

Feng Wei

☎ (+1)213-453-2214 ✉ fengwei@buffalo.edu 🔗 <https://akafengfeng.github.io/>

CAREER OBJECTIVES

To enhance the robustness and trustworthiness of cybersecurity measures in an era of rapidly advancing technology, my career objective is centered on pioneering research in explainable and trustworthy AI. With a specific focus on implementing deep learning techniques in network intrusion detection, I am committed to innovating within this sphere, aiming to significantly contribute to both the academic and practical advancements in cybersecurity.

RESEARCH INTERESTS

- Explainable and Trustworthy AI
- Machine Learning for Cybersecurity
- Network Intrusion Detection Systems

EDUCATION

Ph.D. in Computer Science and Engineering	University at Buffalo(June 2021 - Expected May 2024)
• Advisor: Prof. Hongxin Hu	GPA: 4.0/4.0
M.S. and Ph.D. student in Computer Science	Clemson University(Aug. 2016 - May 2021)
• Advisor: Prof. Hongxin Hu	Transferred
M.E. student in Cybersecurity	University of Science and Technology of China, 2016
B.E. in Automation	Xi'an Jiaotong University, 2014

WORK EXPERIENCE

AI and Security Research Intern	May 2023 - Aug. 2023
Mitsubishi Electric Research Laboratories, Inc. (MERL)	Cambridge MA USA
• Mentor: Dr. Ye Wang, Dr. Toshiaki Koike-Akino, and Dr. Jing Liu	
• Collaborate with MERL researchers on developing robust AI for cybersecurity technology.	

RESEARCH TOPICS

Topic 1: Explainable Deep Learning-based Network Intrusion Detection System (DL-NIDS)

- Designed a novel explanation method specifically for DL-NIDS that outperforms existing baselines (LIME, SHAP, LRP, and Integrated Gradients) in terms of fidelity, sparsity, completeness, and stability. This was achieved by approximating and sampling the historical inputs and capturing feature dependencies using the sparse group lasso.
- Developed a defense rule generation methodology that enables active intrusion responses. This methodology considers defense rule scopes and security constraints to ensure accuracy, and a unified rule representation to make defense rules compatible with different defense tools.
- The project received the **Amazon Research Award (ARA)** and was published in **USENIX Security 2023**.

Topic 2: Robust Deep Learning-based NIDS with Adversarial Training and Data Augmentation

- Designed a data augmentation approach using generative adversarial networks (GANs) to counteract the distribution shifts that naturally occur in DL-NIDS. This reduced the false positive rate of DL-NIDS when tested with new data.
- Proposed a robust training strategy for DL-NIDS to withstand synthetic distribution shifts, including advanced adversarial attacks and evading attacks. This improved the efficiency of DL-NIDS training and increased its detection effectiveness.

Topic 3: Explainable and Robust Transformer-based NIDS

- Designed a transformer-based NIDS that can provide the detection scores of the anomalies as well as the high-level explanations of the malicious events.
- The proposed model is self-explainable and robust against adversarial attacks. Compared to the traditional deep-learning method, our method can detect and explain the attacks at the same time.

PUBLICATIONS

- **Feng Wei**, Hongda Li, Ziming Zhao, and Hongxin Hu. "xNIDS: Explaining Deep Learning-based Network Intrusion Detection Systems for Active Intrusion Responses". In Proceedings of the 32nd USENIX Security Symposium (**USENIX Security**) , Anaheim, CA, USA, August 9-11, 2023
- **Feng Wei***, Hongda Li*, and Hongxin Hu. "Enabling Dynamic Network Access Control with Anomaly-based IDS and SDN". In Proceedings of ACM International Workshop on Security in Software Defined Networks and Network Function Virtualization (SDN-NFV Security 2019), Richardson, Texas, USA, March 27, 2019. (*co-first author)

SELECTED WORKING PAPERS

- **Feng Wei**, Hongxin Hu. “rNIDS: On Training Robust Deep Learning-based Network Intrusion Detection Systems” Target: Proceedings of the ACM SIGSAC Conference on Computer and Communications Security 2024, Status: Final Iterations.
- **Feng Wei**, Hongxin Hu. “tNIDS: Transformer-based NIDS for Active Intrusion Responses” Target: Proceedings of the ACM SIGSAC Conference on Computer and Communications Security 2024, Status: Draft
- **Feng Wei**, Jing Liu, Toshiaki Koike-Akino, and Ye Wang. “On Evaluating and Improving Deep Learning based Log Anomaly Detection Systems” Target: Proceedings of the International Federation for Information Processing (IFIP) Security 2024, Status: In Submission.

