Welcome to the 11th International Conference on Database Systems for Advanced Applications (DASFAA 2006)! This year’s conference was held in Singapore where DASFAAA was last held in 1997. DASFAA 2006 continued the tradition of providing an international forum for technical discussion among researchers, developers and users of database systems from academia, business and industry. Organizing DASFAA 2006 was a very rewarding experience—it gave me an excellent opportunity to work with many fine colleagues both within and outside Singapore.

I would like to thank Kian-Lee Tan and Vilas Wuwongse for putting together a world-class Program Committee. The committee worked very hard to bring a high-quality technical program to the conference. DASFAA 2006 also included an industrial track. David Cheung and Hwee Hwa Pang co-chaired this track and set up a separate committee to assess the quality of the submitted papers.

The conference also featured three tutorials: (1) Database Watermarking by Radu Sion, (2) Multilingual Database Systems by Jayant R. Haritsa, and (3) Video Sequence Indexing and Query Processing by Xiaofang Zhou, and a panel session. I would like to thank Ee Peng Lim and Krithi Ramamritham for their effort in organizing the tutorials and panel, respectively.

This conference would not have been possible without the support of many other colleagues: Tok-Wang Ling (Honorary Conference Chair), Mong Li Lee (Publication Chair), Masatoshi Yoshikawa (Publicity Chair), Sourav Saha Bhowmick and Anthony Kum Hoe Tung (Local Arrangement Co-chairs), Chee Yong Chan (Treasurer), and Mrs. Siew Foong Ho (Secretary).

Finally, I greatly appreciated the support of the National University of Singapore (NUS) and the Nanyang Technological University. In particular, I was happy that DASFAA 2006 participated in the NUS Centennial Celebrations as an event organized by the NUS academic staff.

April 2006

Beng Chin Ooi
Conference Chair
The 11th International Conference on Database Systems for Advanced Applications (DASFAA 2006) was held in Singapore from April 12 to 15, 2006. As an annual international conference in the Pacific Asia region, DASFAA 2006 kept the traditions of the conference in promoting research and development activities in the database field among participants and their institutions from Pacific Asia and the rest of the world.

This year, the conference received 188 (research-track) submissions from over 28 different countries. The submitted papers were rigorously reviewed by the Program Committee members, and 46 full papers and 16 short papers were accepted for presentation.

The papers chosen for presentation spanned a wide range of topics, ranging from well-established areas such as XML, spatial and temporal databases, and data mining to upcoming areas like sensor networks, uncertainty and data streams, and subsequence matching and bioinformatics. The combination of selected papers made the conference interesting and provided the basis for discussion and exchange of ideas and for future development.

The conference was privileged to have keynote addresses delivered by Alon Y. Halevy of Google Inc. and University of Washington, Krithi Ramamritham of IIT Bombay, and Christian Jensen of Aalborg University. They provided insightful thoughts into various research challenges on information management, dissemination of dynamic data and geo-enabled mobile services.

DASFAA 2006 also included an industrial track with the aim of drawing industry practitioners and the academic community to share practical experience and real-world challenges that require research attention, and to advance the state of the art by integrating new techniques and research results in novel systems and applications. This track included a paper on automating the maintenance of the statistics for query optimization in Sybase ASE 15.0, a paper on automatically finding a mapping that transforms an XML schema into a relational counterpart that is tuned to the application workload, and a third paper that treats the problem of missing data from sensors deployed to monitor elderly dementia patients.

The research and industrial tracks were both supported by their respective Technical Program Committees. Both teams comprised renowned and hardworking researchers from around the world. Their invaluable efforts in reviewing the papers ensured the high quality of the accepted papers. We would like to take this opportunity to thank them again!

