YUGENG LIU

Stuhlsatzenhausweg 5, 66123, Saarbrücken, Germany
yugeng.liu@cispa.de ⋄ https://liu.ai

EDUCATION

CISPA Helmholtz Center for Information Security Since 10/2019
Ph.D. Student
Supervisor: Michael Backes and Yang Zhang

Shanghai Jiao Tong University 09/2014 - 07/2018
Bachelor in Computer Science and Technology

The University of Melbourne (Summer School) 06/2015 - 09/2015
Exchange Student in School of Engineering

PUBLICATIONS

Games and Beyond: Analyzing the Bullet Chats of Esports Livestreaming
Yukun Jiang, Xinyue Shen, Rui Wen, Zeyang Sha, Junjie Chu, **Yugeng Liu**, Michael Backes, Yang Zhang. *ICWSM*, 2024.

Comprehensive Assessment of Jailbreak Attacks Against LLMs
Junjie Chu, **Yugeng Liu**, Ziqing Yang, Xinyue Shen, Michael Backes, Yang Zhang. *Preprint*, 2024.

Robustness Over Time: Understanding Adversarial Examples' Effectiveness on Longitudinal Versions of Large Language Models
Yugeng Liu*, Tianshuo Cong*, Zhengyu Zhao, Michael Backes, Yun Shen, Yang Zhang. *Preprint*, 2023 (* equal contribution).

Watermarking Diffusion Model
Yugeng Liu, Zheng Li, Michael Backes, Yun Shen, Yang Zhang. *Preprint*, 2023.

Backdoor Attacks Against Dataset Distillation
Yugeng Liu, Zheng Li, Michael Backes, Yun Shen, Yang Zhang. *NDSS*, 2023.

ML-Doctor: Holistic Risk Assessment of Inference Attacks Against Machine Learning Models
Yugeng Liu*, Rui Wen*, Xinlei He, Ahmed Salem, Zhikun Zhang, Michael Backes, Emiliano De Cristofaro, Mario Fritz, Yang Zhang. *USENIX Security*, 2022 (* equal contribution).

Securing Android Applications via Edge Assistant Third-Party Library Detection
Zhushou Tang, Minhui Xue, Guozhu Meng, Chengguo Ying, **Yugeng Liu**, Jianan He, Haojin Zhu, Yang Liu. *Computers & Security*, 2019.

HoMonit: Monitoring SmartHome Apps from Encrypted Traffic
Wei Zhang*, Yan Meng*, **Yugeng Liu**, Xiaokuan Zhang, Yinqian Zhang, Haojin Zhu. *CCS*, 2018.

SERVICES

Conference PC Member:

- 2025: IEEE SaTML
- 2024: ACM WWW
- 2022: SocInfo

Journal Reviewer:

- 2024: IEEE TCSVT, IEEE JETCAS
- 2023: IEEE TDSC
- 2022: IEEE TDSC

HONORS & AWARDS

Honorable Mention of American Mathematical Contest in Modeling	02/2017
Second Award of Ericsson Hackathon	04/2015
Academic Excellence Scholarship of Shanghai Jiao Tong University	2014 - 2015

INVITED TALKS

2023.09: Secure Machine Learning, Xidian University, China.
2024.06: Trustworthy Large Language Models, Zhejiang University, China.