

EDUCATION

University of Illinois at Urbana-Champaign

Urbana-Champaign, IL

Ph.D. Candidate in Computer Science, Department of Computer Science, The Grainger College of Engineering,

Advisor: Prof. [Tarek Abdelzaher](#)

Aug 2020–May 2025 (Expected)

- **Specialization:** Graph Mining and Multi-modal Large Language Models, especially on Social Network Analysis and Simulation. **Participated in two product-oriented grant-funded projects applying research outcomes.**

University of Illinois at Urbana-Champaign

Urbana-Champaign, IL

M.S. in Computer Science, Department of Computer Science, The Grainger College of Engineering,

Advisor: Prof. [Tarek Abdelzaher](#)

Aug 2022–May 2023

- **Thesis:** A Structured Matrix Factorization Method for Computational Modeling of Hierarchical Polarization in Social Interactions

University of Illinois at Urbana-Champaign

Urbana-Champaign, IL

B.S. in Computer Science & B.A. in Music, Department of Computer Science & School of Music

Aug 2015–May 2020

- **Completed Advanced Elective Tracks:** Intelligence and Big Data; Software Foundations; Media; Scientific, Parallel, and High Performance Computing

AWARDS AND CERTIFICATES

- **Certificate in Foundations of Teaching** UIUC Center for Innovation in Teaching & Learning, May 2024
- **Outstanding Teaching Assistant Award** Spring 2024
- University Honors: Bronze Tablet Plaque May 2020
- John R. Pasta Outstanding Undergraduate Award April 2018

TEACHING INTERESTS

- Computer Science Core Introductory Curricula
- Artificial Intelligence and Machine Learning
- Computer Networks
- Discrete Mathematics, Data Structures, and Algorithms
- Software Engineering
- Multimedia (Web Development and Computer Graphics)

TEACHING EXPERIENCE

Graduate Teaching Assistant & Guest Lecturer

UIUC

CS 438: Communication Networks

Spring 2024

- Guest lectures on the Network layer (Routing and IP).
- Responsible for managing and grading homework and machine problems, answering questions online, and hosting office hours.
- Retooled the autograder for HTTP and Routing programming assignment for fast and automated feedback.
- Mentoring students who registered for extra credit hours with their projects.
- Honored with the Outstanding Teaching Assistant Award.

Course Assistant

UIUC

CS 126: Software Design Studio

Spring 2018

- Moderation of team discussions and code reviewing submitted projects.
- Hosting office hours to resolve questions and advise on students' projects.

Session Lecturer

UIUC

CS@Illinois Sail Event

April 2016

- Taught one hour in a class of 10 high school students (potential incoming first-year students) about speech synthesis in an intuitive and example-and-application-based presentation.
- Guided and gave tour to students around the department, and provided requested information.





RESEARCH INTERESTS

Artificial Intelligence and Social Network Analysis Continue researching graph mining and large language models for social network analysis, emphasizing (1) belief detection and embedding, and (2) community response simulation to social events. Integrate more modalities, such as knowledge graphs, images, audio, and video, into analyzing social media platforms. Expand the research to video and streaming platforms and extend the language models to better process short messages, which usually appear on social media platforms.








Computer Science Education. Investigate new techniques, evaluate existing practices in computer science education, and derive more pedagogical methods from learning and cognitive theories. Build on notional machines and leverage the theory on misconceptions for core computer science concepts to help students develop good mental models and insights. Investigate the power of technologies in classrooms but also ways to increase inclusivity. Develop a course on AI technology basics that provides opportunities for hands-on projects and promotes AI safety awareness. Explore the potential of Generative AI in the context of specialization in computer science classrooms.

PUBLICATIONS

— Journal Articles —


1. **The Paradox of Information Access: On Modeling Polarization in the Age of Information** 
Chao Xu, Jinyang Li, Dachun Sun, Jinning Li, Tarek Abdelzaher, Jesse Graham, Michael Macy, C. Lebiere, B. Szymanski
IEEE Transactions on Control of Network Systems, **TCNS**, 2023
2. **ControlVAE: Tuning, Analytical Properties, and Performance Analysis** 
Huajie Shao, Zhisheng Xiao, Shuochao Yao, Dachun Sun, Aston Zhang, Shengzhong Liu, Tianshi Wang, J. Li, T. Abdelzaher
IEEE Transactions on Pattern Analysis and Machine Intelligence, **PAMI**, 2022
3. **Computational Modeling of Hierarchically Polarized Groups by Structured Matrix Factorization** 
Dachun Sun, Chaoqi Yang, Jinyang Li, Ruijie Wang, Shuochao Yao, Huajie Shao, Dongxin Liu, S. Liu, T. Wang, T. Abdelzaher
Frontiers in Big Data, 2021
4. **Truth Discovery With Multi-Modal Data in Social Sensing** 
Huajie Shao, Dachun Sun, Shuochao Yao, Lu Su, Zhibo Wang, Dongxin Liu, Shengzhong Liu, Lance Kaplan, Tarek Abdelzaher
IEEE Transactions on Computers, 2020

