

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/~steinbac

Research Interests

Data mining, medical- and bio-informatics, pattern recognition 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

2000 -	Research Associate, 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. Helped create and teach a data mining for bioinformatics course at UM for Fall 2006.
2. Co-taught a data mining tutorial at UM for Bioinformatics Summer Institute.
3. Co-taught CSCI 5523: Introduction to Data Mining at UM in Spring 2006.
4. Helped create and teach a data mining course at UM.
5. Co-taught a data mining tutorial at the Army Research Laboratory.
6. Statistics teaching assistant at UM. This included teaching recitations of introductory statistics classes.

Publications

Books

1. Gaurav Pandey, Vipin Kumar, and Michael Steinbach, *Computational Approaches for Protein Function Prediction*, to be published in the Wiley Book Series on Bioinformatics, Fall 2007.
2. Pang-Ning Tan, Michael Steinbach, and Vipin Kumar, *Introduction to Data Mining*, May 2005, Addison-Wesley. (ISBN:0-321-32136-7)

Book Chapters

1. Vipin Kumar, Pang-Ning Tan, and Michael Steinbach, Data Mining, in Handbook of Data Structures and Applications, CRC Press, 2005.
2. 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.

3. 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.
4. 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. X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z.-H. Zhou, D. J. Hand, D. Steinberg, and M. Steinbach, Top 10 Algorithms in Data Mining, *IEEE Transactions on Systems, Man and Cybernetics, Part B* (to appear).
2. 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.
3. M. Steinbach and V. Kumar, Generalizing the Notion of Confidence, *Knowledge and Information Systems*, published online, October 3, 2006.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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 Papers

1. 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) book series (to appear).
2. 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.
3. 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.
4. 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.
5. 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.

6. 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.
7. 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
8. 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.
9. 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.

Workshop Papers

1. 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 (SDM 2007), April 28, 2007 (to appear).
2. Prediction and Enhancements Using Association Analysis 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.
3. 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.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.
9. 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/~steinbac/se/se.php>

Professional Activities

Professional Affiliations

- Member of the ACM
- Member of IEEE

Program Committees

- The 10th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 2006), Berlin, Germany, September 18– 22, 2006
- The 5th IEEE International Conference on Data Mining (ICDM 2005), Houston, TX, November 27– 30, 2005.
- The 6th SIAM International Conference on Data Mining (SDM 2006), Bethesda, MD, April 20-22, 2006.

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, Supercomputing
- **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