

Chris Volinsky

Statistician

180 Park Ave
Florham Park, NJ
USA

1-973-360-8644

volinsky@research.att.com

Experience

- 2004–current **Executive Director, Statistics Research, AT&T Labs-Research**, Florham Park, NJ.
Manages Statistics Research department consisting of 13 statisticians, probabilists, and computer scientists. Responsibilities include developing a research agenda for the organization and maintaining a statistical center of excellence for the rest of the company. Projects include developing recommender systems, media consumption modelling, anomaly and fraud detection, statistical visualization, analysis of social networks, and data mining methodology.
- 1997–2004 **Member of Technical Staff, Statistics Research, AT&T Labs-Research**, Florham Park, NJ.
Research included development of a social network marketing strategy, detection of various types of telecommunications fraud, modelling the social network communications graph, and building new data mining algorithms.
- 1994–1997 **Research Assistant, Statistics Department, University of Washington**, Seattle, WA.
Research as part of dissertation work into Bayesian methodology, model averaging, and graphical models. Developed R software to implement new methodology.

Teaching

- Fall 2009 **Adjunct Assistant Professor, Columbia University**.
Will be teaching Data Mining course to Masters level statistics graduate students and Ph.D. students from other departments.
- July 2009 **Instructor, Explorations in Statistics Research Summer Program**, University of California, Berkeley.
Taught 2-day course in data mining, focusing on recommender systems.
- Summer 2008 **Instructor, Massey University**, New Zealand.
Taught Data Mining in a concentrated block course to grads and undergrads. Topics covered: data collection and quality, data presentation and visualization, classification, clustering, trees, neural networks, and other statistics and machine learning topics.
- 2002–2007 **SCEEP Tutor**.
Tutored for inner-city children through the Suburban Cultural Education and Enrichment Program. Two hours per week spent with an inner city child teaching math, reading, and writing.
- 1996–1997 **Instructor, University of Washington**.
Taught introductory statistics class designed for social science majors.
- 1992–1994 **Teaching Assistant, University of Washington**.
Led recitation sections and held office hours for students of introductory level statistics classes.

Education

- 1992–1997 **Ph.D. Statistics, University of Washington**, Seattle, WA, Thesis: Bayesian Model Averaging for Censored Survival Models. Advisor: Adrian E. Raftery.
- 1988–1992 **B.A. Statistics and Mathematics (double major), State University of New York at Buffalo**, Buffalo, NY.

Publications

Robert Bell, Yehuda Koren, and Chris Volinsky. Matrix factorization techniques for recommender systems. *IEEE Computer*, (to appear), 2009.

Emden Gansner, Yifan Hu, Stephen Kobourov, and Chris Volinsky. Putting recommendations on the map – visualizing clusters and relations. *Proceedings of the 3rd ACM Conference on Recommender Systems*, 2009.

Robert Bell, Jim Bennett, Yehuda Koren, and Chris Volinsky. The million dollar programming prize. *IEEE Spectrum*, 2009(05):28–33, 2009.

Richard A. Becker, Chris Volinsky, and Allan R. Wilks. Fraud detection at AT&T: A historical perspective. *Technometrics*, (to appear), 2009.

Yifan Hu, Yehuda Koren, and Chris Volinsky. Collaborative filtering for implicit feedback datasets. In *IEEE International Conference on Data Mining (ICDM)*, 2008.

Yehuda Koren, Stephen C. North, and Chris Volinsky. Measuring and extracting proximity graphs in networks. *ACM Transactions on Knowledge Discovery from Data*, 1(3):12, 2007.

Robert M. Bell, Yehuda Koren, and Chris Volinsky. Chasing \$1,000,000: How we won the Netflix Progress Prize. *Statistical Computing and Statistical Graphics Newsletter*, 18:4–12, 2007.

Robert M. Bell, Yehuda Koren, and Chris Volinsky. Modeling relationships at multiple scales to improve accuracy of large recommender systems. In *KDD '07: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 95–104, 2007.

