Bin Tan

602 E. Clark St. Apt. 34
Champaign IL, 61820
bintan@uiuc.edu
http://ews.uiuc.edu/~bintan

EDUCATION

Sept. 2003 - Present	**University of Illinois at Urbana-Champaign** Ph.D. candidate in Computer Science Thesis: *A study of language models for exploiting user feedback in Information Retrieval* Adviser: Prof. ChengXiang Zhai GPA: 4.0/4.0 Illiac Fellowship
Sept. 1999 - June 2003	**Nanjing University**, China B.S. in Computer Science GPA: 3.9/4.0 Robert Mundell Scholarship, People's Scholarship, Honored Student

EXPERIENCES

June 2005 - Present **Research Assistant in Prof. ChengXiang Zhai's Text Information Management Group**

- Developed algorithms to exploit explicit/implicit feedback to provide personalized search and recommendation. Used a language-modeling approach to extract relevant search patterns from personal search logs for user interest modeling.
- Built a user-centered adaptive search system (UCAIR) to provide web search personalization using the feedback algorithms and a decision-theoretic framework. Implemented the system as Internet Explorer search toolbar and local proxy server. The real-time implicit user modeling technology became co-patented with SurfCanyon.com.
- Developed a web-based user interaction module for the Bioinformatics Tutoring System using a Python web application framework (TurboGears) and designed schema to store tutoring data in MySQL.
- Participated in NIST's 2005 TREC HARD track and achieved one of the top results with a term-based relevance feedback algorithm.
- Administered Linux servers in the group. Managed software installations and wrote scripts to maintain various research datasets.

Summer 2008 **Intern at Google**

Developed system tools to transfer and process user data for a Google hosting service (Google Research Datasets).

Summer 2007 **Intern at Yahoo!**

Designed scalable algorithms to segment web search queries into semantic units. The work got

patented by Yahoo!.

Summer 2006 **Intern at Nextumi (now named ShareThis)**

Worked on text mining and information retrieval algorithms for the company's social sharing service. This internship was followed by a one-year research assistantship with the company.

Sept. 2003 - May 2005 **Teaching Assistant for CS423 Operating Systems Design**

Duty included designing/grading assignments/exams, holding office hours and giving guest lectures.

PROGRAMMING SKILLS

- Languages: C++ (STL), Python, Java
- Technologies: HTML, XML, AJAX, SQL
- Systems: Linux, Windows

MAJOR PUBLICATIONS

- **B. Tan**, F. Peng, *Unsupervised Query Segmentation using Generative Language Models and Wikipedia*. In WWW'08. (11% acceptance)
- **B. Tan**, A. Velivelli, H. Fang, C. Zhai, *Term Feedback for Information Retrieval with Language Models*. In SIGIR'07. (18% acceptance)
- **B. Tan**, X. Shen, C. Zhai, *Mining long-term search history to improve search accuracy*. In KDD'06. (23% acceptance)
- X. Shen, **B. Tan**, C. Zhai, *Context-Sensitive Information Retrieval with Implicit Feedback*. In SIGIR'05. (19% acceptance)
- X. Shen, **B. Tan**, C. Zhai, *Implicit User Modeling for Personalized Search*. In CIKM'05. (18% acceptance)