

# **2007 IEEE 23rd International Conference on Data Engineering**

**Istanbul, Turkey  
11-15 April 2007**

**Volume 1 of 3**



**IEEE Catalog Number:**  
**ISBN:**

**07EX1614**  
**1-4244-0802-4**

# Table of Contents

<b>Adaptively Reordering Joins during Query Execution</b> .....	1
<i>Quanzhong Li, Minglong Shao, Volker Markl, Kevin Beyer, Latha Colby, and Guy Lohman</i>	
<b>Hypertree Decompositions for Query Optimization</b> .....	11
<i>Lucantonio Ghionna, Luigi Granata, Gianluigi Greco, and Francesco Scarcello</i>	
<b>Unnesting Scalar SQL Queries in the Presence of Disjunction</b> .....	21
<i>Matthias Brantner, Norman May, and Guido Moerkotte</i>	
<b>Efficient Maintenance of Materialized Outer-Join Views</b> .....	31
<i>Per-Ake Larson and Jingren Zhou</i>	
<b>A General Cost Model for Dimensionality Reduction in High Dimensional Spaces</b> .....	41
<i>Xiang Lian and Lei Chen</i>	
<b>Accelerating Profile Queries in Elevation Maps</b> .....	51
<i>Feng Pan, Wei Wang, and Leonard McMillan</i>	
<b>Selecting Stars: The k Most Representative Skyline Operator</b> .....	61
<i>Xuemin Lin, Yidong Yuan, Qing Zhang, and Ying Zhang</i>	
<b>Computing Compressed Multidimensional Skyline Cubes Efficiently</b> .....	71
<i>Jian Pei, Ada Wai-Chee Fu, Xuemin Lin, and Haixun Wang</i>	
<b>t-Closeness: Privacy Beyond k-Anonymity and l-Diversity</b> .....	81
<i>Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian</i>	
<b>Aggregate Query Answering on Anonymized Tables</b> .....	91
<i>Qing Zhang, Nick Koudas, Divesh Srivastava, and Ting Yu</i>	
<b>Worst-Case Background Knowledge for Privacy-Preserving Data Publishing</b> .....	101
<i>David J. Martin, Daniel Kifer, Ashwin Machanavajjhala, Johannes Gehrke, and Joseph Y. Halpern</i>	
<b>On Randomization, Public Information and the Curse of Dimensionality</b> .....	111
<i>Charu C. Aggarwal</i>	
<b>Adaptive Load Diffusion for Multiway Windowed Stream Joins</b> .....	121
<i>Xiaohui Gu, Philip S. Yu, and Haixun Wang</i>	
<b>Efficiently Processing Continuous k-NN Queries on Data Streams</b> .....	131
<i>Christian Bohm, Beng Chin Ooi, Claudia Plant, and Ying Yan</i>	
<b>ICEDB: Intermittently-Connected Continuous Query Processing</b> .....	141
<i>Yang, Zhang, Bret Hull, Hari Balakrishnan, and Samuel Madden</i>	
<b>A Cooperative, Self-Configuring High-Availability Solution for Stream Processing</b> .....	151
<i>Jeong-Hyon Hwang, Ying Xing, Ugur Cetintemel, and Stan Zdonik</i>	
<b><math>\mu</math>BE: User Guided Source Selection and Schema Mediation for Internet Scale Data Integration</b> .....	161
<i>Ashraf Aboulnaga and Kareem El Gebaly</i>	
<b>Source-aware Entity Matching: A Compositional Approach</b> .....	171
<i>Warren Shen, Pedro DeRose, Long Vu, AnHai Doan, and Raghu Ramakrishnan</i>	
<b>A Semantic Approach to Discovering Schema Mapping Expressions</b> .....	181
<i>Yuan An, Alex Borgida, Renee J. Miller, and John Mylopoulos</i>	
<b>Efficient Approximation Algorithms for Repairing Inconsistent Databases</b> .....	191
<i>Andrei Lopatenko and Loreto Bravo</i>	
<b>eXrQuy: Order Indifference in XQuery</b> .....	201
<i>Torsten Grust, Jan Rittinger, and Jens Teubner</i>	
<b>Combined Static and Dynamic Analysis for Effective Buffer Minimization in Streaming XQuery Evaluation</b> .....	211
<i>Michael Schmidt, Stefanie Scherzinger, and Christoph Koch</i>	

