

# Yan Liu

---

**TITLE** Associate Professor, Computer Science Department, Viterbi School of Engineering  
Philip and Cayley MacDonald Endowed Early Career Chair  
Director of USC Machine Learning Center

**CONTACT INFORMATION** Computer Science Department *Office:* SAL 216  
941 Bloom Walk SAL 216 *Phone:* (213)740-4371  
University of Southern California *E-mail:* yanliu.cs@usc.edu  
Los Angeles, CA 90089 USA *Webpage:* <http://www-bcf.usc.edu/~liu32>

**RESEARCH INTERESTS** Developing machine learning and data mining algorithms for analyzing large-scale time-series and spatial time-series data, structured input/output data, and multi-sourced data with applications to climate, social media and biology applications.

**EDUCATION** Ph.D., **Carnegie Mellon University**, May. 2007  
School of Computer Science.  
*Thesis: Conditional Graphical Models for Protein Structure Prediction*

M.Sc., **Carnegie Mellon University**, May. 2004  
School of Computer Science

B.S., **Peking University**, Beijing, China, Jul. 2001  
Department of Computer Science and Technology

**EMPLOYMENT** University of Southern California Log Angeles, CA, USA  
Associate Professor Apr 2016 – Present  
Computer Science Department, Viterbi School of Engineering

University of Southern California Log Angeles, CA, USA  
Assistant Professor Aug 2010 – Apr 2016  
Computer Science Department, Viterbi School of Engineering

IBM T.J. Watson Research Center Yorktown Heights, NY, USA  
Research Staff Member Nov 2006 – Aug 2010  
Machine Learning Group, Math Science and Business Analytics Department

**AWARDS AND HONORS** IBM Faculty Award 2015  
Facebook Faculty Award 2015  
NSF CAREER Award 2013  
Okawa Foundation Award 2013  
Yahoo Faculty Research & Engagement Award 2011, 2015  
James H. Zumberge Research and Innovation Fund Individual Award 2011  
Best Paper Finalist of IEEE Conference on Data Mining 2011  
ACM Doctoral Dissertation Award Honorable Mention 2007  
Best Paper Award in SIAM Conference on Data Mining 2007

Team Member of Winner of KDD Cup 2009 on Fast Scoring on a Large Marketing Database  
Team Member of Winner of KDD Cup 2008 on Early Detection of Breast Cancer

Team Member of Winner of INFORMS 2008 Data Mining Competition  
Team Lead of Winner of KDD Cup 2007 on Netflix Movie Rating Prediction

