Xiaohan **Kang**

MACHINE LEARNING SCIENTIST · SOFTWARE ENGINEER

 (+1) 515-509-6693 | ✉ xkang515@gmail.com | 🏠 veggente.github.io | Veggente | in xiaohankang | 🎓 Xiaohan Kang

Work Experience

University of Illinois at Urbana–Champaign *Urbana, IL*
POSTDOCTORAL RESEARCH ASSOCIATE, DEPT. OF ELECTRICAL AND COMPUTER ENGINEERING, ADVISOR: PROF. BRUCE HAJEK *Mar. 2016–present*

- Conducted cutting-edge research on gene regulatory network reconstruction by analyzing large-scale time-series RNA-seq data, utilizing advanced machine learning techniques to infer regulatory interactions and identify key genes driving soybean flowering and circadian rhythms.
- Explored the frontiers of knowledge by investigating the fundamental limits of binary classification errors and causal network inference, employing rigorous statistical methods and mathematical modeling to unravel the intricacies of complex systems with uncertainty.
- Mentored and instructed a diverse cohort of over 60 students in ECE 313 (Probability with Engineering Applications), employing engaging pedagogical approaches to facilitate a deep understanding of probabilistic concepts and their practical applications in engineering, fostering a positive learning environment and receiving positive feedback from students.

Cisco Systems, Inc. *San Jose, CA*
SOFTWARE ENGINEERING INTERN *May–Aug. 2015*

- Spearheaded the development of a Django-based web application for debugging networking applications, employing modern software engineering practices and technologies to deliver a robust and user-friendly tool that significantly streamlined the debugging process, enhancing the efficiency and effectiveness of the networking team.

Education

Arizona State University *Tempe, AZ*
PH.D. IN ELECTRICAL ENGINEERING, ADVISOR: PROF. LEI YING *2015*

Tsinghua University *Beijing, China*
B.E. IN ELECTRONIC ENGINEERING *2009*

Skills

PYTHON, UNIX, GIT, C/C++, MATLAB, R, DJANGO, HUGO, PYTORCH

Recent Research Projects

My recent research has been focused on addressing both theoretical and computational challenges that arise from, or are inspired by, the analysis of large-scale biological datasets.

Gene regulatory network reconstruction

- Developed CausNet, a novel framework for sparse causal network reconstruction using time-series RNA-seq data, involving advanced statistical techniques and employing a Gaussian approximation of bootstrap aggregating to provide reliability scores for predicted regulatory interactions.
- Customized a range of advanced algorithms to the CausNet causal network inference framework. This included lasso, an optimization-based algorithm; oCSE, an efficient greedy algorithm; as well as sparse Bayesian learning, an EM algorithm. These adaptations exemplified capability and versatility of the CausNet framework.
- Conducted in-depth investigations into the importance of condition diversity in time series RNA-seq experiments with one-shot sampling, analyzing the impact of limited sample size on network inference accuracy and proposing strategies to mitigate the issue.
- Explored the connection between ordinary differential equation (ODE) models and graph models for gene regulatory networks, comparing the interpretability of these models in the context of RNA-seq data analysis.

Fundamental limits on binary classification errors and causal network inference

- Derived a robust and statistically rigorous maximum likelihood estimator of the receiver operating characteristic (ROC) curve for binary classification problems, utilizing advanced mathematical techniques to provide accurate estimates of classification performance.
- Developed mleroc, a Python implementation of the maximum likelihood estimator of the ROC curve, incorporating efficient algorithms and data structures to enable fast and scalable computations for large datasets.
- Made novel contributions to the field of causal network inference by providing a lower bound on the information requirements for accurate and reliable inference of causal relationships among variables, advancing our understanding of the fundamental limits and challenges in inferring causal networks from observational data.

Previous Research Projects

- Conducted extensive investigations into various topics in the field of computer networks, including scheduling algorithms in wireless communication networks, load balancing algorithms in computer networks, performance analysis of peer-to-peer (P2P) streaming networks, and admission control mechanisms for wireless access networks. My work has involved developing novel algorithms, analyzing performance metrics, and proposing mechanisms to enhance the efficiency, reliability, and quality of service in these network systems.

Publications

Conference publications

- [C9] Bruce Hajek and Xiaohan Kang, "Maximum likelihood estimation of optimal receiver operating characteristic curves from likelihood ratio observations," *IEEE International Symposium on Information Theory (ISIT)*, 2022. [DOI] [arXiv]
- [C8] Xiaohan Kang and Bruce Hajek, "Lower bounds on information requirements for causal network inference," *IEEE International Symposium on Information Theory (ISIT)*, 2021. [DOI] [arXiv]
- [C7] Honghao Wei, Xiaohan Kang, Weina Wang, and Lei Ying, "QuickStop: A Markov optimal stopping approach for quickest misinformation detection," *ACM International Conference on Measurement and Analysis of Computer Systems (SIGMETRICS)*, 2019. [DOI] [arXiv]
- [C6] Xiaohan Kang, I-Hong Hou, and Lei Ying, "On the capacity requirement of largest-deficit-first for scheduling real-time traffic in wireless networks," *ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2015. [DOI]
- [C5] Lei Ying, R. Srikant, and Xiaohan Kang, "The power of slightly more than one sample in randomized load balancing," *IEEE Conference on Computer Communications (INFOCOM)*, 2015. **(Best Paper Award)** [DOI]
- [C4] Xiaohan Kang, Juan José Jaramillo, and Lei Ying, "Stability of longest-queue-first scheduling in linear wireless networks with multihop traffic and one-hop interference," *IEEE Conference on Decision and Control (CDC)*, 2013. [DOI]
- [C3] Xiaohan Kang, Weina Wang, Juan José Jaramillo, and Lei Ying, "On the performance of largest-deficit-first for scheduling real-time traffic in wireless networks," *ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2013. [DOI]
- [C2] Xiaohan Kang, Juan José Jaramillo, and Lei Ying, "Impacts of peer churn on P2P streaming networks," *Annual Allerton Conference on Communication, Control and Computing (Allerton)*, 2012. [DOI]
- [C1] Xiaohan Kang, Juan José Jaramillo, "A strategy-proof and non-monetary admission control mechanism for wireless access networks," *International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine)*, 2010. [DOI]

