

# Jimeng Sun

E-Mail: [jsun@cc.gatech.edu](mailto:jsun@cc.gatech.edu)

Update date: June 11, 2014

---

## Summary

Jimeng Sun is an Associate Professor of School of Computational Science and Engineering at College of Computing in Georgia Institute of Technology. His research focuses on health analytics and data mining, especially in applying large-scale predictive modeling and similarity analytics on biomedical applications. Dr. Sun has extensive research records on data mining and machine learning: big data analytics, similarity metric learning, social network analysis, predictive modeling, tensor analysis, and visual analytics. He applies data mining and machine learning to healthcare applications such as phenotyping from electronic health records, heart failure onset prediction and hypertension control management.

Dr. Sun has received ICDM best research paper in 2008, SDM best research paper in 2007, and KDD Dissertation runner-up award in 2008. Dr. Sun received his B.S. and M.Phil. in Computer Science from Hong Kong University of Science and Technology in 2002 and 2003, and PhD in Computer Science from Carnegie Mellon University in 2007. Prior to joining Georgia Tech, He was a research staff member at IBM TJ Watson Research Center. He has published over 80 papers, filed over 20 patents (5 granted) and an h-index of 29 and i10-index 52.

---

## Experience

### Associate Professor

Georgia Institute of Technology, GA, USA

2014-present

- Healthcare informatics
- Big data analytics
- Social network analysis

### Healthcare Analytic Lead

IBM TJ Watson Research Center, NY, USA

2009-2013

- Lead the research and development on predictive modeling technology for personalized disease risk assessment
- Lead the research and development on patient similarity technology for supporting various healthcare analytics
- Research on advanced visual analytics for exploring patient cohorts

### Service Research

IBM TJ Watson Research Center, NY, USA

2007-2009

- Design and implement a distributed data warehouse that host IT delivery data and metrics for over thousands of accounts
- Design data analytic techniques for IT service delivery data and metrics to optimize service processes.
- Research on enterprise social networks, visualization and recommendation.

### Intern, Software Tools and Techniques

IBM TJ Watson Research Center, NY, USA

2006 summer

- Developed techniques to summarize huge volume of monitoring data streams through matrix and tensor techniques
- Developed privacy-preserving publishing methods for sharing data streams

### Senior Research Associate, Center for advanced research

PricewaterhouseCoopers, CA, USA

2004 summer, 2005 summer

- Worked on data analysis of business transactional data and anomaly detection
- Developed techniques to summarize huge volume of financial streams

---

## Education

### Ph.D in Computer Science

Carnegie Mellon University, Pittsburgh, PA, USA.

2007

Thesis title: "Incremental Pattern Discovery on Streams, Graphs and Tensors" Adviser: Prof. Christos Faloutsos

### M.Sc. in Computer Science

Carnegie Mellon University, Pittsburgh, PA, USA.

2006

### M.Phil. in Computer Science,

Hong Kong University of Science and Technology, Hong Kong

2003

Thesis title: "Analysis of predictive spatio-temporal queries" Research adviser: Prof. Dimitris Papadias

# Jimeng Sun

**B.Sc. in Computer Science (minor in Mathematics)**

**Hong Kong University of Science and Technology, Hong Kong      2002**

## Awards

---

- Amazon AWS in Education Research Award 2014
- IBM Master Inventor 2013
- IBM Research Accomplishment Award for Intelligent Care Delivery Analysis 2013
- KDD'12 Best Poster Presentation Award
- AMIA'10 Distinguished Paper Award nominee
- IBM Research Accomplishment Award for Service Quality Research 2009
- ICDM'08 Best Research Paper Award
- KDD'08 Dissertation Award (runner-up)
- SDM'07 Best Research Paper
- HKUST Academic Achievement Awards 2002

## Representative Publications

---

1. Joyce Ho, Joydeep Ghosh, **Jimeng Sun**. Marble: High-throughput Phenotyping from Electronic Health Records via Sparse Nonnegative Tensor Factorization. ACM SIGKDD, 2014
2. **Jimeng Sun**, Candace D McNaughton, Ping Zhang, Adam Perer, Aris Gkoulalas-Divanis, Joshua C Denny, Jacqueline Kirby, Thomas Lasko, Alexander Saip, and Bradley A Malin. "Predicting Changes in Hypertension Control Using Electronic Health Records from a Chronic Disease Management Program." Journal of the American Medical Informatics Association: JAMIA (September 17, 2013). doi:10.1136/amiajnl-2013-002033.
3. Kenney Ng, Amol Ghoting, Steven R Steinhubl, Walter F Stewart, Bradley Malin, and **Jimeng Sun**. PARAMO: A Parallel Predictive Modeling Platform for Healthcare Analytic Research using Electronic Health Records. Journal of Biomedical Informatics
4. Fei Wang, **Jimeng Sun**, Shahram Ebadollahi: Composite distance metric integration by leveraging multiple experts' inputs and its application in patient similarity assessment. Statistical Analysis and Data Mining 5(1): 54-69 (2012) (**journal special issue for the best of SDM'11**)
5. Tamara G. Kolda and **Jimeng Sun**. Scalable Tensor Decompositions for Multi-aspect Data Mining. in Proc. IEEE International Conference on Data Mining (ICDM), 2008. (**Best Paper**). Acceptance rate: 9.7%

## Teaching Experience

---

### Georgia Tech

- VIP team: Predictive Health, [Fall 2014](#)
- CS4001: Computing & Society, [Spring 2014](#)
- CS8001: Big data analytics in healthcare, [Spring 2014](#)

