Communications
in Computer and Information Science

Editorial Board

Simone Diniz Junqueira Barbosa
Pontifical Catholic University of Rio de Janeiro (PUC-Rio),
Rio de Janeiro, Brazil

Phoebe Chen
La Trobe University, Melbourne, Australia

Alfredo Cuzzocrea
ICAR-CNR and University of Calabria, Italy

Xiaoyong Du
Renmin University of China, Beijing, China

Joaquim Filipe
Polytechnic Institute of Setúbal, Portugal

Orhun Kara
TÜBİTAK BİLGEM and Middle East Technical University, Turkey

Igor Kotenko
St. Petersburg Institute for Informatics and Automation
of the Russian Academy of Sciences, Russia

Krishna M. Sivalingam
Indian Institute of Technology Madras, India

Dominik Ślezak
University of Warsaw and Infobright, Poland

Takashi Washio
Osaka University, Japan

Xiaokang Yang
Shanghai Jiao Tong University, China
Preface

NLPCC (CCF Conference on Natural Language Processing and Chinese Computing) is the annual conference of CCF TCCI (Technical Committee of Chinese Information). As a leading conference in the field of NLP and Chinese Computing of CCF, NLPCC is the premier forum for NLP researchers and practitioners from academia, industry, and government in China and Pacific Asia to share their ideas, research results and experiences, which will highly promote the research and technical innovation in these fields domestically and internationally. The papers contained in these proceedings address challenging issues in Web mining and big data, knowledge acquisition, search and ads, social networks, machine translation and multi-lingual information access, question answering and user interface, machine learning for NLP, as well as the fundamentals and applications of Chinese computing.

This year, NLPCC received 203 submissions. After a thorough reviewing process, 31 English papers and 14 Chinese papers were selected for presentation as full papers, with an acceptance rate of 22.17%. Furthermore, this year’s NLPCC also included 11 English papers and 14 Chinese papers as posters, with an acceptance rate of 12.32%. Additionally, five industrial/demo papers were selected. The Chinese full papers together with posters are published by ACTA Scientiarum Naturalium Universitatis Pekinensis, and are not included in these proceedings. This volume contains the 31 English full papers presented at NLPCC 2013 and 13 short papers including the 11 English posters and two industrial/demo papers.

The high-quality program would not have been possible without the authors who chose NLPCC 2013 as a venue for their publications. We are also very grateful to the Program Committee members and Organizing Committee members, who put a tremendous amount of effort into soliciting and selecting research papers with a balance of high quality and new ideas and new applications.

We hope that you enjoy reading and benefit from the proceedings of NLPCC 2013.

November 2013

Guodong Zhou
Juanzi Li
NLPCC 2013 was organized by the Technical Committee of the Chinese Information of CCF, Peking University, and Microsoft Research Asia.

Organizing Committee

Conference Co-chairs
Ming Zhou  Microsoft Research Asia, China
Dan Yang  Chongqing University, China

Program Co-chairs
Guodong Zhou  Soochow University, China
Juanzi Li  Tsinghua University, China

Area Chairs

Fundamentals on CIT
Qiang Zhou  Tsinghua University, China
Tuergen  Xinjiang University, China

Applications on CIT
Donghong Ji  Wuhan University, China
Chenglin Liu  Chinese Academy of Sciences

Machine Translation
Yajuan Lv  Chinese Academy of Sciences
Muyun Yang  Harbin Institute of Technology, China

Web Mining and Big Data
Shiqi Zhao  Baidu, China
Zhongyang Xiong  Chongqing University, China

Machine Learning for NLP
Liwei Wang  Peking University, China
Xiaofei He  Zhejiang University, China
Knowledge Acquisition
Juanzi Li  Tsinghua University, China
Kuo Zhang  Sogou, China

NLP for Social Networks
Jie Tang  Tsinghua University, China
Shoushan Li  Soochow University, China

NLP for Search and Ads
Jun Ma  Shandong University, China
Bin Wang  Chinese Academy of Sciences, China

