Tianwei Zhang

Contact Information

Email: tianwei.zhang@ntu.edu.sg
Webpage: https://personal.ntu.edu.sg/tianwei.zhang/
Address: Blk N4, 02a-15, 50 Nanyang Ave, Singapore, 639798
Phone: +65 67906277

Education

Princeton University, Princeton, NJ, USA Sep.2011 - Aug.2017

- Ph.D., Electrical Engineering
- M.A., Electrical Engineering

Peking University, Beijing, P.R.China Sep.2007 - July.2011

- B.S., Physics
- B.A., Economics

Honors and Awards

- Distinguished Artifact Award, ACM Conference on Computer and Communications Security (CCS), 2024
- Distinguished Artifact Award, Usenix Security Symposium, 2024
- Outstanding Paper Award, Annual Meeting of the Association for Computational Linguistics (ACL), 2024
- Stamatis Vassiliadis Best Paper Award Nominee, International Conference on Field-Programmable Logic and Applications (FPL), 2024
- Distinguished Paper Award, Workshop on Artificial Intelligence System with Confidential Computing (AISCC), 2024
- Distinguished Paper Award, ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), 2023
- Best Paper Award, IEEE International Conference on Big Data Security on Cloud (BigDataSecurity), 2023
- Best Associate Editor Award, IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), 2023
- Best Paper Award, IEEE International Conference on Data Intelligence and Security (ICDIS), 2022
- Most Innovative Paper Award, IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA), 2021

Research Interests

AI Security, Privacy and Safety, with the following applications:

- Large Language Models, and Multimodal Models.
- Agents, Embodied AI, and multi-agent systems
- Data privacy protection in AI
- Autonomous and robotic system

Professional Experience

Nanyang Technological University, Singapore

- Associate Professor, College of Computing and Data Science (CCDS), Sept. 2024 - Now
- Deputy Director, Cyber Security Research Centre @ NTU (CYSREN), Sept. 2024 - Now
- Associate Director, NTU Centre Computational Technologies for Finance (CCTF), Apr. 2023 - Now
- Assistant Professor, School of Computer Science and Engineering (SCSE), Aug. 2019 - Aug. 2024

Amazon, Seattle, WA, USA

- Software Engineering, Sept. 2017 - July. 2019

Princeton University, Princeton, NJ, USA

- Research Assistant, Feb. 2012 - Aug. 2017

Selected Publications

Google Citation: 7723, H-index: 42, as of February 1, 2025

International Conferences:

[17] Junqi Zhang, Shaoyin Cheng, Linqing Hu, Jie Zhang, Chengyu Shi, Xingshuo Han, Tianwei Zhang, Yueqiang Cheng, Weiming Zhang, The Ghost Navigator: Revisiting the Hidden Vulnerability of Localization in Autonomous Driving, USENIX Security Symposium, August, 2025

[16] Haolin Wu, Chang Liu, Jing Chen, Ruiying Du, Kun He, Yu Zhang, Cong Wu, Tianwei Zhang, Qing Guo, Jie Zhang, When Translators Refuse to Translate: A Novel Attack to Speech Translation Systems, USENIX Security Symposium, August, 2025

[15] Boheng Li, Yanhao Wei, Yankai Fu, Zhenting Wang, Yiming Li, Jie Zhang, Run Wang, Tianwei Zhang, Towards Reliable Verification of Unauthorized Data Usage in Personalized Text-to-Image Diffusion Models, IEEE Symposium on Security and Privacy (S&P), May, 2025

[14] Yutong Wu, Jie Zhang, Florian Kerschbaum, Tianwei Zhang, THEMIS: Regulating Textual Inversion for Personalized Concept Censorship, Network and Distributed System Security Symposium (NDSS), February, 2025

[13] Xingshuo Han, Haozhao Wang, Kangqiao Zhao, Gelei Deng, Yuan Xu, Hangcheng Liu, Han Qiu, Tianwei Zhang, VisionGuard: Secure and Robust Visual Perception of Autonomous Vehicles in Practice, ACM Conference on Computer and Communications Security (CCS), October, 2024

[12] Yuan Xu, Gelei Deng, Xingshuo Han, Guanlin Li, Han Qiu, Tianwei Zhang, PhyScout: Detecting Sensor Spoofing Attacks via Spatio-temporal Consistency, ACM Conference on Computer and Communications Security (CCS), October, 2024