INVITED TALKS AND POSTERS

- **Talk:**“xNIDS: Explaining Learning-based Network Intrusion Detection Systems for Active Intrusion Responses”. **VMware Talk**, June 10th 2021.
- **Talk:**“Interpreting leaning based network intrusion detection system for active intrusion response”. **Great Lakes Security Day**, November 12th 2021.
- **Poster:**“Explaining Learning-based Network Intrusion Detection Systems for Active Intrusion Responses”. NSF/VMware Partnership on SDI as a Foundation for Clean-slate Computing Security (SDI-CSCS) Final Annual PI Meeting 2021.
- **Poster:**“Dynamic Defense with Explainable Network Intrusion Detection Systems”. NSF/VMware Partnership on SDI as a Foundation for Clean-slate Computing Security (SDI-CSCS), Annual PI Meeting 2020.
- **Poster:**“Enabling Dynamic Network Access Control with Anomaly-based NIDS and SDN”. NSF/VMware Partnership on SDI as a Foundation for Clean-slate Computing Security (SDI-CSCS), Annual PI Meeting 2019.
- **Poster:**“Explainable Network Intrusion Detection Systems with Deep Learning”. AI for Industry Conference, CUiCAR, Greenville SC, 2018.

AWARDS AND COMPETITIONS

- USENIX Security Student Grant 2023
- **Amazon Research Award** (ARA AI for Information Security) **Award Amount:\$100,000** 2022
- DEFCON AutoDriving Capture the Flag(CTF) Competition **5th and 13th**, 2021 and 2022
- DJI RoboMasters Robotics Competition **Championship** 2015
- Freescale (NXP) Cup Smart Car Competition **Top 3/2,000 teams** 2013
- The Mathematical Contest in Modeling (MCM) **Meritorious Winner** 2013
- Contemporary Undergraduate Mathematical Contest in Modeling (CUMCM) **1st Prize** 2011

TECHNICAL PROGRAM COMMITTEE

- International Workshop on Cyberspace Security and Artificial Intelligence 2023
- ACM CCS Artifact Evaluation Committee 2023
- USENIX Security Symposium (USENIX Security) Artifact Evaluation Committee 2022, 2023, 2024
- Annual Computer Security Applications Conference (ACSAC) Artifacts Evaluation Committee 2021, 2022, 2023
- Poster Program of ACM Conference on Data and Application Security and Privacy (CODASPY) 2020, 2022

CONFERENCE PAPER (SUB) REVIEWER

- International Conference on Machine Learning (ICML) 2023
- ACM Conference on Computer and Communications Security (CCS) 2023
- Conference on Neural Information Processing Systems (NeurIPS) 2022
- AAAI Conference On Artificial Intelligence (AAAI) 2021
- The Web Conference (WWW) 2019-2022
- ACM Symposium on Information, Computer and Communications Security (AsiaCCS) 2019-2022
- Annual Computer Security Applications Conference (ACSAC) 2019-2022
- ACM Conference on Data and Application Security and Privacy (CODASPY) 2019-2022

JOURNAL PAPER REVIEWER

- IEEE Transactions on Machine Learning in Communications and Networking
- IEEE Transactions on Dependable and Secure Computing
- IEEE Transactions on Cloud Computing
- Information Systems Frontiers
- Computers & Security

TEACHING EXPERIENCE

- **Guest Lecture:** CSE 702 Machine Learning and Cybersecurity University at Buffalo, Spring 2023
- **Teaching Assistant:** CPSC 8430 Deep Learning Clemson University, Spring 2020

SKILL

Programming languages: Python, C/C++, JavaScript, Latex

Tools and Frameworks: Tensorflow, Keras, Pytorch, Pandas, Git, Vim, VScode, Wireshark, Tshark

Last Updated on February 3, 2024