

Lan (Anna) Huang

CONTACT INFORMATION Computer Science Department, Private Bag 3105 mobile: (64) 21-1117058
Hillcrest, Hamilton, 3216 e-mail: lh92@cs.waikato.ac.nz
New Zealand homepage: <http://cs.waikato.ac.nz/~lh92>

RESEARCH INTERESTS Text clustering
Information retrieval and digital libraries
Machine learning for natural language processing

EDUCATION **Waikato University**, Hamilton, New Zealand
Ph.D. Candidate, Computer Science September 2006–present
Thesis title: Concept-based Text Clustering
Supervisors: Prof. Ian H. Witten, Assoc. Prof. Eibe Frank
Beijing Normal University, Beijing, China
M.S., Information Technologies and Information Systems September 2003–June 2006
Supervisor: Prof. Guangjian Li
B.S., Information Management and Information Systems September 1999–June 2003

PUBLICATIONS PEER-REVIEWED Learning document similarity.
Lan Huang
New Zealand Computer Science Research Student Conference (NZCSRSC), 2011.
Combining Global Semantic Relatedness and Local Analysis for Document Clustering.
Lan Huang
New Zealand Computer Science Research Student Conference (NZCSRSC), 2010.
Clustering documents using a Wikipedia-based concept representation.
Lan Huang, David Milne and Eibe Frank, Ian H. Witten
The Thirteenth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 2009.
Clustering documents with active learning using Wikipedia.
Lan Huang, David Milne, Eibe Frank, Ian H. Witten
The Eighth IEEE International Conference on Data Mining (ICDM), 2008.
Similarity Measures for Text Document Clustering.
Lan Huang
New Zealand Computer Science Research Student Conference (NZCSRSC), 2008.
Digital Object Identifier:Handle System.
Guangjian Li, Lan Huang
Library Development, 2004 (3). (In Chinese)

PUBLICATIONS TRANSLATION Integrated Broadband Networks
Translators: Guangjian Li, Yongwen Huang, Lan Huang.
Publishing House of Eletronics Industry, 2005.

PROFESSIONAL EXPERIENCE Greenstone Project, Waikato University, Hamilton September 2006–present
Part-time research programmer; Liaison on internationalization
Development and maintenance of the Greenstone digital library software.
Co-ordinate with translators of about seventy languages.
The Eighth IEEE International Conference on Data Mining, Pisa, Italy December 2008
Student volunteer
Help conference organizers.

Heidelberg Institute for Theoretical Studies, Germany October 2008–December 2008
Intern, at the Natural Language Processing Group
Development of a classification algorithm based on Wikipedia.

National Science Library, Chinese Academy of Sciences, Beijing, China 2003–2005
Part-time research programmer, at the Information Technology Development Department
Design and development of the cross-database search system; the web-based intelligent in-
formation collecting system; and the open-source library portal system.

Management school, Beijing Normal University, Beijing, China 2003–2004
English language lecturer
English language lecturer for two semesters, at the public relations department of adult
education.

SELECTED	Doctoral Scholarship, Waikato University	September 2007–April 2011
HONOURS AND	Excellent Master Thesis, BNU	2006
AWARDS	Department student representative, BNU	2003–2006
	Ranked first; Directly matriculated to BNU's master program	2003
	Excellent Graduates of BNU; Excellent Graduates of Beijing	2003
	Lenovo Corporation Scholarship	2002
	BNU scholarships: first class twice; second class once	2001–2003

TALKS AND
PRESENTATIONS Clustering documents with active learning using Wikipedia.
Department of Computational Linguistics, Heidelberg University, Germany, 2008.

Clustering documents using a Wikipedia-based concept representation.
*Knowledge Discovery and Data Mining Laboratory, Istituto di Scienza e Tecnologie
dell'Informazione (ISTI-CNR), Pisa, Italy, 2008.*

REFEREES References available upon request.