More information about this series at http://www.springer.com/series/7409
Database Systems for Advanced Applications

20th International Conference, DASFAA 2015
Hanoi, Vietnam, April 20–23, 2015
Proceedings, Part I
Preface

It is our great pleasure to welcome you to DASFAA 2015, the 20th edition of the International Conference on Database Systems for Advanced Applications (DASFAA 2015), which was held in Hanoi, Vietnam during April 20–23, 2015. Hanoi (Vietnamese: Hà Nội), the capital of Vietnam, is the second largest city in Vietnam and has collected all the essence, unique features, and diversification of Vietnamese culture. The city is preserving more than 4000 historical and cultural relics, architecture and beauty spots, in which nearly 900 relics have been nationally ranked with hundreds of pagodas, temples, architectural works, and sceneries. Handcraft streets and traditional handcraft villages remain prominent and attractive to tourists when visiting Hanoi, many of which centered around the Hoan Kiem Lake in the Old Quarter, close to the conference venue. Hanoi has recently been included on TripAdvisor’s list of best destinations in the world, ranked 8th among the world’s top 25 destinations.

We are delighted to offer an exciting technical program, including two keynote talks by Amr El Abbadi (University of California, Santa Barbara) and Luc Vincent (Google Inc.); one 10-year best paper award presentation; a panel session on “Big Data Search and Analysis;” a poster session with 18 papers; a demo session with 6 demonstrations; an industry session with 3 full paper presentations; 3 tutorial sessions; and of course a superb set of research papers. This year, we received 287 submissions, each of which went through a rigorous review process. That is, each paper was reviewed by at least three Program Committee members, followed by a discussion led by the meta-reviewers, and a final meta-review prepared for each paper. At the end, DASFA 2015 accepted 63 full papers (the acceptance ratio is 22%).

Two workshops were selected by the Workshop Co-chairs to be held in conjunction with DASFAA 2015. They are the Second International Workshop on Big Data Management and Service (BDMS 2015), and the Second International Workshop on Semantic Computing and Personalization (SeCoP 2015). The workshop papers are included in a separate volume of proceedings also published by Springer in its Lecture Notes in Computer Science series.

The conference received generous financial support from the Hanoi University of Science and Technology (HUST). We, the conference organizers, also received extensive help and logistic support from the DASFAA Steering Committee and the Conference Management Toolkit Support Team at Microsoft.

We are grateful to all conference organizers, Han Su (University of Queensland) and many other people, for their great effort in supporting conference organization. Special thanks also go to the DASFAA 2015 Local Organizing Committee: Tuyet-Trinh Vu, Hong-Phuong Nguyen, and Van Thu Truong, all from the Hanoi University of Science
and Technology, Vietnam. Finally, we would like to take this opportunity to thank all the meta-reviewers, Program Committee members, and external reviewers for their expertise and help in evaluating papers, and all the authors who submitted their papers to this conference.

February 2015

Quyet-Thang Huynh
Qing Li
Matthias Renz
Cyrus Shahabi
Xiaofang Zhou
Organization

**General Co-chairs**

Qing Li  
City University of Hong Kong, HKSAR, Hong Kong

Quyet-Thang Huynh  
Hanoi University of Science and Technology, Vietnam

**Program Committee Co-chairs**

Cyrus Shahabi  
University of Southern California, USA

Matthias Renz  
Ludwig-Maximilians-Universität München, Germany

Xiaofang Zhou  
University of Queensland, Australia

**Tutorial Co-chairs**

Arbee L.P. Chen  
NCCU, Taiwan

Pierre Senellart  
Télécom ParisTech, France

**Workshops Co-chairs**

An Liu  
Soochow University, China

Yoshiharu Ishikawa  
Nagoya University, Japan

**Demo Co-chairs**

Haiwei Pan  
Harbin Engineering University, China

Binh Minh Nguyen  
Hanoi University of Science and Technology, Vietnam

**Panel Co-chairs**

Bin Cui  
Peking University, China

Katsumi Tanaka  
Kyoto University, Japan

**Poster Co-chairs**

Sarana Nutanong  
City University of Hong Kong, China

Tieyun Qian  
Wuhan University, China
**Industrial/Practitioners Track Co-chairs**

Mukesh Mohania  IBM, India  
Khai Tran  Oracle, USA

**PhD Colloquium**

Khoat Than  Hanoi University of Science and Technology, Vietnam 
Ge Yu  Northeastern University, China  
Tok Wang Ling  National University of Singapore, Singapore 
Duong Nguyen Vu  John Von Neumann Institute - VNU-HCMUS, Vietnam

