

Ying Xu

Tel: 650-862-6758

xuying@stanford.edu

<http://theory.stanford.edu/~xuying/>

P.O. Box 18329, Stanford, CA 94309

Education

2003.9 - Dept. of Computer Science, Stanford University **Ph.D. candidate**
2000.9 –2003.7 Dept. of Computer Science, Beijing University, China **M.S.**
1996.9 –2000.7 Dept. of Computer Science, Beijing University, China **B.S.**

Research Interests: Algorithm, Database, Web, and Privacy

Publications

(Most papers available at <http://theory.stanford.edu/~xuying/index.html#pub>)

- Random Sampling Algorithms for Efficient Semi-Key Discovery. *Technical Report*. (with R. Motwani)
- Estimating Sum by Weighted Sampling. *Technical Report*. (with Motwani, Panigrahy)
- Evolution of Page Popularity under Random Web Graph Models. *PODS 2006*. (with R. Motwani)
- Estimating corpus size via queries. *CIKM 2006*. (with A. Broder, M. Fontura, V. Josifovski et al)
- Fractional Matching via Balls-and-Bins. *RANDOM 2006*. (with R. Motwani, R. Panigrahy)
- Efficient probabilistic algorithms for detecting pairs of fuzzy duplicates. *Microsoft Technical Report*. (with V. Ganti)
- Two Can Keep A Secret: A Distributed Architecture for Secure Database Services. *CIDR 2005*. (with G. Aggarwal, M. Bawa, P. Ganesan et al.)
- Enabling Privacy for the Paranoids. *VLDB 2004*. (with G. Aggarwal, M. Bawa, P. Ganesan et al)
- An $O(n^{1.5})$ deterministic gossiping algorithm for radio networks, *Algorithmica*, 36, 2003;
- Exact matching of RNA secondary structure patterns. *Theor. Comput. Sci.*, 2005. (with X. Deng, L. Wang)
- SEGID: identifying interesting segments in sequence alignments, *Bioinformatics*, 2003. (with L. Wang)
- A greedy method for inferring tandem duplication history, *Bioinformatics*, 19(12), 2003. (with L. Zhang, B. Ma, L. Wang)
- Haplotype Inference by Parsimony, *Bioinformatics*, 2003. (with L. Wang)

Experience

- **Microsoft Search Labs:** 2006.7-9
Intern. Project: personalized search. Infer personal preference from user query and click history and tailor query results according to user preference.
- **Microsoft Research, Data Mining and Exploration group:** 2004.6-9
Intern. Project: data cleaning. Design and implement efficient randomized algorithms for finding similar records in a database.
- **Teaching Assistant, Stanford:** 2006.9-12
Course: Information Retrieval and Web Mining, by Chris Manning and Prabhakar Raghavan
- **Dept of Computer Science, City University of Hong Kong:** 2002.1-2003.4
Research Assistant. Area: algorithms and bioinformatics. Design and implement algorithms for bioinformatics problems such as tandem duplication history inference, haplotype inference, RNA secondary structure matching etc. Software available on my homepage.
- **Bell Labs China :** 2001.2 - 2001.6 (part-time)
Software Developer. Design and implement an online report scheme for billing systems using html + asp + Oracle/Sybase.

Skills: c++/c#, java, perl, asp, scripts; databases such as SQL server

Honors and Prizes

- 2003-2006 Stanford Graduate Fellowship
- 2001 Outstanding Student of Beijing University (among top 2%, the highest honor)
- 2000 Outstanding Graduate Award of Beijing City (among top 3%, the highest honor)
- 2000 Intel Scholarship for Computer Science and Electronics (top 2%)
- 1999 Lucent Global Science Scholarship (only 5 students in China awarded that year)
- 1998 Outstanding Student of Beijing University (among top 2%, the highest honor)
- 1995 National mathematics competition for high school students, 2nd prize
- 1995 National physics competition for high school students, 2nd prize

Course Work

Mathematics:

Discrete Mathematics, Mathematics Analysis, Advanced Algebra, Stochastic Process, Probability & Statistics

Theoretical Computer Science:

Algorithms & Data Structures, randomized algorithm, Theory of Computation, approximation Algorithms, Cryptography, Game Theory, Petri Nets, Formal Semantics of Language

Systems:

Digital Logic, Computer Architecture, Operating Systems, Compilers, Database, Networks, Artificial Intelligence, Bioinformatics, Computer Graphics, Computational Linguistics

Pass all comprehensive exams required by Stanford CS PhD program in the first quarter, including: Analysis of Algorithms, Automata and Formal Languages, Logic, Computer Architecture, Compilers, Networks, Programming Language, Artificial Intelligence, Database