DAQING HE

BSc, MSc, PhD
Assistant Professor
School of Information Sciences,
The University of Pittsburgh,
135 N. Bellefield Avenue, Pittsburgh, PA 15260, USA
Email: dah44@pitt.edu
Tel: +1 412-624-2477 *Fax:* +1 412-648-7001
Homepage: http://www.sis.pitt.edu/~daqing

RESEARCH INTERESTS AND EXPERTISE

Information retrieval (monolingual and multilingual), adaptive information access and user modeling, interactive retrieval interface design, and collaborative information access.

EDUCATION

PhD in Artificial Intelligence	Division of Informatics (formerly Dept of Artificial Intelligence), University of Edinburgh, Edinburgh, UK	12/95 – 7/01
Thesis Title	References to Graphical Objects in Intelligent Multimodal Interfaces	
Award	Overseas Research Studentship, Colin & Ethel Gordon Scholarship	
Supervisors	Dr Graeme Ritchie and Dr John Lee	
MSc in Computer Science	Dept. of Computer Science & Engineering, Beijing University of Aeronautics and Astronautics, Beijing, China	9/92 – 3/95
Dissertation Title	A Study of Three-value Logic and Its Application to the Detection of Functional Hazards in Combinational Circuits.	
Supervisors	Prof. Qinping Zhao and Prof. Huaimin Sun	
BSc in Computer Science	Dept. of Computer Science & Engineering, Beijing University of Aeronautics and Astronautics, Beijing, China	9/88 – 7/92
Dissertation Title	A Design of a Prolog Interpreter and its C Implementation	
Supervisor	Prof. Huaimin Sun	

PROFESSIONAL EMPLOYMENT SUMMARY

Assistant Professor	School of Information Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania	9/04 – present
Research Scientist	Institute for Advanced Computer Studies, The University of Maryland, College Park, Maryland	2/02 – 8/04

Research Fellow	School of Computing, Robert Gordon University, Aberdeen, United Kingdom	7/99 – 2/02

HONORS AND AWARDS

University of Pittsburgh Provost's Faculty Diversity Seminar (2005)
Provided by University of Pittsburgh

Overseas Research Studentship (1995-1999)
Provided by the UK government (highly competitive among all foreign students in UK).

Colin & Ethel Gordon Scholarship (1995-1999)
Provided by the University of Edinburgh (highly competitive among candidates applying to the university).

The second place award of FengRu Cup student research competition (1995)
Organized by Beijing University of Aeronautics and Astronautics (highly competitive among students in the university).

PUBLICATIONS

Papers in Refereed Journals

1. He, Daqing, Zhen Yue, Jon Walker and Yiling Lin. "Collaborative Search in E-Discovery: An Initial Study." Submitted to Information Processing and Management. Under review from May, 2009.
2. He, Daqing, Ming Mao, Yefei Peng. "Support Information Access in E-Learning by Integrating Digital Libraries and Ontology." Submitted to Online Information Review. Under review from May, 2009.
3. Qiang, Pu, Daqing He. "Semantic Clustering in Relevance Based Language Models." Submitted to ACM Transactions on Information Systems. Under review from June, 2009.
4. Ahn, Jae-wook, Peter Brusilovsky, Jonathan Grady, Daqing He. "Semantic Annotation Based Exploratory Search for Information Analysts." Information Processing and Management. Accepted with minor revision.
5. He, Daqing, Dan Wu. "Enhancing Query Translation with Relevance Feedback in Translingual Information Retrieval." Information Processing and Management. Accepted with minor revision.
6. Wu, Dan, Daqing He. "Signal Boosting for Robust Data Fusion in Speech Retrieval." International Journal of Innovative Computing, Information and Control. To be published in March, 2010.
7. Wu, Dan, Daqing He, Huilin Wang. "Cross-Language Query Expansion Using Pseudo Relevance Feedback." Journal of the Chinese Society for Scientific and Technical Information. (2009)
8. He, Daqing, Graeme Ritchie, John Lee. "Reference to Graphics Objects in Intelligent Multimodal Systems." Knowledge Base Systems 21.7 (2008):617-628
9. Wu, Dan, Daqing He, Huilin Wang, Chongde Shi, Chengzhi Zhang. "Does Query Length Matter: A Comparison of Query Expansion Methods in English-Chinese Cross-Language Information Retrieval." Journal of Computational Information Systems, 4.3 (2008):763-770.