**Publication Chair**

Muhammad Aamir Cheema  Monash University, Australia

**Publicity Co-chairs**

Yunjun Gao  Zhejiang University, China  
Bao-Quoc Ho  VNU-HCMUS, Vietnam  
Jianfeng Si  Institute for Infocomm Research, Singapore  
Wen-Chih Peng  National Chiao Tung University, Taiwan

**Local Organizing Committee**

Tuyet-Trinh Vu  Hanoi University of Science and Technology, Vietnam 
Hong-Phuong Nguyen  Hanoi University of Science and Technology, Vietnam  
Van Thu Truong  Hanoi University of Science and Technology, Vietnam

**Steering Committee Liaison**

Stephane Bressan  National University of Singapore, Singapore

**Webmaster**

Viet-Trung Tran  Hanoi University of Science and Technology, Vietnam
Program Committees

Senior PC members

Ira Assent  Aarhus University, Denmark
Lei Chen  Hong Kong University of Science and Technology (HKUST), China
Reynold Cheng  University of Hong Kong, China
Gabriel Ghinita  University of Massachusetts Boston, USA
Panos Kalnis  King Abdullah University of Science and Technology, Saudi Arabia
Nikos Mamoulis  University of Hong Kong, China
Kyriakos Mouratidis  Singapore Management University, Singapore
Mario Nascimento  University of Alberta, Canada
Dimitris Papadias  Hong Kong University of Science and Technology (HKUST), China
Stavros Papadopoulos  MIT, USA
Torben Bach Pedersen  Aalborg University, Denmark
Jian Pei  Simon Fraser University, Canada
Thomas Seidl  RWTH Aachen University, Germany
Timos Sellis  RMIT University, Australia
Raymond Wong  Hong Kong University of Science and Technology (HKUST), China

