Xin Zhang

Ph.D. student in Statistics

1219 Snedecor Hall
2438 Osborn Dr.
Ames, IA 50011, U.S.A.
Email: xinzhang.nac@gmail.com
Webpage: xinzhang-nac.github.io
Github: github.com/xinzhang-nac
Phone: +1)515-203-7703

Research Interests

Distributed Optimization, Deep Machine Learning, Reinforcement Learning, Differential Privacy.

Education

Iowa State University, Ames, IA *08/2016 - 05/2021*
Ph.D. student, Statistics
Advisors: Prof. Zhengyuan Zhu (STAT Dept.) and Prof. Jia Liu (CS Dept.)

Fudan University, Shanghai, China *09/2012 - 06/2016*
Bachelor of Science, Mathematics and Applied Mathematics

Professional Experience

Facebook, Inc., Menlo Park, CA *05/2020 - 08/2020*
Software Engineer Intern, Machine Learning

Research Projects

Distributed Methods for Large-Scale Machine Learning Models
Supervisor : Prof. Jia Liu *12/2017 - Current*

- Improved distributed machine learning algorithms for large-scale parallel computing and optimization.
- Developed asynchronous stochastic algorithms to utilize the stale information.
- Proposed a multi-stage sign operator to reduce communication cost and increase system robustness.
- Adopted the coordinate-wise median in message aggregation to mitigate Byzantine attacks.

Network Optimization: Efficient Communication and Data Privacy
Supervisor : Prof. Jia Liu *06/2018 - Current*

- Developed message compression algorithms in network to reduce the communication cost.
- Designed a novel compression operator to maximally reduce the communication load.
- Developed differential private network algorithms to protect data privacy.
- Developed stochastic varaince reduction method in network to reduced the computation complexity.

Spatial Coefficient Clustering Analysis for Remote Sensing Data
Supervisors : Prof. Zhengyuan Zhu *01/2019 - Current*

- Applied clustering models to the remote sensing data (MODIS Data) to identify the spatial groups.
- Proposed the forest/tree-based lasso penalties for simultaneous estimation and clustering.
- Improved the spatial non-parametric estimation by adopting the bivariate spline method.
- Designed parallel computing algorithm to speed up the estimation.

Conference Papers

1. **Zhang, X.**, Liu, J. and Zhu, Z., 2020. Taming Convergence for Asynchronous Stochastic Gradient Descent with Unbounded Delay in Non-Convex Learning. In *Proc. IEEE CDC 2020.*
2. Yang, H., **Zhang, X.**, Fang M. and Liu J. Adaptive Multi-Hierarchical signSGD for Communication-Efficient Distributed Optimization. In *Proc. IEEE SPAWC 2020.*
3. **Zhang, X.**, Fang, M., Liu, J. and Zhu, Z. Private and Communication-Efficient Edge Learning: A Sparse Differential Gaussian-Masking Distributed SGD Approach. In *Proc. ACM Mobihoc 2020.* (acceptance rate: 15%)

4. **Zhang, X.**, Liu, J., Zhu, Z., and Bentley, E. S. Communication-Efficient Network Distributed Optimization with Differential-Coded Compressors. In *Proc. IEEE INFOCOM 2020.* (acceptance rate: 19.8%)
5. Yang, H., **Zhang, X.**, Fang M. and Liu J. Byzantine-Resilient Stochastic Gradient Descent for Distributed Learning: A Lipschitz-Inspired Coordinate-wise Median Approach. In *Proc. IEEE CDC 2019.*
6. **Zhang, X.**, Liu, J., Zhu, Z. and Bentley, E.S. Compressed Distributed Gradient Descent: Communication-Efficient Consensus over Networks. In *Proc. IEEE INFOCOM 2019.* (acceptance rate: 19.7%)

Journal Articles

1. **Zhang, X.**, Li, L., Ng, M.K. and Zhang, S., 2017. Drug–target interaction prediction by integrating multiview network data. *Computational biology and chemistry*, 69, pp.185-193.
2. **Zhang, X.**, Gao, W.G. and Su, Y., 2015. Electricity consumer archetypes study based on functional data analysis and k-means algorithm. *Power System Technology*, 39, pp.3153-3162.

In Preparation

- **Zhang, X.**, Liu J. and Zhu, Z.. Distributed Linear Model Clustering over Networks: A Tree-Based Fused-Lasso ADMM Approach. arXiv preprint arXiv:1905.11549 (2019).
- **Zhang, X.**, Liu, J. and Zhu, Z. GT-STORM: Taming Sample and Communication Complexities in Decentralized Optimization.
- **Zhang, X.**, Wang, X. and Zhu, Z. Homogeneity Pursuit of Spatial Partial Linear Model: A Joint Approach of Bivariate Spline and Forest Lasso.
- Liu, Z., **Zhang, X.** and Liu, J. On Convergence and Generalization of Asynchronous SpiderBoost for Non-Convex Optimization

Teaching Experience

Teaching Assistant
Dept. Statistics, Iowa State University, Ames, IA *08/2016 - 05/2017*

- STAT 341(Spring 2017), Instructor: Dr.Kevin Kasper
- STAT 430(Fall 2016), Instructor: Dr. Kris De Brabanter
- STAT 341(Fall 2016), Instructor: Dr.Kevin Kasper

Awards

- **The Holly C. and E. Beth Fryer Award**, Department of Statistics, Iowa State University, 2018. This award is for a top second-year Ph.D. student in the department.
- **Outstanding Graduate of Fudan Unversity**, Fudan University, 2016.
- **First Prize of the scholarship for Outstanding Students at Fudan University** (5%), Fudan University, 2015.
- **Third Prize of the scholarship for Outstanding Students at Fudan University** , Fudan University, 2013-2014.

Skills

Projects in: R, Python, C++, PHP, SQL, Matlab, LATEX, Markdown.
Familiar with: Unix shell, SAS, HTML, Microsoft Office, Pytorch, Tensorflow.
Language: Proficient in English. Native in Chinese.