10. He, Daqing, Peter Brusilovsky, Jaewook Ahn, Jonathan Grady, Rosta Farzan, Yefei Peng, Yiming Yang, Monica Rogati. "An Evaluation of Adaptive Filtering in the Context of Realistic Task-Based Information Exploration." Information Processing and Management 44.2 (2008):511-533.
11. Oard, Douglas W., Daqing He, Jianqiang Wang. "User Assisted Query Translation for Interactive Cross-Language Information Retrieval." Information Processing and Management 44.1 (2008): 181-211.
12. He, Daqing, Douglas W. Oard, Jianqiang Wang, Jun Luo, Dina Demner-Fushman, Kareem Darwish, Philip Resnik, Sanjeev Khudanpur, Michael Nossal and Anton Leuski. "Making MIRACLEs: Interactive Translingual Search for Cebuano and Hindi." ACM Transactions of Asian Language Information Processing 2.3 (2003): 219-244.
13. He, Daqing, Ayse Goker, David Harper. "Combining Evidence for Automatic Web Session Identification." Information Processing and Management 38.5 (2002): 727-742.
14. He, Huacan, Yonghuai Liu, Daqing He, and Hua Cheng. "Generalized Logic in Experience Thinking". Science in China (series E) Vol. 39, No. 3 (1996) 226-234.
15. He, Daqing, Qinping Zhao. "Three-Value Logic Equations and the Detection of Functional Hazards in Combinational Switching Circuits." Journal of Beijing University of Aeronautics & Astronautics. 21.3 (1995): 38-44.

Books

16. Wu, Dan, Huilin Wang and Daqing He. Cross-Language Information Retrieval: Translation Enhancement Theory and Technology. Scientific and Technical Documents Publishing House, Beijing, 2009.

Chapters in Edited Books

17. He, Daqing, Peter Brusilovsky, Jonathan Grady, Jaewook Ahn, Yiming Yang, Monica Rogati . "EDIE: An Evaluation Dataset for Task-Based Information Exploration." Achievements in DARPA GALE Program. Accepted.
18. Oard, Douglas W., Judith Klavans, Dagobert Soergel, Pengyi Zhang, Peter Brusilovsky, Daqing He, Tomasz Loboda, Leiming Qian. "Formative Evaluation for Multilingual Multimedia Search and Sense-Making." Achievements in DARPA GALE Program. Accepted.
19. He, Daqing, Jianqiang Wang. "Cross-Language Information Retrieval." Information Retrieval: Searching in the 21st Century. Eds. Ayse Goker, John Davies. John Wiley and Sons, 2009.
20. He, Daqing, Hong Xu. "Chapter 8: User Studies: Heting Chu & Yin Zhang." Research Fronts in Library and Information Science in the West.. Beijing: Renmin University Press (Series on Research Fronts in the Humanities and Social Sciences in the West), 2007. 210-242.
21. He, Daqing. "Integrating Diversity and Multicultural Education into a Digital Library Course." Diversity Across the Curriculum: A Guide for Faculty in Higher Education. Eds. J. Branche, J.W. Mullennix and E.R. Cohn. Anker Publishing, 2007. 298-302.

Papers in Refereed Conferences

22. Qiang, Pu, Daqing He. "Pseudo Relevance Feedback using Semantic Clustering in Relevance Language Model." ACM 18th Conference on Information and Knowledge Management (CIKM), 2009.

23. Li, Qi, Ming Mao, Daqing He. "A Study of Relation Annotation in Business Environments Using Web Mining." IEEE International Conference on Semantic Computing, 2009.

24. Qiang, Pu, Daqing He, Qi Li. "Query Expansion for Effective Geographic Information Retrieval." Cross-Language Evaluation Forum (CLEF). 2008. Springer.

25. He, Daqing, Dan Wu. "Translation Enhancement: A New Relevance Feedback Method for Cross-Language Information Retrieval." ACM 17th Conference on Information and Knowledge Management. (CIKM), 2008. 585-594.

