NOTE FOR NATIONAL DEFENCE:
Artificial Intelligence: Cybersecurity Challenges

Authors: M. R. Nematollahi [1] and K. Khorasani[2]

[1] Graduate Student, Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada

[2] Professor, Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada

SUMMARY

- Both AI and cybersecurity are capable of complementing and improving each other. Cybersecurity provides higher levels of trustworthiness and resilience for AI systems. At the same time, AI secures an increase in awareness, real-time reaction, and cybersecurity efficacy.

- AI has the growing potential to enable cybersecurity to autonomously detect overt and covert adversarial reconnaissance and attacks at wire speed. For the purpose of integrating AI into computational and cyber-physical systems a fair amount of science, practice, and engineering discipline is required to secure trustworthy use of AI-human systems and environments.

- AI-based reasoning, aligned with cybersecurity priorities, is a promising tool for ensuring the development of automated and trustworthy human-in-the loop systems. For such conditions to occur, experts need to concentrate on the deployment of more reliable software systems and identity management. AI could enhance accuracy within the biometric authentication systems, while, simultaneously, increasing the risk for privacy violations.

- While AI offers several advantageous prospects for enhancing cybersecurity, it can always act as a double-edged sword with certain downsides, the most important of which is the likeliness that it will be used by both attackers and defenders in cyber defensive scenarios. Therefore, both autonomous and semi-autonomous systems need to anticipate and prepare for actual threat occurrences that can affect data owners, service providers, and system operators.

- One of the approaches for better understanding the hackers attack plans and generating potential defenses, is the application of game theory in which the possibility that the 'game' can change quickly is taken into consideration. This theory anticipates a shifting game environment, players with different incentives, or irrational players.

- Since attacking in cyberspace is easier than defending, cyber-attacks are escalating in frequency, impact, and sophistication. AI has the potential to improve the porous defenses and to help reduce cyber-crimes. AI can improve cyber-security through improving system robustness, system resilience, and system responses.

- AI can take software testing to a new level by creating the capability of self-testing and self-healing in software design, which enables them to verify and validate software and to become more robust. It might seem a liberating situation for experts as they will not need to do this tedious job. However, it is imprudent to grant such autonomy to AI as it may increase ethical risks. "Cybersecurity experts need to keep testing systems, for the same reason doctors need to keep reading X-ray scans, so that they still can if AI cannot or gets it wrong".

- Machine learning algorithms are trained and developed by using datasets which ultimately make them vulnerable to unintended bias. Since there is normally a great deal of manual work involved in altering raw data such as age, race, and gender into effective datasets that can be relied on as resources for predictive policing, there is always the possibility and the risk of human social bias that could enter the algorithm training datasets.

- Machine learning algorithms often lack in transparency and they seem to perform as a black box where researchers fail to fully understand how these algorithms come to their arrived conclusions. Consequently, sometimes they do more harm than good as they could cause unprecedented challenges that might have serious unwanted and negative impacts on a person's career or life. Researchers have to try to deal with such challenges ethically and responsibly.

CONTEXT

- AI based tools have emerged to enhance cybersecurity as they are able to quickly identify possible threats. Yet, the intersection of cybersecurity and AI seems to have manifold downsides which reduce their efficacy and need to be addressed by experts and policy makers.

- Due to the current attacker-versus-defender asymmetries, cybersecurity can rely on AI to gain better self-adaptation in the face of current threats, as well as, patterns of identifying the adversary's weaknesses in order to extend levels of protection through developing adaptive and more autonomous responses to attacks.

- AI is looked upon as an asset in securely deploying and operating systems, inspecting for logic errors, monitoring networks, and identifying security vulnerabilities. Moreover, AI could improve identity management and access control by using a history of interactions and expected behavior.

- AI enables attackers to obtain data, execute reconnaissance, and to develop a model of the victim network. Such autonomous analysis and attack must be among the top concerns of experts to categorize attacks and responses through automated isolation which is a form of

behavioral restriction and by defensive agility which uses simulation and updates to strengthen defenses.

- By using data steams or distributed logs of cyber-relevant activity AI can identify adversarial attacks within the early stages of their lifecycle. In this sense, AI acts as a tool for enabling cybersecurity to benefit from predictive analytics and a priori knowledge to identify linkages among datasets that tie together the cyber and human domains.

- Game theory is an approach to model scenarios that imply the necessity of cooperation of multiple AI systems and experts to achieve their goals against an adversary. In this context, multimodal information and probabilistic modeling is incorporated for more effective decision support. However, there remains the possibility of the dual use of game theory as both cyber offense and cyber defense.

