

# Michael Steinbach

Department of Computer Science and Engineering  
University of Minnesota  
4-192 EE/CS Building, 200 Union Street SE  
Minneapolis, MN 55455

telephone: 612-626-7503  
fax: 612-625-0572  
email: steinbac@cs.umn.edu  
webpage: www.cs.umn.edu/~stei0062

## Research Interests

Data mining, machine learning, biomedical informatics, and statistics.

## Education

2005	Ph.D., Computer Science	University of Minnesota, Minneapolis
1992	M.S., Computer Science	University of Minnesota, Minneapolis
1982	M.S., Statistics	University of Minnesota, Minneapolis
1980	B.S., Math	University of Minnesota, Minneapolis

## Professional Experience

2005 -	Research Associate (Researcher 6), University of Minnesota, Minneapolis, MN
2000 - 2005	Research Fellow, University of Minnesota, Minneapolis, MN
1998 - 2000	Research Assistant, University of Minnesota, Minneapolis, MN
1997 - 1998	Senior Software Engineer, Silicon Biology, Minneapolis, MN
1992 - 1997	Manager, Racotek, Minneapolis, MN
1982 - 1992	Senior Principal Systems Engineer/Systems Architect, NCR Comten, St. Paul, MN

## Teaching Experience

1. CSCI 5523: Introduction to Data Mining at UM, helped teach since Spring 2006. Taught solo Fall 2019
2. Helped create and teach a data mining for bioinformatics course at UM for Fall 2006, Spring 2011.
3. Co-taught a data mining tutorial at UM for Bioinformatics Summer Institute.
4. Taught a one-day data mining tutorial to HCBU faculty at NCA&T 2008 and 2009.
5. Co-taught a data mining tutorial at the Army Research Laboratory.

## Professional Activities

### Professional Affiliations

- Member of the ACM
- Member of IEEE

### Program Committees

- The ACM SIGKDD International Conference on Knowledge Discovery and Data Mining
- The IEEE International Conference on Data Mining
- The SIAM International Conference on Data Mining

### Referee

- **Conferences:** ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, SIAM International Conference on Data Mining, IEEE International Conference on Data Mining, Pacific-Asia Conference on Knowledge Discovery and Data Mining, ACM

SIGMOD International Conference on Management of Data, International Conference on Machine Learning, International Conference on Data Engineering,

- **Journals:** Transactions on Knowledge Discovery from Data (TKDD), Transactions on Knowledge and Data Engineering, Transactions on Parallel and Distributed Systems, Journal of American Society of Information Science, Transactions on Geoscience and Remote Sensing, IEEE Transactions on SMCB, VLDB Journal, Data and Knowledge Engineering

### Research Grants/Projects

1. "Collaborative Research: Knowledge Guided Machine Learning: A Framework for Accelerating Scientific Discovery," NSF, 09/01/2019 – 08/31/2021. (co-PI)
2. "BIGDATA: F: Advancing Deep Learning to Monitor Global Change," NSF, 11/01/2018 – 10/31/2021. (co-PI)
3. "Parsing Early Emerging Heterogeneity Related to Autism Spectrum Disorder," NIH, 09/01/2018-08/31/2023. (senior personnel)
4. "MINT: Model Integration through Knowledge Rich Data and Process Composition," DARPA(sub-contract via USC), \$1 Million, January 1, 2018 - December 31, 2022. (senior personnel)
5. "Innovative Methods for Real-Time Risk Modeling of Postoperative Complications," NIH, 04/01/2017 - 03/31/2021. (co-PI)
6. "SCH: EXP: Collaborative Research: Group-Specific Learning to Personalize Evidence-Based Medicine," NSF, \$381,845, 9/1/2016 – 8/31/2020. (PI)
7. "Extracting Typical and Atypical Disease Progression Patterns from Multi-Site EHR," NIH, 09/30/2013 - 06/30/2019. (co-PI)
8. "Understanding Climate Change: A Data-Driven Approach," NSF, \$10,000,000, 09/01/10 - 08/31/18. (senior personnel)
9. "Scalable Analysis of Earth System Data Using Parallelized Graph-Based Approaches," NASA, 10/2015 – 10/2017. (PI)
10. "SCH: EXP: Discovering Patterns to Improve Health to Overcome Health Disparities," NSF, \$497,207, 09/15/13 - 08/31/16. (PI)
11. "Automated Detection of Precursors to Human-Automation Interaction Based Aviation Safety Incidents," NASA, PI: Banerjee, 0/01/2012 - 09/30/2015, (co-PI)
12. "EAGER: Building and analyzing dynamic brain functional networks," NSF, 10/01/2013 – 09/30/2015. (co-PI)
13. "DC: Medium: Collaborative Research: ELLF: Extensible Language and Library Frameworks for Scalable and Efficient Data-Intensive Applications," NSF, 07/01/09 - 06/30/13. (co-PI)