26. He, Daqing, Dan Wu. "Toward a Robust Data Fusion for Document Retrieval." IEEE International Conference on Natural Language Processing and Knowledge Engineering (IEEE NLP-KE), 2008. 338-345. (*Winner of "Excellent Paper Award")*

27. Wu, Dan, Daqing He, Heng Ji, & Ralph Grishman. "The Effects of High Quality Translations of Named Entities in Cross-Language Information Exploration." 2008 IEEE International Conference on Natural Language Processing and Knowledge Engineering (IEEE NLP-KE), 2008. 443-450.

28. Wu, Dan, Daqing He. "Ice-Tea: An Interactive Cross-Language Search Engine with Translation Enhancement." Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2008.

29. Ahn, Jae-wook, Peter Brusilovsky, Daqing He, Jonathan Grady, Qi Li. "Personalized Web Exploration with Task Models." World Wide Web Conference, 2008.

30. He, Daqing, Peter Brusilovsky, Jonathan Grady, Qi Li, and Jae-wook Ahn. "How Up-to-date Should It Be: the Value of Instant Profiling and Adaptation in Information Filtering?" IEEE/WIC/ACM International Conference on Web Intelligence, 2007.

31. Yang, Yiming, Abhimanyu Lad, Ni Lao, Abhay Harpale, Bryan Kisiel, Monica Rogati, Peter Brusilovksy, He, Daqing. "Utility-based Information Distillation Over Temporally Sequenced Documents." Annual International ACM SIGIR Conference on Research & Development on Information Retrieval (SIGIR), 2007.

32. Ahn, Jae-wook, Peter Brusilovsky, Jonathan Grady, Daqing He, Sue Yeon Syn. "Open User Profiles for Adaptive News Systems: Help or Harm?" The International World Wide Web Conference, 2007.

33. Peng, Yefei, Daqing He, Ming Mao. "Geographic Named Entity Disambiguation with Automatic Profile Generation." IEEE/WIC/ACM International Conference on Web Intelligence, 2006.

34. Peng, Yefei, Daqing He. "Direct Comparison of Commercial and Academic Retrieval Systems: An Initial Study." ACM Conference for Information and Knowledge Management (CIKM), 2006.

35. He, Daqing, Ming Mao, and Yefei Peng. "DiLight: a Digital Library Based E-Learning Environment for Learning Digital Libraries." World Conference on E-Learning in Corporate, Government, Healthcare, & Higher Education, 2006.

36. Mao, Ming, Yefei Peng, Daqing He. "DiLight: An Ontology-Based Information Access System for e-Learning Environments." Annual International ACM SIGIR Conference on Research & Development on Information Retrieval (SIGIR), 2006.

37. He, Daqing, Yefei Peng. "Comparing Two Blind Relevance Feedback Techniques." Annual International ACM SIGIR Conference on Research & Development on Information Retrieval (SIGIR), 2006.

38. He, Daqing. "A Study of Self-Organizing Maps in Interactive Relevance Feedback." IEEE International Conference on Information Technology: New Generations (ITNG). Las Vegas, Nevada: April, 2006.

39. Oard, Douglas W., David Doermann, Bonnie Dorr, Daqing He, Philip Resnik, Amy Weinberg, William Byrne, Sanjeev Khudanpur David Yarowsky, Anton Leuski, Philipp Koehn, and Kevin Knight. "Desperately Seeking Cebuano." Human Language Technology-North American Chapter for Association of Computational Linguistics Conference. Edmonton, Canada: 2003.

40. He, Daqing, Jianqiang Wang, Douglas W. Oard, Michael Nossal. "User-Assisted Query Translation for Interactive CLIR." Annual International ACM SIGIR Conference on Research & Development on Information Retrieval (SIGIR). Toronto, Canada: 2003.

41. Goker, Ayse, Daqing He. "Personalization via Collaboration in Web Retrieval System." 76th Annual Conference for American Society for Information Science and Technology (ASIST), 2003.

42. Goker, Ayse, Daqing He. "Analysing Web Search Logs to Determine Session Boundaries for User-Oriented Learning." Adaptive Hypermedia and Adaptive Web-Based Systems, International Conference, Italy: Trento. Eds. Peter Brusilovsky, Oliviero Stock, Carlo Strapparava. LNCS 1892, Springer. Berlin, Germany: 2000. 319-322.