### Before Georgia Tech

- Tutorial on "Healthcare Analytics for Big Data" at KDD at Aug 2013, SDM at Apr 2013
- Rutgers MBA class on "Information Technology for Managers" Spring 2012
- Tutorial on "Distance Metric Learning in Data mining" at SDM, Apr 2012
- Tutorial on "Large-scale Data Mining: Map-Reduce and Beyond" with Spiros Papadimitriou and Rong Yan
  - KDD Aug 2010, ICDM Dec 2009
- Tutorial on "Mining Large Time-Evolving Graphs Using Matrix and Tensor Tools" with Christos Faloutsos and Tammy Kolda
  - SDM Apr 2007, SIGMOD Jun 2007, ICML Jun 2007, KDD Aug 2007
- Fall 2006, Teaching Assistant, 15-415 Database Applications, Carnegie Mellon University, Computer Science, Department for prof. Christos Faloutsos
- Fall 2005, Teaching Assistant, 15-781 Machine Learning, Carnegie Mellon University, Computer Science, Department for prof. Tom Mitchell and Andrew Moore

## Professional activities and services

---

### Organizing committee

- Publication chair for KDD 2013
- Student award chair for KDD 2012, 2011
- Sponsorship chair for SDM 2012

### Guest Editor

- JBI journal special issue on medical privacy
- TKDD journal special issue on Large-scale data mining: theory and applications

### Review Committee

- Review committee for various NSF panels from 2010 to 2014 with a specialty on data mining, big data analysis, health informatics

### Program Chair

- KDD'14 [workshop on health informatics](#) (HI-KDD)
- KDD'13 workshop on data mining for healthcare
- ICML'13 workshop on Role of Machine Learning in Transforming Healthcare
- SAMSI FODAVA workshop: Interactive Visualization and Analysis of Massive Data
- KDD'11 workshops:
  - Large-scale Data Mining: Theory and Applications
  - Visual Analytics and Information Fusion
  - Data Mining for Medicine and Healthcare
- VisWeek'10 workshop: Visual Analytics in Healthcare
- ICDM'10 workshop: Large-scale Analytics for Complex Instrumented Systems
- KDD'10 workshop: LDM: Large-scale Data Mining: theory and applications
- ICDM'09 workshop: LDM: Large-scale Data Mining: theory and applications

### Program Committee Member

- ACM-BCB: [2014](#)
- KDD 2014 Senior PC, 2013 Senior PC, 2012, 2011, 2010, 2009, 2008, 2007
- SDM 2013 Senior PC, 2011 Senior PC, 2010 Senior PC, 2008
- CIKM 2014, 2012, 2011, 2010 2009, 2008
- ICDM 2012, 2011, 2010, 2009
- PKDD 2012

## Refereed Publications

---

### Journals

1. Gkoulalas-Divanis, Aris, Grigorios Loukides, and **Jimeng Sun**. "Publishing Data from Electronic Health Records While Preserving Privacy: A Survey of Algorithms." *Journal of Biomedical Informatics*, June 14, 2014. [doi:10.1016/j.jbi.2014.06.002](https://doi.org/10.1016/j.jbi.2014.06.002).
2. Gkoulalas-Divanis, Aris, Grigorios Loukides, and **Jimeng Sun**. "Toward Smarter Healthcare: Anonymizing Medical Data to Support Research Studies." *IBM Journal of Research and Development* 58, no. 1 (2014). [doi:10.1147/JRD.2013.2288173](https://doi.org/10.1147/JRD.2013.2288173).
3. Ng, Kenney, Amol Ghoting, Steven R. Steinhubl, Walter F. Stewart, Bradley Malin, and **Jimeng Sun**. "PARAMO: A PARALLEL Predictive MOdeling Platform for Healthcare Analytic Research Using Electronic Health Records." *Journal of Biomedical Informatics* 48 (April 2014): 160–70. [doi:10.1016/j.jbi.2013.12.012](https://doi.org/10.1016/j.jbi.2013.12.012).
4. **Sun, Jimeng**, Candace D. McNaughton, Ping Zhang, Adam Perer, Aris Gkoulalas-Divanis, Joshua C. Denny, Jacqueline Kirby, Thomas Lasko, Alexander Saip, and Bradley A. Malin. "Predicting Changes in Hypertension Control Using Electronic Health Records from a Chronic Disease Management Program." *Journal of the American Medical Informatics Association: JAMIA* 21, no. 2 (April 2014): 337–44. [doi:10.1136/amiainl-2013-002033](https://doi.org/10.1136/amiainl-2013-002033).
5. Vijayakrishnan, Rajakrishnan, Steven R. Steinhubl, Kenney Ng, **Jimeng Sun**, Roy J. Byrd, Zahra Daar, Brent A. Williams, null Christopher Defilippi, Shahram Ebadollahi, and Walter F. Stewart. "Prevalence of Heart Failure Signs and Symptoms in a Large

