

Minghao Ye

370 Jay Street, 9th Floor
Brooklyn, NY 11201
United States

Phone: +1 (646) 256-0136
E-mail: minghao.ye@nyu.edu
sites.google.com/nyu.edu/minghao

Research Summary

I have been focusing on designing **intelligent traffic engineering** solutions to **optimize routing strategies** for **software-defined wide-area networks** using **artificial intelligence and machine learning** techniques, aiming to improve network load balancing performance with good QoS provisioning under dynamic traffic scenarios, reduce network disturbance and routing update overhead, accelerate traffic engineering operations in large-scale networks, and deploy robust routing strategies that are resilient to unexpected link failures.

Education

New York University (NYU) *New York, NY*
Ph.D. in Electrical Engineering | GPA: 3.98/4.0 *May 2024*
M.S. in Electrical Engineering | GPA: 4.0/4.0 *May 2019*
Dissertation Title: Leveraging Machine Learning for Intelligent Traffic Engineering Design in Wide-Area Networks (Advisor: Prof. H. Jonathan Chao)
Academic Honors: *The Alexander Hessel Award for the Best Ph.D. Dissertation in Electrical Engineering (2024), The Dante Youla Award for Graduate Research Excellence in Electrical Engineering (2023), ECE MS Student Academic Achievement Award (2018)*


The Hong Kong Polytechnic University (PolyU) *Hong Kong, China*
B.Eng. (Hons.) in Electronic Engineering *Jun. 2017*
Academic Honors: *Best Semester GPA Award, Best Academic Improvement Award (2017)*

Sun Yat-sen University (SYSU) *Guangzhou, China*
B.Eng. in Microelectronic Science and Engineering | 2+2 program with PolyU *Jun. 2017*

Professional Experience

Software Engineer Intern *May 2023 - Aug. 2023*
DreamBig Semiconductor Inc. *San Jose, CA*
Software Group | Supervisor: Mr. Faisal Masood (VP of Engineering)

- Researched critical networking services of major cloud service providers (e.g., Azure, AWS, Google Cloud) with daily sync-up meetings
- Created an internal knowledge base to document, investigate, and analyze technical details of the latest cloud networking technologies (e.g., Private Link, Network Load Balancer)
- Contributed to the research and development progress of DreamBig's novel SmartNIC/DPU design with an emphasis on SONiC-DASH features

Graduate Research Assistant *Sep. 2018 - May 2024*
Department of Electrical and Computer Engineering, NYU *Brooklyn, NY*
High-Speed Networking Lab | Supervisor: Prof. H. Jonathan Chao
(*Ph.D.*) Conducted extensive research in Machine Learning + Traffic Engineering [[Project List](#)]
(*M.S.*) Project Title: *Traffic Engineering with Reinforcement Learning* [[4](#)] 

- Designed and implemented a rule-based heuristic called *Critical Top-K* that reroutes elephant flows from the most congested links to alleviate network congestion effectively
- Conducted extensive experiments on NYU's High Performance Computing (HPC) clusters to evaluate our Reinforcement Learning-based Traffic Engineering (RL-TE) solution that learns to identify and reroute the most critical flows in the network
- Achieved close-to-optimal load balancing performance with up to 12.2% improvement over the best heuristic scheme by rerouting only **10%-21.3%** of total traffic

Research Assistant Intern *Jun. 2018 - Jul. 2018*
Huazhong University of Science and Technology *Wuhan, China*
Cyber-Physical-Social Systems Lab | Supervisor: Prof. Laurence T. Yang & Dr. Huazhong Liu
Project Title: *Network Traffic Prediction with Multivariate Multi-order Markov Models* [[5](#)]

- Constructed multivariate multi-order Markov models and implemented a novel tensor operation called *Unified Product* to perform network traffic prediction
- Utilized a Linux cloud platform to carry out network traffic prediction experiments using real-world network traffic datasets from FiberHome
- Improved the prediction accuracy by up to **38.47%** compared to the Z-eigen based approach

System Engineer Intern

Cisco Systems (China) Networking Technology Co., Ltd.
Service Provider Solution Group | Supervisor: Mr. Mark Li (System Engineer Manager)