[11] Kunsheng Tang, Wenbo Zhou, Jie Zhang, Aishan Liu, Gelei Deng, Shuai Li, Peigui Qi, Weiming Zhang, Tianwei Zhang, Nenghai Yu, GenderCARE: A Comprehensive Framework for Assessing and Reducing Gender Bias in Large Language Models, ACM Conference on Computer and Communications Security (CCS), Distinguished Artifact Award, October, 2024

[10] Gelei Deng, Yi Liu, Víctor Mayoral-Vilches, Peng Liu, Yuekang Li, Yuan Xu, Tianwei Zhang, Yang Liu, Martin Pinzger, Stefan Rass, PentestGPT: Evaluating and Harnessing Large Language Models for Automated Penetration Testing, USENIX Security Symposium, Distinguished Artifact Award, August, 2024

[9] Meng Hao, Weiran Liu, Liqiang Peng, Hongwei Li, Cong Zhang, Hanxiao Chen, Tianwei Zhang, Unbalanced Circuit-PSI from Oblivious Key-Value Retrieval, USENIX Security Symposium, August, 2024

[8] Meng Hao, Hanxiao Chen, Hongwei Li, Chenkai Weng, Yuan Zhang, Haomiao Yang, Tianwei Zhang, Scalable Zero-knowledge Proofs for Non-linear Functions in Machine Learning, USENIX Security Symposium, August, 2024

[7] Xingshuo Han, Yutong Wu, Qingjie Zhang, Yuan Zhou, Yuan Xu, Han Qiu, Guowen Xu, Tianwei Zhang, Backdooring Multimodal Learning, IEEE Symposium on Security and Privacy (S&P), May, 2024

[6] Gelei Deng, Yi Liu, Yuekang Li, Kailong Wang, Ying Zhang, Zefeng Li, Haoyu Wang, Tianwei Zhang, Yang Liu, MASTERKEY: Automated Jailbreaking of Large Language Model Chatbots, Network and Distributed System Security Symposium (NDSS), February, 2024

[5] Chang Liu, Jie Zhang, Tianwei Zhang, Xi Yang, Weiming Zhang, Nenghai Yu, Detecting Voice Cloning Attacks via Timbre Watermarking, Network and Distributed System Security Symposium (NDSS), February, 2024

[4] Gelei Deng, Zhiyi Zhang, Yuekang Li, Yi Liu, Tianwei Zhang, Yang Liu, Guo Yu, Dongjin Wang, NAUTILUS: Automated RESTful API Vulnerability Detection, USENIX Security Symposium, August, 2023

[3] Jialai Wang, Ziyuan Zhang, Meiqi Wang, Han Qiu, Tianwei Zhang, Qi Li, Zongpeng Li, Tao Wei, Chao Zhang, Aegis: Mitigating Targeted Bit-flip Attacks against Deep Neural Networks, USENIX Security Symposium, August, 2023

[2] Ke Jiang, Yuyan Bao, Shuai Wang, Zhibo Liu, Tianwei Zhang, Cache Refinement Type for Side-channel Detection of Cryptographic Software, ACM Conference on Computer and Communications Security (CCS), November, 2022

[1] Gelei Deng, Guowen Xu, Yuan Zhou, Tianwei Zhang, Yang Liu, On the (In)Security of Secure ROS2, ACM Conference on Computer and Communications Security (CCS), November, 2022

Journals

[21] Xiaoyuan Liu, Hongwei Li, Guowen Xu, Xilin Zhang, Tianwei Zhang, Jianying Zhou, Secure and Lightweight Feature Selection for Horizontal Federated Learning, Accepted by IEEE Transactions on Information Forensics and Security

[20] Wenbo Jiang, Hongwei Li, Guowen Xu, Hao Ren, Haomiao Yang, Tianwei Zhang, Shui Yu, Rethinking the Design of Backdoor Triggers and Adversarial Perturbations: A Color Space Perspective, Accepted by IEEE Transactions on Dependable and Secure Computing

[19] Hao Ren, Guowen Xu, Tianwei Zhang, Jianting Ning, Xinyi Huang, Hongwei Li, Rongxing Lu, Efficiency Boosting of Secure Cross-platform Recommender Systems over Sparse Data, Accepted by IEEE Transactions on Dependable and Secure Computing

