

Tao Li

OFFICE ADDRESS:

SCHOOL OF COMPUTING AND INFORMATION SCIENCES

FLORIDA INTERNATIONAL UNIVERSITY

MIAMI, FL 33199, U.S.A.

+1-305-348-6036

FAX: +1-305-348-3549

Email: taoli@cs.fiu.edu

Homepage: <http://www.cs.fiu.edu/~taoli>

RESEARCH AREA

Data Mining, Machine Learning, Computing System Management, Information Retrieval,
Bioinformatics,

EDUCATION

PhD 2004, Computer Science, University of Rochester, Rochester, NY

MS 2000, Mathematics, Oklahoma State University, Stillwater, OK

MS 1998, Computer Science, Chinese Academy of Science, P.R. China

BS 1995, Computer Science, Fuzhou University, P.R. China

CAREER

EMPLOYMENT

- 2009.8 – Associate Professor, School of Computing and Information Science,
Florida International University
- 2004.8 – 2009.7 Assistant Professor, School of Computing and Information Science,
Florida International University

VISITING POSITION

- 2001.5 – 2001.9 Summer Research Intern,
Xerox Research Center
- 2002.5 – 2002.8 Summer Research Intern,
Xerox Research Center
- 2003.5 – 2003.9 Summer Research Intern,
IBM T.J. Watson Research Center
- 2004.6 – 2004.8 Summer Research Intern,
IBM T.J. Watson Research Center

AWARDS

- IBM Scalable Data Analytics Innovation Award, 2010.
- Excellence in Student Mentoring Award, School of Computing and Information Sciences, Florida International University, 2009.
- Excellence in Research, Florida International University, 2009.
- National Science Foundation Faculty CAREER Award, 2006-2011.

- Tao Li, Qi Li, and Mitsunori Ogihara. “Music Feature Extraction Using Wavelet Coefficient Histograms”, US Patent Number 7,091,409. Awarded 8/15/06.
- IBM Faculty Research Award, 2008.
- Excellence in Faculty Scholarship, Florida International University, 2008
- Faculty Fellow for Mathematics of Knowledge and Search Engines at IPAM (Institute for Pure and Applied Mathematics, UCLA)
- IBM Faculty Research Award, 2007.
- Achievement Award at the 2006 International Conference on Machine Learning and Applications, Association for Machine Learning and Applications.
- IBM Faculty Research Award, 2005.
- IBM Shared University Research Award, 2005.
- Xerox University Affair Committee Award, 2005-2008.
- Excellence in Research Award, School of Computer Science at Florida International University, 2005.
- O.H.Hamilton Award, Department of Mathematics, Oklahoma State University, Spring 2000.

PUBLICATION LIST

Quick Summary: I edited 8 conference/workshop proceedings, and hold one patent. I have 170 refereed publications in total: 4 book chapters, 47 journal papers, 119 conference/workshop proceeding papers. Note that, in computer science, conferences are usually at least as important as journals. Among them, 88 papers appeared in top ranked CS conferences in their respective areas (Most conferences are very competitive with rigorous reviews (3 or more reviews) and very low acceptance rate (7-20%).), 27 papers appeared in top ranked journals. According to **Google Scholar**, my top 5 papers were cited over 740 times. I also gave three tutorials at leading data mining conferences.

VOLUMES EDITED

1. TAO LI, CHRIS DING and FEI WANG. **Proceedings of ICDM 2009 Workshop on Optimization Methods for Emerging Data Mining Problems (OEDM'09)**, 2009.
2. CHRIS DING and TAO LI. **Proceedings of the KDD 2009 Workshop on Data Mining using Matrices and Tensors(DMMT'09)**, 2009. ISBN 978-1-60558-673-1.
3. CHRIS DING, TAO LI, and SHENGHUO ZHU. **Proceedings of the KDD 2008 Workshop on Data Mining using Matrices and Tensors(DMMT'08)**, 2008. ISBN 978-1-60558-307-5.
4. M. ARIF WANI, MEHMED KANTARDZIC, TAO LI, and YING LIU. **Proceedings of the Sixth International Conference on Machine Learning and Applications**, ISBN 0-7695-3069-9, IEEE Computer Society, 2007.

5. M. ARIF WANI, TAO LI, LUKASZ KURGAN, JIEPING YE, and YING LIU. **Proceedings of the Fifth International Conference on Machine Learning and Applications**, ISBN 0-7695-2735-3, IEEE Computer Society, 2006.
6. TAO LI, Charles Perng, Haixun Wang, and Carlotta Domeniconi. **Proceedings of the KDD 2006 Workshop on Theory and Practice of Temporal Data Mining (TPTDM 2006)**, 2006.
7. Sheng Ma, TAO LI, and Charles Perng. **Proceedings of the TDM05 Workshop: Temporal Data Mining: Algorithms, Theory and Applications**, ISBN 0-9738918-3-1, IEEE Computer Society, 2005
8. Sheng Ma, TAO LI, and Charles Perng. **Proceedings of the TDM04 Workshop: Temporal Data Mining: Algorithms, Theory and Applications**, 2004.

PATENT

1. TAO LI, Qi Li, and Mitsunori Ogihara. **Music Feature Extraction Using Wavelet Coefficient Histograms**, US Patent Number 7,091,409. Awarded 8/15/06.

BOOK CHAPTERS

1. BRANDT ABSOLU, TAO LI, and MITSUNORI OGIHARA. **Analysis of Chord Progression Data. Advances in Music Information Retrieval**, In Z. W. Ras and A. Wierzchowska, eds., *Advances in Music Information Retrieval*, Springer.

2. TAO LI, MITSUNORI OGIHARA, BO SHAO, DINGDING WANG. **Machine Learning Approaches for Music Information Retrieval**. In M. J. Er and Y. Zhou, eds., *Machine Learning*, In-Tech Education and Publishing, 2009.
3. TAO LI, SHENG MA, and MITSUNORI OGIHARA. **Wavelet Methods in Data Mining**. In O. Maimon and L. Rokach, eds., *The Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*, pages 603–626, Springer-Verlag, 2005.
4. TAO LI and SHENG MA. **An Overview on Clustering Techniques**. In Encyclopedia of Data Warehousing and Mining , Idea Group Inc, 2006.

TUTORIALS

1. CHRIS DING, TAO LI. **Tutorial on Data Mining using Matrices and Tensors**. In 2009 IEEE International Conference on Data Mining (ICDM 2009).
2. FEI WANG, TAO LI, and CHRIS DING. **Tutorial on Data Mining with Graphs and Tensors**. In 2009 SIAM International Conference on Data Mining (SDM 2009).
3. FEI WANG, TAO LI, and CHRIS DING. **Tutorial on Information and Knowledge Management using Matrices and Graphs**. In ACM 17th Conference on Information and Knowledge Management (CIKM 2008).

PUBLICATIONS IN REFEREED JOURNALS

Review Procedure: Below I list articles that have appeared or accepted in various Computer Science journals. All publications below underwent peer-review by an average 3-4 expert reviewers in the area.

1. DINGDING WANG, SHENGHUO ZHU, TAO LI, YUN CHI, and YIHONG GONG. **Integrating Document Clustering and Multi-Document Summarization**, *ACM Transactions on Knowledge Discovery from Data (ACM TKDD)*, accepted, 2010.
2. SHUNZHI ZHU, DINGDING WANG, and TAO LI. **Data Clustering with Size Constraints**, *Knowledge-Based Systems*, **23**(8): 883-889, 2010.
3. FEI WANG, TAO LI, XIN WANG, SHENGHUO ZHU, and CHRIS DING. **Community Discovery Using Nonnegative Matrix Factorization**, *Data Mining and Knowledge Discovery*, in press, 2010. (Impact factor: 2.421)
4. VAGELIS HRISTIDIS, SHU-CHING CHEN, TAO LI, STEVEN LUIS, and YI DENG. **Survey of Data Management and Analysis in Disaster Situations**, *Journal of Systems and Software*, **83**(10): 1701-1714, 2010.
5. DINGDING WANG, TAO LI, SHENGHUO ZHU, and YIHONG GONG. **iHelp: An Intelligent Online Helpdesk System**, *IEEE Transactions on Systems, Man, and Cybernetics Part B*, accepted, 2010.
6. WEI PENG and TAO LI. **Tensor Clustering via Adaptive Subspace Iteration**, *Intelligent Data Analysis Journal*, accepted, 2010.
7. ERLIANG ZENG, CHENGYONG YANG, TAO LI, and GIRI NARASIMHAN. **Clustering Genes using Heterogeneous Data Sources**, *International Journal of Knowledge Discovery in Bioinformatics (IJKDB)*, accepted, 2010.