The technical program also consisted of three tutorials and a panel session. The three tutorials featured were Database Watermarking by Radu Sion, Multilingual Database Systems by Jayant R. Haritsa, and Video Sequence Indexing and Query Processing by Xiaofang Zhou.
The conference would not have been a success without the help and contributions of many individuals, and we would like to acknowledge them here: Tok-Wang Ling, Beng Chin Ooi, Ee Peng Lim, Krithi Ramamritham, Masatoshi Yoshikawa, Mong Li Lee, Sourav Saha Bhowmick, Anthony Kum Hoe Tung, Chee Yong Chan and Mrs. Ho Siew Foong. Finally, we would like to thank the session chairs, tutorial speakers, authors and participants, who contributed to making this conference a success.

April 2006

Kian-Lee Tan and Vilas Wuwongse
Research Track Co-chairs

David Cheung and Hwee Hwa Pang
Industrial Track Co-chairs
# Conference Organization

**Honorary Chair**
Tok Wang Ling  
National University of Singapore, Singapore

**Conference Chair**
Beng Chin Ooi  
National University of Singapore, Singapore

**Program Co-chairs**
Kian Lee Tan  
National University of Singapore, Singapore
Vilas Wuwongse  
Asian Institute of Technology, Thailand

**Tutorial Chair**
Ee Peng Lim  
Nanyang Technological University, Singapore

**Panel Chair**
Krithi Ramamritham  
Indian Institute of Technology, Bombay

**Industrial Program Co-chairs**
David Wai-Lok Cheung  
The University of Hong Kong, China
Hwee Hwa Pang  
Singapore Management University, Singapore

**Publicity Chair**
Masatoshi Yoshikawa  
Nagoya University, Japan

**Publication Chair**
Mong Li Lee  
National University of Singapore, Singapore

**Local Arrangement Co-chairs**
Sourav Saha Bhowmick  
Nanyang Technological University, Singapore
Anthony Kum Hoe Tung  
National University of Singapore, Singapore

**Treasurer**
Chee Yong Chan  
National University of Singapore, Singapore

**Secretary**
Mrs. Ho Siew Foong  
National University of Singapore, Singapore

## Jointly Organized by

School of Computing, National University of Singapore  
School of Computer Engineering, Nanyang Technological University

As part of the National University of Singapore Centennial Celebrations
Program Committee