[18] Hangcheng Liu, Yuan Zhou, Ying Yang, Qingchuan Zhao, Tianwei Zhang, Tao Xiang, Stealthiness Assessment of Adversarial Perturbation: From A Visual Perspective, IEEE Transactions on Information Forensics and Security, Volume: 20, December 2024

[17] Hanxiao Chen, Hongwei Li, Meng Hao, Jia Hu, Guowen Xu, Xilin Zhang, Tianwei Zhang, SecBNN: Efficient Secure Inference on Binary Neural Network, IEEE Transactions on Information Forensics and Security, Volume: 19, November 2024

[16] Yuan Xu, Yungang Bao, Sa Wang, Tianwei Zhang, Function Interaction Risks in Robot Apps: Analysis and Policy-based Solution, IEEE Transactions on Dependable and Secure Computing, Volume: 21, Issue: 4, July 2024

[15] Guowen Xu, Xingshuo Han, Gelei Deng, Tianwei Zhang, Shengmin Xu, Jianting Ning, Anjia Yang, Hongwei Li, VerifyML: Obliviously Checking Model Fairness Resilient to Malicious Model Holder, IEEE Transactions on Dependable and Secure Computing, Volume: 21, Issue: 4, July 2024

[14] Guowen Xu, Xingshuo Han, Tianwei Zhang, Shengmin Xu, Jianting Ning, Xinyi Huang, Hongwei Li, Robert Deng, SIMC 2.0: Improved Secure ML Inference Against Malicious Clients, IEEE Transactions on Dependable and Secure Computing, Volume: 21, Issue: 4, July 2024

[13] Zhirui Zeng, Tao Xiang, Shangwei Guo, Jialing He, Qiao Zhang, Guowen Xu, Tianwei Zhang, Contrast-then-Approximate: Analyzing Keyword Leakage of Generative Language Models, IEEE Transactions on Information Forensics and Security, Volume: 19, April 2024

[12] Renyang Liu, Wei Zhou, Tianwei Zhang, Kangjie Chen, Jun Zhao, Kwok-Yan Lam, Boosting Black-box Attack to Deep Neural Networks with Conditional Diffusion Models, IEEE Transactions on Information Forensics and Security, Volume: 19, April 2024

[11] Wenbo Jiang, Hongwei Li, Guowen Xu, Tianwei Zhang, Rongxing Lu, A Comprehensive Defense Framework against Model Extraction Attacks, IEEE Transactions on Dependable and Secure Computing, Volume: 21, Issue: 2, March 2024

[10] Wenbo Jiang, Tianwei Zhang, Han Qiu, Hongwei Li, Guowen Xu, Incremental Learning, Incremental Backdoor Threats, IEEE Transactions on Dependable and Secure Computing, Volume: 21, Issue: 2, March 2024

[9] Rui Xue, Kaiping Xue, Bin Zhu, Xinyi Luo, Tianwei Zhang, Qibin Sun, Jun Lu, Differentially Private Federated Learning with an Adaptive Noise Mechanism, IEEE Transactions on Information Forensics and Security, Volume: 19, September 2023

[8] Guowen Xu, Xingshuo Han, Shengmin Xu, Tianwei Zhang, Hongwei Li, Xinyi Huang, Robert Deng, Hercules: Boosting the Performance of Privacy-preserving Federated Learning, IEEE Transactions on Dependable and Secure Computing, Volume: 20, Issue: 5, September 2023

[7] Hangcheng Liu, Tao Xiang, Shangwei Guo, Han Li, Tianwei Zhang, Xiaofeng Liao, Erase and Repair: An Efficient Box-Free Removal Attack on High-Capacity Deep Hiding, IEEE Transactions on Information Forensics and Security, Volume: 18, August 2023

[6] Jianfei Sun, Guowen Xu, Tianwei Zhang, Xuehuan Yang, Mamoun Alazab, Robert Deng, Privacy-aware and Security-enhanced Efficient Matchmaking Encryption, IEEE Transactions on Information Forensics and Security, Volume: 18, July 2023

[5] Kaidi Jin, Tianwei Zhang, Chao Shen, Yufei Chen, Ming Fan, Chenhao Lin, Ting Liu, Can We Mitigate Backdoor Attack Using Adversarial Detection Methods? IEEE Transactions on Dependable and Secure Computing, Volume: 20, Issue: 4, July 2023