- Primary Care Population Identified Through the Use of Text and Data Mining of the Electronic Health Record.” *Journal of Cardiac Failure*, April 4, 2014. doi:10.1016/j.cardfail.2014.03.008.
6. Byrd, Roy J., Steven R. Steinhubl, **Jimeng Sun**, Shahram Ebadollahi, and Walter F. Stewart. “Automatic Identification of Heart Failure Diagnostic Criteria, Using Text Analysis of Clinical Notes from Electronic Health Records.” *International Journal of Medical Informatics*, January 10, 2013. doi:10.1016/j.ijmedinf.2012.12.005.
  7. Wang, Fei, Noah Lee, Jianying Hu, **Jimeng Sun**, Shahram Ebadollahi, and Andrew F. Laine. “A Framework for Mining Signatures from Event Sequences and Its Applications in Healthcare Data.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35, no. 2 (February 2013): 272–85. doi:10.1109/TPAMI.2012.111.
  8. Gotz, David, **Jimeng Sun**, and Nan Cao. “Multifaceted Visual Analytics for Healthcare Applications.” *IBM Journal of Research and Development* 56, no. 5 (2012): 6. doi:10.1147/JRD.2012.2199170.
  9. Kang, U., Hanghang Tong, **Jimeng Sun**, Ching-Yung Lin, and Christos Faloutsos. “Gbase: An Efficient Analysis Platform for Large Graphs.” *VLDB J.* 21, no. 5 (2012): 637–50. doi:10.1007/s00778-012-0283-9.
  10. Markatou, Marianthi, Prabani Kuruppumullage Don, Jianying Hu, Fei Wang, **Jimeng Sun**, Robert Sorrentino, and Shahram Ebadollahi. “Case-Based Reasoning in Comparative Effectiveness Research.” *IBM Journal of Research and Development* 56, no. 5 (2012): 4. doi:10.1147/JRD.2012.2198311.
  11. Sondhi, Parikshit, **Jimeng Sun**, ChengXiang Zhai, Robert Sorrentino, and Martin S. Kohn. “Leveraging Medical Thesauri and Physician Feedback for Improving Medical Literature Retrieval for Case Queries.” *Journal of the American Medical Informatics Association: JAMIA* 19, no. 5 (October 2012): 851–58. doi:10.1136/amiajnl-2011-000293.
  12. Sow, Daby M., **Jimeng Sun**, Alain Biem, Jianying Hu, Marion Blount, and Shahram Ebadollahi. “Real-Time Analysis for Short-Term Prognosis in Intensive Care.” *IBM Journal of Research and Development* 56, no. 5 (2012): 3. doi:10.1147/JRD.2012.2197952.
  13. **Sun, Jimeng**, Fei Wang, Jianying Hu, and Shahram Ebadollahi. “Supervised Patient Similarity Measure of Heterogeneous Patient Records.” *SIGKDD Explorations* 14, no. 1 (2012): 16–24. <http://sigkdd.org/sites/default/files/issues/V14-01-03-Sun.pdf>
  14. Tang, Jie, Yuan Zhang, **Jimeng Sun**, Jinghai Rao, Wenjing Yu, Yiran Chen, and Alvis Cheuk M. Fong. “Quantitative Study of Individual Emotional States in Social Networks.” *T. Affective Computing* 3, no. 2 (2012): 132–44.
  15. Wang, Fei, **Jimeng Sun**, and Shahram Ebadollahi. “Composite Distance Metric Integration by Leveraging Multiple Experts’ Inputs and Its Application in Patient Similarity Assessment.” *Statistical Analysis and Data Mining* 5, no. 1 (2012): 54–69. doi:10.1002/sam.11135
  16. Cao, Nan, David Gotz, **Jimeng Sun**, and Huamin Qu. “DICON: Interactive Visual Analysis of Multidimensional Clusters.” *IEEE Transactions on Visualization and Computer Graphics* 17, no. 12 (December 2011): 2581–90. doi:10.1109/TVCG.2011.188.
  17. Lin, Yu-Ru, **Jimeng Sun**, Hari Sundaram, Aisling Kelliher, Paul Castro, and Ravi B. Konuru. “Community Discovery via Metagraph Factorization.” *TKDD* 5, no. 3 (2011): 17. doi:10.1145/1993077.1993081
  18. Cao, Nan, **Jimeng Sun**, Yu-Ru Lin, David Gotz, Shixia Liu, and Huamin Qu. “FacetAtlas: Multifaceted Visualization for Rich Text Corpora.” *IEEE Transactions on Visualization and Computer Graphics* 16, no. 6 (December 2010): 1172–81. doi:10.1109/TVCG.2010.154.
  19. Cao, Nan, **Jimeng Sun**, Yu-Ru Lin, David Gotz, Shixia Liu, and Huamin Qu. “FacetAtlas: Multifaceted Visualization for Rich Text Corpora.” *IEEE Transactions on Visualization and Computer Graphics* 16, no. 6 (December 2010): 1172–81. doi:10.1109/TVCG.2010.154.
  20. Khan, Asheq, Hani Jamjoom, and **Jimeng Sun**. “AIM-HI: A Framework for Request Routing in Large-Scale IT Global Service Delivery.” *IBM Journal of Research and Development* 53, no. 6 (2009): 4. doi:10.1147/JRD.2009.5429032.
  21. **Sun, Jimeng**. “Incremental Pattern Discovery on Streams, Graphs and Tensors.” *SIGKDD Explorations* 10, no. 2 (2008): 28–29.
  22. **Sun, Jimeng**, Dacheng Tao, Spiros Papadimitriou, Philip S. Yu, and Christos Faloutsos. “Incremental Tensor Analysis: Theory and Applications.” *TKDD* 2, no. 3 (2008). <http://doi.acm.org/10.1145/1409620.1409621>.
  23. **Sun, Jimeng**, Charalampos E. Tsourakakis, Evan Hoke, Christos Faloutsos, and Tina Eliassi-Rad. “Two Heads Better than One: Pattern Discovery in Time-Evolving Multi-Aspect Data.” *Data Min. Knowl. Discov.* 17, no. 1 (2008): 111–28. doi:10.1007/s10618-008-0112-3.
  24. **Sun, Jimeng**, Yinglian Xie, Hui Zhang, and Christos Faloutsos. “Less Is More: Sparse Graph Mining with Compact Matrix Decomposition.” *Statistical Analysis and Data Mining* 1, no. 1 (2008): 6–22. doi:10.1002/sam.102.
  25. Tao, Dacheng, Mingli Song, Xuelong Li, Jialie Shen, **Jimeng Sun**, Xindong Wu, Christos Faloutsos, and Stephen J. Maybank. “Bayesian Tensor Approach for 3-D Face Modeling.” *IEEE Trans. Circuits Syst. Video Techn.* 18, no. 10 (2008): 1397–1410. doi:10.1109/TCSVT.2008.2002825.
  26. Hoke, Evan, **Jimeng Sun**, John D. Strunk, Gregory R. Ganger, and Christos Faloutsos. “InteMon: Continuous Mining of Sensor Data in Large-Scale Self-Infrastructures.” *Operating Systems Review* 40, no. 3 (2006): 38–44.
  27. **Sun, Jimeng**, Yufei Tao, Dimitris Papadias, and George Kollios. “Spatio-Temporal Join Selectivity.” *Inf. Syst.* 31, no. 8 (2006): 793–813. doi:10.1016/j.is.2005.02.002.
  28. **Sun, Jimeng**, Huiming Qu, Deepayan Chakrabarti, and Christos Faloutsos. “Relevance Search and Anomaly Detection in Bipartite Graphs.” *SIGKDD Explorations* 7, no. 2 (2005): 48–55.