PC Members

Nikolaus Augsten  University of Salzburg, Austria
Spiridon Bakiras  City University of New York, USA
Zhifeng Bao  University of Tasmania, Australia
Srikanta Bedathur  IBM Research, Delhi, India
Ladjel Bellatreche  University of Poitiers, France
Boualem Benatallah  University of New South Wales, Australia
Bin Cui  Peking University, China
Athman Bouguettaya  Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia
Panagiotis Bouros  Humboldt-Universität zu Berlin, Germany
Selcuk Candan  Arizona State University, USA
Jianneng Cao  A*STAR, Singapore
Marco Casanova  Pontifical Catholic University of Rio de Janeiro, Brazil
Sharma Chakravarthy  University of Texas at Arlington, USA
Jae Chang  Chonbuk National University, Korea
Rui Chen  Hong Kong Baptist University, China
Yi Chen  New Jersey Institute of Technology, USA
James Cheng  The Chinese University of Hong Kong (CUHK), China
Gao Cong  Nanyang Technological University (NTU), Singapore
<table>
<thead>
<tr>
<th>Name</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ugur Demiryurek</td>
<td>University of Southern California (USC), USA</td>
</tr>
<tr>
<td>Prasad Deshpande</td>
<td>IBM Research, India</td>
</tr>
<tr>
<td>Gill Dobbie</td>
<td>University of Auckland, New Zealand</td>
</tr>
<tr>
<td>Eduard Dragut</td>
<td>Temple University, USA</td>
</tr>
<tr>
<td>Cristina Dutra de Aguiar Ciferri</td>
<td>Universidade de São Paulo, Brazil</td>
</tr>
<tr>
<td>Sameh Elnikety</td>
<td>Microsoft Research Redmond, USA</td>
</tr>
<tr>
<td>Tobias Emrich</td>
<td>Ludwig-Maximilians-Universität München, Germany</td>
</tr>
<tr>
<td>Johann Gamper</td>
<td>Free University of Bozen-Bolzano, Italy</td>
</tr>
<tr>
<td>Xin Gao</td>
<td>King Abdullah University of Science and Technology (KAUST), Saudi Arabia</td>
</tr>
<tr>
<td>Chenjuan Guo</td>
<td>Aarhus University, Denmark</td>
</tr>
<tr>
<td>Ralf Hartmut Güting</td>
<td>University of Hagen, Germany</td>
</tr>
<tr>
<td>Takahiro Hara</td>
<td>Osaka University, Japan</td>
</tr>
<tr>
<td>Haibo Hu</td>
<td>Hong Kong Baptist University, China</td>
</tr>
<tr>
<td>Yoshiharu Ishikawa</td>
<td>Nagoya University, Japan</td>
</tr>
<tr>
<td>Mizuho Iwaihara</td>
<td>Waseda University, Japan</td>
</tr>
<tr>
<td>Adam Jatowt</td>
<td>Kyoto University, Japan</td>
</tr>
<tr>
<td>Vana Kalogeraki</td>
<td>Athens University of Economy and Business, Greece</td>
</tr>
<tr>
<td>Panos Karras</td>
<td>Skoltech, Russia</td>
</tr>
<tr>
<td>Norio Katayama</td>
<td>National Institute of Informatics, Japan</td>
</tr>
<tr>
<td>Sang-Wook Kim</td>
<td>Hanyang University, Korea</td>
</tr>
<tr>
<td>Seon Ho Kim</td>
<td>University of Southern California (USC), USA</td>
</tr>
<tr>
<td>Hiroyuki Kitagawa</td>
<td>University of Tsukuba, Japan</td>
</tr>
<tr>
<td>Peer Kröger</td>
<td>Ludwig-Maximilians-Universität München, Germany</td>
</tr>
<tr>
<td>Jae-Gil Lee</td>
<td>Korea Advanced Institute of Science and Technology (KAIST), Korea</td>
</tr>
<tr>
<td>Wang-Chien Lee</td>
<td>Portland State University (PSU), USA</td>
</tr>
<tr>
<td>Sang-Goo Lee</td>
<td>Seoul National University, Korea</td>
</tr>
<tr>
<td>Hou Leong</td>
<td>University of Macau, China</td>
</tr>
<tr>
<td>Guoliang Li</td>
<td>Tsinghua University, China</td>
</tr>
<tr>
<td>Hui Li</td>
<td>Xidian University, China</td>
</tr>
<tr>
<td>Xiang Lian</td>
<td>University of Texas–Pan American (UTPA), USA</td>
</tr>
<tr>
<td>Lipyeow Lim</td>
<td>University of Hawaii, USA</td>
</tr>
<tr>
<td>Sebastian Link</td>
<td>University of Auckland, New Zealand</td>
</tr>
<tr>
<td>Bin Liu</td>
<td>NEC Laboratories, USA</td>
</tr>
<tr>
<td>Changbin Liu</td>
<td>AT &amp; T, USA</td>
</tr>
<tr>
<td>Eric Lo</td>
<td>Hong Kong Polytechnic University, China</td>
</tr>
<tr>
<td>Jiaheng Lu</td>
