# Human Factors in Cybersecurity and the Role for AI

**Ranjeev Mittu[1] & William F. Lawless[2]**

[1]Naval Research Laboratory, 4555 Overlook Ave SW, Washington, DC 20375
[2]Departments of Math and Psychology, Paine College, Augusta, GA 30901
[1]ranjeev.mittu@nrl.navy.mil; [2]wlawless@paine.edu

**Abstract**

In this paper, we review the pervasiveness of cyber threats and the roles of both attackers and defenders (i.e. the targets of the attackers); the lack of awareness of cyber-threats by users; the complexity of the new cyber environment, including cyber risks; engineering approaches and tools to mitigate cyber threats; and research to identify proactive steps that users and teams can take to reduce cyber-threats. In addition, we review the research needed on the psychology of human users that pose risks to all users from cyber-attacks. For the latter, we review the available theory at the individual and group levels that may help individual users, groups and organizations take actions against cyber threats. We end with future research needs and conclusions.

## The cyber problem

This paper has been abstracted from a much longer chapter now being published (i.e., Lawless et al., 2015).

In our approach to understanding cyber threats, we reviewed the increasing complexity of, and risks in, the new cyber environment. We discussed cyber tools used to mitigate cyber threats. More fully, we reviewed and discussed the pervasiveness of cyber-attacks from multiple perspectives: first at the individual level from the perspective of the human attacker and the user, the attacker's target; and second from the perspective of teams and organizations. We closed with future research needs and conclusions.

## Our modern digital age

Briefly, we live and work in a digital age, where access to information of widely varying values is ubiquitous. However, autonomous agent and human users fail to compute or comprehend the value of personal information (e.g., birthdays, on-line browsing behavior, social interactions, etc.) to malicious actors and states. Information has always been important to survival; the original purpose of the internet was to share the information that would improve global social well-being (Glowniak, 1998). The security of information has been defined as "… protecting information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction …" (LII, 2014). Security directly describes how well a system is protected and indirectly the value of the information being protected (Lewis & Baker, 2014). However in the modern digital age, sharing information now competes with protecting private information from unintended recipients; the complexity of security has increased to protect information that is deemed private, and the interaction between the complexity of networks and security defenses has led to increasing opacity in the functioning of networks and computers for typical users. Furthermore, as complexity increases with greater security features, systems and protocols, the "increased use of networked systems introduces [even newer] cyber vulnerabilities …" (Loukas et al., 2013).

Reporting on the number of cyber incidents, the GAO (2014) reported:

> Cyber Incidents Reported by All Federal Agencies to US-CERT, Fiscal Years 2010-2013: The number of incidents by Federal agencies increased from 34,048 in 2010 to 46,160 in 2013.

The findings in the GAO report are illustrated vividly by the recent cyber-attack on the Sony Company as reported in the *New York Times* (Cieply & Barnes, 2014):

> ... the F.B.I. found that the hackers had used digital techniques to steal the credentials and passwords from a systems administrator who had maximum access to Sony's computer systems. Once in control of the gateways those items opened, theft of information was relatively easy. Government investigators and Sony's private security experts traced the hacking through a network of foreign servers and identified malicious software bearing the familiar hallmarks of a hacking gang known as Dark Seoul. ... Now, five weeks into the episode, Sony's internal technology is still impaired. Executives estimate that a return to normal is at least five to seven weeks away.

An op-ed appeared in the *Wall Street Journal* to discuss the value of improved (AI) software to detect cyber threats (Raul, 2015):

> Today U.S. commerce is threatened by digital Barbary pirates. The most sophisticated companies with every incentive to protect their crown jewels—intellectual property, confidential business information or customer records—are being ransacked and held hostage by cyberterrorists, state-sponsored hackers and highly effective organized cybercriminals. No corporation today is immune or can realistically believe itself adequately protected. ... Currently the federal government has a network-inspection tool, called EINSTEIN, to protect certain federal communications. If EINSTEIN is in fact working, the government should make it available more broadly. New technologies also need to be developed and deployed, and the government should make the investments in the necessary research as well as in so-called "active" defense and intelligence measures designed to protect private networks before they are successfully compromised. This means aggressively tracking, tracing, deceiving, disrupting and punishing the cyber bad guys and their state-sponsors or protectors.

A review of the first Einstein software package can be found at DHS (2004), with updates noted there.

## Discussion

In the following, we discuss a few of our findings from our full review.