# Table of Contents

<b>Put a Tree Pattern in Your Algebra</b> .....	221
<i>Philippe Michiels, George A. Mihail, and Jerome Simeon</i>	
<b>XQuery Streaming a la Carte</b> .....	231
<i>Mary Fernandez, Philippe Michiels, Jerome Simeon, and Michael Stark</i>	
<b>Autonomic Query Allocation based on Microeconomics Principles</b> .....	241
<i>Fragkiskos Pentaris and Yannis Ioannidis</i>	
<b>Exploiting Correlation and Parallelism of Materialized-View Recommendation for Distributed Data Warehouses</b> .....	251
<i>Haifeng Jiang, Dengfeng Gao, and Wen-Syan Li</i>	
<b>LIRA: Lightweight, Region-aware Load Shedding in Mobile CQ Systems</b> .....	261
<i>Bugra Gedik, Ling Liu, Kun-Lung Wu, and Philip S. Yu</i>	
<b>SD-Rtree: A Scalable Distributed Rtree</b> .....	271
<i>Cedric du Mouza, Witold Litwin, and Philippe Rigaux</i>	
<b>Categorization and Optimization of Synchronization Dependencies in Business Processes</b> .....	281
<i>Qinyi Wu, Calton Pu, Akhil Sahai, and Roger Barga</i>	
<b>Invalidation Clues for Database Scalability Services</b> .....	291
<i>Amit Manjhi, Phillip B. Gibbons, Anastassia Ailamaki, Charles Garrod, Bruce M. Maggs, Todd C. Mowry, Christopher Olston, Anthony Tomasic, and Haifeng Yu</i>	
<b>Organizing Hidden-Web Databases by Clustering Visible Web Documents</b> .....	301
<i>Luciano Barbosa, Juliana Freire, Altigran Silva</i>	
<b>Programming Patterns for Architecture-Level Software Optimizations on Frequent Pattern Mining</b> .....	311
<i>Mingliang Wei, Changhao Jiang, and Marc Snir</i>	
<b>Efficient Keyword Search Across Heterogeneous Relational Databases</b> .....	321
<i>Mayssam Sayyadian, Hieu LeKhac, AnHai Doan, and Luis Gravano</i>	
<b>Preference-Aware Query and Update Scheduling in Web-databases</b> .....	331
<i>Huiming Qu and Alexandros Labrinidis</i>	
<b>Ontology-Based Constraint Recognition for Free-Form Service Requests</b> .....	341
<i>Muhammed J. Al-Muhammed. and David W. Embley.</i>	
<b>Annotating Structured Data of the Deep Web</b> .....	351
<i>Yiyao Lu, Hai He, Hongkun Zhao, Weiyi Meng, and Clement Yu</i>	
<b>Clustering wavelets to speed-up data dissemination in structured P2P MANETs</b> .....	361
<i>Mihai Lupu, Jianzhong Li, Beng Chin Ooi, and Shengfei Shi</i>	
<b>Load Distribution Fairness in P2P Data Management Systems</b> .....	371
<i>Theoni Pitoura and Peter Triantafillou</i>	
<b>Distributed File Structures in a Peer-to-Peer Environment</b> .....	381
<i>Stratis D. Viglas</i>	
<b>SKYPEER: Efficient Subspace Skyline Computation over Distributed Data</b> .....	391
<i>Akrivi Vlachou, Christos Doulkeridis, Yannis Kotidis, and Michalis Vazirgiannis</i>	
<b>On k-Nearest Neighbor Searching in Non-Ordered Discrete Data Spaces</b> .....	401
<i>Dashiell Kolbe, Qiang Zhu, and Sakti Pramanik</i>	
<b>The Haar Tree: A Refined Synopsis Data Structure</b> .....	411
<i>Panagiotis Karras, Nikos Mamoulis</i>	
<b>Calibration and Profile based Synopses Error Estimation and Synopses Reconciliation</b> .....	421
<i>Yariv Matias and Yossi Matias</i>	
<b>DIKNN: An Itinerary-based KNN Query Processing Algorithm for Mobile Sensor Networks</b> .....	431
<i>Shan-Hung Wu, Kun-Ta Chuang, Chung-Min Chen, and Ming-Syan Chen</i>	