43. He, Daqing, Ayse Goker. "Detecting Session Boundaries from Web User Logs." The 22nd Annual Colloquium for British Computer Society Information Retrieval Specialist Group on Information Retrieval Research. Cambridge, UK: 2000. 57-66,

44. He, Daqing, Graeme Ritchie, John Lee. "Resolving References to Graphical Objects in Multimodal Queries by Constraint Satisfaction." Advances in Multimodal Interfaces, the 3rd International Conference. China: Beijing. Eds. Tieniu Tan, Yuanchun Shi, Wen Gao. LNCS, 1948. Springer. Berlin, Germany: 2000. 8-15.

45. Liu, Boyang, Daqing He, Qinping Zhao. "An Improvement on Self-Organising Feature Map Algorithm." International Conference for Young Computer Scientist. Beijing, China: 1995.

Papers in non-refereed International Conferences

46. Zhen Yue, Jon Walker, Yilin Lin, Daqing He. "Pitt@TREC08: An Initial Study of Collaborative Information Behavior in E-Discovery." Text REtrieval Conference (TREC), 2008.

47. Qiang Pu, Daqing He, Qi Li. "University of Pittsburgh at Geo (CLEF) 2008: Towards Effective Geographic Information Retrieval." Cross-Language Evaluation Forum (CLEF), 2008.

48. He, Daqing, Yefei Peng. "Pitt at TREC 2006: Identifying Experts via Email Discussions." Text REtrieval Conference (TREC), 2006.

49. He, Daqing, Jae-wook Ahn. "Pitt at TREC 2005: HARD and Enterprise." Text REtrieval Conference (TREC), 2005. 896-908.

50. He, Daqing, Jae-wook Ahn. "Pitt at CLEF05: Data Fusion for Spoken Document Retrieval." Cross-Language Evaluation Forum (CLEF) , 2005.

51. He, Daqing, Jianqiang Wang, Jun Luo, Douglas W. Oard. "iCLEF 2004 at Maryland: Summarization Design for Interactive Cross-Language Question Answering." Cross-Language Evaluation Forum (CLEF) , 2004.
52. He, Daqing, Dina Demner-Fushman, Damianos Karakos, Douglas W. Oard, Sanjeev Khudanpur. "Improving Passage Retrieval via Interactive Elicitation and Language Modeling." Text Retrieval Conference (TREC), 2004.
53. He, Daqing, Dina Demner-Fushman. "HARD Experiment at Maryland: From Need Negotiation to Automated HARD Process." Text REtreival Conference (TREC), 2003.
54. Dorr, Bonnie J., Daqing He, Jun Luo, Douglas W. Oard, Richard Schwartz, Jianqiang Wang, and David Zajic. "iCLEF 2003 at Maryland: Translation Selection and Document Selection." Cross-Language Evaluation Forum (CLEF). Trondheim, Norway: 2003.
55. He, Daqing, Hyuk Ro Park, G. Craig Murray, Michael Subotin, and Douglas W. Oard. "TDT2002 Topic Tracking at University of Maryland." Workshop of Topic Detection and Tracking (TDT). Nov. 2002.
56. He, Daqing, Jianqiang Wang, Douglas W. Oard, Michael Nossal. "Comparing User-assisted and Automatic Query Translation." Cross Language Evaluation Forum (CLEF). Sept. 2002.

