

Sanjay Chawla
Associate Professor and Associate Head
School of Information Technologies
University of Sydney
Sydney, NSW, Australia 2006

Education

Phd, 1995, Department of Mathematics, University of Tennessee, Knoxville
MA, 1990, Department of Mathematics, Central Michigan University
BA(Hons), 1987, Mathematics, Delhi University

Employment

Associate Professor (and Associate Head), January 2006-present
School of Information Technologies, University of Sydney

Senior Lecturer, August 2002-December 2005
School of Information Technologies, University of Sydney

Senior Technical Instructor, August 2000-July 2002
Vignette Corporation, Boston

Postdoctoral Associate, August 1997-July 2000
Department of Computer Science, University of Minnesota

Industrial Postdoctoral Fellow, September 1995-August 1997
Institute for Mathematics and its Applications, University of Minnesota

Start Ups

Chief Scientist, January 2004 - October 2006
DTecht Pvt. Ltd.
(Specializes in Data Mining Solutions for Fraud Detection in Health Insurance Transactions)

Consulting

DTecht Pvt. Ltd, January 2004 - present

GE Consumer Finance, April 2004 - September 2004

McKinsey Consulting, August -September 2003

Vignette Corporation, December 2002

Awards and Honours

Best Application Paper Award

Outlier Detection in Sequential Databases

SIAM International Conference on Data Mining, April 2006

Snowy Hydro Prize for *Best Commercialization Research Potential Award*

Electrical and Information Engineering Foundation, University of Sydney, 2004

Vignette All-Star Award for Q2 2002

Vignette Corporation, Waltham, MA.

Competitive Grants

1. Chawla, Sanjay. Cross Market Pattern Mining for Insider Trading. *Capital Markets Commonwealth Research Center (CMCRC)*, A\$55,000; Jan-June 2007
2. Agastya, Murali and Chawla, Sanjay and De'Roos Nick. *University of Sydney, Australian Research Council (ARC) Near-Miss Scheme*, A\$20,000; Jan-Dec 2007
3. Chawla, Sanjay. New Directions in Mining Complex Spatial Relationships in Scientific Databases. *Australian Research Council Discovery Grant*, A\$194,000; 2005-2007. [Australian Research Council is equivalent to NSF]
4. Chawla, Sanjay. DTecht: Health Insurance Data Mining, *Capital Markets Commonwealth Research Centre (CRC)*, A\$ 439,000; 2004-2005 [CRCs have been set up by the Federal government to commercialize university research through spin-offs. This grant led to the formation of DTecht Pvt. Ltd. for which I served as the Chief Scientific Officer till October 2006]
5. Chawla, Sanjay. Outlier Patterns in Health Insurance Data Mining. *Capital Markets Commonwealth Research Center (CMCRC) seed grant* A\$15,000. 2003.
6. Zomaya, Albert and Everitt, David and Chawla, Sanjay and Landfeldt, Bjorn and Quigley Aaron. Storage Area Network *University of Sydney SESQUI Infrastructure Grant*, A\$100,000; 2004
7. Chawla, Sanjay. Extraction and Management of Patterns from Spatially Indexed Time Series. *University of Sydney New Staff Grant*, A\$15,000. 2003

Teaching

I have more than ten years of teaching experience. The first full course I taught was *College Remedial Algebra* at Central Michigan University in the Spring of 1989.

University of Sydney

Period	Course Title	Comments
Second Semester (2006)	Advanced Data Models	An advanced graduate level course which covers postrelational database technology: Spatial, Temporal, XML, Data Mining, P2P, Security
	Database Systems 2	The second undergraduate course in database systems covering transactions, indexing, query optimization
	Theory of Computation	Team Teaching: I am teaching the <i>Approximation Algorithms</i> part
First Semester (2006)	No Teaching	All my teaching was pushed into the second semester
Second Semester (2005)	Advanced Data Models	
	Database Systems 1	First undergraduate course in Databases
First Semester (2005)	Advanced Data Models	
	Data Mining	Team Teaching: Graduate level course in Data Mining covers Association Rule Mining, Classification, Outlier Detection and Clustering
2002-2004	Database and Data Mining Courses	

Research Students Supervised

In Australia, besides Phd and Masters students, a big emphasis is placed on the supervision of fourth year Honours thesis. Good Honours thesis are roughly equivalent to US Masters thesis in terms of original research contribution.