# Jimeng Sun

29. Tao, Yufei, **Jimeng Sun**, and Dimitris Papadias. “Analysis of Predictive Spatio-Temporal Queries.” *ACM Trans. Database Syst.* 28, no. 4 (2003): 295–336.
30. Papadias, Dimitris, Yufei Tao, Jun Zhang, Nikos Mamoulis, Qiongmao Shen, and **Jimeng Sun**. “Indexing and Retrieval of Historical Aggregate Information about Moving Objects.” *IEEE Data Eng. Bull.* 25, no. 2 (2002): 10–17.

## Conferences

31. Jiayu Zhou, Yashu Liu, **Jimeng Sun**, Hu, Jianying, and Jieping Ye. “Patient Risk Prediction Model via Top-K Stability Selection.” In *Proceedings of the 13th SIAM International Conference on Data Mining*, May 2–4, 2013. Austin, Texas, USA, 55–63. SIAM, 2013. doi:10.1137/1.9781611972832.7.
32. Tang, Jie, Sen Wu, and **Jimeng Sun**. “Confluence: Conformity Influence in Large Social Networks.” In *The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2013*, Chicago, IL, USA, August 11–14, 2013, 347–55. ACM, 2013. <http://doi.acm.org/10.1145/2487575.2487691>.
33. Wu, Sen, **Jimeng Sun**, and Jie Tang. “Patent Partner Recommendation in Enterprise Social Networks.” In *Sixth ACM International Conference on Web Search and Data Mining, WSDM 2013*, Rome, Italy, February 4–8, 2013, 43–52. ACM, 2013. <http://doi.acm.org/10.1145/2433396.2433404>.
34. Zhou, Jiayu, Zhaosong Lu, **Jimeng Sun**, Lei Yuan, Fei Wang, and Jieping Ye. “FeaFiner: Biomarker Identification from Medical Data through Feature Generalization and Selection.” In *The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2013*, Chicago, IL, USA, August 11–14, 2013, 1034–42. ACM, 2013. <http://doi.acm.org/10.1145/2487575.2487671>.
35. Gotz, David, Harry Stavropoulos, **Jimeng Sun**, and Fei Wang. “ICDA: A Platform for Intelligent Care Delivery Analytics.” *AMIA* (2012): 264–73.
36. Hu, Jianying, Fei Wang, **Jimeng Sun**, Robert Sorrentino, and Shahram Ebadollahi. “A Healthcare Utilization Analysis Framework for Hot Spotting and Contextual Anomaly Detection.” 2012 (2012): 360–69.
37. Kang, U., Hanghang Tong, and **Jimeng Sun**. “Fast Random Walk Graph Kernel.” In *Proceedings of the Twelfth SIAM International Conference on Data Mining*, Anaheim, California, USA, April 26–28, 2012, 828–38. SIAM / Omnipress, 2012. doi:10.1137/1.9781611972825.71.
38. Luo, Dijun, Fei Wang, **Jimeng Sun**, Marianthi Markatou, Jianying Hu, and Shahram Ebadollahi. “SOR: Scalable Orthogonal Regression for Low-Redundancy Feature Selection and Its Healthcare Applications.” In *Proceedings of the Twelfth SIAM International Conference on Data Mining*, Anaheim, California, USA, April 26–28, 2012, 576–87. SIAM / Omnipress, 2012. doi:10.1137/1.9781611972825.50.
39. Miao, Gengxin, Ziyu Guan, Louise E. Moser, Xifeng Yan, Shu Tao, Nikos Anerousis, and **Jimeng Sun**. “Latent Association Analysis of Document Pairs.” In *The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’12*, Beijing, China, August 12–16, 2012, edited by Qiang Yang, Deepak Agarwal, and Jian Pei, 1415–23. ACM, 2012. <http://doi.acm.org/10.1145/2339530.2339752>.
40. Perer, Adam, and **Jimeng Sun**. “MatrixFlow: Temporal Network Visual Analytics to Track Symptom Evolution during Disease Progression.” *AMIA* 2012 (2012): 716–25.
41. Sondhi, Parikshit, **Jimeng Sun**, Hanghang Tong, and ChengXiang Zhai. “SympGraph: A Framework for Mining Clinical Notes through Symptom Relation Graphs.” In *The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’12*, Beijing, China, August 12–16, 2012, , 1167–75. ACM, 2012. <http://doi.acm.org/10.1145/2339530.2339712>.
42. **Sun, Jimeng**, Jianying Hu, Dijun Luo, Marianthi Markatou, Fei Wang, Shahram Ebadollahi, Steven E. Steinhubl, Zahra Daar, and Walter F. Stewart. “Combining Knowledge and Data Driven Insights for Identifying Risk Factors Using Electronic Health Records.” *AMIA* 2012 (2012): 901–10.
43. Tang, Jie, Sen Wu, **Jimeng Sun**, and Hang Su. “Cross-Domain Collaboration Recommendation.” In *The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’12*, Beijing, China, August 12–16, 2012, 1285–93. ACM, 2012. <http://doi.acm.org/10.1145/2339530.2339730>.
44. Wang, Fei, Jianying Hu, and **Jimeng Sun**. “Medical Prognosis Based on Patient Similarity and Expert Feedback.” In *Proceedings of the 21st International Conference on Pattern Recognition, ICPR 2012*, Tsukuba, Japan, November 11–15, 2012, 1799–1802. IEEE, 2012. [http://ieeexplore.ieee.org/xpl/freeabs\\_all.jsp?arnumber=6460501](http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=6460501).
45. Wang, Fei, Noah Lee, Jianying Hu, **Jimeng Sun**, and Shahram Ebadollahi. “Towards Heterogeneous Temporal Clinical Event Pattern Discovery: A Convolutional Approach.” In *The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’12*, Beijing, China, August 12–16, 2012, 453–61. ACM, 2012. <http://doi.acm.org/10.1145/2339530.2339605>.
46. Cao, Nan, David Gotz, **Jimeng Sun**, Yu-Ru Lin, and Huamin Qu. “SolarMap: Multifaceted Visual Analytics for Topic Exploration.” In *11th IEEE International Conference on Data Mining, ICDM 2011*, Vancouver, BC, Canada, December 11–14, 2011, 101–10. IEEE, 2011. <http://doi.ieeecomputersociety.org/10.1109/ICDM.2011.135>.
47. Gotz, David, **Jimeng Sun**, Nan Cao, and Shahram Ebadollahi. “Visual Cluster Analysis in Support of Clinical Decision Intelligence.” *AMIA* 2011 (2011): 481–90.