JOURNAL  
PUBLICATION

- J1. R. Yu, H. Qiu, Z. Wen, C.-Y. Lin, Y. Liu. **A Survey on Social Media Anomaly Detection.** *ACM SIGKDD Explorations*, 2016.
- J2. R. Yu, Y. Liu. **Spatio-Temporal Analysis of Social Media Data.** *Encyclopedia of GIS*, 2016.
- J3. V. Pham, C. Shahabi, Y. Liu. **An Entropy-Based Model to Infer Social Strength from Spatiotemporal Data.** *ACM Transactions on Database Systems (TODS)*.
- J4. Y. Liu. **Scalable Multivariate Time Series Models for Climate Informatics.** *Computing in Science & Engineering/IEEE Intelligent Systems (CiSE) Magazine*, 2015.
- J5. R. Yu, H. Xin, Y. Liu. **GLAD: Group Anomaly Detection in Social Media Analysis.** *Transactions on Knowledge Discovery from Data (TKDD)*, 10(2):Article No. 18, 2015.
- J6. F. Liu, S. Chakraborty, F. Li, Y. Liu, A. C. Lozano. **Bayesian Regularization via Graph Laplacian.** *Journal of Bayesian Statistics*, 9(2): 449-474, 2014.
- J7. W. Li, S. Kang, C. Liu, S. Zhang, Y. Shi, Y. Liu, X. Zhou. **High-resolution Functional Annotation of Human Transcriptome: Predicting Isoform Functions by a Novel Multiple Instance-based Label Propagation Method,** *Nucleic Acids Research*, 42(6):e39, 2014.
- J8. Y. Chang, L. Tang, Y. Inagaki, Y. Liu. **What is Tumblr: A Statistical Overview and Comparison.** *SIGKDD Explorations*, 16(1): 21-29, 2014.
- J9. G.A. Burns, Y. Gil, N. Villanueva-Rosales, Y. Liu. **The AAAI Fall Symposium on Discovery Informatics: AI Takes a Science-Centered View on Big Data.** *AI Magazine*, 35(2). 2014.
- J10. M. T. Bahadori, Y. Liu, D. Zhang, **A General Framework for Scalable Transductive Transfer Learning,** *Knowledge and Information Systems (KAIS)*, (38):1-61-83, 2013.
- J11. Y. Chang, A. Dong, P. Kolari, R. Zhang, Y. Inagaki, F. Diaz, H. Zha, Y. Liu. **Improving Recency Ranking Using Twitter Data.** *ACM Transactions on Intelligent Systems and Technology (ACM TIST)*, 4(1): Article No. 4. 2013.
- J12. Y. Liu, A. Niculescu-Mizil, A. Lozano, Y. Lu. **Temporal Graphical Models for Cross-Species Gene Regulatory Network Discovery.** *J Bioinform Comput Biol*, 9(2):231-50, 2011.
- J13. J. Sun, Y. Liu, J. Tang, C. Apte. **Introduction to Special Issue on Large-Scale Data Mining.** *Transactions on Knowledge Discovery from Data (TKDD)* 5(2):7, 2011.
- J14. S. Rosset, C. Perlich, G. Swirszcz, P. Melville, Y. Liu. **Medical Data Mining: Insights from Winning Two Competitions.** *Data Min. Knowl. Discov.* 20(3): 439-468, 2010.
- J15. Y. Liu, J. Carbonell, V. Gopalakrishnan, P. Weigele. **Conditional Graphical Models for Protein Structural Motif Recognition.** *Journal of Computational Biology.* 16(5):639-57, 2009.
- J16. J. Yang, R. Yan, Y. Liu, E. P. Xing. **Harmonium Models for Video Classification.** *Journal of Statistical Analysis and Data Mining*, 1(1), 23-37, February 2008.
- J17. C. Perlich, P. Melville, Y. Liu, G. Swirszcz, R. Lawrence. **Winner Report: KDD CUP Breast Cancer Identification.** *SIGKDD Explorations.* 10(2): 39-42, 2008.
- J18. Y. Liu, Z. Kou. **Predicting Who Rated What in Large-Scale Datasets.** *SIGKDD Explorations.* 9(2): 62-65, 2007.
- J19. S. Rosset, C. Perlich, Y. Liu. **Making the Most of Your Data: KDD Cup 2007 “How Many Ratings” Winners Report.** *SIGKDD Explorations.* 9(2): 66-69, 2007.

- J20. Y. Liu, J. Carbonell, J. K. Seetharaman, V. Gopalakrishnan. **Comparison of Probabilistic Combination Methods for Protein Secondary Structure Prediction.** *Bioinformatics*. **20**(17):3099-107.
- J21. Y. Liu, J. Carbonell, P. Weigele, V. Gopalakrishnan. **Protein Fold Recognition Using Segmentation Conditional Random Fields.** *Journal of Computational Biology*. **13**(2):394-406.
- J22. R. Ghani, K. Probst, Y. Liu, M. Krema, A. Fano. **Text Mining to Extract Product Attributes.** *SIGKDD Explorations*. **8**(1): 41-48, 2006.

