyber-stalking legislation in the nation. [4]

Having gained the attention of State legislators and local government officials by tackling a serious problem arising from misuse of public networks, the group has begun focusing on the broader issue of the vulnerability to identity theft of personal and private information housed in systems accessible from the Internet. Through security awareness training, they are attempting a two-pronged approach designed to 1) bring attention to the need to improve network and data management and 2) influence helpful legislative change.

II. Introduction

The news media now regularly features stories about database break-ins that result in the theft of thousands of names with associated credit card numbers, driver's licenses and/or social security numbers, treasure troves for identity thieves. [5, 6, 7, 8] The Federal Trade Commission reports that 1 in every 20 Americans has been a victim of identity theft in the last year. [9]

Barbara Endicott-Popovsky, Lecturer, Seattle University; Ivan Orton, JD, Senior Deputy Prosecuting Attorney with the Fraud Division of the King County Prosecutor's Office in Seattle Kirk Bailey, Chief Information Security Officer, City of Seattle Deb Frincke, Ph.D., Chief Scientist Cybersecurity, Pacific Northwest National Laboratory and Associate Professor (on leave), Computer Science Department, University of Idaho

[1] "Google-hacking" commonly refers to obtaining anything exploitable, including usernames, passwords, credit card numbers and other personal identifiable information using the search engine, Google.

0-7803-9290-6/05/$20.00 ©2005 IEEE.

While most financial institutions don't publish data on the annual amounts of such losses, an FTC report estimates the impact to the U.S. economy in the hundreds of millions of dollars, [10, 11] which the consumer ultimately pays through higher prices and fees.

Although most financial institutions have stepped up to losses resulting from unauthorized purchases on stolen credit cards, they don't take responsibility for costs related to clearing one's credit. Each incident averages $1000 in coping costs. [9]

Thus costs associated with coping following the theft of one's identity are largely born by the victim. This is true regardless of whether the theft occurs as a result of an individual's carelessness in the release of personal data or from a hacker's intrusion into a poorly managed network that allowed theft from a database housing that information. The latter seems particularly unfair since the consumer has little control over his/her personal information once it is stored in private or public databases.

The information outlined above was the impetus behind developing the security awareness event in Seattle. In addition to drawing community leaderships' attention to the vulnerability of personal and private information to unlawful discovery through public networks, the event planners hoped to gain interest in exploring possible legislative reform.

III. Addressing Lack of Awareness

As an object lesson to the community, the Agora proposed inviting business and government leaders and the press to a security awareness event using the Agora meeting forum. A Google-Hacking Contest was chosen as the best means to demonstrate the extent of the problem.

The purpose of the Google Hacking Contest was to highlight some serious security vulnerabilities for the general public's education, which would lead to a desire on the part of attendees to take corrective action.

IV. Google Hacking

Search engines, while conceived for benign purposes, can be double-edged swords. Instant information is a benefit to both the well- and the ill-intended. It is the responsibility of those making information available on the Internet to understand how search engines operate and to take private, sensitive information out of the each of web crawlers.

There is no inherent flaw in Google that led to its being selected for this demonstration, simply its widespread use and familiarity with the general public.

Employed in attacks, the Google search engine is alarmingly simple to use. [12] Many online sites exist that offer advice, instruction, tools to make this an easy-to-learn process. [13] As a means of attack, it lately has gained prominence with the appearance of a MyDoom virus variant and the Santy worm, both of which automate Google attacks. [14] Indeed the appearance of these automated attacks has led security experts to predict a 'massive increase' in such attacks this year.' [14]

While it is not the intent to do a tutorial on Google Hacking[2], a brief description is appropriate. The majority use Google to look up helpful information; however, hackers know that it also can be used "to look for information residing in other people's Web-connected servers--and machines connected to those servers. Stuff that you're not supposed to see." [12]

Scott Granneman, Security Focus, identifies the vulnerability as "uneducated folks putting content on the web they think is hidden from the world." [15]

> *Often Web servers are left configured to list the contents of directories if there is no default Web page in those directories; on top of that, those directories often contain lots of stuff that the website owners don't actually want to be on the Web. That makes such directory lists prime targets for snoopers.* [15]

Fixing the problem is not difficult. It requires awareness and following up with more thoughtfully configured network devices and some caution about what data is made public. The Google Hacking Contest discussed in the balance of the paper was designed to be a dramatic awareness event to raise the consciousness of the community's leadership about the need to mitigate against the vulnerabilities just described.