14. III:Small: Generalization of the Association Analysis Framework,” NSF, 08/01/09 - 11/30/12. (co-PI)
15. "NOAA Interdisciplinary Scientific Environmental Technology (ISET) Cooperative Research and Education Center," NOAA, 09/01/2006 - 08/31/2011. Prime: North Carolina A&T State University. (co-PI on Minnesota subaward)
16. "Collaborative Research: Spatio-Temporal Data Mining For Global Scale Eco-Climatic Data," NSF, August 1, 2007 - July 31, 2010. (co-PI)
17. "Collaborative Research: CRI - Scalable Benchmarks, Software and Data for Data Mining, Analytics and Scientific Discoveries," NSF, 3/15/2006 - 2/28/2010. (co-PI)

## Publications

### Complete List of Published Work in Google Scholar and MyBigliography:

[https://scholar.google.com/citations?hl=en&user=Hi\\_xXwoAAAAJ&sortby=pubdate&view\\_op=list\\_works](https://scholar.google.com/citations?hl=en&user=Hi_xXwoAAAAJ&sortby=pubdate&view_op=list_works)

<http://www.ncbi.nlm.nih.gov/sites/myncbi/michael.steinbach.1/bibliography/40758700/public/?sort=date&direction=ascending>

## Books

1. Pang-Ning Tan, Michael Steinbach, and Vipin Kumar, *Introduction to Data Mining 2<sup>nd</sup> edition*, 2019, Addison-Wesley. (1<sup>st</sup> edition Addison Wesley 2005)

## Book Chapters

1. Anuj Karpatne, James H. Faghmous, Jaya Kawale, Luke Styles, Mace Blank, Varun Mithal, Xi C. Chen, Ankush Khandelwal, Shyam Boriah, Karsten Steinhaeuser, Michael Steinbach, Vipin Kumar, Stefan Liess, Earth Science Applications of Sensor Data. Managing and Mining Sensor Data 2013, 505-530
2. S. Boriah, V. Kumar, M. Steinbach, P.-N. Tan, C. Potter, and S. Klooster, “Detecting ecosystem disturbances and land cover change using data mining,” in Next Generation of Data Mining, editors: H. Kargupta, J. Han, P. Yu, R. Motwani, and V. Kumar, CRC Press, 2009.
3. Michael Steinbach and Pang-Ning Tan, “kNN: k-Nearest Neighbors,” The Top Ten Algorithms in Data Mining, Chapman & Hall/CRC Data Mining and Knowledge Discovery Series, 2009.
4. Vipin Kumar, Pang-Ning Tan, and Michael Steinbach, “Data Mining,” in Handbook of Data Structures and Applications, CRC Press, 2005.
5. Pusheng Zhang, Michael Steinbach, Vipin Kumar, Shashi Shekhar, Pang-Ning Tan, Steve Klooster, and Chris Potter, “Discovery of Patterns of Earth Science Data Using Data Mining,” in Next Generation of Data Mining Applications, Jozef Zurada and Medo Kantardzic(eds), IEEE Press, 2005.
6. Michael Steinbach, Levent Ertoz, and Vipin Kumar, “Challenges of Clustering High Dimensional Data,” in New Vistas in Statistical Physics – Applications in Econophysics, Bioinformatics, and Pattern Recognition, Springer-Verlag, 2004.

7. Levent Ertoz, Michael Steinbach, and Vipin Kumar, "Finding Topics in Collections of Documents: A Shared Nearest Neighbor Approach," in *Clustering and Information Retrieval*, Kluwer Academic Publishers, 2004.

