

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Alfred Kobsa

University of California, Irvine, CA, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

TU Dortmund University, Germany

Madhu Sudan

Microsoft Research, Cambridge, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbruecken, Germany

Yuexian Hou Jian-Yun Nie Le Sun
Bo Wang Peng Zhang (Eds.)

Information Retrieval Technology

8th Asia Information Retrieval
Societies Conference, AIRS 2012
Tianjin, China, December 17-19, 2012
Proceedings



Springer

Volume Editors

Yuexian Hou

Tianjin University, School of Computer Science and Technology
300072 Tianjin, China
E-mail: yxhou@tju.edu.cn

Jian-Yun Nie

University of Montreal, DIRO
CP. 6128, succursale Centre-ville, Montreal, H3C 3J7, QC, Canada
E-mail: nie@iro.umontreal.ca

Le Sun

Chinese Academy of Sciences
Institute of Software, Storage & Information Retrieval Laboratory
100190 Beijing, China
E-mail: sunle@iscas.ac.cn

Bo Wang

Tianjin University, School of Computer Science and Technology
300072 Tianjin, China
E-mail: bo.wang.1979@gmail.com

Peng Zhang

Robert Gordon University, School of Computing
St Andrew Street, Aberdeen, AB25 1HG, UK
E-mail: p.zhang1@rgu.ac.uk

ISSN 0302-9743

e-ISSN 1611-3349

ISBN 978-3-642-35340-6

e-ISBN 978-3-642-35341-3

DOI 10.1007/978-3-642-35341-3

Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012953221

CR Subject Classification (1998): H.3, H.4, F.2.2, I.4-5, E.1, H.2.8

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

© Springer-Verlag Berlin Heidelberg 2012

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

These proceedings contain the refereed papers and posters presented at the 8th Asia Information Retrieval Societies Conference (AIRS 2012). The conference was held during December 17–19, 2012, in Tianjin University (TJU), which is the oldest university in China, founded in 1895 as Peiyang University.

The annual AIRS conference has established its reputation as the main information retrieval (IR) conference for the Asia-Pacific region, where the research community and industry in IR have been growing rapidly. The conference aims to bring together researchers and developers to exchange new ideas and the latest achievements in the broad area of IR, covering topics on theories, systems, technologies, and applications of IR on both text and multimedia data.

This year, we received 77 submissions from Asian countries and all over the world; 22 (28.5%) of them were accepted as regular papers for oral presentations and 26 (33.7%) for poster presentations. All the papers went through a rigorous review process. Each paper was reviewed by three reviewers and a meta-reviewer.

The conference also featured two keynote speeches. Norbert Fuhr, the winner of the Salton Award of SIGIR in 2012, gave a keynote presentation on the calibrated values of probability of relevance and its application on vertical search, optimum clustering framework, and interactive IR. Haifeng Wang, a senior director of Baidu, gave a keynote speech on Web-based machine translation and cross-language IR. In addition, an industrial track was organized, in which several researchers from industry shared their technologies and experiences with the conference participants.

We would like to thank all the authors for submitting their research work to the conference. We are particularly grateful to the members of the Program Committee for their efforts in providing timely and high-quality reviews. We thank Tianjin University and Microsoft Research Asia for their financial support. We also thank the editorial staff of Springer for their assistance in publishing the conference proceedings as part of the *Lecture Notes in Computer Science* (LNCS). Finally, a special thanks goes to all the participants and student volunteers for their contributions and help in making AIRS 2012 a success.

December 2012

Yuexian Hou
Jian-Yun Nie
Le Sun
Bo Wang
Peng Zhang

Organization

AIRS 2012 was organized by Tianjin University, University of Montreal, Institute of Software and Institute of Computing Technology at the Chinese Academy of Sciences, and Robert Gordon University.