Completed Supervision

Student	Type	Status	Topic
Pei Sun	Phd	Submitted	Outlier Detection for High-Dimensional, spatial and Sequential Data
Lior Lukov	MSc	Completed	Unravelling the architecture of membrane proteins with Conditional Random Fields
Aditya Menon	UG Honour	2006	Sparse Random Projections (University Medal)
Tara McIntosh	UG Honours	2005	Extracting Gene Networks-Rules and Pattern based Approach
Florian Verhein	UG Honours	2004	Spatio-Temporal Data Mining (University Medal)
Robert Munro	UG Honours	2003	Natural Language Computation (University Medal)

Currently Supervising

Student	Type	Status	Topic
Bavani Arunasalam	Phd	Will Submit in 2007	Support-Free Techniques In Data Mining
Florian Verhein	Phd	Submit in late 2007	Tentative: Advanced Association Rule Mining Techniques
Ghazi Al-Naymat	Phd		Tentative:Time Series Data Mining
Fadi Dabhar	Honours	2006	Kernel Methods for Outlier Detection for Protein class identification

Administration

1. *Associate Head of School(HOS), January 2006- present*
(In Sydney University, a large department is called a School)
 - (a) Serve as Acting Head of School during HOS's absence.
 - (b) Incharge of assigning teaching and administration duties to faculty members (approx 28 full-time academic staff).
 - (c) Manage, update and revise the School's workload formula to ensure fairness in assignment of teaching and administration duties to academic staff.
 - (d) Chair of the School's Teaching and Learning Committee.
 - (e) Responsible for revamping the curriculum for 2008.
2. *Postgraduate Research Coordinator, January 2004- December 2005*
3. *Seminar Coordinator, January 2003 - December 2004.*

Professional Service

1. Workshops Chair for 2008 PAKDD in Japan.
2. Asia-Pacific Publicity Chair for 2007 SIAM International Conference on Data Mining.
3. Program Committee for IEEE International Conference on Data Mining (ICDM):2005;2006
4. Program Committee for Pacific-Asia Advances in Knowledge Discovery and Data Mining (PAKDD):2004;2005;2006
5. Program Committee for SIAM International Conference on Data Mining (SDM):2004;2005;2006

Research Publications

Books

1. Shekhar, Shashi and Chawla, Sanjay. **Spatial Databases: A Tour**. *Prentice Hall*, 2002. (note: A Chinese and Russian translation appeared in 2004)

In Communication

1. Verhein, Florian and Chawla, Sanjay. Mining Spatio-Temporal Patterns in Object Mobility Databases *Data Mining and Knowledge Discovery (DMKD)*[**Accepted, subject to minor revisions**]
2. Menon, Aditya and Pham, Anh and Chawla, Sanjay and Viglas, Tasos. An Incremental Data-Stream Sketch Using Sparse Random Projection[**Submitted**]

Journals

1. McIntosh, Tara and Chawla, Sanjay. High Confidence Rule Mining for Microarray Analysis. *ACM/IEEE Transactions on Computational Biology and Bioinformatics*[**Accepted**]
2. Chawla, Sanjay and Sun, Pei. SLOM: A new measure for local spatial outliers. *Knowledge and Information Systems*. 9(4):412-429, April 2006
3. Liu, Xuan and Shekhar, Shashi and Chawla, Sanjay. Object-based directional query processing in spatial databases. *IEEE Transactions on Knowledge and Data Engineering*. 15(2):295-304, 2003
4. Shekhar, Shashi and Schrater, Paul and Vatsavai, Ranga Raju and Wu, Wei Li and Chawla, Sanjay. Spatial Contextual Classification and Prediction Models for Mining Geospatial Data. *IEEE Transactions on Multimedia*. 4(2):174-188, 2002
5. Shekhar, Shashi and Liu, Chang Tien and Chawla, Sanjay and Ravada, Sivakumar. Efficient Join-Index-Based Spatial-Join Processing: A Clustering Approach. *IEEE Transactions on Knowledge and Data Engineering*. 14(6):1400-1421, 2002
6. Liu, Xuan and Shekhar, Shashi and Chawla, Sanjay. Maintaining Spatial Constraints Using a Dimension Graph Approach. *International Journal on Artificial Intelligence Tools* 10(4): 639-662 (2001)
7. Shekhar, Shashi and Liu, Xuan and Chawla, Sanjay. An Object Model of Direction and Its Implications. *GeoInformatica* 3(4): 357-379 (1999)
8. Shekhar, Shashi and Chawla, Sanjay and Ravada, Sivakumar and Fetterer, Andrew and Liu, Xuan and Lu Chang-Tien. Spatial Databases - Accomplishments and Research Needs. *IEEE Transactions on Knowledge and Data Engineering* 11(1): 45-55 (1999)
9. Chawla, Sanjay and Lenhart Suzanne. Application of Optimal Control Theory to Bioremediation. *Journal of Computational and Applied Mathematics*, (114):81-102 (1999)

10. Chawla, Sanjay. A minmax problem in parabolic systems with competitive interactions. *Electronic Journal of Differential Equations, Volume (1999):1-18 (1999)*