8. WEI PENG and TAO LI. **On the Equivalence Between Nonnegative Tensor Factorization and Tensorial Probabilistic Latent Semantic Analysis**, *Applied Intelligence*, to appear, 2010. (impact factor: 0.696)
9. YANFANG YE, TAO LI, QINGSHAN JIANG, and YOUYU WANG. **CIMDS: Adapting Postprocessing Techniques of Associative Classification for Malware Detection**, *IEEE Transactions on Systems, Man, and Cybernetics Part C*, to appear, 2010.
10. WEI PENG and TAO LI. **Temporal Relation Co-clustering on Directional Social Network and Author-Topic Evolution**, *Knowledge and Information Systems Journal*, to appear, 2010. (impact factor: 1.733).
11. ZHIYUAN CHEN and TAO LI. **Addressing Diverse User Preferences: A Framework for Query Results Navigation**. *IEEE Data Eng. Bull.*, **32**(4): 41-48, 2009.
12. ZHONG-YUAN ZHANG, TAO LI, CHRIS DING, XIAN-WEN REN, and XIANG-SUN ZHANG. **Binary Matrix Factorization for Analyzing Gene Expression Data**. *Data Mining and Knowledge Discovery*, **20**(1): 28-52, 2010. (Impact factor: 2.421)
13. FEI WANG, XIN WANG, DAOQIANG ZHANG, CHANGSHUI ZHANG, and TAO LI. **marginFace: A Novel Face Recognition Method by Average Neighborhood Margin Maximization**. *Pattern Recognition*, **42**(11): 2863-2875, 2009. (Impact factor: 3.279)
14. YANFANG YE, TAO LI, KAI HUANG, QINGSHAN JIANG, YONG CHEN. **Hierarchical Associative Classifier (HAC) for Malware Detection from the Large and Imbalanced Gray List**. *Journal of Intelligent Information Systems*, to appear, 2010.
15. TAO LI, MITSUNORI OGIHARA and SHENG MA. **A Unified View on Combining Multiple Clusterings**. *Applied Intelligence*, to appear, 2010. (impact factor: 0.696)

16. BO SHAO, DINGDING WANG, TAO LI, and MITSUNORI OGIHARA. **Music Recommendation Based on Acoustic Features and User Access Patterns.** *IEEE Transactions on Audio, Speech and Language Processing*, **17**(8): 1602-1611, 2009. (impact factor: 0.768)
17. FEI WANG, CHANGSHUI ZHANG, and TAO LI. **Clustering with Local and Global Regularization.** *IEEE Transactions on Knowledge and Data Engineering*, **21**(12):1665-1678, 2009. (impact factor: 1.896)
18. CHRIS DING, TAO LI, and MICHAEL I. JORDAN. **Convex and Semi-Nonnegative Matrix Factorizations.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32**(1):45-55, 2010. (impact factor: 3.8)
19. TAO LI, MITSUNORI OGIHARA, WEI PENG, BO SHAO, and SHENGHUO ZHU. **Music Clustering with Features from Different Information Sources.** *IEEE Transactions on Multimedia*, **11**(3): 477-485, 2009. (impact factor: 1.518)
20. TAO LI, WEI PENG, CHARLES PERNG, SHENG MA, and HAIXUN WANG. **An Integrated Data-Driven Framework for Computing System Management.** *IEEE Transactions on Systems, Man, and Cybernetics Part A*, **40**(1): 90-99, 2010. (impact factor: 1.353)
21. YANFANG YE, LIFEI CHEN, DINGDING WANG, TAO LI, and MIN ZHAO. **SBMDS: An Interpretable String Based Malware Detection System using SVM Ensemble with Bagging.** *Journal in Computer Virology*, **5**(4): 283-293, 2009.
22. WEI PENG, TONG SUN, PHILIP ROSE, and TAO LI. **Computation and Applications of Industrial Leading Indicators to Business Process Improvement.** *International Journal of Intelligent Control and Systems (IJICS)*, **13**(3):196-207, 2008.

23. CHRIS DING, TAO LI, and SHENGHUO ZHU. **KDD2008 Workshop Report DMMT.08: Data Mining Using Matrices and Tensors.** *SIGKDD Explorations*, **10**(2): 54-56, 2008.
24. YI ZHANG, DINGDING WANG, and TAO LI. **LIBGS: A MATLAB Software Package For Gene Selection.** *International Journal of Data Mining and Bioinformatics*, to appear, 2010.
25. YI ZHANG, CHRIS DING, and TAO LI. **Gene Selection Algorithm by Combining ReliefF and mRMR.** *BMC Genomics*, **9**: S27, 2008. (impact factor: 4.18)
26. SHENGHUO ZHU, DINGDING WANG, KAI YU, TAO LI and YIHONG GONG. **Feature Selection for Gene Expression using Model-based Entropy.** *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **7**(1): 25-36, 2010. (impact factor: 1.803)
27. TAO LI, SHENGHUO ZHU, and MITSUNORI OGIHARA. **Text Categorization via Generalized Discriminant Analysis.** *Information Processing & Management*, Vol 44: 1684-1697, 2008. (impact factor: 1.500)
28. TAO LI. **Clustering Based on Matrix Approximation.** *Knowledge and Information Systems Journal*, **17**(1): 1-15, 2008. (impact factor: 1.733)
29. RAMAKRISHNA VARADARAJAN, VAGELIS HRISTIDIS, and TAO LI. **Beyond Single-Page Web Search Results.** *IEEE Transactions on Knowledge and Data Engineering*, **20**(3): 411-424, 2008. (impact factor: 1.896)
30. CHRIS DING, TAO LI and WEI PENG. **On the Equivalence Between Non-negative Matrix Factorization and Probabilistic Latent Semantic Indexing.** *Computational Statistics and Data Analysis*, **52**(8): 3913-3927, 2008. (impact factor: 1.029)

31. YANFANG YE, DINGDING WANG, TAO LI and DONGYI YE. **An Intelligent PE-Malware Detection System Based on Association Mining.** *Journal in Computer Virology*, 4(4): 323-334, 2008.
32. TAO LI, CHANG-SHING PERNG, and SHENG MA. **Guest Editorial: Special Issue on Temporal Data Mining: Theory, Algorithms and Applications.** *Data Mining and Knowledge Discovery*, 16(1): 1-3, 2008. (impact factor: 2.42)
33. LI ZHANG, TAO LI, SHIXIA LIU, and YUE PAN. **An Integrated System for Building Enterprise Taxonomies.** *Information Retrieval*, 10(4/5): 365-395, 2007. (impact factor: 0.770)
34. TAO LI, SHENGHUO ZHU, and MITSUNORI OGIHARA. **Hierarchical Document Classification Using Automatically Generated Hierarchy.** *Journal of Intelligent Information Systems*, 29(2): 211-230, 2007. (impact factor: 0.796)
35. WEIXIANG SUN, TAO LI, WEI PENG, and TONG SUN. **Incremental Workflow Mining with Optional Patterns and Its Application to Production Printing Process.** *The International Journal of Intelligent Control and Systems*, 12(1): 45-55, 2007.
36. TAO LI. **A Unified View on Clustering Binary Data.** *Machine Learning*, 62(3): 199-215, 2006. (impact factor: 1.742)
37. TAO LI, SHENGHUO ZHU, and MITSUNORI OGIHARA. **Using Discriminant Analysis for Multi-class Classification: An Experimental Investigation.** *Knowledge and Information Systems*, 10(4): 453-472, 2006. (impact factor: 1.733)
38. TAO LI and Chang-Shing Perng. **KDD-2006 Workshop Report: Theory and Practice of Temporal Data Mining.** *SIGKDD Explorations*, 8(2): 96-97, 2006.

39. TAO LI and MITSUNORI OGIHARA. **Toward Intelligent Music Information Retrieval.** *IEEE Transactions on Multimedia*, **8**(3):564–574, 2006. (impact factor: 1.518)
40. JIEPING YE, TAO LI, TAO XIONG, and RAVI JANARDAN. **Using Uncorrelated Discriminant Analysis for Tissue Classification with Gene Expression Data.** *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **1**(4): 181–190, 2004. (impact factor: 1.803)
41. TAO LI and MITSUNORI OGIHARA. **Semisupervised Learning from Different Information Sources.** *Knowledge and Information Systems*, **7**(3): 289–309 (2005). (impact factor: 1.733)
42. WEI PENG, TAO LI, and SHENG MA. **Mining logs files for Data-Driven System Management.** *SIGKDD Explorations*, **7**(1): 44–51, 2005.
43. TAO LI, CHENGLIANG ZHANG, and MITSUNORI OGIHARA. **A Comparative Study of Feature Selection and Multiclass Classification Methods for Tissue Classification Based on Gene Expression.** *Bioinformatics*, **20**:2421–2428, 2004. (impact factor: 5.039)
44. TAO LI, QI LI, SHENGHUO ZHU, and MITSUNORI OGIHARA. **A Survey on Wavelet Applications in Data Mining.** *SIGKDD Explorations*, **4**(2):49–68, 2003.
45. TAO LI, SHENGHUO ZHU, and MITSUNORI OGIHARA. **Algorithms for Clustering High Dimensional and Distributed Data.** *Intelligent Data Analysis*, **7**(4):305–326, 2003. (impact factor: 0.446)
46. TAO LI, MITSUNORI OGIHARA, and SHENGHUO ZHU. **Association-based Similarity Testing and Its Applications.** *Intelligent Data Analysis*, **7**(3):209–232, 2003. (impact factor: 0.446)

47. XIANGHUI LIU, NOPHILL PARK, TAO LI, and Roger Y. Lee. **Exploiting Both Spatial And Temporal Locality In Page Replacement Algorithm.** *International Journal of Computer & Information Science (IJCIS)*, **3**(1), 2002.

PUBLICATIONS IN PEER-REVIEWED CONFERENCE PROCEEDINGS

Review Procedure: Below I list articles that have appeared or accepted in the proceedings of various Computer Science conferences. All publications underwent peer-review by an average 3-4 expert reviewers in the area. Note that, in several areas of Computer Science, including my own (Data Mining), conferences are usually at least as important as journals. Conference publications are considered to have greater impact and are often more timely and prestigious than journal publications, due to its rigorous review process and low acceptance rates.

1. ZENGGUO SUN, TAO LI, and NAPHTALI RISHE. **Large-Scale Matrix Factorization using MapReduce**, In *Proceedings of the Optimization-based Methods for Emerging Data Mining Workshop associated with ICDM 2010 (OEDM'10)*, 2010.
2. LI ZHENG, TAO LI, and CHRIS DING. **Hierarchical Ensemble Clustering**, In *Proceedings of 2010 IEEE International Conference on Data Mining (ICDM 2010)*, 2010.
3. LIANG TANG and TAO LI. **LogTree: A Framework for Generating System Events from Raw Textual Logs**, In *Proceedings of 2010 IEEE International Conference on Data Mining (ICDM 2010)*, 2010.

