Jing Gao

Assistant Professor
Department of Computer Science and Engineering
State University of New York at Buffalo
350 Davis Hall, Buffalo, NY 14260
jing@buffalo.edu
Tel 716-645-1586
http://www.cse.buffalo.edu/~jing

EDUCATION

Ph.D. 2011, Computer Science, University of Illinois, Urbana-Champaign, IL

Thesis: Exploring the Power of Heterogeneous Information Sources

Advisor: Jiawei Han

M.Eng. 2004, Computer Science, Harbin Institute of Technology, Harbin, China

Thesis: Algorithms and Techniques in Data Stream Mining

Advisor: Jianzhong Li

B.Eng. 2002, Computer Science, Harbin Institute of Technology, Harbin, China

EMPLOYMENT

State University of New York at Buffalo
Assistant Professor, *Department of Computer Science and Engineering*, January 2012-present

IBM TJ Watson Research Center
Research Intern, *Exploratory Stream Analytics Group*, Summer 2009/2010/2011

NEC Labs America
Research Intern, *Robust and Secure Systems Department*, Summer 2008

IBM TJ Watson Research Center
Research Intern, *Predictive Modeling Group*, Summer 2007

RESEARCH INTERESTS

I am broadly interested in data and information analysis with a focus on data mining. In particular, I am interested in information integration, anomaly detection, semi-supervised and unsupervised learning, transfer learning, mining data streams and information networks. I am also interested in various data mining applications in cyber security, health care, bioinformatics and social media.

CITATION H-INDEX

According to Google Scholar (http://scholar.google.com/citations?user=Ftj1h4cAAAAJ&hl=en)

- Citation is 1662 (as of 1/16/2015)

- H-index is 25

PUBLICATIONS

Conferences

VLDB15	Qi Li, Yaliang Li, **Jing Gao**, Lu Su, Bo Zhao, Murat Demirbas, Wei Fan, Jiawei Han. *A Confidence-Aware Approach for Truth Discovery on Long-Tail Data.* Proceedings of the 41st International Conference on Very Large Data Bases (VLDB 2015), Kohala Coast, HI, August 2015, 8(4): 425-436.
SDM15	Houping Xiao, Yaliang Li, **Jing Gao**, Fei Wang, Liang Ge, Wei Fan, Long Vu, Deepak Turaga, *Believe It Today or Tomorrow? Detecting Untrustworthy Information from Dynamic Multi-Source Data.* Proceedings of the 15th SIAM International Conference on Data Mining (SDM 2015), Vancouver, Canada, April 2015, to appear.
SDM15	Bowen Dong, Sihong Xie, **Jing Gao**, Wei Fan, Philip Yu, *OnlineCM: Real-time Consensus Classification with Missing Values.* Proceedings of the 15th SIAM International Conference on Data Mining (SDM 2015), Vancouver, Canada, April 2015, to appear.
SDM15	Kang Li, **Jing Gao**, Suxin Guo, Nan Du, Aidong Zhang, *Functional Node Detection on Linked Data.* Proceedings of the 15th SIAM International Conference on Data Mining (SDM 2015), Vancouver, Canada, April 2015, to appear.
SDM15	Jialu Liu, Chi Wang, Quanquan Gu, Jiawei Han, **Jing Gao**, Charu Aggarwal, Lance Kaplan, *GIN: A Unified Model for Capturing Dual Heterogeneity in Clustering Networked Data.* Proceedings of the 15th SIAM International Conference on Data Mining (SDM 2015), Vancouver, Canada, April 2015, to appear.
ICDM14	Kang Li, **Jing Gao**, Suxin Guo, Nan Du, Xiaoyi Li, Aidong Zhang, *LRBM: A Restricted Boltzmann Machine based Approach for Representation Learning on Linked Data.* Proceedings of the 14th IEEE International Conference on Data Mining (ICDM 2014), Shenzhen, China, December 2014. (acceptance rate: 9.7%)
RTSS14	Lu Su, Qi Li, Shaohan Hu, Shiguang Wang, **Jing Gao**, Hengchang Liu, Tarek Abdelzaher, Jiawei Han, Xue Liu, Yan Gao, Lance Kaplan, *Generalized Decision Aggregation in Distributed Sensing Systems.* Proceedings of the 14th IEEE IEEE Real-Time Systems Symposium (RTSS 2014), Rome, Italy, December 2014.
CIKM14	Xiaowei Jia, Nan Du, **Jing Gao**, Aidong Zhang, *Analysis on Community Variational Trend in Dynamic Networks.* Proceedings of the 23rd ACM International Conference on Information and Knowledge Management (CIKM 2014), Shanghai, China, November 2014, 151-160.
KDD14	Sihong Xie, **Jing Gao**, Wei Fan, Deepak Turaga, Philip S. Yu, *Class-Distribution Regularized Consensus Maximization for Alleviating Overfitting in Model Combination.* Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2014), New York, NY, August 2014, 303-312. (acceptance rate: 151/1036 = 14.6%)
IAAI14	Bahadir Aydin, Yavuz Yilmaz, Yaliang Li, Qi Li, **Jing Gao**, Murat Demirbas, *Crowdsourcing for Multiple-Choice Question Answering.* Proceedings of Annual Conference on Innovative Applications of Artificial Intelligence (IAAI 2014), Quebec City, Canada, July 2014, 2946-2953.
SIGMOD14	Qi Li, Yaliang Li, **Jing Gao**, Bo Zhao, Wei Fan, Jiawei Han. *Resolving Conflicts in Heterogeneous Data by Truth Discovery and Source Reliability Estimation.* Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD 2014), Snowbird, UT, June 2014, 1187-1198. (acceptance rate: 107/421 = 25%)
SDM14	Xiaoyi Li, Nan Du, Hui Li, Kang Li, **Jing Gao**, Aidong Zhang, *A Deep Learning Approach to Link Prediction in Dynamic Networks.* Proceedings of the 14th SIAM International Conference on Data Mining (SDM 2014), Philadelphia, PA, April 2014, 289-297. (acceptance rate: 60/389 = 15%)