— Peer-Reviewed Conference and Workshop Papers —

1. **Fine-grained Control of Generative Data Augmentation in IoT Sensing**
Tianshi Wang, Qikai Yang, Ruijie Wang, Dachun Sun, J. Li, Y. Chen, Y. Hu, C. Yang, T. Kimura, D. Kara, T. Abdelzaher
Advances in Neural Information Processing Systems, **NeurIPS**, 2024
2. **MetaHKG: Meta Hyperbolic Learning for Few-shot Temporal Reasoning** 
Ruijie Wang, Yutong Zhang, Jinyang Li, Shengzhong Liu, Dachun Sun, T. Wang, T. Wang, Y. Chen, D. Kara, T. Abdelzaher
Proceedings of the 47th International ACM SIGIR Conference, **SIGIR**, 2024
3. **LLM-Guided Disentangled Belief Representation Learning on Social Graphs**  
Jinning Li, Ruipeng Han, Chenkai Sun, Dachun Sun, Ruijie Wang, Jingying Zeng, Yuchen Yan, H. Tong, T. Abdelzaher
The 33rd International Conference on Computer Communications and Networks, **ICCCN**, 2024
4. **Data Augmentation for Human Activity Recognition via Condition Space Interpolation within Generative Model** 
Tianshi Wang, Yizhuo Chen, Qikai Yang, Dachun Sun, Ruijie Wang, Jinyang Li, Tomoyoshi Kimura, Tarek Abdelzaher
The 33rd International Conference on Computer Communications and Networks, **ICCCN**, 2024
5. **Influence Pathway Discovery on Social Media** 
Dachun Sun*, Xinyi Liu*, Ruijie Wang*, Jinning Li, Christina Youn, You Lyu, Jianyuan Zhan, Dayou Wu, Xinhe Xu, Mingjun Liu, Xinshuo Lei, Zhihao Xu, Yutong Zhang, Zehao Li, Qikai Yang, Tarek Abdelzaher
IEEE 9th International Conference on Collaboration and Internet Computing, **CIC**, 2023
6. **Noisy Positive-Unlabeled Learning with Self-Training for Speculative Knowledge Graph Reasoning** 
Ruijie Wang, Baoyu Li, Yichen Lu, Dachun Sun, Jinning Li, Yuchen Yan, Shengzhong Liu, Hanghang Tong, Tarek Abdelzaher
Association for Computational Linguistics, **ACL**, 2023
7. **AdaMask: Enabling Machine-Centric Video Streaming w/ Adaptive Frame Masking for DNN Inference Offloading** 
Shengzhong Liu, Tianshi Wang, Jinyang Li, Dachun Sun, Mani Srivastava, Tarek Abdelzaher
Proceedings of the 30th ACM International Conference on Multimedia, **MM**, 2022

