XIANGYU ZHAO

Room 906, EEE Building, Imperial College London, Exhibition Road, London SW7 2AZ

Personal Website: `https://victorzxy.github.io/`

Email: `x.zhao22@imperial.ac.uk` ⋄ GitHub: VictorZXY

EDUCATION

PhD in Electrical and Electronic Engineering *January 2023 – January 2027 (Expected)*
Imperial College London, United Kingdom
Research topic: Representation Learning on Higher-Order Graphs
Award: Electrical and Electronic Engineering PhD Scholarship

BA & MEng (Hons) in Computer Science *October 2018 – June 2022*
Trinity College, University of Cambridge, United Kingdom
BA final year result: 2.i (Overall: 70.5%; Units: 86%; Dissertation: 81%)
MEng result: Distinction (Overall: 81.2%; Modules: 79.7%; Dissertation: 83%; Ranking: 7/25)
Awards: Senior Scholarship; Exam Prize 2022

WORK EXPERIENCE

Department of Computer Science and Technology, University of Cambridge *June – September 2022*
Undergraduate Research Opportunities Programme (UROP) – Research Intern *Cambridge, United Kingdom*

· Research Project: Long and Short-Range Attentions for Complex Hypergraph Data;
· Supervisors: Dr Yiren Zhao, Prof Robert Mullins.

ByteDance *June – September 2021*
Algorithm Engineer Intern *Beijing, China*

· Trained a model based on causal forest and double ML for TikTok Lite's user growth campaign;
· Increased per-user successful invitation count by 43.9% while decreasing its cost by 56.1%.

PUBLICATIONS

· **Xiangyu Zhao**, Zehui Li, Mingzhu Shen, Guy-Bart Stan, Pietro Liò, and Yiren Zhao (2024). Enhancing Node Representations for Real-World Complex Networks with Topological Augmentation. In *The 27th European Conference on Artificial Intelligence (ECAI 2024)*, Frontiers in Artificial Intelligence and Applications. IOS Press.

· Xiandong Zou, **Xiangyu Zhao**, Pietro Liò, and Yiren Zhao (2023). Will More Expressive Graph Neural Networks do Better on Generative Tasks? In *Proceedings of the 2nd Learning on Graphs Conference (LoG 2023)*, volume 231, pages 21:1–21:26. PMLR.

· **Xiangyu Zhao**, Zehui Li, Mingzhu Shen, Guy-Bart Stan, Pietro Liò, and Yiren Zhao (2023). Hybrid Graph: A Unified Graph Representation with Datasets and Benchmarks for Complex Graphs. *arXiv preprint arXiv:2306.05108*.

· **Xiangyu Zhao**, Hannes Stärk, Dominique Beaini, Yiren Zhao, and Pietro Liò (2023). Task-Agnostic Graph Neural Network Evaluation via Adversarial Collaboration. *The 11th International Conference on Learning Representations (ICLR 2023) Machine Learning for Drug Discovery (MLDD) Workshop*.

· **Xiangyu Zhao** and Sean B. Holden (2022). Towards a Competitive 3-Player Mahjong AI using Deep Reinforcement Learning. In *2022 IEEE Conference on Games (CoG)*.

· **Xiangyu Zhao** and Sean B. Holden (2022). Building a 3-Player Mahjong AI using Deep Reinforcement Learning. *arXiv preprint arXiv:2202.12847*.

ACADEMIC SERVICE

Reviewer KDD 2024

SKILLS

Languages	Chinese (Native), English (Fluent, IELTS Band 8.0)
Programming	Python, C, C++, C#, Java, OCaml, Standard ML, Prolog, SQL, LaTeX
ML Libraries	PyTorch, PyTorch Geometric, TensorFlow, Keras
Music	Clarinet (Grade 10, soloist of the Cambridge University Chinese Orchestra Society)
Sports	Kendo (1st Dan, University team), Football (High school team), Badminton (High school team)

AWARDS AND PRIZES

Imperial College London EEE PhD Scholarship	*2023–2027*	Total amount >£150,000
Trinity College Senior Scholarship	*2022*	Distinction in MEng examination
Trinity College Exam Prize	*2022*	Distinction in MEng examination
Google Hash Code Competition	*2020*	Team ranked global top 10%
LeetCode Programming Contest (LCCUP)	*2020*	Team ranked global top 5%
UK & Ireland Programming Contest (UKIEPC)	*2019*	Team ranked regional top 25%
British Informatics Olympiad (BIO)	*2017*	National Finalist (UK top 15, raw mark 96%)
British Physics Olympiad (BPhO)	*2017*	Gold Medal (UK top 5%)
British Chemistry Olympiad (BChO)	*2017*	Gold Medal (UK top 5%)
British Mathematics Olympiad (BMO)	*2016, 2017*	Distinction (UK top 5%)

TEACHING EXPERIENCE

Department of Electrical and Electronic Engineering, Imperial College London   *January 2023 – Present*
Teaching Assistant & Project Supervisor   *London, United Kingdom*

· Course: ELEC70109 Advanced Deep Learning Systems (2023–2024);
· Projects: Investigating GNN Expressiveness in Graph Generative Tasks (UROP Project 2023),
  Neural Weather Forecasting using High-Order Graphs (MEng Final-Year Project 2023–2024).

Department of Computer Science and Technology, University of Cambridge   *January 2023 – Present*
Teaching Assistant, Course Supervisor & Project Supervisor   *Cambridge, United Kingdom*

· Courses: Algorithms 1 (2022–2024), Algorithms 2 (2023–2024), Foundations of Computer Science (2023–2024),
  Introduction to Probability (2022–2024);
· Project: Improving 3D Graph Generative Models with More Expressive GNNs (MPhil Project 2024–2025).