SDM14	Ke Wu, Andrea Edwards, Wei Fan, **Jing Gao**, Kun Zhang, *Classifying Imbalanced Data Streams via Dynamic Feature Group Weighting with Importance Sampling.* Proceedings of the 14th SIAM International Conference on Data Mining (SDM 2014), Philadelphia, PA, April 2014, 722-730. (acceptance rate: 120/389 = 29%)
ICDE14	Manish Gupta, **Jing Gao**, Xifeng Yan, Hasan Cam, Jiawei Han, *Top-K Interesting Subgraph Discovery in Information Networks.* Proceedings of the 30th IEEE International Conference on Data Engineering (ICDE 2014), Chicago, IL, March 2014, 820-831. (acceptance rate: 89/446 = 19.96%)
ICDM13	Kang Li, Suxin Guo, Nan Du, **Jing Gao**, Aidong Zhang, *Learning, Analyzing and Predicting Object Roles on Dynamic Networks.* Proceedings of the 13th IEEE International Conference on Data Mining (ICDM 2013), Dallas, TX, December 2013, 428-437. (acceptance rate: 94/809 = 11.62%)
ICDM13	Nan Du, **Jing Gao**, Aidong Zhang, *Progression Analysis of Community Strengths in Dynamic Networks.* Proceedings of the 13th IEEE International Conference on Data Mining (ICDM 2013), Dallas, TX, December 2013, 1031-1036. (acceptance rate: 159/809 = 19.65%)
ICDM13	Sihong Xie, Xiangnan Kong, **Jing Gao**, Wei Fan, Philip S. Yu, *Multilabel Consensus Classification.* Proceedings of the 13th IEEE International Conference on Data Mining (ICDM 2013), Dallas, TX, December 2013, 1241-1246. (acceptance rate: 159/809 = 19.65%)
CIKM13	Liang Ge, **Jing Gao**, Aidong Zhang, *OMS-TL: A Framework of Online Multiple Source Transfer Learning.* Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (CIKM 2013), San Francisco, CA, October 2013, 2423-2428. (acceptance rate: 250/848 = 29.48%)
CIKM13	Xiaoyi Li, **Jing Gao**, Hui Li, Le Yang, Rohini K. Srihari, *A Multimodal Framework for Unsupervised Feature Fusion.* Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (CIKM 2013), San Francisco, CA, October 2013, 897-902. (acceptance rate: 250/848 = 29.48%)
BCB13	Kang Li, Suxin Guo, **Jing Gao**, Aidong Zhang, *An Ensemble Model for Mobile Device based Arrhythmia Detection.* Proceedings of the 4th ACM Conference on Bioinformatics, Computational Biology and Biomedical Informatics (BCB 2013), Washington, DC, September 2013, 266-275. (acceptance rate: 43/148 = 29%)
ECML13	Manish Gupta, **Jing Gao**, Jiawei Han, *Community Distribution Outlier Detection in Heterogeneous Information Networks.* Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2013), Prague, Czech Republic, September 2013, 557-573. (acceptance rate: 111/443 = 25%)
ECML13	Xian Wu, Ziming Feng, Wei Fan, **Jing Gao**, Yong Yu, *Detecting Marionette Microblog Users for Improved Information Credibility.* Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2013), Prague, Czech Republic, September 2013, 483-498. (acceptance rate: 111/443 = 25%)
ASONAM13	Manish Gupta, **Jing Gao**, Xifeng Yan, Hasan Cam, Jiawei Han, *On Detecting Association-Based Clique Outliers in Heterogeneous Information Networks.* Proceedings of the IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM 2013), Niagara Falls, Canada, August 2013, 108-115. (acceptance rate: 15%)
KDD13	Liang Ge, **Jing Gao**, Xiaoyi Li, Aidong Zhang, *Multi-Source Deep Learning for Information Trustworthiness Estimation.* Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2013), Chicago, IL, August 2013, 766-774. (acceptance rate: 126/726 = 17.4%)