8. **Unsupervised Belief Representation Learning with Information-Theoretic VGAE**  
Jinning Li, Huajie Shao, [Dachun Sun](#), Ruijie Wang, Yuchen Yan, Jinyang Li, Shengzhong Liu, Hanghang Tong, Tarek Abdelzaher
Proceedings of the 45th International ACM SIGIR Conference, SIGIR, 2022
9. **Learning to Sample and Aggregate: Few-shot Reasoning over Temporal Knowledge Graphs** 
Ruijie Wang, Zheng Li, [Dachun Sun](#), Shengzhong Liu, Jinning Li, Bing Yin, Tarek Abdelzaher
Advances in Neural Information Processing Systems, NeurIPS, 2022
10. **DyDiff-VAE: A Dynamic Variational Framework for Information Diffusion Prediction**  
Ruijie Wang, Zijie Huang, Shengzhong Liu, Huajie Shao, Dongxin Liu, Jinyang Li, T. Wang, [Dachun Sun](#), S. Yao, T. Abdelzaher
Proceedings of the 44th International ACM SIGIR Conference, SIGIR, 2021
11. **On Polarization Dynamics in the Age of Information Overload** 
Chao Xu, Jinyang Li, [Dachun Sun](#), Ruijie Wang, Tarek F Abdelzaher, Jesse Graham, Boleslaw K Szymanski
ICWSM Workshops, 2021
12. **Simulating Online Social Response: A Stimulus/Response Perspective** 
Huajie Shao, Tarek Abdelzaher, Jiawei Han, Minhao Jiang, Yuning Mao, Yu Meng, Wenda Qiu, [Dachun Sun](#), Ruijie Wang, Chaoqi Yang, Zhenzhou Yang, Xinyang Zhang, Yu Zhang, Sam Cohen, James Flamino, Gyorgy Korniss, Omar Malik, Aamir Mandviwalla, Boleslaw Szymanski, Lake Yin
Proceedings - Winter Simulation Conference, 2021
13. **ControlVAE: Controllable Variational Autoencoder**   
Huajie Shao, Shuochao Yao, [Dachun Sun](#), Aston Zhang, Shengzhong Liu, Dongxin Liu, Jun Wang, Tarek Abdelzaher
Proceedings of the 37th International Conference on Machine Learning, PMLR, 2020
14. **paperzrepo: GitHub Repository Recommendation for Academic Papers**   
[Dachun Sun](#)*, Huajie Shao*, Jiahao Wu, Zecheng Zhang, Aston Zhang, Shuochao Yao, S. Liu, T. Wang, C. Zhang, T. Abdelzaher
Proceedings of The Web Conference 2020, WWW, 2020

—— Tutorials ——

1. **Towards Efficient Temporal Graph Learning: Algorithms, Frameworks, and Tools** 
Ruijie Wang, Wanyu Zhao, [Dachun Sun](#), Charith Mendis, Tarek Abdelzaher
33rd ACM International Conference on Information and Knowledge Management, CIKM, 2024

—— Under Review ——

1. **SCRaG: A Framework for Simulating Community Responses to Social Media Posts via Large Language Models and Retrieval Techniques**
[Dachun Sun](#), You Lyu, Ruijie Wang, Jianyuan Zhan, Qikai Yang, Jinning Li, Xinyi Liu, Chenkai Sun, Tarek Abdelzaher
2. **Perturbation-based Graph Active Learning for Weakly-Supervised Belief Representation Learning**
[Dachun Sun](#), Ruijie Wang, Jinning Li, Ruipeng Han, Xinyi Liu, You Lyu, and Tarek Abdelzaher
3. **Uncovering Cross-Domain Recommendation Ability of Large Language Models**
Xinyi Liu, Ruijie Wang, [Dachun Sun](#), Dilek Hakkani Tur, Tarek Abdelzaher
4. **Boosting Discernment with Propaganda Detectors: An LLM-based Study on Susceptibility to Opinion Manipulation**
Xinyi Liu, Ruijie Wang, [Dachun Sun](#), Dilek Hakkani Tur, Tarek Abdelzaher

RESEARCH EXPERIENCE

DARPA Program: Influence Campaign Awareness and Sensemaking (INCAS)

Ph.D. Candidate, Advisor: Prof. [Tarek Abdelzaher](#), Collaborators: [Jiawei Han](#) & [Heng Ji](#) & [Hanghang Tong](#)

UIUC

Aug 2020–Present

- Developed techniques and tools to detect, characterize, and track geopolitical influence campaigns with quantified confidence.
- Developed a controllable unsupervised Graph Variational Autoencoder to learn and disentangle belief/ideology representations from social graphs based on the propagation pattern and changes over time. Using the representation to cluster and identify polarized contents.
- Incorporate Large Language Models (LLMs) to embed multimodal contents to enhance the features for graph algorithms, improving the efficiency and effectiveness of influence campaign detection and sensemaking.
- Utilize agentic approach and retrieval augmented generation to simulate the behavior of communities.