# Table of Contents

<b>Materialization Strategies in a Column-Oriented DBMS</b> .....	441
<i>Daniel J. Abadi, Daniel S. Myers, David J. DeWitt, Samuel R. Madden</i>	
<b>Supporting Streaming Updates in an Active Data Warehouse</b> .....	451
<i>Neoklis Polyzotis, Spiros Skiadopoulos, Panos Vassiliadis, Alkis Simitis, and Nils-Erik Frantzell</i>	
<b>DeltaSky: Optimal Maintenance of Skyline Deletions without Exclusive Dominance Region Generation</b> .....	461
<i>Jing Wu, Divyakant Agrawal, Omer Egecioglu, Amr El Abbadi</i>	
<b>Group Linkage</b> .....	471
<i>Byung-Won On, Nick Koudas, Dongwon Lee, Divesh Srivastava</i>	
<b>Reverse Query Processing</b> .....	481
<i>Carsten Binnig, Donald Kossmann, and Eric Lo</i>	
<b>Collecting and Maintaining Just-in-Time Statistics</b> .....	491
<i>Amr El-Helw, Ihab F. Ilyas, Wing Lau, Volker Markl, and Calisto Zuzarte</i>	
<b>Dynamic Materialized Views</b> .....	501
<i>Jingren Zhou, Per-Ake Larson, Jonathan Goldstein and Luping Ding</i>	
<b>A Load Shedding Framework and Optimizations for M-way Windowed Stream Joins</b> .....	511
<i>Bugra Gedik, Kun-Lung Wu, Philip S. Yu, and Ling Liu</i>	
<b>Labeling network motifs in protein interactomes for protein function prediction</b> .....	521
<i>Jin Chen, Wynne Hsu, Mong Li Lee, and See-Kiong Ng</i>	
<b>Topology Search over Biological Databases</b> .....	531
<i>Lin Guo, Jayavel Shanmugasundaram, Golan Yona</i>	
<b>GString: A Novel Approach for Efficient Search in Graph Databases</b> .....	541
<i>Haoliang Jiang, Haixun Wang, Philip S. Yu, and Shuigeng Zhou</i>	
<b>Patterns Discovery on Complex Diagnosis and Biological Data Using Fuzzy Latent Variables</b> .....	551
<i>Zong-Xian Yin and Jung-Hsien Chiang</i>	
<b>Efficient Evaluation of Imprecise Location-Dependent Queries</b> .....	561
<i>Jinchuan Chen and Reynold Cheng</i>	
<b>Representing and Querying Correlated Tuples in Probabilistic Databases</b> .....	571
<i>Prithviraj Sen and Amol Deshpande</i>	
<b>10106 Worlds and Beyond: Efficient Representation and Processing of Incomplete Information</b> .....	581
<i>Lyublena Antova, Christoph Koch, and Dan Olteanu</i>	
<b>Indexing Uncertain Categorical Data</b> .....	591
<i>Sarvjeet Singh, Chris Mayfield, Sunil Prabhakar, Rahul Shah, and Susanne Hambrusch</i>	
<b>Structural Selectivity Estimation for XML Documents</b> .....	601
<i>Damien K. Fisher and Sebastian Maneth</i>	
<b>XPlainer: Visual Explanations of XPath Queries</b> .....	611
<i>Mariano P. Consens, John W. S. Liu, Flavio Rizzolo</i>	
<b>Complex Group-By Queries for XML</b> .....	621
<i>C. Gokhale, N. Gupta, P. Kumar, L.V.S. Lakshmanan, R. Ng, and B.A. Prakash</i>	
<b>Branch Sequencing Based XML Message Broker Architecture</b> .....	631
<i>Aneesh Raj and P.Sreenivasa Kumar</i>	
<b>Rewriting Regular XPath Queries on XML Views</b> .....	641
<i>Wenfei Fan, Floris Geerts, Xibei Jia, and Anastasios Kementsietsidis</i>	
<b>Fast, Secure Encryption for Indexing in a Column-Oriented DBMS</b> .....	651
<i>Tingjian Ge and Stan Zdonik</i>	
<b>Hiding in the Crowd: Privacy Preservation on Evolving Streams through Correlation Tracking</b> .....	661
<i>Feifei Li, Jimeng Sun, Spiros Papadimitriou, George A. Mihaila, and Ioana Stanoi</i>	