Jul. 2016 - Aug. 2016

Guangzhou, China

- Oversaw pre-sales and post-sales technical support for Cisco products
- Performed routine maintenance at China Telecom's data center in Shenzhen
- Investigated emerging software-defined networking (SDN) applications in service provider industries for group presentation and discussion

Teaching & Mentorship

Course Instructor

Department of Electrical and Computer Engineering, NYU

Jan. 2023 - May 2023

Brooklyn, NY

Course: *ECE-GY 6363 Data Center and Cloud Computing (2023 Spring)*

- Provided instructions on course materials to **48 graduate students** enrolled in the course
- Held weekly online Zoom meetings (1 hour per week) for quizzes and Q&A sessions
- Performed course administrative tasks using the NYU Brightspace digital learning platform
- Received good instructor ratings in terms of **clear, effective, inclusive, and receptive**

Graduate Teaching Assistant (TA)

Department of Electrical and Computer Engineering, NYU

Feb. 2019 - Dec. 2023

Brooklyn, NY

Highlights: 5 years of TA experience in teaching 3 different networking courses

Courses: (1) *Data Center and Cloud Computing, 2021 Fall - 2022 Fall*; (2) *High-Speed Networks, 2019 Spring - 2021 Spring and 2023 Fall*; (3) *Communication Networks, 2020 Spring*

- Delivered 2-hour regular lectures to students regarding recent advances in TE and 5G/MEC
- Designed core course components including quizzes, lab structures, and tutorial slides
- Provided short lectures on lab materials and assisted students during lab sessions
- Graded students' exams, quizzes, and lab demos/reports with constructive feedback

Graduate Research Project Mentorship (20+ M.S. Students)

Department of Electrical and Computer Engineering, NYU

Sep. 2019 - Present

Brooklyn, NY

- Huaqiu Liu, Seeger Zou, Po Yen Chen (2024 - Present): "*Networking for Distributed Machine Learning*"
- Yilun Liang, Yuejia Tong (2023 - Present): "*Leveraging Multipath Transport for Improving Application Performance in Dynamic Networks*"
- Po Yen Chen, Zihan Lin (2023 - Present): "*Multi-constrained Distributed QoS Routing*"
- Xing Fang, Xingda Bao, Boyu Han, Tyler Lin, Xiao Xie, Senlin Xiao, Ke Zhou, Xiaocheng Zou, Yao Shen, Runyu Wang (2022 - Present): "*Design and Implementation of SD-WAN Traffic Engineering*"
- Yang Hu (2022): "*QoS-Aware Traffic Engineering with Reinforcement Learning*" [8]
- Yang Hu (2021 - 2022): "*Destination-based Routing Update Overhead Mitigation with Reinforcement Learning*" [3]
- Ke Chen, Han Wang, Shuwen Fang (2020 - 2021): "*Adaptive Forward Error Correction for Real-time Communications Based on Reinforcement Learning*" [10]
- Han Wang, Yang Li (2019 - 2020): "*Traffic Matrix Prediction with Deep Learning Models*"
- Daihui Dou (2019): "*Investigation on 5G and Multi-access Edge Computing Technologies*"

Selected Research Projects

Enabling Lookahead Routing for Scalable TE with Supervised Learning [7]

Graph Neural Networks (GNN) + Supervised Learning (SL)

Dec. 2022 - Oct. 2023

- Formulated a multi-TMs routing optimization problem to derive an optimal routing for multiple future traffic matrices with good load balancing performance
- Designed a scalable TE framework called Roracle based on GNN and SL to efficiently learn from optimal multi-TMs routing targeting future traffic matrices during offline training
- Deployed Roracle in large networks (204 nodes and 964 edges) to quickly infer good routing strategies with up to **36%** worst-case performance improvement and $\geq 71\times$ speedup

Reinforcement Learning-based TE for QoS Provisioning and Load Balancing [8]

Reinforcement Learning (RL) + Linear Programming (LP)

Mar. 2022 - Jun. 2023

- Proposed a TE solution called QoS-RL that achieves promising load balancing performance for network operators and provides good QoS (e.g., low latency) for users simultaneously
- Leveraged RL and LP to intelligently update a few destination-based forwarding entries for different priorities of traffic to reduce management overhead and mitigate service disruption
- Provided at least **95.5%** of optimal delay on average for high priority traffic while achieving above **90%** of optimal load balancing performance in most cases