CONFERENCE  
PUBLICATION

- C1. S. Purushotham, W. Carvalho, T. Nilanon and Y. Liu. **Variational Recurrent Adversarial Deep Domain Adaptation.** *International Conference on Learning Representations (ICLR), 2017*.
- C2. R. Yu, Y. Li, U. Demiryurek, C. Shahabi, Y. Liu. **Deep Learning: A Generic Approach for Extreme Condition Traffic Forecasting.** *SIAM International Conference on Data Mining (SDM)*, 2017
- C3. X. He and Y. Liu, **Not Enough Data? Joint Inferring Multiple Diffusion Networks via Network Generation Priors,** *ACM International Conference on Web Search and Data Mining (WSDM)*, 2017.
- C4. X. He, K. Xu, D. Kempe and Y. Liu. **Learning Influence Functions from Incomplete Observations.** *Advances in Neural Information Processing Systems (NIPS)*, 2016.
- C5. D. Cheng, R. Peng, I. Perros, and Y. Liu. **SPALS: Fast Alternating Least Squares via Implicit Leverage Scores Sampling.** *Advances in Neural Information Processing Systems (NIPS)*, 2016.
- C6. Z. Che, S. Purushotham, R. Khemani, and Y. Liu. **Interpretable Deep Models for ICU Outcome Prediction.** *American Medical Informatics Association Annual Symposium (AMIA)*, 2016.
- C7. D. Deng, C. Shahabi, U. Demiryurek, L. Zhu, R. Yu and Y. Liu. **Latent Space Model for Road Networks to Predict Time-Varying Traffic.** *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'16)*, 2016.
- C8. R. Yu and Y. Liu. **Learning from Multiway Data: Simple and Efficient Tensor Regression.** *International conference on Machine Learning (ICML'16)*, 2016.
- C9. R. Yu, A. Gelfand, S. Rajan, C. Shahabi, Y. Liu. **Geographic Segmentation via Latent Poisson Factor Model.** To appear in *ACM International Conference on Web Search and Data Mining (WSDM)*, 2016.
- C10. M.T. Bahadori, D. Kale, Y. Fan, Y. Liu. **Functional Subspace Clustering with Application to Time Series.** *International Conference on Machine Learning (ICML 2015)*, 2015.
- C11. X. He, T. Rekatsinas, J. Foulds, L. Getoor, Y. Liu. **HawkesTopic: A Joint Model for Network Inference and Topic Modeling from Text-Based Cascades.** *International Conference on Machine Learning (ICML 2015)*, 2015.
- C12. R. Yu, D. Cheng, Y. Liu. **Accelerated Online Low Rank Tensor Learning for Multivariate Spatiotemporal Streams.** *International Conference on Machine Learning (ICML 2015)*, 2015.
- C13. Z. Che, D. Kale, W. Li, M. T. Bahadori, and Y. Liu. **Deep Computational Phenotyping.** *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'15)*, 2015.
- C14. D. Cheng, Y. Cheng, Y. Liu, R. Peng, and S. Teng. **Efficient Sampling for Gaussian Graphical Models via Spectral Sparsification.** *International Conference on Learning Theory (COLT '15)*, 2015.