## Jimeng Sun

48. Kang, U., Spiros Papadimitriou, **Jimeng Sun**, and Hanghang Tong. "Centralities in Large Networks: Algorithms and Observations." In Proceedings of the Eleventh SIAM International Conference on Data Mining, SDM 2011, April 28-30, 2011, Mesa, Arizona, USA, 119–30. SIAM / Omnipress, 2011. doi:10.1137/1.9781611972818.11.
49. Kang, U., Hanghang Tong, **Jimeng Sun**, Ching-Yung Lin, and Christos Faloutsos. "GBASE: A Scalable and General Graph Management System." In Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, August 21-24, 2011, 1091–99. ACM, 2011. <http://doi.acm.org/10.1145/2020408.2020580>.
50. Lee, Noah, Andrew F. Laine, Jianying Hu, Fei Wang, **Jimeng Sun**, and Shahram Ebadollahi. "Mining Electronic Medical Records to Explore the Linkage between Healthcare Resource Utilization and Disease Severity in Diabetic Patients." In 2011 IEEE International Conference on Healthcare Informatics, Imaging and Systems Biology, HISB 2011, San Jose, CA, USA, July 26-29, 2011, 250–57. IEEE, 2011. <http://doi.ieeecomputersociety.org/10.1109/HISB.2011.34>.
51. Wang, Chi, Jie Tang, **Jimeng Sun**, and Jiawei Han. "Dynamic Social Influence Analysis through Time-Dependent Factor Graphs." In International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2011, Kaohsiung, Taiwan, 25-27 July 2011, 239–46. IEEE Computer Society, 2011. doi:10.1109/ASONAM.2011.116.
52. Wang, Fei, Noah Lee, **Jimeng Sun**, Jianying Hu, and Shahram Ebadollahi. "Automatic Group Sparse Coding." In Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2011, San Francisco, California, USA, August 7-11, 2011. <http://www.aaai.org/ocs/index.php/AAAI/AAAI11/paper/view/3717>.
53. Wang, Fei, **Jimeng Sun**, and Shahram Ebadollahi. "Integrating Distance Metrics Learned from Multiple Experts and Its Application in Inter-Patient Similarity Assessment." In Proceedings of the Eleventh SIAM International Conference on Data Mining, SDM 2011, April 28-30, 2011, Mesa, Arizona, USA, 59–70. SIAM / Omnipress, 2011. doi:10.1137/1.9781611972818.6.
54. Wang, Fei, **Jimeng Sun**, Jianying Hu, and Shahram Ebadollahi. "iMet: Interactive Metric Learning in Healthcare Applications." In Proceedings of the Eleventh SIAM International Conference on Data Mining, SDM 2011, April 28-30, 2011, Mesa, Arizona, USA, 944–55. SIAM / Omnipress, 2011. doi:10.1137/1.9781611972818.81.
55. Ebadollahi, Shahram, **Jimeng Sun**, David Gotz, Jianying Hu, Daby Sow, and Chalapathy Neti. "Predicting Patient's Trajectory of Physiological Data Using Temporal Trends in Similar Patients: A System for Near-Term Prognostics." AMIA 2010 (2010): 192–96.
56. Lin, Yu-Ru, **Jimeng Sun**, Nan Cao, and Shixia Liu. "ContexTour: Contextual Contour Analysis on Dynamic Multi-Relational Clustering." In Proceedings of the SIAM International Conference on Data Mining, SDM 2010, April 29 - May 1, 2010, Columbus, Ohio, USA, 418–29. SIAM, 2010. doi:10.1137/1.9781611972801.37.
57. Sondhi, Parikshit, **Jimeng Sun**, ChengXiang Zhai, Robert Sorrentino, Martin S. Kohn, Shahram Ebadollahi, and Yanen Li. "Medical Case-Based Retrieval by Leveraging Medical Ontology and Physician Feedback: UIUC-IBM at ImageCLEF 2010." In CLEF 2010 LABs and Workshops, Notebook Papers, 22-23 September 2010, Padua, Italy, 2010. <http://www.clef-initiative.eu/documents/71612/86374/CLEF2010wn-ImageCLEF-SondhiEt2010.pdf>.
58. Sow, Daby, Alain Biem, **Jimeng Sun**, Jianying Hu, and Shahram Ebadollahi. "Real-Time Prognosis of ICU Physiological Data Streams." Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 2010 (2010): 6785–88. doi:10.1109/IEMBS.2010.5625983.
59. **Sun, Jimeng**, David Gotz, and Nan Cao. "DiseaseAtlas: Multi-Facet Visual Analytics for Online Disease Articles." Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 2010 (2010): 1123–26. doi:10.1109/IEMBS.2010.5627103.
60. **Sun, Jimeng**, Daby M. Sow, Jianying Hu, and Shahram Ebadollahi. "A System for Mining Temporal Physiological Data Streams for Advanced Prognostic Decision Support." In ICDM 2010, The 10th IEEE International Conference on Data Mining, Sydney, Australia, 14-17 December 2010, 1061–66. IEEE Computer Society, 2010. <http://doi.ieeecomputersociety.org/10.1109/ICDM.2010.102>.
61. **Sun, Jimeng**, Daby M. Sow, Jianying Hu, and Shahram Ebadollahi. "Localized Supervised Metric Learning on Temporal Physiological Data." In 20th International Conference on Pattern Recognition, ICPR 2010, Istanbul, Turkey, 23-26 August 2010, 4149–52. IEEE, 2010. doi:10.1109/ICPR.2010.1009.
62. Tan, Chenhao, Jie Tang, **Jimeng Sun**, Quan Lin, and Fengjiao Wang. "Social Action Tracking via Noise Tolerant Time-Varying Factor Graphs." In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, July 25-28, 2010, 1049–58. ACM, 2010. <http://doi.acm.org/10.1145/1835804.1835936>.
63. Xiang, Liang, Quan Yuan, Shiwang Zhao, Li Chen, Xiatian Zhang, Qing Yang, and **Jimeng Sun**. "Temporal Recommendation on Graphs via Long- and Short-Term Preference Fusion." In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, July 25-28, 2010, 723–32. ACM, 2010. <http://doi.acm.org/10.1145/1835804.1835896>.
64. Zhang, Duo, **Jimeng Sun**, ChengXiang Zhai, Abhijit Bose, and Nikos Anerousis. "PTM: Probabilistic Topic Mapping Model for Mining Parallel Document Collections." In Proceedings of the 19th ACM Conference on Information and Knowledge Management, CIKM 2010, Toronto, Ontario, Canada, October 26-30, 2010, 1653–56. ACM, 2010. <http://doi.acm.org/10.1145/1871437.1871696>.

