Binhao Wang

(+86) 180-5082-7575 | binhawang2-c@my.cityu.edu.hk | github.com/wbhtcl2019

Shenzhen / Hong Kong

EDUCATION

City University of Hong Kong   *Expected Start Jan. 2026*
Ph.D. in Data Science, College of Computing

City University of Hong Kong   *Sep. 2024 – Expected Oct. 2025*
M.Sc. in Data Science, College of Computing   **GPA: 3.99/4.0** (Distinction)

Xiamen University   *Sep. 2019 – Jun. 2023*
B.Sc. in Mathematical Statistics, WISE   **GPA: 3.52/4.0** (Top 10%)
B.Sc. in Chemistry, College of Chemistry and Chemical Engineering   GPA: 3.46/4.0

HONORS, AWARDS, AND SERVICE

Conference Reviewer   *2025 – Present*

- ACM KDD 2026 (Datasets and Benchmarks Track)
- ACM CIKM 2025 (Short Papers)

Contemporary Undergraduate Mathematical Contest in Modeling (CUMCM)   *Sep. 2021*
Second Prize, Fujian Provincial Level

China National College Student Mathematics Competition   *Dec. 2020*
Third Prize, Fujian Provincial Level

PUBLICATIONS

Selected Conference Publications

- **FindRec: Stein-Guided Entropic Flow for Multi-Modal Sequential Recommendation**
  Maolin Wang[*], Yutian Xiao[*], **Binhao Wang**[*], Sheng Zhang, Shanshan Ye, Wanyu Wang[†], Hongzhi Yin, Ruocheng Guo, and Zenglin Xu.
  Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data mining (KDD '25).
  [DOI] (CCF-A)

  To address distribution inconsistency in multi-modal recommendation, we proposed (1) a theoretically-grounded **Stein-guided entropic flow** module for robust feature alignment; and (2) a cross-modal **expert routing mechanism** for adaptive signal fusion. The framework achieves a **2-4% relative performance lift** over SOTA models while **reducing inference overhead by 15%**.

- **SPARK: Adaptive Low-Rank Knowledge Graph Modeling in Hybrid Geometric Spaces for Recommendation**
  Binhao Wang, Yutian Xiao, Maolin Wang, Zhiqi Li, Tianshuo Wei, Ruocheng Guo, and Xiangyu Zhao[†].
  ACM International Conference on Information and Knowledge Management (CIKM '25), Accepted. (CCF-B)

  To address challenges in knowledge-aware recommendation, **I developed** (1) a novel **hybrid geometric GNN** leveraging **Hyperbolic spaces** to model noisy hierarchical structures; and (2) an item **popularity-aware adaptive fusion** strategy to specifically target the long-tail problem. Our approach yields a **5-12% relative improvement** in long-tail recommendation accuracy.

Manuscripts under Review

- **DDL: Collaborative Dual-Size Large Language Models with Dual-Stage Deferral Risk Control**
  Maolin Wang, **Binhao Wang**, et al. *(Under review at NeurIPS '25).*

- **SONATA: Synergistic Coreset Informed Adaptive Temporal Tensor Factorization**
  Maolin Wang, Zhiqi Li, **Binhao Wang**, et al. *(Under review at NeurIPS '25).*

- **DynaNS: Dynamic Negative Sampling Strategy for Multi-modal Recommendation**
  Anonymous Author(s) *(Under review at KDD '26, Research Track).*

- **Enhancing Sequential Recommendation with DTW-based Coreset Clustering**
  Anonymous Author(s) *(Under review at AAAI '26).*

- **MARS: Modality-Aligned Retrieval for Sequence Augmented CTR Prediction**
  Yutian Xiao[*], Shukuan Wang[*], **Binhao Wang**, et al. *(Under review at KDD '26, Applied Data Science Track).*

- **CuFun: Stable Temporal Point Processes via Cumulative Distribution Function Modeling**
  Maolin Wang, **Binhao Wang**, et al. *(Under review at IEEE Transactions on Knowledge and Data Engineering (TKDE)).*

RESEARCH & PROJECT EXPERIENCE

Knowledge Graph-based Chemical Reaction Path Finding (Undergraduate Thesis) *Nov. 2022 – May. 2023*

- Automated semantic entity extraction from **1,000+ scientific papers** by designing and implementing a data pipeline with the **GPT API** to construct a structured knowledge graph for reaction path analysis.
- Improved reaction pathway search efficiency by **20%** over keyword-based baselines by developing an intelligent retrieval system leveraging **graph algorithms** and **LLM embeddings**.
- **Tech Stack:** Python, Django, **GPT API**, **MongoDB**, NetworkX.

Application of HSIC for Non-linear Dependency Analysis (Independent Research) *Feb. 2022 – May. 2022*

- Applied the **Hilbert-Schmidt Independence Criterion (HSIC)** as a robust non-parametric test to assess the goodness-of-fit for complex time-series models.
- Successfully demonstrated HSIC's capability to capture complex, **non-linear dependencies** in model residuals, a critical task where traditional linear methods like **Pearson correlation** fail.
- **Core Concepts: HSIC**, Non-parametric Statistics, **R**.

TECHNICAL SKILLS

Programming: Python (**PyTorch, Recbole**, Transformers, Scikit-learn, Pandas), R, SQL, C/C++.

Machine Learning:

- *Recommender Systems:* Sequential Rec, Multi-Modal Rec, KG-aware Rec, Temporal Point Processes.
- *Natural Language Processing:* Foundational Models (BERT, GPT), RAG, Fine-tuning (LoRA).
- *Graph Learning:* GNNs (GCN, GAT, GraphSAGE), Deep Graph Library (DGL).
- *Core Techniques:* Contrastive Learning, Multi-modal Fusion, GNNs.

Languages: English (IELTS 6.5).