- C15. D. Cheng, X. He, Y. Liu. **Model Selection for Topic Models via Spectral Decomposition.** *Proceedings of AISTAT'2015*, 2015.
- C16. D. Kale, M. Ghazvininejad, A. Ramakrishna, J. He, Y. Liu. **Hierarchical Active Transfer Learning.** *Proceedings of SIAM Conference on Data Mining (SDM)*, 2015.
- C17. M. T. Bahadori, R. Yu, Y. Liu, **Fast Multivariate Spatio-temporal Analysis via Low Rank Tensor Learning.** *Spotlight in NIPS (NIPS 2014)*, 2014.
- C18. Y. Chang, M. Yamada, A. Ortega, Y. Liu. **Ups and Downs in Buzzes: Life Cycle Modeling for Temporal Pattern Discovery.** *Proceedings of the IEEE 14th International Conference on Data Mining (ICDM)*, 2014.
- C19. D. Kale, D. Gong, Z. Che, G. Medioni, R. Wetzell, P. Ross, Y. Liu. **An Examination of Multivariate Time Series Hashing with Applications to Health Care.** *Proceedings of the IEEE 14th International Conference on Data Mining (ICDM)*, 2014.
- C20. D. Cheng, Y. Liu. **Parallel Gibbs Sampling for Hierarchical Dirichlet Processes via Gamma Processes Equivalence.** *ACM Conference on Knowledge Discovery & Data Mining (SIGKDD)*, 2014.
- C21. D. Cheng, M. Bahadori, Y. Liu. **FBLG: A Simple and Effective Approach for Temporal Dependence Discovery from Time Series Data.** *ACM Conference on Knowledge Discovery & Data Mining (SIGKDD)*, 2014.
- C22. R. Yu, X. He, Y. Liu. **GLAD: Group Anomaly Detection in Social Media Analysis.** *ACM Conference on Knowledge Discovery & Data Mining (SIGKDD)*, 2014.
- C23. J. He, Y. Liu, Q. Yang. **Linking Heterogeneous Input Spaces with Pivots for Multi-Task Learning.** *SIAM Conference on Data Mining (SDM'14)*, 2014.
- C24. T. Bahadori, Y. Liu, E. P. Xing. **Fast Structure Learning in Generalized Stochastic Processes with Latent Factors.** *ACM Conf. on Knowledge Discovery & Data Mining (SIGKDD)*, 2013.
- C25. M. T. Bahadori, Y. Liu, **An Examination of Large-Scale Granger Causality Inference.** *SIAM Conference on Data Mining (SDM'13)*, 2013.
- C26. V. Pham, C. Shahabi, Y. Liu. **An Entropy-Based Model to Infer Social Strength from Spatiotemporal Data.** *ACM SIGMOD Conference on Management of Data (SIGMOD'13)*, 2013.
- C27. Y. Chang, X. Wang, Q. Mei, Y. Liu. **Towards Twitter Context Summarization with User Influence Models.** *International Conference on Web Search and Data Mining (WSDM' 2013)*, 2013.
- C28. H. Qiu, Y. Liu, N. Subrahmanya, W. Li. **Granger Graphical Models for Time-Series Anomaly Detection.** *International conference on Data Mining (ICDM' 2012)*, 2012.
- C29. Y. Liu, T. Bahadori, H. Li. **Sparse-GEV: Sparse Latent Space Model for Multivariate Extreme Value Time Series Modeling.** *International conference on Machine Learning (ICML'2012)*, 2012.
- C30. S. Purushotham, Y. Liu, C-C. Kuo. **Collaborative Topic Regression with Social Matrix Factorization for Recommendation Systems.** *International conference on Machine Learning (ICML'2012)*, 2012.
- C31. W. Zhou, H. Jin, Y. Liu. **Community Discovery and Profiling with Social Messages.** *ACM Conf. on Knowledge Discovery & Data Mining (SIGKDD)*, 2012.
- C32. M. T. Bahadori, Y. Liu, **Causality Analysis in Irregular Time Series.** *SIAM Conference on Data Mining (SDM'12)*, 2012.
- C33. J.-H. Kang, J. Ma, Y. Liu. **Transfer Topic Modeling with Ease and Scalability.** *SIAM Conference on Data Mining (SDM'12)*, 2012.