Adaptive Range Routing Prediction with Graph Neural Networks [9] 📄

Graph Neural Networks (GNN) + Supervised Learning (SL)

Dec. 2021 - Present

- Proposed a learning-based TE called LARRI to directly predict a routing strategy (instead of predicting future traffic demands) for accommodating highly dynamic network traffic
- Employed SL to learn from optimal path-based range routing for future traffic scenarios and designed a scalable GNN-based framework to facilitate training and inference
- Improved the worst-case load balancing performance by up to **43.3%** under future traffic fluctuations while achieving the lowest end-to-end delay among all baseline methods

Flexible and Disturbance-aware TE with Reinforcement Learning in SDN [2, 12] 📄

Graph Neural Networks (GNN) + Reinforcement Learning (RL)

Jul. 2021 - Nov. 2022

- Proposed a new QoS metric named *network disturbance* to evaluate the negative impact of TE's flow rerouting operations on WANs (such as service disruption)
- Designed a disturbance-aware TE with GNN and RL to intelligently reroute flexible numbers of critical flows under dynamic traffic fluctuations and unexpected single link failures
- Achieved close-to-optimal performance (i.e., above 90% of optimal performance) in **99%** of network scenarios and mitigated network disturbance by up to **38.6%** in five real networks

Mitigating Routing Update Overhead for Destination-based TE [3, 13] 📄 🌐

2-stage Reinforcement Learning (RL) + Linear Programming (LP)

Jul. 2021 - Jul. 2022

- Customized a 2-stage RL approach called FlexEntry to identify critical destination-based forwarding entries for routing updates in different traffic scenarios
- Adopted LP to produce reward signals for RL and optimize traffic split ratios for the selected critical entries to control traffic distribution
- Achieved near-optimal performance in unseen traffic scenarios with at most **99.3%** of average entry update savings in six real networks

Adaptive Forward Error Correction (FEC) for Real-time Communications [10] 📄 🌐

Reinforcement Learning (RL) + Supervised Learning (SL)

May 2020 - Jun. 2022

- Designed a no-reference packet-level video quality assessment method with SL to generate real-time video quality scores as feedback information without reconstructing videos
- Leveraged RL to determine redundancy rates for video frames under dynamic packet loss and added redundant packets accordingly to improve video quality with loss recovery
- Achieved good video quality in **≥95%** cases with **≥50%** savings on additional bandwidth consumption when transmitting redundant packets over the Internet

Federated TE with Supervised Learning in Multi-region Networks [11] 📄

2-layer Graph Neural Networks (GNN) + Supervised Learning (SL)

Feb. 2021 - Nov. 2021

- Designed a 2-layer GNN architecture to model different levels of network abstractions in multi-region networks (intra-region and inter-region)
- Exploited light-weight GNN message exchanges to facilitate collaborations among regions and adopted SL to predict cross-region traffic at border routers for routing optimization
- Boosted distributed TE's performance by up to **28.9%** with low computation cost in large networks (**<1s** execution time in the BRITE network with 204 nodes and 964 links)

**Selected
Course
Projects**

Adaptive Traffic Engineering for Resilient Networks

Apr. 2020 - May 2020

Course Project of *ECE7363 Network Design and Algorithms* at NYU

- Incorporated a self-attention method to process the dependency of traffic matrix and graph-structured network topology using the Transformer architecture
- Customized a Reinforcement Learning approach to learn a critical flow selection policy against various link failures by observing the reward generated from Linear Programming
- Improved network performance with low rerouting impact while being resilient to different link failure scenarios in the Abilene network with real traffic traces

Solving LunarLander-v2 with Deep Q-Networks

Apr. 2020 - May 2020

Course Project of *ECE9243 Optimal Learning and Control for Robotics* at NYU

- Implemented the Deep Q-Network (DQN) algorithm with PyTorch to solve complex control tasks of the LunarLander-v2 environment in OpenAI Gym
- Explored different hyperparameter settings and neural network architectures to learn a good policy to successfully “land the spaceship on the lunar surface”

