Minhao Cheng

CONTACT INFORMATION

College of Information Sciences & Technology
Pennsylvania State University

Tel: (814) 925 4997
E-mail: mmc7149@psu.edu

EDUCATION

University of California Los Angeles, Los Angeles, USA
Ph.D. in Computer Science, March 2021
Advisor: Cho-Jui Hsieh

University of Electronic Science and Technology of China, Chengdu, China
B.Eng. in Computer Science and Technology, July, 2015

RESEARCH INTERESTS

Machine learning, Trustworthy Machine learning, Deep Learning, Optimization, AutoML.

WORK EXPERIENCE

Assistant Professor, College of Information Sciences & Technology
Pennsylvania State University Jan. 2024 - Now

Assistant Professor, Department of Computer Science & Engineering
Hong Kong University of Science and Technology Jan. 2022 - Dec. 2023

HONORS

Outstanding Paper Award, ICLR, 2021

PUBLICATION

Google Scholar Profile: Number of Citations=2600+; h-index = 17, i10-index = 24. Details available at `https://scholar.google.com/citations?user=_LkC1yoAAAAJ&hl=en`

1. Dandan Ni, Sheng Zhang, Cong Deng, Han Liu, Gang Chen, **Minhao Cheng**, Hongyang Chen. Exploring Robustness of GNN against Universal Injection Attack From a Worst-case Perspective, In 33rd ACM International Conference on Information and Knowledge Management (CIKM), 2024.
2. Rui Min, Sen Li, Hongyang Chen, **Minhao Cheng**. A Watermark-Conditioned Diffusion Model for IP Protection, In European Conference on Computer Vision (ECCV), 2024.
3. Yuanhao Ban, Ruochen Wang, Tianyi Zhou, **Minhao Cheng**, Boqing Gong, Cho-Jui Hsieh. When and How do negative prompts take effect?, In European Conference on Computer Vision (ECCV), 2024.
4. Haosen Wang, Can Xu, Chenglong Shi, PengFei Zheng, Shiming Zhang, **Minhao Cheng**, Hongyang Chen. Unsupervised Heterogeneous Graph Rewriting Attack via Node Clustering, In ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2024.
5. Ruochen Wang*, Sohyun An*, **Minhao Cheng**, Tianyi Zhou, Sung Ju Hwang, Cho-jui Hsieh. One Prompt is not Enough: Automated Construction of a Mixture-of-Expert Prompts, In International Conference on Machine Learning (ICML), 2024.

