

Lihang Liu

+1(984)268-6112 | lliu45@ncsu.edu
www.linkedin.com/in/caffcen | <https://caffcen.github.io>

EDUCATION

North Carolina State University
Program: Ph.D. of Computer Science

Raleigh, USA
08/2025-Expected Graduation: 05/2030

Northwestern University
Program: Master of Computer Science

Evanston, USA
09/2019-05/2021

South China University of Technology (SCUT)
Program: Bachelor of Engineering in Computer Science and Technology

Guangzhou, China
09/2015-07/2019

PROFESSIONAL SKILLS

Programming Skills: Proficiency in C/C++, Java, Python, SQL, Shell; **Familiar** with CUDA, Shell, Go
Framework & Techs: Proficiency in Redis, Pulsar, Kafka, ELK, Nginx, Spring Boot; **Familiar** with AWS, Alibaba Cloud
Research Interests: Distributed Systems, Database, Data-Intensive Applications & Systems, Computer Systems, Systems for AI & AI for Systems

WORK EXPERIENCE

Alibaba - ICBU (International Core Business Unit)

10/2023-03/2024, Hangzhou, China

Position: Senior Software Development Engineer

ICBU AI Agent Development Platform—R-Lab

- Accelerated AI agent development by leading the designing and implementing a production-ready agent framework based on [LangChain](#) architecture, enabling **the whole department** to migrate AI tasks to agent-based workflows within **1 month**.

alibaba.com LLM-Powered Search & Recommendation System

- Enhanced search accuracy and reduced query response time to **<200 ms** by architecting and deploying LLM-powered intent recognition and retrieval system using **ReAct** Agent Framework, serving **250K+** daily active users on *alibaba.com*.

Tencent - Tenpay

07/2021-09/2023, Shenzhen, China

Position: Software Development Engineer, WeChat Pay HK

WeChat Pay HK Data Observability Platform

- Transformed team from zero data visibility to comprehensive business intelligence by architecting distributed log aggregation platform, enabling our team to perform multi-dimensional analysis on **30+** metrics across user behavior and payment channels, processing **10M+** events daily with **99.99%** reliability and **<100ms** latency.
- Pioneered team's cloud infrastructure adoption as first engineer to integrate Tencent TDMQ (Pulsar) and Tenpay's new logging platform, authoring **5+** technical documents and establishing architectural patterns for multi-platform log distribution (ELK, data warehouse, BI), adopted by **20+** subsequent microservices.

WeChat Pay HK Native Payment Infrastructure Upgrade

- Reduced payment failure rate by **30%** and improved transaction completion time from seconds to **<1 second** by migrating H5-based payment flow to native WeChat client integration, orchestrating collaboration across 4 teams (WeChat Guangzhou, WeChat Pay HK Shenzhen/Hong Kong, Tenpay Cryptography).
- Enabled offline payment capability for **700K+** users in low-connectivity areas (subways, shopping malls) by architecting an algorithmic code system with pre-loaded cryptographic seeds, achieving **99%** transaction success rate in 3G/offline scenarios and resolving CEO-reported UX issues.

WeChat Pay HK Payment Method Service Modernization

- Improved system maintainability and reduced feature development time by **40%** by leading complete migration of

payment method integration service from legacy C++ service framework to tRPC-Java microservices, supporting **3** new payment integrations within **2** months.

- Enhanced code reusability and decreased code complexity by implementing various design patterns (BFF, Builder, etc.), establishing **4+** reusable components adopted across microservices.

RESEARCH EXPERIENCE

FORECAST - HPC Network Congestion Forecasting

09/2025-Present

- Designed fault-tolerant CR detection pipeline achieving **93%** parallel efficiency at 64 cores through stateless worker architecture, Parquet checkpointing, and Jaccard-based temporal CR merging (0.5 overlap threshold), reducing peak memory from **15GB to 500MB** per processing unit.
- Developed Spatio-Temporal Graph Neural Network (STGNN) for HPC congestion propagation forecasting, combining GRU temporal encoding, GAT spatial convolution, and multi-head attention to predict propagation probability, rate, and severity across **41,472** network links with **>75%** target accuracy.

Advanced Persistent Threat (APT) Defense System Research

09/2020-04/2021

Lab: [Northwestern Lab for Internet and Security Technology](#)

- Achieved **94%** detection accuracy and **<5%** false positive rate by implementing real-time APT defense system using provenance graph and tag propagation techniques, capable of analyzing **100K** kernel events per second.
- Enhanced system resilience by identifying **10+** critical vulnerabilities through simulated APT attack scenarios using **BeEF** penetration testing framework.
- Validated system effectiveness against known APT attack patterns by collecting kernel tracing data via **LTTng** and testing with both simulated and open-source APT datasets, achieving **90%+** threat detection coverage.

Linux Whitelist System Research

07/2018-08/2018

Lab: [Northwestern Lab for Internet and Security Technology](#)

- Improved system security configuration flexibility by developing a kernel-level whitelist system using **ftrace** and proc filesystem, enabling administrators to manage whitelist executables without system restart.
- Prevented unauthorized process executions during testing by implementing **ftrace**-based `sys_execve` hooking mechanism, achieving **300µs** overhead per process creation.