Age-invariant Face Recognition

Oct. 2017 - Dec. 2017

Course Project of *EL9123 Machine Learning* at NYU

- Performed face recognition based on two aging face databases including the MORPH Album 2 (78000 images of 13000 people) and the FG-NET database
- Preprocessed facial images by detecting feature points, cropping faces, and enhancing images with a median filter and histogram stretch
- Applied Principal Component Analysis (PCA) and Local Binary Patterns (LBP) to effectively extract facial features and perform aging face recognition with C++ and OpenCV
- Fine-tuned the parameters of the patch division scheme and the Weighted Chi-square distance to obtain high recognition accuracy on aging faces

Scholarships & Awards

NYU ECE PhD Scholarship

Department of Electrical and Computer Engineering, NYU

Sep. 2019 - May 2024

During my PhD studies, I received several **scholarships** from NYU to cover my tuition expenses. I was also supported by **graduate research/teaching assistantships** in the ECE department and **grants** from Fortinet, Inc., CA, USA.

- PhD Scholarship (2019-2020)
- Shiv Panwar Scholarship (2019-2020)
- Provost TA's Scholarship (2020-2024)
- SOE PhD Program Scholarship (2024)

Alexander Hessel Award for the Best Ph.D. Dissertation in Electrical Engineering

Department of Electrical and Computer Engineering, NYU

May 2024

- This award is given in memory of the late ECE Professor Alexander Hessel to a graduate student for the most outstanding doctoral dissertation in electrical engineering.
- I was honored to receive the award with a completed Ph.D. dissertation draft entitled "Leveraging Machine Learning for Intelligent Traffic Engineering Design in Wide-Area Networks" based on my scholarly contributions over the past 5 years.

The Dante Youla Award For Graduate Research Excellence in Electrical Engineering

Department of Electrical and Computer Engineering, NYU

May 2023

- This award is given in memory of the late ECE Professor Emeritus Dante Youla for timely recognition of important research contributions made by Ph.D. students. It recognizes one impactful paper in which the awardee is the principal author or a significant external award.
- I was selected as the award recipient with a recently published journal paper in the *IEEE/ACM Transactions on Networking (ToN)*, which is entitled "FlexDATE: Flexible and Disturbance-Aware Traffic Engineering With Reinforcement Learning in Software-Defined Networks". A Certificate of Merits was awarded to me at the ceremony.

IEEE ICNP 2021 Student Registration Award

Travel Grant Chairs of IEEE ICNP 2021

Oct. 2021

- The purpose of this award is to encourage students' participation in the conference by funding their conference registration. Recipients of this reward show evidence of a serious interest in networking, as demonstrated by coursework and project experience.
- As a first-time ICNP attendee, author, and presenter, I wrote a proposal to apply for the registration award and was selected as the award recipient by the travel grant chairs.

NYU ECE MS Student Academic Achievement Award

Department of Electrical and Computer Engineering, NYU

Oct. 2018

- The Academic Achievement Award is given to top MS students in the ECE department, which recognizes their outstanding academic work.
- As the award recipient, I am one of the top MS students among 200+ ECE MS students with a perfect GPA (4.0/4.0). A Certificate of Merits was awarded to me at the ceremony.

PolyU EIE Best Semester GPA Award & Best Academic Improvement Award

Department of Electronic and Information Engineering, PolyU

Mar. 2017

- The Best Semester GPA Award is given to the student with the highest semester GPA in the EIE Department at PolyU, while the Best Academic Improvement Award is given to the student with the highest GPA improvements.
- In the 2016-2017 Fall semester at PolyU, I got a 4.2/5.0 GPA with two A+ grades in EIE core elective courses, which was considered a great improvement over previous semesters. Thus, I was selected as the recipient of these two awards.

