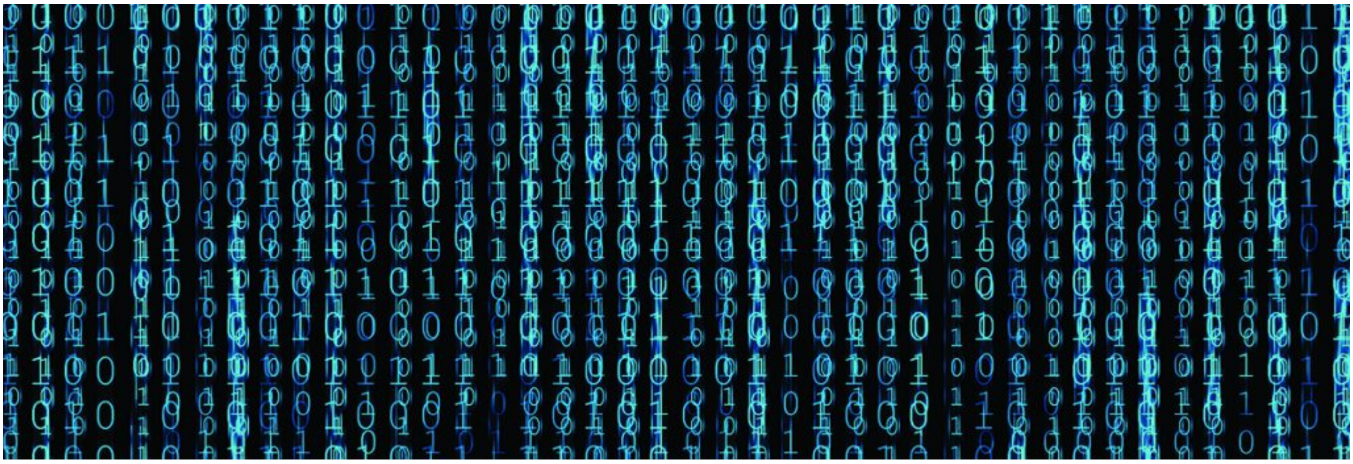
Psyber Security: Thwarting Hackers with Behavioral Science

October 31, 2017



The email included a seemingly normal request from toymaker Mattel's new CEO, Christopher Sinclair, requesting that a vendor in China be paid. Compliantly, the executive who received the email wired more than $3 million to a Chinese bank. But when she later mentioned the payment to Sinclair, he was shocked. He hadn't made the request.

Mattel was a victim of a cyber threat known as the fake CEO scam — a form of electronic fraud that has, according to the US Federal Bureau of Investigation, cost organizations billions in losses over the past 3 years.

From viruses to ransomware and password phishing scams, cyber fraud and other digital security threats are a major worldwide concern. In the month of September alone, some of the biggest names in business — Yahoo!, Equifax, Virgin America, Deloitte — disclosed major data breaches and hacks that affected hundreds of millions of consumers. Now, governments and organizations around the globe are turning not just to computer scientists, but also to psychological scientists to keep their data safe.

Human behavior, just as much as technology, is at the crux of cybersecurity. Hackers and scammers target computer systems, but many of them also attack our biases and cognitive vulnerabilities.

Cyber Inception: Deceiving the Deceiver

The year 2006 marked a major turning point in cybersecurity attacks; spies hacked into military contractor Lockheed Martin's computers and made off with millions of proprietary documents pertaining to the Pentagon's F-35 Joint Strike aircraft. This style of attack, dubbed an Advanced Persistent Threat (APT), has been dramatically increasing and characterizes many of the most high-profile cyberattacks of the past few years.

There isn't an exact definition for an APT, but what these breaches have in common is a diverse set of tactics aimed at incessantly targeting a specific victim — a company, an organization, or even a government. One characteristic that makes these attacks particularly pernicious is their exploitation of our cognitive weaknesses: Unlike some other forms of cyberattack, these onslaughts often rely on simple acts of deception and social manipulation rather than cutting-edge technology.

Most APT attacks are unleashed when a single individual unwittingly opens a “contaminated” link or a document that delivers customized malware to infect the entire computer network. Once an organization’s network is compromised, the attackers can then attempt to access files and data stored on the network. Typically, APT attacks are stealthy — attackers use methods that keep a low digital profile to avoid detection, and malicious code may stay hidden in systems for months or even years, providing attackers with a steady stream of information. Organizations often aren’t even aware of the attack until it’s too late.

Psychological scientist **Nancy Cooke**, left, is part of a team that is utilizing behavioral research findings, computer systems engineering, and game theory in an innovative development of defensive strategies against cyber attacks.

Behavioral scientists Cleotilde (Coty) González and Nancy Cooke are part of a major new research effort to study the psychology of deception in order to counter APT attacks. They are members of an integrative team that was recently awarded a $6.2 million Multidisciplinary University Research Initiative grant from the US Department of Defense. The team will be utilizing research from behavioral science, computer systems engineering, and game theory as part of an innovative effort to develop active defense strategies against APT attacks.