<td>Renmin University of China, China</td>
</tr>
<tr>
<td>Qiong Luo</td>
<td>Hong Kong University of Science and Technology (HKUST), China</td>
</tr>
<tr>
<td>Matteo Magnani</td>
<td>Uppsala University, Sweden</td>
</tr>
<tr>
<td>Silviu Maniu</td>
<td>University of Hong Kong (HKU), China</td>
</tr>
<tr>
<td>Essam Mansour</td>
<td>Qatar Computing Research Institute, Qatar</td>
</tr>
<tr>
<td>Name</td>
<td>Institution</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>Marco Mesiti</td>
<td>University of Milan, Italy</td>
</tr>
<tr>
<td>Yasuhiiko Morimoto</td>
<td>Hiroshima University, Japan</td>
</tr>
<tr>
<td>Wilfred Ng</td>
<td>Hong Kong University of Science and Technology (HKUST), China</td>
</tr>
<tr>
<td>Makoto Onizuka</td>
<td>Osaka University, Japan</td>
</tr>
<tr>
<td>Balaji Palanisamy</td>
<td>University of Pittsburgh, USA</td>
</tr>
<tr>
<td>Stefano Paraboschi</td>
<td>Università degli Studi di Bergamo, Italy</td>
</tr>
<tr>
<td>Sanghyun Park</td>
<td>Yonsei University, Korea</td>
</tr>
<tr>
<td>Dhaival Patel</td>
<td>IIT Roorkee, India</td>
</tr>
<tr>
<td>Evaggelia Pitoura</td>
<td>University of Ioannina, Greece</td>
</tr>
<tr>
<td>Pascal Poncelet</td>
<td>Université Montpellier 2, France</td>
</tr>
<tr>
<td>Maya Ramanath</td>
<td>Indian Institute of Technology, New Delhi, India</td>
</tr>
<tr>
<td>Shazia Sadiq</td>
<td>University of Queensland, Australia</td>
</tr>
<tr>
<td>Sherif Sakr</td>
<td>University of New South Wales, Australia</td>
</tr>
<tr>
<td>Kai-Uwe Sattler</td>
<td>Ilmenau University of Technology, Germany</td>
</tr>
<tr>
<td>Peter Scheuermann</td>
<td>Northwestern University, USA</td>
</tr>
<tr>
<td>Markus Schneider</td>
<td>University of Florida, USA</td>
</tr>
<tr>
<td>Matthias Schubert</td>
<td>Ludwig-Maximilians-Universität München, Germany</td>
</tr>
<tr>
<td>Shuo Shang</td>
<td>China University of Petroleum, Beijing, China</td>
</tr>
<tr>
<td>Kyuseok Shim</td>
<td>Seoul National University, Korea</td>
</tr>
<tr>
<td>Junho Shim</td>
<td>Sookmyung Women’s University, Korea</td>
</tr>
<tr>
<td>Shaoxu Song</td>
<td>Tsinghua University, China</td>
</tr>
<tr>
<td>Atsuhiro Takasu</td>
<td>National Institute of Informatics, Japan</td>
</tr>
<tr>
<td>Kian-Lee Tan</td>
<td>National University of Singapore (NUS), Singapore</td>
</tr>
<tr>
<td>Nan Tang</td>
<td>Qatar Computing Research Institute, Qatar</td>
</tr>
<tr>
<td>Martin Theobald</td>
<td>University of Antwerp, Belgium</td>
</tr>
<tr>
<td>Dimitri Theodoratos</td>
<td>New Jersey Institute of Technology, USA</td>
</tr>
<tr>
<td>James Thom</td>
<td>RMIT University, Australia</td>
</tr>
<tr>
<td>Wolf Tilo-Balke</td>
<td>University of Hannover, Germany</td>
</tr>
<tr>
<td>Hanghang Tang</td>
<td>City University of New York (CUNY), USA</td>
</tr>
<tr>
<td>Yongxin Tong</td>
<td>Hong Kong University of Science and Technology (HKUST), China</td>
</tr>
<tr>
<td>Kristian Torp</td>
<td>Aalborg University, Denmark</td>
</tr>
<tr>
<td>Goce Trajcevski</td>
<td>Northwestern University, USA</td>
</tr>
<tr>
<td>Vincent S. Tseng</td>
<td>National Cheng Kung University, Taiwan</td>
</tr>
<tr>
<td>Stratis Viglas</td>
<td>University of Edinburgh, UK</td>
</tr>
<tr>
<td>Wei Wang</td>
<td>University of New South Wales, Australia</td>
</tr>
<tr>
<td>Huayu Wu</td>
<td>Institute for Infocomm Research (I2R), Singapore</td>
</tr>
<tr>
<td>Yinghui Wu</td>
<td>University of California, Santa Barbara (UCSB), USA</td>
</tr>
<tr>
<td>Xiaokui Xiao</td>
<td>Nanyang Technological University (NTU), Singapore</td>
</tr>
<tr>
<td>Xike Xie</td>
<td>Aalborg University, Denmark</td>
</tr>
<tr>
<td>Jianliang Xu</td>
<td>Hong Kong Baptist University, China</td>
</tr>
<tr>
<td>Bin Yang</td>
<td>Aalborg University, Denmark</td>
</tr>
</tbody>
</table>
XII Organization