SDM13   Liang Ge, **Jing Gao**, Hung Q. Ngo, Kang Li, Aidong Zhang, *On Handling Negative Transfer and Imbalanced Distributions in Multiple Source Transfer Learning.* Proceedings of the 13th SIAM International Conference on Data Mining (SDM 2013), Austin, TX, May 2013, 261-269. (acceptance rate: 50/348 = 14.4%)

SDM13   Jialu Liu, Chi Wang, **Jing Gao**, Jiawei Han, *Multi-view Clustering via Joint Nonnegative Matrix Factorization.* Proceedings of the 13th SIAM International Conference on Data Mining (SDM 2013), Austin, TX, May 2013, 252-260. (acceptance rate: 50/348 = 14.4%)

ICDM12   Liang Ge, **Jing Gao**, Xiao Yu, Wei Fan, Aidong Zhang, *Estimating Local Information Trustworthiness via Multi-Source Joint Matrix Factorization.* Proceedings of the 12th IEEE International Conference on Data Mining (ICDM 2012), Brussels, Belgium, December 2012, 876-881. (acceptance rate: 20%)

RTSS12   Lu Su, Shaohan Hu, Shen Li, Feng Liang, **Jing Gao**, Tarek Abdelzaher, Jiawei Han, *Quality of Information based Data Selection and Transmission in Wireless Sensor Networks.* Proceedings of the 33rd IEEE Real-time Systems Symposium (RTSS 2012), San Juan, Puerto Rico, December 2012, 327-338. (acceptance rate: 22%)

BCB12   Nan Du, Yuan Zhang, Kang Li, **Jing Gao**, Aidong Zhang, *Evolutionary Analysis of Functional Modules in Dynamic PPI Networks.* Proceedings of the 3rd ACM Conference on Bioinformatics, Computational Biology and Biomedicine (BCB 2012), Orlando, Florida, October 2012, 250-257. (acceptance rate: 33/159 = 20.75%)

BCB12   Liang Ge, **Jing Gao**, Nan Du, Aidong Zhang, *Finding Informative Genes for Prostate Cancer: A General Framework of Integrating Heterogeneous Sources.* Proceedings of the 3rd ACM Conference on Bioinformatics, Computational Biology and Biomedicine (BCB 2012), Orlando, Florida, October 2012, 274-281. (acceptance rate: 64/159 = 40.25%)

BIBM12   Nan Du, **Jing Gao**, Vishrawas Gopalakrishna, Aidong Zhang, *De-noise Biological Network from Heterogeneous Sources via Link Propagation.* Proceedings of the 6th IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2012), Philadelphia, PA, October 2012, 1-6. (acceptance rate: 59/299 = 19.93%)

ASIS&T12   Manish Gupta, **Jing Gao**, ChengXiang Zhai, Jiawei Han. *Predicting Future Popularity Trend of Events in Microblogging Platforms.* Proceedings of the 75th Annual Meeting of American Society for Information Science & Technology (ASIS&T 2012), Baltimore, MD, October 2012, 49(1):1-10.

