Fortifying Your Defenses: Techniques to Thwart Adversarial Attacks and Boost Performance of Machine Learning-Based Intrusion Detection Systems

Wenjing Lou
Virginia Tech
Arlington, Virginia, USA
wjlou@vt.edu

ABSTRACT

Machine learning has seen significant advancements in recent years and has proven to be highly effective in a wide range of applications, including intrusion detection systems (IDS). However, while working in adversarial environments, machine learning-based systems are known to be vulnerable to a range of attacks. In this talk, we will discuss techniques aimed at strengthening machine learning-based IDS. On the one hand, we explore techniques for enhancing the performance and robustness of IDS in adversarial environments, where we propose a contrastive learning-based approach that builds highly differentiating IDS. On the other hand, we develop efficient security mechanisms to thwart common attacks, including an adversarial example (AE) detector that filters out suspicious inputs at the model testing time, and a robust model evaluation method that leverages latent space representations to build resiliency in model aggregation against model poisoning attacks in federated learning. This talk will report our research results along this line of research.

CCS CONCEPTS

• **Security and privacy → Network security**; **Intrusion detection systems**.

KEYWORDS

Machine learning; federated learning; contrastive learning; intrusion detection; adversarial examples; model poisoning attacks.

ACM Reference Format:
Wenjing Lou. 2023. Fortifying Your Defenses: Techniques to Thwart Adversarial Attacks and Boost Performance of Machine Learning-Based Intrusion Detection Systems. In *Proceedings of the 2023 ACM Workshop on Wireless Security and Machine Learning (WiseML '23), June 1, 2023, Guildford, United Kingdom.* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/3586209.3591392

BIOGRAPHY

Wenjing Lou is the W. C. English Endowed Professor of Computer Science at Virginia Tech and a Fellow of the IEEE. She holds a Ph.D. in Electrical and Computer Engineering from the University of Florida. Her research interests cover many topics in the cybersecurity field, with her current research interest focusing on wireless networks, blockchain systems, trustworthy machine learning systems, and security and privacy problems in the Internet of Things (IoT) systems. Prof. Lou is a highly cited researcher by the Web of Science Group. She received the Virginia Tech Alumni Award for Research Excellence in 2018, the highest university-level faculty research award. She received the INFOCOM Test-of-Time paper award in 2020. She is the TPC chair for IEEE INFOCOM 2019 and ACM WiSec 2020. She was the Steering Committee Chair for IEEE CNS conference from 2013 to 2020. She is currently a steering committee member of IEEE INFOCOM and IEEE CNS. She served as a program director at US National Science Foundation (NSF) from 2014 to 2017.

WiseML '23, June 1, 2023, Guildford, United Kingdom

ACM ISBN 979-8-4007-0133-7/23/06.
https://doi.org/10.1145/3586209.3591392

ACKNOWLEDGEMENTS

The work presented in this talk was supported by the Office of Naval Research under grant N00014-19-1-2621, and the US National Science Foundation under grants 1837519 and 1916902.

REFERENCES

[1] Ning Wang, Yimin Chen, Yang Xiao, Yang Hu, Wenjing Lou, and Y. Thomas Hou. 2023. MANDA: On Adversarial Example Detection for Network Intrusion Detection System. IEEE Trans. Dependable Secur. Comput. 20, 2 (March-April 2023), 1139–1153. https://doi.org/10.1109/TDSC.2022.3148990
[2] Ning Wang, Yimin Chen, Yang Hu, Wenjing Lou, and Y. Thomas Hou. 2022. FeCo: Boosting Intrusion Detection Capability in IoT Networks via Contrastive Learning. In IEEE INFOCOM 2022 - IEEE Conference on Computer Communications. IEEE Press, 1409–1418. https://doi.org/10.1109/INFOCOM48880.2022.9796926
[3] Ning Wang, Yang Xiao, Yimin Chen, Yang Hu, Wenjing Lou, and Y. Thomas Hou. 2022. FLARE: Defending Federated Learning against Model Poisoning Attacks via Latent Space Representations. In Proceedings of the 2022 ACM on Asia Conference on Computer and Communications Security (ASIA CCS '22). Association for Computing Machinery, New York, NY, USA, 946–958. https://doi.org/10.1145/3488932.3517395