

Bingjun Sun

Department of Computer Science and Engineering
The Pennsylvania State University
University Park, 16802

Phone: (814)360-2567
Email: bsun@cse.psu.edu
Homepage: <http://www.cse.psu.edu/~bsun/>

Research Interests

- Information Retrieval and Digital Library
- Data Mining and Machine Learning, Text and Graph Mining
- Database and Information Systems
- Cheminformatics and Bioinformatics
- Multi-Agent Systems and Artificial Intelligence

Highlights

- **Best Student Paper Nominee** in WWW07.
- Several papers in top conferences of WWW08, WWW07, SIGIR07, ICDM07, AAMAS06, CIKM06, ICPP04.
- Strong interdisciplinary academic experiences on multiple areas.
- Among top **5%** students in **Tsinghua University**, one of the Best Two Universities in China.

Education

Ph.D. (ABD) in Computer Science and Engineering, Minor in Statistics,

The Pennsylvania State University, University Park, PA, Aug. 2008 (Expected). (Major GPA: 3.97/4.0)
Dissertation: Mining, Indexing, and Search Approaches to Chemical Information Retrieval.
Advisor: C. Lee Giles and Prasenjit Mitra.

MS in Computer Science and Engineering,

The Pennsylvania State University, University Park, PA, Dec. 2003. (Major GPA: 3.97/4.0)
Thesis: Energy-Efficient Scheduling Algorithms of Object Retrieval on Indexed Parallel Broadcast Channel.
Advisor: Ali R. Hurson.

MA in Architecture & Urban Planning,

Tsinghua University, Beijing, China, July 2000. (Ranking: 1/44)

BA in Architecture & Urban Planning,

Tsinghua University, Beijing, China, July 1997. (Graduated with Honors, Ranking: 4/80)

Publications

Refereed Journal Papers under Review or in Progress:

1. Bingjun Sun, Prasenjit Mitra, and C. Lee Giles, "Towards Chemical-Entity-Aware Information Retrieval," submitted to ACM Transactions on Information Systems (TOIS). 2008.
2. Ali R. Hurson, Bingjun Sun, John Hannan, and Y. Jiao, "Power-Aware Scheduling of Data Retrieval from Indexed Parallel Broadcast Channels," submitted to IEEE Transactions on Knowledge and Data Engineering (TKDE). 2008.
3. Xiaocong Fan, Michael McNeese, Bingjun Sun, Timothy Hanratty, Laurel Allender, and John Yen, "Human-Agent Collaboration for Time Stressed Multi-Context Decision Makings", submitted to IEEE Transactions on Systems, Man, and Cybernetics (SMC). 2008.

Conference, Workshop, and other Papers:

4. Bingjun Sun, Qingzhao Tan, Prasenjit Mitra, and C. Lee Giles, "Extraction and Search of Chemical Formulae in Text Documents on the Web," in Proceedings of the International World Wide Web Conference (**WWW07**). 2007. **Best Student Paper Nominee**. [acceptance rate: 14%]
5. Bingjun Sun, Prasenjit Mitra, and C. Lee Giles, "Mining, Indexing, and Searching for Textual Chemical Molecule Information on the Web," in Proceedings of the International World Wide Web Conference (**WWW08**). 2008. [acceptance rate: 12%]