Conference Organizing Committee

General Co-chairs

Yuexian Hou	Tianjin University, China
Le Sun	Institute of Software, Chinese Academy of Sciences, China

Program Chair

Jian-Yun Nie	University of Montreal, Canada
--------------	--------------------------------

Local Chair

Ruiguo Yu	Tianjin University, China
-----------	---------------------------

Industry Track Chair

Bin Wang	Institute of Computing Technology, Chinese Academy of Sciences, China
----------	---

Publicity Chair

Bo Wang	Tianjin University, China
---------	---------------------------

Proceedings Chair

Peng Zhang	Robert Gordon University, UK
------------	------------------------------

Senior Program Committee

Wanxiang Che	Harbin Institute of Technology, China
Hsin-Hsi Chen	National Taiwan University, Taiwan
Jianfeng Gao	Microsoft Research, USA
Jimmy Huang	York University, USA
Joemon Jose	University of Glasgow, UK
Noriko Kando	National Institute of Informatics, Japan
Wai Lam	The Chinese University of Hong Kong, Hong Kong

Hang Li
Massimo Melucci
Ian Soboroff

Huawei Noah's Ark Lab, Hong Kong
University of Padua, Italy
National Institute of Standards and
Technology, USA
Tianjin University, China
Institute of Computing Technology, China
Baidu Inc., China
Soochow University, China

Program Committee

Lidong Bing

The Chinese University of Hong Kong,
Hong Kong

Fan Bu

Tsinghua University, China

Ben Carterette

University of Delaware, USA

Michael Chau

The University of Hong Kong, Hong Kong

Pu-Jen Cheng

National Taiwan University, Taiwan

Kevyn Collins-Thompson

Microsoft Research, USA

Alberto Costa

LIX, Ecole Polytechnique, France

Gerard De Melo

ICSI Berkeley, USA

Emanuele Di Buccio

University of Padua, Italy

Kevin Duh

Nara Institute of Science and Technology,
Japan

Georges Dupret

yahoo! Labs, USA

Ingo Frommholz

University of Bedfordshire, UK

Wei Gao

Qatar Computing Research Institute, Qatar

Shlomo Geva

Queensland University of Technology, Australia

Jiafeng Guo

Institute of Computing Technology, China

Cathal Gurrin

Dublin City University, Ireland

Martin Halvey

University of Glasgow, UK

Xianpei Han

ISCAS, China

Ben He

Graduate University of Chinese Academy of
Sciences, China

Ruifang He

Tianjin University, China

Xiaodong He

Microsoft Research, USA

Yulan He

The Open University, UK

Yu Hong

Soochow University, China

Yuexian Hou

Tianjin University, China

Xuanjing Huang

Fudan University, China

Jing Jiang

Singapore Management University, Singapore

Long Jiang

Microsoft Research Asia, China

Hideo Joho

University of Tsukuba, Japan

Gareth Jones

Dublin City University, Ireland

Leszek Kaliciak

Robert Gordon University, UK

Jaap Kamps

University of Amsterdam, The Netherlands

Kazuaki Kishida

Keio University, Japan

Kazuko Kuriyama	Shirayuri College, Japan
Oren Kurland	Technion, Israel Institute of Technology, Israel
Yanyan Lan	Chinese Academy of Sciences, China
Jun Lang	National University of Singapore, Singapore
Raymond Y.K. Lau	City University of Hong Kong, Hong Kong
Gary Lee	POSTECH, South Korea
Teerapong Leelanupab	University of Glasgow, UK
Shoushan Li	Soochow University, China
Wenjie Li	The Hong Kong Polytechnic University, Hong Kong
Zhenghua Li	Harbin Institute of Technology, China
Shasha Liao	New York University, USA
Fei Liu	University of Texas at Dallas, USA
Feifan Liu	University of Wisconsin-Milwaukee, USA
Jingjing Liu	MIT CSAIL, USA
Kang Liu	National Laboratory of Pattern Recognition, Institute of Automation, China
Xiaohua Liu	Microsoft Research Asia, China
Yang Liu	Tsinghua University, China
Yiqun Liu	Tsinghua University, China
Zhiyuan Liu	Tsinghua University, China
Hao Ma	Microsoft Research, USA
Andrew Macfarlane	City University London, UK
David McClosky	Stanford University, USA
Mandar Mitra	Indian Statistical Institute, India
Alistair Moffat	The University of Melbourne, Australia
Jian-Yun Nie	University of Montreal, Canada
Zhengyu Niu	Baidu Inc., China
Manabu Okumura	Tokyo Institute of Technology, Japan
Iadh Ounis	University of Glasgow, UK
Kristen Parton	Columbia University, USA
Tao Qin	Microsoft Research Asia, China
Vijay V. Raghavan	University of Louisiana at Lafayette, USA
Reede Ren	University of Glasgow, UK
Tetsuya Sakai	Microsoft Research Asia, China
Yohei Seki	University of Tsukuba, Japan
Jialie Shen	Singapore Management University, Singapore
Ruihua Song	Microsoft Research Asia, China
Yang Song	Microsoft Research, USA
Valentin Spitzkovsky	Stanford University, USA
Aixin Sun	Nanyang Technological University, Singapore
Mihai Surdeanu	Yahoo! Research Barcelona, Spain
Hisami Suzuki	Microsoft Research, USA
Paul Thomas	CSIRO, Australia
Andrew Trotman	University of Otago, New Zealand
Ming-Feng Tsai	National Chengchi University, Taiwan