65. Zhang, Yuan, Jie Tang, **Jimeng Sun**, Yiran Chen, and Jinghai Rao. "MoodCast: Emotion Prediction via Dynamic Continuous Factor Graph Model." In ICDM 2010, The 10th IEEE International Conference on Data Mining, Sydney, Australia, 14-17 December 2010, 1193–98. IEEE Computer Society, 2010. <http://doi.ieeecomputersociety.org/10.1109/ICDM.2010.105>.
66. Lin, Ching-Yung, Nan Cao, Shixia Liu, Spiros Papadimitriou, **Jimeng Sun**, and Xifeng Yan. "SmallBlue: Social Network Analysis for Expertise Search and Collective Intelligence." In Proceedings of the 25th International Conference on Data Engineering, ICDE 2009, March 29 2009 - April 2 2009, Shanghai, China, 1483–86. IEEE, 2009. <http://doi.ieeecomputersociety.org/10.1109/ICDE.2009.140>.
67. Lin, Yu-Ru, **Jimeng Sun**, Paul Castro, Ravi B. Konuru, Hari Sundaram, and Aisling Kelliher. "Extracting Community Structure through Relational Hypergraphs." In Proceedings of the 18th International Conference on World Wide Web, WWW 2009, Madrid, Spain, April 20-24, 2009, 1213–14. ACM, 2009. <http://doi.acm.org/10.1145/1526709.1526934>.
68. Lin, Yu-Ru, **Jimeng Sun**, Paul Castro, Ravi B. Konuru, Hari Sundaram, and Aisling Kelliher. "MetaFac: Community Discovery via Relational Hypergraph Factorization." In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, June 28 - July 1, 2009, 527–36. ACM, 2009. <http://doi.acm.org/10.1145/1557019.1557080>.
69. Shi, Lei, Nan Cao, Shixia Liu, Weihong Qian, Li Tan, Guodong Wang, **Jimeng Sun**, and Ching-Yung Lin. "HiMap: Adaptive Visualization of Large-Scale Online Social Networks." In IEEE Pacific Visualization Symposium PacificVis 2009, Beijing, China, April 20-23, 2009, 41–48. IEEE Computer Society, 2009. <http://doi.ieeecomputersociety.org/10.1109/PACIFICVIS.2009.4906836>.
70. **Sun, Jimeng**, Spiros Papadimitriou, Ching-Yung Lin, Nan Cao, Shixia Liu, and Weihong Qian. "MultiVis: Content-Based Social Network Exploration through Multi-Way Visual Analysis." In Proceedings of the SIAM International Conference on Data Mining, SDM 2009, April 30 - May 2, 2009, Sparks, Nevada, USA, 1064–75. SIAM, 2009. doi:10.1137/1.9781611972795.91.
71. Tang, Jie, **Jimeng Sun**, Chi Wang, and Zi Yang. "Social Influence Analysis in Large-Scale Networks." In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, June 28 - July 1, 2009, 807–16. ACM, 2009. <http://doi.acm.org/10.1145/1557019.1557108>.
72. Wang, Fei, **Jimeng Sun**, Tao Li, and Nikos Anerousis. "Two Heads Better Than One: Metric+Active Learning and Its Applications for IT Service Classification." In ICDM 2009, The Ninth IEEE International Conference on Data Mining, Miami, Florida, USA, 6-9 December 2009, 1022–27. IEEE Computer Society, 2009. <http://doi.ieeecomputersociety.org/10.1109/ICDM.2009.103>.
73. Boutsidis, Christos, **Jimeng Sun**, and Nikos Anerousis. "Clustered Subset Selection and Its Applications on It Service Metrics." In Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM 2008, Napa Valley, California, USA, October 26-30, 2008, edited by James G. Shanahan, Sihem Amer-Yahia, Ioana Manolescu, Yi Zhang, David A. Evans, Aleksander Kolcz, Key-Sun Choi, and Abdur Chowdhury, 599–608. ACM, 2008. <http://doi.acm.org/10.1145/1458082.1458162>.
74. Kolda, Tamara G., and **Jimeng Sun**. "Scalable Tensor Decompositions for Multi-Aspect Data Mining." In Proceedings of the 8th IEEE International Conference on Data Mining (ICDM 2008), December 15-19, 2008, Pisa, Italy, 363–72. IEEE Computer Society, 2008. <http://doi.ieeecomputersociety.org/10.1109/ICDM.2008.89>.
75. Papadimitriou, Spiros, and **Jimeng Sun**. "DisCo: Distributed Co-Clustering with Map-Reduce: A Case Study towards Petabyte-Scale End-to-End Mining." In Proceedings of the 8th IEEE International Conference on Data Mining (ICDM 2008), December 15-19, 2008, Pisa, Italy, 512–21. IEEE Computer Society, 2008. <http://doi.ieeecomputersociety.org/10.1109/ICDM.2008.142>.
76. Papadimitriou, Spiros, **Jimeng Sun**, Christos Faloutsos, and Philip S. Yu. "Hierarchical, Parameter-Free Community Discovery." In Machine Learning and Knowledge Discovery in Databases, European Conference, ECML/PKDD 2008, Antwerp, Belgium, September 15-19, 2008, Proceedings, Part II, edited by Walter Daelemans, Bart Goethals, and Katharina Morik, 5212:170–87. Lecture Notes in Computer Science. Springer, 2008. doi:10.1007/978-3-540-87481-2\_12.
77. Qu, Huiming, **Jimeng Sun**, and Hani Jamjoom. "SCOOP: Automated Social Recommendation in Enterprise Process Management." In 2008 IEEE International Conference on Services Computing (SCC 2008), 8-11 July 2008, Honolulu, Hawaii, USA, 101–8. IEEE Computer Society, 2008. <http://doi.ieeecomputersociety.org/10.1109/SCC.2008.153>.
78. Sun, Jimeng, Charalampos E. Tsourakakis, Evan Hoke, Christos Faloutsos, and Tina Eliassi-Rad. "Two Heads Better Than One: Pattern Discovery in Time-Evolving Multi-Aspect Data." In Machine Learning and Knowledge Discovery in Databases, European Conference, ECML/PKDD 2008, Antwerp, Belgium, September 15-19, 2008, 5211-22. Lecture Notes in Computer Science. Springer, 2008. doi:10.1007/978-3-540-87479-9\_19.
79. Tao, Dacheng, **Jimeng Sun**, Jialie Shen, Xindong Wu, Xuelong Li, Stephen J. Maybank, and Christos Faloutsos. "Bayesian Tensor Analysis." In Proceedings of the International Joint Conference on Neural Networks, IJCNN 2008, Part of the IEEE World Congress on Computational Intelligence, WCCI 2008, Hong Kong, China, June 1-6, 2008, 1402–9. IEEE, 2008. doi:10.1109/IJCNN.2008.4633981.
80. Tong, Hanghang, Spiros Papadimitriou, **Jimeng Sun**, Philip S. Yu, and Christos Faloutsos. "Colibri: Fast Mining of Large Static and Dynamic Graphs." In Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, Nevada, USA, August 24-27, 2008, 686–94. ACM, 2008. <http://doi.acm.org/10.1145/1401890.1401973>.

