Yuta Nakamura

yutanak6@gmail.com

Address:
20671 Forge Way, Unit 122
Cupertino, CA, 95014

EDUCATION

DePaul University March 2020 – December 2024

Ph.D., Computer Science GPA 3.90/4.00, Researches Program Analysis (Static & Dynamic Analysis)

Coursework:

Distributed Systems I/II	Compiler Design	Advanced Database Concepts
Light Weight Virtual Machines	Database Technologies	Optimized C++

University of Chicago October 2018 – March 2020

M.S., Computer Science GPA 3.38/4.00

Focused on Data Analysis (Machine Learning/Artificial Intelligence)

Coursework:

C Programming	C/C++ for Advanced Programmers	Operating Systems
Introduction to Computer Systems	Foundations of Computational Data Analysis	Time Series Analysis and Stochastic Process
Algorithms	Introduction to Computer Security	Advanced Data Analytics
Machine Learning		High Performance Computing

Hitotsubashi University April 2012 – March 2018

B. A., Economics GPA 3.73/4.00

Relevant Coursework:

Introductory Statistics	Calculus 1B/2	Basic Course on Information Science
Introduction to Probability	Further Calculus	Computer Science
Basic Econometrics	Set And Topology 1/2	Applied Information Technology
Linear Algebra 1B/2	Mathematical Logic 1	

University of Pennsylvania August 2014 – May 2015

Exchange Student

Relevant Coursework:

Econometrics	Game Theory	Advanced Calculus

PROFESSIONAL EXPERIENCE

Meta Platforms Inc. Menlo Park, California

Research Scientist March 2025 – Present

- Incorporated potential multiple conversions to post a more suitable advertisement to customers
- Committed the most number of code changes (diffs) in the team in 2025 H2

DePaul DICE Lab Chicago, Illinois

Researcher June 2019 – March 2025

- Published 6 papers as a first author at peer-reviewed conferences about provenance alignment
- Currently creating software to apply provenance alignment for programs with nondeterminism

Meta Platforms Inc. Menlo Park, California

Software Engineer Intern – Payment Processing Team June 2024 – August 2024

- Worked to create a new database table to handle for the team
- Committed codes over 80 times to the production during the 11-week internship

SKILLS

- Proficient: C, C++, Python 3, x86 assembly, oracle SQL, Java, Vim, Linux, Git, Visual Studio
- Intermediate: OpenMPI, Akka, LLVM, CUDA, HTML, CSS, Scala, Visual Studio

PROJECTS

Ordo Ab Chao (A Private Repository at DePaul Dice Lab) December 2022 – June 2023

Experiment part of the publication "Efficient Differencing of System-level Provenance Graphs"

Defined another way to "diff" 2 sequences of trace while respecting loop iterations, which is vital for program analysis. The asymptotic time of the algorithm is linear (to the length of paths) with optimizations.

Raft Algorithm Implementation May 2022 – June 2022
[https://github.com/nakasan617/Raft]
Implementation of Raft, an algorithm to reach consensus value of a variable in a fault-tolerant way
Implemented with Akka, a toolkit for simplifying concurrent application development on JVM.

ProvScope November 2021 – June 2022
[https://github.com/depaul-dice/ProvScope]
Experiment part of the publication "Provenance-based Workflow Diagnostics Using Program Specification"
Given two traces and program specification (i.e. control flow graphs of each function), maps the traces onto the traces to convert to paths and compares them.

Content-Defined Merkle Tree June 2019 – November 2020
[https://github.com/depaul-dice/CDMT]
Experiment part of the publication "Content-Defined Container Delivery"
Implements Content-Defined Merkle Tree, which is a Merkle tree created on top of data which are divided with content-defined chunking. The Merkle tree is fragile against insertion operations, therefore we do the content-defined chunking on internal nodes of the tree for the robustness.

PUBLICATIONS

- Accurate Differential Analysis using Record and Selective Replay May 2025
  (Proceedings of the 37th International Conference on Scalable Scientific Data Management)
  Shows how to do differential analysis when target applications are parallel/distributed

- Accurate Path Prediction for Provenance-based Differential Analysis October 2024
  (33rd ACM International Conference on Information and Knowledge Management)
  Presents a machine learning approach to find an execution path given an execution trace

- Differential Analysis for System Provenance May 2024
  (2024 IEEE 40th International Conference on Data Engineering)
  Overall descriptions of my Ph.D. work, for debugging and analysis, we compare the execution traces

- Efficient Differencing of System-level Provenance Graphs October 2023
  (32nd ACM International Conference on Information and Knowledge Management)
  Introduced a method to systematically comprehend reasonable loop iterations in execution traces

- Provenance-based Workflow Diagnostics Using Program Specification December 2022
  (IEEE International Conference on High Performance Computing)
  Introduced debugging tools using function call trace of executions, uses numerous optimization techniques to reduce search space for "path explosion problem"

- Content-Defined Container Delivery December 2020
  (IEEE International Conference on High Performance Computing)
  Introduces a new data structure to allow fast access of parts of data divided with content-defined chunking

- Efficient Provenance Alignment in Reproduced Executions June 2020
  (The Theory and Practice of Provenance in Provenance Week 2020)
  Introduced system call sequence alignment

AWARDS

- National Science Foundation and N.A.S.A. Research Grant March 2020 – March 2025
- Josuikai Study Abroad Grant at University of Pennsylvania August 2014 – May 2015

LANGUAGES

- Japanese – Native Speaker
- English – Native Fluency