

Contents

Efficiency

On Structures of Inverted Index for Query Processing Efficiency	3
<i>Xingshen Song, Xueping Zhang, Yuexiang Yang, Jicheng Quan, and Kun Jiang</i>	
Access Time Tradeoffs in Archive Compression	15
<i>Matthias Petri, Alistair Moffat, P.C. Nagesh, and Anthony Wirth</i>	
Large Scale Sentiment Analysis with Locality Sensitive BitHash.	29
<i>Wenhao Zhang, Jianqiu Ji, Jun Zhu, Hua Xu, and Bo Zhang</i>	

Graphs, Knowledge Bases and Taxonomies

Knowledge-Based Query Expansion in Real-Time Microblog Search.	43
<i>Chao Lv, Runwei Qiang, Feifan Fan, and Jianwu Yang</i>	
Enrichment of Academic Search Engine Results Pages by Citation-Based Graphs.	56
<i>Shuhei Shogen, Toshiyuki Shimizu, and Masatoshi Yoshikawa</i>	
Quality-Aware Review Selection Based on Product Feature Taxonomy	68
<i>Nan Tian, Yue Xu, Yuefeng Li, and Gabriella Pasi</i>	

Recommendations

An Author Subject Topic Model for Expert Recommendation.	83
<i>Haikun Mou, Qian Geng, Jian Jin, and Chong Chen</i>	
A Keyword Recommendation Method Using CorKeD Words and Its Application to Earth Science Data	96
<i>Youichi Ishida, Toshiyuki Shimizu, and Masatoshi Yoshikawa</i>	
Incorporating Distinct Opinions in Content Recommender System.	109
<i>Grace E. Lee, Keejun Han, and Mun Y. Yi</i>	

Twitter and Social Media

Detecting Automatically-Generated Arabic Tweets	123
<i>Hind Almerkhi and Tamer Elsayed</i>	

Improving Tweet Timeline Generation by Predicting Optimal Retrieval Depth. 135
Maram Hasanain, Tamer Elsayed, and Walid Magdy

Abstract Venue Concept Detection from Location-Based Social Networks . . . 147
Yi Liao, Shoaib Jameel, Wai Lam, and Xing Xie

Web Search

Snapboard: A Shared Space of Visual Snippets - A Study in Individual and Asynchronous Collaborative Web Search. 161
Teerapong Leelanupab, Hannarin Kruajirayu, and Nont Kanungsukkasem

Improving Ranking and Robustness of Search Systems by Exploiting the Popularity of Documents 174
Ashraf Bah and Ben Carterette

Heading-Aware Snippet Generation for Web Search 188
Tomohiro Manabe and Keishi Tajima

Text Processing, Understanding and Categorization

Smoothing Temporal Difference for Text Categorization 203
Fumiyo Fukumoto and Yoshimi Suzuki

Charset Encoding Detection of HTML Documents: A Practical Experience 215
Shabanali Faghani, Ali Hadian, and Behrouz Minaei-Bidgoli

Structure Matters: Adoption of Structured Classification Approach in the Context of Cognitive Presence Classification 227
Zak Waters, Vitomir Kovanović, Kirsty Kitto, and Dragan Gašević

Topics and Models

A Sequential Latent Topic-Based Readability Model for Domain-Specific Information Retrieval. 241
Wenya Zhang, Dawei Song, Peng Zhang, Xiaozhao Zhao, and Yuexian Hou

Automatic Labelling of Topic Models Using Word Vectors and Letter Trigram Vectors 253
Wanqiu Kou, Fang Li, and Timothy Baldwin

A Study of Collection-Based Features for Adapting the Balance Parameter in Pseudo Relevance Feedback 265
Ye Meng, Peng Zhang, Dawei Song, and Yuexian Hou

Clustering

An MDL-Based Frequent Itemset Hierarchical Clustering Technique to Improve Query Search Results of an Individual Search Engine 279
Diyah Puspitaningrum, Fauzi, Boko Susilo, Jeri Apriansyah Pagua, Aan Erlansari, Desi Andreswari, Rusdi Efendi, and I.S.W.B. Prasetya

Improving Clustering Quality by Automatic Text Summarization 292
Mohsen Pourvali, Salvatore Orlando, and Mehrad Gharagozloo

Tweet Timeline Generation via Graph-Based Dynamic Greedy Clustering . . . 304
Feifan Fan, Runwei Qiang, Chao Lv, Wayne Xin Zhao, and Jianwu Yang

Evaluation

Topic Set Size Design with the Evaluation Measures for Short Text Conversation 319
Tetsuya Sakai, Lifeng Shang, Zhengdong Lu, and Hang Li

Towards Nuanced System Evaluation Based on Implicit User Expectations . . . 332
Paul Thomas, Peter Bailey, Alistair Moffat, and Falk Scholer

Social Media and Recommendation

A Study of Visual and Semantic Similarity for Social Image Search Recommendation 347
Yangjie Yao and Aixin Sun

Company Name Disambiguation in Tweets: A Two-Step Filtering Approach 358
M. Atif Qureshi, Arjumand Younus, Colm O’Riordan, and Gabriella Pasi

Utilizing Word Embeddings for Result Diversification in Tweet Search 366
Kezban Dilek Onal, Ismail Sengor Altingovde, and Pinar Karagoz

Short Papers

Displaying People with Old Addresses on a Map 381
Gang Zhang and Harumi Murakami

A Query Expansion Approach Using Entity Distribution Based on Markov Random Fields 387
Rui Li, Linxue Hao, Xiaozhao Zhao, Peng Zhang, Dawei Song, and Yuexian Hou

Analysis of Cyber Army’s Behaviours on Web Forum for Elect Campaign . . .	394
<i>Man-Chun Ko and Hsin-Hsi Chen</i>	
Beyond tf-idf and Cosine Distance in Documents Dissimilarity Measure	400
<i>Sunil Aryal, Kai Ming Ting, Gholamreza Haffari, and Takashi Washio</i>	
Explorations of Cross-Disciplinary Term Similarity	407
<i>Hanif Sheikhabolkarim and Laurianne Sitbon</i>	
A Signature Approach to Patent Classification	413
<i>Dilesha Seneviratne, Shlomo Geva, Guido Zuccon, Gabriela Ferraro, Timothy Chappell, and Magali Meireles</i>	
The Impact of Using Combinatorial Optimisation for Static Caching of Posting Lists.	420
<i>Casper Petersen, Jakob Grue Simonsen, and Christina Lioma</i>	
Improving Difficult Queries by Leveraging Clusters in Term Graph	426
<i>Rajul Anand and Alexander Kotov</i>	
Demonstrations	
EEST: Entity-Driven Exploratory Search for Twitter	435
<i>Chao Lv, Runwei Qiang, Lili Yao, and Jianwu Yang</i>	
Oyster: A Tool for Fine-Grained Ontological Annotations in Free-Text . . .	440
<i>Hamed Tayebikhorami, Alejandro Metke-Jimenez, Anthony Nguyen, and Guido Zuccon</i>	
TopSig: A Scalable System for Hashing and Retrieving Document Signatures.	447
<i>Timothy Chappell and Shlomo Geva</i>	
Author Index	453



<http://www.springer.com/978-3-319-28939-7>

Information Retrieval Technology

11th Asia Information Retrieval Societies Conference,
AIRS 2015, Brisbane, QLD, Australia, December 2-4,
2015. Proceedings

Zuccon, G.; Geva, S.; Joho, H.; Scholer, F.; Sun, A.;
Zhang, P. (Eds.)

2015, XVIII, 454 p. 84 illus. in color., Softcover

ISBN: 978-3-319-28939-7