81. Li, Feifei, Jimeng Sun, Spiros Papadimitriou, George A. Mihaila, and Ioana Stanoi. "Hiding in the Crowd: Privacy Preservation on Evolving Streams through Correlation Tracking." In Proceedings of the 23rd International Conference on Data Engineering, ICDE 2007, The Marmara Hotel, Istanbul, Turkey, April 15-20, 2007, 686–95. IEEE, 2007. <http://doi.ieeecomputersociety.org/10.1109/ICDE.2007.367914>.
82. **Sun, Jimeng**, Christos Faloutsos, Spiros Papadimitriou, and Philip S. Yu. "GraphScope: Parameter-Free Mining of Large Time-Evolving Graphs." In Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Jose, California, USA, August 12-15, 2007, 687–96. ACM, 2007. <http://doi.acm.org/10.1145/1281192.1281266>.
83. **Sun, Jimeng**, Yinglian Xie, Hui Zhang, and Christos Faloutsos. "Less Is More: Compact Matrix Decomposition for Large Sparse Graphs." In Proceedings of the Seventh SIAM International Conference on Data Mining, April 26-28, 2007, Minneapolis, Minnesota, USA, 366–77. SIAM, 2007. doi:10.1137/1.9781611972771.33.
84. Tao, Dacheng, **Jimeng Sun**, Xindong Wu, Xuelong Li, Jialie Shen, Stephen J. Maybank, and Christos Faloutsos. "Probabilistic Tensor Analysis with Akaike and Bayesian Information Criteria." In Neural Information Processing, 14th International Conference, ICONIP 2007, Kitakyushu, Japan, November 13-16, 2007, 4984:791–801. Lecture Notes in Computer Science. Springer, 2007. doi:10.1007/978-3-540-69158-7\_82.
85. Wei, Xing, **Jimeng Sun**, and Xuerui Wang. "Dynamic Mixture Models for Multiple Time-Series." In IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007, 2909–14, 2007. <http://dli.iit.ac.in/ijcai/IJCAI-2007/PDF/IJCAI07-468.pdf>.
86. Hoke, Evan, **Jimeng Sun**, and Christos Faloutsos. "InteMon: Intelligent System Monitoring on Large Clusters." In Proceedings of the 32nd International Conference on Very Large Data Bases, Seoul, Korea, September 12-15, 2006, 1239–42. ACM, 2006. <http://www.vldb.org/conf/2006/p1239-hoke.pdf>.
87. Papadimitriou, Spiros, **Jimeng Sun**, and Philip S. Yu. "Local Correlation Tracking in Time Series." In Proceedings of the 6th IEEE International Conference on Data Mining (ICDM 2006), 18-22 December 2006, Hong Kong, China, 456–65. IEEE Computer Society, 2006. <http://doi.ieeecomputersociety.org/10.1109/ICDM.2006.99>.
88. **Sun, Jimeng**, Evan Hoke, John D. Strunk, Gregory R. Ganger, and Christos Faloutsos. "Intelligent System Monitoring on Large Clusters." In Proceedings of the 3rd Workshop on Data Management for Sensor Networks, in Conjunction with VLDB, DMSN 2006, Seoul, Korea, September 11, 2006, 47–52. ACM International Conference Proceeding Series. ACM, 2006. <http://doi.acm.org/10.1145/1315903.1315914>.
89. **Sun, Jimeng**, Spiros Papadimitriou, and Christos Faloutsos. "Distributed Pattern Discovery in Multiple Streams." In Advances in Knowledge Discovery and Data Mining, 10th Pacific-Asia Conference, PAKDD 2006, Singapore, April 9-12, 2006, Proceedings, 3918:713–18. Lecture Notes in Computer Science. Springer, 2006. doi:10.1007/11731139\_82.
90. **Sun, Jimeng**, Spiros Papadimitriou, and Philip S. Yu. "Window-Based Tensor Analysis on High-Dimensional and Multi-Aspect Streams." In Proceedings of the 6th IEEE International Conference on Data Mining (ICDM 2006), 18-22 December 2006, Hong Kong, China, 1076–80. IEEE Computer Society, 2006. <http://doi.ieeecomputersociety.org/10.1109/ICDM.2006.169>.
91. **Sun, Jimeng**, Dacheng Tao, and Christos Faloutsos. "Beyond Streams and Graphs: Dynamic Tensor Analysis." In Proceedings of the Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Philadelphia, PA, USA, August 20-23, 2006, 374–83. ACM, 2006. <http://doi.acm.org/10.1145/1150402.1150445>.
92. Papadimitriou, Spiros, **Jimeng Sun**, and Christos Faloutsos. "Streaming Pattern Discovery in Multiple Time-Series." In Proceedings of the 31st International Conference on Very Large Data Bases, Trondheim, Norway, August 30 - September 2, 2005, 697–708. ACM, 2005. <http://www.vldb2005.org/program/paper/thu/p697-papadimitriou.pdf>.
93. **Sun, Jimeng**, Spiros Papadimitriou, and Christos Faloutsos. "Online Latent Variable Detection in Sensor Networks." In Proceedings of the 21st International Conference on Data Engineering, ICDE 2005, 5-8 April 2005, Tokyo, Japan, 1126–27. IEEE Computer Society, 2005. <http://doi.ieeecomputersociety.org/10.1109/ICDE.2005.100>.
94. **Sun, Jimeng**, Huiming Qu, Deepayan Chakrabarti, and Christos Faloutsos. "Neighborhood Formation and Anomaly Detection in Bipartite Graphs." In Proceedings of the 5th IEEE International Conference on Data Mining (ICDM 2005), 27-30 November 2005, Houston, Texas, USA, 418–25. IEEE Computer Society, 2005. <http://doi.ieeecomputersociety.org/10.1109/ICDM.2005.103>.
95. **Sun, Jimeng**, Dimitris Papadias, Yufei Tao, and Bin Liu. "Querying about the Past, the Present, and the Future in Spatio-Temporal." In Proceedings of the 20th International Conference on Data Engineering, ICDE 2004, 30 March - 2 April 2004, Boston, MA, USA, 202–13. IEEE Computer Society, 2004. <http://doi.ieeecomputersociety.org/10.1109/ICDE.2004.1319997>.
96. Tao, Yufei, Dimitris Papadias, and **Jimeng Sun**. "The TPR\*-Tree: An Optimized Spatio-Temporal Access Method for Predictive Queries." In VLDB, 790–801, 2003. <http://www.vldb.org/conf/2003/papers/S24P01.pdf>.
97. Tao, Yufei, **Jimeng Sun**, and Dimitris Papadias. "Selectivity Estimation for Predictive Spatio-Temporal Queries." In Proceedings of the 19th International Conference on Data Engineering, March 5-8, 2003, Bangalore, India, 417–28. IEEE Computer Society, 2003. <http://doi.ieeecomputersociety.org/10.1109/ICDE.2003.1260810>.