4. DINGDING WANG, CHRIS DING, and TAO LI. **Weighted Feature Subset Non-Negative Matrix Factorization and its Applications to Document Understanding**, In *Proceedings of 2010 IEEE International Conference on Data Mining (ICDM 2010)*, 2010.
5. CHAO SHEN, DINGDING WANG, and TAO LI. **Topic Aspect Analysis for Multi-Document Summarization**, In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM 2010)*, 2010.
6. DINGDING WANG and TAO LI. **Document Update Summarization Using Incremental Hierarchical Clustering**, In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM 2010)*, 2010.
7. CHAO SHEN and TAO LI. **Multi-Document Summarization via the Minimum Dominating Set**, In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, 2010.
8. YANFANG YE, TAO LI, and YONG CHEN. **Automatic Malware Categorization Using Cluster Ensemble**, In *Proceedings of the 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD 2010)*, pages 95-104, 2010.
9. LI ZHENG, CHAO SHEN, LIANG TANG, TAO LI, STEVE LUIS, SHU-CHING CHEN, and VAGELIS HRISTIDIS. **Using Data Mining Techniques to Address Critical Information Exchange Needs in Disaster Affected Public-Private Networks**, In *Proceedings of the 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD 2010)*, pages 125-134, 2010.
10. JINGXUAN LI, TAO LI, and MITSUNORI OGIHARA. **Hierarchical Co-Clustering of Artists and Tags**, In *Proceedings of the Eleventh International Society for Music Information Retrieval Conference (ISMIR 2010)*, 2010.

11. DINGDING WANG, TAO LI, and MITSUNORI OGIHARA. **Are Tags Better Than Audio? The Effect of Joint Use of Tags and Audio Content Features for Artistic Style Clustering**, In *Proceedings of the Eleventh International Society for Music Information Retrieval Conference (ISMIR 2010)*, 2010.
12. JINGXUAN LI and TAO LI. **HCC: A Hierarchical Co-clustering Algorithm**, In *Proceedings of the 33rd International ACM SIGIR conference on Research and Development in Information Retrieval (SIGIR 2010)*, Pages 861-862, 2010.
13. DINGDING WANG and TAO LI. **Many Are Better than One: Improving Multi-document Summarization via Weighted Consensus**, In *Proceedings of the 33rd International ACM SIGIR conference on Research and Development in Information Retrieval (SIGIR 2010)*, Pages 809-810, 2010.
14. YI ZHANG, DINGDING WANG, and TAO LI. **VisualSum: An Interactive Multi-document Summarization System using Visualization**, In *Proceedings of the 33rd International ACM SIGIR conference on Research and Development in Information Retrieval (SIGIR 2010)*, Pages 857-858, 2010.
15. LEI LI, DINGDING WANG, CHAO SHEN, and TAO LI. **Ontology-Enriched Multi-document Summarization in Disaster Management**. In *Proceedings of the 33rd International ACM SIGIR conference on Research and Development in Information Retrieval (SIGIR 2010)*, Pages 819-820, 2010.
16. DINGDING WANG, CHRIS DING, and TAO LI. **Feature subset non-negative matrix factorization and its applications to document understanding**, In *Proceedings of the 33rd International ACM SIGIR conference on Research and Development in Information Retrieval (SIGIR 2010)*, Pages 805-806, 2010.

17. TAO LI, VIKAS SINDHWANI, CHRIS DING, and YI ZHANG. **Bridging Domains with Words: Opinion Analysis with Matrix Tri-factorizations**, In *Proceedings of the Tenth SIAM Conference on Data Mining (SDM 2010)*, 2010.
18. YI ZHANG, ERLIANG ZENG, TAO LI, and GIRI NARASIMHAN. **Weighted Consensus Clustering for Identifying Functional Modules in Protein-Protein Interaction Networks**. In *Proceedings of the Eighth International Conference on Machine Learning and Applications (ICMLA 2009)*, 2009.
19. CHRIS DING, TAO LI, and DINGDING WANG. **Label Propagation on K-partite Graphs**. In *Proceedings of the Eighth International Conference on Machine Learning and Applications (ICMLA 2009)*, 2009.
20. FEI WANG, JIMENG SUN, TAO LI, and NIKOS ANEROUSIS. **Two Heads Better Than One: Metric+Active Learning and its Applications for IT Service Classification**. In *Proceedings of 2009 IEEE International Conference on Data Mining (ICDM 2009)*, pages 1022-1027, 2009.
21. FEI WANG, XIN WANG, and TAO LI. **Maximum Margin Clustering on Data Manifolds**. In *Proceedings of 2009 IEEE International Conference on Data Mining (ICDM 2009)*, pages 1028-1033, 2009.
22. DIJUN LUO, CHRIS H. Q. DING, HENG HUANG, and TAO LI. **Non-negative Laplacian Embedding**. In *Proceedings of 2009 IEEE International Conference on Data Mining (ICDM 2009)*, pages 337-346, 2009.
23. FEI WANG, XIN WANG, and TAO LI. **Semi-supervised Multi-task Learning with Task Regularizations**. In *Proceedings of 2009 IEEE International Conference on Data Mining (ICDM 2009)*, pages 562-568, 2009.

24. FEI WANG and TAO LI. **Knowledge Transformation by Cross-Domain Belief Propagation**. In *Proceedings of 2009 IEEE International Conference on Data Mining (ICDM 2009) Workshops*, pages 441-446, 2009.
25. YANFANG YE, TAO LI, QINGSHAN JIANG, ZHIXUE HAN, and LI WAN. **Intelligent File Scoring System for Malware Detection from the Gray List**. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD 2009)*, pages 1385-1394, 2009.
26. FEI WANG, XIN WANG, BO SHAO, TAO LI, and MITSUNORI OGIHARA. **Tag Integrated Multi-Label Music Style Classification with Hypergraph**. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, 2009.
27. DINGDING WANG, SHENGHUO ZHU, TAO LI, and YIHONG GONG: **Comparative Document Summarization via Discriminative Sentence Selection**. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management (CIKM 2009)*: 1963-1966, 2009.
28. DINGDING WANG, CHRIS DING, and TAO LI. **K-Subspace Clustering**. In *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2009)*, pages 506-521, 2009.
29. HELI SUN, JIANBIN HUANG, BOQIN FENG, TAO LI, YINGLIANG ZHAO, and JUN LIU. **OrdRank: Learning to Rank with Ordered Multiple Hyperplanes**. In *Proceedings of the 2009 IEEE/WIC/ACM International Conference on Web Intelligence (WI'2009)*, to appear, 2009.
30. TAO LI, YI ZHANG, and VIKAS SINDHWANI. **A Non-negative Matrix Tri-factorization Approach to Sentiment Classification with Lexical Prior**

- Knowledge.** In *Proceedings of the 47th Annual Meeting of the Association for Computational Linguistics (ACL 2009)*, 2009.
31. DINGDING WANG, SHENGHUO ZHU, TAO LI, and YIHONG GONG. **Multi-document Summarization using Sentence-based Topic Models.** In *Proceedings of the 47th Annual Meeting of the Association for Computational Linguistics (ACL 2009)*, 2009.
32. TAO LI, VIKAS SINDHWANI, CHRIS DING, and YI ZHANG. **Knowledge Transformation for Cross-domain Sentiment Classification.** In *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2009)*, pages 716-717, 2009.
33. DINGDING WANG, LI ZHENG, TAO LI, and YI DENG. **Evolutionary Document Summarization for Disaster Management.** In *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2009)*, pages 680-681, 2009.
34. QI ZHANG, YUANBIN WU, TAO LI, MITSUNORI OGIHARA, JOSEPH JOHNSON, and XUANJING HUANG. **Mining Product Reviews based on Shallow Dependency Parsing.** In *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2009)*, pages 726-727, 2009.
35. FEI WANG, XIN WANG, and TAO LI. **Generalized Cluster Aggregation.** In *The 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1279-1284, 2009.
36. FEI WANG, BIN ZHANG, TA-HSIN LI, WEN JUN YIN, JIN DONG, and TAO LI. **Preference Learning with Extreme Examples.** In *The 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1285-1290, 2009.

37. DAN ZHANG, FEI WANG, LUO SI, and TAO LI. **Maximum Margin Multiple Instance Clustering**. In *The 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1339-1344, 2009.
38. FEI WANG, XIN WANG, and TAO LI. **Beyond the Graphs: Semi-Parametric Semi-Supervised Discriminant Analysis**. In *IEEE Comp. Society Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 2113-2120, 2009.
39. FEI WANG, CHRIS DING, and TAO LI. **Integrated KL (K-means - Laplacian) Clustering: A New Clustering Approach by Combining Attribute Data and Pairwise Relations**. In *the Ninth SIAM Conference on Data Mining (SDM 2009)*, pages 38-48, 2009.
40. CHRIS DING, TAO LI, and MICHAEL I. JORDAN. **Nonnegative Matrix Factorization for Combinatorial Optimization: Spectral Clustering, Graph Matching, and Clique Finding**. In *Proceedings of 2008 IEEE International Conference on Data Mining (ICDM 2008)*, pages 183-192, 2008.
41. DINGDING WANG, SHENGHUO ZHU, TAO LI, YUN CHI, and YIHONG GONG. **Integrating Clustering and Multi-document Summarization to Improve Document Understanding**. In *Proceedings of the Conference on Information and Knowledge Management (CIKM 2008)*, pages 1435-1436, 2008.
42. WEI PENG, TAO LI, and BO SHAO. **Clustering Multiway Data via Adaptive Subspace Iteration**. In *Proceedings of the Conference on Information and Knowledge Management (CIKM 2008)*, pages 1519-1520, 2008.
43. FEI WANG, SHOUCHUN CHEN, CHANGSHUI ZHANG, and TAO LI. **Semi-Supervised Metric Learning by Maximizing Constraint Margin**. In *Proceedings of the*

