Ramamohanarao Kotagiri
P. Radha Krishna Mukesh Mohania
Ekawit Nantajeewarawat (Eds.)

Advances in Databases: Concepts, Systems and Applications

12th International Conference on Database Systems for Advanced Applications, DASFAA 2007
Bangkok, Thailand, April 9-12, 2007
Proceedings

Springer

Table of Contents

Invited Talks

'Socio Sense' and 'Cyber Infrastructure for Information Explosion Era': Projects in Japan 1
Masaru Kitsuregawa

Is (Your) Database Research Having Impact? 3
Guy M. Lohman

Part I: Full Papers

Query Language and Query Optimization - I

Improving Quality and Convergence of Genetic Query Optimizers 6
Victor Muntés-Mulero, Néstor Lafón-Gracia, Josep Aguilar-Saborit, and Josep-L. Larriba-Pey

Cost-Based Query Optimization for Multi Reachability Joins 18
Jiefeng Cheng, Jeffrey Xu Yu, and Bolin Ding

A Path-Based Approach for Efficient Structural Join with Not-Predicates 31
Hanyu Li, Mong Li Lee, Wynne Hsu, and Ling Li

Query Language and Query Optimization - II

RRPJ: Result-Rate Based Progressive Relational Join 43
Wee Hyong Tok, Stéphane Bressan, and Mong-Li Lee

GChord: Indexing for Multi-Attribute Query in P2P System with Low Maintenance Cost 55
Minqi Zhou, Rong Zhang, Weining Qian, and Aoying Zhou

ITREKS: Keyword Search over Relational Database by Indexing Tuple Relationship 67
Jiang Zhan and Shan Wang

Data Mining and Knowledge Discovery

An MBR-Safe Transform for High-Dimensional MBRs in Similar Sequence Matching 79
Yang-Sae Moon

Mining Closed Frequent Free Trees in Graph Databases 91
Peixiang Zhao and Jeffrey Xu Yu

Mining Time-Delayed Associations from Discrete Event Datasets 103
K.K. Loo and Ben Kao

Clustering

A Comparative Study of Ontology Based Term Similarity Measures on PubMed Document Clustering 115
Xiaodan Zhang, Liping Jing, Xiaohua Hu, Michael Ng, and Xiaohua Zhou

An Adaptive and Efficient Unsupervised Shot Clustering Algorithm for Sports Video 127
Jia Liao, Guoren Wang, Bo Zhang, Xiaofang Zhou, and Ge Yu

A Robust Feature Normalization Scheme and an Optimized Clustering Method for Anomaly-Based Intrusion Detection System 140
Jungsuk Song, Hiroki Takakura, Yasuo Okabe, and Yongjin Kwon

Detection and Visualization of Subspace Cluster Hierarchies 152
Elke Achtert, Christian Böhm, Hans-Peter Kriegel, Peer Kröger, Ina Müller-Gorman, and Arthur Zimek

Outlier Detection

Correlation-Based Detection of Attribute Outliers 164
Judice L.Y. Koh, Mong Li Lee, Wynne Hsu, and Kai Tak Lam

An Efficient Histogram Method for Outlier Detection 176
Matthew Gebski and Raymond K. Wong

Privacy Preserving Data Mining

Efficient k-Anonymization Using Clustering Techniques 188
Ji-Won Byun, Ashish Kamra, Elisa Bertino, and Ninghui Li

Privacy Preserving Data Mining of Sequential Patterns for Network Traffic Data 201
Seung-Woo Kim, Sanghyun Park, Jung-Im Won, and Sang-Wook Kim

Privacy Preserving Clustering for Multi-party 213
Weijia Yang and Shangteng Huang

Privacy-Preserving Frequent Pattern Sharing 225
Zhihui Wang, Wei Wang, Baile Shi, and S.H. Boey

Parallel and Distributed Databases

K_nBest - A Balanced Request Allocation Method for Distributed Information Systems 237
Jorge-Arnulfo Quiané-Ruiz, Philippe Lamarre, and Patrick Valduriez

The Circular Two-Phase Commit Protocol 249
Heine Kolltveit and Svein-Olaf Hvasshovd

Towards Timely ACID Transactions in DBMS 262
Marco Vieira, António C. Costa, and Henrique Madeira

Data Warehouse

BioDIFF: An Effective Fast Change Detection Algorithm for Biological Annotations 275
Yang Song, Sourav S. Bhowmick, and C. Forbes Dewey Jr.