Regular Track

Charu Aggarwal IBM T.J. Watson Research Center, USA
Chutiporn Anutariya Shinawatra University, Thailand
Vijay Atluri Rutgers University, USA
Sonia Berman University of Cape Town, South Africa
Sourav S Bhowmick Nanyang Technological University, Singapore
Luc Bouganim INRIA, France
Athman Bouguettaya Virginia Tech, USA
Stephane Bressan National University of Singapore, Singapore
K. Selcuk Candan Arizona State University, USA
Barbara Catania University of Genova, Italy
Arbee L.P. Chen National Tshing Hua University, Taiwan
Ming-Syan Chen National Taiwan University, Taiwan
Ying Chen IBM China Research Lab, China
Brian F Cooper Georgia Institute of Technology, USA
Isabel F. Cruz University of Illinois at Chicago, USA
Manoranjan Dash Nanyang Technological University, Singapore
Amol Deshpande University of Maryland, USA
Klaus Dittrich University of Zurich, Switzerland
Gillian Dobbie University of Auckland, New Zealand
Curtis Dyreson Washington State University, USA
David W. Embley Brigham Young University, USA
Hakan Ferhatosmanoglu Ohio State University, USA
Elena Ferrari University of Insubria at Como, Italy
Ada Waichee Fu Chinese University of Hong Kong, China
Cong Gao University of Edinburgh, UK
Minos Garofalakis Bell Labs-Lucent Technologies, USA
Shahram Ghandeharizadeh University of Southern California, USA
Jonathan Goldstein Microsoft Research, USA
Dimitrios Gunopoulos University of California, Riverside, USA
Theo Haerder TU Kaiserslautern, Germany
Takahiro Hara Osaka University, Japan
Arantza Illarramendi Basque Country University, Spain
Bala Iyer IBM Silicon Valley Lab, USA
H. V. Jagadish University of Michigan, USA
Arnd Christian Konig Microsoft Research, USA
Panagiotis Kalnis National University of Singapore, Singapore
Ibrahim Kamel Zayed University, United Arab Emirates
Hyunchul Kang Chung-Ang University, Korea
George Karypis University of Minnesota, USA
Masaru Kitsuregawa Institute of Industrial Science, Japan
Donald Kossmann University of Heidelberg, Germany
Manolis Koubarakis Technical University of Crete, Greece
Chiang Lee National Cheng-Kung University, Taiwan
Dik Lun Lee, Hong Kong University of Science and Technology, China
Wang-Chien Lee, Penn State University, USA
Yoon-Joon Lee, KAIST, Korea
Jianzhong Li, Harbin Institute of Technology, China
Jinyan Li, Institute for InfoComm Research, Singapore
Wen-Syan Li, IBM Almaden Research Center, USA
Sergio Lifschitz, PUC-Rio, Brazil
Sanjay Madria, University of Missouri-Rolla, USA
Stefan Manegold, CWI, The Netherlands
Ioana Manolescu, INRIA Futurs, France
Volker Markl, IBM Almaden Research Center, USA
Wei-yi Meng, Binghamton University, USA
Xiaofeng Meng, Renmin University of China, China
Mukesh Mohania, IBM India Research Lab, India
Atsuuyuki Morishima, University of Tsukuba, Japan
S. Muthukrishnan, Rutgers University, USA
Mario A Nascimento, University of Alberta, Canada
Wolfgang Nejdl, University of Hannover, Germany
Raymond T. Ng, University of British Columbia, Canada
Aris Ouksel, University of Illinois at Chicago, USA
Sanhyun Park, Yonsei University, Korea
Evaggelia Pitoura, University of Ioannina, Greece
Sunil Prabhakar, Purdue University, USA
Weining Qian, Fudan University, China
Tore Risch, Uppsala University of Sweden
Prasan Roy, IBM India Research Lab, India
Simonas Saltenis, Aalborg University, Denmark
Markus Schneider, University of Florida, USA
Thomas Seidl, RWTH Aachen University, Germany
Oded Shmueli, Technion-Israel Institute of Technology, Israel
Ambuj K. Singh, University of California at Santa Barbara, USA
Dan Suciu, University of Washington, USA
Keishi Tajima, JAIST, Japan
Wang-Chiew Tan, University of California, Santa Cruz, USA
Katsumi Tanaka, Kyoto University, Japan
David Taniar, Monash University, Australia
Yufei Tao, City University of Hong Kong, China
Ozgur Ulusoy, Bilkent University, Turkey
Vasilis A. Vassalos, Athens University of Economics and Business, Greece
Haixun Wang, IBM T.J. Watson Research Center, USA
Wei Wang, University of North Carolina at Chapel Hill, USA
Yan Wang, Macquarie University, Australia
Kyu-Young Whang, KAIST, Korea
Jonker Willem, Philips Research, The Netherlands
Masatoshi Yoshikawa, Nagoya University, Japan
Xu Yu, Jeffrey, Chinese University of Hong Kong, China
Arkady Zaslavsky, Monash University, Australia
Yanchun Zhang, Victoria University, Australia
XII Organization

Jingren Zhou  Microsoft Research, USA
Shuigeng Zhou  Fudan University, China
Xiaofang Zhou  University of Queensland, Australia

Industrial Track

Manish Bhide  IBM India Research Lab, India
Qiong Luo  Hong Kong University of Science and Technology, China
Jussi Myllymaki  Google Inc., USA
Il-Yeol Song  Drexel University, USA
Kam-Fai Wong  The Chinese University of Hong Kong, China
Bai-Hua Zheng  Singapore Management University, Singapore

External Referees

Christoph Brochhaus  RWTH Aachen
Jost Enderle  RWTH Aachen
Ralph Krieger  RWTH Aachen
Guanling Lee  Pennsylvania State University
Ken Lee  Pennsylvania State University
Mei Li  Pennsylvania State University
Jinze Liu  University of North Carolina at Chapel Hill
Feng Pan  University of North Carolina at Chapel Hill
Wen-Chi Peng  Pennsylvania State University
Qingzhao Tan  Pennsylvania State University
Marc Wichterich  RWTH Aachen
Julian Winter  Pennsylvania State University
Table of Contents

