

Pengzhi Gao

CONTACT INFORMATION	Petuum, Inc. 2555 Smallman Street, Suite 120, Pittsburgh, Pennsylvania 15222	<i>Mobile:</i> 215-696-0238 <i>Email:</i> gpengzhi@gmail.com <i>Homepage:</i> https://gpengzhi.github.io/
EDUCATION	Rensselaer Polytechnic Institute , Troy, NY Ph.D., Electrical Engineering, August 2013 - December 2017 <ul style="list-style-type: none">• Advisor: Professor Meng Wang• Thesis: High-dimensional Data Analysis by Exploiting Low-dimensional Models with Applications in Synchrophasor Data Analysis in Power Systems University of Pennsylvania , Philadelphia, PA M.S., Electrical Engineering, August 2011 - May 2013 <ul style="list-style-type: none">• GPA: 3.74/4 Xidian University , China B.S. (with honors), Electronic and Information Engineering, August 2007 - May 2011 <ul style="list-style-type: none">• GPA: 91.4/100 Class Rank: 1st in 114	
RESEARCH INTERESTS	My research interests lie in the intersection of the fields of signal processing, high-dimensional statistics, and machine learning.	
WORK EXPERIENCE	Data Scientist February 2018 to present Machine Learning Team, Petuum, Inc. Supervisor: Hong Wu <ul style="list-style-type: none">• Integrated the distributed gradient boosting algorithm (LightGBM) into Petuum operating system.• Designed and implemented the machine learning library (TensorFlow and DyNet) for Petuum data science product. Research Intern December 2010 to May 2011 Internet Media Group, Microsoft Research Asia Supervisor: Dr. Feng Wu <ul style="list-style-type: none">• Analyzed the data collected from 54 sensors deployed in the Intel Berkeley Research Lab (150 MB of data) to exploit the temporal correlations in sensor readings. Developed a joint source and network coding scheme for approximate data gathering in wireless sensor network.	
SKILL SETS	<ul style="list-style-type: none">• Proficiency with MATLAB, Python, C/C++, and Dynet.• Experienced in Java, R, C#, AMPL, PyTorch, and TensorFlow	
HONORS AND AWARDS	<ul style="list-style-type: none">• North America Finalist of IBM Watson Building Challenge 2017• Paper selected as the runner-up of the Best Paper in Electric Energy Systems Track of Hawaii International Conference on System Sciences 2015• Founders Award of Excellence (top 1%) 2015• Paper selected as one of the Best Conference Papers on Power System Analysis and Modeling of IEEE Power & Energy Society General Meeting 2014• Excellent Graduate of Xidian University (top 1%) 2011• National Scholarship (top 1%) 2010• First prize of the College Academic and Technological Scholarship (top 2%) 2008-2010• Excellent Student Awards (top 1%) 2008	

JOURNAL
PUBLICATIONS

1. **P. Gao**, R. Wang, and M. Wang. “Robust Matrix Completion by Exploiting Dynamic Low-dimensional Structures.” *submitted to IEEE Transactions on Signal Processing*, 2018.
2. **P. Gao**, R. Wang, M. Wang, and J. H. Chow. “Low-rank Matrix Recovery from Noisy, Quantized and Erroneous Measurements.” *IEEE Transactions on Signal Processing*, 2018, 66 (11): 2918-2932.
3. **P. Gao**, M. Wang, J. H. Chow, M. Berger, and L. M. Seversky. “Missing Data Recovery for High-dimensional Signals with Nonlinear Low-dimensional Structures.” *IEEE Transactions on Signal Processing*, 2017, 65 (20): 5421-5436.
4. **P. Gao**, M. Wang, J. H. Chow, S. G. Ghiocel, B. Fardanesh, G. Stefopoulos, and M. P. Razanousky. “Identification of Successive “Unobservable” Cyber Data Attacks in Power Systems Through Matrix Decomposition.” *IEEE Transactions on Signal Processing*, 2016, 64 (21): 5557-5570.
5. **P. Gao**, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos. “Missing Data Recovery by Exploiting Low-dimensionality in Power System Synchrophasor Measurements.” *IEEE Transactions on Power Systems*, 2016, 31 (2): 1006-1013.