Yuen-Hsien Tseng	National Taiwan Normal University, Taiwan
Stephen Wan	CSIRO, Australia
Xiaojun Wan	Peking University, China
Bo Wang	Tianjin University, China
Lei Wang	Robert Gordon University, UK
Mengqiu Wang	Stanford University, USA
Quan Wang	Peking University, China
William Webber	University of Maryland, USA
Tak-Lam Wong	The Hong Kong Institute of Education, Hong Kong
Mingfang Wu	RMIT University, Australia
Yunqing Xia	Tsinghua University, China
Wei Xiang	Baidu Inc., China
Deyi Xiong	Institute for Infocomm Research, Singapore
Jun Xu	Microsoft Research Asia, China
Xiaobing Xue	University of Massachusetts, USA
Zheng Ye	York University, USA
Xing Yi	University of Massachusetts, Amherst, USA
Meishan Zhang	Harbin Institute of Technology, China
Min Zhang	Tsinghua University, China
Peng Zhang	Robert Gordon University, UK
Qi Zhang	Fudan University, China
Zhichang Zhang	Northwest Normal University, China
Hai Zhao	Shanghai Jiao Tong University, China
Jun Zhao	Institute of Automation, China

Sponsors

Tianjin University
University of Montreal
Microsoft Research Asia
Chinese Information Processing Society of China
ACM SIGIR
Springer



The Return of the Probability of Relevance

(Invited Talk)

Norbert Fuhr

University of Duisburg-Essen, Germany

`norbert.fuhr@uni-due.de`

The probability ranking principle (PRP) [6] proves that ranking documents by decreasing probability of relevance yields optimum retrieval quality. Most research on probabilistic models has focused only on producing a probabilistic ranking, without estimating the actual probabilities. For classical IR applications, however, knowing these parameters would e.g. allow for estimating the quality of a query result, or for determining the best cutoff point for document filtering.

This talk discusses models for three types of modern IR applications which rely on calibrated values of the probability of relevance.

In vertical search, there are several collections of different type or media (like e.g. Web pages, news, tweets, videos, images). Given a query, the system has to select the collections containing potentially relevant items, process the query on these collections and then merge the results. The model presented in [2] describes the optimum solution to the selection problem, which is based on the probabilistic estimation of the number of relevant documents per resource; the subsequent merging step is straightforward in case the probabilities of relevance are available.

The optimum clustering framework [4] provides not only the first theoretic foundation for document clustering, it also proves the clustering hypothesis. Its key idea is to base cluster analysis and evaluation on a set of queries, by defining documents as being similar if they are relevant to the same queries. Given a set of queries, a probabilistic retrieval method estimates of the relevance probability for all possible query-document pairs, after which a document similarity metric is applied. It can be shown that more or less all clustering methods are implicitly based on these three components, but that they use heuristic design decisions for some of them.

