Zhengyu Liu

Ph.D. Student
Department of Computer Science
The Johns Hopkins University
Baltimore, MD 21218

zliu192@jhu.edu
jackfromeast.github.io
github.com/jackfromeast
Malone Hall

Research Interests

Web Security, System Security, and AI Security. My research spans two complementary directions:

- **Systematic Vulnerability Research.** I study *(In)security from System Design*: how design choices in widely used systems for functionality, convenience, and efficiency can introduce vulnerabilities into the applications built on them. Such systems include browsers, language runtimes, web infrastructure, and emerging AI-native systems. My work uncovers new vulnerability classes arising from these system designs, characterizes their root causes, exploitability, and real-world prevalence, and develops practical mitigations.
- **Red Teaming with LLMs and Agentic Systems.** Building on this foundation to (i) design reliable and scalable agentic systems that integrate LLMs with program analysis for automated vulnerability discovery, exploit generation, and mitigation, and (ii) develop methods to evaluate and advance the cybersecurity capabilities of the underlying models.

Education

Ph.D. **The Johns Hopkins University** — Jan 2024 - May 2027
Computer Science — Maryland, United States
Advisor: Prof. Yinzhi Cao

M.S. **The Johns Hopkins University** — Sep 2022 - Dec 2023
Computer Science — Maryland, United States

B.E. **Sichuan University** — Sep 2018 - June 2022
Cybersecurity — Sichuan, China
Advisor: Prof. Cheng Huang

Awards and Honors

1st Place (Exploitation), Red Teaming Track, Amazon Nova AI Challenge 2026 (Tournament Two)	June 2026
1st Place, Red Teaming Track, Amazon Nova AI Challenge 2026 (Tournament One)	April 2026
Notable Artifact Reviewer, USENIX Security 2025	Aug 2025
Honorable Mention, USENIX Security 2025	Aug 2025
Distinguished Paper Award, IEEE Symposium on Security and Privacy 2025	May 2025
Best Student Paper Award, ICICS 2021	Dec 2021
9th Place, 2021 ByteDance Security AI Competition, ByteDance (TikTok)	Nov 2021
2nd Place, School of Computing Summer Workshop, National University of Singapore	July 2021
Excellent Thesis, Innovation and Entrepreneurship Training Program for College Students	Sep 2020
¥30,000 Award, the 4th "Qiangwang Cup" National Cybersecurity Challenge	Sep 2020

Scholarships and Honors

"Cybersecurity Elite" Honor, School of Cyber Science and Engineering, Sichuan University	May 2022
The 404 Scholarship, School of Cyber Science and Engineering, Sichuan University	Dec 2021
First Class Scholarship, School of Cyber Science and Engineering, Sichuan University	Sep 2021
Outstanding Student, Sichuan University	2020, 2021
Second Class Scholarship, School of Cyber Science and Engineering, Sichuan University	Sep 2020

AI Security Competitions

Amazon Nova AI Challenge 2026, Team Jayl-Break, *Team Lead* — ***1st Place***

- Lead JHU's Team Jaylbreak, one of ten teams selected worldwide, **supported by a $250,000 research award**. The team placed first in the red-teaming track in the first two tournaments.
- Systematize and empirically evaluate existing red-team harness designs, including task decoupling, architecture, communication, tools, and memory, across effectiveness, efficiency, sustained exploration, and reliability objectives.
- Design and build an **ensemble agentic red-teaming system** combining the complementary strengths of agent reasoning, static analysis, and dynamic testing.
- Designed, built, and deployed **large-scale cloud infrastructure on AWS** for parallel agent execution, with automated environment provisioning, distributed task scheduling, autoscaling compute, persistent experiment storage, and centralized telemetry and cost monitoring.