### Journal Papers

1. G. J. Simon, K. A. Peterson, M. R. Castro, M. S. Steinbach, V. Kumar, and P. J. Caraballo, "Predicting diabetes clinical outcomes using longitudinal risk factor trajectories," *BMC Medical Informatics and Decision Making*, vol. 20, p. 1–9, 2020.
2. J. S. Read, X. Jia, J. Willard, A. P. Appling, J. A. Zwart, S. K. Oliver, A. Karpatne, G. J. A. Hansen, P. C. Hanson, W. Watkins, and others, "Process-guided deep learning predictions of lake water temperature," *Water Resources Research*, 2019.
3. S. Jacob, J. J. Wolff, M. S. Steinbach, C. B. Doyle, V. Kumar, and J. T. Ellison, "Neurodevelopmental heterogeneity and computational approaches for understanding autism," *Translational Psychiatry*, vol. 9, p. 1–12, 2019.
4. W. Oh, M. S. Steinbach, M. R. Castro, K. A. Peterson, V. Kumar, P. J. Caraballo, and G. J. Simona, "Evaluating the Impact of Data Representation on EHR-Based Analytic Tasks.," in *17th World Congress on Medical and Health Informatics, MEDINFO 2019*, 2019.
5. G. Fang, W. Wang, V. Paunic, H. Heydari, M. Costanzo, X. Liu, X. Liu, B. VanderSluis, B. Oatley, M. Steinbach and others, "Discovering genetic interactions bridging pathways in genome-wide association studies," *Nature Communications*, vol. 10, p. 1–18, 2019.
6. S. G. Johnson, L. Pruinelli, A. Hoff, V. Kumar, G. J. Simon, M. Steinbach, and B. L. Westra, "A Framework for Visualizing Data Quality for Predictive Models and Clinical Quality Measures," *AMIA Summits on Translational Science Proceedings*, vol. 2019, p. 630, 2019.
7. P. Yadav, M. Steinbach, V. Kumar, and G. Simon, "Mining Electronic Health Records (EHRs) A Survey," *ACM Computing Surveys (CSUR)*, vol. 50, p. 1–40, 2018.
8. L. Pruinelli, B. L. Westra, P. Yadav, A. Hoff, M. Steinbach, V. Kumar, C. W. Delaney, and G. Simon, "Delay within the 3-hour surviving sepsis campaign guideline on mortality for patients with severe sepsis and septic shock," *Critical care medicine*, vol. 46, p. 500, 2018.
9. A. Karpatne, G. Atluri, J. H. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, S. Shekhar, N. Samatova, and V. Kumar, "Theory-guided data science: A new paradigm for scientific discovery from data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, p. 2318–2331, 2017.
10. B. L. Westra, S. Landman, P. Yadav, and M. Steinbach, "Secondary analysis of an electronic surveillance system combined with multi-focal interventions for early detection of sepsis," *Applied clinical informatics*, vol. 26, p. 47–66, 2017.
11. W. Oh, E. Kim, M. R. Castro, P. J. Caraballo, V. Kumar, M. S. Steinbach and G. J. Simon, "Type 2 diabetes mellitus trajectories and associated risks," *Big data*, vol. 4, p. 25–30, 2016.
12. M. Bhargava, K. J. Viken, S. Dey, M. S. Steinbach, B. Wu, P. D. Jagtap, L. Higgins, A. Panoskaltis-Mortari, D. J. Weisdorf, V. Kumar and others, "Proteome profiling in lung injury after hematopoietic stem cell transplantation," *Biology of Blood and Marrow Transplantation*, vol. 22, p. 1383–1390, 2016.
13. S. Dey, R. Gupta, M. Steinbach, V. Kumar, C. K. Reddy, and C. C. Aggarwal, "Predictive Models for Integrating Clinical and Genomic Data.," *Healthcare Data Analytics*, vol. 36, p. 433, 2015.
14. S. Roark, B. Sandri, S. Dey, M. Steinbach, T. Becker, and C. H. Wendt, "Longitudinal protein expression patterns in bronchiolitis obliterans syndrome," *Pulmonol Respir Res*, vol. 3, p. 1–5, 2015.
15. G. Atluri, M. Steinbach, K. O. Lim, V. Kumar, and A. MacDonald III, "Connectivity cluster analysis for discovering discriminative subnetworks in schizophrenia," *Human brain mapping*, vol. 36, p. 756–767, 2015.