Memberships & Services

Organization Membership

- IEEE Graduate Student Member (2023 - Present)
- IEEE Communications Society (ComSoc) Graduate Student Member (2023 - Present)
- IEEE Young Professionals Member (2023 - Present)
- Science Alliance Member of the New York Academy of Science (2022 - 2023)

Reviewer for International Conferences (Total 6 Reviews)

- IEEE/ACM International Symposium on Quality of Service (IWQoS) - 2 reviews
- IEEE International Conference on Communications (ICC) - 2 reviews
- IEEE Global Communications Conference (GLOBECOM)
- International Teletraffic Congress (ITC)

Reviewer for International Journals (Total 31 Reviews)

- IEEE Journal on Selected Areas in Communications (JSAC) - 2 reviews
- IEEE Transactions on Vehicular Technology
- IEEE Systems Journal - 9 reviews
- IEEE Access - 5 reviews
- Elsevier Computer Networks - 2 reviews
- Elsevier Computers & Operations Research
- Elsevier Journal of Parallel and Distributed Computing (JPDC) - 5 reviews
- EURASIP Journal on Wireless Communications and Networking (JWCN) - 3 reviews
- Springer Telecommunication Systems - 3 reviews

Technical Skills

Programming Languages: Python, C/C++, SQL, PHP, MATLAB, Bash, LaTeX

Frameworks, Libraries & Tools: TensorFlow, PyTorch, Keras, scikit-learn, pandas, NumPy, Matplotlib, OpenCV, Gym, Git, Docker, Kubernetes, Terraform, Azure, AWS, GCP, Hadoop, DDPK, RDMA, VXLAN, OVS, D-ITG, Mininet, Wireshark, tcpdump, eBPF, NetworkX, Gurobi, vim, Linux, HPC

Technologies: Machine Learning (ML), Deep Learning (DL), Reinforcement Learning (RL), Cloud Computing, Data Center, TCP/IP, Computer Networking, Switches and Routers, Software-Defined Networking (SDN), SDWAN, MPTCP, QUIC, MPLS, Tunneling Protocols (VXLAN, GRE, GENEVE, IP-in-IP), Network Virtualization, Traffic Engineering (TE), Routing Protocols, Segment Routing, Graph Algorithms, Network Design and Optimization, Network Security, Data Structure and Algorithms, Computer Architecture, Operating Systems, Database Systems

Publications [\[Google Scholar\]](#) [\[DBLP\]](#)

*I have published **13 papers** in top networking journals and conferences with **200+ citations**.

Journal Papers

[1] (ToN 24) Yuntian Zhang, Ning Han, Tengting Zhu, Junjie Zhang, **Minghao Ye**, Songshi Dou, and Zehua Guo, "Prophet: Traffic Engineering-centric Traffic Matrix Prediction," *IEEE/ACM Transactions on Networking (ToN)*, 2024. [\[URL\]](#) [\[PDF\]](#)

[2] (ToN 23) **Minghao Ye**, Junjie Zhang, Zehua Guo, and H. Jonathan Chao, "FlexDATE: Flexible and Disturbance-Aware Traffic Engineering with Reinforcement Learning in Software-Defined Networks," *IEEE/ACM Transactions on Networking (ToN)*, 2023. [\[URL\]](#) [\[PDF\]](#)

[3] (JSAC 22) **Minghao Ye**, Yang Hu, Junjie Zhang, Zehua Guo, and H. Jonathan Chao, "Mitigating Routing Update Overhead for Traffic Engineering by Combining Destination-based Routing with Reinforcement Learning," *IEEE Journal on Selected Areas in Communications (JSAC)*, 2022. (Impact factor: 16.4) [\[URL\]](#) [\[Codes\]](#) [\[PDF\]](#)

[4] (JSAC 20) Junjie Zhang, **Minghao Ye**, Zehua Guo, Chen-Yu Yen, and H. Jonathan Chao, "CFR-RL: Traffic Engineering with Reinforcement Learning in SDN," *IEEE Journal on Selected Areas in Communications (JSAC)*, 2020. (100+ citations on Google Scholar. Impact factor: 16.4) [\[URL\]](#) [\[Codes\]](#) [\[PDF\]](#) [\[arXiv\]](#)

[5] (TNSM 19) Huazhong Liu, Laurence T. Yang, Jinjun Chen, **Minghao Ye**, Jihong Ding, and Liwei Kuang, "Multivariate Multi-order Markov Multi-modal Prediction with Its Application in Network Traffic Management," *IEEE Transactions on Network and Service Management (TNSM)*, 2019. [\[URL\]](#) [\[PDF\]](#)