# Table of Contents

<b>Preservation Of Patterns and Input-Output Privacy .....</b>	<b>671</b>
<i>Shaofeng Bu, Laks V.S. Lakshmanan, Raymond T. Ng and Ganesh Ramesh</i>	
<b>Mining Colossal Frequent Patterns by Core Pattern Fusion.....</b>	<b>681</b>
<i>Feida Zhu, Xifeng Yan, Jiawei Han, Philip S. Yu, and Hong Cheng</i>	
<b>Discriminative Frequent Pattern Analysis for Effective Classification .....</b>	<b>691</b>
<i>Hong Cheng, Xifeng Yan, Jiawei Han, and Chih-Wei Hsu</i>	
<b>Discovering Relational Patterns across Multiple Databases.....</b>	<b>701</b>
<i>Xingquan Zhu and Xindong Wu</i>	
<b>Distributed Data Stream Clustering: A Fast EM-based Approach.....</b>	<b>711</b>
<i>Aoying Zhou, Feng Cao, Ying Yan, Chaofeng Sha, and Xiaofeng He</i>	
<b>Conditional Functional Dependencies for Data Cleaning .....</b>	<b>721</b>
<i>Philip Bohannon, Wenfei Fan, Floris Geerts, Xibei Jia, and Anastasios Kementsietsidis</i>	
<b>Partial Materialized Views.....</b>	<b>731</b>
<i>Gang Luo</i>	
<b>Updating Recursive XML Views of Relations.....</b>	<b>741</b>
<i>Byron Choi, Gao Cong, Wenfei Fan, Stratis D. Viglas</i>	
<b>Fast Identification of Relational Constraint Violations.....</b>	<b>751</b>
<i>Amit Chandel, Nick Koudas, Ken Q. Pu, Divesh Srivastava</i>	
<b>SpADe: On Shape-based Pattern Detection in Streaming Time Series .....</b>	<b>761</b>
<i>Yueguo Chen, Mario A. Nascimento, Beng Chin Ooi, and Anthony K. H. Tung</i>	
<b>Multi-source Skyline Query Processing in Road Networks .....</b>	<b>771</b>
<i>Ke Deng, Xiaofang Zhou, and Heng Tao Shen</i>	
<b>Continuous Evaluation of Monochromatic and Bichromatic Reverse Nearest Neighbors. ....</b>	<b>781</b>
<i>James M. Kang, Mohamed F. Mokbel, Shashi Shekhar, Tian Xia, and Donghui Zhang</i>	
<b>Index-based Most Similar Trajectory Search .....</b>	<b>791</b>
<i>Elias Frentzos, Kostas Gratsias, and Yannis Theodoridis</i>	
<b>An Online Approach to Physical Design Tuning .....</b>	<b>801</b>
<i>Nicolas Bruno and Surajit Chaudhuri</i>	
<b>Finding Top-k Min-Cost Connected Trees in Databases .....</b>	<b>811</b>
<i>Bolin Ding, Jeffrey Xu Yu, Shan Wang, Lu Qin, Xiao Zhang, and Xuemin Lin</i>	
<b>Adding Context to Preferences.....</b>	<b>821</b>
<i>Kostas Stefanidis, Evaggelia Pitoura and Panos Vassiliadis</i>	
<b>Reducing Order Enforcement Cost in Complex Query Plans.....</b>	<b>831</b>
<i>Ravindra Guravannavar and S Sudarshan</i>	
<b>On Density Based Transforms for Uncertain Data Mining .....</b>	<b>841</b>
<i>Charu C. Aggarwal</i>	
<b>Evaluating Proximity Relations Under Uncertainty.....</b>	<b>851</b>
<i>Zhengdao Xu and Hans-Arno Jacobsen</i>	
<b>Efficient Top-k Query Evaluation on Probabilistic Data.....</b>	<b>861</b>
<i>Christopher Re, Nilesh Dalvi, Dan Suciu</i>	
<b>Top-k Query Processing in Uncertain Databases .....</b>	<b>871</b>
<i>Mohamed A. Soliman, Ihab F. Ilyas, and Kevin Chen-Chuan Chang</i>	
<b>Personalizing XML Search in PIMENTO.....</b>	<b>881</b>
<i>Sihem Amer-Yahia, Iirini Fundulaki, and Laks V.S. Lakshmanan</i>	
<b>X<sup>3</sup>: A Cube Operator for XML OLAP.....</b>	<b>891</b>
<i>Nuwee Wiwatwattana, H. V. Jagadish, Laks V. S. Lakshmanan, and Divesh Srivastava</i>	

# Table of Contents

<b>MultiMap: Preserving disk locality for multidimensional datasets</b> .....	901
<i>Minglong Shao, Steven W. Schlosser, Stratos Papdomanolakis, Jiri Schindler, Anastassia Ailamaki, and Gregory R. Ganger</i>	
<b>Indexing Point Triples Via Triangle Geometry</b> .....	911
<i>Charles B. Cranston and Hanan Samet</i>	
<b>APLA: Indexing Arbitrary Probability Distributions</b> .....	921
<i>Vebjorn Ljosa and Ambuj K. Singh</i>	
<b>Online Bulk Deletion</b> .....	931
<i>Timo Lilja, Riku Saikkonen, Seppo Sippu, and Eljas Soisalon-Soininen</i>	
<b>TreePi: A Novel Graph Indexing Method</b> .....	941
<i>Shijie Zhang, Meng Hu, and Jiong Yang</i>	
<b>Graph Database Indexing Using Structured Graph Decomposition</b> .....	951
<i>David W. Williams, Jun Huan, and Wei Wang</i>	
<b>Many-to-Many Aggregation for Sensor Networks</b> .....	961
<i>Adam Silberstein and Jun Yang</i>	
<b>Proof Sketches: Verifiable In-Network Aggregation</b> .....	971
<i>Minos Garofalakis, Joseph M. Hellerstein, and Petros Maniatis</i>	
<b>Norm, Point, and Distance Estimation Over Multiple Signals Using Max - Stable Distributions</b> .....	981
<i>Stilian A. Stoev, Marios Hadjieleftheriou, George Kollios, and Murad S. Taqqu</i>	
<b>Tree-Pattern Similarity Estimation for Scalable Content-based Routing</b> .....	991
<i>Raphael Chand, Pascal Felber, and Minos Garofalakis</i>	
<b>Space Efficient Streaming Algorithms for the Maximum Error Histogram</b> .....	1001
<i>Chiranjeeb Buragohain, Nisheeth Shrivastava, and Subhash Suri</i>	
<b>Conquering the Divide: Continuous Clustering of Distributed Data Streams</b> .....	1011
<i>Graham Cormode, S. Muthukrishnan, and Wei Zhuang</i>	
<b>Stream Monitoring under the Time Warping Distance</b> .....	1021
<i>Yasushi Sakurai, Christos Faloutsos, and Masashi Yamamuro</i>	
<b>Efficient Evaluation of All-Nearest-Neighbor Queries</b> .....	1031
<i>Yun Chen and Jignesh M. Patel</i>	
<b>Pointwise-Dense Region Queries in Spatio-temporal Databases</b> .....	1041
<i>Jinfeng Ni and Chinya V. Ravishankar</i>	
<b>Top-k Spatial Preference Queries.</b> .....	1051
<i>Man Lung Yiu, Xiangyuan Dai, Nikos Mamoulis, and Michail Vaitis</i>	
<b>Similarity Match Over High Speed Time-Series Streams</b> .....	1061
<i>Xiang Lian, Lei Chen, Jeffrey Xu Yu, Guoren Wang, and Ge Yu</i>	
<b>Scalable Peer-to-Peer Web Retrieval with Highly Discriminative Keys</b> .....	1071
<i>Ivana Podnar, Martin Rajman, Toan Luu, Fabius Klemm, and Karl Aberer</i>	
<b>SPRITE: A Learning-Based Text Retrieval System in DHT Networks</b> .....	1081
<i>Yingguang Li, H.V. Jagadish, and Kian-Lee Tan</i>	
<b>Distributed Processing of Moving K-Nearest-Neighbor Query on Moving Objects</b> .....	1091
<i>Wei Wu, Wenyuan Guo, and Kian-Lee Tan</i>	
<b>Efficient Skyline Query Processing on Peer-to-Peer Networks</b> .....	1101
<i>Shiyuan Wang, Beng Chin Ooi, Anthony K. H. Tung, and Lizhen Xu</i>	
<b>Increasing Buffer-Locality for Multiple Relational Table Scans through Grouping and Throttling</b> .....	1111
<i>Christian A. Lang, Bishwaranjan Bhattacharjee, Tim Malkemus, Sriram Padmanabhan, and Kwai Wong</i>	