QA and User Interaction
Zhengtao Yu  Kunming University of Science and Technology, China
Furu Wei  Microsoft Research Asia, China

Panel Chair
Chengqing Zong  Chinese Academy of Sciences

Demo Co-chairs
Qingsheng Li  Anyang Normal University, China
Xiaojun Wan  Peking University, China

ADL/Tutorial Chair
Min Zhang  Soochow University, China
Mu Li  Microsoft Research Asia, China

Organizing Chair
Dongyan Zhao  Peking University, China

Local Arrangements Chair
Zhongyang Xiong  Chongqing University, China

Publication Chair
Yansong Feng  Peking University, China

Evaluation Chair
Xiaojun Wan  Peking University, China
**Financial Chair**
Zhengyu Zhu
China Computer Federation, China

**Sponsor Chair**
Tang Zhi
Peking University, China

**Website Chair**
Aixia Jia
Peking University, China

**Program Committee**

Jiajun Chen
Nanjing University

Hsin-Hsi Chen
National Taiwan University

Zhumin Chen
Shandong University

Xiaoqing Ding
Tsinghua University

Xinghua Fan
Chongqing University of Posts and Telecommunications

Yang Feng
Sheffield University, UK

Guanglai Gao
Inner Mongolia University

Zhiqiang Gao
Southeast University

Tingting He
Huazhong Normal University

Yulan He
Open University, UK

Yu Hong
Soochow University

Hongxu Hou
Inner Mongolia University

Degen Huang
Dalian University of Technology

Heming Huang
Qinghai Normal University

Xiaojun Huang
Founder Apabi

Xuanjing Huang
Fudan University

Wenbin Jiang
Institute of Computing Technology, CAS

Jing Jiang
Singapore Management University

Long Jiang
Alibaba Corp.

Shengyi Jiang
Guangdong University of Foreign Studies

Peng Jin
Leshan Normal University

Fang Kong
Soochow University

Baoli Li
Henan University of Technology

Fenghua Li
Institute of Information Engineering, CAS

Ning Li
Beijing Information Science and Technology University

Wenjie Li
Hong Kong Polytechnic University

Zhenxing Li
Agile Century

Henry Li
Microsoft Research Asia

Qingsheng Li
Anyang Normal University

Zhouhui Lian
Peking University
Bangyong Liang STCA, Microsoft
Hongfei Lin Dalian University of Technology
Wai Lam Chinese University of Hong Kong
Bingquan Liu Harbin Institute of Technology
Dexi Liu Jiangxi University of Finance and Economics
Pengyuan Liu Beijing Language and Culture University
Xiaohua Liu Microsoft Research Asia
Yang Liu Shandong University
Yang Liu Tsinghua University
Changqing Liu Ningxia University
Xueqiang Lyu TRS Inc., China
Jie Ma Nankai University
Qing Ma RyuKoku University, Japan
Shaoping Ma Tsinghua University
Yanjun Ma Baidu Inc.
Jian-Yun Nie University of Montreal, Canada
Zhendong Niu Beijing Institute of Technology
Ngodrup Tibet University
Longhua Qian Soochow University
Bing Qin Harbin Institute of Technology
Tao Qin Microsoft Research Asia
Weiguang Qu Nanjing Normal University
Liyun Ru Sohu Inc., China
Junichi Tsujii Microsoft Research Asia
Xiaodong Shi Xiamen University
Jinsong Su Xiamen University
Chengjie Sun Harbin Institute of Technology
Weiwei Sun Peking University
Bo Wang Tianjin University
Haifeng Wang Baidu Inc.
Houfeng Wang Peking University
Mingwen Wang Jiangxi Normal University
Shuaiqiang Wang Shandong University of Finance and Economics
Xiaojie Wang Beijing University of Posts and Telecommunications
Gang Wu Northeastern University
Yunqing Xia Tsinghua University
Deyi Xiong Singapore I2R, Singapore
Jinan Xu Beijing Jiaotong University
Jun Xu Huawei
Endong Xun Beijing Language and Culture University
Wenchuan Yang Beijing University of Posts and Telecommunications
Tianfang Yao Shanghai Jiao Tong University
Jianmin Yin Huaguang Inc.
Masaaki Nagata NTT, Japan
Hongzhi Yu Northwest University for Nationalities
Yusufu Xinjiang Normal University
Dongdong Zhang Microsoft Research Asia
Jiajun Zhang Automation Institute, CAS
Jianguo Zhang Founder Electronics
Jianpei Zhang Harbin Engineering University
Junsong Zhang Xiamen University
Min Zhang Tsinghua University
Qi Zhang Fudan University
Ruiqiang Zhang Yahoo!, USA
Yangsen Zhang Beijing Information Science Technology University
Yujie Zhang Beijing Jiaotong University
Hai Zhao Shanghai Jiao Tong University
Jun Zhao Institute of Automation, CAS
Tiejun Zhao Harbin Institute of Technology
Yanyan Zhao Harbin Institute of Technology
Deyu Zhou Southeast University
Guomin Zhou Agricultural Information Institute, CAAS
Junsheng Zhou Nanjing Normal University
Wen Zhou Shanghai University
Xueguang Zhou Navy Engineering University
Jingbo Zhu Northeastern University
Qiaoming Zhu Soochow University
Qing Zhu Renmin University of China
Tse ring rgyal Qinghai Normal University
XII Organization