The interactive PRP [3] generalizes the classical PRP for interactive retrieval. It characterizes interactive retrieval as a sequence of situations, where, in each situation, the user is confronted with a list of choices. Each choice is described by three parameters, namely the effort for evaluating it, the probability that the user will accept it, and the benefit resulting from acceptance. While the user effort and expected benefit can be derived from appropriate user studies involving e.g. gaze tracking [7], the acceptance probabilities depend on the quality of the choices presented by the system. The probabilities of relevance of the documents in the result list are just one example of these parameters, other choice lists for which the acceptance parameters have to be estimated are e.g. lists of query expansion terms or clickthrough rates for result list items. Once these parameters

are available, the expected benefit of a choice can e.g. be estimated as the time saved for reaching the search goal.

For estimating the actual probabilities of relevance required for applying these models, there are three general methods.

1. Direct estimation of the relevance probabilities from the underlying model is very difficult, since it requires the estimation of other parameters (like e.g. the generality of the current query) for which no appropriate observation data exists.
2. Some 'learning to rank' methods [5] (like e.g. logistic regression) aim at a direct estimation of the probability of relevance.
3. Score distribution models [1] regard the distribution of retrieval scores in relevant and nonrelevant documents, from which appropriate transformations onto the probability of relevance can be derived.

By combining these estimation methods with the models described above, it becomes possible to implement approaches based on solid theoretic foundations, which are more transparent than heuristic approaches, thus allowing for theory-guided adaptation and tuning.

References

1. Arampatzis, A., Robertson, S.: Modeling score distributions in information retrieval. *Inf. Retr.* 14(1), 26–46 (2011)
2. Fuhr, N.: A decision-theoretic approach to database selection in networked IR. *ACM Transactions on Information Systems* 17(3), 229–249 (1999)
3. Fuhr, N.: A probability ranking principle for interactive information retrieval. *Information Retrieval* 11(3), 251–265 (2008), <http://dx.doi.org/10.1007/s10791-008-9045-0>
4. Fuhr, N., Lechtenfeld, M., Stein, B., Gollub, T.: The optimum clustering framework: Implementing the cluster hypothesis. *Information Retrieval* 15, 93–115 (2012), doi:10.1007/s10791-011-9173-9
5. Liu, T.-Y.: *Learning to Rank for Information Retrieval*. Springer (2011)
6. Robertson, S.E.: The probability ranking principle in IR. *Journal of Documentation* 33, 294–304 (1977)
7. Tran, V.T., Fuhr, N.: Using eye-tracking with dynamic areas of interest for analyzing interactive information retrieval. In: Hersh, W.R., Callan, J., Maarek, Y., Sanderson, M. (eds.) *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1165–1166. ACM (2012)

Web-Based Machine Translation and Cross-Language Information Retrieval (Invited Talk)

Haifeng Wang

Baidu Campus,
No. 10 Shangdi 10th Street,
Beijing, 100085, China
wanghaifeng@baidu.com

Machine translation (MT) aims to translate text or speech from one language to another, while Cross-Language Information Retrieval (CLIR) tries to retrieve documents in a language different from that of the query. These two are well connected. A CLIR system has the potential to extend the searchable information from a single language to multiple languages. MT system enables users to understand the essential contents of a retrieved document in a foreign language and it makes it easier to implement CLIR by adding an MT system for query translation on top of an IR system. There are many potential web-based applications for MT and CLIR, such as translation of web-page, translation of instant messages, cross-language search, etc. On the one hand, more web data can be crawled to train these systems. Both the scale and quality of the web data poses major challenges to web-based MT and CLIR applications. In order to build better web-based applications, it is essential to develop techniques to adapt state-of-the-art MT and CLIR technologies to the web scenario. In this talk, I will introduce our work on web-based machine translation and cross-language information retrieval.