An Efficient Implementation for MOLAP Basic Data Structure and Its Evaluation 288
K.M. Azharul Hasan, Tatsuo Tsuji, and Ken Higuchi

Information Retrieval

Monitoring Heterogeneous Nearest Neighbors for Moving Objects Considering Location-Independent Attributes 300
Yu-Chi Su, Yi-Hung Wu, and Arbee L.P. Chen

Similarity Joins of Text with Incomplete Information Formats 313
Shaoxu Song and Lei Chen

Self-tuning in Graph-Based Reference Disambiguation 325
Rabia Nuray-Turan, Dmitri V. Kalashnikov, and Sharad Mehrotra

Probabilistic Nearest-Neighbor Query on Uncertain Objects 337
Hans-Peter Kriegel, Peter Kunath, and Matthias Renz

Indexing and Caching Databases

Making the Most of Cache Groups 349
Andreas Bühmann and Theo Härder

Construction of Tree-Based Indexes for Level-Contiguous Buffering Support 361
Tomáš Skopal, David Hoksza, and Jaroslav Pokorný

A Workload-Driven Unit of Cache Replacement for Mid-Tier Database Caching 374
Xiaodan Wang, Tanu Malik, Randal Burns, Stratos Papadomanolakis, and Anastassia Ailamaki

J$^+$-Tree: A New Index Structure in Main Memory 386
Hua Luan, Xiaoyong Du, Shan Wang, Yongzhi Ni, and Qiming Chen

CST-Trees: Cache Sensitive T-Trees 398
Ig-hoon Lee, Junho Shim, Sang-goo Lee, and Jonghoon Chun

Security and Integrity Maintenance

Specifying Access Control Policies on Data Streams 410
Barbara Carminati, Elena Ferrari, and Kian Lee Tan

Protecting Individual Information Against Inference Attacks in Data Publishing 422
Chen Li, Houtan Shirani-Mehr, and Xiaochun Yang

Quality Aware Privacy Protection for Location-Based Services 434
Zhen Xiao, Xiaofeng Meng, and Jianliang Xu

Implementation of Bitmap Based Incognito and Performance Evaluation 447
Hyun-Ho Kang, Jae-Myung Kim, Gap-Joo Na, and Sang-Won Lee

Prioritized Active Integrity Constraints for Database Maintenance 459
Luciano Caroprese, Sergio Greco, and Cristian Molinaro

Image and Ontology-Based Databases

Using Redundant Bit Vectors for Near-Duplicate Image Detection 472
Jun Jie Foo and Ranjan Sinha

OLYBIA: Ontology-Based Automatic Image Annotation System Using Semantic Inference Rules 485
Kyung-Wook Park, Jin-Woo Jeong, and Dong-Ho Lee

OntoDB: An Ontology-Based Database for Data Intensive Applications 497
Hondjack Dehainsala, Guy Pierra, and Ladjel Bellatreche

Sensor and Scientific Database Applications

Continuously Maintaining Sliding Window Skylines in a Sensor Network 509
Junchang Xin, Guoren Wang, Lei Chen, Xiaoyi Zhang, and Zhenhua Wang

Bayesian Reasoning for Sensor Group-Queries and Diagnosis 522
Ankur Jain, Edward Y. Chang, and Yuan-Fang Wang

Telescope: Zooming to Interesting Skylines 539
Jongwuk Lee, Gae-won You, and Seung-won Hwang

Eliciting Matters – Controlling Skyline Sizes by Incremental Integration of User Preferences 551
Wolf-Tilo Balke, Ulrich Güntzer, and Christoph Lofi

Mobile Databases

Optimizing Moving Queries over Moving Object Data Streams 563
Dan Lin, Bin Cui, and Dongqing Yang

MIME: A Dynamic Index Scheme for Multi-dimensional Query in Mobile P2P Networks 576
Ping Wang, Lidan Shou, Gang Chen, and Jinxiang Dong

Temporal and Spatial Databases

Interval-Focused Similarity Search in Time Series Databases 586
Johannes Aßfalg, Hans-Peter Kriegel, Peer Kröger, Peter Kunath, Alexey Pryakhin, and Matthias Renz

Adaptive Distance Measurement for Time Series Databases 598
Van M. Chhieng and Raymond K. Wong

Clustering Moving Objects in Spatial Networks 611
Jidong Chen, Caifeng Lai, Xiaofeng Meng, Jianliang Xu, and Haibo Hu

Data Streams

The Tornado Model: Uncertainty Model for Continuously Changing Data 624
Byunggu Yu, Seon Ho Kim, Shayma Alkobaisi, Wan D. Bae, and Thomas Bailey

ClusterSheddy: Load Shedding Using Moving Clusters over Spatio-temporal Data Streams 637
Rimma V. Nehme and Elke A. Rundensteiner

Evaluating MAX and MIN over Sliding Windows with Various Size Using the Exemplary Sketch 652
Jiakui Zhao, Dongqing Yang, Bin Cui, Lijun Chen, and Jun Gao