ECML12   Manish Gupta, **Jing Gao**, Yizhou Sun, Jiawei Han, *Community Trend Outlier Detection Using Soft Temporal Pattern Mining.* Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2012), Bristol, UK, September 2012, 692-708. (acceptance rate: 105/443 = 23.7%)

KDD12   Manish Gupta, **Jing Gao**, Yizhou Sun, Jiawei Han, *Integrating Community Matching and Outlier Detection for Mining Evolutionary Community Outliers.* Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2012), Beijing, China, August 2012, 859-867. (acceptance rate: 133/755 = 17.6%)

ICDM11   **Jing Gao**, Wei Fan, Deepak Turaga, Srinivasan Parthasarathy, Jiawei Han, *A Spectral Framework for Detecting Inconsistency across Multi-Source Object Relationships.* Proceedings of the 11th IEEE International Conference on Data Mining (ICDM 2011), Vancouver, Canada, December 2011, 1050-1055. (acceptance rate: 148/822 = 18%)

ICDM11   Mohammad Masud, Tahseen Al-Khateeb, Latifur Khan, Charu Aggarwal, **Jing Gao**, Jiawei Han, *Detecting Recurring and Novel Classes in Concept-Drifting Data Streams.* Proceedings of the 11th IEEE International Conference on Data Mining (ICDM 2011), Vancouver, Canada, December 2011, 1176-1181. (acceptance rate: 148/822 = 18%)

SenSys11   Lu Su, Yong Yang, Bolin Ding, **Jing Gao**, Tarek F. Abdelzaher, Jiawei Han, *Hierarchical Aggregate Classification with Limited Supervision for Data Reduction in Wireless Sensor Networks*. Proceedings of the 9th ACM International Conference on Embedded Networked Sensor Systems (SenSys 2011), Seattle, WA, November 2011, 40-53. (acceptance rate: 24/123 = 19.5%)

INFOCOM11   **Jing Gao**, Wei Fan, Deepak Turaga, Olivier Verscheure, Xiaoqiao Meng, Lu Su, Jiawei Han, *Consensus Extraction from Heterogeneous Detectors to Improve Performance over Network Traffic Anomaly Detection*. Proceedings of the 30th IEEE International Conference on Computer Communications (INFOCOM 2011 Mini-Conference), Shanghai, China, April 2011, 181-185. (acceptance rate: 427/1823 = 23.4%)

ICDM10   Mohammad Masud, Qing Chen, Latifur Khan, Charu Aggarwal, **Jing Gao**, Jiawei Han, Bhavani Thuraisingham, *Addressing Concept-Evolution in Concept-Drifting Data Streams*. Proceedings of the 10th IEEE International Conference on Data Mining (ICDM 2010), Sydney, Australia, December 2010, 929-934. (acceptance rate: 155/797 = 19.4%)

ECML10   Ming Ji, Yizhou Sun, Marina Danilevsky, Jiawei Han, and **Jing Gao**, *Graph Regularized Transductive Classification on Heterogeneous Information Networks*. Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2010), Barcelona, Spain, September 2010, 570-586. (acceptance rate: 120/658 = 18.2%)

ECML10   Mohammad Masud, Qing Chen, **Jing Gao**, Latifur Khan, Jiawei Han, Bhavani Thuraisingham, *Classification and Novel Class Detection of Data Streams in a Dynamic Feature Space*. Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2010), Barcelona, Spain, September 2010, 337-352. (acceptance rate: 120/658 = 18.2%)

KDD10   **Jing Gao**, Feng Liang, Wei Fan, Chi Wang, Yizhou Sun, Jiawei Han, *On Community Outliers and their Efficient Detection in Information Networks*. Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2010), Washington, DC, July 2010, 813-822. (full presentation, acceptance rate: 77/578 = 13.3%)

PAKDD10   Mohammad Masud, **Jing Gao**, Latifur Khan, Jiawei Han, Bhavani Thuraisingham, *Classification and Novel Class Detection in Data Streams with Active Mining*. Proceedings of the 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2010), Hyderabad, India, June 2010, 311-324. (full presentation, acceptance rate: 42/412 = 10.2%)

NIPS09   **Jing Gao**, Feng Liang, Wei Fan, Yizhou Sun, Jiawei Han, *Graph-based Consensus Maximization among Multiple Supervised and Unsupervised Models*. Advances in Neural Information Processing Systems (NIPS), 22, 2009, 585-593. (acceptance rate: 263/1105 = 23.8%)

ICDM09   Yizhou Sun, Jiawei Han, **Jing Gao**, Yintao Yu, *iTopicModel: Information Network-Integrated Topic Modeling*. Proceedings of the 9th IEEE International Conference on Data Mining (ICDM 2009), Miami, FL, December 2009, 493-502. (full presentation, acceptance rate: 70/786 = 8.9%)

MM09   Chandrasekar Ramachandran, Rahul Malik, Xin Jin, **Jing Gao**, Klara Nahrstedt, Jiawei Han, *VideoMule: A Consensus Learning Approach to Multi-Label Classification from Noisy User-Generated Videos*. Proceedings of the ACM International Conference on Multimedia (ACM MM 2009), Beijing, China, October 2009, 721-724. (acceptance rate: 58/204 = 28%)

