




# Yifeng Xiao

 [github.com/yifengxiao1](https://github.com/yifengxiao1)  [yifengxiao1.github.io](https://yifengxiao1.github.io)  [yifengxiao.xjtu@gmail.com](mailto:yifengxiao.xjtu@gmail.com)

## EDUCATION

**Xi'an Jiaotong University, Xi'an China**

*Master of Science - Mathematics*

Sep 2021 - June 2024

*Current average score: 85.86*

**Xi'an Jiaotong University, Xi'an China**

*Bachelor of Science - Honors Science Program (Mathematics and applied Mathematics)*

Sep 2017 - June 2022

*Average score: 88.55*

**Georgia Institute of Technology**

*Intensive English Program*

Jan 2020 - May 2020

*GPA: 3.95; Statistical Estimation: A; Information Theory: A*

## INTERNSHIP AND TEACHING EXPERIENCE

**Shenzhen Research Institute of Big Data, (SRIBD)**

*Research Assistant*

May 2023 - Present

**Xi'an Jiaotong University (XJTU)**

Teaching Assistant: Calculus (XJTU courses for undergraduate students).

Sep 2021 - Feb 2022

## SKILLS

**Coding:** Python (3+ yrs), Matlab (1+ yrs),  $\text{\LaTeX}$  (2+ yrs), Linux (6 mth), HTML/CSS (3 mth)

**Fields:** Machine learning (3+ yrs), Applied mathematics (3+ yrs), Operations research (2+ yrs), Wireless communication (1+ yrs)

**English:** Toefl: 100, GRE: 320 + 3.5

## RESEARCH EXPERIENCE

**Optimization Solver Development Lab, SRIBD**

May 2023 - Nov 2023

- 2-stage Stochastic Programming

Used graph neural networks (GNNs) to approximate the optimal objective of second-stage programming under a given first-stage solution, and made this approximation generalizable between instances with a problem. The approximation networks were used for following step by 1) fixing the networks and directly solving the nonlinear optimization problems 2) or embedding the networks into mixed-integer linear programming problems (MILP).

- Symmetry in Mixed-Integer Linear Programming

Proposed a pre-processing approach to handle the symmetry in MILP when using GNNs to predict solutions. Before using GNNs, use a weak learning model for this task and adjust the dataset labels (solutions) by symmetric permutation meanwhile. After this, datasets become more smooth and suitable for later training of GNNs.

**National Engineering Laboratory for Big Data Analytics, XJTU**

Sep 2020 - June 2024

- Distributed Algorithms for Computing and Clustering Analysis

Proposed a distributed clustering algorithm based on learning to hash technique. Specifically, distributedly training a hashing network unsupervisedly and mapping massive high-dimensional data to a small amount of fixed-length hash codes. The hash codes are transmitted to a global site and clustered by a spectral method.

- Beamforming in Full-Duplex Wireless Communication System

Used stochastic successive convex approximation to suppress interferences while maintaining power in a target direction.

- Semantic Communication System Based on Transformer

Achieved semantic joint encoding and decoding in the physical layer of communication system based on transformer network.

- A Numerical Approach for eigenvalues of high-oscillating integral equations

Proposed an approach: factorizing the integral equations on the modified Fourier basis so that the analytical problem can be converted into an algebraic one. Then, find a surrogate approximation by the property of high-oscillating and solve the surrogate problem.

## PUBLICATIONS

**Hashing-Based Distributed Clustering for Massive High-Dimensional Data** IEEE Transactions on Big Data

- Status: Undergoing review

arXiv: <https://arxiv.org/abs/2306.17417>