# Table of Contents

<b>Propagating Updates in SPIDER</b> .....	1121
<i>Nick Koudas, Amit Marathe, and Divesh Srivastava</i>	
<b>Optimizing Concurrency Through Automated Lock Memory Tuning in DB2</b> .....	1129
<i>Sam S. Lightstone, Chris Eaton, Yun Han Lee, and Adam J. Storm</i>	
<b>Optimizing Similar Scalar Subqueries for XML Processing in Microsoft SQL Server</b> .....	1139
<i>Adrian Baras, César A. Galindo-Legaria, Torsten Grabs, Babu Krishnaswamy, and Shankar Pal</i>	
<b>Fine Grained Authorization Through Predicated Grants</b> .....	1149
<i>Surajit Chaudhuri, Tanmoy Dutta, and S. Sudarshan</i>	
<b>RFID Data Processing with a Data Stream Query Language</b> .....	1159
<i>Yijian Bai, Fusheng Wang, Peiya Liu, Carlo Zaniolo, and Shaorong Liu</i>	
<b>XMLTable Index— An Efficient Way of Indexing and Querying XML Property Data</b> .....	1169
<i>Zhen Hua Liu, Muralidhar Krishnaprasad, Hui J. Chang, and Vikas Arora</i>	
<b>Transaction Isolation and Lazy Commit</b> .....	1179
<i>Vishal Kathuria, Robin Dhamankar, and Hanuma Kodavalla</i>	
<b>Parts Traceability for Manufacturers</b> .....	1187
<i>Christine Robson, Yuji Watanabe, and Masayuki Numao</i>	
<b>Information System Construction and Maintenance Based on Business Terminology: Experiences from the Reinsurance Industry</b> .....	1197
<i>Hans Wegener, Swiss Re</i>	
<b>Web-based Collaborative Information Integration for Scientific Research</b> .....	1207
<i>Fusheng Wang, Peiya Liu, Cornelius Rabsch, Patrick Kling, and John Pearson</i>	
<b>Object Distinction: Distinguishing Objects with Identical Names</b> .....	1217
<i>Xiaoxin Yin, Jiawei Han, and Philip S. Yu</i>	
<b>Monkey: Approximate Graph Mining Based on Spanning Trees</b> .....	1222
<i>Shijie Zhang, Jiong Yang, and VenuMadhav Cheedella</i>	
<b>Distance Based Subspace Clustering with Flexible Dimension Partitioning</b> .....	1225
<i>Guimei Liu, Jinyan Li, Kelvin Sim, and Limsoon Wong</i>	
<b>SQL Queries Over Unstructured Text Databases</b> .....	1230
<i>Alpa Jain, AnHai Doan, Luis Gravano</i>	
<b>Disambiguation Algorithm for People Search on the Web</b> .....	1233
<i>Dmitri V. Kalashnikov, Sharad Mehrotra, Zhaoqi Chen, Rabia Nuray-Turan, and Naveen Ashish</i>	
<b>Collaborative Wrapping: A Turbo Framework for Web Data Extraction</b> .....	1236
<i>Shui-Lung Chuang, Kevin Chen-Chuan Chang, and ChengXiang Zhai</i>	
<b>Finding Skyline and Top-k Bargaining Solutions</b> .....	1238
<i>Mohamed A. Soliman, Ihab F. Ilyas, and Nick Koudas</i>	
<b>Cost-Conscious Cleaning of Massive RFID Data Sets</b> .....	1243
<i>Hector Gonzalez, Jiawei Han, and Xuehua Shen</i>	
<b>Parsimonious Explanations of Change in Hierarchical Data</b> .....	1248
<i>Dhiman Barman, Flip Korn, Divesh Srivastava, Dimitrios Gunopulos, Neal Young and Deepak Agarwal</i>	
<b>The Multi-Relational Skyline Operator</b> .....	1251
<i>Wen Jin, Martin Ester, Zengjian Hu and Jiawei Han</i>	
<b>Robust Heuristics for Scalable Optimization of Complex SQL Queries</b> .....	1256
<i>Gopal Chandra Das and Jayant R. Haritsa</i>	
<b>Semantic Prefetching of Correlated Query Sequences</b> .....	1259
<i>Ivan T. Bowman and Kenneth Salem</i>	

