

Xiaoyong Chai

Department of Computer Science, Office 4351
University of Wisconsin
1210 W. Dayton St. WI, 53706

Office: 608-262-5596
xchai@cs.wisc.edu
<http://www.cs.wisc.edu/~xchai>

RESEARCH INTERESTS

- Databases and AI, with an emphasis on information extraction and information integration;
- Unstructured data management and community information management, with an emphasis on leveraging user interaction to improve data quality.

EDUCATION

Ph.D. Student	Computer Science, University of Wisconsin-Madison Advisor: AnHai Doan	2011 (expected)
M.S.	Computer Science, Hong Kong University of Science and Technology Dissertation: Sensor-Based Multiple-Goal Recognition Advisor: Qiang Yang	2005
B.S.	Computer Science, Fudan University , China	2002

RESEARCH AND WORK EXPERIENCE

Research Assistant, University of Wisconsin-Madison *Jan. 2007 – present*

Conducting research on data integration and unstructured data management:

- Studying the problem of efficiently incorporating user feedback into information extraction and integration programs.
- Co-developed MadWiki, a machine-human partnership approach for building community Web portals where community users can provide and edit structured data via a Wiki UI.
- Developed mSeer, a solution for analyzing and revising data integration schemas to improve their amenability to matching.

Research Assistant, Hong Kong University of Science and Technology *Sept. 2003 – July 2005*

Conducted thesis research on probabilistic indoor location estimation and human behavior recognition:

- Applied machine learning techniques to improve the accuracy of indoor location estimation systems and to reduce calibration effort in building such systems.
- Designed probabilistic graphical models for recognizing complex human behaviors based on low-level sensory data.

Intern, Microsoft Research Asia *July 2003 – Sept. 2003*

Designed and implemented a prototype troubleshooting tool for Microsoft Office products to facilitate online technical support.

Undergraduate Research Assistant, Fudan University, China *July 2001 – July 2002*

Implemented a compression-domain face detection module in a real-time video processing system.

PROFESSIONAL SERVICES

External Reviewer for VLDB 2009, VLDB 2008, SIGMOD 2006
Administrator of DBWorld

PUBLICATIONS

Conference Publications

1. **Xiaoyong Chai**, Ba-Quy Vuong, AnHai Doan, and Jeffrey F. Naughton, “Efficiently Incorporating User Feedback into Information Extraction and Integration Programs,” *Proceedings of the ACM Conference on Management of Data (SIGMOD-09)*, 2009.
2. Eric Chu, Akanksha Baid, **Xiaoyong Chai**, AnHai Doan, and Jeffrey F. Naughton, “Combining Keyword Search and Forms for Ad Hoc Querying of Databases,” *Proceedings of the ACM Conference on Management of Data (SIGMOD-09)*, 2009.
3. **Xiaoyong Chai**, Maysam Sayyadian, AnHai Doan, Arnon Rosenthal, and Len Seligman, “Analyzing and Revising Data Integration Schemas to Improve Their Matchability,” *Proceedings of the Int. Conference on Very Large Databases (VLDB-08)*, 2008. 49/296 = 16.5% accepted.
4. Pedro DeRose, **Xiaoyong Chai**, Byron Gao, Warren Shen, AnHai Doan, Philip Bohannon, and Jerry Zhu, “Building Community Wikipedias: A Human-Machine Approach,” *Proceedings of the IEEE Int. Conference on Data Engineering (ICDE-08)*, 2008. 75/617 = 12.1% accepted.
5. Xiaopeng Xiong, Hicham G. Elmongui, **Xiaoyong Chai**, and Walid G. Aref, “Place: A Distributed Spatio-Temporal Data Stream Management System for Moving Objects,” *Proceedings of the Eighth Int. Conference on Mobile Data Management (MDM-07)*, 2007. 28.5% accepted.
6. **Xiaoyong Chai** and Qiang Yang, “Multiple-Goal Recognition From Low-Level Signals,” *Proceedings of the Twentieth National Conference on Artificial Intelligence (AAAI-05)*, 2005. 223/803 = 27.7% accepted.
7. **Xiaoyong Chai** and Qiang Yang, “Reducing Calibration Effort for Location Estimation Using Unlabeled Samples,” *Proceedings of the IEEE Int. Conference on Pervasive Computing and Communications (PERCOM-05)*, 2005. 31/233 = 13.3% accepted.
8. Jie Yin, **Xiaoyong Chai**, and Qiang Yang, “High-Level Goal Recognition in a Wireless LAN,” *Proceedings of the Nineteenth Conference on Artificial Intelligence (AAAI-04)*, 2004. 120/453 = 26.5% accepted.
9. **Xiaoyong Chai**, Ling Deng, Qiang Yang, and Charles Ling, “Test-Cost Sensitive Naïve Bayesian Classification,” *Proceedings of IEEE Int. Conference on Data Mining (ICDM-04)*, 2004. 39/445 = 8.7% accepted.
10. Qingzhao Tan, **Xiaoyong Chai**, Wilfred Ng, and Dik Lun Lee, “Applying Co-training to Clickthrough Data for Search Engine Adaptation,” *Proceedings of the Ninth Int. Conference on Database Systems for Advanced Applications (DASFAA-04)*, 2004. 60/272 = 22% accepted.