Conference and Workshop Papers

- [6] (IWQoS '24) Haowen Zhu, Minghao Ye, and Zehua Guo, "Toward Determined Service for Distributed Machine Learning," Accepted. *The 32nd IEEE/ACM International Symposium on Quality of Service (IWQoS)*, 2024. (Acceptance rate: 24.8%, 81/326)
- [7] (ICNP '23) Minghao Ye, Junjie Zhang, Zehua Guo, and H. Jonathan Chao, "Roracle: Enabling Lookahead Routing for Scalable Traffic Engineering with Supervised Learning," *The 31st IEEE International Conference on Network Protocols (ICNP)*, 2023. (Acceptance rate: 18.8%, 34/181) [URL] [PDF]
- [8] (IWQoS '23) Minghao Ye, Yang Hu (co-first author), Junjie Zhang, Zehua Guo, and H. Jonathan Chao, "Reinforcement Learning-based Traffic Engineering for QoS Provisioning and Load Balancing," *The 31st IEEE/ACM International Symposium on Quality of Service (IWQoS)*, 2023. (Acceptance rate: 23.5%, 62/264) [URL] [PDF]
- [9] (INFOCOM '23) Minghao Ye, Junjie Zhang, Zehua Guo, and H. Jonathan Chao, "LARRI: Learning-based Adaptive Range Routing for Highly Dynamic Traffic in WANs," *IEEE International Conference on Computer Communications (INFOCOM)*, 2023. (Selected as **one of the five fast-tracked papers out of 252 accepted papers** for submission to a top-tier journal: *IEEE/ACM Transactions on Networking*. Acceptance rate: 19.2%, 252/1312) [URL] [PDF]
- [10] (MMSys '22) Ke Chen, Han Wang, Shuwen Fang, Xiaotian Li, Minghao Ye, and H. Jonathan Chao, "RL-AFEC: Adaptive Forward Error Correction for Real-time Video Communication Based on Reinforcement Learning," *The 13th ACM Multimedia Systems Conference (MMSys)*, 2022. [URL] [Codes] [PDF]
- [11] (ICNP '21) Minghao Ye, Junjie Zhang, Zehua Guo, and H. Jonathan Chao, "Federated Traffic Engineering with Supervised Learning in Multi-region Networks," *The 29th IEEE International Conference on Network Protocols (ICNP)*, 2021. (**One of the 11 pre-accepted papers without conditional acceptance**. Acceptance rate: 24.7%, 38/154) [URL] [PDF]
- [12] (IWQoS '21) Minghao Ye, Junjie Zhang, Zehua Guo, and H. Jonathan Chao, "DATE: Disturbance-Aware Traffic Engineering with Reinforcement Learning in Software-Defined Networks," *The 29th IEEE/ACM International Symposium on Quality of Service (IWQoS)*, 2021. (Acceptance rate: 25%, 64/256) [URL] [PDF]
- [13] (NetAI '20) Junjie Zhang, Zehua Guo, Minghao Ye, and H. Jonathan Chao, "SmartEntry: Mitigating Routing Update Overhead with Reinforcement Learning for Traffic Engineering," *ACM SIGCOMM Workshop on Network Meets AI & ML (NetAI)*, 2020. [URL] [Slides] [PDF]

Presentations

1. "Reinforcement Learning-based Traffic Engineering for QoS Provisioning and Load Balancing," *The 31st IEEE/ACM International Symposium on Quality of Service (IWQoS)*, Orlando, FL, USA, Jun. 19, 2023.
2. "LARRI: Learning-based Adaptive Range Routing for Highly Dynamic Traffic in WANs," *IEEE International Conference on Computer Communications (INFOCOM)*, Hoboken, NJ, USA, May 17, 2023.
3. "Federated Traffic Engineering with Supervised Learning in Multi-region Networks," *The 29th IEEE International Conference on Network Protocols (ICNP)*, Online, Nov. 2, 2021. [Recording] [Teaser]
4. "DATE: Disturbance-Aware Traffic Engineering with Reinforcement Learning in Software-Defined Networks," *The 29th IEEE/ACM International Symposium on Quality of Service (IWQoS)*, Online, Jun. 27, 2021. [Recording]
5. "SmartEntry: Mitigating Routing Update Overhead with Reinforcement Learning for Traffic Engineering," *ACM SIGCOMM Workshop on Network Meets AI & ML (NetAI)*, Online, Aug. 10, 2020. [Recording]