Yin Yang
Man-Lung Yiu
Haruo Yokota
Jeffrey Yu
Zhenjie Zhang
Xiuzhen Zhang
Kevin Zheng
Wenchao Zhou
Bin Zhou
Roger Zimmermann
Lei Zou
Andreas Züfle
Advanced Digital Sciences Center, Singapore
Hong Kong Polytechnic University, China
Tokyo Institute of Technology, Japan
The Chinese University of Hong Kong (CUHK), China
Advanced Digital Sciences Center (ADSC), Singapore
RMIT University, Australia
University of Queensland, Australia
Georgetown University, USA
University of Maryland, Baltimore County, USA
National University of Singapore (NUS), Singapore
Beijing University, China
Ludwig-Maximilians-Universität München, Germany

External Reviewers

Yeonchan Ahn
Cem Aksoy
Ibrahim Alabdulmohsin
Yoshitaka Arahori
Zhuowei Bao
Thomas Behr
Jianneng Cao
Brice Chardin
Lei Chen
Jian Dai
Ananya Dass
Aggeliki Dimitriou
Zhaoan Dong
Hai Dong
Zoé Faget
Qiong Fang
ZiQiang Feng
Ming Gao
Azadeh Ghari-Neiat
Zhihan He
Yuzhen Huang
Stéphane Jean
Selma Khouri
Hanbit Lee
Sang-Chul Lee
Seoul National University, Korea
New Jersey Institute of Technology, USA
King Abdullah University of Science and Technology, Saudi Arabia
Tokyo Institute of Technology, Japan
Facebook, USA
University of Hagen, Germany
A*STAR, Singapore
LIAS/ISAE-ENSMA, France
Hong Kong Baptist University, China
The Chinese Academy of Sciences, China
New Jersey Institute of Technology, USA
National Technical University of Athens, Greece
Renmin University of China, China
RMIT University, Australia
LIAS/ISAE-ENSMA, France
Hong Kong University of Science and Technology (HKUST), China
Hong Kong Polytechnic University, China
East China Normal University, China
RMIT University, Australia
Hong Kong Polytechnic University, China
The Chinese University of Hong Kong, China
LIAS/ISAE-ENSMA, France
LIAS/ISAE-ENSMA, France
Seoul National University, Korea
Carnegie Mellon University, USA
<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feng Li</td>
<td>Microsoft Research, Redmond, USA</td>
</tr>
<tr>
<td>Yafei Li</td>
<td>Hong Kong Baptist University, China</td>
</tr>
<tr>
<td>Jinfeng Li</td>
<td>The Chinese University of Hong Kong, China</td>
</tr>
<tr>
<td>Xin Lin</td>
<td>East China Normal University, China</td>
</tr>
<tr>
<td>Yu Liu</td>
<td>Renmin University of China, China</td>
</tr>
<tr>
<td>Yi Lu</td>
<td>The Chinese University of Hong Kong, China</td>
</tr>
<tr>
<td>Nguyen Minh Luan</td>
<td>A*STAR, Singapore</td>
</tr>
<tr>
<td>Gerasimos Marketos</td>
<td>University of Piraeus, Greece</td>
</tr>
<tr>
<td>Jun Miyazaki</td>
<td>Tokyo Institute of Technology, Japan</td>
</tr>
<tr>
<td>Bin Mu</td>
<td>City University of New York, USA</td>
</tr>
<tr>
<td>Johannes Niedermayer</td>
<td>Ludwig-Maximilians-Universität München, Germany</td>
</tr>
<tr>
<td>Konstantinos Nikolopoulos</td>
<td>City University of New York, USA</td>
</tr>
<tr>
<td>Sungchan Park</td>
<td>Seoul National University, Korea</td>
</tr>
<tr>
<td>Youngki Park</td>
<td>Samsung Advanced Institute of Technology, Korea</td>
</tr>
<tr>
<td>Jianbin Qin</td>
<td>University of New South Wales, Australia</td>
</tr>
<tr>
<td>Kai Qin</td>
<td>RMIT University, Australia</td>
</tr>
<tr>
<td>Youhyun Shin</td>
<td>Seoul National University, Korea</td>
</tr>
<tr>
<td>Hiroaki Shiokawa</td>
<td>NTT Software Innovation Center, Japan</td>
</tr>
<tr>
<td>Masumi Shirakawa</td>
<td>Osaka University, Japan</td>
</tr>
<tr>
<td>Md. Anisuzzaman Siddique</td>
<td>Hiroshima University, Japan</td>
</tr>
<tr>
<td>Reza Soltanpoor</td>
<td>RMIT University, Australia</td>
</tr>
<tr>
<td>Yifang Sun</td>
<td>University of New South Wales, Australia</td>
</tr>
<tr>
<td>Erald Troja</td>
<td>City University of New York, USA</td>
</tr>
<tr>
<td>Fabio Valdés</td>
<td>University of Hagen, Germany</td>
</tr>
<tr>
<td>Jan Vosecky</td>
<td>Hong Kong University of Science and Technology</td>
</tr>
<tr>
<td></td>
<td>(HKUST), China</td>
</tr>
<tr>
<td>Jim Jing-Yan Wang</td>
<td>King Abdullah University of Science and Technology, Saudi Arabia</td>
</tr>
<tr>
<td>Huanhuan Wu</td>
<td>The Chinese University of Hong Kong, China</td>
</tr>
<tr>
<td>Xiaoying Wu</td>
<td>Wuhan University, China</td>
</tr>
<tr>
<td>Chen Xu</td>
<td>Technische Universität Berlin, Germany</td>
</tr>
<tr>
<td>Jianqiu Xu</td>
<td>Nanjing University of Aeronautics and Astronautics, China</td>
</tr>
<tr>
<td>Takeshi Yamamuro</td>
<td>NTT Software Innovation Center, Japan</td>
</tr>
<tr>
<td>Da Yan</td>
<td>The Chinese University of Hong Kong, China</td>
</tr>
<tr>
<td>Fan Yang</td>
<td>The Chinese University of Hong Kong, China</td>
</tr>
<tr>
<td>Jongheum Yeon</td>
<td>Seoul National University, Korea</td>
</tr>
<tr>
<td>Seongwook Youn</td>
<td>University of Southern California, USA</td>
</tr>
<tr>
<td>Zhou Zhao</td>
<td>Hong Kong University of Science and Technology</td>
</tr>
<tr>
<td></td>
<td>(HKUST), China</td>
</tr>
<tr>
<td>Xiaoling Zhou</td>
<td>University of New South Wales, Australia</td>
</tr>
</tbody>
</table>
Contents – Part I