16. M. Bhargava, K. Viken, S. Dey, M. Steinbach, B. Wu, P. Jagtap, L. Higgins, A. Panoskaltis-Mortari, D. J. Weisdorf, V. Kumar and others, "B24 Mechanisms of Injury in Sepsis: Inflammation, Oxidation, and Endothelial Damage: Proteome Profiling In Lung Injury Following Hematopoietic Stem Cell Transplantation," *American Journal of Respiratory and Critical Care Medicine*, vol. 191, p. 1, 2015.
17. J. H. Faghmous, A. Banerjee, S. Shekhar, M. Steinbach, V. Kumar, A. R. Ganguly and N. Samatova, "Theory-guided data science for climate change," *Computer*, vol. 47, p. 74–78, 2014.
18. Landman, Sean R., Tae Hyun Hwang, Kevin AT Silverstein, Yingming Li, Scott M. Dehm, Michael Steinbach, and Vipin Kumar. "SHEAR: sample heterogeneity estimation and assembly by reference." *BMC Genomics* 15, no. 1 (2014): 84.
19. M. Bhargava, T. L. Becker, K. J. Viken, P. D. Jagtap, S. Dey, M. S. Steinbach, B. Wu, V. Kumar, P. B. Bitterman, D. H. Ingbar and others, "Proteomic profiles in acute respiratory distress syndrome differentiates survivors from non-survivors," *PloS one*, vol. 9, 2014.
20. Been, Raha A., Michael A. Linden, Courtney J. Hager, Krista J. DeCoursin, Juan E. Abrahante, Sean R. Landman, Michael Steinbach, Aaron L. Sarver, David A. Largaespada, and Timothy K. Starr. "Genetic Signature of Histiocytic Sarcoma Revealed by a Sleeping Beauty Transposon Genetic Screen in Mice." *PloS one* 9, no. 5, (2014).
21. Bhargava, Maneesh, Sanjoy Dey, Trisha Becker, Michael Steinbach, Baolin Wu, Sang Mee Lee, LeeAnn Higgins, et al. "Protein expression profile of rat type two alveolar epithelial cells during hyperoxic stress and recovery." *American Journal of Physiology-Lung Cellular and Molecular Physiology* 305, no. 9 (2013): L604-L614.
22. Atluri, Gowtham, Kanchana Padmanabhan, Gang Fang, Michael Steinbach, Jeffrey R. Petrella, Kelvin Lim, Angus MacDonald III, Nagiza F. Samatova, P. Murali Doraiswamy, and Vipin Kumar. "Complex biomarker discovery in neuroimaging data: Finding a needle in a haystack." *NeuroImage: Clinical* 3 (2013): 123-131.
23. Menary, Kyle, Paul F. Collins, James N. Porter, Ryan Muetzel, Elizabeth A. Olson, Vipin Kumar, Michael Steinbach, Kelvin O. Lim, and Monica Luciana. "Associations between cortical thickness and general intelligence in children, adolescents, and young adults." *Intelligence* 41, no. 5 (2013): 597-606.
24. Jaya Kawale, Stefan Liess, Arjun Kumar, Michael Steinbach, Peter K. Snyder, Vipin Kumar, Auroop R. Ganguly, Nagiza F. Samatova, Fredrick H. M. Semazzi: A graph-based approach to find teleconnections in climate data. *Statistical Analysis and Data Mining* 6(3): 158-179 (2013)
25. Michael Steinbach, Haoyu Yu, and Vipin Kumar, Identification of Co-occurring Insertions in Cancer Genomes Using Association Analysis, *International Journal of Data Mining and Bioinformatics special issue for 2nd International Workshop on Data Mining for Biomarker Discovery (DMBD 2012)*.
26. Tracy L. Bergemann, Timothy K. Starr, Haoyu Yu, Michael Steinbach, Jesse Erdmann, Yun Chen, Robert T. Cormier, David A. Largaespada, and Kevin A. T. Silverstein, New methods for finding common insertion sites and co-occurring common insertion sites in transposon- and virus-based genetic screens, *Nucleic Acids Res.* 2012 May; 40(9): 3822–3833.
27. Gang Fang, Majda Haznadar, Wen Wang, Haoyu Yu, Michael Steinbach, Tim Church, William Oetting, Brian Van Ness, and Vipin Kumar, High-order SNP Combinations Associated with Complex Diseases: Efficient Discovery, Statistical Power and Functional Interactions, *PLoS ONE*, 7(4): e33531. doi:10.1371/journal.pone.0033531, 2012.
28. Gang Fang, Gaurav Pandey, Wen Wang, Manish Gupta, Michael Steinbach, Vipin Kumar: Mining Low-Support Discriminative Patterns from Dense and High-Dimensional Data. *IEEE Trans. Knowl. Data Eng.* 24(2): 279-294 (2012)
29. Varun Mithal, Ashish Garg, Shyam Boriah, Michael Steinbach, Vipin Kumar, Christopher Potter, Steven A. Klooster, Juan Carlos Castilla-Rubio: Monitoring global forest cover using data mining. *ACM TIST* 2(4): 36 (2011)

