

# YUTONG CHENG

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

## EDUCATION

Department of Computer Science, Virginia Tech  
*Ph.D. in Computer Science*

VA, United States  
2024 - present

- Advisor: Dr. Peng Gao
- Research area:
  - LLM/MLLM specialization and customization for applied settings
  - Automated and AI-assisted security operations (SecOps)

School of Software Engineering, Beijing Jiaotong University  
*B.E. in Software Engineering*

Beijing, China  
2018 - 2022

## PUBLICATIONS

1. **Yutong Cheng**, Osama Bajaber, Saimon Tsegai, Dawn Song, Peng Gao. CTINexus: Leveraging Optimized LLM In-Context Learning for Constructing Cybersecurity Knowledge Graphs Under Data Scarcity. *arXiv:2410.21060 (Under peer review)*, 2024.
2. Junjielong, Xu and Qiulai, Fu and Zhouruixing, Zhu and **Yutong, Cheng** and Zhijing, Li and Yuchi, Ma and Pinjia, He. 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 (ESEC/FSE)*, 2023.

## INTERNSHIPS & EXPERIENCE

**Llama Family** | Algorithm Developer

2023.09 - 2023.12

- Fine-tuned Llama2-Chat model with Chinese datasets to enhance Chinese conversational abilities.
- Integrated LangChain framework to improve model performance in Chinese applications.

**AtomEcho** | Algorithm Developer

2023.06 - 2023.09

- Developed a framework (AtomTool) for integrating tool utilization skills into LLMs.
- Implemented API integrations to enhance LLM's tool usage capability, improving LLM accuracy in handling complex API parameter settings.
- Reduced the hallucination phenomenon by refining prompt design and retrieval methods, improving LLM response accuracy and reliability.

**Chinese University of Hong Kong (Shenzhen)** | Research Assistant 2022.09-2023.06

- Collaborated with a cloud company to collect and curate log data from multiple departments, enabling comprehensive log analysis.
- Developed a two-step log grouping algorithm combining coarse- and fine-grained techniques, utilizing a multinomial search tree for efficient template extraction.
- Designed a template updating mechanism that incorporated human feedback to resolve ambiguous tokens, enhancing overall system accuracy.

**Yuewen Group, Tencent** | Backend Developer

2020.06 - 2020.09

- Developed dynamic UI customized to user interests, integrating real-time content recommendation statistics.
- Created responsive framework to adjust interface content based on user click feedback.

## AWARDS & HONORS

- **Bitshares Fellowship**, for smart-contract-related research project 2024.09
- **Virginia Tech CS Departmental Travel Grant** 2024.06
- **CCI SWVA Cyber Innovation Scholarship**, for CTI-related research project 2024.03

## PROFESSIONAL SERVICES

- **Student Organizer**, for the 2024 DMV Security Workshop.
- **External Reviewer**, for NDSS 2025, ACM CCS 2024, DLSP 2024, TOPS 2023, IEEE EuroS&P 2024