Feng Wei

(+1)213-453-2214 fengwei@buffalo.edu https://akafengfeng.github.io/

Career Objectives

Ph.D. in Computer Science and Engineering focusing on trustworthy machine learning and ML for computer security. My recent research includes designing generative models and foundation models for network security, as well as improving the explainability and robustness of LLMs.

Research Interests

- Machine Learning for Cybersecurity
- GenAI and Foundation Models (LLM)
- Explaninbale and Trustworthy AI
- Network Intrusion Detection Systems

Education

Ph.D. in Computer Science and Engineering — **University at Buffalo(June 2021 - Expected Spring. 2025)**
- **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 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 the scope of defense rules and security constraints to ensure precision and a unified rule representation to make defense rules compatible with different defense tools.
- The project received the **Amazon Research Award (ARA)** .

Topic 2: Robust Deep Learning-based NIDS with Data Augmentation and Adversarial Training
- 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: Security Foundation Models for NIDS
- Design and train a security foundation model for NIDS that can provide the detection scores of the anomalies as well as the high-level explanations of the malicious events.
- Integrated with LLM, the proposed method is explainable and robust against advanced attack. Compared to the traditional deep learning method, our method can detect and mitigate attacks simultaneously.

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 USENIX Security Symposium 2025, Status: In submission.
- **Feng Wei**, Hongxin Hu. "tNIDS: Transformer-Based NIDS for Active Intrusion Responses" Target: Proceedings of the USENIX Security Symposium 2025, Status: Draft
- **Feng Wei**, Jing Liu, Toshiaki Koike-Akino, and Ye Wang. "On Evaluating and Improving Deep Learning-Based Log Anomaly Detection Systems" Target: 10th ACM International Workshop on Security and Privacy Analytics(IWSPA 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 Learning-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

- USENIX Security Symposium (USENIX Security) Artifact Evaluation Committee 2022, 2023, 2024
- ACM CCS Artifact Evaluation Committee 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

- ACM Conference on Computer and Communications Security (CCS) 2023-2024
- The Network and Distributed System Security Symposium (NDSS) 2024-2025
- International Conference on Machine Learning (ICML) 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) 2019-2024

Journal Paper Reviewer

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

Teaching Experience

- **Guest Lecture**: CSE 565 Computer Security — University at Buffalo, Spring 2024
- **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 August 2, 2024