PROJECT EXPERIENCE

Enhancing Real-World Complex Network Representations with Topological Augmentation   *2024*

· Introduce a novel graph augmentation method incorporating higher-order node relations for complex networks;
· Develop 23 novel real-world graph datasets across various domains to facilitate evaluation;
· Our method consistently and significantly outperforms GNN baselines and other graph augmentation methods.
· Paper published at the *27th European Conference on Artificial Intelligence (ECAI 2024)*.
· Source code: `https://github.com/VictorZXY/TopoAug`
  Publication: `https://arxiv.org/abs/2402.13033`

Investigating GNN Expressiveness in Graph Generation Tasks   *2023*

· Improve GNN-based graph generative models with more advanced GNNs;
· Show that GNN expressiveness in graph prediction does not correlate to its performance in graph generation.
· Paper published at the *2nd Learning on Graphs Conference (LoG 2023)*.
· Source code: `https://github.com/Yqcca/graph-generative-models`
  Publication: `https://openreview.net/pdf?id=aBL9SfWVJb`

Unifying Higher-Order Graph Representation with Real-World Datasets and Benchmarks   *2023*

· Introduce the concept of hybrid graphs, a unified definition for higher-order graphs;
· Build a collection of real-world hybrid graph datasets with full utility support (e.g. graph samplers), and an
  extensible evaluation framework;
· Develop a hybrid graph learning baseline outperforming purely simple and higher-order graph learning methods.
· Project homepage: `https://zehui127.github.io/hybrid-graph-benchmark/`
  Publication: `https://arxiv.org/abs/2306.05108`

Task-Agnostic Graph Neural Network Evaluation via Adversarial Collaboration   *2022*
MEng Dissertation

· Develop a conceptually novel, principled, task-agnostic, and stable framework for evaluating GNNs through con-
  trastive self-supervision, without needing handcrafted augmentations;
· Distinguish GNNs of different expressiveness by having them directly compete against each other on unlabelled
  graphs, and ensure that more expressive GNNs can win by producing more informative graph embeddings.
· Paper published at the *ICLR 2023 Machine Learning for Drug Discovery (MLDD) Workshop*.
· Source code: `https://github.com/VictorZXY/GraphAC`
  Publication: `https://openreview.net/pdf?id=NXfiEdmA1t`

Multi-Agent Deep Q-Learning for the Berry Poisoning Game 2022

· Perform a multi-agent DQN method on the Berry Poisoning Games, and investigate on the agent performance with respect to different game environment parameters;
· Successfully train agents for the game, with high transferability across different game environment parameters.
· Source code: `https://github.com/VictorZXY/dqn-berry-poisoning`

Function Autoencoders: A Neural Network Approach to Gaussian Processes 2022

· Investigate a neural network alternative to Gaussian processes, and introduce the function autoencoders that preserve GPs' own advantages and avoid their weaknesses with NNs' benefits;
· Models successfully learn distributions over random functions, and perform decently on a 1-d regression task.
· Source code: `https://github.com/VictorZXY/function-autoencoder`

Building a Simulator and Emulator for Traffic Signaling 2022

· Carry out simulation of an urban traffic signalling system, and build an emulator to search for optimal traffic signal scheduling using Bayesian optimisation;
· Introduce multiple scheduling schemes to optimise search space without sacrificing flexibility or descriptiveness.
· Source code: `https://github.com/VictorZXY/traffic-sim`

A Neural Network Approach to Named Entity Recognition on Noisy User-Generated Texts 2021

· Investigate a bidirectional LSTM structure for named-entity recognition on social media texts, and explore various data-processing techniques in order to improve the model's performance;
· Models trained with data-processing techniques applied achieve significant improvements on prediction accuracy on the W-NUT 2017 shared task, compared to the same model trained without any data optimisation.
· Source code: `https://github.com/VictorZXY/nlp-assignments-ner`

Deep Reinforcement Learning for Mahjong 2021
BA Dissertation

· Design an informative and compact data structure for encoding observable Mahjong states, build an AI for Mahjong by pre-training CNN models for each action, and improve the major action's model with self-play RL;
· Models achieve near state-of-the-art level, and RL significantly improve the agent's win rate from SL;
· Paper published at the *2022 IEEE Conference on Games (CoG 2022)*.
· Source code: `https://github.com/VictorZXY/meowjong`
  Publication: `https://ieeexplore.ieee.org/document/9893576`

Investigating Adversarial Examples for Deep Residual Networks 2021

· Investigate on targeted/black-box adversarial attacks on ResNets using one-shot/iterative FGSM algorithms, and experiment the transferability of each FGSM attack to other image inputs and neural networks.
· Source code: `https://github.com/VictorZXY/dnn-assignments`

An Analysis of Clickstream Data for Online Shopping 2020

· Analyse a clickstream dataset on an online shopping website, train various ML models to predict the customers' potential willingness to pay a premium price, and identify the most important features for ML prediction.
· Source code: `https://github.com/VictorZXY/datasci-pnp-final-practical`

Trading Assistant for IMC 2020

· This is a web service which holds market data and responds to queries on demand in a human-like manner, by automating the trader side of the process using speech recognition and NLP;
· Responsible for writing speech-to-text and text-to-speech libraries;
· Efficiently built the libraries with high accuracies, and provided clean interfaces that were used conveniently by the NLP and back-end teammates, improving the system's overall performance.