ECML09   Mohammad Masud, **Jing Gao**, Latifur Khan, Jiawei Han, Bhavani Thuraisingham, *Integrating Novel Class Detection with Classification for Concept-Drifting Data Streams*. Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2009), Bled, Slovenia, September 2009, 79-94. (acceptance rate: 105/422 = 24.9%)

KDD09   **Jing Gao**, Wei Fan, Yizhou Sun, Jiawei Han, *Heterogeneous Source Consensus Learning via Decision Propagation and Negotiation.* Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2009), Paris, France, June 2009, 339-347. (acceptance rate: 112/538 = 20.8%)

ICDCS09   **Jing Gao**, Guofei Jiang, Haifeng Chen, Jiawei Han, *Modeling Probabilistic Measurement Correlations for Problem Determination in Large-Scale Distributed Systems.* Proceedings of the 29th International Conference on Distributed Computing Systems (ICDCS 2009), Montreal, Canada, June 2009, 623-630. (full presentation, acceptance rate: 74/455 = 16%)

PAKDD09   Mohammad Masud, **Jing Gao**, Latifur Khan, Jiawei Han, Bhavani Thuraisingham, *A Multi-Partition Multi-Chunk Ensemble Technique to Classify Concept-Drifting Data Streams.* Proceedings of the 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2009), Bangkok, Thailand, April 2009, 363-375. (full presentation, acceptance rate: 39/338 = 11.5%)

ICDM08   Mohammad Masud, **Jing Gao**, Latifur Khan, Jiawei Han, *A Practical Approach to Classify Evolving Data Streams: Training with Limited Amount of Labeled Data.* Proceedings of the 8th IEEE International Conference on Data Mining (ICDM 2008), Pisa, Italy, December 2008, 929-934. (acceptance rate: 144/724 = 19.9%)

KDD08   **Jing Gao**, Wei Fan, Jing Jiang, Jiawei Han, *Knowledge Transfer via Multiple Model Local Structure Mapping.* Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2008), Las Vegas, NV, August 2008, 283-291. (acceptance rate: 95/510 = 18.6%)

KDD08   Wei Fan, Kun Zhang, Hong Cheng, **Jing Gao**, Xifeng Yan, Jiawei Han, Philip S. Yu, Olivier Verscheure, *Direct Mining of Discriminative and Essential Graphical and Itemset Features via Model-based Search Tree.* Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2008), Las Vegas, NV, August 2008, 230-238. (full presentation, acceptance rate: 50/510 = 9.8%)

ICDM07   **Jing Gao**, Wei Fan, Jiawei Han, *On Appropriate Assumptions to Mine Data Streams: Analysis and Practice.* Proceedings of the 7th IEEE International Conference on Data Mining (ICDM 2007), Omaha, NE, October 2007, 143-152. (full presentation, acceptance rate: 40/573 = 7.5%)

SDM07   **Jing Gao**, Wei Fan, Jiawei Han, Phillip S. Yu, *A General Framework for Mining Concept-Drifting Data Streams with Skewed Distributions.* Proceedings of the 7th SIAM International Conference on Data Mining (SDM 2007), Minneapolis, MN, April 2007, 3-14. (full presentation, acceptance rate: 36/302 = 11.9%)

ICDM06   **Jing Gao**, Pang-Ning Tan, *Converting Output Scores from Outlier Detection Algorithms into Probability Estimates.* Proceedings of the 6th IEEE International Conference on Data Mining (ICDM 2006), Hong Kong, China, December 2006, 212-221. (full presentation, acceptance rate: 73/776 = 9.4%)

SDM06   **Jing Gao**, Pang-Ning Tan, Haibin Cheng, *Semi-Supervised Clustering with Partial Background Information.* Proceedings of the 6th SIAM International Conference on Data Mining (SDM 2006), Bethesda, MD, April 2006, 489-493.

SDM06   **Jing Gao**, Haibin Cheng, Pang-Ning Tan, *A Novel Framework for Incorporating Labeled Examples into Anomaly Detection.* Proceedings of the 6th SIAM International Conference on Data Mining (SDM 2006), Bethesda, MD, April 2006, 594-598.

PAKDD06   Haibin Cheng, Pang-Ning Tan, **Jing Gao**, Jerry Scripps, *An Empirical Study on Multistep-ahead Time Series Prediction.* Proceedings of the 10th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2006), Singapore, April 2006, 765-774.