6. Bingjun Sun, Prasenjit Mitra, Hongyuan Zha, C. Lee Giles, and John Yen, "Topic Segmentation with Shared Topic Detection and Alignment of Multiple Documents," in Proceedings of the International ACM SIGIR Conference on Research & Development in Information Retrieval (**SIGIR07**). 2007. [acceptance rate: 17%]
7. Bi Chen, Qiankun Zhao, Bingjun Sun, and Prasenjit Mitra, "Temporal and Social Network Based Blogging Behavior Prediction In BlogSpace," in Proceedings of the International Conference on Data Mining (**ICDM07**). 2007. [acceptance rate: 12%]
8. Xiaocong Fan, Bingjun Sun, Shuang Sun, Michael McNeese, John Yen, "RPDEnabled Agents Teaming with Human for MultiContext Decision Making," in Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems (**AAMAS06**). 2006. [acceptance rate: 17%]
9. Bingjun Sun, Ding Zhou, Hongyuan Zha, and John Yen, "Multi-Task Text Segmentation and Alignment Based on Weighted Mutual Information," in Proceedings of the Conference on Information and Knowledge Management (**CIKM06**). 2006. [acceptance rate: 25%]
10. Xiaocong Fan, Rui Wang, Bingjun Sun, Shuang Sun, and John Yen. "Multi-Agent Information Dependence," in Proceedings of the International Conference Integration of Knowledge Intensive Multi-Agent Systems (**KIMAS05**). 2005.
11. Xiaocong Fan, Shuang Sun, Bingjun Sun, Guruprasad Airy, Michael McNeese, John Yen, "Collaborative RPD-enabled Agents Assisting - The Three-Block Challenge in C²CUT, " in Proceedings of the Conference on Behavior Representation in Modeling and Simulation (**BRIMS05**). 2005.
12. Bingjun Sun, Ali R. Hurson, and John Hannan, "Energy-Efficient Scheduling Algorithms of Object Retrieval on Indexed Parallel Broadcast Channels," in Proceedings of the International Conference on Parallel Processing (**ICPP04**). 2004. [acceptance rate: 34%]
13. K. T. Mueller, P. Mitra, C. L. Giles, B. J. Garrison, J. D. Kubicki, S. L. Brantley, Bingjun Sun, Y. Liu, W. J. Brouwer, S. Nangia, and J. Z. Bandstra. "ChemXSeer: Cyber-tools for researchers in environmental chemistry," Division of Chemical Information for the 235th American Chemical Society National Meeting, 2008.
14. C. L. Giles, P. Mitra, K. Muller, J. Kubicki, B. Garrison, J. Z. Wang, B. Sun, Y. Liu, Q. Tan, L. Bolelli, X. Lu, A. Jaiswal, K. Bai, I. Councill, W. Brouwer, J. Fernandez, and J. Bandstra, "ChemXSeer: An eScience Web Search Engine and Repository for Chemistry," in iConference08. 2008.
15. C. L. Giles, P., K. Mueller, J. Z. Wang, B. Sun, L. Bolelli, X. Lu, Y. Liu, I. Councill, W. Brower, Q. Tan, A. Jaiswal, J. Kubicki, B. Garrison and J. Bandstra, "ChemXSeer: An eChemistry Web Search Engine and Repository," in NSF Symposium on Next Generation of Data Mining and Cyber-Enabled Discovery for Innovation. 2007.
16. Prasenjit Mitra, C. Lee Giles, Bingjun Sun, Ying Liu, "ChemXSeer: A Digital Library and Data Repository for Chemical Kinetics," in CIKM07 Workshop on CyberInfrastructure: Information Management in eScience (**CIMS07**). 2007.
17. Bolelli, L., Lu, X., Liu, Y., Jaiswal, A., Bai, K., Sun, B., Councill, I., Mitra, P., Wang, J.Z., Mueller, K., Kubicki, J., Garrison, B., Bandstra J., Giles, C.L. "ChemXSeer: A Chemistry Web Portal for Scientific Literature and Datasets," Open Repositories Conference, San Antonio, Texas, 2007.
18. Fu Li, Bingjun Sun, "Specialization of Social Issues – Providing Rentable Residence for the Floating Population". (Chinese). World Architecture. 1997 (5).

Conference Papers under Review or in Progress:

19. Bingjun Sun, Prasenjit Mitra, and C. Lee Giles, "Irredundant Informative Subgraph Mining for Graph Information Retrieval," submitted to the International ACM SIGIR Conference on Research & Development in Information Retrieval (**SIGIR08**). 2008.
20. Bingjun Sun, Prasenjit Mitra, and C. Lee Giles, "Machine Learning Approaches to Chemical Entity Mining and Search," submitted to the 23rd AAAI Conference on Artificial Intelligence (**AAAI08**) NECTAR. 2008.