- Conference on Information and Knowledge Management (CIKM 2008)*, pages 1457-1458, 2008.
44. BO SHAO, TAO LI, and MITSUNORI OGIHARA. **Quantify Music Artist Similarity Based on Style and Mood**. In *Proceeding of the 10th ACM workshop on Web Information and Data Management (WIDM 2008)*, pages 119-124, 2008.
 45. MITSUNORI OGIHARA and TAO LI. **N-Gram Profiles for Composer Style Identification**. In *Proceedings of the ninth International Conference on Music Information Retrieval (ISMIR 2008)*, pages 671-676, 2008.
 46. HENG HUANG, CHRIS DING, DIJUN LUO, and TAO LI. **Simultaneous Tensor Subspace Selection and Clustering: The Equivalence of High Order SVD and K-Means Clustering**. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD 2008)*, pages 327-355, 2008.
 47. SHENGHUO ZHU, TAO LI, ZHIYUAN CHEN, DINGDING WANG, and YIHONG GONG. **Dynamic Active Probing of Helpdesk Databases**. In *Proceedings of the 34th International Conference on Very Large Data Bases (VLDB 2008)*, to appear.
 48. TAO LI and CHRIS DING. **Weighted Consensus Clustering**. In *Proceedings of 2008 SIAM International Conference on Data Mining (SDM 2008)*, pages 798-809, 2008.
 49. FEI WANG, TAO LI, and CHANGSHUI ZHANG. **Semi-Supervised Clustering via Matrix Factorization**. In *Proceedings of 2008 SIAM International Conference on Data Mining (SDM 2008)*, pages 1-12, 2008.

50. TAO LI, CHRIS DING, YI ZHANG and BO SHAO. **Knowledge Transformation from Word Space to Document Space**. In *Proceedings of The 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2008)*, pages 187-194, 2008.
51. DINGDING WANG and TAO LI. **Multi-Document Summarization via Sentence-Level Semantic Analysis and Symmetric Matrix Factorization**. In *Proceedings of The 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2008)*, pages 307-314, 2008.
52. WEI PENG and TAO LI. **Author-Topic Evolution Analysis using Three-way Non-negative Paratucker**. In *Proceedings of The 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2008)*, pages 819-820, 2008.
53. DIJUN LUO, TAO LI, CHRIS DING, and WEI PENG. **Posterior Probabilistic Clustering using NMF**. In *Proceedings of The 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2008)*, pages 831-832, 2008.
54. FEI WANG, TAO LI, GANG WANG, and CHANGSHUI ZHANG. **Semi-supervised Classification Using Local and Global Regularization**. In *Proceedings of The 23rd AAAI Conference on Artificial Intelligence (AAAI 2008)*, pages 726-731, 2008.
55. DAN ZHANG, FEI WANG, CHANGSHUI ZHANG, and TAO LI. **Multi-view Local Learning**. In *Proceedings of The 23rd AAAI Conference on Artificial Intelligence (AAAI 2008)*, pages 752-757, 2008.
56. RICARDO KOLLER, RAJU RANGASWAMI, JOSEPH MARRERO, IGOR HERNANDEZ, GEOFFREY SMITH, MANDY BARSILAI, SILVIU NECULA, S. MASOUD SADJADI,

- TAO LI, and KRISTA MERRILL. **Anatomy of a Real-time Intrusion Prevention System**. In *Proceedings of the IEEE International Conference on Autonomic Computing (ICAC 2008)*.
57. KHALID SALEEM, STEVE LUIS, YI DENG, SHU-CHING CHEN, VAGELIS HRISTIDIS, and TAO LI. **Towards a Business Continuity Information Network for Rapid Disaster Recovery**. In *Proceedings of the 9th Annual International Conference on Digital Government Research*, pages 107-116, 2008.
58. TAO LI, CHRIS DING, and MICHAEL I. JORDAN. **Solving Consensus and Semi-supervised Clustering Problems Using Nonnegative Matrix Factorization**. In *Proceedings of 2007 IEEE International Conference on Data Mining (ICDM 2007)*, pages 577-582, 2007.
59. ZHONGYUAN ZHANG, TAO LI, CHRIS DING, and XIANG-SUN ZHANG. **Binary Matrix Factorization with Applications**. In *Proceedings of 2007 IEEE International Conference on Data Mining (ICDM 2007)*, pages 391-400, 2007.
60. WEI PENG, CHRIS DING, TAO LI, and TONG SUN. **Finding Hotspots in Document Collection**. In *Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*, pages 313-320, 2007.
61. FEI WANG and TAO LI. **Gene Selection via Matrix Factorization**. In *Proceedings of 7th IEEE Conference on Bioinformatics and Bioengineering (BIBE2007)*, pages 1046-1050, 2007.
62. FEI WANG, XIN WANG, and TAO LI. **Efficient Label Propagation for Interactive Image Segmentation**. In *Proceedings of The Sixth International Conference on Machine Learning and Applications (ICMLA 2007)*.

63. ERLIANG ZENG, CHENGYONG YANG, TAO LI, and GIRI NARASIMHAN. **On the Effectiveness of Constraints Sets in Clustering Genes.** In *Proceedings of 7th IEEE Conference on Bioinformatics and Bioengineering (BIBE2007)*, pages 79-86, 2007.
64. YI ZHANG, CHRIS DING, and TAO LI. **A Two-Stage Gene Selection Algorithm by Combining ReliefF and mRMR.** In *Proceedings of 7th IEEE Conference on Bioinformatics and Bioengineering (BIBE2007)*, pages 164-171, 2007.
65. WEI PENG, TONG SUN, PHILIP ROSE, and TAO LI. **A Semi-automatic System with an Iterative Learning Method for Discovering the Leading Indicators in Business Processes,** In *Proceedings of SIGKDD'07 workshop on Domain Driven Data Mining (DDDM'07)*.
66. WEI PENG, TAO LI, and MITSUNORI OGIHARA. **Music Clustering With Constraints.** In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*.
67. YANFANG YE, DINGDING WANG, TAO LI, and DONGYI YE. **IMDS: Intelligent Malware Detection System.** In *Proceedings of ACM Int'l Conf. on Knowledge Discovery and Data Mining (SIGKDD 2007)*, pages 1043-1047, 2007.
68. CHRIS DING, RONG JIN, TAO LI, and HORST D. SIMON. **A Learning Framework using Green's Function and Kernel Regularization with Application for Recommender System.** In *Proceedings of ACM Int'l Conf. on Knowledge Discovery and Data Mining (SIGKDD 2007)*, pages 260-269, 2007.
69. WEI PENG, CHARLES PERNG, TAO LI, and HAIXUN WANG. **Event Summarization for System Management.** In *Proceedings of ACM Int'l Conf. on Knowledge Discovery and Data Mining (SIGKDD 2007)*, pages 1028-1032, 2007.

70. CHRIS DING and TAO LI. **Adaptive Dimension Reduction Using Discriminant Analysis and K-means Clustering**. In *Proceedings of International Conference on Machine Learning (ICML 2007)*, pages 521-528, 2007.
71. FEI WANG, CHANGSHUI ZHANG, and TAO LI. **Clustering with Local and Global Regularization**. In *Proceedings of The 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2007)*, pages 95-102, 2007.
72. FEI WANG, CHANGSHUI ZHANG, and TAO LI. **Regularized Clustering for Documents**. In *Proceedings of the 22nd National Conference on Artificial Intelligence (AAAI 2007)*, pages 657-662, 2007.
73. ZHIYUAN CHEN and TAO LI. **Addressing Diverse User Preferences in SQL-Query-Result Navigation**. In *Proceeding of the ACM SIGMOD 2007 International Conference*, pages 641-652, 2007.
74. TAO LI, MITSUNORI OGIHARA, and SHENGHUO ZHU. **Integrating Features from Different Sources for Music Information Retrieval**. In *Proceedings of the 6th IEEE International Conference on Data Mining (ICDM'06)*, pages 372-381, 2006.
75. TAO LI and CHRIS DING. **The Relationships Among Various Nonnegative Matrix Factorization Methods for Clustering**. In *Proceedings of the 6th IEEE International Conference on Data Mining (ICDM'06)*, pages 362-371, 2006.
76. CHRIS DING, YA ZHANG, TAO LI, and STEVE HOLBROOK. **Biclustering Protein Complex Interactions with a Biclique Finding Algorithm**. In *Proceedings of the 6th IEEE International Conference on Data Mining (ICDM'06)*, pages 178-187, 2006.

77. FEI WANG, SHENG MA, LIUZHONG YANG, and TAO LI. **Item Recommendations on Graph**. In *Proceedings of the 6th IEEE International Conference on Data Mining (ICDM'06)*, pages 1119-1123, 2006.
78. CHRIS DING, TAO LI, WEI PENG, and HAESUN PARK. **Orthogonal Nonnegative Matrix Tri-factorizations for Clustering**. In *Proceedings of The Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD 2006)*, pages 126–135, 2006.
79. WEI PENG and TAO LI. **Interval Data Clustering with Applications**. In *Proceedings of the 18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2006)*, pages 355–362, 2006.
80. TAO LI, CHENGLIANG ZHANG, and SHENGHUO ZHU. **Empirical Studies on Multilabel Classification**. In *Proceedings of the 18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2006)*, pages 86–92, 2006.
81. CHRIS DING, TAO LI, and WEI PENG. **Nonnegative Matrix Factorization and Probabilistic Latent Semantic Indexing: Equivalence, Chi-square Statistic, and a Hybrid Method**. In *Proceeding of National Conference on Artificial Intelligence (AAAI-06)*, 2006.
82. WEI PENG and TAO LI. **IntClust: A Software Package for Clustering Replicated Microarray Data**. In *Proceedings of the IEEE 6th Symposium on Bioinformatics & Bioengineering (BIBE 2006)*, pages 103–109, 2006.
83. WEI PENG, TAO LI, and GIRI NARASIMHAN. **Mining the Database of Transcription Binding Sites**. In *Proceedings of the IEEE 6th Symposium on Bioinformatics & Bioengineering (BIBE 2006)*, pages 61–64, 2006.

