

# Lipyew Lim

## Home Address

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

## Office Address

19 Skyline Dr,  
Hawthorne, NY 10532  
(919) 784-6156

**email:** [liplim@us.ibm.com](mailto:liplim@us.ibm.com)

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

---

## Education

**Ph.D. Candidate, Computer Science, 1999-2004**  
Duke University, Durham, NC.

- Thesis topic: *Online Methods for Database Optimization.*

**M.Sc. in Information System and Computer Science, 1998-2000**  
National University of Singapore.

- Thesis topic: *A theoretical look at pixel ordering.*

**B.Sc.(Honors) in Information System and Computer Science, 1995-1998**  
National University of Singapore.

- Thesis/Project topic: *Implementation of the mobile IPv4 configuration option for PPP IPCP.*

## Publications

**Submitted for publication** CXHist: An On-line Classification-Based Histogram for XML String Selectivity Estimation. Lipyew Lim, Min Wang, Jeffrey S. Vitter.

**VLDB 2003** SASH: A Self-Adaptive Histogram Set for Dynamically Changing Workloads. Lipyew Lim, Min Wang, Jeffrey S. Vitter.

**WWW 2003** Dynamic Maintenance of Web Indexes Using Landmarks. Lipyew Lim, Min Wang, Sriram Padmanabhan, Jeffrey S. Vitter, Ramesh C. Agarwal.

**VLDB 2002** XPathLearner: An On-Line, Self-Tuning Markov Histogram for XML Path Selectivity Estimation. Lipyew Lim, Min Wang, Sriram Padmanabhan, Jeffrey S. Vitter, Ronald Parr.

**SSTD 2001.** Wavelet-Based Cost Estimation for Spatial Queries. Min Wang, Jeffrey S. Vitter, Lipyew Lim, Sriram Padmanabhan. Symposium on Spatial and Temporal Databases.

**WAIM 2001.** Characterizing Web Document Change. Lipyew Lim, Min Wang, Sriram Padmanabhan, Jeffrey S. Vitter, Ramesh C. Agarwal. Web-Age Information Management 2001.

## References

**Jeffrey S. Vitter**, Purdue University, (765) 494-1730, [jsv@purdue.edu](mailto:jsv@purdue.edu).

**Min Wang**, IBM T. J. Watson, (914) 784-6268, [min@us.ibm.com](mailto:min@us.ibm.com).

**Ronald Parr**, Duke University, (919) 660-6537, [parr@cs.duke.edu](mailto:parr@cs.duke.edu).

**Jun Yang**, Duke University, (919) 660-6587, [junyang@cs.duke.edu](mailto:junyang@cs.duke.edu).

**Sriram Padmanabhan**, IBM Silicon Valley Labs, (408) 463-2712, [srp@us.ibm.com](mailto:srp@us.ibm.com).

## Experience

**IBM Research Staff Member.** Current employment.

**IBM DB2XML Product Development.** Internship project at IBM T. J. Watson Research Center in summer (June-September) 2003. Prototyped the runstats routine in DB2XML, IBM's new native XML DBMS. The runstats routine collects statistics that are used for query optimization. Mentor: Min Wang (IBM).

**Self-Adaptive Set of Histograms.** On-going database research on using statistical graphical models to learn which attributes to build histograms on and how to build these histograms for estimating query result sizes without the requirement to scan the data. Collaborators: my advisor Jeffrey S. Vitter (Purdue) and Min Wang (IBM).

**XML Path Selectivity.** On-going database research on estimating the selectivity of XML path expressions in XML data using query feedback. This project started during an internship at IBM T. J. Watson Research Center in summer 2001. Part of this research produced XPathLearner (published in VLDB 2002), an online method for estimating the selectivity of XML path expressions. Collaborators: my advisor Jeffrey S. Vitter (Purdue), Min Wang (IBM) and Ronald Parr (Duke).

**Updating Web Indexes.** Research on techniques for keeping web indexes up-to-date. This project started during an internship at IBM T. J. Watson Research Center in summer (May-August) 2000 and produced the landmark-diff method (published in WWW 2003) for keeping web indexes up-to-date. Collaborators: my advisor Jeffrey S. Vitter (Purdue) and Min Wang (IBM), Sriram Padmanabhan (IBM), Ramesh Agarwal (IBM). Patent pending.

**Selectivity Estimation in Spatial Database.** Worked on using wavelets in selectivity estimation. Collaborators: my advisor Jeffrey S. Vitter (Duke) and Min Wang (IBM).

**Image Compression Research.** Worked on image compression research using discrete cosine transform (DCT), wavelet transforms and hybrid transforms during my Masters studies (1998-1999) under Philip M. Long.

**Mobile-IP Research** Implemented RFC 2290 *MIPv4 Option for PPP IPCP*. The scope of the project covers issues related to the inter-working of the MIPv4 protocol and the PPP protocol. The MIPv4 option was implemented on Linux's PPP source. This is the first implementation of this RFC in the world.

## Honors and Awards

**Awarded the IBM PhD Fellowship 2003.**

**Outstanding Service to the Department Award.** Duke University Computer Science Dept.

**Dean's List 1995/1996.** National University of Singapore.

**Dean's List 1996/1997.** National University of Singapore.

## Language Ability

- English, Mandarin, German, Cantonese and the Teochew dialect.