Robert M. Bell, Patrick G. Haffner, and Chris Volinsky. Modifying boosted trees to improve performance on Task 1 of the 2006 KDD Challenge Cup. *SIGKDD Explorations*, 8:47–52, 2006.

Yehuda Koren, Stephen C. North, and Chris Volinsky. Measuring and extracting proximity in networks. In *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 245–255, New York, NY, USA, 2006. ACM Press.

Shawndra Hill, Deepak Agarwal, Robert Bell, and Chris Volinsky. Building an effective representation for dynamic networks. *Journal of Computational and Graphical Statistics*, 15(3):584–608, 2006.

Shawndra Hill, Foster Provost, and Chris Volinsky. Network-based marketing: Identifying likely adopters via consumer networks. *Statistical Science*, 21:256–276, 2006.

Shawndra Hill, Foster Provost, and Chris Volinsky. The gift of gab: Evidence tele-commerce firms can profit from viral marketing. In *Statistical Challenges in ECommerce: The First Interdisciplinary Symposium Between Information Systems, Statistics, and Related Fields*, 2005.

Adrian E. Raftery, Ian Painter, and Chris Volinsky. BMA: An R package for Bayesian Model Averaging. *R News*, 5:2–8, 2005.

Shona Douglas, Deepak Agarwal, Tirso Alonso, Robert Bell, Mazin Rahim, Deborah Swayne, and Chris Volinsky. Mining customer care dialogs for "daily news". *IEEE Transactions on Speech and Audio Processing (Special Data Mining Issue)*, 13:652–660, 2005.

Brendan Wintle, M.A. McCarthy, Chris Volinsky, and R. P. Kavanagh. The use of Bayesian model averaging to better represent uncertainty in the predictions of ecological models. *Conservation Biology*, 17:1579–1590, 2003.

Corinna Cortes, Daryl Pregibon, and Chris Volinsky. Computational methods for dynamic graphs. *Journal of Computational and Graphical Statistics*, 12:950–970, 2003.

Antony Unwin, Chris Volinsky, and Sylvia Winkler. Parallel coordinates for exploratory modelling analysis. *Computational Statistics and Data Analysis*, 28:553–564, 2002.

Scott R. Millis and Chris Volinsky. Assessment of response bias in mild head injury: Beyond malingering tests. *Journal of Clinical and Experimental Neuropsychology*, 23:809–828, 2001.

Corinna Cortes, Daryl Pregibon, and Chris Volinsky. Communities of interest. In *Proceedings of Intelligent Data Analysis 2001*, pages 105–114, 2001.

Chris T. Volinsky and Adrian E. Raftery. Bayesian information criterion for censored survival models. *Biometrics*, 56:256–262, 2000.

Bill DuMouchel, Chris Volinsky, Theodore Johnson, Corinna Cortes, and Daryl Pregibon. Squashing flat files flatter. In *KDD '99: Proceedings of the 5th Int. Conf. on Knowledge Discovery and Data Mining*, pages 6–15. AAAI Press, 1999.

Jennifer A. Hoeting, David Madigan, Adrian Raftery, and Chris T. Volinsky. Bayesian Model Averaging: A tutorial. *Statistical Science*, 14:382–401, 1999.

Chris T. Volinsky and Adrian E. Raftery. Bayesian information criterion for censored survival models. Technical Report 349, Department of Statistics, University of Washington, 1999. (<http://www.stat.washington.edu/tech.reports/tr349.ps>).

Chris Volinsky. *Bayesian Model Averaging for Censored Survival Models*. PhD thesis, University of Washington, 1997.

Chris T. Volinsky, David Madigan, Adrian E. Raftery, and Richard A. Kronmal. Bayesian Model Averaging in proportional hazard models: Assessing the risk of a stroke. *Applied Statistics*, 46(3):443–448, 1997.

David Madigan, Adrian E. Raftery, Chris Volinsky, and Jennifer Hoeting. Bayesian Model Averaging. In P. Chan, S. Stofofo, and D. Wolpert, editors, *Integrating Multiple Learned Models (IMLM96)*, 1996.

