AI vs. AI: Exploring the Intersections of AI and Cybersecurity

Ian Molloy
IBM T. J. Watson Research Center
Yorktown Heights, New York, USA
molloyim@us.ibm.com

Josyula R. Rao
IBM T. J. Watson Research Center
Yorktown Heights, New York, USA
jrrao@us.ibm.com

Marc Ph. Stoecklin
IBM Zurich Research Lab
CH–8803 Rüschlikon, Switzerland
mtc@zurich.ibm.com

ABSTRACT

The future of cybersecurity will pit AI against AI.

In this talk, we explore the role of AI in strengthening security defenses as well as the role of security in protecting AI services. We expect that the scale, scope and frequency of cyber attacks will increase disruptively with attackers harnessing AI to develop attacks that are even more targeted, sophisticated and evasive. At the same time, analysts in security operations centers are being increasingly overwhelmed in their efforts to keep up with the tasks of detecting, managing and responding to attacks. To cope, the security industry and practitioners are experimenting with the application of AI and machine learning technologies in different areas of security operations. These include a diverse set of areas such as detecting (mis)behaviors and malware, extracting and consolidating threat intelligence, reasoning over security alerts, and recommending countermeasures and/or protective measures.

At the same time, adversarial attacks on machine learning systems have become an indisputable threat. Attackers can compromise the training of machine learning models by injecting malicious data into the training set (so-called poisoning attacks), or by crafting adversarial samples that exploit the blind spots of machine learning models at test time (so-called evasion attacks). Adversarial attacks have been demonstrated in a number of different application domains, including malware detection, spam filtering, visual recognition, speech-to-text conversion, and natural language understanding. Devising comprehensive defenses against poisoning and evasion attacks by adaptive adversaries is still an open challenge. Thus, gaining a better understanding of the threat by adversarial attacks and developing more effective defense systems and methods are paramount for the adoption of machine learning systems in security-critical real-world applications.

The talk will provide an industrial research perspective and will cover research conducted at IBM Security Research over several years.

CCS CONCEPTS

• **Security and privacy → Intrusion/anomaly detection and malware mitigation**.

KEYWORDS

Cybersecurity, Artificial Intelligence, Big Data Analytics, Threat Intelligence, Natural Language Processing, Adversarial AI, Poisoning Attacks, Evasion Attacks

ACM Reference Format:
Ian Molloy, Josyula R. Rao, and Marc Ph. Stoecklin. 2021. AI vs. AI: Exploring the Intersections of AI and Cybersecurity. In *Proceedings of the 2021 ACM International Workshop on Security and Privacy Analytics (IWSPA'21), April 28, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/3445970.3456286

IWSPA'21, April 28, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8320-2/21/04.
https://doi.org/10.1145/3445970.3456286