DARPA AI Cyber Challenge (AIxCC), Team 42-b3yond-6ug, *Team Member* — ***Finalist***

- Led JHU's effort in developing an LLM-driven Cyber Reasoning System for autonomous vulnerability discovery and patching in critical open-source infrastructure. The team was selected as **one of seven finalist teams** and received a **$2,000,000 award**.
- Designed and implemented an **LLM-based fuzzing pipeline for Java** that uses execution feedback to adaptively synthesize harnesses and semantically structured inputs, targeting program states difficult to reach through mutation alone.

Publications

Peer-Reviewed Papers

2027 **Call Back Later: Studying and Detecting Reentrant Interpretation Vulnerability in Python Interpreters**
Zhengyu Liu, Zhuohao Zhang, Yingfei Xu, Yu Sun, Jiacheng Zhong, Letao Zhao, Jianjia Yu, Ziyang Li, and Yinzhi Cao
to appear in the Proceedings of the IEEE Symposium on Security and Privacy (S&P), 2027

Minnie: Dynamic Privacy Leak Detection in WeChat Miniapps via Unified Shadow Memory
Jianjia Yu, ***Zhengyu Liu***, Zhihan Xia, Penghui Li, Zifeng Kang, Junfeng Yang, and Yinzhi Cao
Network and Distributed System Security Symposium (NDSS), 2027

2026 **Comment and Control: Hijacking Agentic Workflows via Context-Grounded Evolution**
Zhengyu Liu*, Neil Fendley*, Aonan Guan, Jiacheng Zhong, and Yinzhi Cao (* equal contribution)
to appear in the Proceedings of the ACM Conference on Computer and Communications Security (CCS), 2026
Featured by The Register, VentureBeat, SecurityWeek, The Next Web, Cybernews, SANS ISC Stormcast, CSA Labs, Security Boulevard, Techzine, CyberSecurity News, GBHackers, Xataka, and FreeBuf

The First Large-Scale Systematic Study of Python Class Pollution Vulnerability
Zhengyu Liu, Jiacheng Zhong, Jianjia Yu, Muxi Lyu, Zifeng Kang, and Yinzhi Cao
In the Proceedings of the IEEE Symposium on Security and Privacy (S&P), 2026
Wiki: class-pollution.github.io / Code: python-class-pollution

Poisoned by the Host: Large-Scale Measurement of Host Name Poisoning in Web Applications
Rui Yang, Haoyu Wang, Zhicheng Sun, ***Zhengyu Liu***, and Yinzhi Cao
In the Proceedings of the IEEE Symposium on Security and Privacy (S&P), 2026

2025 **The DOMino Effect: Detecting and Exploiting DOM Clobbering Gadgets via Concolic Execution with Symbolic DOM**
Zhengyu Liu, Theo Lee, Jianjia Yu, Zifeng Kang, and Yinzhi Cao
In the Proceedings of the USENIX Security Symposium, 2025
Artifact Badges: Artifacts Available, Artifacts Functional, Results Reproduced / Code: TheHulk
Honorable Mention (6% of accepted papers)

Follow My Flow: Unveiling Client-Side Prototype Pollution Gadgets from One Million Real-World Websites
Zifeng Kang, Muxi Lyu, ***Zhengyu Liu***, Jianjia Yu, Runqi Fan, Song Li, and Yinzhi Cao
In the Proceedings of the IEEE Symposium on Security and Privacy (S&P), 2025
Distinguished Paper Award

2024 **Undefined-oriented Programming: Detecting and Chaining Prototype Pollution Gadgets in Node.js Template Engines for Malicious Consequences**
Zhengyu Liu, Kecheng An, and Yinzhi Cao
In the Proceedings of the IEEE Symposium on Security and Privacy (S&P), 2024
Nominee, Top 10 Web Hacking Techniques of 2024 (PortSwigger)

2022 **Coreference Resolution for Cybersecurity Entity: Towards Explicit, Comprehensive Cybersecurity Knowledge Graph with Low Redundancy**
Zhengyu Liu, Haochen Su, Nannan Wang, and Cheng Huang
In the Proceedings of the 18th EAI International Conference on Security and Privacy in Communication Networks (SecureComm), 2022
Code: CyberCoref

