

Lipyeow Lim

Curriculum Vitae

Home Address

194 Nob Hill Dr.
Elmsford, NY 10523
(914) 347-3915

Office Address

19 Skyline Dr.
Hawthorne, NY 10532
(914) 784-6156

email: liplim@us.ibm.com

<http://www.research.ibm.com/people/l/lipyeow/>

Education

Aug 2004 *Ph.D., Computer Science*

Department of Computer Science, Duke University, NC, U.S.A.

Dissertation: Online methods for database optimization.

Advisor: Jeffrey Scott Vitter (Purdue)

Committee: Min Wang (IBM), Ronald Parr, and Jun Yang

GPA: 3.9

Sep 2000 *M.Sc., Information Systems and Computer Science*

School of Computing, National University of Singapore, Singapore.

Dissertation: A theoretical look at pixel ordering.

Advisor: Philip M. Long

Committee: Ghim-Hwee Ong, and Wee-Sun Lee

May 1999 *B.Sc. (accelerated 3-year honors), Information Systems and Computer Science*

School of Computing, National University of Singapore, Singapore.

Thesis Project: Implementation of the mobile IPv4 configuration option for PPP IPCP (RFC 2290).

Advisor: Yong-Chiang (Y. C.) Tay

Committee: Kam-Hong Shum & Kwok-Yan Lam

GPA: 3.5

Skills

Programming Languages: C, C++, PERL, Lisp, Prolog, Bash, Tcsh, Matlab, Java.

APIs: Standard Template Library (STL), OpenGL, Message Passing Interface (MPI), X/Motif, Visual C++/MFC.

Operating Systems: Microsoft Windows NT/98/2K/XP, Unix, Sun Solaris, Linux.

Research Experience

- Oct 2004–present *Research Staff Member.*
IBM T. J. Watson Research Center, NY, U.S.A.
Research Area: Database Optimization
- Sep 2000–Sep 2004 *Research Assistant* (Advisor: Prof. Jeffrey Scott Vitter).
Department of Computer Science, Duke University, NC, U.S.A.
Research Area: On-line Methods for Database Optimization
- Jun–Sep 2003 *Research Intern* (Mentor: Min Wang)
IBM T. J. Watson Research Center, NY, U.S.A.
Research Topic: Statistics Collection for IBM DB2XML Optimization.
- Jun–Aug 2001 *Research Intern* (Mentor: Min Wang)
IBM T. J. Watson Research Center, NY, U.S.A.
Research Topic: On-line Statistics for Native XML Databases. (Published in VLDB 2002).
- May–Aug 2000 *Research Intern* (Mentor: Min Wang)
IBM T. J. Watson Research Center, NY, U.S.A.
Research Topic: Keeping Web Indexes Up-to-date Using The Landmark-Diff Method. (Published in WWW 2003). Patent pending.
- Jul 1998–Aug 1999 *Research Assistant* (Advisor: Prof. Philip M. Long).
School of Computing, National University of Singapore, Singapore.
Research Area: Pixel Ordering in Image Compression.
- Jul 1997–Jun 1998 *Member of the Mobile-IP Research Group* (Advisor: Prof. Y. C. Tay).
School of Computing, National University of Singapore, Singapore.
Research Area: Inter-operability of the PPP and MobileIP networking protocols.

Teaching Experience

- Jan–Apr 2001 *Teaching Assistant* (to Prof. Curry Guinn)
Department of Computer Science, Duke University, NC, U.S.A.
CPS 170: Artificial Intelligence (Undergraduate level).
- Sep–Dec 2000 *Teaching Assistant* (to Prof. Lars Arge)
Department of Computer Science, Duke University, NC, U.S.A.
CPS 230: Analysis of Algorithms (Graduate level).
- Jan–Jun 1999 *Teaching Assistant* (to Prof. Kian-Lee Tan)
School of Computing, National University of Singapore, Singapore.
CS2102: Introduction to Database Systems
- Jan–Jun 1998 *Teaching Assistant* (to Prof. Weng-Fai Wong)
School of Computing, National University of Singapore, Singapore.
IC1273: C Programming & Unix Tools

Honors and Awards

- 2003–2004 *IBM PhD Fellowship.*
International Business Machines.
- 2002/2003 *Outstanding Service to the Department Award.*
Department of Computer Science, Duke University, NC, U.S.A.
- 1999/2000 *Graduate Fellowship.*
Department of Computer Science, Duke University, NC, U.S.A.
- 1995/1996 *On the Dean's List.*
School of Computing, National University of Singapore, Singapore.
- 1996/1997 *On the Dean's List.*
School of Computing, National University of Singapore, Singapore.