SAC06	**Jing Gao**, Haibin Cheng, Pang-Ning Tan, *Semi-Supervised Outlier Detection.* Proceedings of the 2006 ACM Symposium on Applied Computing (SAC 2006), Dijon, France, April 2006, 635-636.
PAKDD05	**Jing Gao**, Jianzhong Li, Zhaogong Zhang, Pang-Ning Tan, *An Incremental Data Stream Clustering Algorithm Based on Dense Units Detection.* Proceedings of the 9th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2005), Hanoi, Vietnam, May 2005, 420-425.

Books and Book Chapters

MC14	Manish Gupta, **Jing Gao**, Charu Aggarwal, Jiawei Han, *Outlier Detection for Temporal Data.* Morgan & Claypool, Synthesis Lectures on Data Mining and Knowledge Discovery, 2014.
DC14	Yaliang Li, **Jing Gao**, Qi Li, Wei Fan, *Ensemble Learning.* in C. Aggarwal (eds.), Data Classification: Algorithms and Applications, Taylor & Francis, 2014, 483-510.
GEPA12	**Jing Gao**, Nan Du, Wei Fan, Deepak Turaga, Srinivasan Parthasarathy, Jiawei Han, *A Multi-graph Spectral Framework for Mining Multi-source Anomalies.* in Y. Fu, et al., (eds.), Graph Embedding for Pattern Analysis, Springer, 2012, 205-228.
NGDM09	Jiawei Han, **Jing Gao**, *Research Challenges for Data Mining in Science and Engineering.* in H. Kargupta, et al., (eds.), Next Generation of Data Mining, Chapman & Hall, 2009, 3-28.

Journals and Magazines

TKDE14	Manish Gupta, **Jing Gao**, Charu Aggarwal, Jiawei Han:, *Outlier Detection for Temporal Data: A Survey.* IEEE Transactions on Knowledge and Data Engineering (TKDE), 26(9): 2250-2267, September 2014.
SAM14	Liang Ge, **Jing Gao**, Hung Q. Ngo, Kang Li, Aidong Zhang, *On Handling Negative Transfer and Imbalanced Distributions in Multiple Source Transfer Learning.* Statistical Analysis and Data Mining (SAM), 7(4): 254-271, April 2014.
TKDE13	Mohammad M. Masud, Qing Chen, Latifur Khan, Charu C. Aggarwal, **Jing Gao**, Jiawei Han, Ashok Srivastava, Nikunj C. Oza, *Classification and Adaptive Novel Class Detection of Feature-Evolving Data Streams.* IEEE Transactions on Knowledge and Data Engineering (TKDE), 25(7): 1484-1497, July 2013.
TKDE13	**Jing Gao**, Feng Liang, Wei Fan, Yizhou Sun, Jiawei Han, *A Graph-based Consensus Maximization Approach for Combining Multiple Supervised and Unsupervised Models.* IEEE Transactions on Knowledge and Data Engineering (TKDE), 25(1): 15-28, January 2013.
KAIS11	Mohammad Masud, Clay Woolam, **Jing Gao**, Latifur Khan, Jiawei Han, Kevin W. Hamlen, Nikunj C. Oza, *Facing the Reality of Data Stream Classification: Coping with Scarcity of Labeled Data.* Knowledge and Information Systems, 33(1): 213-244, October 2011.
TMIS11	Mohammad M. Masud, Tahseen Al-Khateeb, Kevin W. Hamlen, **Jing Gao**, Latifur Khan, Jiawei Han, Bhavani M. Thuraisingham, *Cloud-based Malware Detection for Evolving Data Streams.* ACM Transactions on Management Information Systems, 2(3): 16, October 2011.
TKDE11	Mohammad Masud, **Jing Gao**, Latifur Khan, Jiawei Han, Bhavani Thuraisingham, *Classification and Novel Class Detection in Concept-Drifting Data Streams under Time Constraints.* IEEE Transactions on Knowledge and Data Engineering (TKDE), 23(6): 859-874, June 2011.

IEEEIC08   **Jing Gao**, Bolin Ding, Wei Fan, Jiawei Han, Philip S. Yu, *Classifying Data Streams with Skewed Class Distributions and Concept Drifts.* IEEE Internet Computing, Special Issue on Data Stream Management, 12(6): 37-49, Nov./Dec. 2008.