6. Kuan Li, YiWen Chen, Yang Liu, Jin Wang, Qing He, **Minhao Cheng**, Xiang Ao. Boosting the Adversarial Robustness of Graph Neural Networks: An OOD Perspective, In International Conference on Learning Representations (ICLR), 2024.
7. Jianqiu Wu, Hongyang Chen, **Minhao Cheng**, Haoyi Xiong. CurvAGN: Curvature-based Adaptive Graph Neural Networks for Predicting Protein-Ligand Binding Affinity, In BMC Bioinformatics 24.
8. Lichang Chen, Heng Huang, **Minhao Cheng**. PTP: Boosting Stability and Performance of Prompt Tuning with Perturbation-Based Regularizer, In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023.
9. Rui Min*, Zeyu Qin*, Li Shen, **Minhao Cheng**. Stable Backdoor Purification with Feature Shift Tuning, In Advances in Neural Information Processing Systems (NeurIPS), 2023.
10. Zeyu Qin, Liuyi Yao, Daoyuan Chen, Yaliang Li, Boling Ding, **Minhao Cheng**. Revisiting Personalized Federated Learning: Robustness Against Backdoor Attacks, In ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2023.
11. **Minhao Cheng**, Rui Min, Haochen Sun, Pin-Yu Chen. Identification of the Adversary from a Single Adversarial Example, In International Conference on Machine Learning (ICML), 2023.
12. Bo Huang, Mingyang Chen, Yi Wang, Junda Lu, **Minhao Cheng**, Wei Wang. Boosting Accuracy and Robustness of Student Models via Adaptive Adversarial Distillation, In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023.
13. Yuanhao Xiong*, Ruochen Wang*, **Minhao Cheng**, Felix Yu, Cho-Jui Hsieh. Communication-Efficient Federated Learning via Dataset Distillation, In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023. (* Equal Contribution)
14. Joseph Lavond, **Minhao Cheng**, Yao Li. Trusted Aggregation (TAG): Model Filtering Backdoor Defense In Federated Learning, In NeurIPS 2022 Workshop on Federated Learning: Recent Advances and New Challenges.
15. Xingling Li, Yao Li, **Minhao Cheng**. Defend Against Textual Backdoor Attacks By Token Substitution, In NeurIPS Workshop on Robustness in Sequence Modeling, 2022.
16. Yong Liu, Siqi Mai, **Minhao Cheng**, Xiangning Chen, Cho-Jui Hsieh, Yang You. Random Sharpness-Aware Minimization. In Advances in Neural Information Processing Systems (NeurIPS), 2022.
17. Ruochen Wang, Yuanhao Xiong, **Minhao Cheng**, Cho-Jui Hsieh. Efficient Non-Parametric Optimizer Search for Diverse Tasks. In Advances in Neural Information Processing Systems (NeurIPS), 2022.
18. **Minhao Cheng**, Qi Lei, Pin-Yu Chen, Inderjit Dhillon, Cho-Jui Hsieh. CAT: Customized Adversarial Training for Improved Robustness. In International Joint Conference on Artificial Intelligence (IJCAI), 2022.
19. Yong Liu, Xiangning Chen, **Minhao Cheng**, Cho-Jui Hsieh, Yang You. Concurrent Adversarial Learning for Large-Batch Training, In International Conference on Learning Representations (ICLR), 2022.
20. Chenxi Liu, Zhu Xiao, Dong Wang, **Minhao Cheng**, Hongyang Chen, Jiawei Cai. Foreseeing private car transfer between urban regions with multiple graph-based generative adversarial networks. The World Wide Web Journal, 2022.
21. Yao Li, **Minhao Cheng**, Cho-Jui Hsieh, Thomas Lee. A Review of Adversarial Attack and Defense for Classification Methods. In The American Statistician, 2021
22. Ruochen Wang, Xiangning Chen, **Minhao Cheng**, Xiaocheng Tang, Cho-Jui Hsieh. RANKNOSH: Efficient Predictor-Based NAS via Non-Uniform Successive Halving. In International Conference on Computer Vision (ICCV), 2021

23. **Minhao Cheng**. On the Robustness of Neural Network: Attacks and Defenses. PhD Dissertation.
24. Ruochen Wang, **Minhao Cheng**, Xiangning Chen, Xiaocheng Tang, Cho-Jui Hsieh. Rethinking Architecture Selection in Differentiable NAS. In International Conference on Learning Representations (ICLR), 2021. (Outstanding Paper Award)
25. Xiangning Chen*, Ruochen Wang*, **Minhao Cheng***, Xiaocheng Tang, Cho-Jui Hsieh. DrNAS: Dirichlet Neural Architecture Search. In International Conference on Learning Representations (ICLR), 2021. (* Equal Contribution)
26. **Minhao Cheng**, Pin-Yu Chen, Sijia Liu, Shiyu Chang, Cho-Jui Hsieh, Payel Das. Self-Progressing Robust Training. In AAAI Conference on Artificial Intelligence (AAAI), 2021.
27. Xiaoqing Zheng, Jiehang Zeng, Yi Zhou, Cho-Jui Hsieh, **Minhao Cheng**, Xuanjing Huang. Evaluating and enhancing the robustness of neural network-based dependency parsing models with adversarial examples. In Proceedings of Association for Computational Linguistics (ACL), 2020.
28. **Minhao Cheng***, Simranjit Singh*, Patrick H. Chen, Pin-Yu Chen, Sijia Liu, Cho-Jui Hsieh. Sign-OPT: A Query-Efficient Hard-label Adversarial Attack. In International Conference on Learning Representations (ICLR), 2020. (* Equal Contribution)
29. **Minhao Cheng**, Jinfeng Yi, Huan Zhang, Pin-Yu Chen, Cho-Jui Hsieh. Seq2Sick: Evaluating the Robustness of Sequence-to-Sequence Models with Adversarial Examples. In AAAI Conference on Artificial Intelligence (AAAI), 2020.
30. Yu-Lun Hsieh, **Minhao Cheng**, Da-Cheng Juan, Wei Wei, Wen-Lian Hsu, Cho-Jui Hsieh. On the Robustness of Self-Attentive Models. In Proceedings of Association for Computational Linguistics (ACL), 2019.
31. **Minhao Cheng**, Wei Wei, Cho-Jui Hsieh: Evaluating and Enhancing the Robustness of Dialogue Systems: A Case Study on a Negotiation Agent. In Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT), 2019.
32. **Minhao Cheng**, Thong Le, Pin-Yu Chen, Huan Zhang, Jinfeng Yi, Cho-Jui Hsieh: Query-Efficient Hard-label Black-box Attack: An Optimization-based Approach. In International Conference on Learning Representations (ICLR), 2019.
33. Huang Fang, **Minhao Cheng**, Cho-Jui Hsieh, Michael Friedlander: Fast Training for Large-Scale One-versus-All Linear Classifiers using Tree-Structured Initialization. In SIAM International Conference on Data Mining (SDM), 2019.
34. Yao Li, **Minhao Cheng**, Kevin Fujii, Fushing Hsieh, Cho-Jui Hsieh. Learning from Group Comparisons: Exploiting Higher Order Interactions. In Advances in Neural Information Processing Systems (NIPS), 2018.
35. Xuanqing Liu, **Minhao Cheng**, Huan Zhang, Cho-Jui Hsieh. Towards Robust Neural Networks via Random Self-ensemble. European Conference on Computer Vision (ECCV), 2018.
36. **Minhao Cheng**, Ian Davidson, Cho-Jui Hsieh. Extreme Learning to Rank via Low Rank Assumption. In International Conference on Machine Learning (ICML), 2018.
37. **Minhao Cheng**, Cho-Jui Hsieh. Distributed Primal-Dual Optimization for Non-uniformly Distributed Data. In International Joint Conference on Artificial Intelligence (IJCAI), 2018.
38. Huang Fang, **Minhao Cheng**, Cho-Jui Hsieh. A Hyperplane-based Algorithm for Semi-supervised Dimension Reduction. IEEE International Conference on Data Mining (ICDM), 2017.