Publications and Papers

Refereed Articles

1. *CXHist: An On-line Classification-Based Histogram for XML String Selectivity Estimation*. Lipyeow Lim, Min Wang, and Jeffrey S. Vitter. Submitted for publication, 2004.
2. *SASH: A Self-Adaptive Histogram Set for Dynamically Changing Workloads*. Lipyeow Lim, Min Wang, and Jeffrey S. Vitter. *Very Large Databases*, pages 369–380, 2003.
3. *Dynamic Maintenance of Web Indexes Using Landmarks*. Lipyeow Lim, Min Wang, Sriram Padmanabhan, Jeffrey S. Vitter, and Ramesh C. Agarwal. *Proceedings of the 12th Intl. World Wide Web Conf.*, pages 102–111, 2003.
4. *XPathLearner: An On-Line, Self-Tuning Markov Histogram for XML Path Selectivity Estimation*. Lipyeow Lim, Min Wang, Sriram Padmanabhan, Jeffrey S. Vitter, and Ronald Parr. *Very Large Databases*, pages 442–453, 2002.
5. *Wavelet-Based Cost Estimation for Spatial Queries*. Min Wang, Jeffrey S. Vitter, Lipyeow Lim, and Sriram Padmanabhan. *Symposium on Spatial and Temporal Databases*, pages 175–196, 2001.
6. *Characterizing Web Document Change*. Lipyeow Lim, Min Wang, Sriram Padmanabhan, Jeffrey S. Vitter, and Ramesh C. Agarwal. *Web-Age Information Management*, pages 133–144, 2001.
7. *A Theoretical Look at Pixel Ordering*. Lipyeow Lim. M.Sc. Thesis, School of Computing, National University of Singapore, Singapore., 1999.

Unpublished Articles

1. *An Efficient Method to Find Motifs from Expression and Sequence Data*. Allister Bernard, Lipyeow Lim, and Ryan Deering. 2002.
2. *Pixel Ordering with Markov Decision Processes*. Lipyeow Lim. 2000.
3. *Wavelets: An Overview*. Sanjay Banajee and Lipyeow Lim. 2000.
4. *Cache Efficient Algorithms*. Lipyeow Lim and Vijay Natarajan. 1999.
5. *Applying Linear Cryptanalysis to Reduced Rounds DES*. Lipyeow Lim. 1997.

Research Projects

- Sep 2003–present *Classification-based Histograms.*
Research on a histogram technique that uses Bayesian classifiers to remember the values associated with a bucket approximately. Histograms are used to store statistics for database query optimization.
- Jun–Sep 2003 *IBM DB2XML Product Development.*
Internship project at IBM T. J. Watson Research Center. Prototyped the runstats routine in DB2XML, IBM’s new native XML DBMS. The runstats routine collects statistics used in query optimization.
- Jul 2002–present *Self-Adaptive Set of Histograms.*
Research on building a set of histograms to store statistics for relational database query optimization without scanning the underlying data. Statistical graphical models are used to learn which attributes to build histograms on and how to build these histograms. Collaborators: Jeffrey S. Vitter (Purdue) and Min Wang (IBM).
- Jun 2001–present *XML Path Selectivity.*
Research on collecting statistics for query optimization in native XML databases without scanning the underlying data. Started as an internship project at IBM T. J. Watson Research Center, this research produced XPathLearner (VLDB 2002), an on-line method for estimating the selectivity of XML path expressions. Collaborators: Jeffrey S. Vitter (Purdue), Min Wang (IBM) and Ronald Parr (Duke).
- Mar–May 2002 *Finding Motifs from Gene Expression and Sequence Data.*
Research on an efficient algorithm to find candidate motifs in upstream regions by combining information from both gene expression data and gene sequencing data. Course project for graduate level course *Computational Functional Genomics*. Collaborators: Allister Bernard and Ryan Deering.
- May 2000–May 2001 *Updating Web Indexes.*
Research on how to keep web indexes up-to-date. Started as an internship project at IBM T. J. Watson Research Center, this research produced the novel landmark-diff method (published in WWW 2003) for updating web indexes. Collaborators: Jeffrey S. Vitter (Purdue), Min Wang (IBM), Sriram Padmanabhan (IBM), Ramesh Agarwal (IBM). Patent pending.