JCRD04   **Jing Gao**, Jianzhong Li, Zhaogong Zhang, *Mining Maximal Frequent Patterns over Numerical Data Streams.* Journal of Computer Research and Development, Vol. 41, Suppl.: 282-288, 2004. (in Chinese)

JCS03   **Jing Gao**, Jianzhong Li, Zhaogong Zhang, Yanqiu Zhang, *An Efficient Data Stream Clustering Algorithm Based on Grids and Density.* Journal of Computer Science, Vol. 30, Suppl.: 256-260, 2003. (in Chinese)

Tutorials

CIKM13   Manish Gupta, **Jing Gao**, Charu Aggarwal, Jiawei Han *Outlier Detection for Temporal Data.* The 22nd ACM International Conference on Information and Knowledge Management (CIKM 2013), San Francisco, CA, October 2013.

SDM13   Manish Gupta, **Jing Gao**, Charu Aggarwal, Jiawei Han *Outlier Detection for Temporal Data.* The 13th SIAM International Conference on Data Mining (SDM 2013), Austin, TX, May 2013.

PAKDD11   Latifur Khan, Wei Fan, Jiawei Han, **Jing Gao**, Mohammad M. Masud, *Data Stream Mining: Challenges and Techniques.* The 15th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2011), Shenzhen, China, May 2011.

SDM10   **Jing Gao**, Wei Fan, Jiawei Han, *On the Power of Ensemble: Supervised and Unsupervised Methods Reconciled.* The 10th SIAM International Conference on Data Mining (SDM 2010), Columbus, Ohio, May 2010.

TEACHING EXPERIENCE

State University of New York at Buffalo
Instructor

CSE722: Selected Topics in Data Mining, Spring 2012, Spring 2013, Spring 2014, Spring 2015

CSE601: Data Mining and Bioinformatics, Fall 2012, Fall 2013, Fall 2014

UE141: Discovery Seminar on Data Mining, Spring 2013

University of Illinois at Urbana-Champaign
Teaching Assistant

CS512: Data Mining: Principles and Algorithms, Spring 2007

CS101: Introduction to Computing, Fall 2006

HONORS and AWARDS

2013   IBM Faculty Award

2013   One of Best 10 papers of SIAM International Conference on Data Mining (SDM), Austin, TX Selected to be included in a special issue of Statistical Analysis and Data Mining (SAM)

2010-2011   IBM Ph.D. Fellowship

PRESENTATIONS

- Inferring Information Trustworthiness from Multiple Sources of Heterogeneous Data
  Invited talks on Workshop on Big Data, Streams and Heterogeneous Source Mining in conjunction with KDD'14 conference, New York City, NY, August 2014; NEC Labs America, Princeton, August 2014; Workshop on Heterogeneous Learning in conjunction with SDM'14 Conference, Philadelphia, PA, April 2014; Houghton College, Houghton, NY, March 2014; CSE Industry Advisory Board Meeting, Buffalo, NY, March 2014.

- An Overview of Data Mining and Potential Applications to Cancer Research
  Roswell Park Cancer Institute, Buffalo, NY, February 2013.

- Exploring the Power of Heterogeneous Information Sources
  Stony Brook University, 2012; EITA-YIC conference, 2013.

- Industrial lab presentations, Summer 2010, 2009, 2008, 2007:
  at various research groups in IBM T.J. Watson Research Center including stream analysis, security, network sciences, predictive modeling groups and NEC Labs America.

- Bipartite Graph-based Consensus Maximization among Supervised and Unsupervised Models
  Yahoo!-DAIS seminar, University of Illinois, Urbana-Champaign, November 2009.

- Knowledge Transfer via Multiple Model Local Structure Mapping
  Yahoo!-DAIS seminar, University of Illinois, Urbana-Champaign, September 2008.

- Mining Concept-Drifting Data Streams: Analysis and Practice
  Computer Science Department, Simon Fraser University, Vancouver, Canada, January 2008.

PROFESSIONAL SERVICE

- Program Committee Member
  - The 15th SIAM International Conference on Data Mining (**SDM 2015**), Vancouver, Canada, April 2015
  - The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML/PKDD 2014**), Nancy, France, September 2014
  - The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2014**), New York, NY, August 2014
  - The 31st International Conference on Machine Learning (**ICML 2014**), Beijing, China, June 2014
  - The 14th SIAM International Conference on Data Mining (**SDM 2014**), Philadelphia, PA, April 2014
  - Neural Information Processing Systems (**NIPS 2013**), Lake Tahoe, NV, December 2013
  - The 13th IEEE International Conference on Data Mining (**ICDM 2013**), Austin, TX, December 2013
  - The 22nd ACM International Conference on Information and Knowledge Management (**CIKM 2013**), San Francisco, CA, October 2013
  - The 1st IEEE International Conference on Big Data (**IEEE BigData 2013**), Santa Clara, CA, October 2013