84. RAMAKRISHNA VARADARAJAN, VAGELIS HRISTIDIS, and TAO LI. **Searching the Web using Composed Pages**. In *Proceedings of The Twenty-Ninth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2006)*, pages 713–714, 2006.
85. CHRIS DING, TAO LI, and WEI PENG. **NMF and PLSI: Equivalence and a Hybrid Algorithm**. In *Proceedings of The Twenty-Ninth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2006)*, pages 641–642, 2006.
86. WEIXIANG SUN, TAO LI, WEI PENG, and TONG SUN. **Incremental Workflow Mining with Optional Patterns**. In *Proceedings of 2006 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2006)*, 2006.
87. TAO LI and WEI PENG. **A Clustering Model Based on Matrix Approximation with Applications to Cluster System Log Files**. In *Proceedings of the 16th European Conference on Machine Learning (ECML 2005)*, pages 625–632, 2005.
88. TAO LI. **A General Model for Clustering Binary Data**. In *Proceedings of The Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD 2005)*, pages 188–197, 2005.
89. TAO LI, FENG LIANG, SHENG MA, and WEI PENG. **An Integrated Framework on Mining Logs Files for Computing System Management**. In *Proceedings of The Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD 2005)*, pages 776–781, 2005.
90. CHENGYONG YANG, ERLIANG ZENG, TAO LI, and GIRI NARASIMHAN. **Clustering Genes Using Gene Expression and Text Literature Data**. In *Proceedings of*

the 2005 IEEE Computational Systems Bioinformatics Conference (CSB 2005), pages 329–340, 2005.

91. CHENGYONG YANG, ERLIANG ZENG, TAO LI, and GIRI NARASIMHAN. **A Knowledge-Driven Method to Evaluate Multi-source Clustering**. In *Proceedings of the ISPA Workshops 2005*, pages 196–202, 2005.
92. WEI PENG, TAO LI, and SHENG MA. **Mining Logs Files for Computing System Management**. In *Proceedings of The 2nd IEEE International Conference on Autonomic Computing (ICAC 2005)*, pages 309–310, 2005.
93. TAO LI and SHENGHUO ZHU. **On Clustering Binary Data**. In *Proceedings of the 2005 SIAM International Conference on Data Mining (SDM 2005)*, 2005.
94. TAO LI and SHENGHUO ZHU. **Hierarchical Document Classification Using Automatically Generated Hierarchy**. In *Proceedings of the 2005 SIAM International Conference on Data Mining (SDM 2005)*, 2005.
95. TAO LI and MITSUNORI OGIHARA. **Music Genre Classification with Taxonomy**. In *Proceedings of the 2005 International Conference on Acoustic, Speech, and Signal Processing (ICASSP'05)*, pages V198-201. IEEE Computer Society Press, 2005.
96. TAO LI, SHENG MA, and MITSUNORI OGIHARA. **Document Clustering via Adaptive Subspace Iteration**. In *Proceedings of the 27th Annual International ACM Conference on Research and Development in Information Retrieval (SIGIR'04)*, pages 218–225.
97. TAO LI and SHENG MA. **Mining Temporal Patterns Without Pre-defined Time Windows**. In *Proceedings of The Fourth International Conference on Data Mining (ICDM 2004)*, pages 451–454, 2004.

98. TAO LI and MITSUNORI OGIHARA. **Semi-supervised Learning for Music Artist Classification.** In *Proceedings of the IEEE Conference on Information Knowledge Management (CIKM'04)*, pages 152–153, 2004.
99. TAO LI, SHENG MA, and MITSUNORI OGIHARA. **On Combining Multiple Clusterings.** In *Proceedings of the IEEE Conference on Information Knowledge Management (CIKM'04)*, pages 294–303, 2004.
100. TAO LI and SHENG MA. **IFD: Iterative Feature and Data Clustering.** In *Proceedings of The 2004 SIAM International Conference on Data Mining (SDM 2004)*, 2004.
101. TAO LI and MITSUNORI OGIHARA. **Music Artist Style Identification by Semisupervised Learning from Both Lyrics and Content.** In *Proceedings of the 12th ACM International Conference on Multimedia (MM'04)*, pages 364–367, 2004.
102. TAO LI, SHENG MA, and MITSUNORI OGIHARA. **Entropy-based Criterion for Categorical Data Clustering.** In *Proceedings of the 21st International Conference on Machine Learning (ICML'04)*, pages 536–543.
103. TAO LI and MITSUNORI OGIHARA. **Content-based Music Similarity Search and Emotion Detection.** In *Proceedings of the 2004 International Conference on Acoustic, Speech, and Signal Processing (ICASSP'04)*, IEEE Computer Society Press, pages V705–V708, 2004.
104. TAO LI, SHENGHUO ZHU, and MITSUNORI OGIHARA. **Using Discriminant Analysis for Multi-class Classification.** In *Proceedings of the 2003 International Conference on Data Mining (ICDM'03)*, pages 589–592, IEEE Computer Society Press, 2003.

105. TAO LI and MITSUNORI OGIHARA. **Detecting Emotion in Music**. In *Proceedings of the Fifth International Symposium on Music Information Retrieval (ISMIR'03)*, pages 239–240, 2003.
106. TAO LI, SHENGHUO ZHU, and MITSUNORI OGIHARA. **Efficient Multi-way Text Categorization via Generalized Discriminant Analysis**. In *Proceedings of the Twelfth International Conference on Information and Knowledge Management (CIKM'03)*, pages 317–324. ACM Press, 2003.
107. TAO LI, MITSUNORI OGIHARA, and QI LI. **A Comparative Study on Content-based Music Genre Classification**. In *Proceedings of the 26th Annual International ACM Conference on Research and Development in Information Retrieval (SIGIR'03)*, pages 282–289. ACM Press, 2003.
108. TAO LI, SHENGHUO ZHU, and MITSUNORI OGIHARA. **Topic Hierarchy Generation via Linear Discriminant Projection**. In *Proceedings of the 26th Annual International ACM Conference on Research and Development in Information Retrieval (SIGIR'03)*, pages 421–422. ACM Press, 2003.
109. TAO LI, SHENGHUO ZHU, and MITSUNORI OGIHARA. **A New Distributed Data Mining Model Based on Similarity**. In *Proceedings of the 18th Annual ACM Symposium on Applied Computing (SAC'03)*, pages 432–436, 2003.
110. TAO LI, SHENGHUO ZHU, QI LI, and MITSUNORI OGIHARA. **Gene Functional Classification by Semi-supervised Learning from Heterogeneous Data**. In *Proceedings of the 18th Annual ACM Symposium on Applied Computing (SAC'03)*, pages 78–82, 2003.
111. QI LI, CHRIS BROWN, CHANDRA KAMBHAMETTU, TAO LI, and SHENGHUO ZHU. **A Framework of Individually-Focused Teleconferencing (IFT) via an Efficient**

- 3D Reprojection Technique.** In *Proceedings of The 18th Annual ACM Symposium on Applied Computing(SAC 2003)*, pages 951–955, 2003.
112. TAO LI and GEORGE TZANETAKIS. **Factors in Automatic Musical Genre Classification of Audio Signals.** In *Proceedings of 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2003)*, pages 143–146, 2003.
113. QI LI, TAO LI, and SHENGHUO ZHU. **Improving Medical/Biological Data Classification Performance by Wavelet Preprocessing.** In *Proceedings of The 2002 International Conference on Data Mining (ICDM 2002)*, pages 657–660, 2002.
114. TAO LI, SHENGHUO ZHU, MITSUNORI OGIHARA, and YINHE CHENG. **Estimating Joint Probabilities from Marginal Ones.** In *4th International Conference on Data Warehousing and Knowledge Discovery (DaWaK'02)*, pages 31–41. Springer-Verlag Lecture Notes in Computer Science #2454, 2002.
115. SHENGHUO ZHU, TAO LI, and MITSUNORI OGIHARA. **CoFD: An Algorithm for Non-distance Based Clustering in High Dimensional Spaces.** In *4th International Conference on Data Warehousing and Knowledge Discovery (DaWaK'02)*, pages 52–62. Springer-Verlag Lecture Notes in Computer Science #2454, 2002.
116. QI LI, TAO LI, SHENGHUO ZHU, and CHANDRA KAMBHAMETTU. **How well can wavelet denoising improve the accuracy of computing fundamental matrices?** In *Proceedings of IEEE Workshop on Motion and Video Computing*, pages 247–252, 2002.
117. SHENGHUO ZHU and TAO LI. **A Non-distance Based Clustering Algorithm.** In *Proceedings of The 2002 International Joint Conference on Neural Networks (IJCNN 2002)*, 2002.

118. TAO LI, SHENGHUO ZHU, and MITSUNORI OGIHARA. **Mining Patterns from Case-based Analysis**. In *ICDM 2001 Workshop on Integrating Data Mining and Knowledge Management*.
119. TAO LI and CHEN DING. **Instruction Balance and its Relation to Program Energy Consumption**. In *Proceedings of The 14th International Workshop on Languages and Compilers for Parallel Computing of(LCPC 2001)*, pages 71–85, 2001.