Selective Conference and Workshop Papers

1. Al-Naymat, Ghazi and Chawla, Sanjay and Gudmundsson, Joachim. Dimensionality Reduction of Long Duration and Complex Spatio-Temporal Queries. **To Appear** In *Proceedings of the ACM Symposium on Applied Computing, 2007*
2. Verhein, Florian and Chawla, Sanjay. Geometrically Inspired Itemset Mining. **To Appear** in *Proceedings of the Sixth IEEE International Conference on Data Mining (ICDM 2006), December 18-22 2006, Hong Kong.*
3. Arunasalam, Bavani and Chawla, Sanjay. CCCS: a top-down associative classifier for imbalanced class distribution. In *Proceedings of the Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 517-522, Philadelphia, PA, USA, August 20-23, 2006. ACM 2006*
4. Sun, Pei and Chawla, Sanjay and Arunasalam, Bavani. Mining for Outliers in Sequential Databases. In *Proceedings of the Sixth SIAM International Conference on Data Mining, April 20-22, pp 94-106, Bethesda, MD, USA. SIAM 2006 Best Application Paper Award*
5. Verhein, Florian and Chawla, Sanjay. Mining Spatio-temporal Association Rules, Sources, Sinks, Stationary Regions and Thoroughfares in Object Mobility Databases. In *Proceedings of Database Systems for Advanced Applications, 11th International Conference (DASFAA), pp 187-201, Singapore, April 12-15, 2006, Springer Lecture Notes in Computer Science 3882*
6. Lukov, Lior and Chawla, Sanjay and Church Brett. Conditional Random Fields for Transmembrane Helix Prediction. In *Proceedings of the ninth Pacific-Asia Advances in Knowledge Discovery and Data Mining (PAKDD 2005), pp 155-161, Hanoi, Vietnam, May 18-20, 2005, Springer Lecture Notes in Computer Science 3518*
7. Arunasalam, Bavani and Chawla, Sanjay and Sun, Pei. Striking Two Birds with One Stone: Simultaneous Mining of Positive and Negative Spatial Patterns. In *Proceedings of the Fifth SIAM International Conference on Data Mining, April 21-23, pp, Newport Beach, CA, USA, SIAM 2005*
8. Chawla, Sanjay and Davis, Joseph and Pandey, Gaurav. On Local Pruning of Association Rules Using Directed Hypergraphs. In *Proceedings of IEEE the 20th International Conference on Data Engineering (ICDE 2004) pp 832, 30 March - 2 April 2004, Boston, MA, USA*
9. Sun, Pei and Chawla, Sanjay. On Local Spatial Outliers. In *Proceedings of the 4th IEEE International Conference on Data Mining (ICDM 2004), pp 209-216, 1-4 November 2004, Brighton, UK*
10. Munro, Robert and Chawla, Sanjay and Sun, Pei. Complex Spatial Relationships. In *Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM 2003), pp 227-234, 19-22 December 2003, Melbourne, Florida, USA*

11. Chawla, Sanjay and Arunasalam, Bavani and Davis, Joseph. Mining Open Source Software (OSS) Data Using Association Rules Network. *In Proceedings of the seventh Pacific-Asia Advances in Knowledge Discovery and Data Mining (PAKDD 2003) pp 461-466, Seoul, Korea, April 30 - May 2, 2003, Springer Lecture Notes in Computer Science 2637*
12. Chawla, Sanjay and Shekhar, Shashi and Wu, Weili. A Comparison of Markov Random Field and Spatial Regression Models for Mining Geospatial Data. *In Proceedings of the 6th Joint Conference on Information Science, pp 245-250, March 8-13, 2002, Research Triangle Park, North Carolina, USA. JCIS / Association for Intelligent Machinery, Inc. 2002*
13. Chawla, Sanjay and Shekhar, Shashi and Wu Weili and Ozesmi, Uygur. Extending Data Mining for Spatial Applications: A Case Study in Predicting Nest Locations. *ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery 2000 pp 70-77*
14. Chawla, Sanjay and Shekhar, Shashi and Wu, Weili and Ozesmi, Uygur. Predicting Locations Using Map Similarity (PLUMS): A Framework for Spatial Data Mining. *In Proceedings of the International Workshop on Multimedia Data Mining, MDM/KDD 2000, pp 14-24, August 20th, 2000, Boston, MA, USA. University of Alberta*
15. Liu, Xuan and Shekhar, Shashi and Chawla, Sanjay. Processing object-orientation-based direction queries: a summary of results. *Proceedings of the Eighth ACM Symposium on Advances in Geographic Information Systems, pp 69-76, November 10-11, 2000, Washington D.C., USA. ACM 2000*
16. Liu, Xuan and Shekhar, Shashi and Chawla, Sanjay. Consistency checking for Euclidean spatial constraints: a dimension graph approach. *In Proceedings of 12th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2000), pp 333-339, 13-15 November 2000, Vancouver, BC, Canada*
17. Ravada, Sivakumar and Shekhar, Shashi, Liu, Chang-Tien and Chawla, Sanjay. Optimizing Join Index Based Join Processing: A Graph Partitioning Approach. *In Proceedings of the Seventeenth Symposium on Reliable Distributed Systems, pp 302-308, October 20-22, 1998, West Lafayette, Indiana, USA*