CONTENTS

KEYNOTES

Databases in a Wireless World 3
David Yach (Research in Motion)

Structures, Semantics and Statistics 4
Alon Halevy (Univ. of Washington)

TEN YEAR BEST PAPER AWARD

Whither Data Mining? 9
Rakesh Agrawal and Ramakrishnan Srikant (IBM Almaden)

RESEARCH SESSIONS

Research Session 1: Compression & Indexing

Compressing Large Boolean Matrices using Reordering Techniques 13
David S. Johnson, Shankar Krishnan (AT&T Labs-Research), Jatin Chhugani, Subodh Kumar (John Hopkins Univ.), Suresh Venkatasubramanian (AT&T Labs-Research)

On the Performance of Bitmap Indices for High Cardinality Attributes 24
Kesheng Wu, Ekow Otoo, Arie Shoshani (Lawrence Berkeley National Laboratory)

Practical Suffix Tree Construction 36
Sandeep Tata, Richard A. Hankins, Jignesh M. Patel (Univ. of Michigan)

Research Session 2: XML Views and Schemas

Answering XPath Queries over Networks by Sending Minimal Views 48
Keishi Tajima, Yoshiki Fukui (JAIST)

A Framework for Using Materialized XPath Views in XML Query Processing 60
Andrey Balmin, Fatma Özcan, Kevin Beyer, Roberta Cochrane, Hamid Pirahesh (IBM Almaden)

Schema-Free XQuery 72
Yunyao Li, Cong Yu, H. V. Jagadish (Univ. of Michigan)

Research Session 3: Controlling Access

Client-Based Access Control Management for XML Documents 84
Luc Bouganim (INRIA Rocquencourt), François Dang Ngoc, Philippe Pucheral (PRiSM Laboratory)

Secure XML Publishing without Information Leakage in the Presence of Data Inference 96
Xiaochun Yang (Northeastern Univ.), Chen Li (Univ. of California, Irvine)

Limiting Disclosure in Hippocratic Databases 108
Kristen LeFevre (IBM Almaden and Univ. of Wisconsin-Madison), Rakesh Agrawal (IBM Almaden), Vuk Ercegovac, Raghu Ramakrishnan (Univ. of Wisconsin-Madison), Yirong Xu (IBM Almaden), David DeWitt (Univ. of Wisconsin-Madison)

Research Session 4: XML (I)

On Testing Satisfiability of Tree Pattern Queries 120
Laks V.S. Lakshmanan, Ganesh Ramesh, Hui (Wendy) Wang, Zheng (Jessica) Zhao (Univ. of British Columbia)

Containment of Nested XML Queries 132
Xin Dong, Alon Halevy, Igor Tatarinov (Univ. of Washington)

Efficient XML-to-SQL Query Translation: Where to Add the Intelligence? 144
Rajasekar Krishnamurthy (IBM Almaden), Raghav Kaushik (Microsoft Research), Jeffrey Naughton (Univ. of Wisconsin-Madison)

Taming XPath Queries by Minimizing Wildcard Steps 156
Chee-Yong Chan (National Univ. of Singapore), Wenfei Fan (Univ. of Edinburgh & Bell Laboratories), Yiming Zeng (National Univ. of Singapore)

The NEXT Framework for Logical XQuery Optimization 168
Alin Deutsch, Yannis Papakonstantinou, Yu Xu (Univ. of California, San Diego)

Research Session 5: Stream Mining

Detecting Change in Data Streams 180
Daniel Kifer, Shai Ben-David, Johannes Gehrke (Cornell Univ.)

Stochastic Consistency, and Scalable Pull-Based Caching for Erratic Data Stream Sources 192
Shanzhong Zhu, Chinya V. Ravishankar (Univ.of California, Riverside)

False Positive or False Negative: Mining Frequent Itemsets from High Speed Transactional Data Streams 204
Jeffrey Xu Yu (The Chinese Univ. of Hong Kong), Zhihong Chong (Fudan Univ.), Hongjun Lu (The Hong Kong Univ. of Science and Technology), Aoying Zhou (Fudan Univ.)