Preprint

Minhao Cheng, Zeyu Qin. Class-wise Visual Explanations for Deep Neural Networks.

Minhao Cheng, Zhe Gan, Yu Cheng, Shuohang Wang, Cho-Jui Hsieh, Jingjing Liu. Adversarial Masking: Towards Understanding Robustness Trade-off for Generalization.

Huan Zhang, **Minhao Cheng**, Cho-Jui Hsieh. Enhancing Certifiable Robustness via a Deep Model Ensemble.

Xiaoyun Wang, **Minhao Cheng**, Joe Eaton, Cho-Jui Hsieh, Felix Wu. Attack graph convolutional networks by adding fake nodes.

Liu Liu, **Minhao Cheng**, Cho-Jui Hsieh, Dacheng Tao. Stochastic Zeroth-order Optimization via Variance Reduction method.

Patents

Minhao Cheng, Pin-Yu Chen, Sijia Liu, Shiyu Chang, Payel Das. Method and System of Training Robust Machine Learning Models. US Patent 11,416,775 B2.

Minhao Cheng, Xiaocheng Tang, Chu-Cheng Hsieh. Hierarchical Classification Using Neural Networks. US Patent 2019/0171913 A1.

Talks

"Generating Class-wise Visual Explanations for Deep Neural Networks". ICLR TML4H Workshop, May 2023.

"The Secret Sauce in ChatGPT". HKUST CSE Undergraduate Seminar, April 2023.

"Towards Trustworthy Machine Learning: Training-time and Test-time Integrity". HKUST CSE Departmental Seminars, October 2022.

"Towards Automated and Trustworthy Machine Learning". SUSTech CSE Seminar, January 2022.

Professional Services

- Senior Programming Committee: AAAI
- Paper Reviewer/Programming Committee: ICML, IJCAI, ACL, NeurIPS, ICLR, AISTATS, Neurocomputing, IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Information Forensics & Security, IEEE Wireless Communications Letters, IEEE Transactions on Cybernetics, JMLR, IEEE TNNLS, TMLR.

Teaching Experience

Hong Kong University of Science and Technology, Hong Kong

- COMP 6211I Trustworthy Machine Learning. — Spring 2023
- COMP 5212 Machine Learning. — Fall 2022, Fall 2023

University of California, Los Angeles, Los Angeles, USA

- CS 180 Introduction to Algorithms. — Spring 2020
- CS 33 Introduction to Computer Organization. — Fall 2019
- CS 260 Machine Learning Algorithms. — Winter 2019