# Jimeng Sun

## Patents

---

1. Patent number: 7505876, Systems and methods for simultaneous summarization of data cube streams
2. Patent number: 7853545, Preserving privacy of one-dimensional data streams using dynamic correlations
3. Patent number: 8204988, Content-based and time-evolving social network analysis
4. Patent number: 7840516, Preserving privacy of one-dimensional data streams by perturbing data with noise and using dynamic autocorrelation
5. Patent number: US8583586 B2, Mining temporal patterns in longitudinal event data using discrete event matrices and sparse coding

## Advising Students

---

### Georgia Tech

#### PhD students

- Robert Chen, 2014 – present, MD-PhD Emory/GT, Phenotyping from electronic health records and predictive modeling
- Hang Su, 2014 – present, PhD
- Zhaoming Wu, 2014 – present, PhD
- Ioakeim (Joachim) Perros, 2014 – present, PhD

#### Postdocs

- Mohammed Khalilia, 2014 – present

#### Master students

- Apoorva Prabhu Verlekar, 2014 – present, MS in CSE

### IBM Research

- Joyce Ho, University of Texas at Austin, 2013, Unsupervised Phenotyping using Tensor Factorization
- Jiayu Zhou, Arizona State University, 2012, Stable Feature Selection and its applications in healthcare analytics
- Parikshit Sondhi, University of Illinois at Urbana-Champaign, 2011, Text mining on Clinical data
- Dijun Luo, University of Texas at Arlington, 2011, Feature Selection and predictive modeling on clinical data
- Lei Li, Carnegie Mellon University, 2010, Time series analysis on ECG data
- Nan Cao, Hong Kong University of Science & Technology, 2010, Visualization on healthcare data
- U Kang, Carnegie Mellon University, 2010, Graph mining on big data using Hadoop
- Duo Zhang, University of Illinois at Urbana-Champaign, 2009, Text mining on IT support data
- Yu-Ru Lin, Arizona State University, 2008, Social Media Analytics using Tensor analysis
- Christos Boutsidis, Rensselaer Polytechnic Institute, 2008, Column selection algorithm and its application to IT support data