Data Mining I

Leveraging Homomorphisms and Bitmaps to Enable the Mining of Embedded Patterns from Large Data Trees ............................................. 3
   Xiaoying Wu and Dimitri Theodoratos

Cold-Start Expert Finding in Community Question Answering via Graph Regularization ................................................................. 21
   Zhou Zhao, Furu Wei, Ming Zhou, and Wilfred Ng

Mining Itemset-based Distinguishing Sequential Patterns with Gap Constraint ................................................................................. 39
   Hao Yang, Lei Duan, Guozhu Dong, Jyrki Nummenmaa,
   Changjie Tang, and Xiaosong Li

Mining Correlations on Massive Bursty Time Series Collections .......... 55
   Tomasz Kusmierczyk and Kjetil Nørvåg

Data Streams and Time Series

Adaptive Grid-Based k-median Clustering of Streaming Data with Accuracy Guarantee ................................................................. 75
   Jianpeng Cao, Yongluan Zhou, and Min Wu

Grouping Methods for Pattern Matching in Probabilistic Data Streams ...... 92
   Kento Sugiura, Yoshiharu Ishikawa, and Yuya Sasaki

Fast Similarity Search of Multi-Dimensional Time Series via Segment Rotation ........................................................................... 108
   Xudong Gong, Yan Xiong, Wenchao Huang, Lei Chen, Qiwei Lu,
   and Yiqing Hu

Measuring the Influence from User-Generated Content to News via Cross-dependence Topic Modeling ..................................... 125
   Lei Hou, Juanzi Li, Xiao-Li Li, and Yu Su

Database Storage and Index I

SASS: A High-Performance Key-Value Store Design for Massive Hybrid Storage ................................................................. 145
   Jiangtao Wang, Zhiliang Guo, and Xiaofeng Meng
An Efficient Design and Implementation of Multi-Level Cache for Database Systems ................................................................. 160
    Jiangtao Wang, Zhiliang Guo, and Xiaofeng Meng

A Cost-aware Buffer Management Policy for Flash-based Storage Devices ........................................................................ 175
    Zhiwen Jiang, Yong Zhang, Jin Wang, and Chunxiao Xing

The Gaussian Bloom Filter ........................................................................................................................................... 191
    Martin Werner and Mirco Schönfeld

Spatio-Temporal Data I

Detecting Hotspots from Trajectory Data in Indoor Spaces ...................... 209
    Peiquan Jin, Jiang Du, Chuanglin Huang, Shouhong Wan, and Lihua Yue

On Efficient Passenger Assignment for Group Transportation .................. 226
    Jiajie Xu, Guanfeng Liu, Kai Zheng, Chengfei Liu, Haoming Guo,
    and Zhiming Ding

Effective and Efficient Predictive Density Queries for Indoor Moving Objects ................................................................. 244
    Miao Li, Yu Gu, and Ge Yu

Efficient Trip Planning for Maximizing User Satisfaction ............................... 260
    Chenghao Zhu, Jiajie Xu, Chengfei Liu, Pengpeng Zhao,
    An Liu, and Lei Zhao

Modern Computing Platform

Accelerating Search of Protein Sequence Databases using CUDA-Enabled GPU ............................................................... 279
    Lin Cheng and Greg Butler

Fast Subgraph Matching on Large Graphs using Graphics Processors ....... 299
    Ha-Nguyen Tran, Jung-jae Kim, and Bingsheng He

On Longest Repeat Queries Using GPU .................................................. 316
    Yun Tian and Bojian Xu

Process-driven Configuration of Federated Cloud Resources ......................... 334
    Denis Weerasiri, Boualem Benatallah, and Moshe Chai Barukh
Social Networks I

An Integrated Tag Recommendation Algorithm Towards Weibo User Profiling. ................................................. 353
   Deqing Yang, Yanghua Xiao, Hanghang Tong, Junjun Zhang, and Wei Wang

An Efficient Approach of Overlapping Communities Search .............. 374
   Jing Shan, Derong Shen, Tiezheng Nie, Yue Kou, and Ge Yu

A Comparative Study of Team Formation in Social Networks ............. 389
   Xinyu Wang, Zhou Zhao, and Wilfred Ng

Inferring Diffusion Networks with Sparse Cascades by Structure Transfer... 405
   Senzhang Wang, Honghui Zhang, Jiawei Zhang, Xiaoming Zhang, Philip S. Yu, and Zhoujun Li

Information Integration and Data Quality

Scalable Inclusion Dependency Discovery .................................. 425
   Nuhad Shaabani and Christoph Meinel

Repairing Functional Dependency Violations in Distributed Data .......... 441
   Qing Chen, Zijing Tan, Chu He, Chaofeng Sha, and Wei Wang