- C34. D. Zhang, Y. Liu, S. Luo, J. Zhang and R. Lawrence. **Multiple Instance Learning on Structured Data.** *Proceedings of Twenty-Fifth Annual Conference on Neural Information Processing Systems (NIPS '11)*, 2011.
- C35. M. T. Bahadori, Y. Liu, D. Zhang. **Learning with Minimum Supervision: A General Framework for Transductive Transfer Learning.** *Proceedings of IEEE International Conference on Data Mining (ICDM' 11)*, 2011.
- C36. M. Haider, Y. Gil, Y. Liu. **A Framework for Efficient Text Analytics through Automatic Configuration and Customization of Scientific Workflows.** *Proceedings of e-Science*, 2011.
- C37. Y. Liu, P. Hseuh, R. Lawrence, S. Meliksetian, C. Perlich, A. Veen. **Latent Graphical Models for Quantifying and Predicting Patent Quality.** *Proceedings of the 17th ACM Conf. on Knowledge Discovery & Data Mining (SIGKDD)*, 2011.
- C38. D. Zhang, Y. Liu, Luo Si. **Serendipitous Learning: Learning Beyond the Predefined Label Space.** *Proceedings of the 17th ACM Conf. on Knowledge Discovery & Data Mining (SIGKDD)*, 2011.
- C39. D. Zhang, J. He, Y. Liu, L. Si, R. D. Lawrence. **Multi-View Transfer Learning with a Large Margin Approach.** *Proceedings of the 17th ACM Conf. on Knowledge Discovery & Data Mining (SIGKDD)*, 2011.
- C40. Y. Chang, R. Zhang, S. Reddy, Y. Liu. **Detecting Multilingual and Multi-Regional Query Intent in Web Search.** *Proceedings of the 25th AAAI Conference on Artificial Intelligence (AAAI)*, 2011.
- C41. D. Zhang, Y. Liu, R. D. Lawrence, V. Chenthamarakshan. **Transfer Latent Semantic Learning: Microblog Mining with Less Supervision.** *Twenty-Fifth AAAI Conference on Artificial Intelligence (AAAI)*, 2011.
- C42. Y. Liu, A. Niculescu-Mizil, A. Lozano, Y. Lu. **Learning Temporal Graphs for Relational Time-Series Analysis.** *International conference on Machine Learning (ICML' 10)*, 2010.
- C43. X. Chen, Y. Liu, H. Liu, J.G. Carbonell. **Learning Spatial-Temporal Varying Graphs with Applications to Climate Data Analysis.** *Proceedings of the Twenty-Fourth Conference on Artificial Intelligence (AAAI)*, 2010.
- C44. Y. Liu, A. Niculescu-Mizil, A. Lozano, Y. Lu. **Temporal Graphical Models for Cross-Species Gene Regulatory Network Discovery.** *International conference on Computational Systems Bioinformatics (CSB' 10)*, 2010.
- C45. Q. Wang, H. Jin. Y. Liu, **Collaboration Analytics: Mining Work Patterns from Collaboration Activities.** *To appear in International Conference on Information and Knowledge Management (CIKM' 10)*, 2010.
- C46. Y. Liu, A. Niculescu-Mizil, W. Gryc. **Topic-Link LDA: Joint Models of Topic and Author Community.** *International conference on Machine Learning (ICML' 09)*, 2009.
- C47. Y. Liu, J.R. Kalagnanam, O. Johnsen. **Learning Dynamic Temporal Graphs for Oil-production Equipment Monitoring System.** *International Conference on Knowledge Discovery and Data Mining (KDD' 09)*, 2009.
- C48. J. He, Y. Liu, R. Lawrence. **Graph-based Transfer Learning.** *Proc. of ACM Conference on Information and Knowledge Management (CIKM'09)*, 2009.
- C49. A. Lozano, H. Li, A. Niculescu-Mizil, Y. Liu, C. Perlich, J. Hosking, N. Abe. **Spatial-temporal causal modeling for climate change attribution.** *International Conference on Knowledge Discovery and Data Mining (KDD' 09)*, 2009.
- C50. A.C. Lozano, N. Abe, Y. Liu, S. Rosset. **Grouped Graphical Granger Modeling Methods for Temporal Causal Modeling.** *International Conference on Knowledge Discovery and Data Mining (KDD' 09)*, 2009.