V. Agora's Google Hacking Contest Rules

Holding a Google Hacking contest requires carefully designed contest rules to ensure that no laws are violated and that everyone behaves ethically regarding what is discovered.

As responsible professionals, Agora members are obligated to ethical codes of conduct established by professional organizations, their organizations and the Agora, itself. For that reason the leadership circulated the

[2] Readers are referred to those sites mentioned in the preceding paragraph for details on the methods and tools of Google-Hacking.

0-7803-9290-6/05/$20.00 ©2005 IEEE.

following rules before the contest. These were also available as handouts the day of the event. In addition, selected monitors/contest judges were assigned to each group of contestants to ensure that rules were adhered to.

Participants were read, and agreed to abide by, the following rules while participating in the contest. Each rule is stated first, followed by an explanation that provides elaboration.

A. Rule #1: Information Protection

Rule: *All contest participants must be **VERY CAREFUL** to manage and protect any sensitive information they discover from further disclosure beyond the current exposure the data already has online.*

Explanation: The intent of the contest was to learn more as professionals about a provocative and troubling public problem. The intent was not to break any laws, harm any individual or organization, or embarrass anybody or any institution. Information found during the contest was described as 'for demonstration and instructional purposes only.' There were several law enforcement officials attending the meeting who served as reminders of participants' obligations to follow this rule above all others.

B. Rule #2: Required Gear for Competing Gear Heads

Rule: *Teams must bring their own 'stuff' in order to play. Teams will need as many mobile computing devices as they feel are necessary to win the contest, depending on what Google hacking strategy they decide to employ. Each team will need at least one box that has 802.11x connection capability, or some CDMA-type enabled service. There will be wireless access to the Internet available.*[3] *Teams also should bring at least one standard-size (8½" x 11") notepad and several manual writing devices for keeping score. In the event that we may want to show some of the hacking discoveries to the larger meeting audience, teams should bring a USB-flash drive or a CD burner.*

Explanation: The Agora collective does not have the resources to provide laptops, paper, and pens for those who participate and did not wish to presume on the hosting institution.

C. Rule #3: Respecting our Host's Internet Connection and Network

Rule: *Everyone who even thinks about using the Internet access provided by Seattle University for this Contest **WILL NOT ABUSE THIS SERVICE IN ANY WAY.** You all know what this means. If you don't know what this means, you can't play.*

Explanation: Internet access was provided by Seattle University. This service was not intended for any purpose except the Google hacking activities. No personal use or any inappropriate activity was allowed on the network. The University, as a sponsor of this and other past events, had agreements and understandings in place with Agora that needed to be honored.

D. Rule #4: Judging

Rule: *Each team will be assigned a Contest Judge before the contest begins. The assigned Judge has absolute authority over the team's information discoveries, discovery claims and scoring. The judge will also act as an observer of the team's activities to ensure all rules are observed. All judges will be briefed and prepared to apply uniform oversight and scoring tabulation.*

Explanation: This control was necessary to ensure that each team abided by the rules and that fair and equal rigor was applied to scoring tabulation for all teams.

E. Rule #5: Time Allowed for Hacking and What is to be Considered.

Rule: *Teams will be given **45 minutes** to 'have at it' with the Google search engine. When the time is up, the team with the most points wins the contest. Teams must use time wisely and efficiently. It isn't just about locating the targeted information, the discoveries have to be accurately documented and scored during the hacking timeframe. **Discovered data has to be documented (On that 8 ½" x 11" notepad) with a listing of the associated URL, brief description of the discovered document or data file.** In addition, the information content has to be reviewed for scoring by the judge. So prepare for how this might best be done before starting the actual event.*

Explanation: The amount of time allowed was closer to an hour to allow enough time to cull sufficiently interesting and thought-provoking discoveries. Additional time was allowed at the end of the contest for discussing what was found and identifying the key issues, mitigations and next steps.

[3] The Seattle University IT department provided a wireless LAN connectivity, separated from the University network.

0-7803-9290-6/05/$20.00 ©2005 IEEE.

F. Rule #6: Scoring

Rule: *Points will be awarded by judges based on the scoring criteria listed below.* (See Section G.) *Team judges only can allow points for documented discoveries made during the timed Google hacking period. If participants have been practicing their Google hacking skills prior to the contest and have previously found stuff, they are going to have to find it again during the contest with an assigned judge observing their search and discovery processes.*

Explanation: There were many ways to conduct the contest. The score card developed is similar to ones found on the Internet. The process and rules of the contest were how those assembled decided to follow. Having procedures outlined in advance helped avoid questions and challenges to the results.