Organizers

Organized by

China Computer Federation, China

Hosted by

Chongqing University

State Key Laboratory of Digital Publishing

In Cooperation with:

Tsinghua University

ACTA Scientiarum Naturalium

Universitatis Pekinensis

Sponsoring Institutions

Microsoft Research Asia

Sogou Inc.

Sina Weibo

Tencent Weibo

Sichuan Institute of Computer Science

Mingbo Education Technology
Keynote Talks
Knowledge Mining and Semantic Search

Wei-Ying Ma

Microsoft Research Asia, Beijing, China
wyma@microsoft.com

Abstract. Today’s search engines are primarily operated based on terms and string matching. While this term-based paradigm has been pretty successful in the past 15 years, it has started to show many limitations given the rapid evolution of the Web. The Web today contains not only static documents but also dynamic information about real world entities. It is becoming a digital copy of the world, capturing data from people, products, services, locations, and even objects. This new Web requires a new paradigm that can directly fulfill people’s information needs and empower them with knowledge. In this talk, I will introduce a new entity-based search paradigm and various knowledge mining techniques to construct knowledge graphs from different types of data. I will show how we can use the knowledge to understand queries, enable semantic matching, and provide direct answers to natural language queries.

Bio: Dr. Wei-Ying Ma is an Assistant Managing Director at Microsoft Research Asia, where he oversees multiple research groups including Web Search and Mining, Natural Language Computing, Data Management and Analytics, and Internet Economics and Computational Advertising. He and his team of researchers have developed many key technologies that have been transferred to Microsoft’s Online Services Division, including Bing Search Engine and Microsoft Advertising. He has published more than 250 papers at international conferences and journals. He is a Fellow of the IEEE and a Distinguished Scientist of the ACM. He currently serves on the editorial boards of ACM Transactions on Information System (TOIS) and is a member of the International World Wide Web (WWW) Conferences Steering Committee. In recent years, he has served as program co-chair of WWW 2008 and as general co-chair of ACM SIGIR 2011. He received a Bachelor of Science in electrical engineering from the National Tsing Hua University in Taiwan in 1990. He earned both a Master of Science and a doctorate in electrical and computer engineering from the University of California at Santa Barbara in 1994 and 1997, respectively. More information about him can be found at http://research.microsoft.com/en-us/people/wyma/
Domestic Chinese Treebanks Need to Strengthen the Predicate-Argument Structure Description

Changning Huang

Department of Computer Science and Technology, Tsinghua University
Beijing, 100084, China
cnhuang0908@126.com

