

Maosong Sun · Xiaojie Wang  
Baobao Chang · Deyi Xiong (Eds.)

# Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data

16th China National Conference, CCL 2017  
*and* 5th International Symposium, NLP-NABD 2017  
Nanjing, China, October 13–15, 2017  
Proceedings

# Contents

## Fundamental Theory and Methods of Computational Linguistics

Arabic Collocation Extraction Based on Hybrid Methods . . . . .	3
<i>Alaa Mamdouh Akef, Yingying Wang, and Erhong Yang</i>	
Employing Auto-annotated Data for Person Name Recognition in Judgment Documents. . . . .	13
<i>Limin Wang, Qian Yan, Shoushan Li, and Guodong Zhou</i>	
Closed-Set Chinese Word Segmentation Based on Convolutional Neural Network Model . . . . .	24
<i>Zhipeng Xie</i>	
Improving Word Embeddings for Low Frequency Words by Pseudo Contexts. . . . .	37
<i>Fang Li and Xiaojie Wang</i>	
A Pipelined Pre-training Algorithm for DBNs. . . . .	48
<i>Zhiqiang Ma, Tuya Li, Shuangtao Yang, and Li Zhang</i>	
Enhancing LSTM-based Word Segmentation Using Unlabeled Data . . . . .	60
<i>Bo Zheng, Wanxiang Che, Jiang Guo, and Ting Liu</i>	

## Machine Translation and Multilingual Information Processing

Context Sensitive Word Deletion Model for Statistical Machine Translation . . . . .	73
<i>Qiang Li, Yaqian Han, Tong Xiao, and Jingbo Zhu</i>	
Cost-Aware Learning Rate for Neural Machine Translation . . . . .	85
<i>Yang Zhao, Yining Wang, Jiajun Zhang, and Chengqing Zong</i>	

## Knowledge Graph and Information Extraction

Integrating Word Sequences and Dependency Structures for Chemical-Disease Relation Extraction. . . . .	97
<i>Huiwei Zhou, Yunlong Yang, Zhuang Liu, Zhe Liu, and Yahui Men</i>	
Named Entity Recognition with Gated Convolutional Neural Networks . . . . .	110
<i>Chunqi Wang, Wei Chen, and Bo Xu</i>	

Improving Event Detection via Information Sharing Among Related Event Types . . . . .	122
<i>Shulin Liu, Yubo Chen, Kang Liu, Jun Zhao, Zhunchen Luo, and Wei Luo</i>	
Joint Extraction of Multiple Relations and Entities by Using a Hybrid Neural Network. . . . .	135
<i>Peng Zhou, Suncong Zheng, Jiaming Xu, Zhenyu Qi, Hongyun Bao, and Bo Xu</i>	
A Fast and Effective Framework for Lifelong Topic Model with Self-learning Knowledge. . . . .	147
<i>Kang Xu, Feng Liu, Tianxing Wu, Sheng Bi, and Guilin Qi</i>	
Collective Entity Linking on Relational Graph Model with Mentions. . . . .	159
<i>Jing Gong, Chong Feng, Yong Liu, Ge Shi, and Heyan Huang</i>	
XLink: An Unsupervised Bilingual Entity Linking System . . . . .	172
<i>Jing Zhang, Yixin Cao, Lei Hou, Juanzi Li, and Hai-Tao Zheng</i>	
Using Cost-Sensitive Ranking Loss to Improve Distant Supervised Relation Extraction . . . . .	184
<i>Daojian Zeng, Junxin Zeng, and Yuan Dai</i>	
Multichannel LSTM-CRF for Named Entity Recognition in Chinese Social Media . . . . .	197
<i>Chuanhai Dong, Huijia Wu, Jiajun Zhang, and Chengqing Zong</i>	
<b>Language Resource and Evaluation</b>	
Generating Chinese Classical Poems with RNN Encoder-Decoder . . . . .	211
<i>Xiaoyuan Yi, Ruoyu Li, and Maosong Sun</i>	
Collaborative Recognition and Recovery of the Chinese Intercept Abbreviation . . . . .	224
<i>Jinshuo Liu, Yusen Chen, Juan Deng, Donghong Ji, and Jeff Pan</i>	
Semantic Dependency Labeling of Chinese Noun Phrases Based on Semantic Lexicon. . . . .	237
<i>Yimeng Li, Yanqiu Shao, and Hongkai Yang</i>	
<b>Information Retrieval and Question Answering</b>	
Bi-directional Gated Memory Networks for Answer Selection. . . . .	251
<i>Wei Wu, Houfeng Wang, and Sujian Li</i>	