Research Session 6: XML (II)

Indexing Temporal XML Documents 216
Alberto Mendelzon, Flavio Rizzolo (Univ. of Toronto), Alejandro Vaisman (Univ. of Buenos Aires)

Schema-based Scheduling of Event Processors and Buffer Minimization for Queries on Structured Data Streams 228
Christoph Koch, Stefanie Scherzinger (Technische Univ. Wien), Nicole Schweikardt (Humboldt Univ. Berlin), Bernhard Stegmaier (Technische Univ. München)

Bloom Histogram: Path Selectivity Estimation for XML Data with Updates 240
Wei Wang (Univ. of NSW), Haifeng Jiang, Hongjun Lu (Hong Kong Univ. of Science and Technology), Jeffrey Xu Yu (The Chinese Univ. of Hong Kong)

Research Session 7: XML and Relations

XQuery on SQL Hosts 252
Torsten Grust, Sherif Sakr, Jens Teubner (Univ. of Konstanz)

ROX: Relational Over XML 264
Alan Halverson (Univ. of Wisconsin-Madison), Vanja Josifovski, Guy Lohman, Hamid Pirahesh (IBM Almaden), Mathias Mörschel

From XML View Updates to Relational View Updates: Old Solutions to a New Problem 276
Vanessa Braganholo (Universidade Federal do Rio Grande do Sul), Susan Davidson (Univ. of Pennsylvania and INRIA-FUTURS), Carlos Heuser (Universidade Federal do Rio Grande do Sul)

Research Session 8: Stream Mining (II)

XWAVE: Optimal and Approximate Extended Wavelets for Streaming Data 288
Sudipto Guha (Univ. of Pennsylvania), Chulyun Kim, Kyuseok Shim (Seoul National Univ.)

REHIST: Relative Error Histogram Construction Algorithms 300
Sudipto Guha (Univ. of Pennsylvania), Kyuseok Shim, Jungchul Woo (Seoul National Univ.)

Distributed Set-Expression Cardinality Estimation 312
Abhinandan Das (Cornell Univ.), Sumit Ganguly (IIT Kanpur), Minos Garofalakis, Rajeev Rastogi (Bell Labs, Lucent Technologies)

Research Session 9: Stream Query Processing

Memory-Limited Execution of Windowed Stream Joins 324
Utkarsh Srivastava, Jennifer Widom (Stanford Univ.)

Resource Sharing in Continuous Sliding-Window Aggregates 336
Arvind Arasu, Jennifer Widom (Stanford Univ.)

Remembrance of Streams Past: Overload-Sensitive Management of Archived Streams 348
Sirish Chandrasekaran, Michael J. Franklin (UC Berkeley)

WIC: A General-Purpose Algorithm for Monitoring Web Information Sources 360
Sandeep Pandey, Kedar Dhamdhere, Christopher Olston (Carnegie Mellon Univ.)

Research Session 10: Managing Web Information Sources

Similarity Search for Web Services 372
Xin Dong, Alon Halevy, Jayant Madhavan, Ema Nemes, Jun Zhang (Univ. of Washington)

AWESOME: A Data Warehouse-based System for Adaptive Website Recommendations 384
Andreas Thor, Erhard Rahm (Univ. of Leipzig)

Accurate and Efficient Crawling for Relevant Websites 396
Martin Ester (Simon Fraser Univ.), Hans-Peter Kriegel, Matthias Schubert (Univ. of Munich)

Instance-based Schema Matching for Web Databases by Domain-specific Query Probing 408
Jiying Wang (Hong Kong Univ. of Science and Technology), Ji-Rong Wen (Microsoft Research Asia), Fred Lochovsky (Hong Kong Univ. of Science and Technology), Wei-Ying Ma (Microsoft Research Asia)

Research Session 11: Distributed Search and Query Processing

Computing PageRank in a Distributed Internet Search System 420
Yuan Wang, David DeWitt (Univ. of Wisconsin-Madison)

Enhancing P2P File-Sharing with an Internet-Scale Query Processor 432
Boon Thau Loo (UC Berkeley), Joseph M. Hellerstein (UC Berkeley and Intel Research Berkeley), Ryan Huebsch (UC Berkeley), Scott Shenker (UC Berkeley and International Computer Science Institute), Ion Stoica (UC Berkeley)

Online Balancing of Range-Partitioned Data with Applications to Peer-to-Peer Systems 444
Prasanna Ganesan, Mayank Bawa, Hector Garcia-Molina (Stanford Univ.)

Network-Aware Query Processing for Distributed Stream-based Applications 456
Yanif Ahmad, Uğur Çetintemel (Brown Univ.)

