Jinrui (Jerry) Gou

Computer Science and Engineering
Tandon School of Engineering
New York University

Tel : (347) 200 5159
Tel : +86 189 8325 3062
Email: jg6226@nyu.edu

EDUCATION

New York University — Brooklyn, NY
Ph.D. in Computer Science — 09/2021-Present
Advisor: Professor Torsten Suel
Overall GPA: 3.97/4.0

M.S. in Computer Science — 09/2019-05/2021
Overall GPA: 4.0/4.0

Sun Yat-Sen University — Guangzhou, China
B.E. in Computer Science & Technology — 09/2015-06/2019
Overall GPA: 90/100 (3.8/4.0)
Mathematics Courses GPA: 95/100

KNOWLEDGE

- **Languages:** C/C++, Python, SQL, AMPL/CPLEX, Latex, MATLAB, Bash, HTML5, JavaScript
- **Tools:** Scikit-learn, PyTorch, TensorFlow, Numpy, Matplotlib, TBB, AVX/AVX2, CUDA, OpenCL
- **Courses Enrolled:** Computer Architecture, Natural Language Processing, Machine Learning (Rank 1^{st} in Class of 43 students), Deep Learning, Web Search Engines, Design & Analysis of Algorithms (Rank 1^{st} in Class of 47 students), Network Design and Algorithms (Rank 2^{nd} in Class of 20 students), Algorithmic Machine Learning and Data Science

RESEARCH EXPERIENCE

Ph.D. Research Assistant for Prof. Torsten Suel, NYU — 09/2021-Present

1. Effective Candidate Generation and Threshold Estimation for Fast Top-K Query Processing
   - Build specialized first-layer index structures that storing high-scoring postings
   - Fine-tune the storage of index structures and fit them into main memory for fast access
   - Reduce cache line miss by utilizing bloom filter to speed up lookups
   - Implement the entire process with 2000+ lines C++ code to to provide high quality estimate of top-k threshold and promissing candidates documents
2. Graph Based Approximate Near-neighbor Search
   - Propose navigable graph that can preform greedy routing for nearest neighbor search
   - Design randomized algorithms of adding links between nodes in high dimension space to create effective navigable graphs
   - Explore and bound the max degree and average degree of navigable graphs with proof

- Write efficient code to compute distances between high dimensional points and create routing tree

3. Sparse Learned Index Structures Optimization
    - Expand documents by generating related questions, using LLM Llama 2
    - Improve document expansion quality by filtering unrelated questions
    - Train BERT model CoCondenser to predict score of terms in documents for better top-k query processing
    - Fuse the knowledge from cross-encoder to improve training by hard negatives and distillation

Graduate Research Assistant for Prof. Yong Liu, Networked Systems Lab, NYU 09/2019-05/2021

1. Realtime Mobile Network QoS Prediction:
    - Collect mobile 4G/LTE bandwidth traces in Metro Area of NYC
    - Develop models for realtime bandwidth prediction and network handoffs prediction
2. Joint Traffic Routing and Computation Server Placement in Edge Cloud Networks:
    - Develop quasi-convex optimization models to obtain optimal traffic routing and computation resource allocation to simultaneously balance the traffic and computation load distribution in edge cloud networks
    - Use AMPL/CPLEX and Python optimization packages to obtain numerical solutions for comparative case studies

PUBLICATIONS

1. Lifan Mei, **Jinrui Gou**, Yujin Cai, Houwei Cao, and Yong Liu, "Realtime mobile bandwidth and handoff predictions in 4G/5G networks", *Computer Networks, Feb 2022*.
2. Soyuj Basnet, **Jinrui Gou**, Antonio Mallia, and Torsten Suel, "DeeperImpact: Optimizing Sparse Learned Index Structures", *in ReNeuIR at SIGIR 2024: The Third Workshop on Reaching Efficiency in Neural Information Retrieval.*
3. Lifan Mei, **Jinrui Gou**, Jingrui Yang, Yujin Cai, and Yong Liu, "On Routing Optimization in Networks with Embedded Computational Services", *IEEE Transactions on Network and Service Management, accepted 2024.*
4. **Jinrui Gou**, Yifan Liu, Minghao Shao, and Torsten Suel, "Beyond Quantile Methods: Improved Top-K Threshold Estimation for Traditional and Learned Sparse Indexes", *submitted to IEEE International Conference on Big Data 2024.*
5. Haya Diwan, **Jinrui Gou**, Cameron Musco, Christopher Musco, and Torsten Suel, "Navigable Graphs for High-Dimensional Nearest Neighbor Search: Constructions and Limits", *submitted to NeurIPS 2024: Conference on Neural Information Processing Systems, under Review.*

TEACHING ASSISTANT EXPERIENCE

Teaching Assistant, ECE-GY 9343 Data Structures and Algorithms, NYU Spring, Summer, Fall 2020

- Create homework questions, and grade homework submissions and exams;
- Help the professor answer students' questions on Piazza, and hold online TA office hours.