Ren Pang

(484)747-2401 | ain-soph@live.com | ain-soph.github.io

RESEARCH INTEREST

My research focuses on understanding and tackling the challenges arising in the advances of deep learning and artificial intelligence in general, especially about security risks in image classification task.

EDUCATION

Ph.D.,	Information Sciences and Technology,	Pennsylvania State University	2019–2023
BSc.,	Mathematics,	Nankai University	2014–2018

EXPERIENCE

Software Engineer (Intern), Meta 2022 Summer

The work focuses on in-stream ad breaks demonetization during creator onboarding, which detects malicious creators/pages that violate Facebook regulations. During the development, I fix several existing bugs which introduce good but incorrect metric results, extend labels from binary violation to concrete violated policies. SQL and FBlearner are used in the work.

PUBLICATION

1. A Tale of Evil Twins: Adversarial Inputs versus Poisoned Models,
   R. Pang, H. Shen, X. Zhang, S. Ji, Y. Vorobeychik, X. Luo, A. Liu, and T. Wang,
   Proceedings of *the ACM Conference on Computer and Communications Security* (**CCS**), 2020.
2. AdvMind: Inferring Adversary Intent of Black-Box Attacks,
   R. Pang, X. Zhang, S. Ji, X. Luo, and T. Wang,
   Proceedings of *the ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (**KDD**), 2020.
3. i-Algebra: Towards Interactive Interpretability of Deep Neural Networks,
   X. Zhang, **R. Pang**, S. Ji, F. Ma, and T. Wang,
   Proceedings of *the AAAI Conference on Artificial Intelligence* (**AAAI**), 2021.
4. Graph Backdoor,
   Z. Xi, **R. Pang**, S. Ji, and T. Wang,
   Proceedings of *the USENIX Security Symposium* (**USENIX**), 2021.
5. On the Security Risks of AutoML,
   R. Pang, Z. Xi, S. Ji, X. Luo, and T. Wang,
   Proceedings of *the USENIX Security Symposium* (**USENIX**), 2022.
6. TrojanZoo: Towards Unified, Holistic, and Practical Evaluation of Neural Backdoors,
   R. Pang, Z. Zhang, X. Gao, Z. Xi, S. Ji, P. Cheng, and T. Wang,
   Proceedings of *the IEEE European Symposium on Security and Privacy* (**EuroS&P**), 2022.
7. The Dark Side of AutoML: Towards Architectural Backdoor Search,
   R. Pang, C. Li, Z. Xi, S. Ji, T. Wang,
   Arxiv Preprint, 2022.
8. Demystifying Self-supervised Trojan Attacks,
   C. Li, **R. Pang**, Z. Xi, T. Du, S. Ji, Y. Yao, T. Wang,
   Arxiv Preprint, 2022.
9. Reasoning over Multi-view Knowledge Graphs,
   Z. Xi, **R. Pang**, C. Li, T. Du, S. Ji, F. Ma, T. Wang,
   Arxiv Preprint, 2022.

OPEN-SOURCED ARTIFACT

1. TrojanZoo
   `https://ain-soph.github.io/trojanzoo`
   A universal, flexible PyTorch platform to conduct security analysis of attacks and defenses (e.g., adversarial evasion, backdoor injection, model poisoning) on deep neural network models.

2. TrojanZoo Sphinx Theme
   `https://ain-soph.github.io/trojanzoo_sphinx_theme`
   A light-weight, customizable theme that generalizes pytorch_sphinx_theme.

3. AlpsPlot
   `https://ain-soph.github.io/alpsplot`
   A customizable Python plotting library based on matplotlib.