Third CIPS-SIGHAN Joint Conference on Chinese Language Processing

(CLIP 2014)

Wuhan, China
20 - 21 October 2014

ISBN: 978-1-63439-414-7
# Table of Contents

## Invited Talks

*Research on Chinese discourse rhetorical structure representation scheme and corpus annotation*
  Guodong Zhou ................................................................. 1

*The Construction of language Resource and Knowledge Base for Chinese Language Computing*
  Zhifang Sui ................................................................. 2

## Research Papers

*COV Model and its Application in Chinese Part-of-Speech Tagging*
  Xing Fukun and Song Rou ...................................................... 3

*Maximum Entropy for Chinese Comma Classification with Rich Linguistic Features*
  Xiaojuan Li, Hua Yang and JiangPing Huang ............................. 11

*Improving Bilingual Lexicon Extraction Performance from Comparable Corpora via Optimizing Translation Candidate Lists*
  Shaoqi Wang, Miao Li, Zede Zhu, Zhenxin Yang and Shizhuang Weng ......................................................... 18

*Ch2R: A Chinese Chatter Robot for Online Shopping Guide*
  Peijie Huang, Xianmao Lin, Zeqi Lian, De Yang, Xiaoling Tang, Li Huang, Qiang Huang, Xiupeng Wu, Guisheng Wu and Xinrui Zhang ......................................................... 26

*Improving Chinese Sentence Polarity Classification via Opinion Paraphrasing*
  Guohong Fu, Yu He, Jiaying Song and Chaoyue Wang ..................... 35

*Problematic Situation Analysis and Automatic Recognition for Chinese Online Conversational System*
  Yang Xiang, Yaoyun Zhang, Xiaqianguang Zhou, Xiaolong Wang and Yang Qin ......................................................... 43

*Segment-based Fine-grained Emotion Detection for Chinese Text*
  Odbal and Zengfu Wang .......................................................... 52

*Bilingual Product Name Dictionary Construction Using a Two Stage Method*
  Yatian Shen and Xuanjing Huang .............................................. 61

*Detection on Inconsistency of Verb Phrase in TreeBank*
  Chaoqun Duan, Dequan Zheng, Conghui Zhu, Sheng Li and Hongye Tan ......................................................... 70

*Local Phrase Reordering Model for Chinese-English Patent Machine Translation*
  Xiaodie Liu, Yun Zhu and Yaohong Jin ....................................... 76

*Sentence Alignment of Historical Classics based on Mode Prediction and Term Translation Pairs*
  Chao Che and Xiaojun Zheng .................................................... 85

## Bake-off Papers

*The CIPS-SIGHAN CLP 2014 Chinese Word Segmentation Bake-off*
  Huiming Duan, Zhifang Sui and Tao Ge ........................................ 90

  Lu Xiang, Xiaqing Li and Yu Zhou ............................................. 96
Leveraging Rich Linguistic Features for Cross-domain Chinese Segmentation
Guohua Wu, Dezhu He, Keli Zhong, Xue Zhou and Caixia Yuan ................................. 101

Personal Attributes Extraction in Chinese Text Bakeoff in CLP 2014: Overview
Ruifeng Xu, Shuai Wang, Feng Shi and Jian Xu ......................................................... 108

Personal Attributes Extraction Based on the Combination of Trigger Words, Dictionary and Rules
Kailun Zhang, Mingyin Wang, Xiaoyue Cong, Fang Huang, Hongfa Xue, Lei Li and Zhiqiao Gao
114

An Introduction to BLCU Personal Attributes Extraction System
Dong YU, Cheng YU, Qin QU, Gongbo TANG, Chunhua LIU, Yue TIAN and Jing YI ........ 120

Overview of SIGHAN 2014 Bake-off for Chinese Spelling Check
Liang-Chih Yu, Lung-Hao Lee, Yuen-Hsien Tseng and Hsin-Hsi Chen ......................... 126

Extended HMM and Ranking Models for Chinese Spelling Correction
Jinhua Xiong, Qiao Zhang, Jianpeng Hou, Qianbo Wang, Yuanzhuo Wang and Xueqi Cheng . 133

Chinese Word Spelling Correction Based on Rule Induction
Jui-Feng Yeh, Yun-Yun Lu, Chen-Hsien Lee, Yu-Hsiang Yu and Yong-Ting Chen .......... 139

Qiang Zhou .................................................................................................................. 146

Generative CCG Parsing with OOV Prediction
Huijia Wu .................................................................................................................... 153

An Improved Graph Model for Chinese Spell Checking
Yang Xin, Hai Zhao, Yuzhu Wang and Zhongye Jia .................................................... 157

Introduction to NJUPT Chinese Spelling Check Systems in CLP-2014 Bakeoff
Lei Gu, Yong Wang and Xitao Liang ............................................................................ 167

Chinese Spelling Check System Based on Tri-gram Model
Qiang Huang, Peijie Huang, Xirnui Zhang, Weijian Xie, Kaiduo Hong, Bingzhou Chen and Lei
Huang .......................................................................................................................... 173

Introduction to BIT Chinese Spelling Correction System at CLP 2014 Bake-off
Min Liu, Ping Jian and Heyan Huang ........................................................................... 179

A mixed approach for Chinese word segmentation
Zhen Wang .................................................................................................................. 186

Extraction system for Personal Attributes Extraction of CLP2014
Zhen Wang .................................................................................................................. 189

A Study on Personal Attributes Extraction Based on the Combination of Sentences Classifications and Rules
Nan-chang Cheng, Cheng-qing Zong, Min Hou and Yong-lin Teng ............................. 192

Chinese Spell Checking Based on Noisy Channel Model
Hsun-wen Chiu, Jian-cheng Wu and Jason S. Chang .................................................... 202

NTOU Chinese Spelling Check System in CLP Bake-off 2014
Wei-Cheng Chu and Chuan-Jie Lin ............................................................................. 210
NCTU and NTUT's Entry to CLP-2014 Chinese Spelling Check Evaluation
Yih-Ru Wang and Yuan-Fu Liao .................................................. 216

Chinese Spelling Error Detection and Correction Based on Language Model, Pronunciation, and Shape
Junjie Yu and Zhenghua Li .................................................. 220