The researchers’ ultimate goal is to thwart cyberattacks by learning how to “deceive the deceivers.”

González is a Professor of Decision Sciences and the founding director of the Dynamic Decision Making Laboratory at Carnegie Mellon University, where she conducts studies using cognitive computational models to support decision-making in dynamic environments. Cooke is a cognitive psychologist in the Human Systems Engineering program in the Ira A. Fulton Schools of Engineering at Arizona State University and studies teamwork and human performance in complex environments. Her lab will gather and evaluate data from teams of participants engaged in an advanced cyberattack simulation.

“Advanced persistent threats originate from humans,” Cooke explained. “These threats can therefore only be understood and mitigated through understanding of humans and human factors.”

González and her team will develop cognitive models of attackers based on Instance-Based Learning Theory, which draws from memory research, decision science, and machine learning. When a person assesses a situation in which they must make a decision, they retrieve memories of past events and experiences and compare them with the current situation. This process has some memory biases that defense mechanisms can take advantage of in order to deceive the attackers.

In the example of a cyberattack, experiences and memories of past threats determine how aggressively a security analyst might respond to the ambiguous early signs of danger. In APT campaigns, attackers are persistent and likely to learn and adapt to defenses over time. González, Cooke, and colleagues will focus on developing their own deceptive tactics that are just as adaptable and versatile. They call this methodology Cyber Inception.

“This new approach to cybersecurity will exploit the psychology of deception to lure attackers into believing that they have successfully compromised a system, while keeping our systems safe,” González said.

Eventually, the data from behavioral experiments and cognitive models will be used to fine-tune sophisticated algorithms capable of detecting APT attacks.

Computer Scientists Embrace Social Psychology

By training, Jason I. Hong is a computer scientist. He helped found a startup cybersecurity company called Wombat Security Technologies, and his Computer Human Interaction: Mobility Privacy Security lab at Carnegie Mellon University researches usable privacy and security. However, he’s recently been at the helm of a fascinating series of large-scale social psychology experiments.

“The ‘light bulb’ moment for me happened one day at my startup,” Hong explained. “Two women were talking to each other about a recent event. One said, ‘Did you hear what happened to Moe? He slipped on the ice [and dropped his laptop], and now can’t access the files on it.’ The other women said, ‘I’m going to back up my data right now.’ And she did!

“It immediately struck me that this was a positive example of social influence and behavior change for cybersecurity. I had heard my colleagues in the behavioral sciences talk about concepts like social proof, commitment, and reciprocity for years, and it all crystallized in my head based on this one event that we could also use these kinds of techniques to solve hard problems in cybersecurity.”

Hong’s interest in social psychology emerged from working as an associate professor in the Human Computer Interaction Institute at Carnegie Mellon University.

“Psychologists, designers, and computer scientists are all sitting next to each other,” Hong said. “So over the past few years, I’ve slowly absorbed many of the theories and methods used by these other disciplines.”

With a recent grant from the National Science Foundation (NSF), Hong and Laura Dabbish are the principal investigators on a project exploring the use of social influence to encourage safer cybersecurity behaviors.

And Hong and colleagues collaborated with social media giant Facebook on a massive experiment inspired by social-proof research from APS Fellow Robert Cialdini.

In his 2006 book *Influence: The Psychology of Persuasion*, Cialdini explains how social influences play a vital role in how we make decisions. When we are unsure of the appropriate course of action — say, adopting a security feature versus using a stronger password — we look to people around us for what Cialdini has dubbed “social proof.” One of the biggest challenges in convincing people to adopt safer cybersecurity practices is that people simply don’t have much opportunity to observe each other’s behavior.

“Our experiment with Facebook was based on two insights,” Hong said. “The first is that cybersecurity has low observability. I don’t know how good your passwords are or what security settings you have, and vice versa. This lack of observability makes it hard for good practices to diffuse through a social network.

“The second is that we could use social proof to positively influence people’s awareness, knowledge, and motivation to be secure. Facebook already had data about who was using various security features.”

Facebook’s Site Integrity team wanted to encourage users to take advantage of more of the platform’s security features, such as activating Login Notifications, Login Approvals, and Trusted Contacts.

A team led by Hong’s student, Sauvik Das, wanted to see whether increasing the observability of cybersecurity social norms could persuade more users to adopt these security features. The research team showed a sample of 50,000 active Facebook users one of eight possible security announcements prompting them to adopt these security features.

The seven social-proof messages informed users that their Facebook friends were already using these security features. These messages varied in specificity and phrasing — from showing the exact number of friends to just saying “some” friends. A control group received a message without any social-proof framing (i.e., “You can use security settings to protect your account and make sure it can be recovered if you ever lose access”).

“We found that while all of our social-proof-based interventions were effective, simply showing people the specific number of their friends that used security features without any subjective framing was most effective — driving 37% more viewers to explore the promoted security features compared to the non-social announcement,” the researchers wrote.

Over the following 5 months, both conditions continued to generate more views of the security features compared with controls.