G. Score Card

Points were awarded based on the following scale:

Table 1 Google Hacking Score Card

Personally Identifiable Information	**Points**
Name and Social Security Number (SSN) together	1 pt
Name, SSN, Date of Birth (DOB) together:	2 pts
Name, Credit Card number (CCN#) together	1 pt
Name, CCN#, Expiration date together	2 pts
Name, CCN#, Exp. Date, and 3-digit security code (aka CID#) together	3 pts
Name, Bank Account Number or Brokerage Account Number	1 pt
Name, Bank Account Number and PIN	3 pt
Additional data associated with each CCN# & SSN (e.g. address, phone)	0.5 pt
Name, password, and related online account identifier to anything	5 pts
Bonus points for anything above associated with a Washington State Citizen	10 pts

In addition to the points described above, an additional 500-point bonus was offered for the "Most Sensitive Document." Each team was asked to select what they thought was their most provocative and sensitive document discovered and to submit it to their Contest Judge. The judges presented these to the audience at the end of the contest for a vote.

VI. Results

Results can be divided into two categories. The results of the contest itself and the results in terms of community awareness created by this event.

A. The Contest

Eight teams were fielded in the contest. Each team consisted of between eight and twelve contestants. Not all team members had laptops; some were simply observers or 'coaches.'

There were three teams of students, one from a local technical college and two from local universities. The remaining five teams were fielded from different companies or industries. One of the five was a group of attorneys with significant knowledge of computers and information assurance.

The remainder of the audience of approximately 300 roamed the ballroom where the event was held, observing the results and learning from the experience.

The follow is a partial list of contest results [2]. Due to the sensitivity of some of the information uncovered, it will not be reproduced here. Anything of a sensitive nature was referred to the appropriate parties following the event.

1) Credit card numbers of military personnel,
2) A million Social Security numbers of recent immigrants, their tax records and addresses,
3) Names, birth dates, Social Security numbers, race and religion of deceased military personnel,
4) Names, credit card numbers, birth dates and home phone numbers of 388 Americans who appeared to have ordered pornographic movies online from a Brazilian web site,
5) More than one hundred million death certificates with Social Security numbers, dates of birth and city of last residence,
6) Highly personal information of two individuals, along with their level of government security clearance[4]. One was an expert in virology investigations and the other a responder to nuclear emergencies,
7) Personal information about people on terrorist watch lists,

The winning group was a pick-up team of lawyers and computer-security specialists. They won with over 190 million points. It was their group that discovered the database containing millions of names and social security numbers of deceased persons. They also won the bonus for most sensitive information--the highly personal

[4] Clearance information isn't classified, up through fairly high levels. It is therefore permissible to have it on resumes, for instance. Nevertheless, that information linked with additional data (such as found in this case) could be problematic in the hands of the wrong individuals.

0-7803-9290-6/05/$20.00 ©2005 IEEE.

information and clearances of individuals working sensitive government projects (See above).

Coming in at a distant second was a group of penetration testers from a local network security firm, scoring 13 million points. The student groups clustered near the bottom.

From the results, it appeared that having the experience to know where to look was an advantage for the attorneys.

B. Community Awareness

The event was attended by a local reporter with an interest in cyber crime. This individual had followed Agora's work with the cyber stalking initiative and was intrigued and concerned about what he learned at the event. [2] A front-page article was published the next day in his morning newspaper that drew both local and national attention. [2, 3] A syndicated columnist is featuring the article on her daily blog. [3]

Those attending, representing both the public and private sector, had their eyes opened. As each team's report was read before the audience, audible gasps could be heard when the quantity and sensitivity of the information discovered was particularly significant.

One of the attendees, a CEO of a local network security firm summed up the experience by saying:

> "*The problem is not with Google, but with corporate cultures with the attitude, "Nobody is going to find me, nobody cares what's on my computer." These companies allow Google to enter into the public portion of their networks, sometimes called the DMZ, and index all the information contained there.*" [2]

An information security specialist added that

> "*Google doesn't need to be fixed. Companies need to understand that they are leaving themselves exposed by posting sensitive information in public places .... If they're performing proper security, then their intranet shouldn't be vulnerable to a Google search engine.*" [2]

These were the lessons those planning the event hoped participants would take from the meeting.

