Proceedings of the 54th Hawaii International Conference on System Sciences | 2021

Introduction to the HICSS-54 Minitrack on Cyber Deception and Cyber Psychology for Defense

Kimberly Ferguson-Walter
Laboratory for Advanced Cybersecurity Research
kjfergu@niwc.navy.mil

Sunny Fugate
Naval Information Warfare Center, Pacific
fugate@niwc.navy.mil

Cliff Wang
Army Research Office
xiaogang.wang.civ@mail.mil

Matt Bishop
University of California, Davis
mabishop@ucdavis.edu

The defense and defensibility of our systems and networks is often conflated with the culpability of human decision-makers. This is reflected in the often-quoted turn of phrase that “the human is the weakest link”. Understanding the analogy is nearly automatic. The weakest link in a chain will fail before the others. Humans often make bad security choices which cause failures in security. This faulty reasoning leads to the conclusion that we should simply remove human decisions from security and that will make our systems more secure.

This philosophy is emblematic of modern cyber defenses where corporations and government organizations seek turnkey solutions which promise improved security while requiring only minimal manning and maintenance. More recently focus has shifted from mere automation to the application of machine learning and other artificial intelligence techniques. However, the human is simply not a removable component. And the user is not the only human implicated in a cyberattack.

The user may fall prey to deception in the form of a phishing email, the email client being used hiding security relevant information and placing the threat alongside emails from known sources, biasing the user to trust its content simply through its adjacency to trusted sources. Automated detection tools quickly identify the malicious nature of files transiting the network, but instilled with the biases of the security engineer, these tools are not capable of asking the user how the malicious file was obtained. The cyber defender triages events as they are received and preemptively locks the infected computer against further use, unaware that this action will directly inform the attacker of the detectability of the malicious file. The attacker continues with impunity by sending a new malicious file to an as yet unaffected user, achieving simultaneously a defender-imposed denial of service and eventual success in implanting undetectable malware within the network. Every component of this scenario is a weak link.

Over the last decade, our respective communities have begun to recognize the need to study the human decision makers in cyber environments and to do so from all facets and perspectives as they relate to the security and defensibility of our systems. Extending this endeavor, our communities have also realized the value in understanding, informing, protecting, and exploiting human psychological biases.

By leveraging knowledge of a user’s deficiencies, automated systems might better inform them of threats. By exploiting an attacker’s inherent biases, we might create systems which are able to enjoy the advantage of Stackelberg’s first mover’s advantage, delaying an attacker’s progress and deterring further action through direct and intentional deception. By exposing the biases and preferences of an attacker, we might enable defenders to understand the motivation for an attack and to thus take actions which prevent the goals of the attacker rather than solely mitigating the immediate perceived threat. By demonstrating the relationship between user, system, and network, we may embolden security engineers to incorporate human perception and cognition rather than eschewing them – refocusing the purpose of security to achieving the user’s mission rather than simply preventing the immediate attack or falling headlong into misdirection.

Lastly, and most importantly, this could enable identification of an insider who is ever present. Authorized users of systems and networks, or masqueraders of these users, may act as attackers. These insiders can leak sensitive information deliberately or accidentally. Detecting and thwarting these attacks is more difficult than dealing with external attackers because the users are often

URI: https://hdl.handle.net/10125/70852
978-0-9981331-4-0

authorized to access the information, or the systems and networks the data is on. As a defensive technique, deception in this context must take into account the psychology of the attacker and the organizational, political, and societal environments in which the attack occurs.

Looking back on decades of work by cognitive psychologists, and highlighted by current events, it is easily realized that defender, attacker, and even bystander are often the victims of cognitive bias and the incredulous sycophants of fallacious reasoning. It is unarguable that both attackers and defenders are susceptible to decision-making bias. Bias, in some situations, can cause as much harm as direct and conscious action. Thus, it should be our goal to protect the defender from the insider threat of their own inherent cognitive biases, while adversely affecting the attacker using the same.

The principal open questions for our community are whether and how we are able to incorporate knowledge of this bias into the manner in which we defend our systems and users. How should these systems incorporate knowledge of human behavior and psychology for defensive purposes? What are useful metrics and measures to quantify the effectiveness of these techniques and systems? To what extent do the capabilities of defensive deception and the characteristic features of cyber psychology fundamentally change how we secure our systems, protect our networks, and minimize harm to people?

Cyber deception is a collection of defensive techniques that considers the human component of cyberattacks. Deception holds promise as a successful tactic for making an attacker's job harder by moving beyond mere perimeter defenses. It can disrupt or delay progress of a persistent attacker by wasting their time, resources, and effort. Moreover, deception can be used strategically by a defender to develop specific incorrect beliefs in the attacker, the effects of which can persist over time. Understanding the cognition and behavior of both the cyber defender and cyber attacker is a critical component. Cyber psychology research advances the science of human behavior and decision making in cyberspace to understand, anticipate, and influence attacker behavior. It also seeks to ensure scientific rigor and quantify the effectiveness of our defensive methods.

These research efforts require an interdisciplinary approach and the mini-track is pleased to present papers across multiple disciplines. This year the minitrack features five papers. The first group of papers focus on cyber psychology: (1) providing analysis to identify gaps in cyber incident response teams, and (2) providing guidelines towards building automated cybersecurity technologies based on established human factors theory.

1) *Identifying Expertise Gaps in Cyber Incident Response: Cyber Defender Needs vs. Technological Development* (by Megan Nyre-Yu)
2) *Human Factors in Automating Cyber Operations* (by Robert Gutzwiller and Dirk Van Bruggen)

The next set of papers features in this minitrack is focused on building cyber deception technologies: (3) using a game-theoretic framework to automatically select which emulated software stack will provide the highest defender payoff, and (4) using an autonomic reasoning approach for a system to perform automated countermeasures using adaptive deception techniques.

3) *Software Deception Steering through Version Emulation* (by Frederico Araujo, Sailik Sengupta, Jiyong Jang, Adam Doupé, Kevin Hamlen, Subbarao Kambhampati)
4) *Towards Self-Adaptive Cyber Deception for Defense* (by Jason Landsborough, Braulio Coronado, Luke Carpenter, Sunny Fugate, Kimberly Ferguson-Walter, Dirk Van Bruggen)

And finally, we introduce the paper nominated from this minitrack for Best Paper this year, which is a lovely pairing of both cyber deception and cyber psychology: (5) details the results from a design thinking workshop conducted using experts from different fields including critical analysis of design provocations for cyber deception and a journey map providing considerations for operationalization of cyber deception technologies.

5) *Design Thinking for Cyber Deception* (by Debi Ashenden, Rob Black, Iain Reid, Simon Henderson)

We believe that our community is at the cusp of rapid advances in defensible human-machine systems. The multidisciplinary nature of the work represented in the minitrack this year is representative of the novel intersection of psychology, human-machine interaction, artificial intelligence, and cyber defense. We are eager to facilitate discussions based on these new and exciting contributions to the community.