30. Hui Xiong, Michael Steinbach, Arifin Ruslim, Vipin Kumar: Characterizing pattern preserving clustering. *Knowl. Inf. Syst.* 19(3): 311-336, 2009.
31. Brian Van Ness, Christine Ramos, Majda Haznadar, Antje Hoering, Jeff Haessler, John Crowley, Susanna Jacobus, Martin Oken, Vincent Rajkumar, Philip Greipp, Bart Barlogie, Brian Durie, Michael Katz, Gowtham Atluri, Gang Fang, Rohit Gupta, Michael Steinbach, Vipin Kumar, Richard Mushlin, David Johnson, Gareth Morgan, "Genomic Variation in Myeloma: Design, content, and initial application of the Bank On A Cure SNP Panel to the Analysis of survival," *BMC Medicine*, 2008, doi:10.1186/1741-7015-6-26
32. Xindong Wu, Vipin Kumar, J. Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan, Angus F. M. Ng, Bing Liu, Philip S. Yu, Zhi-Hua Zhou, Michael Steinbach, David J. Hand, Dan Steinberg: Top 10 algorithms in data mining. *Knowl. Inf. Syst.* 14(1): 1-37, 2008.
33. Christopher Potter, Shyam Boriah, Michael Steinbach, Vipin Kumar and Steven Klooster Terrestrial vegetation dynamics and global climate controls, *Climate Dynamics* 2008.
34. C. Potter, S. Boriah, M. Steinbach, V. Kumar, and S. Klooster, Terrestrial Vegetation Dynamics and Global Climate Controls in North America: 2001--2005, *Earth Interactions*, 2008.
35. Potter, C., V. Genovese, P. Gross, S. Boriah, M. Steinbach, and V. Kumar, Revealing Land Cover Change in California With Satellite Data, *Eos Trans. AGU*, 88(26), 269, 2007.
36. Hui Xiong, Michael Steinbach, Vipin Kumar, Privacy Leakage in Multi-relational Databases: A Semi-supervised Learning Perspective, *VLDB Journal Special Issue on Privacy Preserving Data Management*, Vol. 15, No. 4, pp. 388-402, November 2006.
37. Michael Steinbach and Vipin Kumar, Generalizing the Notion of Confidence, *Knowledge and Information Systems (KAIS)*, Vol 12(3), August 2007.
38. Hui Xiong, Gaurav Pandey, Michael Steinbach, Vipin Kumar, Enhancing Data Analysis with Noise Removal, *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 18(3), pp. 304-319, March 2006.
39. C. Potter, S. Klooster, M. Steinbach, P. Tan, V. Kumar, S. Shekhar, and C. Carvalho, Understanding Global Teleconnections of Climate to Regional Model Estimates of Amazon Ecosystem Carbon Fluxes, *Global Change Biology*, 2004.
40. C. Potter, P.-N.Tan, M. Steinbach, S. Klooster, V. Kumar, R. Myneni, and V. Genovese, Major Disturbance Events in Terrestrial Ecosystems Detected using Global Satellite Data Sets, *Global Change Biology*, 2003.
41. C. Potter, S. Klooster, M. Steinbach, P. Tan, V. Kumar, S. Shekhar, R. Nemani, and R. Myneni, Global Teleconnections of Ocean Climate to Terrestrial Carbon Flux, *J. of Geophysical Research*, Vol. 108, No. D17, 4556, 2003.
42. C. Potter, S. Klooster, M. Steinbach, P. Tan, V. Kumar, R. Myneni, V. Genovese, Variability in Terrestrial Carbon Sinks Over Two Decades: Part 1-North America, *Earth Interactions*, 2003.

### **Conference and Workshop Papers**

1. X. Jia, J. Willard, A. Karpatne, J. Read, J. Zwart, M. Steinbach, and V. Kumar, "Physics guided RNNs for modeling dynamical systems: A case study in simulating lake temperature profiles," in *Proceedings of the 2019 SIAM International Conference on Data Mining*, 2019.
2. P. Yadav, M. Steinbach, M. R. Castro, P. J. Caraballo, V. Kumar, and G. Simon, "Frequent Causal Pattern Mining: A Computationally Efficient Framework For Estimating Bias-Corrected Effects," in *2019 IEEE International Conference on Big Data (Big Data)*, 2019.
3. J. Li, M. Wang, M. S. Steinbach, V. Kumar, and G. J. Simon, "Don't Do Imputation: Dealing with Informative Missing Values in EHR Data Analysis," in *2018 IEEE International Conference on Big Knowledge (ICBK)*, 2018.