- The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2013**), Chicago, IL, August 2013
- IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (**ASONAM 2013**), Niagara Falls, Canada, August 2013
- IEEE Intelligence and Security Informatics (**ISI 2013**), Seattle, WA, June 2013
- The 13th SIAM International Conference on Data Mining (**SDM 2013**), Austin, TX, May 2013
- The 12th IEEE International Conference on Data Mining (**ICDM 2012**), Brussels, Belgium, December 2012
- The 8th International Conference on Advanced Data Mining and Applications (**ADMA 2012**), Nanjing, China, December 2012
- The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2012**), Beijing, China, August 2012
- Workshop on Big Data Mining, in conjunction with KDD 2012, Beijing, China, August 2012
- IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (**ASONAM 2012**), Istanbul, Turkey, August 2012
- The 12th SIAM International Conference on Data Mining (**SDM 2012**), Anaheim, CA, May 2012
- The 1st International Conference on Health Information Science (**HIS 2012**), Beijing, China, April 2012
- The 7th International Conference on Advanced Data Mining and Applications (**ADMA 2011**), Beijing, China, December 2011
- Workshop on Handling Concept Drift in Adaptive Information Systems, in conjunction with ICDM 2011, Vancouver, Canada, December 2011
- The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML/PKDD 2011**), Athens, Greece, September 2011
- Workshop on Transfer Mining, in conjunction with ICDM 2009, Miami, FL, December 2009
- Workshop on Transfer Mining for Structured Data, in conjunction with NIPS 2009, Vancouver, Canada, December 2009

- Journal Reviewer

  - Transactions on Knowledge and Data Engineering (**TKDE**)
  - ACM Transactions on Knowledge Discovery from Data (**TKDD**)
  - ACM Transactions on Intelligent Systems and Technology (**TIST**)
  - Data Mining and Knowledge Discovery (**DAMI**)
  - Knowledge and Information Systems (**KAIS**)
  - Statistical Analysis & Data Mining (**SAM**)
  - IEEE Transactions on Neural Networks and Learning Systems (**TNNLS**)
  - IEEE Transactions on Cybernetics (**Cybernetics**)
  - Pattern Recognition Letters (**PRLetters**)
  - IEEE Signal Processing Letters (**SPL**)
  - Information Sciences (**INS**)
  - Information Processing & Management (**IPM**)
  - Journal of Computer Science and Technology (**JCST**)
  - Information Fusion (**IF**)

- External Reviewer

  - SIGKDD (2007, 2008), SIGMOD (2007), WWW (2007, 2009), WSDM (2009),
  - ICDM (2005, 2007, 2008), SDM (2006, 2007, 2008, 2009), WI (2005), CIDU (2010)

- Session Chair and Panelist

  - Panelist, National Science Foundation, Information & Intelligent Systems Division, Washington, DC, three times in 2013 and 2014
  - Session Chair, Session CP5: Multi-task/Transfer Learning, the 14th SIAM International Conference on Data Mining (**SDM 2014**), Philadelphia, PA, April 2014
  - Session Chair, Session 1B: Classification, the 13th IEEE Conference on Data Mining (**ICDM 2013**), Dallas, TX, December 2013
  - Panelist, Future Faculty Panel, School of Engineering and Applied Science, State University of New York at Buffalo, Buffalo, NY, May 2012 & May 2013
  - Session Chair, Session 19: Databases, Knowledgebases & Ontologies, the 3rd ACM Conference on Bioinformatics, Computational Biology and Biomedicine (**BCB 2012**), Orlando, FL, October 2012
  - Panelist, Forum: Women in Bioinformatics, the 3rd ACM Conference on Bioinformatics, Computational Biology and Biomedicine (**BCB 2012**), Orlando, FL, October 2012

- Conference and Workshop Organization

  - Proceedings Co-chair, the 5th ACM Conference on Bioinformatics, Computational Biology and Health Informatics (**BCB 2014**), Newport Beach, CA, September 2014
  - Publication Co-chair, the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2014**), New York, NY, August 2014
  - Publicity Co-chair, the 14th SIAM International Conference on Data Mining (**SDM 2014**), Philadelphia, PA, April 2014
  - Proceedings Chair, the 4th ACM Conference on Bioinformatics, Computational Biology and Biomedicine (**BCB 2013**), Washington, DC, September 2013
  - Proceedings Chair, the 3rd ACM Conference on Bioinformatics, Computational Biology and Biomedicine (**BCB 2012**), Orlando, FL, October 2012
  - Organizer, Workshop on Big Data Mining, in conjunction with KDD 2012, 2013, 2014