CLAIM: An Efficient Method for Relaxed Frequent Closed Itemsets Mining over Stream Data 664
Guojie Song, Dongqing Yang, Bin Cui, Baihua Zheng, Yunfeng Liu, and Kunqing Xie

P2P and Grid-Based Data Management

Capture Inference Attacks for K-Anonymity with Privacy Inference Logic 676
Xiaojun Ye, Zude Li, and Yongnian Li

Schema Mapping in P2P Networks Based on Classification and Probing 688
Guoliang Li, Beng Chin Ooi, Bei Yu, and Lizhu Zhou

ABIDE: A Bid-Based Economic Incentive Model for Enticing Non-cooperative Peers in Mobile-P2P Networks 703
Anirban Mondal, Sanjay Kumar Madria, and Masaru Kitsuregawa

XML Databases

An Efficient Encoding and Labeling for Dynamic XML Data 715
Jun-Ki Min, Jihyun Lee, and Chin-Wan Chung

An Original Semantics to Keyword Queries for XML Using Structural Patterns 727
Dimitri Theodoratos and Xiaoying Wu

Lightweight Model Bases and Table-Driven Modeling 740
Hung-chih Yang and D. Stott Parker

XML Indexing

An Efficient Index Lattice for XML Query Evaluation 753
Wilfred Ng and James Cheng

A Development of Hash-Lookup Trees to Support Querying Streaming XML 768
James Cheng and Wilfred Ng

Efficient Integration of Structure Indexes of XML 781
Taro L. Saito and Shinichi Morishita

Efficient Support for Ordered XPath Processing in Tree-Unaware Commercial Relational Databases 793
Boon-Siew Seah, Klarinda G. Widjanarko, Sourav S. Bhowmick, Byron Choi, and Erwin Leonardi

XML Query Processing

On Label Stream Partition for Efficient Holistic Twig Join 807
Bo Chen, Tok Wang Ling, M. Tamer Özsu, and Zhenzhou Zhu

Efficient XML Query Processing in RDBMS Using GUI-Driven Prefetching in a Single-User Environment 819
Sandeep Prakash, Sourav S. Bhowmick, Klarinda G. Widjanarko, and C. Forbes Dewey Jr.

Efficient Holistic Twig Joins in Leaf-to-Root Combining with Root-to-Leaf Way 834
Guoliang Li, Jianhua Feng, Yong Zhang, and Lizhu Zhou

TwigList: Make Twig Pattern Matching Fast 850
Lu Qin, Jeffrey Xu Yu, and Bolin Ding

Part II: Short Papers

Query Language and Query Optimization

CircularTrip: An Effective Algorithm for Continuous kNN Queries 863
Muhammad Aamir Cheema, Yidong Yuan, and Xuemin Lin

Optimizing Multiple In-Network Aggregate Queries in Wireless Sensor Networks 870
Huei-You Yang, Wen-Chih Peng, and Chia-Hao Lo

Visible Nearest Neighbor Queries 876
Sarana Nutanong, Egemen Tanin, and Rui Zhang

On Query Processing Considering Energy Consumption for Broadcast Database Systems 884
Shinya Kitajima, Jing Cai, Tsutomu Terada, Takahiro Hara, and Shojiro Nishio

Data Mining and Knowledge Discovery

Mining Vague Association Rules 891
An Lu, Yiping Ke, James Cheng, and Wilfred Ng

An Optimized Process Neural Network Model 898
Guojie Song, Dongqing Yang, Yunfeng Liu, Bin Cui, Ling Wu, and Kunqing Xie

Clustering XML Documents Based on Structural Similarity 905
Guangming Xing, Zhonghang Xia, and Jinhua Guo

The Multi-view Information Bottleneck Clustering 912
Yan Gao, Shiwen Gu, Jianhua Li, and Zhining Liao

Web and Information Retrieval

Web Service Composition Based on Message Schema Analysis 918
Aiqiang Gao, Dongqing Yang, and Shiwei Tang

SQORE: A Framework for Semantic Query Based Ontology Retrieval... 924
Chutiporn Anutariya, Rachanee Ungrangsi, and Vilas Wuwongse

Graph Structure of the Korea Web 930
In Kyu Han, Sang Ho Lee, and Soowon Lee

EasyQuerier: A Keyword Based Interface for Web Database Integration System 936
Xian Li, Weiyi Meng, and Xiaofeng Meng

Database Applications and Security

Anomalies Detection in Mobile Network Management Data..... 943
Marco Anisetti, Claudio A. Ardagna, Valerio Bellandi, Elisa Bernardoni, Ernesto Damiani, and Salvatore Reale

