YI YANG

(917) 972 8286 | yiyang5@andrew.cmu.edu | github.com/2042third | https://pdm.pw

EDUCATION

Carnegie Mellon University — Pittsburgh, PA
M.S. in Information Security Policy and Management — *Aug 2023 - Dec 2025*

Rensselaer Polytechnic Institute — Troy, NY
B.S. in Computer Science; B.S. in Physics — *Sep 2018 - May 2023*

WORK EXPERIENCE

Network Tree — Remote
Software Engineer — *Aug 2024 - Jan 2025*

- Led **7-engineer team** in Agile environment as technical lead and **the primary code reviewer**, implementing **CI/CD best practices** and directing sprint-level task allocation for successful iOS app store release with more than 420 users at launch.
- Designed and implemented RESTful microservices using **Node.js** and **MongoDB** aggregation pipeline, and optimized for team accessibility and a streamlined, faster CI/CD pipeline.
- Built **AWS ECS** notification system using **EventBridge** schedules to trigger **Lambda** functions for Apple push notifications.

Network Tree — New York, NY
Fullstack Software Engineer Intern — *June 2024 - Aug 2024*

- Personally developed **12+ new features** in **React Native** and **ExpressJS**, where 9 are currently in the production build.
- Optimized microservices architecture with Express.js & Node.JS backend (1256ms to 45ms) through data normalization.
- Enhanced React Native app performance using **Redux Toolkit** and custom middleware, improving rendering from 3s to 136ms.

Rensselaer Polytechnic Institute, IBM Research — Troy, NY
Researcher & Software Engineer — *Jan 2022 - May 2022*

- Engineered high-throughput data pipeline analyzing 2.8M Python source files across **20,000+ repositories**, utilizing local MySQL and custom-built data-mining software.
- **First-author** of and developed all software used in award-winning paper in IEEE MSR '22 providing foundational insights for Python language optimization and future compiler designs.
- Built **Python AST analyzer** and parser to extract language features from source code, concluding that dynamic feature usage (e.g. *eval*, *exec*, ...) are related to other complex feature usage (e.g. *lmabda*, list comprehension, decorators, ...) in existing Python repositories, and 12.97% Python projects use dynamic features.
- Developed ML models using scikit-learn for Python Github repository's feature prediction based on existing characteristic (e.g. functional programming calls etc.), contributing to **IBM**'s Python static analysis/compilation research.

PROJECTS

PDM (Personal Data Manager) — Aug 2019 - Present

- Developed a multi-threaded, cross-platform **C++ encryption library** for PDM applications, supporting arbitrary data lengths with constant memory footprint per thread, having performance same as OpenSSL commend line with the same thread count.
- Built and self-hosts end-to-end encrypted data platform with **Go backend**, implementing custom authentication with SHA3 and zero-knowledge PostgreSQL storage using custom xchacha20-poly1305 encryption implementation, with **3300+ GitHub commits**.
- Implemented distributed system using **Redis** for caching frequently accessed data, **RabbitMQ** for async operations, and **PostgreSQL** with read replicas, reducing latencies by 85%.
- Implemented C++ mutex-free ring buffers for scalable parallel processing, architecture-agnostic SIMD intrinsics for ChaCha20 operations, and memory-mapped files for efficient disk I/O on desktop platforms (Windows, Linux, macOS) and **WebAssembly**.
- Wrote and scaled **Go server** supporting **3000+ concurrent users** through goroutine pooling and connection management, deployed with **Docker compose** and **GitHub Actions** for automated CI/CD pipeline on self-hosted Linux machines.
- Developed Qt desktop app with custom encrypted **SQLite** VFS implementation, **NUXT 3** web client with WebAssembly module, and React Native mobile app sharing core C++ library.

PUBLICATIONS

Y. Yang, A. Milanova and M. Hirzel, "Complex Python Features in the Wild," in 2022 IEEE/ACM 19th International Conference on Mining Software Repositories (MSR), Pittsburgh, PA, USA, 2022, pp. 282-293, doi: 10.1145/3524842.3528467.

SKILLS

Languages & Compilers : C/C++, Go, Java, Python, Javascript, TypeScript, Swift, Bash, Shell, gcc, clang, llvm, GraalVM

Frameworks: Go Standard library, Echo, Gorm, Mongoose, Vue, React Native, Nuxt, Angular, Spring, Spring Security, Data JPA, JDBC

Infrastructure: AWS (EC2, Lambda, EventBridge, S3), Docker, Redis, RabbitMQ, Fafka, Pulsar, MinIO, Red Hat Linux Enterprise, Alma Linux

Database: SQLite, PostgreSQL, MongoDB, Oracle SQL, MySQL, DynamoDB, Firebase, Cassandra