Data Sharing Through Query Translation in Autonomous Sources 468
Anastasios Kementsietsidis, Marcelo Arenas (Univ. of Toronto)

Research Session 12: Stream Data Management Systems

Linear Road: A Stream Data Management Benchmark 480
Arvind Arasu (Stanford Univ.), Mitch Cherniack, Eduardo Galvez (Brandeis Univ.), David Maier (Oregon Health & Science Univ.), Anurag Maskey, Esther Ryvkina (Brandeis Univ.), Michael Stonebraker, Richard Tibbetts (MIT)

Query Languages and Data Models for Database Sequences and Data Streams 492
Yan-Nei Law (UCLA), Haixun Wang (IBM T. J. Watson Res. Ctr.), Carlo Zaniolo (UCLA)

Research Session 13: Auditing

Tamper Detection in Audit Logs 504
Richard Snodgrass, Shilong Stanley Yao, Christian Collberg (Univ. of Arizona)

Auditing Compliance with a Hippocratic Database 516
Rakesh Agrawal, Roberto Bayardo, Christos Faloutsos, Jerry Kiernan, Ralf Rantzau, Ramakrishnan Srikant (IBM Almaden)

Research Session 14: Data Warehousing

High-Dimensional OLAP: A Minimal Cubing Approach 528
Xiaolei Li, Jiawei Han, Hector Gonzalez (Univ. of Illinois at Urbana-Champaign)

The Polynomial Complexity of Fully Materialized Coalesced Cubes 540
Yannis Sismanis, Nick Roussopoulos (Univ. of Maryland at College Park)

Research Session 15: Link Analysis

Relational Link-based Ranking 552
Floris Geerts (Univ. of Edinburgh), Heikki Mannila, Evimaria Terzi (Univ. of Helsinki)

ObjectRank: Authority-Based Keyword Search in Databases 564
Andrey Balmin (IBM Almaden), Vagelis Hristidis (Florida International Univ.), Yannis Papakonstantinou (UC San Diego)

Combating Web Spam with TrustRank 576
Zoltan Gyöngyi, Hector Garcia-Molina (Stanford Univ.), Jan Pedersen (Yahoo! Inc.)

Research Session 16: Sensors, Grid, Pub/Sub

Model-Driven Data Acquisition in Sensor Networks 588
Amol Deshpande (UC Berkeley), Carlos Guestrin (Intel Research Berkeley), Samuel Madden (MIT and Intel Research Berkeley), Joseph Hellerstein (UC Berkeley and Intel Research Berkeley), Wei Hong (Intel Research Berkeley)

GridDB: A Data-Centric Overlay for Scientific Grids 600
David Liu, Michael J. Franklin (UC Berkeley)

Towards an Internet-Scale XML Dissemination Service 612
Yanlei Diao, Shariq Rizvi, Michael J. Franklin (UC Berkeley)

Research Session 17: Top-K Ranking

Efficiency-Quality Tradeoffs for Vector Score Aggregation 624
Pavan Kumar C. Singitham, Mahathi Mahabhashyam (Stanford Univ.), Prabhakar Raghavan (Verity Inc.)

Merging the Results of Approximate Match Operations 636
Sudipto Guha (Univ. of Pennsylvania), Nick Koudas, Amit Marathe, Divesh Srivastava (AT&T Labs-Research)

Top-k Query Evaluation with Probabilistic Guarantees 648
Martin Theobald, Gerhard Weikum, Ralf Schenkel (Max-Planck Institute of Computer Science)

Research Session 18: DBMS Architecture and Performance

STEPS Towards Cache-resident Transaction Processing 660
Stavros Harizopoulos, Anastassia Ailamaki (Carnegie Mellon Univ.)

Write-Optimized B-Trees 672
Goetz Graefe (Microsoft Corp.)

Cache-Conscious Radix-Decluster Projections 684
Stefan Manegold, Peter Boncz, Niels Nes, Martin Kersten (CWI)

Clotho: Decoupling Memory Page Layout from Storage Organization 696
Minglong Shao, Jiri Schindler, Steven Schlosser, Anastassia Ailamaki, Gregory R. Ganger (Carnegie Mellon Univ.)

Research Session 19: Privacy

Vision Paper: Enabling Privacy for the Paranoids 708
Gagan Aggarwal, Mayank Bawa, Prasanna Ganesan, Hector Garcia-Molina, Krishnaram Kenthapadi, Nina Mishra, Rajeev Motwani, Utkarsh Srivastava, Dilys Thomas, Jennifer Widom, Ying Xu (Stanford Univ.)