GB-JER: A Graph-Based Model for Joint Entity Resolution ............... 458
   Chenchen Sun, Derong Shen, Yue Kou, Tiezheng Nie, and Ge Yu

Provenance-Aware Entity Resolution: Leveraging Provenance to Improve Quality ............................................. 474
   Qing Wang, Klaus-Dieter Schewe, and Woods Wang

Information Retrieval and Summarization

A Chip off the Old Block – Extracting Typical Attributes for Entities Based on Family Resemblance ........................................ 493
   Silviu Homoceanu and Wolf-Tilo Balke

Tag-based Paper Retrieval: Minimizing User Effort with Diversity Awareness ....................................................... 510
   Quoc Viet Hung Nguyen, Son Thanh Do, Thanh Tam Nguyen, and Karl Aberer

Feedback Model for Microblog Retrieval .................................... 529
   Ziqi Wang and Ming Zhang
Efficient String Similarity Search: A Cross Pivotal Based Approach

*Fei Bi, Lijun Chang, Wenjie Zhang, and Xuemin Lin*

**Security and Privacy**

Authentication of Top-k Spatial Keyword Queries in Outsourced Databases

*Sen Su, Han Yan, Xiang Cheng, Peng Tang, Peng Xu, and Jianliang Xu*

Privacy-Preserving Top-k Spatial Keyword Queries over Outsourced Database

*Sen Su, Yiping Teng, Xiang Cheng, Yulong Wang, and Guoliang Li*

Bichromatic Reverse Nearest Neighbor Query without Information Leakage

*Lu Wang, Xiaofeng Meng, Haibo Hu, and Jianliang Xu*

Authentication of Reverse k Nearest Neighbor Query

*Guohui Li, Changyin Luo, and Jianjun Li*

**Author Index**
Contents – Part II

Outlier and Imbalanced Data Analysis

A Synthetic Minority Oversampling Method Based on Local Densities in Low-Dimensional Space for Imbalanced Learning .......................... 3
Zhipeng Xie, Liyang Jiang, Tengju Ye, and Xiao-Li Li

Fast and Scalable Outlier Detection with Approximate Nearest Neighbor Ensembles ................................................................. 19
Erich Schubert, Arthur Zimek, and Hans-Peter Kriegel

Rare Category Exploration on Linear Time Complexity ......................... 37
Zhenguang Liu, Hao Huang, Qinming He, Kevin Chiew, and Yunjun Gao

Probabilistic and Uncertain Data

FP-CPNNQ: A Filter-Based Protocol for Continuous Probabilistic Nearest Neighbor Query ................................................................. 57
Yinuo Zhang, Anand Panangadan, and Viktor K. Prasanna

Efficient Queries Evaluation on Block Independent Disjoint Probabilistic Databases ................................................................. 74
Biao Qin

Parallel Top-k Query Processing on Uncertain Strings Using MapReduce ..................................................................................... 89
Hui Xu, Xiaofeng Ding, Hai Jin, and Wenbin Jiang

Tracing Errors in Probabilistic Databases Based on the Bayesian Network .......................................................................................... 104
Liang Duan, Kun Yue, Cheqing Jin, Wenlin Xu, and Weiyi Liu

Data Mining II

Mining Frequent Spatial-Textual Sequence Patterns .............................. 123
Krishan K. Arya, Vikram Goyal, Shamkant B. Navathe, and Sushil Prasad

Effective and Interpretable Document Classification Using Distinctly Labeled Dirichlet Process Mixture Models of von Mises-Fisher Distributions ................................................................. 139
Ngo Van Linh, Nguyen Kim Anh, Khoat Than, and Nguyen Nguyen Tat
XX Contents – Part II

MPTM: A Topic Model for Multi-Part Documents ........................................... 154
    Zhipeng Xie, Liyang Jiang, Tengju Ye, and Zhenying He

Retaining Rough Diamonds: Towards a Fairer Elimination
of Low-Skilled Workers ................................................................. 169
    Kinda El Maarry and Wolf-Tilo Balke

Spatio-temporal Data II

Skyline Trips of Multiple POIs Categories ............................................. 189
    Saad Aljubayrin, Zhen He, and Rui Zhang

Keyword-Aware Dominant Route Search for Various User Preferences .......... 207
    Yujiao Li, Weidong Yang, Wu Dan, and Zhipeng Xie

Spatial Keyword Range Search on Trajectories ....................................... 223
    Yuxing Han, Liping Wang, Ying Zhang, Wenjie Zhang, and Xuemin Lin

TOF: A Throughput Oriented Framework for Spatial Queries Processing
in Multi-core Environment .................................................................. 241
    Zhong-Bin Xue, Xuan Zhou, and Shan Wang