DARPA Program: Social Simulation System for Social Streams (SocialSim)

Research Assistant, Advisor: Prof. [Tarek Abdelzaher](#), Collaborator: [Jiawei Han](#)

UIUC

Aug 2017–2021

- Developed an innovative system for high-fidelity computational simulation of online social behavior, focusing on information spread and evolution.
- We applied our system to analyze how external events stimulate topic discussions, and how information propagates on social platforms (e.g., Twitter, YouTube, Reddit). We ranked the top in internal challenges multiple times.
- We built a new class of Non-negative Matrix Tri-factorization (NMTF) approach to disentangle hierarchical belief structures from collected texts and propose a universal multi-belief structure under specific social topics.

NCSA SPIN Program: Computer Vision for Crop Growth Stage Classification

Research Intern, Advisor: Prof. [Kaiyu Guan](#)

NCSA, UIUC

May 2018–Dec 2018

- Worked on the classification system that identifies critical crop growth stages from field cameras using Computer Vision techniques. On identifying the VT stage of corn growth, the approach of VAE with patches achieved 81% testing accuracy.

INDUSTRY EXPERIENCE

TripAdvisor

Full-stack Software Developer Intern

Needham, MA

May–Aug 2017 & in 2019

- Maintained full-stack code-base for Restaurant Business-to-Business services, including owners' dashboards and feedback aggregation.
- Developed a new product checkout pipeline that assists restaurant owners in setting up customer WiFi hotspots.
- Proposed the idea of personal itinerary designers with crafted attraction spot recommendations and transportation supports during travel on mobile.

Spatial Light Interference Microscopy (SLIM)

Research Intern & Software Developer, Advisor: Dr. [Gabriel Popescu](#)

QLI Labs, UIUC

Aug 2016–Feb 2017

- Developed new image gradient processing techniques to augment phase-gradient microscopy.
- Engineered control software using C++/Qt for more precise maneuvering of the microscope.

SERVICE AND OUTREACH

- Group session host for TA Training at UIUC 2024
- Reviewer at AAAI 2024, 2025 2023 & 2024
- Reviewer at TIST Journal 2024
- Reviewer at CVPR 2024 Workshop SynData4CV 2024
- Reviewer at ACL 2024 Workshop ALVR and KnowledgeLM 2024
- Reviewer at ICML 2024 Workshop FM-Wild 2024
- Director at Jasmine Field Orchestra Summer 2021–Present
Arranger, orchestrator, and conductor for the ensemble.
- Volunteer at IEEE INFOCOM 2021 2021
Co-hosting the online session.
- Member at ACM-ICPC 2017–2018
Learn competitive programming and attend weekly internal rounds. Participate in developing a new judge system as the university's platform for better user experience and administrative convenience.

PROJECTS

Apollo: Anatomy of Conflict (Accompanying Website for the INCAS Research Project)

Tech Lead, Advisor: Prof. [Tarek Abdelzaher](#)

UIUC

Oct 2021–Present

- Designed and implemented the comprehensive analysis pipeline and research outcomes in the project capability showcasing website, supporting Influence Campaign Awareness and Sensemaking program.
- Developed systems to visualize ideological embeddings and information pathways, providing insights into influence campaigns.
- Created an automated data collection and analysis pipeline, streamlining the gathering and processing of large-scale data.
- Collaborated with interdisciplinary teams to ensure the pipeline met both technical requirements and research objectives.

DiffSinger: Singing Voice Synthesis via Shallow Diffusion Mechanism

Core Contributor

Open-source
Aug 2022–Present

An open-source system with high fidelity that is expressive, controllable, flexible, and agnostic of languages.

- Developed the acoustic, auto-pitch, and other time-dependent generation models such as breathiness, energy, and tension.
- Trained the models for fundamental pitch extraction and harmonic-noise signal separation.
- Accelerated the training process by supporting multi-GPU and multi-node distributed strategies.
- Developed various GUI tools for dataset creation and editor.

MicroChromo

Advisor: Prof. [Heinrich Taube](#) & Prof. [Eli Fieldsteel](#)

UIUC, Open-source
Feb 2020–May 2020

Capstone project for the degree of B.A. in Music. MicroChromo is a virtual instrument wrapper to aid microtonal composers in digital audio workstations (DAW).

- Enabled all virtual instruments to support microtonal passage with instance duplication and pitch shift.
- Efficient scheduling of microtonal notes to instrument instances while automatically synchronizing states between instances.

PolygonSynth

Advisor: Prof. [Heinrich Taube](#) & Prof. [Eli Fieldsteel](#)

UIUC, Open-source
Feb 2020–Apr 2020

A personal project on creating a software synthesizer using polygonal shapes, which creates extraordinary stereo effects.