

---

**Zhiyuan Chen's Curriculum Vitae**

---

Department of Computer Sciences  
Cornell University  
4110 Upson Hall  
Ithaca, NY 14853

Email: [zhychen@cs.cornell.edu](mailto:zhychen@cs.cornell.edu)  
Office: (607) 255-9124  
Home: (607) 257-9230  
<http://www.cs.cornell.edu/zhychen>

**Research Interests**

Database systems in general. Specific interests include database compression, physical database design, E-Commerce, XML, and data warehousing.

**Education**

**Ph.D.** Cornell University, Computer Science Aug 2002 (expected)  
Dissertation: Compressed Database Systems  
Advisor: Prof. Johannes Gehrke  
GPA: 4.01/4.0

**M.S.** Fudan University, China, Computer Science Aug 1997  
GPA: 4.0/4.0

**B.S.** Fudan University, China, Computer Science Aug 1995  
GPA: 3.9/4.0

**Research Experience**

**Research Assistant, Cornell University** Jan 1998 - present  
Advisor: Prof. Johannes Gehrke and Prof. Praveen Seshadri

For my thesis research, I worked on development of *compressed database systems* that use compression techniques to overcome the exponentially increasing gap between CPU speed and disk access speed. My thesis has the following four contributions: (1) I proposed strategies to select appropriate compression methods; (2) I developed compression-aware query optimization algorithms that choose optimal query execution strategies for queries on compressed data; (3) I proposed a general automatic physical database design approach that designs algorithms for a particular physical design problem based on investigation of the shape of the search space of the problem; (4) I applied this approach to solve the problem of selecting what attributes to compress in compressed database systems given a query workload.

**Research Intern, AT&T Laboratories** Aug 1999 - Sep 2000  
Mentor: Dr. Flip Korn and Dr. Nick Koudas

I proposed techniques that use succinct summary structures to give quick and accurate estimation of the result size of queries on string or hierarchical data (e.g., XML) that is abundant in E-Commerce applications.

**Research Intern, IBM T.J. Watson Research Center**

May 1998 - Aug 1998

Mentor: Dr. Murray Campbell.

I worked on mining long sequence patterns.

**Teaching Experience****Teaching Assistant, Cornell University**

Aug 1997 - Dec 1997

Designed and set up assignments and projects, held office hours, and graded projects for a senior level introductory class on database systems (CS432).

**Teaching Assistant, Fudan University**

Aug 1997 - Dec 1997

Graded homework and held discussion sessions for a Calculus class.

**Referred Conference Papers**

Zhiyuan Chen, Johannes Gehrke, and Flip Korn. Query optimization in compressed database systems. In *Proc. of ACM SIGMOD International Conference on Management of Data (SIGMOD'2001)*, 2001. (Acceptance rate: 15%.)

Zhiyuan Chen, H.V. Jagadish, Flip Korn, Nick Koudas, S. Muthukrishnan, Raymond Ng, and Divesh Srivastava. Counting twig matches in a tree. In *Proc. of the 17th International Conference on Data Engineering (ICDE'2001)*, 2001. (Acceptance rate: 17%.)

Zhiyuan Chen, Flip Korn, Nick Koudas, and S. Muthukrishnan. Selectivity estimation for Boolean queries. In *Proc. of the 19th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS'2000)*, pages 216–225, 2000. Also invited to the *Journal of Computer and System Sciences (JCSS)*. (Acceptance rate: 22%.)

Zhiyuan Chen and Praveen Seshadri. An algebraic compression framework for query results. In *Proc. of the 16th International Conference on Data Engineering (ICDE'2000)*, 2000. Preliminary version: Zhiyuan Chen and Praveen Seshadri. A compression framework for query results. *Computer Science Dept, Cornell University, Technical Report TR98-1720, Sep 1998*. (Acceptance rate: 14%.)

Philippe Bonnet, Kyle Buza, Zhiyuan Chen, Victor Cheng, Randolph Chung, Takako M. Hickey, Ryan Kennedy, Daniel Mahashin, Tobias Mayr, Ivan Oprencak, Praveen Seshadri, and Hubert Siu. The cornell jaguar system: Adding mobility to predator. In *Proc. of ACM SIGMOD International Conference on Management of Data (SIGMOD'1999)*, pages 580–581, 1999.

**Patents**

Zhiyuan Chen, Flip Korn, Nick Koudas, and S. Muthukrishnan: "Multi-dimensional Sub-string Selectivity Estimation". US patent pending.

Zhiyuan Chen, Flip Korn, Nick Koudas, S. Muthukrishnan: "Method and Apparatus for Estimating the Selectivity of Boolean Queries". US patent pending.

---

Zhiyuan Chen, H.V. Jagadish, Flip Korn, Nick Koudas, S., Muthukrishnan, Raymond Ng, and Divesh Srivastava: "Counting Twig Matches in a Tree". US patent pending.

**Professional Activity**

Reviewer for submissions to SIGMOD'2001, the Journal of Computer and System Sciences (JCSS), and the Journal Informatica.

**Visa Status**

Citizen of China, currently on F1 visa.

**References**

References are available upon request from:

Professor Johannes Gehrke  
Department of Computer Sciences  
Cornell University  
4105B Upson Hall  
Ithaca, NY 14853  
Email: johannes@cs.cornell.edu  
Office: 607-255-1045; fax: 607-255-4428

Dr. Praveen Seshadri  
Microsoft Corporation  
3/1155 Microsoft  
One Microsoft Way  
Redmond, WA 98052  
Email: pravse@microsoft.com  
Office:(425) 705-3231

Dr. Divesh Srivastava  
AT&T Laboratories  
Rm A163, 180 Park Ave, Bldg 103  
Florham Park, NJ 07932  
Email: divesh@research.att.com  
Phone: (973) 360-8776

Dr. Flip Korn  
AT&T Laboratories  
Rm A101, 180 Park Ave, Bldg 103  
Florham Park, NJ 07932  
Email: flip@research.att.com  
Phone: (973) 360-8722

Professor Jayavel Shanmugasundaram  
4105A Upson Hall  
Department of Computer Sciences  
Cornell University  
Ithaca, NY 14853  
Email: jai@cs.cornell.edu  
Phone: (607) 255-4117; fax: (607)255-4428