RUOMENG DING

ruomeng@cs.unc.edu Google Scholar Github Homepage

Research Interest: LLM Reasoning · Active Learning · Trustworthy AI

EDUCATION

Ph.D. in Computer Science, University of North Carolina at Chapel Hill **Aug, 2025 - Now**
- Advised by Prof. Zhun Deng.

M.S. in Computer Science, Georgia Institute of Technology **Aug, 2022 - May, 2025**
- GPA: 4.00/4.00, Specialization: Machine Learning.

B. Eng. in Computer Science and Technology, Tianjin University **Aug, 2018 - May, 2022**
- GPA: 3.75/4.00, Major Rank: 11/169 (6.5%).

PREPRINT

1. **R. Ding***, Y. Pang*, H. Sun, Y. Wang, Z. S. Wu, and Z. Deng, "Rubrics as an attack surface: Stealthy preference drift in llm judges," 2026 [paper] [code]
2. **R. Ding***, T. Gao*, T. P. Zollo, E. Bachmat, R. Zemel, and Z. Deng, "Whom to query for what: Adaptive group elicitation via multi-turn llm interactions," 2026 [paper] [code]

PUBLICATIONS

1. [AAAI 2026] **R. Ding**, W. Cheng, M. Shao, and C. Zhao, "Skillgen: Learning domain skills for in-context sequential decision making," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2026 **(Oral)** [paper] [code]
2. [ACL 2024] **R. Ding**, C. Zhang, L. Wang, Y. Xu, M. Ma, W. Zhang, S. Qin, S. Rajmohan, Q. Lin, and D. Zhang, "Everything of thoughts: Defying the law of penrose triangle for thought generation," in *Findings of the Association for Computational Linguistics (ACL)*, 2024 [paper] [code]
3. [NeurIPS 2024] R. Yang, **R. Ding**, Y. Lin, H. Zhang, and T. Zhang, "Regularizing hidden states enables learning generalizable reward model for LLMs," in *The Thirty-eighth Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024 [paper] [code]
4. [FSE 2023] **R. Ding**, C. Zhang, L. Wang, Y. Xu, M. Ma, X. Wu, M. Zhang, Q. Chen, X. Gao, X. Gao, H. Fan, S. Rajmohan, Q. Lin, and D. Zhang, "Tracediag: Adaptive, interpretable, and efficient root cause analysis on large-scale microservice systems," in *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE)*, 2023 [paper] [slides]
5. [KDD 2023] L. Wang, C. Zhang, **R. Ding**, Y. Xu, Q. Chen, W. Zou, Q. Chen, M. Zhang, X. Gao, H. Fan, S. Rajmohan, Q. Lin, and D. Zhang, "Root cause analysis for microservice systems via hierarchical reinforcement learning from human feedback," in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2023 [paper]
6. [SDM 2025] **R. Ding**, X. Zhao, C. Zhao, M. Shao, Z. Chen, and H. Chen, "Evidence-based out-of-distribution detection on multi-label graphs," in *Proceedings of the 2025 SIAM International Conference on Data Mining (SDM)*, pp. 588–597, SIAM, 2025 [paper]

7. [VLDB 2023] Y. Chen, C. Zhang, M. Ma, Y. Liu, **R. Ding**, B. Li, S. He, S. Rajmohan, Q. Lin, and D. Zhang, "Imdiffusion: Imputed diffusion models for multivariate time series anomaly detection," in *Proceedings of the VLDB Endowment (VLDB)*, 2023 [paper]

8. [TCYB 2023] T. Li, W. Wang, P. Jiao, Y. Wang, **R. Ding**, H. Wu, L. Pan, and D. Jin, "Exploring temporal community structure via network embedding," in *IEEE Transactions on Cybernetics (TCYB)*, 2023 [paper]

9. [UDM-AAAI 2023] **R. Ding**, X. Zhao, C. Zhao, and M. Shao, "Detecting multi-label out-of-distribution nodes on graphs," in *AAAI Workshop on Uncertainty Reasoning and Quantification in Decision Making*, 2023 [paper]

RESEARCH EXPERIENCE

- Research Intern, **Microsoft Research** **May, 2024 - Aug, 2024**
Advised by Dr. Minghua Ma and Dr. Ze Li Redmond, WA (Onsite)

- Focus on Large Language Model based Multi-Agents system for Incident Triage.
- Responsible for methodology design, experiments, and prototyping; presented the work to Microsoft Research and Azure.

- Research Assistant, **University of Illinois Urbana-Champaign** **Mar, 2024 - Jul 2024**
Advised by Dr. Tong Zhang, working with Rui Yang Urbana, IL (Remote)

- Focus on enhancing the Reward Model's generalization ability against distribution shifts for Reinforcement Learning from Human Feedback (RLHF). (*NeurIPS 2024, 2nd author)*
- Responsible for literature review and RLHF experiments. Contribute to paper writing and code repository.

- Research Intern, **Microsoft Research Asia** **Nov, 2022 - Aug, 2023**
Advised by Dr. Lu Wang and Dr. Chaoyun Zhang Beijing, China (Hybrid)

- Focus on Boost LLM reasoning with domain knowledge via Monte Carlo Tree Search (MCTS). Responsible for methodology design, experiments, paper writing, and code open-sourcing. (*ACL 2024, 1st author*)
- Focus on Root Cause Analysis for large-scale microservices systems, leveraging RL and RLHF to detect the root causes of anomalies. The method related to the paper has been deployed in the M365 system. (*KDD 2023, 3rd author; FSE 2023, 1st author*)

- Research Assistant, **Tianjin University** **Aug, 2021 - Aug, 2022**
Advised by Dr. Minglai Shao and Dr. Wenjun Wang Tianjin, China (Onsite)

- Focus on Multi-label Out-of-Distribution detection on graphs. Co-advised by Dr. Xujiang Zhao and Dr. Chen Zhao. *(UDM-AAAI 2022, 1st author)*
- Focus on dynamic community detection on graphs. Responsible for part of the experiments. *(TCYB 2023)*

HONORS AND FELLOWSHIPS

1. **Doctoral Merit Fellowship**, University of North Carolina at Chapel Hill 2025 - 2026
2. **Merit Scholarship**, Georgia Institute of Technology 2022 - 2023