David Madigan, Steen A. Andersson, Michael Perlman, and Chris T. Volinsky. Bayesian Model Averaging and model selection for Markov equivalence classes of acyclic digraphs. *Communications in Statistics: Theory and Methods*, 25:2493–2520, 1996.

Adrian E. Raftery, David Madigan, and Chris T. Volinsky. Accounting for model uncertainty in survival analysis improves predictive performance (with discussion). In J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M Smith, editors, *Bayesian Statistics 5*, pages 323–349. Oxford University Press, 1995.

Selected Presentations at Conferences and Departments

- 2009 **Invited Talk**, *NYU Workshop on Managing Data Risk*, New York University, "Risk and Reward in Data Analytics".
- Invited Talk**, *Center for Statistics and the Social Sciences*, University of Washington, "Recommender Systems for Fun and Profit".
- Panel Organizer/Moderator**, *Fifth Symposium on Statistical Challenges in Electronic Commerce Research*, Carnegie Mellon University, "Social Media Research in Practice".
- 2008 **Plenary Talk**, *9th INFORMS Telecommunications Conference*, University of Maryland, MD, "What Data Mining Isn't".
- Plenary Talk**, *17th Annual Frontiers in Service Conference*, Washington, DC, "What Data Mining Isn't".
- Keynote Talk**, *SAS Global Forum, Section on Communications, Media, Entertainment, and Travel*, San Antonio, TX, "What Data Mining Isn't".
- Invited Talk**, *Center for Data Analysis and Statistics*, US Military Academy, West Point, NY, "Recommender Systems for Fun and Profit".

- Invited Talk**, *Fourth Symposium on Statistical Challenges in Electronic Commerce Research*, Stern School of Business, New York, NY, “Recommender Systems for Media Consumption: Netflix and More”.
- Invited Talk**, *Massey University, Institute of Information and Mathematical Sciences*, Auckland, NZ, “Recommender Systems for Fun and Profit”.
- 2007 **Invited Talk**, *9th Annual AT&T Cyber Security Conference*, Bridgewater, NJ, “Communities of Interest, Concepts and Applications”.
- Graduate Student Talk**, *Wharton School of Business, Statistics Department*, Philadelphia, PA, “Statistics Research at AT&T”.
- Invited Talk**, *Blair Academy*, Blairstown, NJ, “Data for Fraud Detection: Catching Bad Guys Using Statistics”.
- 2006 **Invited Talk**, *12th SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2006)*, Philadelphia, PA, “Measuring and Extracting Proximity in Networks”.
- Invited Talk**, *Joint Statistical Meetings*, Seattle, WA, “Modelling Massive Dynamic Graphs”.
- Invited Talk**, *IWAP: International Workshop on Applied Probability*, Storrs, CT, “Modelling Massive Dynamic Graphs”.
- 2005 **Invited Talk**, *New York University, Statistics Department*, New York, NY, “Modelling Massive Dynamic Graphs”.
- Invited Talk**, *Yale University, Statistics Department*, New Haven, CT, “Modelling Massive Dynamic Graphs”.
- Invited Talk**, *Workshop on Data Mining, Fields Institute*, Toronto, ON, “Modelling Massive Dynamic Graphs”.
- 2004 **Invited Talk**, *Columbia University, Statistics Department*, New York, NY, “Modelling Massive Dynamic Graphs”.
- 2003 **Invited Talk**, *Workshop on Statistical Inference, Computing and Visualization for Graphs*, Stanford University, CA, “Exploratory Data Analysis with Graphs”.
- 2002 **Invited Talk**, *Joint Statistical Meetings*, New York, NY, “Mining of Communication Networks”.

Conferences and Conference Sessions Chaired / Organized

- 2008 **Program Co-Chair**, *KDD 2008, Industrial/Government Applications Track*, Las Vegas, NV.
- 2008 **Session Organizer**, *Joint Statistical Meetings, Invited Session*, Denver, CO, “Analysis of Online Communities”.
- 2006 **Session Organizer**, *Joint Statistical Meetings, Invited Session*, Seattle, WA, “Personal Networks: Applications Using Data on Social Relationships”.
- 2004 **Committee Chair and Workshop Organizer**, *IMS New Researchers Conference*, Toronto, ON.
- 2003 **Organizer and Workshop Co-Chair**, *Workshop on Statistical Inference, Computing and Visualization for Graphs*, Stanford, CA, www.research.att.com/~volinsky/Graphs.

