CiCi (Yutong) Cheng Ph.D.

(+1) 540-666-1970
yutongcheng@vt.edu
ccsnow127.github.io

Education

Department of Computer Science, Virginia Tech Blacksburg, VA, United States
Ph.D. in Computer Science 2024 - 2028 *(expected)*

- Advisor: Prof. Peng Gao
- Research area: Large Language Model and its Applications, Security, Software Engineering
- Selected courses: Natural Language Processing, System and Software Security, Trustworthiness of Large Language Models

School of Science and Engineering, Chinese Uni. of Hong Kong Shenzhen, China
MPhil in Computer and Information Engineering 2022 - 2024

- Advisor: Prof. Pinjia He
- Research area: Large Language Model and its Applications, Software Engineering
- Selected courses: Deep Learning Foundations and their Applications, Image Processing and Computer Vision, Mobile Computing and Internet of Things

School of Software Engineering, Beijing Jiaotong University Beijing, China
B.E. in Software Engineering 2018 - 2022

- Selected courses: Software System Analysis and Design, Software Architecture, Software Testing Technology and Practice, Computer Network, Database Systems, Reinforcement Learning

Publications

1. **Yutong Cheng**, Osama Bajaber, Peng Gao, Dawn Song. CTINEXUS: Unleash Cyber Threat Intelligence via In-Context Learning of Large Language Models. *Under peer review*, 2024.
2. Xu, Junjielong and Fu, Qiuai and Zhu, Zhouruixing and **Cheng, Yutong** and Li, Zhijing and Ma, Yuchi and He, Pinjia. Hue: A user-adaptive parser for hybrid logs. *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2023.

Projects

Characterizing, Detecting, and Correcting Comment Errors in Smart Contract Functions via In-Context Learning of Large Language Model
Github Repo 2023.06 - 2023.10

This project presents the first empirical study on 253 verified contracts encompassing 16,620 functions from Etherscan, uncovering that 87% of the smart contract functions have NatSpec comment errors and pinpointing prevalent error patterns. Based on our findings, we propose CETerminator, an automated approach for detecting and rectifying comment errors.

CTINEXUS: Unleash Cyber Threat Intelligence via In-Context Learning of Large Language Models.
Github Repo 2023.10 - 2024.7

This project presents CTINEXUS, a novel framework leveraging in-context learning (ICL) of large language models (LLMs) for data-efficient CTI knowledge ex- traction and high-quality cybersecurity knowledge graph (CSKG) construction.

Internships

Huawei PaaS Lab | Shenzhen, China 2022.06 - 2022.12

- Develop an adaptive ramp up algorithm for load testing
- Develop an heuristic algorithm for parsing hybrid logs in cloud system.

Tencent Yuewen Group | Beijing, China 2020.06 - 2020.09

- Implemented a dynamic user interface tailored to individual users' interests, collaborating with the recommendation algorithm team to integrate real-time content suggestion statistics.
- Developed a responsive framework that adjusts interface content based on users' real-time click feedback.

Awards

- **Bitshares Fellowship**, for smart-contract-related research project 2024.09
- **CCI SWVA Cyber Innovation Scholarship**, for CTI-related research project 2024.03

Services

- **Lab Lead**, for VT Security & Intelligence Lab 2024.08
- **Student Organizer**, for the 2024 DMV Security Workshop.
- **External Reviewer**, for NDSS 2025 (Top-tier Conference in Security) 2024.08
- **External Reviewer**, for CCS 2024 (Top-tier Conference in Security) 2024.03
- **External Reviewer**, for DLSP 2024 2024.02
- **External Reviewer**, for TOPS 2023 2024.01
- **External Reviewer**, for EuroS&P 2024 2023.12