Papers in Refereed Workshops

57. Yue, Zhen, Abhay Harpale, Daqing He, Jonathan Grady, Yiling Lin, Jon Walker, Siddharth Gopal, Yiming Yang. "CiteEval for Evaluating Personalized Social Web Search." The Future of Information Retrieval Evaluation: A Workshop at ACM SIGIR, 2009.
58. Wu, Dan, Daqing He, Heng Ji, & Ralph Grishman. "A Study of Using an Out-Of-Box Commercial MT System for Query Translation in CLIR." 2nd International ACM Workshop on Improving Non-English Web Searching (iNEWS08): A Workshop at ACM CIKM, 2008.
59. Jiang, Tingting, Daqing He. "Redefining Social Network Services: A Solution to Personal Information and Knowledge Management." Social Media Analysis: A Workshop at International Conference on Web Intelligence, 2007.
60. Mao, Ming, Yefei Peng, Daqing He. "Ontology-based Content Management and Access Framework for Supporting E-Learning Systems." 2nd International Workshop on Web Support Systems: A Workshop at Web Intelligence, 2006.
61. He, Daqing, Douglas W. Oard Lynne, Plettenberg. "Studying the Use of Interactive Multilingual Information Retrieval." New Directions of Multilingual Information Access: A Workshop at SIGIR, 2006.
62. Han, Shen, Ayse Goker, Daqing He. "Web User Search Pattern Analysis for Modeling Query Topic Changes." User Modeling for Context-Aware Applications: A Workshop at 8th International Conference on User Modeling, July 2001.
63. He, Daqing, Graeme Ritchie, and John Lee. "Disambiguation Between Visual Display and Represented Domain in Multimodal Interfaces." Combining AI & Graphics for the Interface of the Future: A Workshop at European Conference on Artificial Intelligence. Brighton, New York: 1998.

64. He, Daqing, Graeme Ritchie, and John Lee. "Referring to Displays in Multimodal Interfaces." Referring Phenomena in a Multimedia Context and Their Computational Treatment: A Workshop at EACL/ACL Madrid, Spain: 1997.

Papers in non-published format

65. He, Daqing. "References to Graphical Objects in Intelligent Multimodal Interfaces." PhD thesis. Division of Informatics, University of Edinburgh, 2000.
66. He, Daqing. "A Study of Three-value Logic and Its Application to the Detection of Functional Hazards in Combinational Circuits." Master's of Science Dissertation. Department of Computer Science and Engineering, Beijing University of Aeronautics and Astronautics. Beijing, China: 1995.

INVITED TALKS AND PRESENTATIONS

1. Towards Intelligent Information Access. Palo Alto: SAP, July 2008.
2. Towards Useful Interactive Multilingual Information Access Systems. Beijing: Department of Information Management, Beijing University, Jan. 2007.
3. Cross Language Information Access: Problems, Techniques, and Future. Xian, China: School of Software Engineering, Xian University of Science and Technology, Dec. 2006.
4. Cross Language Information Access: Problems, Techniques, and Future. Xian, China: School of Computer Science, Northwestern Polytechnic University, Dec. 2006.
5. Cluster-Based Query Expansion and Passage Extent Determination. Research talk: DARPA Tides PI Meeting, 2004.
6. From Need Negotiation to Automated HARD Process. Team site visit: TIDES project Maryland, Oct. 2003.
7. Making MIRACLEs: Interactive Translingual Search for Cebuano and Hindi. Stanford University: CLIR Minisummit, Center for the Study of Language and Information, Oct. 2003.
8. Interactive Cross Language Information Retrieval. University of Maryland: Research Review Day, May 2003.
9. TDT2002 Topic Tracking at University of Maryland. Gaithersburg, Maryland: Topic Detection and Tracking TDT2002, Nov. 2002.
10. TDT2002 Dry-run at University of Maryland. Los Angles, California: TDT 2002 Dry-run, April 2002.
11. Comparing User-Assisted and Automatic Query Translation. Santa Monica, California: Information Science Institute Seminar, University of Southern California, April 2002.
12. Personalization in a Role-Based Web Retrieval System. College Park, Maryland: Linguistic Consortium of Cross Language and Information Processing Lab, University of Maryland, Oct. 2001.
13. Personalization in Web Retrieval. Aberdeen, Scotland: Burns Retreat, School of Computing, Robert Gordon University, June 2001.
14. References Ambiguities in Intelligent Multimodal Interfaces. Aberdeen, Scotland: School of Computing, Robert Gordon University, Feb. 1999.

AWARDED GRANTS

Principal Investigator

2007-2010 "Collaborative Research: User Centric, Adaptive and Collaborative Information Filtering." NSF III-COR. PI at University of Pittsburgh: Collaborative project with Yiming Yang at Carnegie Mellon University. Award Amount: $829,940.

2005-2006 "Building a Digital Library for Learning Digital Libraries." Chancellor's Innovation of Teaching Award. PI, University of Pittsburgh. Award amount: $20,000.