In addition, Attendees from government made a report to the new Washington State Attorney General who has made cracking down on cyber crime one of his main agenda items. Agora members expect to continue contact with his office to explore possible legislative avenues for increasing data security protections in the State. This will be covered more fully in the Conclusions and Future work section.

VII. Lessons Learned

These are several lessons learned from this experience:

1) While we ordinarily think of security awareness training as appropriate in the work place, it is possible to design and deliver an effective security awareness event to begin the education process for a specifically identified community regarding information assurance vulnerabilities.
2) A Google-Hacking contest is an effective means of communicating to lay people the vulnerabilities companies potentially have when they make information available on the web. Google hacking is easier to understand than some of the complex kinds of attacks that can occur online and it can deliver an object lesson in the possible consequences of ignoring vulnerabilities to this kind of attack.
3) Such a contest is easy to stage. It requires a simple wireless LAN set up independent of the host organization's network. It was fun for the participants, who were willing to provide their own equipment and supplies.
4) It is helpful to notify attendees well in advance so that teams can form and work logistics issues in advance (numbers of computers, etc.) This also allows participants to familiarize themselves with Google hacking before coming. (The student groups used the opportunity as a school project for the term. They learned while having fun.)

VIII. Conclusions and Future Work

The intent of the event outlined in this paper was to raise awareness among a community's leadership of the vulnerabilities of data held in private and public databases to Internet attacks, specifically attacks generated by using a well known search engine. The planners concluded that from this point of view the event was a success.

In addition, planners hoped to interest members in influencing the adoption of further legislation addressing the protection of personal and sensitive data stored in databases over which the owner of that information has little or no control. Avenues for follow up have resulted along both lines of inquiry.

A. Future Awareness Training Events

The awareness training conducted by the Agora is transferable to the University of Washington's Center of Information Assurance and Cyber Security, a newly designated NSA Center of Academic Excellence, as an outreach project to the community. This possibility is under consideration. The Center also offers more in depth education in information assurance through certificate

programs for those interested in increasing their knowledge after attending an awareness session.

Any additional security awareness events would benefit from the collection of pre- and post-training data that would be useful, quantifiable measures of the effectiveness of the content and approach.

B. The Fairness Proposition

As a result of the meeting, it was apparent to some that having the victim bear the associated coping costs due to the misuse of personal information stored in trust on public and private databases is unfair.

Figure 1 captures the unfairness of the current situation.

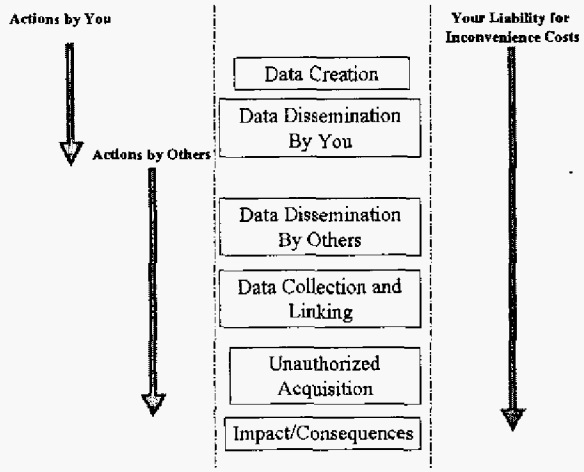

Figure 1 The Unfairness Proposition

An individual's data is shown moving through several different states, from its creation (receiving a social security number when born) to its dissemination, either with or without consent. Activities resulting in state changes are largely under the control of others, while liability for the inconvenience costs and impacts[5] due to misuse are born entirely by the individual.

A simple fairness proposition would propose the following:

• Individuals **should** bear the inconvenience costs associated with misuse of the portions of the creation and distribution of any personal information that they control

• Individuals should **not** bear the inconvenience costs associated with misuse of the portions of the distribution of their personal information that they **do not** control

While the fairness proposition appears obvious, it is not reflected in current law. Some in attendance at the event intend to pursue the possibility of influencing the development of legislative remedies. The desired outcome is shown in Figure 2 where the parties responsible for storing and forwarding data are the ones held liable for costs arising from misuse.