Keynote Addresses

Dataspaces: A New Abstraction for Information Management
   Alon Y. Halevy, Michael J. Franklin, David Maier ......................... 1

Dissemination of Dynamic Data: Semantics, Algorithms, and Performance
   Krithi Ramamritham ................................................................. 3

Geo-Enabled, Mobile Services—A Tale of Routes, Detours, and Dead Ends
   Christian S. Jensen ................................................................. 6

Sensor Networks

Processing Multiple Aggregation Queries in Geo-Sensor Networks
   Ken C.K. Lee, Wang-Chien Lee, Baihua Zheng,
   Julian Winter ................................................................. 20

In-Network Processing of Nearest Neighbor Queries for Wireless Sensor Networks
   Yuxia Yao, Xueyan Tang, Ee-Peng Lim .................................... 35

Associated Load Shedding Strategies for Computing Multi-joins in Sensor Networks
   Xiaochun Yang, Lin Li, Yiu-Kai Ng, Bin Wang, Ge Yu ............... 50

Subsequence Matching and Repeating Patterns

Using Multiple Indexes for Efficient Subsequence Matching in Time-Series Databases
   Seung-Hwan Lim, Hee-Jin Park, Sang-Wook Kim ......................... 65

DAPSS: Exact Subsequence Matching for Data Streams
   Yasuhiro Fujiwara, Yasushi Sakurai, Masashi Yamamuro ............ 80

An Efficient Approach for Mining Top-K Fault-Tolerant Repeating Patterns
   Jia-Ling Koh, Yu-Ting Kung .................................................... 95
Spatial-temporal Databases

Querying Multi-granular Compact Representations
Romāns Kasperovičs, Michael Böhlen ............................... 111

The COST Benchmark—Comparison and Evaluation of Spatio-temporal Indexes
Christian S. Jensen, Dalia Tiešytė,
Nerius Tradišauskas ........................................... 125

Efficient Maintenance of Ephemeral Data
Albrecht Schmidt, Christian S. Jensen ......................... 141

Data Mining

Mining Outliers in Spatial Networks
Wen Jin, Yuelong Jiang, Weining Qian, Anthony K.H. Tung .......... 156

Summarizing Frequent Patterns Using Profiles
Gao Cong, Bin Cui, Yingxin Li, Zonghong Zhang .................. 171

Mining Spatio-temporal Association Rules, Sources, Sinks, Stationary Regions and Thoroughfares in Object Mobility Databases
Florian Verhein, Sanjay Chawla ......................... 187

XML Compression and Indexing

Document Decomposition for XML Compression: A Heuristic Approach
Byron Choi .................................................. 202

An Efficient Co-operative Framework for Multi-query Processing over Compressed XML Data
Juzhen He, Wilfred Ng, Xiaoling Wang, Aoying Zhou ............ 218

Adaptively Indexing Dynamic XML
Damien K. Fisher, Raymond K. Wong .......................... 233