[4] Meng Hao, Hongwei Li, Hanxiao Chen, Pengzhi Xing, Tianwei Zhang, FastSecNet: An Efficient Cryptographic Framework for Private Neural Network Inference, IEEE Transactions on Information Forensics and Security, Volume: 18, March 2023

[3] Hanxiao Chen, Hongwei Li, Yingzhe Wang, Meng Hao, Guowen Xu, Tianwei Zhang, PriVDT: An Efficient Two-Party Cryptographic Framework for Vertical Decision Trees, IEEE Transactions on Information Forensics and Security, Volume: 18, December 2022

[2] Jianfei Sun, Guowen Xu, Tianwei Zhang, Xuehuan Yang, Mamoun Alazab, Robert Deng, Verifiable, Fair and Privacy-preserving Broadcast Authorization for Flexible Data Sharing in Clouds, IEEE Transactions on Information Forensics and Security, Volume: 18, December 2022

[1] Jianfei Sun, Guowen Xu, Tianwei Zhang, Mamoun Alazab, Robert H. Deng, A Practical Fog-based Privacy-preserving Online Car-hailing Service System, IEEE Transactions on Information Forensics and Security, Volume: 17, August 2022

Professional Services

Organization Committee:

- General co-chairs: 7th International Conference on Frontiers in Cyber Security (FCS) : 2024

- Publicity co-chair: 27th Conference on Innovation in Clouds, Internet and Networks (ICIN) : 2024
- Program co-chair: Hardware and Architecture Support for Security and Privacy (HASP) : 2023 - 2024
- Demos co-chair: 25th Conference on Innovation in Clouds, Internet and Networks (ICIN): 2023
- General co-chair: International Conference on Knowledge Science, Engineering and Management (KSEM): 2022
- Publicity co-chair: Asian Hardware Oriented Security and Trust Symposium (AsianHost): 2022
- Registration chair: IEEE International Symposium on Secure and Private Execution Environment Design (SEED): 2021

Technical Program Committee:

- Neural Information Processing Systems (NeurIPS) Track on Datasets and Benchmark: 2024 (Area Chair)
- AAAI Conference on Artificial Intelligence (AAAI): 2023 - 2024 (Senior PC)
- IEEE International Conference on Data Security and Privacy Protection (DSPP): 2024
- International Conference on Information and Communications Security (ICICS): 2019 - 2024
- IEEE International Symposium on Secure and Private Execution Environment Design (SEED): 2021 - 2024
- IEEE International Conference on Data Mining (ICDM): 2023
- Asian Hardware Oriented Security and Trust Symposium (AsianHOST): 2023
- ACM ASIA Conference on Computer and Communications Security (AsiaCCS): 2021 - 2022
- EAI International Conference on Security and Privacy in Communication Networks (SecureComm): 2022
- Hardware and Architectural Support for Security and Privacy (HASP): 2018 - 2021
- IEEE International Conference on Distributed Computing Systems (ICDCS): 2020
- ACM Cloud Computing Security Workshop (CCSW): 2019 - 2020
- IEEE International Conference on Computational Science and Engineering (CSE), 2019
- IEEE International Conference on Big Data Security on Cloud (BigDataSecurity), 2019
- IEEE International Conference on Cyber Security and Cloud Computing (CSCloud), 2019
- ACM International Symposium on Blockchain and Secure Critical Infrastructure (BSCI), 2019

Journal Editorial:

- Associate Editor: IEEE Transactions on Circuits and Systems for Video Technology: 2022-now
- Guest Editor: IEEE Internet of Things Journal Special Issue on Security and Privacy in Large Language Models for Internet of Things (IoT), 2024
- Guest Editor: ACM Transactions on Sensor Networks Special Issue on Energy-efficient and Secure Computing for Artificial Intelligence and Beyond, 2021

Teaching Experience

Nanyang Technological University

- SC6115: Introduction to Cyber Security, 2024
- SC3010: Computer Security, 2020 – 2024
- SE6011: Network Security, 2023 – 2024
- SC6108: Blockchain & Smart Contract Security, 2023-2024
- SC6113: Development of Decentralized Applications, 2023
- CZ3007: Compiler Techniques, 2019 – 2021