# Table of Contents

<b>On Suspending and Resuming Dataflows</b> .....	1264
<i>Badrish Chandramouli, Christopher N. Bond, Shivnath Babu, and Jun Yang</i>	
<b>A Lightweight Online Framework For Query Progress Indicators</b> .....	1267
<i>Chaitanya Mishra and Nick Koudas</i>	
<b>Efficient Hash Probes on Modern Processors</b> .....	1272
<i>Kenneth A. Ross</i>	
<b>SAO: A Stream Index for Answering Linear Optimization Queries</b> .....	1277
<i>Gang Luo, Kun-Lung Wu and Philip S. Yu</i>	
<b>Boundary-Based Lower-Bound Functions for Dynamic Time Warping and Their Indexing</b> .....	1282
<i>Mi ZHOU and Man Hon WONG</i>	
<b>On MBR Approximation of Histories for Historical Queries: Expectations and Limitations</b> .....	1287
<i>Reza Sherkat and Davood Rafiei</i>	
<b>Monitoring Regular Expressions on Out-of-Order Streams</b> .....	1290
<i>Theodore Johnson, S. Muthukrishnan and Irina Rozenbaum</i>	
<b>Efficient Detection of Distributed Constraint Violations</b> .....	1295
<i>Shipra Agrawal, Supratim Deb, K. V. M. Naidu, Rajeev Rastogi</i>	
<b>Data-quality Guided Load Shedding for Expensive In-Network Data Processing</b> .....	1300
<i>Lina Peng and K. Selcuk Candan</i>	
<b>Summarizing Order Statistics over Data Streams with Duplicates</b> .....	1304
<i>Ying Zhang, Xuemin Lin, Yidong Yuan, Masaru Kitsuregawa, Xiaofang Zhou and Jeffrey Xu Yu,</i>	
<b>Optimizing Timestamp Management in Data Stream Management Systems</b> .....	1309
<i>Yijian Bai, Hetal Thakkar, Haixun Wang, and Carlo Zaniolo</i>	
<b>Multiple Query Optimization for Wireless Sensor Networks</b> .....	1314
<i>Shili Xiang, Hock Beng Lim, and Kian-Lee Tan</i>	
<b>TCAM-conscious Algorithms for Data Streams</b> .....	1317
<i>Nagender Bandi, Ahmed Metwally, Divyakant Agrawal, Amr El Abbadi</i>	
<b>Scoop: An Adaptive Indexing Scheme for Stored Data in Sensor Networks</b> .....	1320
<i>Thomer M. Gil and Samuel Madden</i>	
<b>CPS-tree: A Compact Partitioned Suffix Tree for Disk-based Indexing on Large Genome Sequences</b> .....	1325
<i>Swee-Seong Wong, Wing-Kin Sung, and Limsoon Wong</i>	
<b>Location-based Spatial Queries with Data Sharing in Wireless Broadcast Environments</b> .....	1330
<i>Wei-Shinn Ku, Roger Zimmermann, and Haixun Wang</i>	
<b>Super Peer Web Service Discovery Architecture</b> .....	1335
<i>Evren Ayorak and Dr. Ayse Basar Bener</i>	
<b>Oscar: A Data-Oriented Overlay For Heterogeneous Environments</b> .....	1340
<i>Sarunas Girdzijauskas, Anwitaman Datta, and Karl Aberer</i>	
<b>Making Applications Persistent at Run-time.</b> .....	1343
<i>Angela Nicoara and Gustavo Alonso</i>	
<b>Efficient Filter Based Replication for LDAP Directories</b> .....	1348
<i>Apurva Kumar</i>	
<b>Building a Network-Aware and Load-Balanced Structured Peer-to-Peer System for Range Query</b> .....	1351
<i>Yuh-Jzer Joung and Yi-Fang Chou</i>	
<b>Support Concurrent Queries in Multiuser CBIR Systems</b> .....	1354
<i>Danzhou Liu and Kien A. Hua</i>	
<b>Generalized Precis Queries for Logical Database Subset Creation</b> .....	1357
<i>Alkis Simitsis, Georgia Koutrika, and Yannis Ioannidis</i>	