Interestingly, getting people to click through to the promoted features didn’t necessarily mean that people were ready to adopt them: There was no difference in the actual adoption rate of those who viewed a social prompt compared with a nonsocial announcement.

A follow-up survey confirmed that the social announcements raised viewers’ awareness of available security features. However, individuals in the control condition who clicked through for more information may have had higher intrinsic motivation for using security features, the research team points out.

Personality Traits and Risk

We are all vulnerable to cybersecurity attacks, but research indicates that a small segment of the population appears to be particularly at risk. Carl Weems, a professor of human development and family studies at Iowa State University, is part of an integrative team investigating whether certain personality traits predispose individuals to higher risk of careless cyber behaviors.

Weems’s main research area is emotional development and traumatic stress, but he has always been interested in the translation of basic psychological science into answering important applied questions. Along with University of New Orleans computer science professors Irfan Ahmed and Golden Richard III, Weems recently received a grant from the NSF to investigate personality factors in cyber security.

“The goal of this project was to utilize the methods of psychological science to build a platform and techniques for predicting secure versus insecure cyber behavior,” Weems said.

Weems and colleagues collected an initial set of psychometric data from a socioeconomically and ethnically diverse sample of 210 adults. Participants reported how often they carried out 20 security-related tasks drawn from a pool of cybersecurity recommendations.

The researchers found that most participants seemed highly engaged in security-enhancing behaviors, while only a few reported engaging heavily in the kinds of practices that compromise security.

Participants also completed a series of personality measures. As might be expected, highly conscientious people were less likely to engage in insecure behaviors. However, contrary to their hypothesis, the researchers found no link between high neuroticism scores and secure behaviors. Participants scoring high on aggression, depression, and trait anxiety also scored significantly higher on the insecure behavior scale.

Weems and colleagues have been developing a highly customizable research tool that other researchers can use to study the relationship between personality traits and cybersecurity behavior. The Software Package for Investigating Computer Experiences is a script-based product that provides an easily modified platform for analyzing security behavior and personality. It’s designed to capture data detailing the personality traits and cyber behaviors of a large population of users, and to create data sets for studying the variations of cyber behavior across different personality types.

To start, Weems and colleagues have been examining the link between risky cyber behavior and both trait anxiety and the callousness–unemotional trait. The researchers are using a standard cognitive assessment — emotional dot-probe tasks (see sidebar) — to assess the personality traits.

After completing the personality trait assessment, participants engage in a realistic scenario, assuming the role of a new employee at an accounting firm. As they complete mundane office tasks like reading emails, checking stocks, and completing accounting math problems, they're simultaneously prompted with realistic decisions about cybersecurity: phishing emails, software update requests, and antivirus scanning.

This multitasking environment allows researchers to collect fine-grained data on what people actually do when faced with cybersecurity decisions.

Weems and colleagues deliberately made it easy to modify the contents and manage the flow of different events in the scenario, allowing other researchers to highly customize their own experiments.

"An important initial step in actualizing the benefits of psychological research on cybersecurity is to empirically establish the ability to measure the dependent variable of secure and insecure behavior so that typical cognitive and behavioral experiments on predictors can be conducted," Weems said.

References

Cialdini, R. B. (2006). *Influence: The Psychology of Persuasion*. New York City, NY: HarperCollins.

Cooke, N. J., Champion, M., Rajivan, P., & Jariwala, S. (2014). *Cyber Situation Awareness and Teamwork*. ICST (Institute for Computer Science, Social Informatics, and Telecommunications Engineering) Transactions of Security and Safety. Special Section on: The Cognitive Science of Cyber Defense. doi:10.4108/trans.sesa.01-06.2013.e5

Das, S., Kramer, A. D., Dabbish, L. A., & Hong, J. I. (2014). Increasing security sensitivity with social proof: A
large-scale experimental confirmation. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security* (pp. 739–749). ACM. doi:10.1145/2660267.2660271

Dutt, V., Ahn, Y. S., & Gonzalez, C. (2013). Cyber situation awareness: modeling detection of cyber attacks with instance-based learning theory. *Human Factors, 55,* 605–618. doi:10.1177/0018720812464045

No author. (2016, March 29). Mattel vs. Chinese cyberthieves: It's no game. *Associated Press*. Retrieved from www.cbsnews.com/news/mattel-vs-chinese-cyberthieves-its-no-game

Russell, J. D., Weems, C. F., Ahmed, I., & Richard III, G. G. (Advance online publication). Self-reported secure and insecure cyber behaviour: factor structure and associations with personality factors. *Journal of Cybersecurity Technology*, 1–12. doi:10.1080/23742917.2017.1345271

Tamrakar, A., Russell, J. D., Ahmed, I., Richard III, G. G., & Weems, C. F. (2016, March). SPICE: A software tool for bridging the gap between end-user’s insecure cyber behavior and personality traits. In *Proceedings of the Sixth ACM Conference on Data and Application Security and Privacy* (pp. 124–126). ACM. doi:10.1145/2857705.2857744