2005-2006 "Combining Multiple Resources for Improving Enterprise Retrieval: An Initial Study," PI. School of Information Sciences' Dean's Entrepreneurial Award. Award amount: $15,000.

2005 "Improving Digital Library Teaching with Diversity Issues." PI, University of Pittsburgh Diversity Seminar Award. Award amount: $1,500.

Co- Investigator

2005-2010 "Rosetta: An Analyst Co-Pilot" Co-PI of Pittsburgh team (PI of Pittsburgh Team: P. Brusilovsky). Award amount: $993,896.

TEACHING EXPERIENCE

Assistant Professor at School of Information Sciences, University of Pittsburgh

Doctoral and Research Seminars

LIS3600 Web Information Systems

Master Level Core Courses

LIS2002 Retrieving Information

LIS2600 Introduction to Information Technology

LIS2670 Digital Libraries

INFSCI2140 Information Storage and Retrieval

Other Courses

LIS2680 Database Design and Applications,

Adjunct Faculty at College of Information Studies, University of Maryland at College Park

Master Level Core Courses

LIS690 Information Technology

STUDENT ADVISING

Current PhD Advisees

Library and Information Science Program:

Jiepu Jiang, Jongdo Park, Zhen Yue

Information Science and Technology Program:

Qi Li, Yilin Lin

Graduated PhD Advisees

Library and Information Science Program:

Demetrios Ioannides

Current and previous sponsored PhD Students

Library and Information Science Program:

Zhen Yue, Jongdo Park, Minsoo Park

Information Science and Technology Program:
Qi Li, Ming Mao, Yefei Peng, Jae-wook Ahn

International Visiting PhD Advisees

Information Science Area:
Dan Wu (Beijing University, China, one year), Peng Qu (Beijing University, China, one year)

Computer Science Area:
Qiang Pu (Science and Technology University of China, China, two years), Rui Liu (Beijing University of Aeronautics and Astronautics, China, one year),

PhD Dissertation Committee Member

Library and Information Science Program:
Minsoo Park, Mary Jo Dorsey, Barbara Zaborowski

Information Science and Technology Program:
Jae-wook Ahn, Ming Mao, Worasit Choochaiiwattana

Intelligent System Program
Rosta Farzan, Behrang Mohit

PhD Comprehensive Committee Member

Library and Information Science Program:
Minsoo Park, Mary Jo Dorsey, Barbara Zaborowski

Information Science and Technology Program:
Jae-wook Ahn, Ming Mao, Worasit Choochaiiwattana, Jon Walker

Intelligent System Program
Rosta Farzan, Behrang Mohit

SERVICES

University, School and Program Committees

- University Senate Election Committee (2007-current)
- School of Information Science (SIS) Council member (2009-current)
- SIS i-fest student research poster competition co-organizer (2006,2007, 2009)
- SIS international student organization faculty advisor (2006-current)
- SIS Faculty and Postdoc search committee (2006, 2009)
- PhD committee member of Library and information Science Program at SIS, Information Science and Technology Program at SIS, and Intelligent System Program at the College of Arts and Sciences.
- Committees for Catherine Ofiesh Orner Award and Robert R. Korfhage Award

Member of the Program Committees

- European Conference on Information Retrieval (ECIR) Program Committee: 2008, 2009
- Conference on Information and Knowledge Management (CIKM) Program Committee: 2008, 2009

- Asian Information Retrieval Symposium (AIRS) Program Committee: 2009
- ACL-HLT General Program committee, Poster and Demo program committee: 2008
- ACM SIGIR Annual Conference, General Program committee, Poster and Demo Program Committee: 2006, 2007, 2008
- American Society for Information Science and Technology, SIG International Information Issues Paper Contest: 2009
- International Conference on Knowledge Management and Information Systems (KMIS): 2009
- Web Information Retrieval Support Systems Workshop (WIRSS): 2008, 2009
- Natural Language Processing and Ontology Engineering (NLPOE): 2008, 2009
- "Evaluation of Exploratory Search Systems," A workshop of SIGIR 2006
- "Web Search Technologies," International Conference on Information Technology: 2005, 2006
- "User Modeling, Machine Learning and Information Retrieval", a workshop of User Modeling 2003.
- The Student Section of COLING-ACL98.