2021 **CyberRel: Joint Entity and Relation Extraction for Cybersecurity Concepts**
Yongyan Guo, ***Zhengyu Liu***, Cheng Huang, and Jiayong Liu
In the Proceedings of the International Conference on Information and Communication Security (ICICS), 2021
Best Student Paper Award

A Framework for Threat Intelligence Extraction and Fusion
Yongyan Guo, ***Zhengyu Liu***, Cheng Huang, Nannan Wang, Hai Min, Wenbo Guo, and Jiayong Liu
Computers & Security

A Sybil Detection Method in OSN based on DistilBERT and Double-SN-LSTM for Text Analysis
Xiaojie Xu, Jian Dong, ***Zhengyu Liu***, Jin Yang, Bin Wang, and Zhaoyuan Wang
In the Proceedings of the 17th EAI International Conference on Security and Privacy in Communication Networks (SecureComm), 2021

Preprints / Under Review

2026 **SoK: When Safe Agents Fail Together: The Security of Multi Agent LLM Systems**
Rui Yang, Junjie Xu, ***Zhengyu Liu***, Neil Fendley, Yang Hong, Ziyang Li, and Yinzhi Cao
arXiv:2609.00595, 2026

Automatic Exploitation of Sandboxed Interpreter Injection Vulnerabilities via Attacker Capability Chaining
Bofei Chen, ***Zhengyu Liu***, Yu Sun, Zhuohao Zhang, Jianjia Yu, Jiacheng Zhong, Lei Zhang, Ziyang Li, and Yinzhi Cao
Under review

Buzz to Boom: Detecting Message Progression Vulnerabilities in Electron Applications via Segmented Directed Fuzzing
Jianjia Yu, ***Zhengyu Liu***, Ziyang Li, and Yinzhi Cao
Under review

When Refusal Does Not Hold: Cybersecurity Assistance under Adversarial Conversational Contexts
Rui Yang, Yang Hong, Yichao Xu, ***Zhengyu Liu***, Yinzhi Cao, and Ziyang Li
Under review

Talks

2026 **LaunchBreak: a Slip of Tea, a Click, a Full Multi-stage Desktop Takeover**
Speaker, DEFCON 34, Las Vegas, Aug 2026
Link / *Slides*

Get Set, Exploit! Unveiling Python Class Pollution In-the-Wild
Speaker, DEFCON 34, Las Vegas, Aug 2026
Link / *Slides* / *Video*

Remote Server, Local Root. Welcome to MCP.
Contributor, BlackHat Asia 2026, Singapore, April 2026
Link

New and Overlooked Attack Surfaces: Detecting and Exploiting DOM Clobbering Gadgets
Speaker, RSAC 2026, San Francisco, March 2026
Link / *Slides*

Not My Vibe: When AI Coding Agents Go Off the Rails
Speaker, BSidesSF 2026, San Francisco, March 2026
Link / *Slides* / *Video*

2025 **The DOMino Effect: Automated Detection and Exploitation of DOM Clobbering Vulnerability at Scale**
Speaker, DEFCON 33, Las Vegas, Aug 2025
Link / *Slides* / *Video*

From Static to Smart: LLM-enhanced Static Analysis on Web Application Vulnerability Detection
Speaker, Ant Group SRC Annual Celebration, June 2025
Link

Professional Experience

Security Research Intern, supervised by Dr. Anton Kocheturov — May 2025 - Dec 2025
Siemens, Princeton, New Jersey, United States

Ph.D. Research Assistant, advised by Dr. Yinzhi Cao — Jan 2024 - Current
SecLab, The Johns Hopkins University, Baltimore, Maryland, United States

Research Intern, advised by Dr. Yinzhi Cao — June 2023 - Dec 2023
SecLab, The Johns Hopkins University, Baltimore, Maryland, United States

Research Assistant, advised by Dr. Cheng Huang — Aug 2020 - June 2022
Web Attack and Defense Lab, Sichuan University, Chengdu, Sichuan, China

