CHEJIAN XU

Tel: {+1} 217-979-7063 | **Email:** chejian2@illinois.edu | **Homepage:** xuchejian.com

EDUCATION

University of Illinois Urbana-Champaign (UIUC) *Aug. 2022 –*

- Ph.D. in Computer Science
- **Advisor**: *Prof.* Bo Li
- **Research interests**: Trustworthy ML, Natural Language Processing, Reinforcement Learning

Zhejiang University (ZJU) *Sept. 2017 – Jun. 2021*

- B.E. in Computer Science and Technology
- **Advisor**: *Prof.* Shouling Ji, *Prof.* Siliang Tang
- **GPA**: 3.94/4.0

PUBLICATIONS

DecodingTrust: A Comprehensive Assessment of Trustworthiness in GPT Models
Boxin Wang, Weixin Chen, Hengzhi Pei, Chulin Xie, Mintong Kang, Chenhui Zhang, ***Chejian Xu****, Zidi Xiong, Ritik Dutta, Rylan Schaeffer, Sang T. Truong, Simran Arora, Mantas Mazeika, Dan Hendrycks, Zinan Lin, Yu Cheng, Sanmi Koyejo, Dawn Song, Bo Li.*
Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS), 2023 (Oral)

DiffScene: Diffusion-Based Safety-Critical Scenario Generation for Autonomous Vehicles
Chejian Xu*, Ding Zhao, Alberto Sangiovanni-Vincentelli, Bo Li*
Workshop on New Frontiers in Adversarial Machine Learning at ICML 2023

SafeBench: A Benchmarking Platform for Safety Evaluation of Autonomous Vehicles
Chejian Xu**, Wenhao Ding*, Weijie Lyu, Zuxin Liu, Shuai Wang, Yihan He, Hanjiang Hu, Ding Zhao, Bo Li*
Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS), 2022

A Survey on Safety-Critical Driving Scenario Generation – A Methodological Perspective
Wenhao Ding, ***Chejian Xu****, Mansur Arief, Haohong Lin, Bo Li, Ding Zhao*
IEEE Transactions on Intelligent Transportation Systems (T-ITS), 2023

SemAttack: Natural Textual Attacks via Different Semantic Spaces
Boxin Wang,* ***Chejian Xu*****, Xiangyu Liu, Yu Cheng, Bo Li*
North American Chapter of the Association for Computational Linguistics (NAACL), 2022 (Findings)

Copy Motion From One to Another: Fake Motion Video Generation
Zhenguang Liu, Sifan Wu, ***Chejian Xu****, Xiang Wang, Lei Zhu, Shuang Wu, Fuli Feng*
31st International Joint Conference on Artificial Intelligence (IJCAI), 2022

COPA: Certifying Robust Policies for Offline Reinforcement Learning against Poisoning Attacks
Fan Wu, Linyi Li*,* ***Chejian Xu****, Huan Zhang, Bhavya Kailkhura, Krishnaram Kenthapadi, Ding Zhao, Bo Li*
The Tenth International Conference on Learning Representations (ICLR), 2022

Adversarial GLUE: A Multi-Task Benchmark for Robustness Evaluation of Language Models

Boxin Wang, **Chejian Xu***, Shuohang Wang, Zhe Gan, Yu Cheng, Jianfeng Gao, Ahmed Hassan Awadallah, Bo Li.*

Thirty-fifth Conference on Neural Information Processing Systems (NeurIPS), 2021 (Oral)

RESEARCH EXPERIENCES

Secure Learning Lab, University of Illinois Urbana-Champaign *May. 2020 – Aug. 2022*
Research Intern

Network System Security & Privacy Research Lab, ZJU *Jun. 2021 – Aug. 2022*
Research Assistant

Network System Security & Privacy Research Lab, ZJU *Mar. 2020 – Jun. 2021*
Research Intern

ACADEMIC SERVICES

Conference Reviewer	NeurIPS 2022-2023, AAAI 2023-2024, AACL 2022
Journal Reviewer	IEEE T-ITS
Organizer	CVPR 2023 SSAD Workshop, NeurIPS 2022 DMLW Workshop

HONORS, SCHOLARSHIP and AWARDS

NeurIPS Scholar Award	*2022*
Third-Class Scholarship for Outstanding Students of ZJU	*2020*
Third-Class Scholarship for Outstanding Students of ZJU	*2019*
First Prize for College Physics Competition of Zhejiang Province, China	*2019*
First-Class Scholarship for Outstanding Students of ZJU	*2018*

INVITED TALKS

Trustworthy Reinforcement Learning and Autonomous Driving *Oct. 2023*
AI Times, China (online)

Trustworthy Reinforcement Learning and Autonomous Driving *Oct. 2022*
Machine Heart, China (online)