Table of Contents

Session 1: IR Models

Exploring and Exploiting Proximity Statistic for Information Retrieval Model	1
<i>Yadong Zhu, Yuanhai Xue, Jiafeng Guo, Yanyan Lan, Xueqi Cheng, and Xiaoming Yu</i>	

A Category-integrated Language Model for Question Retrieval in Community Question Answering	14
<i>Zongcheng Ji, Fei Xu, and Bin Wang</i>	

Session 2: Evaluation and User Studies

The Reusability of a Diversified Search Test Collection	26
<i>Tetsuya Sakai, Zhicheng Dou, Ruihua Song, and Noriko Kando</i>	

One Click One Revisited: Enhancing Evaluation Based on Information Units	39
<i>Tetsuya Sakai and Makoto P. Kato</i>	

A Comparison of Action Transitions in Individual and Collaborative Exploratory Web Search	52
<i>Zhen Yue, Shuguang Han, and Daqing He</i>	

Session 3: NLP for IR

Query-Oriented Keyphrase Extraction	64
<i>Minghui Qiu, Yaliang Li, and Jing Jiang</i>	

Using Lexical and Thematic Knowledge for Name Disambiguation	76
<i>Jinpeng Wang, Wayne Xin Zhao, Rui Yan, Haitian Wei, Jian-Yun Nie, and Xiaoming Li</i>	

Sells Out or Piles Up? A Sentiment Autoregressive Model for Predicting Sales Performance	89
<i>Xueni Li, Shaowu Zhang, Liang Yang, and Hongfei Lin</i>	

Session 4: Machine Learning and Data Mining

Are Human-Input Seeds Good Enough for Entity Set Expansion? Seeds Rewriting by Leveraging Wikipedia Semantic Knowledge	103
<i>Zhenyu Qi, Kang Liu, and Jun Zhao</i>	

Combining Modifications to Multinomial Naive Bayes for Text
Classification 114
Antti Puurula

Organizing Information on the Web through Agreement-Conflict
Relation Classification 126
Junta Mizuno, Eric Nichols, Yotaro Watanabe, and Kentaro Inui

Session 5: Social Media

Exploiting Twitter for Spiking Query Classification 138
Mitsuo Yoshida and Yuki Arase

Where Are You Settling Down: Geo-locating Twitter Users Based on
Tweets and Social Networks 150
Kejiang Ren, Shaowu Zhang, and Hongfei Lin

Detecting Informative Messages Based on User History in Twitter 162
*Chang-Woo Chun, Jung-Tae Lee, Seung-Wook Lee, and
Hae-Chang Rim*

Session 6: IR Applications

A Study on Potential Head Advertisers in Sponsored Search 174
Changhao Jiang, Min Zhang, Bin Gao, and Tie-Yan Liu

Feature Transformation Method Enhanced Vandalism Detection
in Wikipedia 187
Tianshu Chang, Hongfei Lin, and Yuan Lin

PLIDMiner: A Quality Based Approach for Researcher’s Homepage
Discovery 199
Junting Ye, Yanan Qian, and Qinghua Zheng

Session 7: Multimedia IR and Indexing

Exploiting and Exploring Hierarchical Structure in Music
Recommendation 211
Kai Lu, Guanyuan Zhang, Rui Li, Shuai Zhang, and Bin Wang

Enhancing Music Information Retrieval by Incorporating Image-Based
Local Features 226
*Leszek Kaliciak, Ben Horsburgh, Dawei Song,
Nirmalie Wiratunga, and Jeff Pan*

Yet Another Sorting-Based Solution to the Reassignment of Document
Identifiers 238
Liang Shi and Bin Wang

Session 8: Collaborative and Federated Search

- What Snippets Say about Pages in Federated Web Search 250
*Thomas Demeester, Dong Nguyen, Dolf Trieschnigg,
 Chris Develder, and Djoerd Hiemstra*
- CoFox: A Synchronous Collaborative Browser 262
Jesus Rodriguez Perez, Teerapong Leelanupab, and Joemon M. Jose