Teaching Experience

Course Assistant, Practical Cryptographic Systems (EN.445/645), JHU — 2023 - 2025
Course Assistant, Web Security (EN.601.340/440/640), JHU — 2023 - 2025

Professional Services

Program Committee

PC, ACM Asia Conference on Computer and Communications Security (AsiaCCS) — 2027
PC, The Fourth International Workshop on Large Language Models for Code (LLM4Code '27) — 2027
PC, The Third International Workshop on Large Language Models for Code (LLM4Code '26) — 2026

Artifact Evaluation Committee

AEC, Network and Distributed System Security Symposium (NDSS) — 2026

AEC, ACM Conference on Computer and Communications Security (CCS)	2025, 2026
AEC, USENIX Security Symposium	2025

External Reviewer

IEEE Symposium on Security and Privacy (S&P)	2025, 2026
USENIX Security Symposium	2024–2026
ACM Conference on Computer and Communications Security (CCS)	2026
Network and Distributed System Security Symposium (NDSS)	2027
European Conference on Computer Systems (EuroSys)	2026
Annual Computer Security Applications Conference (ACSAC)	2025, 2026
Symposium on Research in Attacks, Intrusions, and Defenses (RAID)	2025, 2026
Workshop on Measurements, Attacks, and Defenses for the Web (MADWeb)	2025
IEEE Computer Security Foundations Symposium (CSF)	2024
Conference on Detection of Intrusions and Malware & Vulnerability Assessment (DIMVA)	2024

Capture the Flag

Team Member, **TheHackersCrew** team — Jan 2023 - Dec 2025
17 medals in 2024: 7 Gold, 5 Silver, and 5 Bronze; Top 10 on CTFTime.org

Team Lead, **Z0D1AC**, The Johns Hopkins University academic team — Oct 2022 - Dec 2025
2nd place, Raymond James CTF 2024 ($5,000 prize)
3rd place, Raymond James CTF 2023 ($2,500 prize)

StarBugs

Websites. Identified high-impact vulnerabilities across major websites, including www.nvidia.com, www.cnn.com, www.hp.com, intuit.com, ibm.com, ebay.com, okta.com, nist.gov, w3.org, ubuntu.com, espn.com, forbes.com, theguardian.com, f5.com, asus.com, mediafire.com, and iqiyi.com.

Critical Software Infrastructure. Discovered 114 memory-safety bugs in Python interpreters (CPython, MicroPython, and RustPython), including use-after-free, type confusion, and out-of-bounds access. Developed one into a full-chain exploit escaping ByteDance's Coze production sandbox.

CVEs. Discovered vulnerabilities across popular open-source projects, resulting in 50+ CVEs (full vulnerability list). Selected examples include Microsoft Azure CLI (CVE-2025-24049), Jupyter Notebook/JupyterLab (CVE-2024-43805), Webpack (CVE-2024-43788), Hugging Face Transformers (CVE-2025-5120), and AutoGPT (CVE-2024-10457).

Bug Bounty and Responsible Disclosure. Acknowledged and rewarded by Google, Microsoft, Anthropic, GitHub, Snowflake, Meta, ByteDance, Vercel, Hugging Face, Xiaomi, iQIYI, and Ant Group for reporting vulnerabilities that led to security fixes. Named on the Microsoft Researcher Recognition Program quarterly leaderboard in 2025 and 2026.

Technical Skills

Languages: Python, JavaScript/TypeScript, C/C++, Java, Assembly, CodeQL, SQL, PHP, Bash

AI/ML: PyTorch, Hugging Face Transformers, LLM fine-tuning and post-training (SFT, DPO/RLHF), agentic frameworks (MCP, LangChain), model red teaming, and evaluation harnesses

Program Analysis: CodeQL, Jalangi, Jazzer/JQF, AFL, Valgrind, IDA, Ghidra, and V8/Chromium instrumentation

Security and Infrastructure: Pwntools, Burp Suite, Wireshark, Docker/Kubernetes, QEMU, and AWS