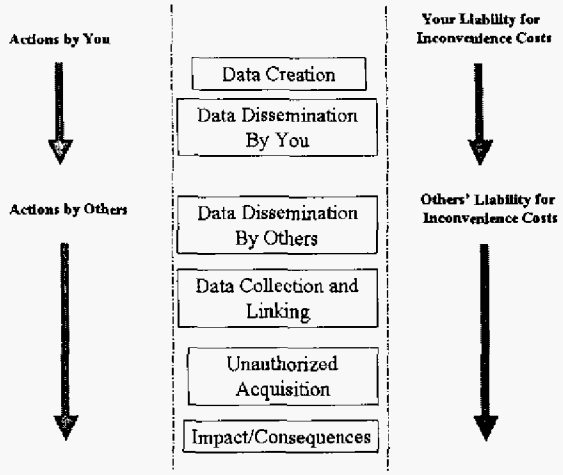

Figure 2 The Fairness Proposition

From the preceding discussion, the event achieved its goals to: (1) alert community leaders to take appropriate measures to ensure protection of personal and private information stored in their organizations' databases and (2) to open the way to influencing legislative change in the State, where the event sponsors contend the statutes are outpaced by technological advances.

IX. REFERENCES

[1] Wilson, M. and Hash, J. (2003). "Building an Information Technology Security Awareness Training Program." U.S. Department of Commerce, NIST Special Publication 800-50.

[2] Shukovsky, P. " 'Good guys' show just how easy it is to steal ID" Seattle Post Intelligencer. March 5, 2005. (Retrieved from the Web March 19, 2005). http://seattlepi.nwsource.com/local/214663_googlehack05.html

[3] Malkin, M., "Google Hacking," Michelle Malkin's Blog. (Retrieved from the Web March 19, 2005). http://michellemalkin.com/archives/001680.htm

[5] Impacts associated with misuse of private data are significant. [10] Aside from the out-of-pocket coping costs identified above, a resulting loss of credit can interfere with the ability to get a job, apply for a mortgage, buy a car, etc.

[4] Shukovsky, P. "Law banning cyberstalking is a victory for a victim." Seattle Post Intelligencer. March 25, 2004. Retrieved from the Web March 19, 2005. http://seattlepi.nwsource.com/local/166204_cyberstalk25.html

[5] InfoSec News. "Audit: State voter system left information vulnerable." March 18, 2005. (Retrieved from the Web March 19, 2005) isn@c4i.org

[6] CBS News.com "Alleged Database Hacker Arrested" Sept. 10, 2003. (Retrieved from the Web March 19, 2005) http://www.cbsnews.com/stories/2003/09/10/tech/main572449.shtml

[7] Wired.com. "How to Foil Data Thieves, Hackers," Jan. 20, 2003. (Retrieved from the Web March 19, 2005). http://www.wired.com/news/infostructure/0,1377,57302,00.html

[8] McWilliams, B. "Hackers Arrested for E-Commerce Site Break-ins" Internet news.com March 24, 2000. (Retrieved from the Web March 19, 2005). http://www.internetnews.com/ec-news/article.php/4_327181

[9] United Nations report of the Secretary General. "'The State of Crime and Criminal Justice World Wide." 10th United Nations Congress on the Prevention of Crime and the Treatment of Offenders. Vienna (April, 2000).

[10] Synovate. Federal Trade Commission Identity Theft Survey Report. September 2003. (Retrieved from the Web March 19, 2005). http://www.consumer.gov/idtheft/stats.html

[11] Identity Theft Statistics. (Retrieved from the Web March 19, 2005) http://www.bhs.k12.nj.us/coltech/statistics.htm .

[12] Ong Boon Kiat "Google hacking for beginners." Cnet Asia, November 8, 2004. (Retrieved from the Web March 19, 2005). http://www.zdnet.co.uk/zdnetuk/comment/other/0,39020682,39172957,00.htm

[13] Googledorks. (Retrieved from the Web March 19, 2005).http://johnny.ihackstuff.com/index.php?module=prodreviews

[14] Kotadia, M. (1977). "Protect yourself from 'Google hacking' ". Silicon.com, Jan. 14, 2005. (Retrieved from the Web March 19, 2005). http://networks.silicon.com/webwatch/0,39024667,39127080,00.htm.

[15] Granneman, S. "The Perils of Googling" Security Focus (Retrieved from the Web March 19, 2005). http://www.theregister.co.uk/2004/03/10/the_perils_of_googling/

0-7803-9290-6/05/$20.00 ©2005 IEEE.