



Distributed and Parallel Data Analysis Workshop

September 21-23, 2016

SCHEDULE

Wednesday, September 21, 2016

NCSU Centennial Campus

- | | |
|-------------|--|
| 8:30-8:50 | Registration |
| 8:50-9:00 | Welcome and Introductory Remarks
Sujit Ghosh , SAMSI |
| 9:00-9:35 | <i>Scalable Probabilistic Inference from Big and Complex Data</i>
David Dunson , Duke |
| 9:40-10:15 | <i>Asynchronous Parallel Coordinate Update Algorithms</i>
Wotao Yin , UCLA |
| 10:15-10:50 | Break |
| 10:50-11:25 | <i>Distributed Hyper-Parameter Optimization for Machine Learning</i>
Yan Xu , SAS |
| 11:30-12:05 | <i>Interaction Selection and Screening for High Dimensional Data</i>
Helen Zhang , University of Arizona |
| 12:05-1:30 | Lunch |
| 1:30-2:05 | <i>Privacy-Preserving Methods for Handling Missing Data in Distributed Health Data Networks</i>
Qi Long , Emory |
| 2:10-2:45 | <i>DPDA Application in Predix Ecosystem for Real-time Monitoring and Diagnostics of Energy Assets</i>
Xiaomo Jiang , GE Power |
| 2:50-3:25 | <i>A Sequential Split-Conquer-Combine Approach for Gaussian Process Model in Analysis of Big Spatial Data</i>
Min-ge Xie , Rutgers |
| 3:25-4:00 | Break |
| 4:00-4:15 | <i>Funding Opportunities at NSF</i>
Yong, Zeng , NSF |

4:15-5:00 *Lightning Talks*
Moderator: **Sujit Ghosh**, SAMSI

5:00-6:30 Poster Session and Reception

6:30 Shuttle to Hotel

Thursday, September 22, 2016
NCSU Centennial Campus

8:50-9:00 Introductory Remarks – Program Organizers

9:00-9:35 *Distributed Estimation and Inference with Statistical Guarantees*
Jianqing Fan, Princeton

9:40-10:15 *HPDA Growth Constraints in Digital Marketing*
Samuel Franklin, 360i

10:15-10:50 Break

10:50-11:25 *Bayesian Aggregation for Extraordinarily Large Dataset*
Guang Cheng, Purdue

11:30-12:05 *Blessing of Massive Scale*
Han Liu, Princeton

12:05-1:30 Lunch

1:30-2:05 *Bayesian Neural Networks for High Dimensionsl Nonlinear Variable Selection*
Faming Liang, University of Florida

2:10-2:45 *Strategies & Principles for Distributed Machine Learning*
Eric Xing, Carnegie Mellon

2:50-3:25 *Challenges and Opportunities in Automated Driving and Connected Vehicles*
Yilu Zhang, GM; **Wei Tong**, GM

3:30-4:00 Break

4:00-4:35 *Parallel Local Graph Clustering*
Kimon Fountoulakis, UC Berkeley

4:35-5:00 Panel

5:00 Shuttle to Hotel

Friday, September 23, 2016

NCSU Centennial Campus

- | | |
|-------------|--|
| 8:50-9:00 | Introductory Remarks – Program Organizers |
| 9:00-9:35 | <i>Distributed Data Analysis on the Open Source HPC Systems Platform Stack</i>
Flavio Villanustre , LexisNexis Risk Solutions / RELX |
| 9:40-10:15 | <i>Scalable and Robust Statistical Estimation: a tale of the geometric median</i>
Stas Minsker , USC |
| 10:15-10:50 | Break |
| 10:50-11:25 | <i>Uncover Customer Insights with Apache Spark and ML</i>
Bo Zhang , IBM |
| 11:30-12:05 | <i>Some Recent Development in Spatial Statistics for Large Datasets</i>
Raj Guhaniyogi , University of California, Santa Cruz |
| 12:05-1:30 | Discussion and Wrap-up |
| 1:30-> | Box Lunch and Shuttle to RDU Airport |