Journal Publications

11. AnHai Doan, Philip Bohannon, Raghu Ramakrishnan, **Xiaoyong Chai**, Pedro DeRose, Byron J. Gao, and Warren Shen, “User-Centric Research Challenges in Community Information Management Systems,” *IEEE Data Engineering Bulletin*, Special Issue on Data Management in Social Networks, 2007. Invited paper.

12. AnHai Doan, Jeffrey Naughton, Raghu Ramakrishnan, Akanksha Baid, **Xiaoyong Chai**, Fei Chen, Ting Chen, Eric Chu, Pedro DeRose, Byron Gao, Chaitanya Gokhale, Jiansheng Huang, Warren Shen, and Ba-Quy Vuong, "Information Extraction Challenges in Management Unstructured Data," SIGMOD Record, Special Issue on Managing Information Extraction, 2008.
13. **Xiaoyong Chai** and Qiang Yang, "Reducing the Calibration Effort for Probabilistic Indoor Location Estimation," *IEEE Transactions on Mobile Computing*, 6(6), 2007.
14. Qiang Yang, Charles Ling, **Xiaoyong Chai**, and Rong Pan, "Test-Cost Sensitive Classification on Data with Missing Values," *IEEE Transactions on Knowledge and Data Engineering*, 18(5), 2006.
15. Yiqiang Chen, Qiang Yang, Jie Yin, and **Xiaoyong Chai**, "Power-Efficient Access-Point Selection for Indoor Location Estimation," *IEEE Transactions on Knowledge and Data Engineering*, 18(7), 2006.

Other Publications

16. AnHai Doan, Jeffrey Naughton, Akanksha Baid, **Xiaoyong Chai**, Fei Chen, Ting Chen, Eric Chu, Pedro DeRose, Byron Gao, Chaitanya Gokhale, Jiansheng Huang, Warren Shen, and Ba-Quy Vuong, "The Case for a Structured Approach to Managing Unstructured Data," *Proceedings of the Conference on Innovative Data Systems Research (CIDR-09)*, 2009.
17. Lin Deng, Wilfred Ng, **Xiaoyong Chai**, and Dik Lun Lee, "Spying Out Accurate User Preferences for Search Engine Adaptation," *Proceedings of the Sixth Int. Workshop on Knowledge Discovery from the Web (WebKDD-04)*, 2004.
18. Qiang Yang, Yiqiang Chen, Jie Yin, and **Xiaoyong Chai**, "LEAPS: A Location Estimation and Action Prediction System in a Wireless LAN Environment," *the NPC Workshop on Building Intelligent Sensor Networks (BISON-04)*, 2004.