Abstract. Base on Penn English Treebank (PTB-II), 1998–2000 the University of Pennsylvania built the Penn Chinese Treebank (CTB) with the news corpus of Xinhua News Agency. CTB not only sets the predicate-argument structure description as its important goal in the very beginning, but also defines a standard hierarchy structure of phrases based on the X-bar Theory. According to this hierarchy structure CTB strives to make each bracket pair or sub-tree dominated by a phrase node to represent only one abstract grammatical relation. A comparative review has been made on domestic and oversea Chinese treebanks. Although domestic treebanks generally annotates the phrasal notes, such as NP, VP, PP, etc., with internal structures information, such as subject-predicate, verb-object, and modifier-head, etc., and in order to indicate the head of the phrase an integer or other symbol is attached to the phrase note. Apparently they did not set the predicate-argument structure description as an important goal of treebank construction. Starting from the X-bar Theory about the representation of general phrase structure, the article illustrates the terms of complement, adjunct and specifier in Contemporary Syntax, explains the applications of empty category and co-indexing in the syntactic structure. Then, by the CTB specific examples the article describes the predicate-argument structure on the syntactic trees.

Above is the first part of my speech, in the following part I’ll briefly reviews the research of Chinese information processing in the past three decades and looking to the near future. I will be happy to discuss with my colleagues and graduate students on my teaching and research experience.

Bio: Professor Chang-Ning Huang graduated from Department of Automatic Control, Tsinghua University, Beijing, China, on 1961. He joined Microsoft Research Asia (MSRA) as a senior researcher and the research manager of the NLC group on April 1999. Before that he was the professor of the Department of Computer Science and Technology, Tsinghua University, and the founder and Director of the NLP research group in Tsinghua from 1982 to 1999. He was retired from Microsoft on August 2012. His research interest includes Chinese word segmentation, POS tagging, named entity recognition, word sense disambiguation, coreference resolution, syntactic parsing and machine translation, etc.
Some Mathematical Models to Turn Social Media into Knowledge

Xiaojin Zhu
Department of Computer Sciences,
University of Wisconsin-Madison, WI, USA
jerryzhu@cs.wisc.edu

Abstract. Social media data-mining opens up many interesting research questions, whose answers correspond to elegant mathematical models that go beyond traditional NLP techniques. In this talk we present two examples, namely estimating intensity from counts and identifying the most chatty users. In both examples, naive heuristic methods do not take full advantage of the data. In contrast, there are mathematical models, in the first case inhomogeneous Poisson process and in the second case multi-arm bandit, with provable properties that better extract knowledge from social media.

Bio: Xiaojin Zhu is an Associate Professor in the Department of Computer Sciences at the University of Wisconsin-Madison, with affiliate appointments in the Departments of Electrical and Computer Engineering and Psychology. Dr. Zhu received his B.S. and M.S. degrees in Computer Science from Shanghai Jiao Tong University in 1993 and 1996, respectively, and a Ph.D. degree in Language Technologies from Carnegie Mellon University in 2005. He was a research staff member at IBM China Research Laboratory from 1996 to 1998. Dr. Zhu received the National Science Foundation CAREER Award in 2010. His research interest is in machine learning, with applications in natural language processing, cognitive science, and social media. http://pages.cs.wisc.edu/jerryzhu/
## Table of Contents

