

Jinrui (Jerry) Gou

Computer Science and Engineering
Tandon School of Engineering
New York University

Tel : (347) 200 5159
Email: jg6226@nyu.edu
HomePage: j9rrygou.github.io

RESEARCH INTERESTS

Web Search Efficiency, Information Retrieval, Nearest Neighbor Search

EDUCATION		
	New York University Ph.D. in Computer Science Advisor: Professor Torsten Suel Cumulative GPA: 3.97/4.0	Brooklyn, NY 09/2021-Present
	M.S. in Computer Science Cumulative GPA: 4.0/4.0	09/2019-05/2021
	Sun Yat-Sen University B.E. in Computer Science & Technology Cumulative GPA: 90/100 (3.8/4.0) Mathematics Courses GPA: 95/100	Guangzhou, China 09/2015-06/2019

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
 - Develop efficient and effective early termination algorithms for simple ranking functions.
 - Design specialized prefixes index structures and fine-tune its storage for efficient access.
 - Perform fast lookups using highly optimized index structures to improve the quality of results.
 - Implement efficient C++ code for experiments on different collections and ranking functions.
2. Graph Based Approximate Near-neighbor Search
 - Study the theoretical limits of constructing sparse navigable graphs for nearest neighbor search in high-dimensional spaces.
 - Design efficient randomized algorithms to create effective sparse navigable graphs.
 - Bound the maximum degree and average degree of navigable graphs using mathematical proof.
 - Develop efficient heuristic graph pruning algorithms to reduce the average degree of a graph while maintaining navigability.
 - Propose Distance Adaptive Beam Search, with a simple new termination condition that replaces the standard fixed beam width.
 - Prove guaranteed accuracy on navigable graphs and demonstrate consistent performance improvements on real-world graphs (HNSW, Vamana).

3. Sparse Learned Index Structures Optimization

- Assist dataset preparation and experiment reproducibility.
- Conduct evaluation experiments using PISA search engine toolkit.
- Contribute to preparing project presentation materials.

Graduate Research Assistant for Prof. Yong Liu, Networked Systems Lab, NYU

09/2019-05/2021

1. Realtime Mobile Network Prediction for High QoS

- 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

- Build quasi-convex optimization model to balance traffic and computation resource allocation.
- Use AMPL/CPLEX and Python optimization packages to obtain optimal feasible solutions.

PUBLICATIONS

1. Yousef Al-Jazzazi*, Haya Diwan*, **Jinrui Gou***, Cameron Musco*, Christopher Musco*, Torsten Suel*, “Distance Adaptive Beam Search for Provably Accurate Graph-Based Nearest Neighbor Search”, *NeurIPS 2025: Advances in Neural Information Processing Systems, December 2025*¹
2. **Jinrui Gou**, Antonio Mallia, Yifan Liu, Minghao Shao, and Torsten Suel, “Fast and Effective Early Termination for Simple Ranking Functions”, *ACM SIGIR Conference on Research and Development in Information Retrieval, July 2025*.
3. Haya Diwan*, **Jinrui Gou***, Cameron Musco*, Christopher Musco*, and Torsten Suel*, “Navigable Graphs for High-Dimensional Nearest Neighbor Search: Constructions and Limits”, *NeurIPS 2024: Conference on Neural Information Processing Systems, December 2024*.²
4. **Jinrui Gou**, Yifan Liu, Minghao Shao, and Torsten Suel, “Beyond Quantile Methods: Improved Top-K Threshold Estimation for Traditional and Learned Sparse Indexes”, *IEEE International Conference on Big Data, December 2024*.
5. 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, July 2024*.
6. 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, September 2024*.
7. Lifan Mei, **Jinrui Gou**, Yujin Cai, Houwei Cao, and Yong Liu, “Realtime mobile bandwidth and handoff predictions in 4G/5G networks”, *Computer Networks, February 2022*.

¹All authors contributed equally.

²All authors contributed equally.

EMPLOYMENT EXPERIENCE

Intern, Cantor Fitzgerald, New York

Summer 2025

- Process and analyze high-frequency equity market data to support short-term trading analysis.
- Build data pipelines in kdb+/q for efficient preprocessing and generation of financial indicators.
- Integrate proprietary transaction data to enhance the quality and coverage of internal market signals.

Course Assistant, CS-GY 6913 Web Search Engines, NYU

Fall 2024, Fall 2025

- Conduct demo sessions and grade students' assignments and exams.
- Assist the professor by answering students' questions on Piazza and holding weekly TA office hours.

Course Assistant, CS-GY 6083 Principles of Database Systems, NYU

Spring 2025

- Design homework problems, conduct demo sessions, and grade students' assignments and exams.
- Assist the professor by answering students' questions on Piazza and holding weekly TA office hours.

Teaching Assistant, ECE-GY 9343 Data Structures and Algorithms, NYU

Spring, Summer, Fall 2020

- Design homework problems and grade students' assignments and exams.
- Assist the professor by answering students' questions on Piazza and holding weekly TA office hours.

ADDITIONAL INFORMATION

- **Computer Skills:** Python, C++, MySQL, AMPL/CPLEX
- **Languages:** English (Fluent), Mandarin Chinese (Native)

REFERENCES

Prof. Torsten Suel

NYU Tandon School of Engineering
torsten.suel@nyu.edu

Prof. Christopher Musco

NYU Tandon School of Engineering
cmusco@nyu.edu

Prof. Yong Liu

NYU Tandon School of Engineering
yongliu@nyu.edu