First, the cyber environment is becoming more and more complex along with the cyber-threats. For example, "By 2020 Cisco estimates that 99% of devices (50 billion) will be connected to the Internet. In contrast, currently only around 1% is connected today" (e.g., Rosenquist, 2014). Even defenses are becoming complex, whether a defense is passive or active (e.g., despite our lengthy review of cyber defenses, we omitted numerous defenses, such as the use of encrypting emails, randomly generating passwords, using peer networks to increase security, hardening websites, etc.; from FIPSP, 1993; Intel, 2014; respectively). One of the problems with defending a website against cyber threats is that the relative value of what is being protected increases to cyber-attackers as the defenses they face improve, fueling the arms race between cyber hackers and cyber defenders (Schwartz, 2014).

This review was not inclusive of all potential cyber threats. We omitted many threats, such as those for businesses that must handle private personnel information.

For example, from the Washington *Post* (Somashekhar, 2014):

> But unlike Settles's other [business] experiments … [with Obamacare] he is still trying to figure some things out – for example, how to safeguard employee information that must now be reported to the Internal Revenue Service, such as the Social Security numbers of children who are covered under their parents' health plans. "We don't want to be liable for that," he said. "What if we get hacked?"

From the *Wall Street Journal* (Thompson, 2014), "The health care info that was hacked (and bank account info) may have affected contractors as well as both former and current employees. Their names, addresses, birth dates, Social Security numbers and dates of service were also included in the mix."

Second, time criticality may be important. Actions can occur at wire speed in cyber, but 'slow and low' attacks like APTs are very difficult to detect and often sit until pre-specified conditions are met (an APT is an advanced persistent threat, like the Stuxnet virus; from Zetter, 2014). A metric to watch is the cost of the defensive decisions per unit of time per unit of defense resource (from Walters, 2014). The implication is that too much cost for cyber defense leads some businesses to settle instead of to defend (i.e., the example we used above where "ransomware" is used by cybercriminals to encrypt a firm's proprietary information and then seeking a fee to decrypt exemplifies the cost of a failed strategy; from Urbina, 2014). Instead of settling, businesses and others must be persuaded that a better strategy is possible with improved defenses (Friedman, 2014).

Third, APTs are becoming a larger threat to national defense. For example, Naji (2004), Zarqawi's Islamist strategist, "proposed a campaign of constant harassment of Muslim states that exhausted the states' will to resist" (Wright, 2014). Harassment is a characteristic of cyber-attacks against businesses such as when the attackers hold computer assets hostage until their ransom demands are met.

Fourth, a list of open cognitive science questions needs to be addressed. For example, we need to know, based on cognitive science, the characteristics of good cyber defenders. We need to know the biases of attackers, users and user groups. And we need to explore the personal, organizational and computational (AI) steps that can be taken to counter biases to better defend users from cyber-attackers.

Fifth, questions exist also for the interdependence in teams, organizations and systems. From interdependence theory, we need to know how to make teams into better cyber defenders; e.g., based on theory, maintaining the boundaries of good teams should generate less entropy—the evidence, supporting our hypothesis, indicates that the best teams generate less noise, but this evidence is anecdotal (Lawless et al., 2013). We have also found that internally cooperative teams compete better in increasingly competitive environments. (Indirectly supporting our conclusion, HHS reported "…that more competition among health plans tends to lower prices …"; from Goldstein, 2014.)

To further develop interdependence theory, we need to better understand the limits of teamwork as cooperation, competition, boundaries, training and technology interact in interdependent environments. We have found that in a competitive world, as teams cooperate to improve their competitiveness, a team's boundaries are strengthened and better maintained (Lawless et al., 2013).

Interdependence theory informs us that boundaries can be maintained by searching for organizational vulnerabilities. Using attacks by "red" teams (Robinson, 2014) to search across cyber defenses for vulnerabilities in "blue" teams helps organizations to defend against cyber-attacks (Schwartz, 2014). We agree with Martinez (2014) that system predictions and assessments are currently weak or nonexistent; system defenses need to be practiced, improved, automated where possible (with AI), and metrics established, measured and reported. Even though we warned that predictions made under interdependent states are clouded by uncertainty, predictions must be made of expected system performance during cyber games, followed by assessments of the metrics for the systems that suffered from red attacks. Comparative analyses of all of the teams playing cyber games need to be assessed and compared against real systems affected by actual cyber-attacks. But, in addition, we want to understand how malicious agents interdependently select targets – not just watch them do it.  We should be able to create a system that predicts a malicious action before a red team composed of humans or autonomous AI agents enact a threat. Based on data sets of past cyber threats and defensive actions, predictive cyber threat analytics that predict future threats should become a part of the AI tool kit used by defenders against malicious actors.