# Table of Contents

<b>Abstract Process Data Warehousing</b> .....	1362
<i>Fabio Casati, Malu Castellanos, Norman Salazar and Umeshwar Dayal</i>	
<b>Substring Matching in P2P Publish/Subscribe Data Management Networks</b> .....	1365
<i>Ioannis Aekaterinidis, Peter Triantafyllou</i>	
<b>Dynamic Batch Nearest Neighbor Search in Video Retrieval</b> .....	1370
<i>Jie Shao, Zi Huang, Heng Tao Shen, Xiaofang Zhou, and Yijun Li</i>	
<b>XML Routing in Data Dissemination Networks</b> .....	1375
<i>Guoli Li, Shuang Hou, and Hans-Arno Jacobsen</i>	
<b>Protocol Discovery from Imperfect Service Interaction Logs</b> .....	1380
<i>H.R. Motahari-Nezhad, R. Saint-Paul, B. Benatallah, and F. Casati</i>	
<b>PASS Middleware for Distributed and Autonomous XML Message Processing</b> .....	1385
<i>Dirceu Cavendish and K. Selcuk Candan</i>	
<b>K-Anonymization as Spatial Indexing: Toward Scalable and Incremental Anonymization</b> .....	1389
<i>Tochukwu Iwuchukwu, David J. DeWitt, AnHai Doan, and Jeffrey F. Naughton</i>	
<b>MultiRelational k-Anonymity</b> .....	1392
<i>M. Ercan Nergiz, Chris Clifton, and A. Erhan Nergiz</i>	
<b>On Multidimensional k-Anonymity with Local Recoding Generalization</b> .....	1397
<i>Yang Du, Tian Xia, Yufei Tao, Donghui Zhang, and Feng Zhu</i>	
<b>Blocking Inference Channels in Frequent Pattern Sharing</b> .....	1400
<i>Zhihui Wang, Wei Wang, and Baile Shi</i>	
<b>QPIAD: Query Processing over Incomplete Autonomous Databases</b> .....	1405
<i>Hemal Khatri, Jianchun Fan, Yi Chen, and Subbarao Kambhampati</i>	
<b>LiveConcept: Web Search using Structured Query</b> .....	1408
<i>Pu-Jeng Cheng, Ching-Hsiang Tsai, and Chen-Ming Hung</i>	
<b>Semantic Data Management: Towards Querying Data with their Meaning</b> .....	1413
<i>Lipyeow Lim, Haixun Wang, Min Wang</i>	
<b>Mapping a database into an ontology: an interactive relational learning approach</b> .....	1418
<i>Benjamin Habegger</i>	
<b>Efficiently Detecting Inclusion Dependencies</b> .....	1423
<i>Jana Bauckmann, Ulf Leser, Felix Naumann, Veronique Tietz</i>	
<b>Using Queries to Associate Metadata with Data</b> .....	1426
<i>Divesh Srivastava and Yannis Velegarakis</i>	
<b>Piggyback Optimization of XML Data Dissemination</b> .....	1429
<i>Chee-Yong Chan and Yuan Ni</i>	
<b>SIGOPT: Using Schema to Optimize XML Query Processing</b> .....	1431
<i>Stelios Pappas, Jignesh M. Patel and H. V. Jagadish</i>	
<b>Bringing Precision to Desktop Search: A Predicate-based Desktop Search Architecture</b> .....	1436
<i>Jens-Peter Dittrich, Cristian Duda, Björn Jarisch, Donald Kossmann, Marcos Antonio Vaz Salles</i>	
<b>RoXSum: Leveraging Data Aggregation and Batch Processing for XML Routing</b> .....	1441
<i>Zografoula Vagena, Mirella M. Moro, and Vassilis J. Tsotras</i>	
<b>X-Time: Schema Versioning and Cross-Version Querying in Data Warehouses</b> .....	1446
<i>Stefano Rizzi and Matteo Golfarelli</i>	
<b>GPress: Towards Effective GML Documents Compression</b> .....	1448
<i>Jihong Guan and Shuigeng Zhou</i>	
<b>Database Summarization: The SaintEtiQ System</b> .....	1450
<i>Regis Saint-Paul, Guillaume Raschia, Nouredine Mouaddib</i>	