Query Processing

Identifying and Caching Hot Triples for Efficient RDF Query Processing ........ 259
    Wei Emma Zhang, Quan Z. Sheng, Kerry Taylor, and Yongrui Qin

History-Pattern Implementation for Large-Scale Dynamic
Multidimensional Datasets and Its Evaluations ....................................... 275
    Masafumi Makino, Tatsuo Tsuji, and Ken Higuchi

Scalagon: An Efficient Skyline Algorithm for All Seasons ....................... 292
    Markus Endres, Patrick Roocks, and Werner Kießling

Towards Order-Preserving SubMatrix Search and Indexing ....................... 309
    Tao Jiang, Zhanhuai Li, Qun Chen, Kaiwen Li, Zhong Wang, and Wei Pan

Database Storage and Index II

Large-Scale Multi-party Counting Set Intersection Using a Space
Efficient Global Synopsis .................................................................. 329
    Dimitrios Karapiperis, Dinusha Vatsalan, Vassilios S. Verykios, and Peter Christen
Improved Weighted Bloom Filter and Space Lower Bound Analysis of Algorithms for Approximated Membership Querying

Xiujun Wang, Yusheng Ji, Zhe Dang, Xiao Zheng, and Baohua Zhao

Tree Contraction for Compressed Suffix Arrays on Modern Processors

Takeshi Yamamuro, Makoto Onizuka, and Toshimori Honjo

Scalable Top-$k$ Spatial Image Search on Road Networks

Pengpeng Zhao, Xiaopeng Kuang, Victor S. Sheng, Jiajie Xu, Jian Wu, and Zhiming Cui

**Social Networks II**

An Efficient Method to Find the Optimal Social Trust Path in Contextual Social Graphs

Guanfeng Liu, Lei Zhao, Kai Zheng, An Liu, Jiajie Xu, Zhixu Li, and Athman Bouguettaya

Pricing Strategies for Maximizing Viral Advertising in Social Networks

Bolei Zhang, Zhuzhong Qian, Wenzhong Li, and Sanglu Lu

Boosting Financial Trend Prediction with Twitter Mood Based on Selective Hidden Markov Models

Yifu Huang, Shuigeng Zhou, Kai Huang, and Jihong Guan

k-Consistent Influencers in Network Data

Enliang Xu, Wynne Hsu, Mong Li Lee, and Dhaval Patel

**Industrial Papers**

Analyzing Electric Vehicle Energy Consumption Using Very Large Data Sets

Benjamin Krogh, Ove Andersen, and Kristian Torp

Interactive, Flexible, and Generic What-If Analyses Using In-Memory Column Stores

Stefan Klauck, Lars Butzmann, Stephan Müller, Martin Faust, David Schwalb, Matthias Uflacker, Werner Sinzig, and Hasso Plattner

Energy Efficient Scheduling of Fine-Granularity Tasks in a Sensor Cloud

Rashmi Dalvi and Sanjay Kumar Madria

**Demo**

Invariant Event Tracking on Social Networks

Sayan Unankard, Xue Li, and Guodong Long
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>EmoTrend: Emotion Trends for Events</td>
<td>522</td>
</tr>
<tr>
<td>Yi-Shin Chen, Carlos Argueta, and Chun-Hao Chang</td>
<td></td>
</tr>
<tr>
<td>A Restaurant Recommendation System by Analyzing Ratings and Aspects in Reviews</td>
<td>526</td>
</tr>
<tr>
<td>Yifan Gao, Wenzhe Yu, Pingfu Chao, Rong Zhang, Aoying Zhou, and Xiaoyan Yang</td>
<td></td>
</tr>
<tr>
<td>ENRS: An Effective Recommender System Using Bayesian Model</td>
<td>531</td>
</tr>
<tr>
<td>Yingyuan Xiao, Pengqiang Ai, Hongya Wang, Ching-Hsien Hsu, and Yukun Li</td>
<td></td>
</tr>
<tr>
<td>EPSCS: Simulating and Measuring Energy Proportionality of Server Clusters</td>
<td>536</td>
</tr>
<tr>
<td>Jiazhuang Xie, Peiquan Jin, Shouhong Wan, and Lihua Yue</td>
<td></td>
</tr>
<tr>
<td>MAVis: A Multiple Microblogs Analysis and Visualization Tool</td>
<td>541</td>
</tr>
<tr>
<td>Changping Wang, Chaokun Wang, Jingchao Hao, Hao Wang, and Xiaojun Ye</td>
<td></td>
</tr>
<tr>
<td>Author Index</td>
<td>547</td>
</tr>
</tbody>
</table>