From an individual perspective, cognitive biases form individual vulnerabilities that cyber-attackers attempt to exploit. However, from an interdependent perspective, team training offsets these biases (Lawless et al., 2013). The more competitive is a team, the more able it is to control its biases or limit the extent of their effectiveness (e.g., as with varying levels of cyber defenses; as with checks and balances; in Lawless et al., 2013; or as with the use of "red" teams; in Schwartz, 2014).

Finally, to optimize defenses against cyber-threats, we must shift our focus from an individual to an interdependent perspective. According to methodological individualism, cooperation produces superior social welfare even if punishment is necessary to replace competition with cooperation (Axelrod, 1992, p. 8). But, taken to its logical extreme justifies the savagery used by the Islamic State when it forces its people to be more cooperative (e.g., Naji, 2004). Moreover, this theoretical perspective cannot wish away the threats and risks posed by cyber-attacks. In contrast, the realism of interdependence theory confirms that competition will remain ever present in the struggle for survival, driving the need for disruptive technology. From Kello (2013, p, 31):

> The revolutionary impact of technological change upsets this basic political framework of international society, whether because the transforming technology empowers unrecognized players with subversive motives and aims or because it deprives states of clear "if-then" assumptions necessary to conduct a restrained rivalry.

Competition and disruptive technology combine to create the very real present and future dangers we face in cyberspace; again from Kello (p. 32):

> The cyber domain is a perfect breeding ground for political disorder and strategic instability. Six factors contribute to instrumental instability: offense dominance, attribution difficulties, technological volatility, poor strategic depth, and escalatory ambiguity. Another—the "large-N" problem—carries with it fundamental instability as well.

Staying ahead in the race for new technology is important. In the future, Martinez (2014, p. 8) foresees two things for cyber defenders, that tying speech to visual data needs to be improved (e.g., vocal interactions with recommender displays); and that:

> The development of the recommender system ... is an area of future research applicable to a broad range of applications, including ... cyber anomaly detection … Such an approach will incorporate multiple disciplines in data aggregation, machine learning techniques, augmented cognition models, and probabilistic estimates in reaching the shortest decision time within the courses of action function of a decision support system.

Awareness of the dangers in cyberspace is increasing to Americans along with the need to prepare to face those dangers. Recently in the *Wall Street Journal*, Tom Kean and Lee Hamilton (Kean & Hamilton, 2014), the former chair and vice chair of the 9/11 Commission, respectively, and now co-chairs of the Bipartisan Policy Center's Homeland Security Project, spoke to these dangers:

> A growing chorus of national-security experts describes the cyber realm as the battlefield of the future. American life is becoming ever more dependent on the Internet. At the same time, government and private computer networks in the U.S. are under relentless cyber-attack. This is more than an academic concern—attacks in the digital world can inflict serious damage in the physical world. Hackers can threaten the control systems of critical facilities like dams, water-treatment plants and the power grid. A hacker able to remotely control a dam, pumping station or oil pipeline could unleash large-scale devastation. As terrorist organizations grow and become more sophisticated, the threat of cyber-attack increases as well.

To remain competitive and in business, organizations must be able to defend the proprietary information that they oversee for themselves and their customers in cyberspace (Finch, 2014):

> The real game change for many CIOs [Chief Information Officer] is the emerging movement to consider a company's cybersecurity posture when making procurement decisions. To put it bluntly, companies with weaker cybersecurity are increasingly being viewed as less attractive vendors. … Already companies that have suffered successful cyber-attacks are finding themselves cut off from revenue streams. Just ask USIS, which performs background investigations for the U.S. government. USIS recently suffered a serious data breach, resulting in the personal information of tens of thousands of government employees being compromised. The response from its federal customers, the Department of Homeland Security and the Office of Personnel and Management, was swift: it was issued "stop-work" orders. The "stop-work" means no money coming in from either DHS or OPM. Worse yet, OPM announced earlier this week that it was not renewing its background check contract with USIS.

