

# **Eighth SIGHAN Workshop on Chinese Language Processing 2015**

**Co-located with ACL-IJCNLP 2015**

**Beijing, China  
30 - 31 July 2015**

**ISBN: 978-1-5108-0949-9**

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2015) by the Association for Computational Linguistics  
All rights reserved.

Printed by Curran Associates, Inc. (2015)

For permission requests, please contact the Association for Computational Linguistics  
at the address below.

Association for Computational Linguistics  
209 N. Eighth Street  
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006  
Fax: 1-570-476-0860

[acl@aclweb.org](mailto:acl@aclweb.org)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2634  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>Sequential Annotation and Chunking of Chinese Discourse Structure</i> Frances Yung, Kevin Duh and Yuji Matsumoto .....	1
<i>Create a Manual Chinese Word Segmentation Dataset Using Crowdsourcing Method</i> Shichang Wang, Chu-Ren Huang, Yao Yao and Angel Chan .....	7
<i>Chinese Named Entity Recognition with Graph-based Semi-supervised Learning Model</i> Aaron Li-Feng Han, Xiaodong Zeng, Derek F. Wong and Lidia S. Chao .....	15
<i>Sentence selection for automatic scoring of Mandarin proficiency</i> Jiahong Yuan, Xiaoying Xu, Wei Lai, Weiping Ye, Xinru Zhao and Mark Liberman .....	21
<i>ACBiMA: Advanced Chinese Bi-Character Word Morphological Analyzer</i> Ting-Hao Huang, Yun-Nung Chen and Lingpeng Kong .....	26
<i>Introduction to SIGHAN 2015 Bake-off for Chinese Spelling Check</i> Yuen-Hsien Tseng, Lung-Hao Lee, Li-Ping Chang and Hsin-Hsi Chen .....	32
<i>HANSpeller++: A Unified Framework for Chinese Spelling Correction</i> Shuiyuan Zhang, Jinhua Xiong, Jianpeng Hou, Qiao Zhang and Xueqi Cheng .....	38
<i>Word Vector/Conditional Random Field-based Chinese Spelling Error Detection for SIGHAN-2015 Evaluation</i> Yih-Ru Wang and Yuan-Fu Liao .....	46
<i>Introduction to a Proofreading Tool for Chinese Spelling Check Task of SIGHAN-8</i> Tao-Hsing Chang, Hsueh-Chih Chen and Cheng-Han Yang .....	50
<i>Overview of Topic-based Chinese Message Polarity Classification in SIGHAN 2015</i> Xiangwen Liao, Binyang Li and Liheng Xu .....	56
<i>A Joint Model for Chinese Microblog Sentiment Analysis</i> Yuhui Cao, Zhao Chen, Ruifeng Xu, Tao Chen and Lin Gui .....	61
<i>Learning Salient Samples and Distributed Representations for Topic-Based Chinese Message Polarity Classification</i> Xin Kang, Yunong Wu and Zhifei Zhang .....	68
<i>An combined sentiment classification system for SIGHAN-8</i> Qiuchi Li, Qiyu Zhi and Miao Li .....	74
<i>Linguistic Knowledge-driven Approach to Chinese Comparative Elements Extraction</i> MinJun Park and Yulin Yuan .....	79
<i>A CRF Method of Identifying Prepositional Phrases in Chinese Patent Texts</i> Hongzheng Li and Yaohong Jin .....	86
<i>Emotion in Code-switching Texts: Corpus Construction and Analysis</i> Sophia Lee and Zhongqing Wang .....	91
<i>Chinese in the Grammatical Framework: Grammar, Translation, and Other Applications Anonymous</i> Aarne Ranta, Tian Yan and Haiyan Qiao .....	100

<i>KWB: An Automated Quick News System for Chinese Readers</i>	
Yiqi Bai, Wenjing Yang, Hao Zhang, Jingwen Wang, Ming Jia, Roland Tong and Jie Wang . . . .	110
<i>Chinese Semantic Role Labeling using High-quality Syntactic Knowledge</i>	
Gongye Jin, Daisuke Kawahara and Sadao Kurohashi . . . . .	120
<i>Chinese Spelling Check System Based on N-gram Model</i>	
Weijian Xie, Peijie Huang, Xinrui Zhang, Kaiduo Hong, Qiang Huang, Bingzhou Chen and Lei Huang . . . . .	128
<i>NTOU Chinese Spelling Check System in Sighan-8 Bake-off</i>	
Wei-Cheng Chu and Chuan-Jie Lin . . . . .	137
<i>Topic-Based Chinese Message Sentiment Analysis: A Multilayered Analysis System</i>	
hongjie li, zhongqian sun and wei yang . . . . .	144
<i>Rule-Based Weibo Messages Sentiment Polarity Classification towards Given Topics</i>	
Hongzhao Zhou, Yonglin Teng, Min Hou, Wei He, Hongtao Zhu, Xiaolin Zhu and Yanfei Mu .	149
<i>Topic-Based Chinese Message Polarity Classification System at SIGHAN8-Task2</i>	
Chun Liao, Chong Feng, Sen Yang and Heyan Huang . . . . .	158
<i>CT-SPA: Text sentiment polarity prediction model using semi-automatically expanded sentiment lexicon</i>	
Tao-Hsing Chang, Ming-Jhih Lin, Chun-Hsien Chen and Shao-Yu Wang . . . . .	164
<i>Chinese Microblogs Sentiment Classification using Maximum Entropy</i>	
Dashu Ye, Peijie Huang, Kaiduo Hong, Zhuoying Tang, Weijian Xie and Guilong Zhou . . . . .	171
<i>NDMSCS: A Topic-Based Chinese Microblog Polarity Classification System</i>	
Yang Wang, Yaqi Wang, Shi Feng, Daling Wang and Yifei Zhang . . . . .	180
<i>NEUDM: A System for Topic-Based Message Polarity Classification</i>	
Yaqi Wang, Shi Feng, Daling Wang and Yifei Zhang . . . . .	185