XPath Query Evaluation

TwigStackList¬: A Holistic Twig Join Algorithm for Twig Query with Not-Predicates on XML Data
Tian Yu, Tok Wang Ling, Jiaheng Lu ............................. 249
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficient Schemes of Executing Star Operators in XPath Query Expressions</td>
<td>Young Chul Park, Je Hyun Cho, Geum Ji Cha, Peter Scheuermann</td>
<td>264</td>
</tr>
<tr>
<td>Exploit Sequencing to Accelerate XML Twig Query Answering</td>
<td>Qian Qian, Jianhua Feng, Jianyong Wang, Lizhu Zhou</td>
<td>279</td>
</tr>
<tr>
<td>Uncertainty and Streams</td>
<td>Probabilistic Similarity Join on Uncertain Data</td>
<td>295</td>
</tr>
<tr>
<td></td>
<td>Hans-Peter Kriegel, Peter Kunath, Martin Pfeifle, Matthias Renz</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Handling Uncertainty and Ignorance in Databases: A Rule to Combine Dependent Data</td>
<td>310</td>
</tr>
<tr>
<td></td>
<td>Sunil Choenni, Henk Ernst Blok, Erik Leertouwer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMJoin: Optimizing Distributed Multi-way Stream Joins by Stream Partitioning</td>
<td>325</td>
</tr>
<tr>
<td></td>
<td>Yongluan Zhou, Ying Yan, Feng Yu, Aoying Zhou</td>
<td></td>
</tr>
<tr>
<td>Peer-to-Peer and Distributed Networks</td>
<td>Clustering Peers Based on Contents for Efficient Similarity Search</td>
<td>342</td>
</tr>
<tr>
<td></td>
<td>Yanfeng Shu, Bei Yu</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Optimizing Peer Virtualization and Load Balancing</td>
<td>357</td>
</tr>
<tr>
<td></td>
<td>Wanxia Xie, Shamkant B. Navathe, Sushil K. Prasad, David Fisher, Yong Yang</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Distributed Network Querying with Bounded Approximate Caching</td>
<td>374</td>
</tr>
<tr>
<td></td>
<td>Badrish Chandramouli, Jun Yang, Amin Vahdat</td>
<td></td>
</tr>
<tr>
<td>Performance and Authentication</td>
<td>Type-Level Access Pattern View: A Technique for Enhancing Prefetching Performance</td>
<td>389</td>
</tr>
<tr>
<td></td>
<td>Wook-Shin Han, Woong-Kee Loh, Kyu-Young Whang</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The Dynamic Sweep Scheme Using Slack Time in the Zoned Disk</td>
<td>404</td>
</tr>
<tr>
<td></td>
<td>Sungchae Lim</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Authentication of Outsourced Databases Using Signature Aggregation and Chaining</td>
<td>420</td>
</tr>
<tr>
<td></td>
<td>Maithili Narasimha, Gene Tsudik</td>
<td></td>
</tr>
</tbody>
</table>
# Table of Contents

## XML Query Processing

A Practitioner’s Approach to Normalizing XQuery Expressions  
*Ki-Hoon Lee, Seo-Young Kim, Euijong Whang, Jae-Gil Lee*  
437

Hidden Conditioned Homomorphism for XPath Fragment Containment  
*Yuguo Liao, Jianhua Feng, Yong Zhang, Lizhu Zhou*  
454

Efficient Query Processing for Streamed XML Fragments  
*Huan Huo, Guoren Wang, Xiaoyun Hui, Rui Zhou, Bo Ning, Chuan Xiao*  
468

## OLAP and Data Warehouse

An Efficient Algorithm for Computing Range-Groupby Queries  
*Young-Koo Lee, Woong-Kee Loh, Yang-Sae Moon, Kyu-Young Whang, Il-Yeol Song*  
483

Ag-Tree: A Novel Structure for Range Queries in Data Warehouse Environments  
*Yaokai Feng, Akifumi Makinouchi*  
498

An XML Document Warehouse Model  
*Vicky Nassis, Tharam S. Dillon, Rajugan Rajagopalapillai, Wenny Rahayu*  
513

## Web and Web Services

An Evaluation of Concurrency Control Protocols for Web Services Oriented E-Commerce  
*Hong-Ren Chen*  
530

COWES: Clustering Web Users Based on Historical Web Sessions  
*Ling Chen, Sourav S. Bhowmick, Jinyan Li*  
541

A Precise Metric for Measuring How Much Web Pages Change  
*Shin Young Kwon, Sang Ho Lee, Sung Jin Kim*  
557

## Query Processing

Similarity Search in Transaction Databases with a Two-Level Bounding Mechanism  
*Jo-Chun Chuang, Chung-Wen Cho, Arbee L.P. Chen*  
572
RAF: An Activation Framework for Refining Similarity Queries Using Learning Techniques
   Yiming Ma, Sharad Mehrotra, Dawit Yimam Seid, Qi Zhong .......... 587

Query Optimization for a Graph Database with Visual Queries
   Greg Butler, Guang Wang, Yue Wang, Liqian Zou ................. 602