Poster Session

- FISER: An Effective Method for Detecting Interactions between Topic
 Persons 275
*Yung-Chun Chang, Pi-Hua Chuang, Chien Chin Chen, and
 Wen-Lian Hsu*
- Cross-Lingual Knowledge Discovery: Chinese-to-English Article Linking
 in Wikipedia 286
Ling-Xiang Tang, Andrew Trotman, Shlomo Geva, and Yue Xu
- Unified Recommendation and Search in E-Commerce 296
Jian Wang, Yi Zhang, and Tao Chen
- Identifying Local Questions in Community Question Answering 306
Long Chen, Dell Zhang, and Mark Levene
- Improving Content-Based Image Retrieval by Identifying Least and
 Most Correlated Visual Words 316
Leszek Kaliciak, Dawei Song, Nirmalie Wiratunga, and Jeff Pan
- Topical Relevance Model 326
Debasis Ganguly, Johannes Leveling, and Gareth J.F. Jones
- Adaptive Data Fusion Methods for Dynamic Search Environments 336
Shengli Wu, Yuping Xing, Jieyu Li, and Yaxin Bi
- NMiner: A System for Finding Related Entities by Mining a Bimodal
 Network 346
*VenkataSwamy Martha, Stephen Wallace, Halil Bisgin, Xiaowei Xu,
 Nitin Agarwal, and Hemant Joshi*
- Combining Signals for Cross-Lingual Relevance Feedback 356
Kristen Parton and Jianfeng Gao
- A Conceptual Model for Retrieval of Chinese Frequently Asked
 Questions in Healthcare 366
Rey-Long Liu and Shu-Ling Lin

LDA-Based Topic Formation and Topic-Sentence Reinforcement for Graph-Based Multi-document Summarization	376
<i>Dehong Gao, Wenjie Li, You Ouyang, and Renxian Zhang</i>	
Query Reformulation Based on User Habits for Query-by-Humming Systems	386
<i>Guanyuan Zhang, Kai Lu, and Bin Wang</i>	
Query Expansion Based on User Quality in Folksonomy	396
<i>Qing Guo, Wenfei Liu, Yuan Lin, and Hongfei Lin</i>	
Semantically Enriched Clustered User Interest Profile Built from Users' Tweets	406
<i>Harshit Kumar and Hong-Gee Kim</i>	
Adaptive Weighting Approach to Context-Sensitive Retrieval Model	417
<i>Xiaochun Wang, Muyun Yang, Haoliang Qi, Sheng Li, and Tiejun Zhao</i>	
Effective Keyword-Based XML Retrieval Using the Data-Centric and Document-Centric Features	427
<i>Tsubasa Tanabe, Toshiyuki Shimizu, and Masatoshi Yoshikawa</i>	
Unsupervised Extraction of Popular Product Attributes from Web Sites	437
<i>Lidong Bing, Tak-Lam Wong, and Wai Lam</i>	
Aggregating Opinions on Hot Topics from Microblog Responses	447
<i>Han-Chih Liu and Jenq-Haur Wang</i>	
Topic Sequence Kernel	457
<i>Jian Xu, Qin Lu, Zhengzhong Liu, and Junyi Chai</i>	
Opening Machine Translation Black Box for Cross-Language Information Retrieval	467
<i>Yanjun Ma, Jian-Yun Nie, Hua Wu, and Haifeng Wang</i>	
Topic Based Author Ranking with Full-Text Citation Analysis	477
<i>Jinsong Zhang, Chun Guo, and Xiaozhong Liu</i>	
Vertical Classification of Web Pages for Structured Data Extraction	486
<i>Long Li, Dandan Song, and Lejian Liao</i>	
Grid-Based Interaction for Exploratory Search	496
<i>Hideo Joho and Tetsuya Sakai</i>	
Analyzing the Spatiotemporal Effects on Detection of Rain Event Duration	506
<i>Jyun-Yu Jiang, Yi-Shiang Tzeng, Pei-Ying Huang, and Pu-Jen Cheng</i>	

Outpatient Department Recommendation Based on Medical
Summaries 518
Hen-Hsen Huang, Chia-Chun Lee, and Hsin-Hsi Chen

Actively Mining Search Logs for Diverse Tags 528
Lun Yan, Congrui Huang, and Yan Zhang

Author Index 539