Cybercriminals: Things are about to get a lot more confusing for you

12 January 2020, by Daniel Tkacik

Credit: CC0 Public Domain

There are three boxes on a table. Two are made of cardboard and sealed with packaging tape. The third is made of steel with a series of locks blocking entry. *Obviously*, you think, as an imaginary criminal, *the goods are in the steel box*. After successfully picking the locks, you realize there's nothing inside. As you stare into the empty box, authorities grab your arms from behind and, all of a sudden, you're in handcuffs.

The steel box, you quickly realize, was a trap. The valuables were actually in one of the cardboard boxes, but the locked box looked more attractive. It's too late; you got duped.

This kind of deception is being applied to the elusive world of cybersecurity, and CyLab's Cleotilde Gonzalez is leading the charge.

"You don't necessarily need more resources for security defense," says Gonzalez, a professor in Social and Decision Sciences at Carnegie Mellon. "You just need to strategically manipulate the information that is shown to the attacker."

This concept of cyber deception is the focus of two papers, authored by Gonzalez and colleagues, that are being presented at this week's Hawaii International Conference on System Sciences (HICSS). These studies were funded by a Multi-University Research Initiative grant from the Army Research Office.

While cyber deception isn't totally new—researchers have been looking into this technique for a few years now—Gonzalez's group takes a unique approach: using cognitive science to inform how to deceive attackers effectively.

"Prior to our group's work, researchers operated on the assumption that attackers are rational," says Gonzalez. "But humans are not rational. Humans get biased by the frequency and recency of events, and other cognitive factors. The defense algorithms can take advantage of the attackers' cognitive biases to become more effective."

In the first paper being presented at HICSS, Gonzalez and her team show that *signaling* – a strategic technique used to make something seem like something it isn't—can be an effective way of throwing attackers off their game.

"Through signaling, we can manipulate the information that is sent to the attacker," says Gonzalez. "We try to make them believe that there is something really valuable in this node of the network, for example, so we can trace their steps in the system. Then, our defenders will have more time to detect that they are under attack."

In a second paper being presented at HICSS, Gonzalez and her team show that a previously-developed signaling technique can be improved using an advanced cognitive model, and back their results with human experiments.

The researchers used a video game to study the effectiveness of different signaling schemes on

human attackers. In the game, players try to score points by attacking computers, but they must tread carefully as some computers may be monitored by defenders; attacking those computers result in a deduction of points, and in the real world, attacking such computers result in getting caught. When a player selects which computer they'd like to attack, a signaling algorithm determines whether to send a truthful or a deceptive signal.

A truthful signal might indicate that a node is monitored by the defender, which will deter the attack. A deceptive signal might indicate that a node is not monitored by the defender which will motivate an attack, when indeed it is protected.

"Balancing the rate and the timing at which an attacker is sent deceptive signals is crucial," says Gonzalez. "We have to maintain their trust in order for deception to work."

The paper has been nominated for a Best Paper Award at the conference.

Provided by Carnegie Mellon University