- C51. A.C. Lozano, N. Abe, Y. Liu, S. Rosset. **Grouped Graphical Granger Modeling for Gene Expression Regulatory Networks Discovery.** *International Conference on Intelligent Systems for Molecular Biology (ISMB'09)*, 2009.
- C52. Y. Liu, P. Hsueh, J. Lai, M. Sangin, M.A. Nussli, P. Dillenbourg. **Who is the Expert? Analyzing Gaze Data to Predict Expertise Level in Collaborative Applications.** *International Conference on Multimedia & Expo (ICME'09)*, 2009.
- C53. T. Ide, A. Lozano, N. Abe, Y. Liu. **Proximity-Based Anomaly Detection using Sparse Structure Learning.** *SIAM Conference on Data Mining (SDM'09)*, 2009.
- C54. J. He, Y. Liu, R. Lawrence. **Graph-based Rare Category Detection.** *International Conference on Data Mining (ICDM'08)*, 2008.
- C55. A. Arnold, Y. Liu, N. Abe. **Temporal Causal Modeling with Graphical Granger Methods.** *ACM Conference on Knowledge Discovery and Data Mining (KDD'07)*, 2007.
- C56. Y. Liu, J. Carbonell, V. Gopalakrishnan, P. Weigele. **Protein Quaternary Fold Recognition Using Conditional Graphical Models.** *International Joint Conference in Artificial Intelligence (IJCAI'07)*, 2007.
- C57. J. He, J. Carbonell, Y. Liu. **Graph-Based Semi-Supervised Learning as a Generative Model.** *International Joint Conference in Artificial Intelligence (IJCAI'07)*, 2007.
- C58. K. Probst, R. Ghani, M. Krema, A. Fano, Y. Liu. **Semi-supervised Learning of Attribute-Value Pairs from Product Descriptions.** *International Joint Conference in Artificial Intelligence (IJCAI'07)*, 2007.
- C59. J. Yang, Y. Liu, E. P. Xing, A. Hauptmann. **Harmonium Models for Semantic Video Representation and Classification.** *SIAM Conference on Data Mining (SDM'07)*, 2007.
- C60. Y. Liu, J. Yang, A. Hauptmann. **Undirected Graphical Models for Video Analysis and Classification.** *International Conference on Multimedia & Expo (ICME'07)*, 2007.
- C61. Y. Liu, E. Xing, J. Carbonell. **Predicting Protein Folds with Structural Repeats Using a Chain Graph Model.** *International conference on Machine Learning (ICML'05)*, 2005.
- C62. Y. Liu, J. Carbonell, P. Weigele, V. Gopalakrishnan. **Segmentation Conditional Random Fields (SCRFs): A New Approach for Protein Fold Recognition.** *ACM Conference on Research in Computational Molecular Biology (RECOMB'05)*, 2005.
- C63. Y. Liu, J. Carbonell, J. K. Seetharaman, V. Gopalakrishnan. **Context Sensitive Vocabulary and Its Application in Protein Secondary Structure Prediction.** *ACM International Conference on Research and Development in Information Retrieval (SIGIR'04)*, 2004.
- C64. J. Lafferty, X. Zhu, Y. Liu. **Kernel Conditional Random Fields: Representation and Clique Selection.** *International Conference on Machine Learning (ICML'04)*, 2004.
- C65. Y. Liu, J. Carbonell, J. K. Seetharaman, V. Gopalakrishnan. **Prediction of Parallel and Antiparallel Beta-sheets using Conditional Random Fields.** *Biological Language Conference (BLC'03)*, 2003.
- C66. Y. Liu, J. Carbonell, R. Jin. **A New Pairwise Ensemble Approach for Text Classification.** *Proc. of 14th European Conference on Machine Learning (ECML'03)*, 2003.
- C67. R. Jin, Y. Liu, S. Luo, J. Carbonell, A. Hauptmann. **A New Boosting Algorithm Using Input-Dependent Regularizer.** *Proc. of 20th International Conference on Machine Learning (ICML'03)*, 2003.
- C68. R. Yan, Y. Liu, R. Jin, A. Hauptmann. **On Predicting Rare Class with SVM Ensemble in Scene Classification.** *Proc. of IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)*, 2003.
- C69. Y. Liu, Y. Yang, J. Carbonell. **Boosting to Correct the Inductive Bias for Text Classification.** *Proc. of ACM Conference on Information and Knowledge Management (CIKM'02)*, 2002.