Invited Referee for journals

- *Journal for Information Science*
- *The International Information and Library Review*
- *ACM Computer Surveys*
- *Information Processing and Management*
- *Journal of the American Society for Information Science and Technology*
- *ACM Transactions on Asian Language and Information Processing*
- *User Modeling and User-Adapted Interaction*

Invited External Referee for Research Proposals

- *Collaborative Incentive Research Grant (CIRG) Program, City University of New York*
- *The Innovation and Technology Commission, a funding agency for applied research of the Government of the Hong Kong Special Administrative Region, China*

Member of Professional Societies

- ACM SIGIR
- Association of Computational Linguistics
- Association of Computing Machinery, Special Interest Group on Information Retrieval
- American Society for Information Science and Technology
- IEEE Computer Society
- British Computer Society

PROFESSIONAL ACHIEVEMENTS

Assistant Professor (9/04 – present)
School of Information Sciences, University of Pittsburgh (Pitt), Pittsburgh, Pennsylvania

❖ PI for Pitt team: NSF Collaborative Project "User-Centered Adaptive and Collaborative Filtering" with Professor Yiming Yang at Carnegie Mellon University (2007-2010).

- This project's goal is to significantly improve Adaptive Filtering (AF) technologies by addressing several fundamental weaknesses in current systems that limit their usability in real-world settings.

AF collaboration is our major emphasis. Current AF research focuses on incremental learning of topic models, but does not take into account the information needs of different users and cannot leverage parallel, multi-user relevance feedback. Although information sharing among users has been a focus of research in the field of collaborative filtering (CF), applying these research findings to AF is far from straightforward because CF algorithms are primarily designed for batch learning with large volumes of training examples. AF problems require incremental learning with extremely sparse initial training examples. This project bridges the technical gap between CF and AF, and focuses on the development of new methods that can be learned "on the fly" with minimum relevance feedback from each user and that can "borrow" information effectively from other similar users in constructing the profile for a specific user.

- Co-PI for Pitt team: DARPA Research Project, "Rossetta: An Analyst Co-Pilot" (2005-2010).
  - This project's goal is in developing and demonstrating an integrated set of techniques for building an adaptive, robust and distilled information access module for the Distillation Engine. The module will be able to handle electronic newswire, transcriptions from broadcast news, telephone conversations, talk shows, and web documents from newsgroups and weblogs -- no matter whether they are originally in English or in Chinese, Arabic or unexpected languages.

- PI for the University Of Pittsburgh's "Chancellor's Innovation of Teaching Award": "Building a Digital Library for Learning Digital Libraries." (2005-2006)
  - The goal of this project is to provide students with an interactive, integrated (and active) learning environment for studying digital library topics. My unique approach is that the learning environment is based on an open source digital library system. The focus of the research is on how to design effective semantic based information access methods to enable students to explore the course materials based on keywords, timeline, ontological relationships, and topical relevance. The proposed system can improve the quality of the student's learning experience and increase the effectiveness of that experience.

- PI for School of Information Sciences, "Dean's Entrepreneurial Award:" "Combining Multiple Resources for Improving Enterprise Retrieval: An Initial Study." (2005-2006)
  - The goal of this project is to explore multiple evidence identification and combination tools to improve search effectiveness. We use both CLEF and TREC experiments as the realistic evaluation framework to test the ideas and systems developed for evidence identification and combination.

Research Scientist (Feb. 2002 – Aug. 2004)
Institute for Advanced Computer Studies, The University of Maryland (UMD), College Park, Maryland

- Team leader of UMD's participation in the High Accuracy Retrieval of Documents (HARD) track of Text REtrieval Conference (TREC), 2004. TREC is organized by the National Institute of Standards and Technology (NIST), and is the most prestigious gathering of the best scientists and teams in the world that are working on problems/techniques in Information Retrieval (IR).
  - Responsible for the design of the overall approach taken by the UMD HARD team.
  - Designed and implemented several key components of the system, including the automatic query expansion module, the document re-ranking module, and the ranked-list merging module.
  - Led the effort in analyzing results, and writing reports about the effort of the UMD HARD team.
  - Participated in the TREC Conference to present the work of the team.