### Fundamentals on Language Computing

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text Window Denoising Autoencoder: Building Deep Architecture for Chinese Word Segmentation</td>
<td>1</td>
</tr>
<tr>
<td><em>Ke Wu, Zhiqiang Gao, Cheng Peng, and Xiao Wen</em></td>
<td></td>
</tr>
<tr>
<td>Language Model for Cyrillic Mongolian to Traditional Mongolian Conversion</td>
<td>13</td>
</tr>
<tr>
<td><em>Feilong Bao, Guanglai Gao, Xueliang Yan, and Hongwei Wang</em></td>
<td></td>
</tr>
<tr>
<td>Sentence Compression Based on ILP Decoding Method</td>
<td>19</td>
</tr>
<tr>
<td><em>Hongling Wang, Yonglei Zhang, and Guodong Zhou</em></td>
<td></td>
</tr>
<tr>
<td>Chinese Negation and Speculation Detection with Conditional Random Fields</td>
<td>30</td>
</tr>
<tr>
<td><em>Zhancheng Chen, Bowei Zou, Qiaoming Zhu, and Peifeng Li</em></td>
<td></td>
</tr>
<tr>
<td>Chinese Argument Extraction Based on Trigger Mapping</td>
<td>41</td>
</tr>
<tr>
<td><em>Yuan Huang, Peifeng Li, and Qiaoming Zhu</em></td>
<td></td>
</tr>
<tr>
<td>Exploring Multiple Chinese Word Segmentation Results Based on Linear Model</td>
<td>50</td>
</tr>
<tr>
<td><em>Chen Su, Yujie Zhang, Zhen Guo, and Jinan Xu</em></td>
<td></td>
</tr>
</tbody>
</table>