**Design: Modeling and Dependencies**

A Four Dimensional Petri Net Approach for Workflow Management
   Ping-Yu Hsu, Yen-Liang Chen, Yuan-Bin Chang...................... 617

Containment of Conjunctive Queries over Conceptual Schemata
   Andrea Calì .............................................................. 628

Data Tables with Similarity Relations: Functional Dependencies, Complete Rules and Non-redundant Bases
   Radim Bělohlávek, Vilém Vychodil ............................... 644

**Labeling Scheme and Graph Queries in XML**

Reuse or Never Reuse the Deleted Labels in XML Query Processing Based on Labeling Schemes
   Changqing Li, Tok Wang Ling, Min Hu ............................... 659

Fast Reachability Query Processing
   Jiefeng Cheng, Jeffrey Xu Yu, Nan Tang .......................... 674

**Document Retrieval**

Relation-Based Document Retrieval for Biomedical Literature Databases
   Xiaohua Zhou, Xiaohua Hu, Xia Lin, Hyoil Han,
   Xiaodan Zhang ............................................................. 689

Effective Keyword Search in XML Documents Based on MIU
   Jianjun Xu, Jiaheng Lu, Wei Wang, Baile Shi ...................... 702

**Industrial Papers**

Assessing the Completeness of Sensor Data
   Jit Biswas, Felix Naumann, Qiang Qiu ............................. 717
Intelligent Statistics Management in Sybase ASE 15.0
Satya Sreenivasan, Xiao Ming Zhou, Tat Keong Loh 733

Holistic Schema Mappings for XML-on-RDBMS
Priti Patil, Jayant R. Haritsa 741

Short Papers
Semi-supervised Classification Based on Smooth Graphs
Xueyuan Zhou, Chunping Li 757

Compacting XML Data
Shuohao Zhang, Curtis Dyreson, Zhe Dang 767

Fast Structural Join with a Location Function
Nan Tang, Jeffrey Xu Yu, Kam-Fai Wong, Haifeng Jiang 777

Adapting Prime Number Labeling Scheme for Directed Acyclic Graphs
Gang Wu, Kuo Zhang, Can Liu, Juanzi Li 787

KEYNOTE: Keyword Search by Node Selection for Text Retrieval on DHT-Based P2P Networks
Zheng Zhang, Shuigeng Zhou, Weining Qian, Aoying Zhou 797

How to BLAST Your Database — A Study of Stored Procedures for BLAST Searches
Uwe Röhm, Thanh-Mai Diep 807

DTD-Diff: A Change Detection Algorithm for DTDs
Erwin Leonardi, Tran T. Hoai, Sourav S. Bhowmick, Sanjay Madria 817

Mining Models of Composite Web Services for Performance Analysis
Aiqiang Gao, Dongqing Yang, Shiwei Tang, Ming Zhang 828

Modeling Multimedia Data Semantics with MADS
Oleksandr Drutska, Stefano Spaccapietra 838

STIL: An Extended Resource Description Framework and an Advanced Query Language for Metadatabases
Benjamin Buffereau, Philippe Picouet 849

Communication-Efficient Implementation of Range-Joins in Sensor Networks
Aditi Pandit, Himanshu Gupta 859
Efficient k-Nearest Neighbor Searches for Parallel Multidimensional Index Structures

Kyoung Soo Bok, Seok Il Song, Jae Soo Yoo

Efficient Non-Blocking Top-k Query Processing in Distributed Networks

Bo Deng, Yan Jia, Shuqiang Yang

Continuous Expansion: Efficient Processing of Continuous Range Monitoring in Mobile Environments

Xiaoyuan Wang, Wei Wang

Effective Low-Latency K-Nearest Neighbor Search Via Wireless Data Broadcast

KwangJin Park, MoonBae Song, Ki-Sik Kong, Sang-Won Kang, Chong-Sun Hwang, Kwang-Sik Chung, SoonYoung Jung

Nearest Neighbor Queries for R-Trees: Why Not Bottom-Up?

MoonBae Song, KwangJin Park, SeokJin Im, Ki-Sik Kong

Author Index