- Team leader for the CLIR team in UMD's participation in the interactive track of the Cross Language Evaluation Forum (CLEF), 2004. CLEF is an European conference equivalent to TREC and is a specialized gathering for CLIR researchers.
  - Responsible for the overall design of the experiments run by the UMD team and conducted major parts of the experiments.
  - Responsible for the design and extension of the user-assisted query translation approach for interactive Cross Language Information Retrieval (CLIR). This was the main theme of UMD's participation in iCLEF in 2000 and 2001.
- Team leader of UMD's participation in the High Accuracy Retrieval of Documents (HARD) track of Text REtrieval Conference (TREC), 2003.
  - Responsible for the design of the overall approach taken by the UMD HARD team.
  - Designed and implemented several key components of the system, including the automatic query expansion module, the document re-ranking module, and the ranked-list merging module.
  - Led effort in analyzing results, and wrote the report about the efforts of the UMD HARD team.
  - Participated in the TREC conference to present the work of the team.
- Team leader of the Cross Language Information Retrieval (CLIR) team in UMD's participation in the surprise language experiment and dry-run of Translingual Information Detection, Extraction, and Summarization (TIDES) project, which is a research project funded by the Defense Advanced Research Projects Agency (DARPA).
  - Responsible for the overall design of a rapidly retarget able CLIR interactive system (called MIRACLE) for processing previously-unknown languages.
  - Led the team to complete the implementation of the MIRACLE system for Hindi language within a month. MIRACLE was the first interactive CLIR system for Hindi in the world.
  - Led the team to complete the design and implementation of the batch CLIR systems for Hindi language within a month and for Cebuano language within 10 days. Initial test results showed that our batch CLIR systems achieved comparable results to monolingual Hindi retrieval.
  - Wrote a journal paper presenting the UMD CLIR team's effort in surprise language experiments. Listed as the first author of the paper.
- Team leader of the CLIR team in UMD's participation in the interactive track of the Cross Language Evaluation Forum (iCLEF) in 2003 and 2002.
  - Responsible for the overall design of the experiments run by the UMD team, and conducted major parts of the experiments.
  - Responsible for the design and extension of the user-assisted query translation approach for interactive Cross Language Information Retrieval (CLIR), which was the main theme of UMD's participation in iCLEF in 2000 and 2001.
  - Led the effort in analyzing results; wrote reports and presentations about UMD's iCLEF effort for those two years. Listed as the first author of the reports.
- Team leader of the UMD team's participation in the Topic Detection and Tracking (TDT) track of TREC, 2002.
  - Responsible for the overall design and implementation of a topic tracking system using language modeling and information retrieval techniques. The system achieved comparable tracking results to the state of art in TDT's 2002 experiments.

- Led effort in analyzing results, writing reports, and presenting at the TDT workshop of the TREC conference. Listed as the first author of the report.

Research Fellow (July, 1999 – Feb., 2002)
School of Computing, The Robert Gordon University, Aberdeen, UK.

- Responsible for the design of a Web user context learning model that uses both the user's personal search history and collaborative information from people who share the same interests to improve search effectiveness. Responsible for the implementation of the key context learning components in a Web search engine. Published several conference papers on this topic.
- Researched, surveyed, and compared theories and models about context, context information of users, and context models in communications, linguistic processes, and information retrieval. Wrote a white paper about the outcomes of the research, and, based on the report, extended a context learning model in traditional information retrieval to a context learning model in Web information retrieval.
- Responsible for the design of a role-based context model for modeling the user in information retrieval. The role model extends the traditional context model by viewing the user's actions from a task-oriented view.
- Conducted text mining in several Web search log collections. Identified the problem of lacking sensible session boundaries in Web logs, which is a problem for context learning in Web Information Retrieval.
- Designed and implemented time interval based session identification method. Conducted experiments to test the method on several Web log collections. The results show that the method can provide reasonable boundaries for our context learning purpose. Wrote several conference papers about this work.
- Designed and implemented an evidence combination method, based on the Dempster-Shafer theory, for Web search session identification. The method uses as input the outcomes of time interval and search pattern methods for Web search session identification as input. It generates statistical-based inferences about the possibility of having session delimiters at a given point. Conducted an experiment and demonstrated the effectiveness of the method. I am the first author of a journal paper published in one of the most prestigious journals of IR research.