### Applications on Language Computing

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Unified Framework for Emotional Elements Extraction Based on Finite State Matching Machine</td>
<td>60</td>
</tr>
<tr>
<td><em>Yunzhi Tan, Yongfeng Zhang, Min Zhang, Yiqun Liu, and Shaoping Ma</em></td>
<td></td>
</tr>
<tr>
<td>Structure-Based Web Access Method for Ancient Chinese Characters</td>
<td>72</td>
</tr>
<tr>
<td><em>Xiaoqing Lu, Yingmin Tang, Zhi Tang, Yujun Gao, and Jianguo Zhang</em></td>
<td></td>
</tr>
<tr>
<td>Simulated Spoken Dialogue System Based on IOHMM with User History</td>
<td>83</td>
</tr>
<tr>
<td><em>Changliang Li, Bo Xu, XiuYing Wang, WenDong Ge, and HongWei Hao</em></td>
<td></td>
</tr>
<tr>
<td>A Mixed Model for Cross Lingual Opinion Analysis</td>
<td>93</td>
</tr>
<tr>
<td><em>Lin Gui, Ruifeng Xu, Jun Xu, Li Yuan, Yuanlin Yao, Jiyun Zhou, Qiaoyun Qiu, Shuwei Wang, Kam-Fai Wong, and Ricky Cheung</em></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td><strong>Machine Learning for NLP</strong></td>
<td></td>
</tr>
<tr>
<td>Semi-supervised Text Categorization by Considering Sufficiency and</td>
<td>105</td>
</tr>
<tr>
<td>Diversity</td>
<td></td>
</tr>
<tr>
<td><em>Shoushan Li, Sophia Yat Mei Lee, Wei Gao, and Chu-Ren Huang</em></td>
<td></td>
</tr>
<tr>
<td>Pseudo In-Domain Data Selection from Large-Scale Web Corpus for</td>
<td>116</td>
</tr>
<tr>
<td>Spoken Language Translation</td>
<td></td>
</tr>
<tr>
<td><em>Shixiang Lu, Xingyuan Peng, Zhenbiao Chen, and Bo Xu</em></td>
<td></td>
</tr>
<tr>
<td>Discriminative Latent Variable Based Classifier for Translation Error</td>
<td>127</td>
</tr>
<tr>
<td>Detection</td>
<td></td>
</tr>
<tr>
<td><em>Jinhua Du, Junbo Guo, and Fei Zhao</em></td>
<td></td>
</tr>
<tr>
<td>Incorporating Entities in News Topic Modeling</td>
<td>139</td>
</tr>
<tr>
<td><em>Linmei Hu, Juanzi Li, Zhihui Li, Chao Shao, and Zhixing Li</em></td>
<td></td>
</tr>
<tr>
<td><strong>Machine Translation and Multi-Lingual Information Access</strong></td>
<td></td>
</tr>
<tr>
<td>An Efficient Framework to Extract Parallel Units from Comparable</td>
<td>151</td>
</tr>
<tr>
<td>Data</td>
<td></td>
</tr>
<tr>
<td><em>Lu Xiang, Yu Zhou, and Chengqing Zong</em></td>
<td></td>
</tr>
<tr>
<td>A Method to Construct Chinese-Japanese Named Entity Translation</td>
<td>164</td>
</tr>
<tr>
<td>Equivalents Using Monolingual Corpora</td>
<td></td>
</tr>
<tr>
<td><em>Kuang Ru, Jinan Xu, Yujie Zhang, and Peihao Wu</em></td>
<td></td>
</tr>
<tr>
<td>Collective Corpus Weighting and Phrase Scoring for SMT Using</td>
<td>176</td>
</tr>
<tr>
<td>Graph-Based Random Walk</td>
<td></td>
</tr>
<tr>
<td><em>Lei Cui, Dongdong Zhang, Shujie Liu, Mu Li, and Ming Zhou</em></td>
<td></td>
</tr>
<tr>
<td>A Simple, Fast Strategy for Weighted Alignment Hypergraph</td>
<td>188</td>
</tr>
<tr>
<td><em>Zhaopeng Tu, Jun Xie, Yajuan Lv, and Qun Liu</em></td>
<td></td>
</tr>
<tr>
<td><strong>NLP for Social Media and Web Mining</strong></td>
<td></td>
</tr>
<tr>
<td>Research on Building Family Networks Based on Bootstrapping and</td>
<td>200</td>
</tr>
<tr>
<td>Coreference Resolution</td>
<td></td>
</tr>
<tr>
<td><em>Jinghong Gu, Ya'nan Hu, Longhua Qian, and Qiaoming Zhu</em></td>
<td></td>
</tr>
<tr>
<td>Learning Sentence Representation for Emotion Classification on</td>
<td>212</td>
</tr>
<tr>
<td>Microblogs</td>
<td></td>
</tr>
<tr>
<td><em>Duoyu Tang, Bing Qin, Ting Liu, and Zhenghua Li</em></td>
<td></td>
</tr>
<tr>
<td>Every Term Has Sentiment: Learning from Emoticon Evidences for</td>
<td>224</td>
</tr>
<tr>
<td>Chinese Microblog Sentiment Analysis</td>
<td></td>
</tr>
<tr>
<td><em>Fei Jiang, Anqi Cui, Yiqun Liu, Min Zhang, and Shaoping Ma</em></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Active Learning for Cross-Lingual Sentiment Classification</td>
<td>236</td>
</tr>
<tr>
<td>Shoushan Li, Rong Wang, Huanhuan Liu, and Chu-Ren Huang</td>
<td></td>
</tr>
<tr>
<td>Expanding User Features with Social Relationships in Social</td>
<td>247</td>
</tr>
<tr>
<td>Recommender Systems</td>
<td></td>
</tr>
<tr>
<td>Chengjie Sun, Lei Lin, Yuan Chen, and Bingquan Liu</td>
<td></td>
</tr>
<tr>
<td><strong>Knowledge Acquisition</strong></td>
<td></td>
</tr>
<tr>
<td>A Hybrid Approach for Extending Ontology from Text</td>
<td>255</td>
</tr>
<tr>
<td>Wei He, Shuang Li, and Xiaoping Yang</td>
<td></td>
</tr>
<tr>
<td>Linking Entities in Short Texts Based on a Chinese Semantic</td>
<td>266</td>
</tr>
<tr>
<td>Knowledge Base</td>
<td></td>
</tr>
<tr>
<td>Yi Zeng, Dongsheng Wang, Tielin Zhang, Hao Wang, and Hongwei Hao</td>
<td></td>
</tr>
<tr>
<td>Entity Linking from Microblogs to Knowledge Base Using ListNet</td>
<td>277</td>
</tr>
<tr>
<td>Algorithm</td>
<td></td>
</tr>
<tr>
<td>Yan Wang, Cheng Luo, Xin Li, Yiqun Liu, Min Zhang, and Shaoping Ma</td>
<td></td>
</tr>
<tr>
<td><strong>NLP for Search Technology and Ads</strong></td>
<td></td>
</tr>
<tr>
<td>Automatic Assessment of Information Disclosure Quality in Chinese</td>
<td>288</td>
</tr>
<tr>
<td>Annual Reports</td>
<td></td>
</tr>
<tr>
<td>Xin Ying Qiu, Shengyi Jiang, and Kebin Deng</td>
<td></td>
</tr>
<tr>
<td>A Fast Matching Method Based on Semantic Similarity for Short</td>
<td>299</td>
</tr>
<tr>
<td>Texts</td>
<td></td>
</tr>
<tr>
<td>Jiaming Xu, Pengcheng Liu, Gaowei Wu, Zhengya Sun, Bo Xu, and Hongwei Hao</td>
<td></td>
</tr>
<tr>
<td>Query Generation Techniques for Patent Prior-Art Search in Multiple Languages</td>
<td>310</td>
</tr>
<tr>
<td>Dong Zhou, Jianxun Liu, and Sanrong Zhang</td>
<td></td>
</tr>
<tr>
<td>Improve Web Search Diversification with Intent Subtopic Mining</td>
<td>322</td>
</tr>
<tr>
<td>Aymeric Damien, Min Zhang, Yiqun Liu, and Shaoping Ma</td>
<td></td>
</tr>
<tr>
<td>Understanding Temporal Intent of User Query Based on Time-Based</td>
<td>334</td>
</tr>
<tr>
<td>Query Classification</td>
<td></td>
</tr>
<tr>
<td>Pengjie Ren, Zhumin Chen, Xiaomeng Song, Bin Li, Haopeng Yang,</td>
<td></td>
</tr>
<tr>
<td>and Jun Ma</td>
<td></td>
</tr>
</tbody>
</table>
## Short Papers I: NLP Fundamentals