Security-Conscious XML Indexing..... 949
Yan Xiao, Bo Luo, and Dongwon Lee

Framework for Extending RFID Events with Business Rule 955
Mikyeong Moon, Seongjin Kim, Keunhyuk Yeom, and Heeseok Choi

Ontology and Data Streams

Approximate Similarity Search over Multiple Stream Time Series 962
Xiang Lian, Lei Chen, and Bin Wang

WT-Heuristics: A Heuristic Method for Efficient Operator Ordering 969
Jun-Ki Min

An Efficient and Scalable Management of Ontology..... 975
Myung-Jae Park, Jihyun Lee, Chun-Hee Lee, Jiexi Lin, Olivier Serres, and Chin-Wan Chung

Estimating Missing Data in Data Streams..... 981
Nan Jiang and Le Gruenwald

XML Databases

AB-Index: An Efficient Adaptive Index for Branching XML Queries 988
Bo Zhang, Wei Wang, Xiaoling Wang, and Aoying Zhou

Semantic XPath Query Transformation: Opportunities and Performance 994
Dung Xuan Thi Le, Stephane Bressan, David Taniar, and Wenny Rahayu

TGV: A Tree Graph View for Modeling Untyped XQuery 1001
Nicolas Travers, Tuyêt Trâm Dang Ngoc, and Tianxiao Liu

Indexing Textual XML in P2P Networks Using Distributed Bloom Filters 1007
Clement Jamard, Georges Gardarin, and Laurent Yeh

Towards Adaptive Information Merging Using Selected XML Fragments 1013
Ho-Lam Lau and Wilfred Ng

Part III: Posters

Data Warehouse and Data Mining

LAPIN: Effective Sequential Pattern Mining Algorithms by Last Position Induction for Dense Databases 1020
Zhenglu Yang, Yitong Wang, and Masaru Kitsuregawa

Spatial Clustering Based on Moving Distance in the Presence of Obstacles 1024
Sang-Ho Park, Ju-Hong Lee, and Deok-Hwan Kim

Tracing Data Transformations: A Preliminary Report 1028
Gang Qian and Yisheng Dong

Query Processing

QuickCN: A Combined Approach for Efficient Keyword Search over Databases 1032
Jun Zhang, Zhaohui Peng, and Shan Wang

Adaptive Join Query Processing in Data Grids: Exploring Relation Partial Replicas and Load Balancing 1036
Donghua Yang, Jianzhong Li, and Hong Gao

Efficient Semantically Equal Join on Strings 1041
Juggapong Natwichai, Xingzhi Sun, and Maria E. Orlowska

Database Modeling and Information Retrieval

Integrating Similarity Retrieval and Skyline Exploration Via Relevance Feedback 1045
Yiming Ma and Sharad Mehrotra

An Image-Semantic Ontological Framework for Large Image Databases 1050
Xiaoyan Li, Lidan Shou, Gang Chen, and Kian-Lee Tan

Flexible Selection of Wavelet Coefficients for Continuous Data Stream Reduction 1054
Jaehoon Kim and Seog Park

Versioned Relations: Support for Conditional Schema Changes and Schema Versioning 1058
Peter Sune Jørgensen and Michael Böhlen

Network and XML Databases

Compatibility Analysis and Mediation-Aided Composition for BPEL Services 1062
Wei Tan, Fangyan Rao, Yushun Fan, and Jun Zhu

Expert Finding in a Social Network 1066
Jing Zhang, Jie Tang, and Juanzi Li

Efficient Reasoning About XFDs with Pre-image Semantics 1070
Sven Hartmann, Sebastian Link, and Thu Trinh

Part IV: Industrial Track

Context RBAC/MAC Access Control for Ubiquitous Environment 1075
Kyu Il Kim, Hyuk Jin Ko, Hyun Sik Hwang, and Ung Mo Kim

Extending PostgreSQL to Support Distributed/Heterogeneous Query Processing 1086
Rubao Lee and Minghong Zhou

Geo-WDBMS: An Improved DBMS with the Function of Watermarking Geographical Data 1098
Min Huang, Xiang Zhou, Jiaheng Cao, and Zhiyong Peng

Part V: Demonstrations Track

TinTO: A Tool for the View-Based Analysis of Streams of Stock Market Data 1110
Andreas Behrend, Christian Dorau, and Rainer Manthey

Danaïdes: Continuous and Progressive Complex Queries on RSS Feeds 1115
Wee Hyong Tok, Stéphane Bressan, and Mong-Li Lee

OntoDB: It Is Time to Embed Your Domain Ontology in Your Database 1119
Stéphane Jean, Hondjack Dehainsala, Dung Nguyen Xuan, Guy Pierra, Ladjel Bellatreche, and Yamine Aït-Ameur

Author Index 1123