The Future of Cyber Systems: Human-AI Reinforcement Learning with Adversarial Robustness

Anonymous Authors[1]

Abstract

Integrating adversarial machine learning (AML) with cyber data representations that support reinforcement learning would unlock human-ai systems with a capacity to dynamically defend against novel attacks, robustly, at machine speed, and with human intelligence.

All machine learning (ML) has an underpinning need for robustness to natural errors and malicious tampering. However, unlike many consumer/commercial models, all ML systems built for cyber will be operating in an inherently adversarial environment with skilled adversaries taking advantage of any flaw.

This paper outlines the research challenges, integration points, and programmatic importance of such a system, while highlighting the social and scientific benefits of pursuing this ambitious program.

1. Problem Definition

Large language models (LLM's) have captured the public's imagination for their wide-reaching impact on society. Other model types, including RL, have similar potential to change how global communities interact. Dramatic advances have occurred in playing complex strategy games with autonomous RL agents. Cyber security is a real world extension of such strategy games.

All ML has an underpinning need for robustness to natural errors and malicious tampering. However, entertainment products that drive mass-market revenue of for-profit companies are not safety critical. Consequently, there has been little investment in the research necessary to address malicious manipulation of ML in the real world. Cyber security is possibly the only domain where you must assume an adversary is working against you and will use any means available. Fundamental research in ML robustness is dominated by demonstrations in the image domain, however, the security and privacy community is leading the integration of these techniques to deployed real-world tools because other commercial applications have insufficient financial incentive. An ambitious program to apply ML to the cyber domain will co-advance ML robustness in a way that can be re-applied to other security domains where ML is essential, but under-developed.

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Under review by *AdvML-Frontiers'23: The 2nd Workshop on New Frontiers in Adversarial Machine Learning@ICML'23*, Honolulu, Hawaii, USA. Do not distribute.

2. The Case for Adversarial ML in Cyber

Cyber warfare is the invisible front line in global security. Economies depend on utilities, banks, and healthcare, which cyber-attacks disrupt as effectively as traditional weapons. ML and automation are needed to defend human-initiated attacks at scale. That need is compounded when attackers use automation and ML themselves, to launch increasingly sophisticated, stealthy, and large-scale attacks.

Short-term gains are made from providing ML tools to the human defender, such as automated alert triage. However, the human cognitive capacity needs to be augmented and scaled with fully autonomous agents, observing and defending real-time networks in tandem with humans. This is an ambitious, risky, and multi-disciplinary research agenda.

3. Critical Research Elements

Many research elements need to be co-developed in order to address the current gaps preventing RL autonomous cyber agents from operating robustly in the real world.

Establish a framework integrating the data types needed to train an autonomous RL cyber agent and evaluate its performance in dynamic and realistic data. This environment should be standardized for agent comparison. The framework must have sufficient detail to ultimately integrate with real-world systems, but simplified representation to allow RL agent training. Additionally, the environment needs to be human playable and observable to provide meaningful feedback and correction to the autonomous agent. (Molina-Markham et al., 2021; Piplai et al., 2022; Applebaum et al.,

2022)

Creating RL algorithms which can perform prediction tasks on large dynamic graphs, is state-of-the-art research. These bounds will need to be pushed further to perform tasks and in a timely manner on real-world, enterprise-scale cyber networks. The indicators of malicious behavior are often subtle, and sparse temporal events. Detection and correlation of these subtle events over days, weeks and months of voluminous data make the cyber security domain particularly challenging. (Ren et al., 2022; You et al., 2022) Creating a standardized playable environment for cyber security is needed to enable the domain performance.

Establish and extend robustness measures needed for human-computer teaming, in training and deployment of RL graph-based systems (Wirth et al., 2017). RL systems are vulnerable to a variety of adversarial attacks. (Chen et al., 2019). Machine unlearning techniques, used to remove adversarial effects, needs further study to extend to RL models and shifting data distributions that occur in streaming cyber data. (Nguyen et al., 2022) A variety of quality measures specific to RL have been developed, (Chan et al., 2019), but extensions to reflect adversarial attack success in RL will need additional research.

These research challenges are further tied together by human-computer interactions in all phases of ML/RL development and deployment. It is likely the RL agent will need to be trained from direct observation of actions taken by human defenders in an identical game play. Providing meaningful feedback interactions at the speed and scale the RL agent can operate will be a unique research challenge. The RL agents must also demonstrate robust performance against human and computer opponents that may use adversarial machine learning techniques.

4. Outcomes

Successful outcomes will require technical exchange between diverse experts, to align the solutions developed in each track. The most important enabler of this research will be the creation of the autonomous cyber agent environment. Without an environment, no RL experiments can be conducted and the realism of the environment will directly determine the ability to apply the methods in real world scenarios. Environment creation will require equally advanced, cutting-edge knowledge in both cyber and RL domains, to reflect the attack kill chain, relevant network, endpoint, and alert data signals used by human analysts. A standardized environment will also enable uniform comparison and evaluation of the robustness of agents developed by independent research groups.

System security is limited by the robustness of the weakest component. Currently, many systems are under-defended because of a shortage of technical talent. If autonomous RL cyber agents capable of work in tandem with humans becomes widely available, it will revolutionize the defense of critical infrastructure.

References

Applebaum, A., Dennler, C., Dwyer, P., Moskowitz, M., Nguyen, H., Nichols, N., Park, N., Rachwalski, P., Rau, F., Webster, A., et al. Bridging automated to autonomous cyber defense: Foundational analysis of tabular q-learning. In *Proceedings of the 15th ACM Workshop on Artificial Intelligence and Security*, pp. 149–159, 2022.

Chan, S. C., Fishman, S., Canny, J., Korattikara, A., and Guadarrama, S. Measuring the reliability of reinforcement learning algorithms. *arXiv preprint arXiv:1912.05663*, 2019.

Chen, T., Liu, J., Xiang, Y., Niu, W., Tong, E., and Han, Z. Adversarial attack and defense in reinforcement learning-from ai security view. *Cybersecurity*, 2:1–22, 2019.

Molina-Markham, A., Miniter, C., Powell, B., and Ridley, A. Network environment design for autonomous cyberdefense. *arXiv preprint arXiv:2103.07583*, 2021.

Nguyen, T. T., Huynh, T. T., Nguyen, P. L., Liew, A. W.-C., Yin, H., and Nguyen, Q. V. H. A survey of machine unlearning. *arXiv preprint arXiv:2209.02299*, 2022.

Piplai, A., Anoruo, M., Fasaye, K., Joshi, A., Finin, T., Ridley, A., et al. Knowledge guided two-player reinforcement learning for cyber attacks and defenses. In *International Conference on Machine Learning and Applications*, 2022.

Ren, H., Dai, H., Dai, B., Chen, X., Zhou, D., Leskovec, J., and Schuurmans, D. Smore: Knowledge graph completion and multi-hop reasoning in massive knowledge graphs. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1472–1482, 2022.

Wirth, C., Akrour, R., Neumann, G., Fürnkranz, J., et al. A survey of preference-based reinforcement learning methods. *Journal of Machine Learning Research*, 18(136): 1–46, 2017.

You, J., Du, T., and Leskovec, J. Roland: graph learning framework for dynamic graphs. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 2358–2366, 2022.

A. You *can* have an appendix here.

You can have as much text here as you want. The main body must be at most 6 pages long. For the final version, one more page can be added. If you want, you can use an appendix like this one, even using the one-column format.