Research on Semantic-Based Passive Transformation in Chinese-English Machine Translation

Wenfei Chang, Zhiying Liu, and Yaohong Jin

Research of an Improved Algorithm for Chinese Word Segmentation Dictionary Based on Double-Array Trie Tree

Wenchuan Yang, Jian Liu, and Miao Yu

Study on Tibetan Word Segmentation as Syllable Tagging

Yachao Li and Hongzhi Yu

The Spoken/Written Language Classification of English Sentences with Bilingual Information

Kuan Li, Zhongyang Xiong, Yufang Zhang, Xiaohua Liu, Ming Zhou, and Guanghua Zhang

## Short Papers II: NLP Applications

Design and Implementation of News-Oriented Automatic Summarization System Based on Chinese RSS

Jie Wang, Jie Ma, and Yingjun Li

Grey Relational Analysis for Query Expansion

Junjie Zou, Zhengtao Yu, Huanyun Zong, Jianyi Guo, and Lei Su

A Comprehensive Method for Text Summarization Based on Latent Semantic Analysis

Yingjie Wang and Jun Ma

## Short Papers III: NLP for Social Media (Industrial Paper Included)

A Time-Sensitive Model for Microblog Retrieval

Cunhui Shi, Bo Xu, Hongfei Lin, and Qing Guo

Feature Analysis in Microblog Retrieval Based on Learning to Rank

Zhongyuan Han, Xuwei Li, Muyun Yang, Haoliang Qi, and Sheng Li

Opinion Sentence Extraction and Sentiment Analysis for Chinese Microblogs

Hanxiao Shi, Wei Chen, and Xiaojun Li

Research on the Opinion Mining System for Massive Social Media Data

Lijun Zhao, Yingjie Ren, Ju Wang, Lingsheng Meng, and Cunlu Zou
Grammatical Phrase-Level Opinion Target Extraction on Chinese Microblog Messages .............................................. 432  
   Haochen Zhang, Yiqun Liu, Min Zhang, and Shaoping Ma

Simple Yet Effective Method for Entity Linking in Microblog-Genre Text ........................................................... 440  
   Qingliang Miao, Huayu Lu, Shu Zhang, and Yao Meng

Author Index .......................................................... 449