4. W. Oh, P. Yadav, V. Kumar, P. J. Caraballo, M. R. Castro, M. S. Steinbach, and G. J. Simon, "Estimating Disease Onset Time by Modeling Lab Result Trajectories via Bayes Networks," in 2017 IEEE International Conference on Healthcare Informatics (ICHI), 2017.
5. S. Agrawal, K. Das, S. Liess, G. Atluri, M. Steinbach, K. Steinhaeuser, and V. Kumar, "Anomaly Detection in Time-Evolving Climate Graphs," in 2016 AGU Fall Meeting, 2016.
6. S. Liess, S. Agrawal, K. Das, G. Atluri, M. Steinbach, K. Steinhaeuser, and V. Kumar, "Anomaly Detection in Time-Evolving Climate Graphs," in AGU Fall Meeting Abstracts, 2016.
7. P. Yadav, M. Steinbach, L. Pruinelli, B. Westra, C. Delaney, V. Kumar, and G. Simon, "Forensic style analysis with survival trajectories," in 2015 IEEE International Conference on Data Mining, 2015.
8. S. Dey, G. Simon, B. Westra, M. Steinbach, and V. Kumar, "Mining interpretable and predictive diagnosis codes from multi-source electronic health records," in Proceedings of the 2014 SIAM International Conference on Data Mining, 2014.
9. Vanja Paunic, Michael Steinbach, Abeer Madbouly, Vipin Kumar, Evaluation of Label Dependency for the Prediction of HLA Genes. BCB 2013
10. Igor Melnyk, Pranjul Yadav, Michael Steinbach, Jaideep Srivastava, Vipin Kumar, Arindam Banerjee, Detection of Precursors to Aviation Safety Incidents Due to Human Factors, ICDM Workshops, 407-412, 2013.
11. Sanjoy Dey, Kelvin O. Lim, Gowtham Atluri, Angus W. MacDonald III, Michael Steinbach, Vipin Kumar: A pattern mining based integrative framework for biomarker discovery. BCB 2012: 498-505.
12. Anuj Karpatne, Mace Blank, Michael Lau, Shyam Boriah, Karsten Steinhaeuser, Michael Steinbach, Vipin Kumar: Importance of vegetation type in forest cover estimation. CIDU 2012: 71-78
13. Xi Chen, Anuj Karpatne, Yashu Chamber, Varun Mithal, Michael Lau, Karsten Steinhaeuser, Shyam Boriah, Michael Steinbach, Vipin Kumar, Christopher Potter, Steven A. Klooster, Teji Abraham, J. D. Stanley, Juan Carlos Castilla-Rubio: A new data mining framework for forest fire mapping. CIDU 2012: 104-111
14. Gowtham Atluri, Sanjoy Dey, Gang Fang, Sean Landman, Vanja Paunic, Wen Wang, Michael Steinbach, Vipin Kumar: Invited: Discovering combinatorial biomarkers. ICCABS 2012: 1-2
15. Vanja Paunic, Michael Steinbach, Martin Maiers, Vipin Kumar, Prediction of HLA Genes from SNP Data and HLA Haplotype Frequencies, Best Paper in the Ph.D. Student Forum, IEEE International Conference on Data Mining, 2012.
16. Jaya Kawale, Stefan Liess, Arjun Kumar, Michael Steinbach, Auroop R. Ganguly, Nagiza F. Samatova, Fredrick H. M. Semazzi, Peter K. Snyder, Vipin Kumar: Data Guided Discovery of Dynamic Climate Dipoles. CIDU 2011: 30-44 **(Best Student Paper)**
17. Bonnie Westra, Sanjoy Dey, Gang Fang, Michael Steinbach, Kay Savik, Cristina Oancea, and Vipin Kumar, Interpretable Predictive Models for Knowledge Discovery from Home Care Electronic Health Records, Journal of Healthcare Engineering, pp. 55-74, Volume 2, Number 1 / March 2011, <http://multi-science.metapress.com/content/c45r7407w3262p20/>
18. Varun Mithal, Ashish Garg, Ivan Brugere, Shyam Boriah, Vipin Kumar, Michael Steinbach, Christopher Potter, Steven A. Klooster: Incorporating Natural Variation into Time Series-Based Land Cover Change Detection. CIDU 2011: 45-59
19. Ashish Garg, Lydia Manikonda, Shashank Kumar, Vikrant Krishna, Shyam Boriah, Michael Steinbach, Durga Toshnival, Vipin Kumar, Christopher Potter, Steven A. Klooster: A Model-Free Time Series Segmentation Approach for Land Cover Change Detection. CIDU 2011: 144-158
20. Jaya Kawale, Snigdhasu Chatterjee, Arjun Kumar, Stefan Liess, Michael Steinbach, Vipin Kumar: Anomaly Construction in Climate Data: Issues and Challenges. CIDU 2011: 189-203