GRANT SUPPORT

Quick Summary: So far I have brought about 1.8 million research grant as **the sole PI**. I received several sole-PI awards: a CAREER Award from the National Science Foundation (2006-2011), and three IBM Faculty Research Awards (2005, 2007 & 2008). These are highly prestigious research awards. I am the first faculty member in School of Computing and Information Science (SCIS) at FIU to receive the prestigious NSF CAREER Award in 2006 (\$482K). I also received Army Research Office Award for \$300K in 2010 (note the award was originally funded as a young investigator career award, but was later upgraded as a regular grant to give me more money). I am also serving as a sole-PI of an NSF DBI Award (\$325K), a sole-PI of an NSF DMS grant (\$150K), a sole-PI of an NSF BPC supplement grant (\$96K), a PI of NSF DMS grant (\$44K), a PI of an NSF CCF grant (\$50K), a Co-PI on an NSF MRI grant (\$1000K), a Co-PI on an DHS grant (\$200K), a Co-PI on an NSF I/UCRC grant (\$350K), a Co-PI on an NIH grant (\$293K), a Co-PI on the NSF REU (Research Experience for Undergraduates) grant (\$300K), a Co-PI on NSF CREST supplement grant (\$100K), a senior investigator on NSF PIRE grant (\$2300K) and NSF CREST grant (\$5000K). I also received many prestigious industry research awards including IBM Faculty Research Awards, IBM Shared University Research Award, Xerox Research Award, and NEC research Award.

- Project Title: Knowledge Discovery and Information Delivery for Disaster Management

Funding Agency & Grant No.: IBM Scalable Data Analytics for A Smarter Planet Innovation Award

Title on Grant: PI

Period Covered: October 20, 2010 — October 20, 2011

Amount: \$20,000

- Project Title: Unsupervised Learning from Multiple Information Sources Based on Non-negative Matrix Factorization

Funding Agency & Grant No.: Army Research Office, W911NF-10-1-0366

Title on Grant: PI (sole PI)

Period Covered: August 19th, 2010 — August 31, 2013

Amount: \$300,000
- Project Title: A Data Mining Framework for Enhancing Emergency Response Situation Reports with Multi-Agency, Multi-Partner, Multimedia Data

Funding Agency & Grant No.: Purdue VACCINE/DHS 4112-35822

Title on Grant: Co-PI (PI: Shu-ching Chen)

Period Covered: April 1, 2010 — March 31, 2011

Amount: \$45,000
- Project Title: NSF/DBI: Gene Functional Prediction from Protein-Protein Interaction Data

Funding Agency & Grant No.: National Science Foundation (NSF), DBI-0850203

Title on Grant: PI (sole PI)

Period Covered: September 1, 2009 — August 31, 2012

Amount: \$325,000
- Project Title: NSF/CCF: Collaborative Research: Cross-Domain Knowledge Transformation via Matrix Decompositions

Funding Agency & Grant No.: National Science Foundation (NSF), DMS-0939179

Title on Grant: PI (sole PI)

Period Covered: September 1, 2009 — August 31, 2010

Amount: \$46,000

- Project Title: Information Delivery and Knowledge Discovery for Hurricane Disaster Management

Funding Agency & Grant No.: Department of Homeland Security, 2009-ST-062-000016

Title on Grant: Co-PI (PI: Vagelis Christidis)

Period Covered: May 1, 2009 — April 30, 2011

Amount: \$200,000

- Project Title: Collaborative Research: Non-negative Matrix Factorizations for Data Mining: Algorithms and Applications

Funding Agency & Grant No.: National Science Foundation (NSF), DMS-0915110

Title on Grant: PI (sole PI)

Period Covered: September 1, 2009 — August 31, 2012

Amount: \$150,000

- Project Title: Collaborative Research: Non-negative Matrix Factorizations for Data Mining: Algorithms and Applications

Funding Agency & Grant No.: National Science Foundation (NSF), DMS-0844513

Title on Grant: PI (sole PI)

Period Covered: September 15, 2008 — September 30, 2009

Amount: \$44,000

- Project Title: Collaborative Research: Matrix-Model Machine Learning: Unifying Machine Learning and Scientific Computing

Funding Agency & Grant No.: National Science Foundation (NSF), CCF-0830659

Title on Grant: PI (sole PI)

Period Covered: October 1, 2008 — September 30, 2011

Amount: \$62,000

- Project Title: CAREER: Mining Log Data for Computing System Management

Funding Agency & Grant No.: National Science Foundation (NSF), IIS-0546280

Title on Grant: PI (Sole PI)

Period Covered: January 1, 2006 — December 31, 2010

Amount: \$482,719

- Project Title: Integrated Genomic Databases and Microarray Data Analysis

Funding Agency & Grant No.: National Institute of Health (NIH),

Title on Grant: Co-PI (PI: Dr. Giri Narasimhan)

Period Covered: April, 2006 — April, 2008.

Amount: \$293,478

- Project Title: MRI:Development of a High-Performance Database Appliance for Geospatial Applications

Funding Agency & Grant No.: National Science Foundation (NSF), CNS-0821345

Title on Grant: Co-PI (PI: Dr. Naphtali D. Rishe)

Period Covered: August 1, 2008 — July 31, 2013.

Amount: \$1,052,000

- Project Title: CREST: Center for Innovative Information Systems Engineering

Funding Agency & Grant No.: National Science Foundation (NSF),

Title on Grant: Subcomponent Co-PI (Subcomponent PI: Dr. Shu-ching Chen)

Period Covered: August 1, 2008 — July 31, 2013.

Amount: \$5,000,000

- Project Title: Center for Advanced Knowledge Enablement

Funding Agency & Grant No.: National Science Foundation (NSF), IIP-0829576

Title on Grant: Co-PI (PI: Dr. Naphtali D. Rishe)

Period Covered: July 15, 2008 — June 30, 2013.

Amount: \$395,000

- Project Title: Research Experiences for Undergraduates: Autonomic Computing Research at FIU

Funding Agency & Grant No.: National Science Foundation (NSF), IIS-0552555

Title on Grant: co-PI (PI: Dr. Masoud Milani)

Period Covered: February 15, 2006 — February 14, 2009

Amount: \$300,000

- Project Title: Collaborative Research: Science of Search: Data Search, Analytics, and Architectures Center (DSAAC)

Funding Agency & Grant No.: National Science Foundation (NSF), IIP-0630326

Title on Grant: co-PI (PI: Dr. Naphtali Rishe)

Period Covered: August 15, 2006 — August 14, 2007

Amount: \$20,000

- Project Title: BPC Supplement for CAREER Award
Funding Agency & Grant No.: National Science Foundation (NSF)
Title on Grant: PI (Sole PI)
Period Covered: August, 2007 — December, 2010
Amount: \$96,000
- Project Title: Learning from Multi-relation Data
Funding Agency & Grant No.: NEC Laboratories America
Title on Grant: PI (Sole PI)
Period: October, 2008 — October, 2009
Amount: \$25,000
- Project Title: Intelligent Call Center
Funding Agency & Grant No.: NEC Laboratories America
Title on Grant: PI (Sole PI)
Period: October, 2006 — October, 2007
Amount: \$24,000
- Project Title: An Integrated Platform for Adaptive Workflow System — Combining knowledge-based and data-driven approaches
Funding Agency & Grant No.: Xerox Research.
Title on Grant: PI (Sole PI)
Period Covered: July, 2005 — July, 2008
Amount: \$60,000

- Project Title: Distributed Flexible Resource Scheduling in Production Systems
Funding Agency & Grant No.: NSF SBIR IIA subcontract from LS Optimal, Inc.
Title on Grant: co-PI (PI: Dr. Xudong He)
Period Covered: August, 2006 — July, 2007
Amount: \$105,000
- Project Title: REU Supplement for CAREER AWARD
Funding Agency & Grant No.: National Science Foundation (NSF), IIS-0618493
Title on Grant: PI (Sole PI)
Period Covered: April 13, 2006 — April 12, 2007
Amount: \$12,000
- Project Title: Machine Learning Approaches for Power Management
Funding Agency & Grant No.: IBM Research
Title on Grant: PI (Sole PI)
Period Covered: July, 2008 — July, 2009
Amount: \$28,000
- Project Title: Knowledge Discovery for Autonomic Computing
Funding Agency & Grant No.: IBM Research
Title on Grant: PI (Sole PI)
Period Covered: July, 2005 — July, 2006
Amount: \$40,000

- Project Title: Combining Knowledge Discovery and Adaptive Software Techniques to Build Autonomic Systems
Funding Agency & Grant No.: IBM Research.
Title on Grant: PI (Co-PI: Dr. Masoud Sadjadi)
Period Covered: June, 2005 — June, 2006
Amount: \$50,000 (Equipment)
- Project Title: Pattern Discovery for Autonomic Computing Research
Funding Agency & Grant No.: IBM Research.
Title on Grant: PI (Sole PI)
Period Covered: July, 2007 — July, 2008
Amount: \$30,000
- Project Title: Software Toolkits for Analyzing Heterogeneous Gene Expression Datasets for Cancer Diagnosis
Funding Agency & Grant No.: FIU Office for Access to Biomedical Research
Title of grant: PI (Sole PI)
Period Covered: Nov. 2006, — February, 2007
Amount: \$3,500
- Project Title: PIRE: A Global Living Laboratory for Cyberinfrastructure Application Enablement
Funding Agency & Grant No.: NSF OISE-0730065
Title on Grant: Senior Investigator (PI: Dr. Yi Deng)
Period Covered: September, 2007 — August, 2012

Amount: \$2,300,000

PROFESSIONAL ACTIVITIES

Quick Summary: I served as guest editors for the data mining and knowledge discovery journal, a conference chair or program co-chair for 13 international conferences or workshops. In addition, I have served on the program committees for more than 100 international conferences/workshops including all major conferences in data mining and information retrieval. I have also been a reviewer for more than 40 top quality journals.