Service

- Associate Editor 1999–2007 *Journal of Computational and Graphical Statistics*.
- Program Committees International Chinese Statistical Association (2010)
- Committees KDD Research Track
KDD Industrial Track
WWW2009: Social Networks and Web 2.0 Track
ACM Conference on Electronic Commerce
KDD Workshop on Web Mining and Social Network Analysis
ACM Recommender Systems
- Referee Biometrika
Biostatistics
Data and Knowledge Engineering
IEEE Transactions on Knowledge and Data Engineering
Journal of Computational and Graphical Statistics
Journal of Machine Learning Research
Journal of the American Statistical Association
Optimization Methods in Software
Statistical Methodology
Transactions on Knowledge Discovery in Data
- Committees 2009-2011 Member, ASA - JCGS Management Committee
2004 Chair, IMS New Researchers Committee
2002–2004 Member, IMS New Researchers Committee

Supervision

- Ph.D. Committee Co-chair of dissertation committee for Shawndra Hill (Assistant Professor, Wharton School of Business)
- Advisor Summer Intern advisor for Valerie Hyde (Accenture, Inc., New York, NY)
- Mentor Mentor, AT&T Labs Fellowship Program, for Jennifer Neville (Assistant Professor, Purdue University)

Awards and Prizes

- 2007-2008 **Netflix Prize**, *Winner of 2007 and 2008 \$50,000 Progress Prize.*, Prize awarded to the team with the best recommender system at the one year point of the US\$1M Netflix Prize competition. Joint with Yehuda Koren and Bob Bell (see www.research.att.com/~volinsky/netflix.html).
- 2006 **KDD Challenge Cup**, Winner of Task 1 of the annual data mining competition at the KDD conference, Joint with Bob Bell and Patrick Haffner.
- 1997 **ZW Birnbaum Award, University of Washington Statistics Department**, Annual award given to the graduate student with the best General Exam.
- 1994 **Teaching Award, University of Washington Statistics Department**, Annual award given to the best teacher among the graduate students..

Patents

- US 6,539,391 Method and System for Squashing a Large Data Set.
- Filed 2009 Health Monitoring and Onset Analysis.
- Filed 2008 Analytical Entertainment Services.
- Filed 2008 Social Health Tracker.
- Filed 2007 Modeling Relationships at Multiple Scales to Improve Accuracy.
- Filed 2007 Mobile Speech Search Using Community of Interest.
- Filed 2007 Method and Apparatus to Identify Influencers.
- Filed 2006 Method for Identifying Likely Adopters via Consumer Networks.

Media Coverage

- 2008 **New York Times Magazine**, "*If You Liked This, You're Sure to Love That*", November 28, 2008, Online version available at www.tinyurl.com/5byw73. Accompanying video segment available at www.tinyurl.com/59nd54.
- 2008 **CNBC**, "*Profiting from Collaboration*", Broadcast October, 2008., Part 3 of the 5 part "Collaboration Now" series. Online video available at www.tinyurl.com/8tozfa (Netflix segment begins at 39:15 into the show) .
- 2007 **AT&T Tech Channel**, *Interview with Chris Volinsky and Bob Bell*, Online video available at www.tinyurl.com/9gmecv.

References

- Wolfgang Jank Associate Professor, Department of Decisions, Operations, and Information Technologies, Robert H. Smith School of Business, University of Maryland. Email: wjank@rhsmith.umd.edu
- Foster Provost Professor, Department of Information, Operations and Management Sciences, Leonard N. Stern School of Business, NYU. Email: fprovost@stern.nyu.edu
- Daryl Pregibon Google, Inc.. Email: daryl@google.com
- Adrian Raftery Blumstein-Jordan Professor of Statistics and Sociology, University of Washington. Email: raftery@stat.washington.edu