Generating Textual Entailment Using Residual LSTMs . . . . .	263
<i>Maosheng Guo, Yu Zhang, Dezhi Zhao, and Ting Liu</i>	
Unsupervised Joint Entity Linking over Question Answering Pair with Global Knowledge . . . . .	273
<i>Cao Liu, Shizhu He, Hang Yang, Kang Liu, and Jun Zhao</i>	
Hierarchical Gated Recurrent Neural Tensor Network for Answer Triggering . . . . .	287
<i>Wei Li and Yunfang Wu</i>	
Question Answering with Character-Level LSTM Encoders and Model-Based Data Augmentation . . . . .	295
<i>Run-Ze Wang, Chen-Di Zhan, and Zhen-Hua Ling</i>	
Exploiting Explicit Matching Knowledge with Long Short-Term Memory . . .	306
<i>Xinqi Bao and Yunfang Wu</i>	
<b>Text Classification and Summarization</b>	
Topic-Specific Image Caption Generation. . . . .	321
<i>Chang Zhou, Yuzhao Mao, and Xiaojie Wang</i>	
Deep Learning Based Document Theme Analysis for Composition Generation . . . . .	333
<i>Jiahao Liu, Chengjie Sun, and Bing Qin</i>	
UIDS: A Multilingual Document Summarization Framework Based on Summary Diversity and Hierarchical Topics . . . . .	343
<i>Lei Li, Yazhao Zhang, Junqi Chi, and Zuying Huang</i>	
Conceptual Multi-layer Neural Network Model for Headline Generation . . . .	355
<i>Yidi Guo, Heyan Huang, Yang Gao, and Chi Lu</i>	
<b>Social Computing and Sentiment Analysis</b>	
Local Community Detection Using Social Relations and Topic Features in Social Networks . . . . .	371
<i>Chengcheng Xu, Huaping Zhang, Bingbing Lu, and Songze Wu</i>	
<b>NLP Applications</b>	
DIM Reader: Dual Interaction Model for Machine Comprehension . . . . .	387
<i>Zhuang Liu, Degen Huang, Kaiyu Huang, and Jing Zhang</i>	

Multi-view LSTM Language Model with Word-Synchronized Auxiliary Feature for LVCSR . . . . .	398
<i>Yue Wu, Tianxing He, Zhehuai Chen, Yanmin Qian, and Kai Yu</i>	
Memory Augmented Attention Model for Chinese Implicit Discourse Relation Recognition . . . . .	411
<i>Yang Liu, Jiajun Zhang, and Chengqing Zong</i>	
Natural Logic Inference for Emotion Detection . . . . .	424
<i>Han Ren, Yafeng Ren, Xia Li, Wenhe Feng, and Maofu Liu</i>	
<b>Minority Language Information Processing</b>	
Tibetan Syllable-Based Functional Chunk Boundary Identification . . . . .	439
<i>Shumin Shi, Yujian Liu, Tianhang Wang, Congjun Long, and Heyan Huang</i>	
Harvest Uyghur-Chinese Aligned-Sentences Bitexts from Multilingual Sites Based on Word Embedding . . . . .	449
<i>ShaoLin Zhu, Xiao Li, YaTing Yang, Lei Wang, and ChengGang Mi</i>	
Language Model for Mongolian Polyphone Proofreading . . . . .	461
<i>Min Lu, Feilong Bao, and Guanglai Gao</i>	
End-to-End Neural Text Classification for Tibetan. . . . .	472
<i>Nuo Qun, Xing Li, Xipeng Qiu, and Xuanjing Huang</i>	
<b>Author Index</b> . . . . .	481