1. Editorial Positions:

- **Guest Editor**, *A Special Issue on Temporal Data Mining with DMKD(Data Mining and Knowledge Discovery) Journal*
- **Guest Editor**, *A Special Issue on Data Mining using Matrices and Graphs with DMKD(Data Mining and Knowledge Discovery) Journal*

2. Conference and Workshop Organizations:

- **Workshop Co-Chair**, *ICDM 2010 Workshop on Optimization-based Methods for Emerging Data Mining Problems*
- **Workshop Co-Chair**, *Workshop on Recommender System with The 5th International Conference on Computer Science & Education (ICCSE10)*
- **Workshop Co-Chair**, *ICDM 2009 Workshop on Optimization-based Methods for Emerging Data Mining Problems*
- **Workshop Co-Chair**, *KDD 2009 Workshop on Data Mining using Matrices and Tensors*
- **Workshop Co-Chair**, *KDD 2008 Workshop on Data Mining using Matrices and Tensors*

- **Program Co-Chair**, *The Sixth International Conference on Machine Learning and Applications(ICMLA07)*
- **Conference Chair**, *The Fifth International Conference on Machine Learning and Applications(ICMLA06)*
- **Workshop Co-Chair**, *KDD 2006 Workshop on Theory and Practice of Temporal Data Mining*
- **Workshop Co-Chair**, *ICDM 2005 Workshop on Temporal Data Mining: Algorithms, Theory and Applications*
- **Workshop Co-Chair**, *ICDM 2004 Workshop on Temporal Data Mining: Algorithms, Theory and Applications*
- **Publicity Co-Chair**, *The Fourth International Conference on Advanced Data Mining and Applications (ADMA08)*
- **Local Arrangement Chair**, *The 2009 IEEE International Conference on Data Mining (ICDM 2009)*
- **Local Arrangement Chair**, *The Eighth International Conference on Machine Learning and Applications (ICMLA 2009)*

3. Technical Program Committees:

- *The 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL/HLT 2011)*
- *The 20th International World Wide Web Conference (WWW 2011)*
- *2011 SIAM International Conference on Data Mining (SDM 2011)*
- *The 2011 International Conference on Advances in Social Network Analysis and Mining (ASONAM 2011)*

- *International Conference on Machine Learning and Data Mining (MLDM 2011)*
- *The 24th International FLAIRS Conference (FLAIRS' 24)*
- *The 15th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2011)*
- *The 2nd Asian Conference on Machine Learning (ACML 2010)*
- *The 5th International Conference on Computer Science & Education (ICCSE10)*
- *The 19th ACM International Conference on Information and Knowledge Management (CIKM 2010)*
- *The 11th International Conference on Web-Age Information Management (WAIM 2010)*
- *The 2010 IEEE International Conference on Granular Computing (GrC 2010)*
- *International Workshop on Topic Feature Discovery and Opinion Mining (TF-DOM 2010)*
- *The 10th IEEE International Conference on Data Mining (ICDM 2010)*
- *The Ninth International Conference on Machine Learning and Applications (ICMLA 2010)*
- *The 22nd International Conference on Software Engineering and Knowledge Engineering (SEKE 2010)*
- *The 2010 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2010)*
- *ACM Fourth International Workshop on Data and Text Mining in Biomedical Informatics (DTMBIO'10)*
- *IEEE Symposium on Foundations and Practices of Data Mining (FPDM2010)*

- *MultiCust: 1st International Workshop on Discovering, Summarizing and Using Multiple Clusterings*
- *The 2010 International Conference on Database Theory and Application (DTA 2010)*
- *The 9th Mexican International Conference on Artificial Intelligence (MICAI 2010)*
- *Workshop on Dynamic Networks and Knowledge Discovery (DyNak 2010)*
- *The 33rd ACM SIGIR International Conference on Research and Development in Information Retrieval (SIGIR 2010)*
- *The Second International Symposium on Data, Privacy, & E-Commerce (ISDPE 2010)*
- *The 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2010)*
- *The 26th IEEE International Conference on Data Engineering (ICDE 2010)*
- *Data Mining track of the 8th Mexican International Conference on Artificial Intelligence (MICAI 2009)*
- *The Eighth International Conference on Machine Learning and Applications (ICMLA '09)*
- *The 21st IEEE Int'l Conference on Tools with Artificial Intelligence (ICTAI 2009)*
- *The 23rd International FLAIRS Conference (FLAIRS-23)*
- *International Conference on Knowledge Discovery and Information Retrieval (KDIR2009)*
- *The 24th International Symposium on Computer and Information Sciences (IS-CIS 2009)*

- *The Eleventh International Conference on Software Engineering and Knowledge Engineering (SEKE 2009)*
- *The European Conference on Data Mining (ECDM'09)*
- *The 2009 IEEE International Conference on Data Mining (ICDM 2009)*
- *The 32nd ACM SIGIR International Conference on Research and Development in Information Retrieval (SIGIR 2009)*
- *The 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD 2009)*
- *The 2009 SIAM International Conference on Data Mining 2009 (SDM 2009)*
- *The 1st Asian Conference on Machine Learning (ACML'09)*
- *The 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2009)*
- *2009 International Symposium on Database Theory and Application (DTA'09)*
- *International Conference on Machine Learning and Data Mining (MLDM'2009)*
- *2009 ACM Third International Workshop on Data and Text Mining in Bioinformatics*
- *The Joint International Conferences on Asia-Pacific Web Conference (APWeb) and Web-Age Information Management (WAIM) (APWeb-WAIM 2009)*
- *The 2009 IEEE / WIC / ACM International Conference on Web Intelligence (WI 2009)*
- *The 2009 IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2009)*
- *The 2009 IEEE International Conference on Granular Computing (GrC 2009).*

- *The 22nd International FLAIRS Conference (FLAIRS-22)*
- *The 2008 IEEE International Conference on Data Mining (ICDM 2008)*
- *The 2nd CIKM workshop of Data and Text Mining Methods in Bioinformatics (DTMbio)*
- *The 20th IEEE Int'l Conference on Tools with Artificial Intelligence (ICTAI 2008)*
- *The Seventh International Conference on Machine Learning and Applications (ICMLA'08)*
- *The 2008 IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2008)*
- *The 5th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2008)*
- *The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2008)*
- *The 2008 IEEE / WIC / ACM International Conference on Web Intelligence (WI 2008)*
- *The 31st ACM SIGIR International Conference on Research and Development in Information Retrieval (SIGIR 2008)*
- *The 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2008)*
- *The 21st International FLAIRS Conference (FLAIRS-21)*
- *The 2008 SIAM International Conference on Data Mining (SDM 2008)*
- *The 12th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2008)*

- *The Tenth International Conference on Software Engineering and Knowledge Engineering (SEKE 2008)*
- *The 2007 IEEE International Conference on Granular Computing (IEEE GrC 2007)*
- *The 2007 IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2007)*
- *The 2007 IEEE/WIC/ACM International Conference on Web Intelligence (WI-07)*
- *The 19th IEEE International Conference on Tools with Artificial Intelligence (IC-TAI 2007)*
- *The Second International Workshop on Data and Text Mining in Bioinformatics (TDMBio) in conjunction with ACM CIKM 2007*
- *Mining and Management of Biological Data (MMBD) in conjunction with IEEE ICDM 2007*
- *KDD 2007 Workshop on Mining Multiple Information Sources in conjunction with SIGKDD 2007*
- *The Thirteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(SIGKDD 2007)*
- *The Fourth IEEE International Conference on Autonomic Computing (ICAC 2007)*
- *The 2007 SIAM International Conference on Data Mining 2007 (SDM 2007)*
- *The 11th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2007)*

- *The 2007 International Workshop on High Performance Data Mining and Applications (HPMDA '07)*
- *The Nineteenth International Conference on Software Engineering and Knowledge Engineering (SEKE 2007)*
- *the 4th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)*
- *Workshop on Text Data Mining and Management (TDMM) In conjunction with 23rd IEEE ICDE Conference*
- *International Conference on Machine Learning and Data Mining (MLDM2007)*
- *The 20th International FLAIRS Conference (FLAIRS-20)*
- *IEEE Symposium on Computational Intelligence and Data Mining (CIDM 2007)*
- *The 2006 International Conference on Data Mining (ICDM 2006)*
- *The 18th IEEE International Conference on Tools with Artificial Intelligence (IC-TAI 2006)*
- *The 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI-06)*
- *The Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD2006)*
- *The 2006 IEEE International Conference on Multimedia & Expo (ICME 2006)*
- *2006 Workshop on Spatial Data Mining: Consolidation and Renewed Bearing*
- *SIAM International Conference on Data Mining 2006 (SDM 2006)*
- *The 10th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2006)*

- *The First International Conference on Scalable Information Systems (INFOSCALE)*
- *The 44th ACM Southeast Conference (ACMSE 2006)*
- *The 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2005)*
- *The 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval(SIGIR 2005)*
- *The Fourteenth ACM Conference on Information and Knowledge Management (CIKM 2005)*
- *The Thirteenth ACM Conference on Information and Knowledge Management (CIKM 2004)*

4. **Journal Referee:** Served as an ad hoc referee for the following journals:

- *ACM Transactions on Autonomous and Adaptive Systems*
- *Applied Intelligence*
- *ACM Transactions on Information Systems*
- *ACM Transactions on Intelligent Systems and Technology*
- *ACM Transactions on Knowledge Discovery from Data*
- *Bioinformatics*
- *Computational Intelligence*
- *Computational Statistics and Data Analysis*
- *Computer Music Journal*
- *Data and Knowledge Engineering*