CONFERENCE
PUBLICATIONS

1. G. Mijolla, S. Konstantinoupolos, **P. Gao**, J. H. Chow, and M. Wang. “An Evaluation of Low-Rank Matrix Completion Algorithms for Synchrophasor Missing Data Recovery.” *Proc. of the 20th Power Systems Computation Conference (PSCC)*, Dublin, Ireland, Jun. 2018.
2. **P. Gao**, and M. Wang. “Dynamic Matrix Recovery from Partially Observed and Erroneous Measurements.” *Proc. of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, Canada, Apr. 2018.
3. M. Wang, J. H. Chow, **P. Gao**, Y. Hao, W. Li, and R. Wang. “Recent Results of PMU Data Analytics by Exploiting Low-dimensional Structures.” *Proc. of the 10th Bulk Power Systems Dynamics and Control Symposium (IREP)*, Espinho, Portugal, Aug. 2017.
4. **P. Gao**, R. Wang, and M. Wang. “Quantized Low-rank Matrix Recovery with Erroneous Measurements: Application to Data Privacy in Power Grids.” *Proc. of Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2016.
5. **P. Gao**, M. Wang, and J. H. Chow. “Matrix Completion with Columns in Union and Sums of Subspaces.” *Proc. of IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Orlando, FL, Dec. 2015.
6. M. Wang, J. H. Chow, **P. Gao**, X. T. Jiang, Y. Xia, S. G. Ghiocel, B. Fardanesh, G. Stefopoulos, Y. Kokai, N. Saito, and M. P. Razanousky. “A Low-Rank Matrix approach for the Analysis of Large Amounts of Synchrophasor Data.” *Proc. of Hawaii International Conference on System Sciences (Runner-up of Best Paper in Electric Energy Systems Track)*, Kauai, Hawaii, Jan. 2015.
7. M. Wang, **P. Gao**, S. G. Ghiocel, J. H. Chow, B. Fardanesh, G. Stefopoulos, and M. P. Razanousky. “Identification of “Unobservable” Cyber Data Attacks on Power Grids.” *Proc. of IEEE SmartGridComm*, Venice, Italy, Nov. 2014.
8. **P. Gao**, M. Wang, S. G. Ghiocel, and J. H. Chow. “Modeless Reconstruction of Missing Synchrophasor Measurements.” *Proc. of IEEE Power & Energy Society General Meeting (selected in Best Conference Paper sessions)*, Washington, DC, Jul. 2014.

PATENTS

1. Meng Wang, **Pengzhi Gao**, and Joe H. Chow. “A low-rank-based missing PMU data recovery method.” Application No.: 62/445305, Filed January 12, 2017.

RESEARCH
EXPERIENCE

Research Assistant

August 2013 to present

ECSE Department,
Rensselaer Polytechnic Institute
Supervisor: Professor Meng Wang

- Analyzed the Phasor Measurement Unit (PMU) data (> 200 MB of data) to exploit the temporal and spatial correlations (low dimensionality) of the data.
- Proposed an identification method that can detect the cyber data attack in the power system. Tested our method on the actual PMU data from Central New York Power System.
- Developed an on-line algorithm to estimate the missing PMU data in real time manner. Built the corresponding action adapter in OpenPDC by C# code, reducing the computational time by 50%.
- Proposed a novel model to characterize the practical nonlinear datasets. Developed convex-optimization-based methods to recover missing data under this model. Tested our method on simulated power system data in IEEE 39-bus New England Power System.
- Proposed a novel method to recover the original data from quantized measurements even when partial measurements are corrupted. Developed a projected gradient method to solve the non-convex problem approximately. Tested our method on actual PMU data from Central New York Power System.

Research Assistant

May 2012 to May 2013

Department of Bioengineering,
University of Pennsylvania
Supervisor: Professor Gershon Buchsbaum

- Analyzed the EEG data from IIEG Portal for epilepsy detection.
- Proposed a new dictionary for EEG dataset. Improved the reconstruction performance of the EEG data by 20%.

PROJECTS

DyNet: The Dynamic Neural Network Toolkit

Machine Learning Team,
Petuum, Inc.

- DyNet is a neural network library developed by Carnegie Mellon University, Petuum, and many others. It is written in C++ and is designed to be efficient when run on either CPU or GPU, and to work well with networks that have dynamic structures that change for every training instance. I constantly contribute to this open sourced project.

Online Algorithm for PMU Data Processing (OLAP)

ECSE Department,
Rensselaer Polytechnic Institute

- We implemented OLAP by C# based on Project Alpha for the real-time application. Project Alpha is the elite version of Open PDC. The code developed on Project Alpha can be run on Open PDC as an action adapter.

PROFESSIONAL
ACTIVITIES &
SERVICE

- Student Member of IEEE, 2013 - 2017. Member of IEEE, 2018 - present.
- Teaching Assistant (Rensselaer Polytechnic Institute):
Modeling and Analysis of Uncertainty, Fall 2017,
Distributed Systems and Sensor Networks, Fall 2017.
- Reviewer:
IEEE Transactions on Smart Grid,
IEEE Transactions on Automatic Control,
IEEE/ACM Transactions on Networking,
IEEE Signal Processing Letters,
Annals of Mathematics and Artificial Intelligence,
American Control Conference,
International Symposium on Antennas and Propagation.