# Table of Contents

<b>R-Cubes: OLAP Cubes Contextualized with Documents .....</b>	<b>1452</b>
<i>Juan Manuel Perez, Rafael Berlanga, Marya Jose Aramburu and Torben Bach Pedersen</i>	
<b>MayBMS: Managing Incomplete Information with Probabilistic World-Set Decompositions .....</b>	<b>1454</b>
<i>Lyublena Antova, Christoph Koch, and Dan Olteanu</i>	
<b>MMS: Using Queries As Data Values for Metadata Management.....</b>	<b>1456</b>
<i>Divesh Srivastava and Yannis Velegarakis</i>	
<b>STAR: A System for Tuple and Attribute Ranking of Query Answers.....</b>	<b>1458</b>
<i>Nishant Kapoor, Gautam Das, Vagelis Hristidis, S. Sudarshan and Gerhard Weikum</i>	
<b>Pushing Business Data Processing Towards the Periphery .....</b>	<b>1460</b>
<i>Pablo Guerrero, Kai Sachs, Mariano Cilia, Christof Bornhovd, Alejandro Buchmann</i>	
<b>Creating Nested Mappings with Clío .....</b>	<b>1462</b>
<i>Mauricio A. Hernandez, Howard Ho, Lucian Popa, Takeshi Fukuda, Ariel Fuxman, Renee J. Miller, Paolo Papotti</i>	
<b>Faceted Browsing over Large Databases of Text-Annotated Objects.....</b>	<b>1464</b>
<i>Wisam Dakka, Panagiotis G. Ipeirotis, and Kenneth R. Wood</i>	
<b>Using Control Theory for Load Shedding in Data Stream Management .....</b>	<b>1466</b>
<i>Yi-Cheng Tu, Song Liu, Sunil Prabhakar, Bin Yao and William Schroeder</i>	
<b>Odysseus: a High-Performance ORDBMS Tightly-Coupled with Spatial Database Features .....</b>	<b>1468</b>
<i>Kyu-Young Whang, Jae-Gil Lee, Min-Soo Kim, Min-Jae Lee and Ki-Hoon Lee</i>	
<b>Theseos: A Query Engine for Traceability across Sovereign, Distributed RFID Databases .....</b>	<b>1470</b>
<i>Alvin Cheung, Karin Kailing and Stefan Schonauer</i>	
<b>Service Mosaic: Interactive Analysis and Manipulation of Service Conversations .....</b>	<b>1472</b>
<i>H. Motahari-Nezhad, R. Saint-Paul, B. Benatallah, F. Casati, J. Ponge, F. Toumani</i>	
<b>The New Casper: A Privacy-Aware Location-Based Database Server.....</b>	<b>1474</b>
<i>Mohamed F. Mokbel, Chi-Yin Chow, and Walid G. Aref</i>	
<b>P2PTester: a tool for measuring P2P platform performance .....</b>	<b>1476</b>
<i>Bogdan Butnaru, Florin Dragan, Georges Gardarin, Ioana Manolescu, Benjamin Nguyen, Radu Pop, Nicoleta Preda and Laurent Yeh</i>	
<b>UniStore: Querying a DHT-based Universal Storage.....</b>	<b>1478</b>
<i>Marcel Karnstedt, Kai-Uwe Sattler, Manfred Hauswirth, Roman Schmidt, Martin Richtarsky, Jessica Muller, Renault John</i>	
<b>Adapting Relational Database Engine to Accommodate Moving Objects in SpADE.....</b>	<b>1480</b>
<i>Beng Chin Ooi, Zhiyong Huang, Dan Lin, Hua Lu, Linhao Xu</i>	
<b>FireStream: Sensor Stream Processing for Monitoring Fire Spread.....</b>	<b>1482</b>
<i>Venkatesh Raghavan, Elke Rundensteiner, John Woycheese and Abhishek Mukherji</i>	
<b>Attention Attractor:efficient video stream similarity query processing in real time.....</b>	<b>1484</b>
<i>Ira Assent, Ralph Krieger and Thomas Seidl</i>	
<b>Dewex: An Exploration Facility for Enabling the Deep Web Integration.....</b>	<b>1486</b>
<i>Govind Kabra, Zhen Zhang and Kevin Chen-Chuan Chang</i>	
<b>µBE: User Guided Source Selection and Schema Mediation for Internet Scale Data Integration.....</b>	<b>1488</b>
<i>Ashraf Aboulnaga, Kareem El Gebaly, Daniel Wong</i>	
<b>DSphere: A Source-Centric Approach to Crawling, Indexing and Searching the World Wide Web.....</b>	<b>1490</b>
<i>Bhuvan Bamba, Ling Liu, James Caverlee, Vaibhav Padliya, Mudhakar Srivatsa, Tushar Bansal, Mahesh Palekar, Joseph Patrao, Suiyang Li and Aameek Singh</i>	
<b>Evolution of Distributed Web Data: An Application of the Reactive Language XChange.....</b>	<b>1492</b>
<i>Francois Bry, Michael Eckert, Hendrik Grallert and Paula-Lavinia Patranjan</i>	

# Table of Contents

<b>PIBE: Manage Your Images the Way You Want!</b> .....	1494
<i>Ilaria Bartolini, Paolo Ciaccia, Marco Patella</i>	
<b>ProVeR: Probabilistic Video Retrieval using the Gauss-Tree</b> .....	1496
<i>Christian Bohm, Michael Gruber, Peter Kunath, Alexey Pryakhin and Matthias Schubert</i>	
<b>The Video Event Awareness System</b> .....	1498
<i>Dimitrios Georgakopoulos and Donald Baker</i>	
<b>A Tool for Information Retrieval Research in Peer-to-Peer File Sharing Systems</b> .....	1500
<i>Linh Thai Nguyen, Wai Gen Yee, Dongmei Jia and Ophir Frieder</i>	