- *Data Mining and Knowledge Discovery*
- *Distributed and Parallel Databases*
- *Engineering Applications of Artificial Intelligence*
- *EURASIP Journal on Advances in Signal Processing*
- *Knowledge and Information Systems*
- *Knowledge-based Systems*
- *IEEE Intelligent Systems*
- *IEEE Transactions on Audio, Speech, and Language Processing*
- *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*
- *IEEE Transactions on Knowledge and Data Engineering*
- *IEEE Transactions on Multimedia*
- *IEEE Transactions on Neural Network*
- *IEEE Transactions on Pattern Analysis and Machine Intelligence*
- *IEEE Transactions on Systems, Man and Cybernetics*
- *Information Processing and Management*
- *Image and Vision Computing Journal*
- *Information Sciences*
- *Information Systems*
- *International Journal of Computers and Their Applications*
- *International Journal of Pattern Recognition and Artificial Intelligence*
- *Journal of Classification*
- *Journal of Combinatorial Optimization*

- *Journal of Computational Biology*
- *Journal of Machine Learning Research*
- *Journal of Systems and Software*
- *Machine Learning Journal*
- *Neural Processing Letters*
- *Neurocomputing*
- *Pattern Recognition Letters*
- *Signal Processing*
- *Soft Computing*
- *Statistical Analysis and Data Mining*
- *Theoretical Computer Science*
- *VLDB Journal*

5. **Book Review:** Reviewed book proposals or proofread book drafts for the following publishers:

- *Chapman and Hall/CRC Press, Pearson, Springer, Wiley.*

6. **Grant Proposal Review:**

Ad hoc reviewer/panelist for: National Science Foundation (NSF), The Kentucky Science and Engineering Foundation, Qatar Science Foundation

7. **Other Activities:**

- Mentor for 32th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2009)

- Mentor for 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2007)
- Reviewer for The Eighteenth International Conference on Software Engineering and Knowledge Engineering (SEKE 2006)
- Session Chair at The Thirteenth ACM Conference on Information and Knowledge Management (CIKM 2004)
- Reviewer for International Conference on Bioinformatics and its Applications (ICBA'04)
- Reviewer for The 2002 International Conference on Data Mining (ICDM 2002)

TEACHING AND SERVICE

COURSE PROPOSALS

I developed/revised the following course proposals. I first introduced the *Principles of Data Mining* course at FIU in Fall 2004, the *Advanced Data Mining* course at FIU in Spring 2007, the *Introduction to Data Mining* course at FIU in Fall 2007, and the *Introduction to Machine Learning* course at FIU in Spring 2008.

COP4990 Introduction to Data Mining

COP5577 Principles of Data Mining,

CAP5610 Introduction to Machine Learning

CAP6778 Advanced Data Mining

COURSES TAUGHT

Fall 2004 COP5592: Principles of Data Mining
Spring 2005 COP4540: Database Management
Fall 2005 COP5577: Principles of Data Mining
Spring 2006 COP4540: Database Management
Spring 2006 COP3530: Data Structure
Fall 2006 COT5407: Introduction to Algorithms
Fall 2006 COP3530: Data Structure
Spring 2007 CAP6778: Advanced Data Mining
Spring 2007 COP4540: Database Management
Fall 2007 COP5577: Principles of Data Mining
Fall 2007 COP4990: Introduction to Data Mining
Spring 2008 CAP5610: Introduction to Machine Learning
Fall 2008 COP4990: Introduction to Data Mining
Fall 2008 COP5577: Principles of Data Mining
Fall 2009 COT5407: Introduction to Algorithms
Fall 2009 COP5577: Principles of Data Mining
Fall 2010 CAP4770: Introduction to Data Mining
Fall 2010 COP5577: Principles of Data Mining

TEACHING EVALUATION

My average ratings are summarized below:

Year 2004-2005 3.86/5.00

Year 2005-2006 4.44/5.00

Year 2006-2007 4.66/5.00

Year 2007-2008 4.84/5.00

Year 2008-2009 4.81/5.00

GRADUATE ADVISING

Quick Summary: So far, I have graduated two Ph.D. students: David Kaiser (defended in Fall 2007, currently the director of institutional research at Miami-Dade College), Wei Peng (defended in March 2008, currently a research scientist at Xerox Research Center). 9 students are in the pipeline (Dingding Wang, Yi Zhang, Bo Shao, Jingxuan Li, Li Zheng, Liang Tang, Lei Li, Chao Shen, Yexi Jiang). Three of my students, Wei Peng, Dingding Wang and Yi Zhang, have won the best graduate student research award in 2006, 2008, & 2009, an honor given to the top graduate student at SCIS who has done outstanding research during the corresponding academic year.

• Visiting Researcher

1. Lin Yuan (November 2010 - November 2011)

- Supported by China Scholarship Council
- Associate Professor, Northwestern Polytechnical University, Xi'an, Shaan Xi, P.R. China.

• Postdoc Advisee

1. Fei Wang (Fall 2008 - Summer 2009)

- Ph.D. from Tsinghua University , 2008

• Ph.D. Advisees

1. David Kaiser (Ph.D., Dec. 2007)
 - Dissertation title: The Structure of Games
 - Current Employment: Director of Institutional Research, Miami-Dade College
 - Second Place, AAAI-04 General Game Playing Competition
 - Fourth Place, AAAI-05 General Game Playing Competition
2. Yexi Jiang (Expected defense: Summer 2014)
3. Lei Li (Expected defense: Summer 2013)
4. Jingxuan Li (Expected defense: Summer 2012)
5. Wei Peng (Ph.D., April 2008)
 - Dissertation title: Actionable Pattern Analysis in Computing and BI Systems
 - Current Employment: Xerox Research
 - Outstanding PhD Graduate, College of Engineering and Computing, Florida International University (2008)
 - The Best Graduate Student Research Award, School of Computer Science, Florida International University (2006)
 - Excellence Award, School of Computer Science, 2006-07
 - Summer Intern at Xerox Research, 2006
 - Summer Intern at Xerox Research, 2007
 - FIU Presidential Fellowship 2003-2005
6. Bo Shao (Expected defense: Fall 2010)
 - Summer Intern at IBM Almaden Research Center, 2006
 - Summer Intern at IBM Autonomic Computing, 2007
 - Thesis proposal title: User-centric Music Information Retrieval

7. Chao Shen (Expected defense: Summer 2013)
8. Liang Tang (Expected defense: Summer 2013)
9. Dingding Wang (Expected defense: Spring 2010)
 - Summer Intern at NEC Research, 2009
 - Summer Intern at NEC Research, 2008
 - The Best Graduate Student Research Award, School of Computer Science, Florida International University (2008)
 - Summer Intern at NEC Research, 2007
 - Thesis proposal title: Learning to Understand Documents
10. Yi Zhang (Expected Defense: Summer 2011)
 - The Best Graduate Student Research Award, School of Computer Science, Florida International University (2009)
11. Li Zheng (Expected Defense: Summer 2012)

• **Visiting Ph.D. Students**

1. Fei Wang (June 2007 - July 2007)
 - Department of Automation, Tsinghua University, P. R. China
2. Zhongyuan Zhang (November 2007 - December 2007)
 - Institute of Mathematics and System Science, Chinese Academy of Sciences, P. R. China
3. Wenting Lu (August 2010 - March 2012)
 - Supported by China Scholarship Council
 - School of Information and Communication, Beijing University of Posts and Telecommunications, P. R. China

- **Ph.D. Thesis Committee**

- Chengyong Yang (Ph.D., Summer 2006)
- Andres Parra (Ph.D., Fall 2009)
- Erliang Zeng (Ph.D., Summer 2008)
- Ramakrishna Varadarajan (Ph.D., Spring 2009)
- Oscar Valdivia (Expected defense: Fall 2010)
- Zhongping Zhai (Expected defense: Spring 2010)
- Ting Li (Expected defense: Fall 2012)

- **Master Thesis Committee**

- Yichen Xu (Master of Industrial Engineering, 2006)

- **Ph.D. Qualifying Committee**

- Irene Polycarpou (Fall 2004)
- Jose Cartaya (Spring 2005)
- Xiaosi Zhou (Spring 2005)
- Wei Peng (Fall 2005)
- Erliang Zeng (Fall 2005)
- Haifeng Wang (Spring 2006)
- Ramakrishna Varadarajan (Fall 2006)
- Jing Zhang (Fall 2006)
- Khalid Saleem (Fall 2006)
- Bo Shao (Spring 2007)

- Khalid Saleem (Spring 2007)
- Dingding Wang (Fall 2007)
- Oscar Valdivia (Spring 2008)
- Michael Robinson (Fall 2008)
- Michael Armella (Fall 2009)
- Guangqiang Zhao (Fall 2009)
- Eduardo Ruiz (Spring 2010)

- **Independent Study**

- Xiaoli Liang (Summer 2006)
- Jing Zhang (Fall 2006)

UNDERGRADUATE ADVISING

- **Undergraduate Independent Research Supervision** Didier Garcia (Spring '06, Fall '06), Ruskin Miller (Summer '08, Fall '08)

DEPARTMENTAL COMMITTEE WORK

- 2004-2005 Equipment Committee
- 2006-2007 Annual Award Committee
- 2005-2006 Faculty Recruiting Committee
- 2006-2007 Research Colloquium Coordinator
- 2006-2007 Graduate Committee
- 2007-2008 Faculty Search Committee
- 2008-2009 Graduate Committee
- 2008-2009 Faculty Search Committee
- 2009-2010 Faculty Search Committee
- 2010-2011 Graduate Committee
- 2010-2011 Faculty Search Committee (Chair)
- 2010-2011 Annual Award Committee

SERVICE TO THE UNIVERSITY COMMUNITY

- 2005-Current Faculty Advisor at Honor College
- 2009-Current Member of the Dissertation Advisor Status Evaluation Committee
- 2009-Current Alternate Rep of SCIS for Faculty Council