Research and Professional Experiences

Google Inc. *New York Office*, Software Engineer Intern May. 2007-Aug. 2007

- Webpage classification for blog search to identify blogs by training classifiers to filter spam blogs and non-blogs.

Pennsylvania State University. *Intelligent Information Systems Research Lab, College of IST*, Jan. 2006-present

- Graph mining, indexing, and search with the application of cheminformatics.
- Chemical entity tagging, indexing, and search with full document indexing and search, an integral part of an NSF-funded project of an academic digital library, Chem_xSeer, the Chemical Cyber-infrastructure.
- Blog data mining on prediction of blogging behavior using temporal and social network analysis.
- Multi-document topic segmentation and alignment and shared topic detection.
- Automatic scoring project on Biology reports by semi-structured text classification in the Lion Writer project.

Pennsylvania State University. *Intelligent Agents Lab, College of IST*, Jan. 2004-Jan. 2006

- Multi-agent system framework based on the computational collaborative Recognition-Primed Decision model.
- Combat simulation for Army Research Lab. Spatial and temporal limited resource allocation under uncertainty.

Pennsylvania State University. *Global Information Systems Research Group, Dept of CSE*, Aug. 2002- Dec. 2003

- Energy-efficient object retrieval from parallel broadcast channels on mobile database using wireless networking.

Pennsylvania State University. *CAD & Immersive Environments Lab, Dept of Architecture*, Aug. 2000–July 2002

- Human computer interaction, virtual reality, and web-based collaborative virtual design studio.

Tsinghua University, Beijing, China. *School of Architecture* Jan. 1992–Jun. 2000

- Computer graphic applications, 3d modeling, rendering, and animation, AutoLisp programming, and web design.
- Collaborative urban design project between Tsinghua Univ. and MIT.
- Outstanding undergraduate diploma project on urban planning. Statistic analysis and GIS application.
- Project of the New Campus of National College of Fine Art and the New Campus of Shanghai University.
- Project among top 3 in Cities with Soul Sino-British Students' Architecture and Urban Design Competition 1998.
- Project among top 1% in ACSA/OTIS Elevator International Student Design Competition 1996.
- Project won top 1 in Shanghai International Design Competition on Residence 1996.

Teaching Experiences

• Pennsylvania State University. *Department of Computer Science and Engineering* Jan. 2003–May. 2003
Teaching assistant and grader of Logical Design of Digital Systems.

• Pennsylvania State University. *School of Architecture* Aug. 2000–May. 2002
Instructor of Computer Aided Design Labs.
Teaching assistant of Introduction to Architecture, Materials and Building Construction.

• Tsinghua University, Beijing, China. *School of Architecture* Sep. 1998–Jan. 1999
Instructor of the course of Architecture Design.

Professional Services and Activities

- Reviewer: WWW07, SIGIR07, ICDL07, AAAI07, ICML07, ICDM07, WIDM07, SDM08.
- Conference presentation: WWW 2007 and CIKM 2006.
- Talk: IST Graduate Symposium 2008.

Honors

- Best Student Paper Nominee and Student travel award of WWW2007 2007
- Tsinghua University Outstanding Graduate (top 3%) 2000
- Tsinghua University Graduated with Honors (top 5%) 1997
- Tsinghua University Society & Club Activities Awards 1996–1999
- Tsinghua University *Nitianzeng* Fellowship, *12.9* Fellowship, *Guanghua* Fellowships 1995–1999
- Tsinghua University Outstanding Undergraduate 1994–1996
- 1st Class Prize in Physics, 2nd Class Prize in Mathematics, Olympiad of High Schools, Sichuan Province 1991

Technical Skills

Programming Languages: C, C++, Java, Perl, SQL, XML, HTML, VRML, Matlab, Lisp, Prolog.

Platforms: Windows/DOS, Linux, Solaris Unix, MAC.

Applications and Tools: Lucene & Nutch, SAS, Photoshop, Form-Z, 3DMax, AutoCAD, MS Office.

References

References will be provided upon request.