

## **Yifan He**

Centre for Next Generation Localisation  
Dublin City University  
Glasnevin  
Dublin 9

### *Education*

Sep. 2008 – Now: PhD Student, Centre for Next Generation Localisation, Dublin City University  
Sep. 2005 – Jun. 2008: M. A. in Computational Linguistics Lab, School of Humanities and Social Sciences, Tsinghua University  
Sep. 2000 – Jul. 2005: Bachelor of Management in Management Information System, Department of Management Science and Engineering, Tongji University

### *Publications*

- Hongyu Gao, Bingsheng He, Xiangye Xiao, **Yifan He**, Qiong Luo and Bo Peng. Frequent Pattern Mining on Graphics Processors. ACM SIGMOD 2008 Undergraduate Poster Competition
- Ang Sun, Minghu Jiang, **Yifan He**, Lin Chen and Baozong Yuan. Chinese Question Answering Based on Syntax Analysis and Answer Classification. Acta Electronica Sinica 36(5), 2008 (in Chinese)
- **Yifan He** and Minghu Jiang. Text Categorization Using Distributional Clustering and Concept Extraction. IEEE ICIC 2007 (LNCS 4681)
- **Yifan He** and Minghu Jiang. Automatic Analysis of Chinese Dialect Tones Based on Self-Organizing Maps. IEEE ICIC 2007
- Zhan Wang, **Yifan He** and Minghu Jiang. A Comparison among Three Neural Networks in Text Classification. IEEE ICSP 2006

### *Research Projects*

**Sep. 2008 – Now: Integrated Language Technologies 1, Centre for Next Generation Localisation**

- Developing novel evaluation metrics for Machine Translation

**Jul. 2006 – Jun. 2008: China NSF Project: Document Categorization based on Cognitive Mechanism and Semantic Hierarchy (No. F02010625)**

- Developed the infrastructure of the Documentation Categorization System with Java, which features text processing, a categorization algorithm interface, evaluation and cross-validation
- Tuned the performance of mature feature selection and machine learning algorithms on the implemented system
- Put forward new feature selection algorithms based on experimental results

**Aug. 2007 – Jan. 2008: Frequent Pattern Mining on GPU**

- Surveyed major FP-Growth CPU implementations and implemented paralleled versions with OpenMP
- Implemented and modified some primitive algorithms with NVIDIA CUDA, including bitonic sort on the GPU

**Apr. 2005 – Jul. 2005: Retrieval System for IT Documents (Final-Year Project)**

- Expanded Lucene's compatibility to fit the need of parsing/indexing of Chinese HTML files

**Apr. 2004 – Sep. 2004: Management Information System for a Metal Company (Third-Year Project)**

- Led the development of the Decision Support module. Worked closely with group leader in the overall design and documentation
- The project is chosen as reference for future students

### *Working Experience*

#### **May 2008 – Aug. 2008: Intern Developer, Adobe System (China)**

- Developing test automation tools for desktop products
- Developing reporting tools for internal use, data collected from heterogeneous data source

#### **Jul. 2006 – Jun. 2008: Research Assistant, Computational Linguistics Lab, Tsinghua University**

- Worked on research projects of the lab (listed above)
- Administered the hardware/network environment of the lab

#### **Jul. 2007 - Aug. 2007: Social Service Intern, Food Bureau of Kun-Shan Municipal Government**

- Helped investigate the food supply chain in Kun-Shan (a city near Suzhou and Shanghai)
- Helped providing reports on the city's food security challenges

#### **Sep. 2006 – Feb. 2007: Instructor, Overseas Chinese Writing, Department of Chinese Language and Literature, Tsinghua University**

- Worked with other teachers to design the materials of the course
- Gave lectures to a class of foreign students 2h/week, prepared homework and exams

### *Awards and Honors*

- Guanghua Scholarship, Tsinghua U.: 1st Class (2006)
- Scholarship for Top Students, Tongji U.: 1st Class (2004), 2nd or 3rd Classes (2001 - 2003)
- Honor for Excellent Students, Tongji U.: 2005

### *Skills*

#### **(Programming) Languages**

- Extensive experience with C / C++ / Java / C# / Python / ActionScript
- Knowledge of LISP / Ruby / Matlab / PHP

#### **(Human) Languages**

- English: TOEFL iBT: 110/120; Certified Senior Interpreter in Shanghai
- German: HSPD-4 (College German Test in China): 92/100

#### **System Administration**

- OS: Debian / Ubuntu server administration. Familiar with CVS and Samba
- DB: MySQL / SQL Server / Access