

Xiaolei Li

me@xiaolei.org

520 4th St. Unit 39
Kirkland, WA 98033

(217) 390-8478
U.S. Citizen

Objectives

Seeking a position where I can maximize my technical and leadership skills while utilizing my knowledge in the fields of data mining, anomaly detection, and machine learning.

Employment

- July 2008–Present* Developer and Applied Researcher, Microsoft, Redmond, WA.
Developed algorithms in AdCenter for monitoring, diagnostics, and simulation.
Invited to MS Bench Program (top 4% of MS Employees) and received Gold Star.
- May 2007–Aug 2007* Summer intern, Google, Mountain View, CA.
Designed algorithms to detect click spam and fraud in online advertising.
- Aug 2006–May 2008* Contract research programmer, General Motors.
Created algorithms and UI to detect multi-dimensional outliers in large time series datasets.
- May 2006–Aug 2006* Summer intern, Google, Mountain View, CA.
Designed algorithms to detect click spam and fraud in online advertising.
- May 2005–Aug 2005* Summer intern, HRL Laboratories, Malibu, CA.
Designed algorithms for natural language processing and social network planning.
- May 2003–May 2005* Research assistant, National Center for Supercomputing Applications (NCSA).
Developed machine learning and information retrieval software.
- 1996–2001* Summer intern, database and web programmer, EnSite, Inc., Omaha, NE.
Developed software for gas and electricity supply management.

Education

- Aug 2004–May 2008* University of Illinois at Urbana-Champaign, Ph.D. in Computer Science.
Thesis: Multidimensional Analysis of Moving Object Data (Advisor: Jiawei Han).
- Aug 2002–May 2004* University of Illinois at Urbana-Champaign, M.S. in Computer Science.
Thesis: High-Dimensional OLAP: A Minimal Cubing Approach.
- Aug 1999–May 2002* University of Illinois at Urbana-Champaign, B.S. in Computer Science.
Mathematics Minor. 3.73 overall GPA, 3.90 technical GPA.

Entrepreneurship

- Nov 1996–2002* Founder and operator of Wickedland.com, a music blog.
Reached over 10k daily uniques.
Sold in 2000 and stayed on as day-to-day operator.
- Aug 2001–Sept 2002* Founder and operator of Polytope Records.
Funded and released one record in December of 2001.

Publications

All listed in chronological order of publication date.

Refereed Journals

- [1] “Modeling Massive RFID Data Sets: A Gateway-Based Movement Graph Approach.” Hector Gonzalez, Jiawei Han, Hong Cheng, Xiaolei Li, Diego Klabjan, Tianyi Wu. IEEE Transactions on Knowledge and Data Engineering (TKDE), Vol 22 (1), 2010.
- [2] “Computing Iceberg Cubes by Top-Down and Bottom-Up Integration: The Star-Cubing Approach.” Dong Xin, Jiawei Han, Xiaolei Li, Zheng Shao, Benjamin W. Wah. IEEE Transactions on Knowledge and Data Engineering (TKDE), Vol 19 (1), 2007.
- [3] “Making SVMs Scalable to Large Data Sets using Hierarchical Cluster Indexing.” Hwanjo Yu, Jiong Yang, Jiawei Han, Xiaolei Li. Data Mining and Knowledge Discovery: An International Journal (DMKD). 2005.