Research Projects (Continued)

- Sep–Nov 2000 *Pixel Ordering with Markov Decision Processes (MDPs).*
Research on using MDPs to find an optimal path to traverse all the pixels of an image for optimal compression using a predictive encoder. Course project for graduate-level course *Markov Decision Processes*.
- Mar–May 2000 *Using Prediction to Manage Power-Aware Memory.*
Research on using different prediction methods to predict memory usage patterns for powering up/down of memory chips. Course project for graduate-level course *Operating Systems*. Collaborators: Martin Gilbert.
- Nov 1999–Mar 2000 *Selectivity Estimation in Spatial Database.*
Research on using wavelets to compress sparse 2D histograms that are used in query optimization in spatial databases. Collaborators: Jeffrey S. Vitter (Duke) and Min Wang (IBM).
- Jul 1998–Aug 1999 *Image Compression Research.*
Research on image compression techniques using predictive-differential coding, discrete cosine transform (DCT), wavelet transforms, hybrid transforms and pixel ordering. Advisor: Philip M. Long.
- Jul 1997–Jun 1998 *Mobile-IP Research.*
Implemented RFC 2290 *MIPv4 Option for PPP IPCP* on Linux's PPP source. The scope of the project covers issues related to the interworking of the MIPv4 protocol and the PPP protocol. This is the first implementation of this RFC in the world.
- Jul–Dec 1997 *Cryptanalysis Research.*
Implemented and benchmarked the linear cryptanalysis method (Mitsuru Matsui) and the differential cryptanalysis method on a reduced rounds DES. Course project for the graduate-level course *Cryptographic systems*.
- Jul–Dec 1997 *Parallel Matrix Multiplication.*
Implemented and benchmarked parallel versions of a block-matrix multiplication algorithm and the Strassen's matrix multiplication algorithm using the Message Passing Interface (MPI) standard. Course project for the graduate-level class *Parallel and Distributed Algorithms*.

Service

- 2004 *Reviewer for World Wide Web conference posters.*
- 2003 *Reviewer for IEEE Intelligent Systems magazine.*
- 2002/2003 *Graduate Student Liaison to the Faculty.*
Department of Computer Science, Duke University, NC, U.S.A.
Represented the students in faculty meetings and coordinated the resolution of issues.
- 2002/2003 *Member of Leadership Team*
Graduate Inter-Varsity Christian Fellowship, Duke University, NC, U.S.A.
- 2001–2004 *Founder and Instructor of the Duke Taijiquan Club.*
Duke University, NC, U.S.A.

References

1. *Jeffrey S. Vitter*, Purdue University, (765) 494-1730, jsv@purdue.edu.
2. *Min Wang*, IBM T. J. Watson, (914) 784-6268, min@us.ibm.com.
3. *Ronald Parr*, Duke University, (919) 660-6537, parr@cs.duke.edu.
4. *Jun Yang*, Duke University, (919) 660-6587, junyang@cs.duke.edu.
5. *Sriram Padmanabhan*, IBM Silicon Valley Labs, (408) 463-2712, srp@us.ibm.com.

Personal

- Born March 15 1974 in Singapore; Citizen of Singapore; Married.
- U.S. Visa Status: F-1
- Language Ability: English, Mandarin, German, Cantonese and the Teochew dialect.