- As AI enables better targeted, faster, and more impactful attacks through recognizing and detecting the vulnerabilities of system, it ironically facilitates the escalation process of attacks and the exploitation of potential vulnerabilities. Therefore, as with the pervasive distribution and fast-paced execution of AI systems unforeseen consequences increase and relying on AI to control such errors becomes less effective.

- AI is employed for threat and anomaly detection to improve system resilience. In many cases, AI is able to flag and prioritize threats based on their risk level, as well as to analyze malware and viruses and to quarantine and further investigate them. For this purpose, AI often tracks and monitors human data and their device interactions. It also monitors their behavior and generates biometric profiles. Such extensive monitoring and comprehensive data collection adversely impact human subjects' privacy and increases the risk for the breach of data confidentiality through cyber-attacks. It also creates a mass-surveillance effect which is mostly undesirable.

- Algorithms heavily depend on large data collections in order to progress in their recognition and detection capabilities. However, the advent of large datasets has increased the risk of unintentional bias such as race or gender stereotyping. Therefore, it is required that deliberate work be done by a diverse group of subject experts to reduce such ethical risks.

- In order to develop and incorporate ethical behavior into AI scientists must overcome several technical and social obstacles. Different cultures value a various set of behavioral codes and as a result it is challenging to come up with a unified set of codes that applies to all cultural standards and ethical systems. Therefore, with cultures being fundamentally different and constantly changing over time, crowdsourcing in the context of machine morality faces serious limitations.

CONSIDERATIONS

- Current adversaries employ AI to develop sophisticated attacks which involve multiple stages before the ultimate vulnerable target is compromised. Therefore, cybersecurity needs to similarly use AI for developing strategic approaches that reveal the attacker's goals and current

status by coordinating the available defensive resources in order to intervene and generate a defense plan in the early stages of the attack, before the network is fully compromised.

- It is therefore necessary to develop high standards of validation in the field of AI analysis for cyber-attacks. It is possible through multimodal analysis and cross validation which monitors and tracks false flags and gaps in the data sets to prevent misattribution or collateral damage. Besides, AI analysis can assist human operators to make fewer mistakes and to gain more confidence in outcomes.

- With the growing complexity and severity of cyber-attacks the coordination between human-AI interfaces gains significance. Decision making requires hybrid approaches that leverage human and AI perspectives, so as to reduce the risk of data manipulation, misattribution, and misinformation. However, the trustworthiness of AI and building trust between systems and humans comes into light.

- Humans need to better comprehend where they fit in the decision-making process in order to maximize outcomes and minimize hazards. In order to accommodate humans in the loop it is necessary to slow the AI systems. It allows humans to intervene where necessary and to provide more accountability and safety, yet it reduces the agility of the system. “the right level of trust requires that humans can identify a system’s state and predict its behavior under various circumstances. Over trust could lead to a reluctance to overrule a misbehaving system; under trust could lead to the abandonment of an otherwise effective system. Determining the right level of trust requires human-readable, rule-based specifications based on approximating system behavior, and consideration of cognitive and other biases”.

- It is crucial that more investment be done in research testbeds and datasets to establish proper AI community standards and metrics. In critical domains, such as autonomous vehicles and medical diagnosis, there must be wider and more in-depth evaluation of threat detection mechanisms.

- The creation of realistic simulation environments can is an asset in producing valuable information and more transparent perspectives. In addition, it is crucial to inform and educate both the public and the necessary workforce to better comprehend the advantages and shortcomings of AI. The goal must be to reach a balance between AI’s benefits and challenges in the field of cybersecurity.

- Developing and enforcing regulations and policies are essential to ensure proportionality of responses, legitimate targets, and responsible behavior both in private and public sectors. It is also necessary to establish an authority with the ability to convene international policies and norms regarding responsible state behavior and compliance in cyberspace.

- As Murat Sonmez, the director of the World Economic Forum, suggested in a 2020 interview, there must be an approach to address cybersecurity threats called an ethics switch on an international level, claiming that “we have a concept of an ethic switch where countries define their ethics rules. We download these rules to the smart devices. When they’re asked to do something that’s harmful, the switch says no”.

REFERENCES

- A report by the networking & information technology research and development subcommittee and the machine learning & artificial intelligence subcommittee of the national science & technology council, (2020), “Artificial intelligence and cybersecurity: opportunities and challenges: Technical workshop summary report”.

- Taddeo, Mariarosaria (2019), “Three ethical challenges of applications of artificial intelligence in cybersecurity”, Minds and Machines 29:187-191, published online, Springer Nature.

- Gearheart, Frank (2020), “The ethical use of machine learning in cybersecurity”, 14-ISSA Journal.