Refereed Conferences

- [4] “Graph Cube: On Warehousing and OLAP Multidimensional Networks.” Peixiang Zhao, Xiaolei Li, Dong Xin, Jiawei Han. ACM Int. Conf. on Management of Data (SIGMOD), Athens, Greece, June 2011.
- [5] “Incremental Clustering for Trajectories.” Zhenhui Li, Jae-Gil Lee, Xiaolei Li, Jiawei Han. Database Systems for Advanced Applications (DASFAA), Tsukuba, Japan, April 2010.
- [6] “Temporal Outlier Detection in Vehicle Traffic Data.” Xiaolei Li, Zhenhui Li, Jiawei Han, Jae-Gil Lee. IEEE Int. Conf. on Data Engineering (ICDE), Shanghai, China, April 2009.
- [7] “TraClass: Trajectory Classification Using Hierarchical Region-Based and Trajectory-Based Clustering.” Jae-Gil Lee, Jiawei Han, Xiaolei Li, and Hector Gonzalez, Int. Conf. on Very Large Data Base (VLDB), Auckland, New Zealand, Aug. 2008.
- [8] “Sampling Cube: A Framework for Statistical OLAP over Sampling Data.” Xiaolei Li, Jiawei Han, Zhijun Yin, Jae-Gil Lee, and Yizhou Sun. ACM Int. Conf. on Management of Data (SIGMOD), Vancouver, BC, Canada, June 2008.
- [9] “Trajectory Outlier Detection: A Partition-and-Detect Framework.” Jae-Gil Lee, Jiawei Han, and Xiaolei Li. IEEE Int. Conf. on Data Engineering (ICDE), Cancun, Mexico, April 2008.
- [10] “Mining Approximate Top- K Subspace Anomalies in Multi-Dimensional Time-Series Data.” Xiaolei Li, Jiawei Han. Int. Conf. on Very Large Data Bases (VLDB), Vienna, Austria, Sept 2007.
- [11] “DataScope: Viewing Database Contents in Google Maps’ Way.” Tianyi Wu, Xiaolei Li, Dong Xin, Jiawei Han, Jacob Lee, Ricardo Redder. Int. Conf. on Very Large Data Bases (VLDB), Vienna, Austria, Sept 2007.
- [12] “Adaptive Fastest Path Computation on a Road Network: A Traffic Mining Approach.” Hector Gonzalez, Jiawei Han, Xiaolei Li, Margaret Myslinska, John Paul Sondag. Int. Conf. on Very Large Data Bases (VLDB), Vienna, Austria, Sept 2007.

- [13] “Traffic Density-based Discovery of Hot Routes in Road Networks.” Xiaolei Li, Jiawei Han, Jae-Gil Lee, Hector Gonzalez. Int. Symp. on Spatial and Temporal Databases (SSTD), Boston, MA, Jul 2007.
- [14] “ROAM: Rule- and Motif-Based Anomaly Detection in Massive Moving Object Data Sets.” Xiaolei Li, Jiawei Han, Sangkyum Kim, Hector Gonzalez. SIAM Int. Conf. on Data Mining (SDM), Minneapolis, MN, April 2007.
- [15] “Mining Compressed Commodity Workflows From Massive RFID Data Sets.” Hector Gonzalez, Jiawei Han, Xiaolei Li. ACM Int. Conf. on Information and Knowledge Management (CIKM), Arlington, VA, Nov. 2006.
- [16] “Answering Top- K Queries with Multi-Dimensional Selections: The Ranking Cube Approach.” Dong Xin, Jiawei Han, Hong Cheng, Xiaolei Li. Int. Conf. on Very Large Data Bases (VLDB), Seoul, Korea, Sept 2006.
- [17] “FlowCube: Constructing RFID FlowCubes for Multi-Dimensional Analysis of Commodity Flows.” Hector Gonzalez, Jiawei Han, Xiaolei Li. Int. Conf. on Very Large Data Bases (VLDB), Seoul, Korea, Sept 2006.
- [18] “Motion-Alert: Automatic Anomaly Detection in Massive Moving Objects.” Xiaolei Li, Jiawei Han, Sangkyum Kim. IEEE Intelligence and Security Informatics Conference (ISI), San Diego, CA, May 2006.
- [19] “Warehousing and Analyzing Massive RFID Data Sets.” Hector Gonzalez, Jiawei Han, Xiaolei Li, Diego Klabjan. IEEE Int. Conf. on Data Engineering (ICDE), Atlanta, GA, April 2006. Winner of best student paper.
- [20] “Mining Evolving Customer-Product Relationships in Multi-Dimensional Space.” Xiaolei Li, Jiawei Han, Xiaoxin Yin, Dong Xin. IEEE Int. Conf. on Data Engineering (ICDE), Tokyo, Japan, April 2005.
- [21] “Scalable Construction of Topic Directory with Nonparametric Closed Termset Mining.” Hwanjo Yu, Duane Sears Smith, Xiaolei Li, Jiawei Han. IEEE Int. Conf. on Data Mining (ICDM), Brighton, UK, November 2004.
- [22] “High-Dimensional OLAP: A Minimal Cubing Approach.” Xiaolei Li, Jiawei Han, Hector Gonzalez. Int. Conf. on Very Large Data Bases (VLDB), Toronto, Canada, August 2004.
- [23] “Star-Cubing: Computing Iceberg Cubes by Top-Down and Bottom-Up Integration.” Dong Xin, Jiawei Han, Xiaolei Li, Benjamin W. Wah. Int. Conf. on Very Large Data Bases (VLDB), Berlin, Germany, September 2003.

Languages and Skills

Languages C#, C/C++, Java, Python, Perl, PHP, JavaScript.

Technologies MS SQL Server, MySQL, HTML, .NET, ASP.NET, WCF, Unix/Linux, Photoshop.

References

Available upon request.