PATENTS

- P1. J. He, R. Lawrence, Y. Liu. A System and Method for Graph-based Transfer Learning. US patent files ad Docket YOR9-2009-0424-US2.
- P2. N. Duan, P.S. Hsueh, Y. Liu. Systems and Methods for Identifying Content-Sensitive Authorities from Very Large Scale Networks. US patent files ad Docket YOR9-2009-0630-US1.
- P3. W. Gryc, R. Lawrence, Y. Liu. System and Method for Extracting Patterns from Graph and Unstructured Data. US patent files ad Docket YOR8-2008-0976-US1.
- P4. B. Dietrich J. Lai, Y. Liu. Method and System for Controlling Skill Acquisition. US patent filed as Docket YOR9-2008-0048-US1.
- P5. S. Carty, Y. Liu, T. Kumar, G. Parijia. A Goal Programming Approach for Optimal Budget Allocation for National Analysis of Wildland Fire Management. US patent filed as Docket YOR9-2007-0755-US1.
- P6. U. Chitnis, T. Bowden, Khabibrakhmanov, R. Lawrence, Y. Liu, P. Melville. A Method and System for Identifying Companies with Specific Business Objectives. US patent filed as Docket YOR9-2007-0060-US1.
- P7. A. Fano, R. Ghani, M. Krema, Y. Liu, K. Probst. Extraction of Attributes and Values from Natural Language Documents. US patent filed as Docket No. 33836.00.0145.

PROFESSIONAL SERVICES

**Workshop Co-organizer**

- KDD Workshop on Time Series Learning and Mining, 2015-2016
- 4th International Workshop on Climate Informatics, 2015
- AAAI Discovery Informatics Symposium, 2013
- ICML Workshop on Machine Learning with Workflows, 2013
- KDD Workshop on Large-scale Data Mining: Theory and Applications (LDMTA), 2010-2012
- ICDM Workshop on Large-scale Analytics for Complex Instrumented Systems (LACIS), 2010
- ICDM Workshop on Large-scale Data Mining: Theory and Applications (LDMTA), 2009
- IBM Workshop on Statistical Learning and its Application (SMILE), 2009

**Conference Organizing Committee**

- Publicity Chair of SDM 2013
- Poster Chair of KDD 2012
- Local Chair of SDM 2012
- Publicity Chair of KDD 2011

**Tutorials**

- Challenges and Solutions: Social Media Anomaly Detection. KDD 2015.
- Causality Analysis in Large-scale Time Series Data. CIKM 2013.

**Program Co-chair**

- Third International Workshop on Climate Informatics, 2013
- High Performance Computing Conference (HiPC) Workshop on Massive Data Analytics on Scalable Systems, 2012

**Senior Program Committee (or Area Chair)**

- KDD 2016, SDM 2016, CIKM 2015, WSDM 2015, ICML 2014, WSDM 2013

**Program Committee (or equivalent)**

- KDD (2008, 2010, 2012-2015); NIPS (2009, 2010, 2014-2016); ICML (2009, 2011, 2014-2015); SDM (2010-2014, 2017); AAAI (2010-2013, 2015-2016); IJCAI (2013, 2015-2016); UAI (2015); AI-Stat (2016-2017); ECML (2011, 2014, 2015); CIKM (2008-2011, 2013-2014); ICDM (2009-2014); EMNLP (2010-2011); WWW (2011); COLING (2010); SIGIR (2009); PAKDD (2012); NAACL (2012); ACL (2012); WSDM (2012, 2013-2014); IEEE BigData (2013-2015)

ICDM Workshop on Mining Graphs and Complex Structures, 2007; NIPS workshop on Structured Input and Structured Output, 2008; NIPS workshop on cost-sensitive learning, 2008; ICDM International Workshop on Knowledge Discovery Using Cloud and Distributed Computing Platforms, 2011; KDD SNA Workshop, 2012; KDD Workshop on Cross Domain Knowledge Discovery in Web and Social Network Mining, 2012; AMIA Summit on Translational Bioinformatics, 2012

**Guest Editor**

TKDD Special Issue on Large-Scale Data Mining

**Reviewer**

Journal of Machine Learning Research (JMLR), Transactions on Knowledge Discovery and Data Mining (TKDD), Transactions on Knowledge and Data Engineering (TKDE), Journal of Artificial Intelligence Research (JAIR), Artificial Intelligence, Annals of Applied Statistics, Statistical Analysis and Data Mining, Information Retrieval, Journal of Intelligent Information Systems (JIIS), ACM Transactions on Information Systems, Information Processing & Management (IP&M), Bioinformatics., Journal of Computational Biology (JCB), FEBS letters, Concurrency and Computation: Practice and Experience, Advances in Data Analysis and Classification (ADAC)