21. Yashu Chamber, Ashish Garg, Varun Mithal, Ivan Brugere, Vipin Kumar, Michael Lau, Michael Steinbach, Christopher Potter, Steven A. Klooster, Vikrant Krishna, Shyam Boriah: A Novel Time Series Based Approach to Detect Gradual Vegetation Changes in Forests. CIDU 2011: 248-262
22. Ashish Garg, Varun Mithal, Yashu Chamber, Ivan Brugere, Vijay Chaudhari, Marc Dunham, Vikrant Krishna, Sairam Krishnamurthy, Sruthi Vangala, Shyam Boriah, Michael Steinbach, Vipin Kumar, Albert Cho, J. D. Stanley, Teji Abraham, Juan Carlos Castilla-Rubio, Christopher Potter, Steven A. Klooster: Gopher: Global observation of Planetary Health and Ecosystem Resources. IGARSS 2011: 1449-1452
23. Michael Steinbach, Haoyu Yu, Gang Fang, Vipin Kumar: Using Constraints to Generate and Explore Higher Order Discriminative Patterns. PAKDD (1) 2011: 338-350
24. Jaya Kawale, Michael Steinbach, Vipin Kumar: Discovering Dynamic Dipoles in Climate Data. SDM 2011: 107-118
25. Steinbach, M., Yu, H., and Kumar, V., 2010, December. Identification of co-occurring insertions in cancer genomes using association analysis. In 2010 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW) (pp. 494-499). IEEE.
26. Shyam Boriah, Varun Mithal, Ashish Garg, Vipin Kumar, Michael Steinbach, Christopher Potter, Steven A. Klooster: A Comparative Study Of Algorithms For Land Cover Change. CIDU 2010: 175-188
27. Caitlin Race, Michael Steinbach, Auroop R. Ganguly, Fredrick H. M. Semazzi, Vipin Kumar: A Knowledge Discovery Strategy for Relating Sea Surface Temperatures to Frequencies of Tropical Storms and Generating Predictions of Hurricanes Under 21st-century Global Warming Scenarios. CIDU 2010: 204-212
28. Shyam Boriah, Varun Mithal, Ashish Garg, Michael Steinbach, Vipin Kumar, Christopher Potter, Steven A. Klooster, Juan Carlos Castilla-Rubio: Automated detection of forest cover changes. IGARSS 2010: 44-47
29. Gang Fang, Rui Kuang, Gaurav Pandey, Michael Steinbach, Chad L. Myers, Vipin Kumar: Subspace Differential Coexpression Analysis: Problem Definition and a General Approach. Pacific Symposium on Biocomputing 2010: 145-156
30. Hui Xiong, Michael Steinbach, Arifin Ruslim, Vipin Kumar: Characterizing pattern preserving clustering. Knowl. Inf. Syst. 19(3): 311-336 (2009)
31. Gowtham Atluri, Rohit Gupta, Gang Fang, Gaurav Pandey, Michael Steinbach, Vipin Kumar: Association Analysis Techniques for Bioinformatics Problems. BICoB 2009: 1-13
32. Gaurav Pandey, Gowtham Atluri, Gang Fang, Rohit Gupta, Michael Steinbach, Vipin Kumar: Association analysis techniques for analyzing complex biological data sets. GENSiPS 2009: 1-4
33. Gaurav Pandey, Gowtham Atluri, Gang Fang, Rohit Gupta, Michael Steinbach, and Vipin Kumar, Association Analysis Techniques For Analyzing Complex Biological Data Sets, Proceedings of the IEEE International Workshop on Genomic Signal Processing and Statistics (GENSIPS), May 2009.
34. Gaurav Pandey, Gowtham Atluri, Michael Steinbach, Chad L. Myers, Vipin Kumar: An association analysis approach to biclustering. KDD 2009: 677-686
35. Gaurav Pandey, Lakshmi N. Ramakrishnan, Michael Steinbach, and Vipin Kumar, Systematic Evaluation of Scaling Methods for Gene Expression Data, Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM) November 2008.
36. Rohit Gupta, Gang Fang, Blayne Field, Michael Steinbach and Vipin Kumar, Quantitative Evaluation of Approximate Frequent Pattern Mining Algorithms, Proceedings of ACM SIGKDD 2008, August 2008.
37. S. Boriah, V. Kumar, M. Steinbach, C. Potter, and S. Klooster, Land cover change detection: A case study, Proceedings of ACM SIGKDD 2008, August 2008, pp 857-865.



