



Bingxiang He

✉ hbx20@mails.tsinghua.edu.cn · 📍 Homepage: <https://hbx-hbx.github.io>

☎ (+86) 134-1598-2735 · 🏢 Dept. of Computer Science, Tsinghua University, China

🎓 EDUCATION

Tsinghua University, Beijing 09, 2020 – 06, 2024

Undergraduate Dept. of Computer Science, graduate on 06, 2024

Tsinghua University, Beijing 09, 2024 – present

PhD Candidate Dept. of Computer Science, graduate on 06, 2029

Main courses and academic performance

GPA: 3.97/4.0, **Rank:** 4/181

A+ : English for Academic Purposes (A): Spoken Communication, Probability and Statistics, Physics for Scientists and Engineers B(2), **Artificial Neural Networks**, Diploma Project(Thesis)

A : Fundamentals of Computer Science, **Fundamentals of Programming, Programing and Training, Foundation of Object-Oriented Programming, Assembly Language Programming, Calculus A(1), Calculus A(2), Linear Algebra, Advanced Topics in Linear Algebra, Discrete Mathematics(1)**, Physics for Scientists and Engineers B(1), Introduction to Artificial Intelligence, Software Engineering, **Digital Logic Circuit, Digital Logic Experimentation**, Fundamentals of Computer Graphics, Introduction to Modern Cryptography, **Principles of Signal Processing, Principles and Practice of Compiler Construction**, Introduction to High Performance Computing, **Operating Systems**, Database Special Topic Training, **Cybersecurity Fundamentals**

A- : Introduction to Complex Analysis, **Data Structures, Numerical Analysis, Computer Network Security Technology, Computer Architecture**

🏆 HONORS & AWARDS

- | | |
|---|----------|
| 1. Second Prize in Freshmen Scholarship, Tsinghua University | 09, 2020 |
| 2. Comprehensive Excellence Award for the 2020-2021 school year, Dept. of CST | 10, 2021 |
| 3. Third Prize in THU Challenge Cup Academic Competition, Tsinghua University | 04, 2022 |
| 4. December-9th Scholarship, highest scholarship in Dept. of CST | 01, 2023 |
| 5. Comprehensive Excellence Award for the 2022-2023 school year, Dept. of CST | 10, 2023 |

- | | |
|---|----------|
| 6. Five Star ZiJing Volunteer Award, Tsinghua University Communist Youth League Committee | 05, 2024 |
| 7. Outstanding Paper Award for Diploma Project, Tsinghua University | 06, 2024 |
| 8. Outstanding Graduate Award, Beijing Municipal Education Commission | 06, 2024 |

</> PROJECTS

1. OpenBackdoor: An open-source toolkit for textual backdoor attack and defense 04, 2022 – 09, 2022
 - Directed by Associate Professor **Zhiyuan Liu**, THUNLP
 - Summarize three practical scenarios of attack methods based on their accessibility and goals
 - Conclude novel metrics for three evaluation dimensions and recommend scenario-specified evaluation methodologies
 - Develop an open-source toolkit OpenBackdoor and conduct extensive benchmark experiments
 - Propose **CUBE**, a simple yet strong baseline method targeting purifying poisoned datasets
 - Second author. **Paper** submitted to NeurIPS 2022 D&B

⚙️ PUBLICATIONS

* indicates equal contribution.

- **A Unified Evaluation of Textual Backdoor Learning: Frameworks and Benchmarks [paper]**
Ganqu Cui*, Lifan Yuan*, **Bingxiang He**, Yangyi Chen, Zhiyuan Liu, Maosong Sun.
NeurIPS 2022 Datasets and Benchmarks Track (Spotlights)
- **Beat LLMs at Their Own Game: Zero-Shot LLM-Generated Text Detection via Querying ChatGPT [paper]**
Biru Zhu, Lifan Yuan, Ganqu Cui, Yangyi Chen, Chong Fu, **Bingxiang He**, Yangdong Deng, Zhiyuan Liu, Maosong Sun, Ming Gu
EMNLP 2023 Main
- **ULTRAFEEDBACK: Boosting Language Models with Scaled AI Feedback [paper]**
Ganqu Cui*, Lifan Yuan*, Ning Ding, Guanming Yao, **Bingxiang He**, Wei Zhu, Yuan Ni, Guotong Xie, Ruobing Xie, Yankai Lin, Zhiyuan Liu, Maosong Sun
ICML 2024 Poster
- **Tell Me More! Towards Implicit User Intention Understanding of Language Model Driven Agents [paper]**
Cheng Qian*, **Bingxiang He***, Zhong Zhuang, Jia Deng, Yujia Qin, Xin Cong, Zhong Zhang, Jie Zhou, Yankai Lin, Zhiyuan Liu, Maosong Sun
ACL 2024 Main
- **Zero-Shot Generalization during Instruction Tuning: Insights from Similarity and Granularity [paper]**
Bingxiang He*, Ning Ding*, Cheng Qian*, Jia Deng, Ganqu Cui, Lifan Yuan, Huan-ang Gao, Huimin Chen, Zhiyuan Liu, Maosong Sun
NeurIPS 2024 Submission

🔧 IT SKILLS

- Computer Language: C/C++ == Python > JavaScript == TypeScript > SQL
- Platform: Linux, Windows
- Tool: Git, SSH, Make, Tmux, Markdown, LaTeX
- Full-stack development: React, Django, Nginx, MySQL
- CET6: 571