Bo HAN

CONTACT INFORMATION	Room 8.19, Level 8, Doug McDonell Building Department of Computing and Information System Melbourne University, Victoria, Australia	Office: +61-3-8344-1357 Email: hanb (at) student.unimelb.edu.au Homepage: http://ww2.cs.mu.oz/hanb/

RESEARCH TOPIC

Improve the utility of Natural Language Processing in social media, particularly on noisy text normalisation and user geolocation prediction

EDUCATION

Melbourne University, Melbourne, Victoria, Australia

PhD Candidate   **July 2010 – present**
- Expected thesis submission date: January 2014
- Supervisors: Tim Baldwin, Paul Cook
- Committee members: Tim Baldwin, Paul Cook, Justin Zobel, Steven Bird

Dalian University of Technology, Dalian, Liaoning, China

Master of Engineering   **September 2007 – December 2009**
Bachelor of Engineering   **September 2003 – July 2007**

SELECTED PUBLICATIONS

Bo Han, Paul Cook and Timothy Baldwin, "A Stacking-based Approach to Twitter User Geolocation Prediction", In *ACL 2013*, 7–12, Sofia, Bulgaria.

Bo Han, Paul Cook and Timothy Baldwin "Geolocation Prediction in Social Media Data by Finding Location Indicative Words", In *COLING 2012*, 1045–1062, Mumbai, India.

Bo Han, Paul Cook and Timothy Baldwin, "Lexical Normalisation of Short Text Messages", In *ACM Transactions on Intelligent Systems and Technology*, 4(1), 5:1–5:27, 2013.

Yang Liao, Masud Moshtaghi, Bo Han, Shanika Karunasekera, Ramamohanarao Kotagiri, Timothy Baldwin, Aaron Harwood and Philippa Pattison, "Mining Micro-Blogs: Opportunities and Challenges", In *Social Networks: Computational Aspects and Mining*, (to appear), Ajith Abraham and Aboul Ella Hassanien (eds.), Springer.

Bo Han, Paul Cook and Timothy Baldwin, "Automatically Constructing a Normalisation Dictionary for Microblogs", In *Proceedings of EMNLP-CoNLL 2012*, 421–432, Jeju, Republic of Korea, 2012.

Timothy Baldwin, Paul Cook, Bo Han, Aaron Harwood, Shanika Karunasekera and Masud Moshtaghi, "A Support Platform for Event Detection using Social Intelligence", In *EACL 2012*, Demo Session, 69–72, Avignon, France, 2012.

Bo Han, Marco Lui and Timothy Baldwin, "Melbourne Language Group Microblog Track Report", In *TREC 2011*, NIST Special Publication: SP 500–295.

Bo Han and Timothy Baldwin, "Lexical normalisation of short text messages: Makn sens a #twitter", In *ACL 2011*, 368–378, Portland, OR, USA, 2011.

Xiaohua Liu, Bo Han, Kuan Li, Stephan Hyeonjun Stiller, and Ming Zhou, "SRL-based verb selection for ESL", In *EMNLP 2010*, 1068–1076, Cambridge, MA, USA, 2010.

Xiaohua Liu, Kuan Li, Bo Han, Ming Zhou, Long Jiang, Zhongyang Xiong, and Changning Huang, "Semantic role labeling for news tweets", In *Coling 2010*, 698–706, Beijing, China, 2010.

PROFESSIONAL EXPERIENCE

Melbourne University, Melbourne, Victoria, Australia

Teaching Assistant   **July 2013 – October 2013**
COMP90049 Knowledge Technologies, taught by Justin Zobel and Karin Verspoor

Research Assistant   **July 2011 – January 2012**
IBES project for "Early detection and mitigation of disastrous events with broadband-enabled social interaction technologies"

Microsoft Research Asia, Beijing, China

Research Intern   **September 2009– June 2010**

Microsoft Bing dictionary: Bilingual short text mining; Incremental web page crawling; English as a
Second Language learner writing assistance; Short text classification
Text processing in Twitter: Semantic Role Labelling for news tweets

Research Intern   **January 2007 – August 2007**
Data scientist: Web data cleansing, Web 2.0 design for internal projects, bilingual short text mining

ACADEMIC
EXPERIENCE

Reviewers: SIGIR 2011 (secondary), *SEM 2013, EMNLP 2013 and EACL 2014 (SRW)
Presentations and Invited Talks:
13/08/2013, Improving the Utility of NLP on Social Media, DFKI/COLI, Saarbrucken, Germany.
05/08/2013, A Stacking-based Approach to Twitter User Geolocation Prediction, Sofia, Bulgaria.
01/08/2013, Improving the Utility of NLP on Social Media, TU-Darmstadt, Germany.
28/08/2012, Event Detection over Social Media, Kyoto University, Japan.
12/07/2012, Automatically Constructing a Normalisation Dictionary for Microblogs, Jeju, Korea.
21/07/2011, Lexical Normalisation for Twitter Messages, Portland, USA.

SKILLS

General: Fast modeling and prototyping; Data analysis; Familiar with Twitter data processing
Programming: Python, Unix/Linux shell scripting, R, HPC, C/C++, Java

DEMOS

Twitter user geolocation prediction demo: `http://hum.csse.unimelb.edu.au:9000/geo.html`
Locations Indicative Words in Tweets: `http://ww2.cs.mu.oz.au/~hanb/coling2012.html`
Twitter text normalisation lexicon: `http://ww2.cs.mu.oz.au/~hanb/emnlp.tgz`

HONOURS AND
AWARDS

JASSO Machine Learning Summer School Scholarship, 2012.

Melbourne University Faculty Travelling Scholarship, ACL, 2012.

Google Australia Travel Scholarship, EMNLP, 2012.

ACL Student Travel Scholarship, ACL, 2011.

EII PhD School Scholarship, EII, 2011.

NICTA Top-up Scholarship, NICTA, 2010 - Now.

MIFRS, MIRS Scholarships, University of Melbourne, 2010 - Now.

Outstanding graduate (Summa Cum Laude), Dalian University of Technology, 2009.

Excellent graduate scholarship, Dalian University of Technology, 2009.

Bronze medal at ACM/ICPC regional programming competition, 2007.

"Tomorrow Star" award at Microsoft Research Asia, 2007 and 2010.

Outstanding undergraduate (Summa Cum Laude), Liaoning, China, 2007.

Excellent undergraduate scholarships, Dalian University of Technology, 2004–2006

REFEREES

A/Professor Tim Baldwin
PhD, A/Professor, ARC future fellow
Melbourne University
Email: *tb (at) ldwin.net*

Dr. Paul Cook
PhD, McKenzie postdoctoral fellow
Melbourne University
Email: *paulcook (at) unimelb.edu.au*

Dr. Xiaohua LIU
PhD, Researcher
Microsoft Research Asia
Email: *xiaoliu (at) microsoft.com*

Professor Xianchao ZHANG
PhD, Professor
Dalian University of Technology
Email: *xczhang (at) dlut.edu.cn*