

WENYI HUANG

CONTACT INFORMATION

College of Information Science & Technology
313D IST Building
The Pennsylvania State University
University Park, PA 16803

Phone: (814) 308-2156
Email: harrywy@gmail.com
<http://www.personal.psu.edu/wzh112/>

RESEARCH INTEREST

Information Retrieval, Recommendation System, Text Mining, Natural Language Processing

EDUCATION

The Pennsylvania State University
Ph.D. Candidate

August 2011 - Present
University Park, PA

- College of Information Science & Technology
- Advisor: Dr. C. Lee Giles

Tsinghua University
Bachelor of Engineering

August 2007 - July 2011
Beijing, China

- Department of Computer Science & Technology

PUBLICATIONS AND PATENTS

- **Wenyi Huang**, Zhaohui Wu, Chen Liang, Prasenjit Mitra and C. Lee Giles. A Neural Probabilistic Model for Context Based Citation Recommendation. To appear in the *Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI 2015)*.
- **Wenyi Huang**, Ingmar Weber, Sarah Vieweg. Inferring Nationalities of Twitter Users and Studying Inter-National Linking. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media (Hypertext 2014)*.
- **Wenyi Huang**, Zhaohui Wu, Prasenjit Mitra and C. Lee Giles. RefSeer: A Citation Recommendation System. In *Proceedings of ACM/IEEE Joint Conference on Digital Libraries 2014 (DL 2014)*.
- Zhaohui Wu, **Wenyi Huang**, Chen Liang and C. Lee Giles. Crowd-sourcing Web Knowledge for Metadata Extraction. In *Proceedings of ACM/IEEE Joint Conference on Digital Libraries 2014 (DL 2014)*.
- Zhaohui Wu, Jian Wu, Madian Khabza, Kyle Williams, Hung-Hsuan Chen, **Wenyi Huang**, Supawong Tuarob, Sagnik Ray Choudhury, Alexander Ororbia, Prasenjit Mitra and C. Lee Giles. Towards Building a Scholarly Big Data Platform: Challenges, Lessons and Opportunities. In *Proceedings of ACM/IEEE Joint Conference on Digital Libraries 2014 (DL 2014)*.
- Lior Rokach Prasenjit Mitra, Saurabh Kataria, **Wenyi Huang**, C. Lee Giles. A Supervised Learning Method for Context-Aware Citation Recommendation in a Large Corpus. *The 10th Workshop on Large-Scale Distributed Systems for Information Retrieval (LSDS-IR 2013)*.
- **Wenyi Huang**, Saurabh Kataria, Cornelia Caragea, Prasenjit Mitra, C. Lee Giles, Lior Rokach. Recommending Citations: Translating Papers into References. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM 2012)*.

- Zhiyuan Liu, **Wenyi Huang**, Yabin Zheng, Maosong Sun. Automatic Keyphrase Extraction via Topic Decomposition. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing (EMNLP 2010)*.
- Yongjin Liu, Xi Luo, Zeyu Jin, **Wenyi Huang**. Image type labyrinth design method based on Hamilton path. *Application No.: CN 200910079786.X. Public No.: CN 101504686*.

EXPERIENCE

The Pennsylvania State University

Research Assistant in CiteSeerX Group

August 2011 - Present

University Park, PA

- Designing RefSeerX (<http://refseer.ist.psu.edu/>)- a citation recommendation system. The research is focused on content-based citation recommendation. Given an abstract or a paper draft, we try to recommend a list of citations.
- I have investigated using translation model to bridge relationships between citation contexts and cited papers. The result was published in CIKM'12 conference.
- I am currently working on using deep learning and representation learning techniques to model citation context and cited documents.

PARC Xerox

Summer Research Intern

May 2014 - July 2014

Webster, NY

- Research on detecting personal event from Twitter profiles.

Qatar Compute Research Institute

Research Intern in Social Computing Group

June 2013 - August 2013

Doha, Qatar

- Research on ethnicity detection for Twitter profiles. We collected the twitter user profiles located in Qatar (self-stated locations or tweet from Qatar), and try to classify these users by their ethnicity.
- Study factors that make a person be perceived as American.

Tsinghua University

Research Assistant in National Lab for Information Science and Technology

September 2009 - July 2011

Beijing, China

- Research on keyphrase extraction for news articles and research papers. The research involves using topic model and graph-based random walk algorithms. The results were made towards a publication in EMNLP'10 conference.

Sogou.com, Inc

Engineering Intern in Web-NLP Group

July 2010 - October 2010

Beijing, China

- Building one part of query analysis system to identify whether the intention of a search query is to find a product.

TECHNICAL SKILLS

Computer Languages

Python, Java, C++, C, C#

Databases

MongoDB, MySQL

Operating System

Ubuntu, Redhat, Windows