Journal publications

- [J6] Xiaohan Kang, Bruce Hajek, and Yoshie Hanzawa, "From graph topology to ODE models for gene regulatory networks," *PLOS ONE*, vol. 15, no. 6, pp. e0235070, 2020. [DOI]
- [J5] Faqiang Wu, Xiaohan Kang, Minglei Wang, Waseem Haider, William B. Price, Bruce Hajek, and Yoshie Hanzawa, "Transcriptome-enabled network inference revealed the *GmCOL1* feed-forward loop and its roles in photoperiodic flowering of soybean," *Frontiers in Plant Science*, vol. 10, pp. 1221, 2019. [DOI]
- [J4] Xiaohan Kang, Bruce Hajek, Faqiang Wu, and Yoshie Hanzawa, "Time series experiment design under one-shot sampling: The importance of condition diversity," *PLOS ONE*, vol. 14, no. 10, pp. e0224577, 2019. [DOI]
- [J3] Lei Ying, R. Srikant, and Xiaohan Kang, "The power of slightly more than one sample in randomized load balancing," *Mathematics of Operations Research*, vol. 42, no. 3, pp. 692–722, 2017. [DOI]
- [J2] Xiaohan Kang, Weina Wang, Juan José Jaramillo, and Lei Ying, "On the performance of largest-deficit-first for scheduling real-time traffic in wireless networks," *IEEE/ACM Transactions on Networking*, vol. 24, pp. 72–84, Feb. 2016. [DOI]
- [J1] Xiaohan Kang, Juan José Jaramillo, Lei Ying, "Stability of longest-queue-first scheduling in linear wireless networks with multihop traffic and one-hop interference," *Queueing Systems*, vol. 80, no. 3, pp. 273–291, Jul. 2015. [DOI]

Selected talks

- [T10] "Finite-sample lower bounds on information requirements for causal network inference," *BIRS CMO Workshop on Learning in Networks: Performance Limits and Algorithms*, invited talk, Oaxaca, Mexico, 2022.
- [T9] "Lower bounds on information requirements for causal network inference," *INFORMS Annual Meeting*, invited talk, Anaheim, CA, 2021.
- [T8] "On modeling the circadian clock gene regulatory network in soybean," *Finding Your Inner Modeler Workshop IV (FYIM)*, University of Illinois at Chicago, Chicago, IL (virtual), 2021.
- [T7] "Time series experimental design under one-shot sampling: The importance of condition diversity," *Energy & Information Systems Seminar*, invited talk, Carnegie Mellon University, Pittsburgh, PA, 2019.
- [T6] "On the challenge of gene regulatory network reconstruction from high-throughput sequencing data," *Network Science Seminar Series*, invited talk, Arizona State University, Tempe, AZ, 2018.
- [T5] "CausNet: a causal inference algorithm for gene regulatory network reconstruction," *The Plant and Animal Genome XXVI Conference (PAG 2018)*, San Diego, CA, 2018.
- [T4] "The power of slightly more than one sample in randomized load balancing," *SINE Seminar*, invited talk, University of Illinois at Urbana–Champaign, Urbana, IL, 2016.
- [T3] "The power of slightly more than one sample in randomized load balancing," guest lecture (hosted by Prof. Rhonda Righter), University of California, Berkeley, Berkeley, CA, 2016.
- [T2] "The power of slightly more than one sample in randomized load balancing," *INFORMS Annual Meeting*, invited talk, Philadelphia, PA, 2015.
- [T1] "On the performance of largest-deficit-first for scheduling real-time traffic in wireless networks," invited talk (hosted by Prof. Eytan Modiano), Massachusetts Institute of Technology, MA, 2015.

Honors & Awards

2019	**Helmsley Fellowship**, Frontiers and Techniques in Plant Science Course, Cold Spring Harbor Laboratory	*Laurel Hollow, NY*
2015	**The First Place Team**, Cisco Intern Hackathon	*San Jose, CA*
2015	**Best Paper Award**, IEEE Conference on Computer Communications (INFOCOM)	*Hong Kong, China*
2014	**Exemplary Reviewer**, IEEE Communications Letters	

Professional Service

- Reviewer for IEEE/ACM Transactions on Networking, Queueing Systems, IEEE Transactions on Mobile Computing, IEEE Communications Letters, IEEE Transactions on Vehicular Technology, IEEE Signal Processing Letters, IEEE Transactions on Network Science and Engineering, and IEEE International Symposium on Information Theory.
- Technical Program Committee member for ACM MobiHoc 2019–2022, and WiOpt 2021.