38. M. Steinbach, P.-N. Tan, H. Xiong, and V. Kumar, Objective Measures for Association Pattern Analysis, Proceedings of the Joint Summer Research Conference on Machine and Statistical Learning: Prediction and Discovery, AMS Contemporary Mathematics (CONM/443) Prediction and Discovery, AMS, 2007.
39. R. Gupta, T. Garg, G. Pandey, M. Steinbach, and V. Kumar, Comparative Study of Various Genomic Data Sets for Protein Function, Data Mining for Biomedical Informatics Workshop, Seventh SIAM International Conference on Data Mining, April 28, 2007.
40. M. Steinbach and V. Kumar, Generalizing the Notion of Confidence, in Proc. of the Fifth IEEE International Conference on Data Mining (ICDM 2005), Houston, TX, November 27-30, 2005.
41. Hui Xiong, Michael Steinbach, and Vipin Kumar, Privacy Leakage in Multi-relational Databases via Pattern-based Semi-supervised Learning, in Proceedings of the ACM International Conference on Information and Knowledge Management (ACM CIKM 2005), pp. 355-356, 2005.
42. M. Steinbach, P.-N. Tan, H. Xiong, and V. Kumar. Generalizing the Notion of Support, in Proc. of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2004), pp. 689-694, Seattle, WA, August 22-25, 2004.
43. M. Steinbach, P.-N. Tan, and V. Kumar, Support Envelopes: A Technique for Exploring the Structure of Association Patterns, in Proc. of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2004), pp. 296-305, Seattle, WA, August 22-25, 2004.
44. M. Steinbach, P.-N. Tan, V. Kumar, S. Klooster, C. Potter, Discovery of climate indices using clustering, in Proc. of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2003), pp. 446-455, Washington, D.C., August 24-27, 2003.
45. H. Xiong, M. Steinbach, P.-N. Tan, and V. Kumar, HICAP: Hierarchical Clustering with Pattern Preservation, in Proc. of the 2004 SIAM International Conf. on Data Mining (SDM 2004), pp. 279 - 290, Lake Buena Vista, Florida, 2004
46. L. Ertöz, M. Steinbach, and V. Kumar, Finding Clusters of Different Sizes, Shapes, and Densities in Noisy, High Dimensional Data, in Proc. of the 2003 SIAM International Conf. on Data Mining (SDM 2003), San Francisco, CA, 2003.
47. Vipin Kumar, Mahesh V. Joshi, Eui-Hong (Sam) Han, Pang-Ning Tan, and Michael Steinbach, High Performance Data Mining, High Performance Computing for Computational Science, VECPAR 2002, Palma, J. M.L.M., Dongarra, J., Hernandez, V., and Sousa, A. A. (Eds.) 5th International Conference, Porto, Portugal, June 26-28, 2002.
48. M. Steinbach, P.-N. Tan, V. Kumar, C. Potter, and S. Klooster, Temporal Data Mining for the Discovery and Analysis of Ocean Climate Indices, KDD Workshop on Temporal Data Mining, 2002.
49. M. Steinbach, P.-N. Tan, V. Kumar, C. Potter, and S. Klooster, Data Mining for the Discovery of Ocean Climate Indices, Fifth Workshop on Scientific Data Mining, 2nd SIAM International Conference on Data Mining, 2002.
50. V. Kumar, M. Steinbach, P.-N. Tan, S. Klooster, C. Potter, and A. Torregrosa, Mining Scientific Data: Discovery of Patterns in the Global Climate System, Joint Statistical Meeting, 2001.
51. M. Steinbach, P.-N. Tan, V. Kumar, C. Potter, S. Klooster, and A. Torregrosa, Clustering Earth Science Data: Goals, Issues and Results, KDD Workshop on Mining Scientific Datasets, 2001.
52. P.-N. Tan, M. Steinbach, V. Kumar, C. Potter, S. Klooster, and A. Torregrosa, Finding Spatio-Temporal Patterns in Earth Science Data, KDD Workshop on Temporal Data Mining, 2001.
53. M. Steinbach, G. Karypis, and V. Kumar, Efficient Algorithms for Creating Product Catalogs, Web Mining Workshop, First SIAM International Conference on Data Mining, Chicago, IL, 2001.

54. L. Ertöz, M. Steinbach, and V. Kumar, Finding Topics in Collections of Documents: A Shared Nearest Neighbor Approach, Text Mine'01, Workshop on Text Mining, 1st SIAM International Conference on Data Mining, Chicago, IL, April 2001.
55. Michael Steinbach, George Karypis, and Vipin Kumar, A Comparison of Document Clustering Techniques, Text Mining Workshop, in Proc. of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2000), Boston, MA, August 20-23, 2000.

**Software**

MATLAB code for finding and visualizing support envelopes, which capture the overall structure of association patterns, is available at <http://www-users.cs.umn.edu/~stei0062/se/se.php>