## Conclusion

In this brief review of our forthcoming chapter on cybersecurity (Lawless et al., 2015), we first agreed that cyber threats are making cyber environments more complex and uncomfortable for average users; second, we concluded that various factors are important (e.g., timely actions are often necessary in cyber space to counter the threats of the attacks that commonly occur at internet speeds, but also the 'slow and low' advanced persistent threats (APTs) attacks that are difficult to detect, threats that occur only after pre-specified conditions have been satisfied that trigger an unsuspecting attack). Third, we concluded that APTs pose a risk to users but also to national security (viz., the persistent threats posed by other Nations). Fourth, we contend that using "red" teams to search cyber defenses for vulnerabilities encourages users and organizations to better defend themselves. Fifth, the current state of theory leaves many questions unanswered that researchers must pursue to mitigate or neutralize present and future threats. Lastly, we agree with the literature that cyber space has had a dramatic impact on American life and that the cyber domain is a breeding ground for disorder. However, we also believe that actions by users and AI researchers can be taken to stay safe and ahead of existing and future threats.

## References

Axelrod, R. (1984). The evolution of cooperation. New York, Basic.

Cieply, M. & Barnes, B. (2014, 12/30), "Sony Cyberattack, First a Nuisance, Swiftly Grew Into a Firestorm", *New York Times*, from http://www.nytimes.com/2014/12/31/business/media/sony-attack-first-a-nuisance-swiftly-grew-into-a-firestorm-.html

DHS (2004, September), Privacy Impact Assessment EINSTEIN Program. Collecting, Analyzing, and Sharing Computer Security Information Across the Federal Civilian Government, Department of Homeland Security, National Cyber Security Division, United States, Computer Emergency Readiness Team (US-CERT). The software for Einstein 1 was designed to examine network traffic; in 2008, Einstein 2 expanded to look at content. (As an example of updates, see: http://www.hsgac.senate.gov/search/?q=Einstein &search-button=Search&access=p&as_dt=i&as_epq=&as_

eq=&as_lq=&as_occt=any&as_oq=&as_q=&as_sitesearch=&client=hsgac&sntsp=0&filter=0&getfields=&lr=&num=15&numgm=3&oe=UTF8&output=xml_no_dtd&partialfields=&proxycustom=&proxyreload=0&proxystylesheet=default_frontend&requiredfields=&sitesearch=&sort=date%3AD%3AS%3Ad1&start=0&ud=1 )

GAO (2014, April), "INFORMATION SECURITY. Agencies Need to Improve Cyber Incident Response Practices", Report GAO-14-354, from http://www.gao.gov/assets/670/662901.pdf

Finch, B.E. (2014, 9/11), CIOs Spur Revenue Generation Through Smart Cybersecurity, the *Wall Street Journal*, http://blogs.wsj.com/cio/2014/09/11/cios-spur-revenue-generation-through-smart-cybersecurity/?KEYWORDS=cyber+threats

FIPSP (1993, 10/5), Federal Information Processing Standards Publication 181, from http://csrc.nist.gov/publications/fips/fips181/fips181.txt.

Friedman, F. (2014, 6/30), "Cyber Specter Mandates New CFO-IT Dynamic;" *Wall Street Journal*, from http://deloitte.wsj.com/riskandcompliance/2014/06/30/cyber-specter-mandates-new-cfo-it-dynamic/?KEYWORDS=cyber+threat

Glowniak J., (1998), "History, structure, and function of the Internet; Semin Nucl Med., 28(2):135-44; from http://www.ncbi.nlm.nih.gov/pubmed/9579415

Goldstein, A. (2014, 6/18), "Federal insurance exchange subsidies cut premiums by average of 76%, HHS reports", *Washington Post*, from http://www.washingtonpost.com/national/health-science/federal-insurance-exchange-subsidies-cut-premiums-by-average-of-76percent-hhs-reports/2014/06/17/4f31b502-f650-11e3-a3a5-42be35962a52_story.html

Intel (2014), Intel cyber-security briefing trends, challenges and leadership opportunities, Cyberstrat 14; https://communities.intel.com/community/itpeernetwork/blog/2014/02/13/intel-cyber-security-briefing-trends-challenges-and-leadership-opportunities--cyberstrat14

Kean, T. & Hamilton, L. (2014, 9/10), "A New Threat Grows Amid Shades of 9/11. The nation remains largely unaware of the potential for disaster from cyberattacks", Wall Street Journal,from http://online.wsj.com/articles/tom-kean-and-lee-hamilton-a-new-threat-grows-amid-shades-of-9-11-1410390195

Kello, L. (2013), "The Meaning of the Cyber Revolution. Perils to Theory and Statecraft", International Security, 38(2): 7-40.

