

HAOYU HE

he.haoyu1@northeastern.edu | <https://he-h.github.io/>

EDUCATION

Northeastern University

Boston, MA

Doctor of Philosophy

May. 2023 - present

- **Major:** Civil and Environmental Engineering
- **Advisors:** Dr. Ryan Wang and Dr. Kelsey Pieper

Rensselaer Polytechnic Institute (RPI)

Troy, NY

Bachelor of Science

Aug. 2018 – Dec. 2021

- **Major:** Computer Science and Mathematics
- **GPA:** 3.86/4.00

COMPUTER SKILLS

Programming languages: Python, Java, MATLAB, R, C++, C, SQL, and JavaScript

Machine learning libraries: NumPy, SciPy, Scikit-learn, PyTorch, PyTorch-Geometric, and NetworkX

Others: Git, LATEX, and AWS

RESEARCH AND PUBLICATIONS

Percolation of Temporal Hierarchical Mobility Networks during COVID-19 (First Author)

Published by Philosophical Transactions of the Royal Society A on Nov 22, 2021; IF: 4.2

Research Assistant. Supervisors: Qi Wang (Northeastern University) & Jianxi Gao (RPI) Aug. 2020 – May. 2021

- Proposed a temporal hierarchical model of human mobility with Metropolitan Statistical Areas (MSAs) as mid-level subunits to locate critical links in different levels of a large complex network
- Leveraged crowd-sourced and large-scale human mobility data to construct temporal hierarchical networks to examine percolation on both levels and demonstrated the changes of network metrics and the connected components under the influence of COVID-19
- Reduced the time of model construction and percolation process conduction of one-day US-based human mobility network from 10 hrs. to 30 mins to enable analysis on the long-term effect of disease propagation
- Provided data-driven insights into understanding the dynamical community structure of mobility networks during disruptions and contributed to more effective infectious disease control at multiple scales

RESEARCH PROJECTS

Scalable Graph Training: Leverage Score Sampling with Performance Guarantee

Northeastern University

Core Researcher

Sept. 2022 – now

- Sparsified graphs utilizing effective resistance inspired from electric network to sample essential edges to the graph so that can maintain spectral property of the graph (largest singular value)
- Speeded up the Graph Neural Networks (GNNs) training process with and achieved faster performance compared with same models without sparsification
- Harnessed ensemble learning to train multiple sparsified versions of a graph and beat the benchmark accuracy by ~3% with significant less time
- Partitioned large graphs to enable trained with limited memory size in an streaming setting

Safety and Traffic Prediction using GNNs with Calibration

Northeastern University

Core Researcher

Nov. 2022 – now

- Created a novel traffic collision dataset with Caltrans highway sensor data and California collision data from SWITRS
- Improving random oversampling algorithm to combat severely imbalanced classification problem and output a calibrated confidence to match the real-world scenario

Trajectory Prediction in Mobility Network

RPI

Core Researcher

May. 2021 – May. 2022

- Improved current machine learning approaches of spatio-temporal trajectory generation, next location prediction and daily trajectory prediction using feature engineering and LSTM
- Identified ways to completely utilize defective data by predict temporal pattern using unsupervised learning, adding radius of gyration, and assessing mobility entropy of each user

COVID Spread Network Analysis

RPI

Core Researcher

Aug. 2020 – May. 2021

- Harnessed percolation theory of edge removal to analyze and track the vulnerability of hierarchical mobility network model of U.S before and after the national emergency
- Determined critical bottlenecks of the propagation of COVID amid the complex U.S network, and carried out strategies along with other non-pharmaceutical intervention under CDC guidance to slow the spread

CPU Scheduler Simulator

Team Leader

RPI

Apr. 2021

- Led the group project of simulating the workflow of the CPU scheduler along with CPU memory and I/O subsystem and implementing some configurable CPU scheduling algorithms
- Proposed a new algorithm that combined and present better performance than round robin (RR) algorithm

COVID Infection Predictive Modeling and Reopening Strategies

Institute of Data Exploration and Applications, RPI

Core Researcher of COVID Analysis and Modeling Team

May 2020 – Aug. 2020

- Constructed machine learning models, include linear model and quadratic model, as well as status quo and phase reopening, to predict time-series infection rate curve based on social-distancing metric of different locations in COVID WarRoom
- Created a tool for schools to simulate new infection rate, impact of part size, etc. based on their basic statistics with previous models in COVID back-to-school

TEACHING ASSISTANT

Teaching Assistant, *Machine Learning and Data Mining*, Northeastern University

Jan. 2023 – now

- Assisted with curriculum development and led section discussions
- Graded assignments and provided individualized support for students

Mentor, *Computer Science I* and *Foundation of Comp Sci*, RPI

Aug. 2019 – Jun. 2021

- Selected for extraordinary performance while taking these 2 courses
- Led students in computer science lab once a week and answer students' questions during office hour

INTERNSHIP

Software Engineer Intern, NEUSOFT, Wuhan, Hubei, China

May – Aug. 2019

- Designed and developed campus web applications using Java and Python and allowed students to get statistics about Mechanical Engineering Department and compared data with previous year
- Updated and revised web pages based on students' needs weekly

EXTRACURRICULAR ACTIVITIES

Member, Upsilon Pi Epsilon (UPE), RPI

Dec. 2020 – now

- Assisted in facilitating the personal and professional growth of UPE students by financially supporting scholarships and other creative endeavors that include the discovery, integration, and application of knowledge
- Established significant partnerships with external constituencies and seek opportunities to extend individual memberships and chapter activities into additional environments yet to be identified

Member, Google Developer Student Club, RPI

Aug. 2020 – Dec. 2021

- Learned cutting edge technical skills through peer-to-peer learning environment and hands-on workshops
- Applied computer science knowledge to solve local problems to build great solutions with other members

SERVICES

Reviewer: AISTATS 2023, WSDM 2023

HONORS AND AWARDS

Dean's Honor List, RPI

Dec. 2018 – Dec. 2021

- Academic awards used to recognize RPI's high scholarship levels (every semester)

Letter of Recognition, RPI

May. 2021

- Recognized as one of the top few students with exceptional performance in CSCI-4210, Operating Systems

ADDITIONAL INFORMATION

Interests: Saxophone, Kendo, Go Game, Chess, Badminton and Table Tennis

Languages: English (fluent) and Chinese (native)