Lawless, W. F., Llinas, James, Mittu, Ranjeev, Sofge, Don, Sibley, Ciara, Coyne, Joseph, & Russell, Stephen (2013). "Robust Intelligence (RI) under uncertainty: Math. and conceptual foundations of autonomous hybrid (human-machine-robot) teams, org.s and systems." Structure & Dynamics 6(2).

Lawless, W.F., Mittu, Ranjeev, Marble, Julie, Coyne, Joseph, Abramson, Myriam, Sibley, Ciara & Gu, Wei (2015, forthcoming), The Human Factor in Cybersecurity: Robust & Intelligent Defense. To be published by Springer.

LII (2014), 44 U.S. Code § 3544 - Federal agency responsibilities, Title 44, Chapter 35, Subchapter III, Legal Information Institute at Cornell University Legal School, from 44 USC §3542; see http://www.law.cornell.edu/uscode/text/44/3544

Loukas, G., Gan, D. & Vuong, T. (2013, 3/22), A taxonomy of cyber attack and defence mechanisms for emergency management, 2013, Third International Workshop on Pervasive Networks for Emergency Management, IEEE, San Diego.

Martinez, D., Lincoln Laboratory, Massachusetts Institute of Technology (2014, invited presentation), Architecture for Machine Learning Techniques to Enable Augmented Cognition in the Context of Decision Support Systems. Invited paper for presentation at HCI.

Naji, A.B. (2004), The management of savagery. http://azelin.files.wordpress.com/2010/08/abu-bakr-naji-the-management-of-savagery-the-most-critical-stage-through-which-the-umma-will-pass.pdf

Raul, A.C. (2015, 1/5), "Cyberdefense Is a Government Responsibility. The Navy fought Barbary pirates to protect U.S. commerce. Digital pirates have much less to fear", *Wall Street Journal*, from http://www.wsj.com/articles/alan-charles-raul-cyberdefense-is-a-government-responsibility-1420502942

Robinson, F. (2014, 4/28), "Europe Begins Its Largest-Ever Cyberwar Stress Test", *Wall Street Journal,* from http://blogs.wsj.com/digits/2014/04/28/europe-begins-its-largest-ever-cyberwar-stress-test/?KEYWORDS=cyber+threat

Rosenquist, M. (2014, 2/8), IT Peer Network; from https://communities.intel.com/community/itpeernetwork/blog/2014/02/08/cyber-security-is-not-prepared-for-the-growth-of-internet-connected-devices

Schwartz, C. (2014, 6/10), "Program overview/challenges"; presentation to the 2014 Computational methods for decision making gathering, Arlington, VA, 10-12 June 2014.

Somashekhar, S. (2014, 6/23), "As health-care law's employer mandate nears, firms cut worker hours, struggle with logistics", *Washington Post,* http://www.washingtonpost.com/national/health-science/as-health-care-lawles-employer-mandate-nears-firms-cut-worker-hours-struggle-with-logistics/2014/06/23/720e197c-f249-11e3-914c-1fbd0614e2d4_story.html

Thompson, J. *Wall Street Journal* (2014, 6/26), "Montana Breach Affects Up To 1.3 Million As Health Care Data Gets Hacked", http://www.wallstreetotc.com/montana-breach-affects-1-3-million-health-care-data-gets-hacked/24807/

Urbina, I. (2014, 6/21), "Hacker Tactic: Holding Data Hostage. Hackers Find New Ways to Breach Computer Security", *New York Times,* from http://www.nytimes.com/2014/06/22/sunday-review/hackers-find-new-ways-to-breach-computer-security.html?_r=0

Walters, J.P. (2014, 6/12), "Heterogeneous cloud services"; presentation to the 2014 Computational methods for decision making gathering, Arlington, VA, 10-12 June 2014.

Wright, L. (2014, 6/17), "ISIS's savage strategy in Iraq", *The New Yorker*, from www.newyorker.com/online/blogs/comment/2014/06/isis-savage-strategy-in-iraq.html

Zetter, K. (2014, 11/3), "An unprecedented look at Stuxnet, the world's first digital weapon", *Wired*, from http://www.wired.com/2014/11/countdown-to-zero-day-stux