A Privacy-Preserving Index for Range Queries 720
Bijit Hore, Sharad Mehrotra, Gene Tsudik (Univ. of California, Irvine)

Resilient Rights Protection for Sensor Streams 732
Radu Sion, Mikhail Atallah, Sunil Prabhakar (Purdue Univ.)

Research Session 20: Nearest Neighbor Search

Reverse kNN Search in Arbitrary Dimensionality 744
Yufei Tao (City Univ. of Hong Kong), Dimitris Papadias, Xiang Lian (Hong Kong Univ. of Science and Technology)

GORDER: An Efficient Method for KNN Join Processing 756
Chenyi Xia (National Univ. of Singapore), Hongjun Lu (Hong Kong Univ. of Science and Technology), Beng Chin Ooi, Jing Hu (National Univ. of Singapore)

Query and Update Efficient B+-Tree Based Indexing of Moving Objects 768
Christian S. Jensen (Aalborg Univ.), Dan Lin, Beng Chin Ooi (National Univ. of Singapore)

Research Session 21: Similarity Search and Applications

Indexing Large Human-Motion Databases 780
Eamonn Keogh, Themistoklis Palpanas, Victor B. Zordan, Dimitrios Gunopulos (Univ. of California, Riverside), Marc Cardle (Univ. of Cambridge)

On The Marriage of Lp-norms and Edit Distance 792
Lei Chen (Univ. of Waterloo), Raymond Ng (Univ. of British Columbia)

Approximate NN queries on Streams with Guaranteed Error/performance Bounds 804
Nick Koudas (AT&T Labs-Research), Beng Chin Ooi, Kian-Lee Tan, Rui Zhang (National Univ. of Singapore)

Object Fusion in Geographic Information Systems 816
Catriel Beeri, Yaron Kanza, Eliyahu Safra, Yehoshua Sagiv (The Hebrew Univ.)

Maintenance of Spatial Semijoin Queries on Moving Points 828
Glenn Iwerks, Hanan Samet (Univ. of Maryland at College Park), Kenneth Smith (MITRE Corp.)

Voronoi-Based K Nearest Neighbor Search for Spatial Network Databases 840
Mohammad R. Kolahdouzan, Cyrus Shahabi (Univ. of Southern California)

A Framework for Projected Clustering of High Dimensional Data Streams 852
Charu Aggarwal (T.J. Watson Res. Ctr.), Jiawei Han, Jianyong Wang (Univ. of Illinois at Urbana-Champaign), Philip Yu (T.J. Watson Res. Ctr.)

Research Session 22: Query Processing

Efficient Query Evaluation on Probabilistic Databases 864
Nilesh Dalvi, Dan Suciu (Univ. of Washington)

Efficient Indexing Methods for Probabilistic Threshold Queries over Uncertain Data 876
Reynold Cheng, Yuni Xia, Sunil Prabhakar, Rahul Shah, Jeffrey S. Vitter (Purdue Univ.)

Probabilistic Ranking of Database Query Results 888
Surajit Chaudhuri, Gautam Das (Microsoft Research), Vagelis Hristidis (Florida International Univ.), Gerhard Weikum (MPI Informatik)

Research Session 23: Novel Models

An Annotation Management System for Relational Databases 900
Deepavali Bhagwat, Laura Chiticariu, Wang-Chiew Tan, Gaurav Vijayvargiya (Univ. of California, Santa Cruz)

Symmetric Relations and Cardinality-Bounded Multisets in Database Systems 912
Kenneth Ross, Julia Stoyanovich (Columbia Univ.)

Algebraic Manipulation of Scientific Datasets 924
Bill Howe, David Maier (Oregon Health & Science Univ.)

Research Session 24: Query Processing and Optimization

Multi-objective Query Processing for Database Systems 936
Wolf-Tilo Balke (Univ. of California, Berkeley), Ulrich Güntzer (Univ. of Tübingen)

Lifting the Burden of History from Adaptive Query Processing 948
Amol Deshpande (Univ. of California, Berkeley), Joseph M. Hellerstein (Univ. of California, Berkeley and Intel Research, Berkeley)

A Combined Framework for Grouping and Order Optimization 960
Thomas Neumann, Guido Moerkotte (Univ. of Mannheim)

The Case for Precision Sharing 972
Sailesh Krishnamurthy, Michael J. Franklin (UC Berkeley), Joseph M. Hellerstein (UC Berkeley and Intel Research Berkeley), Garrett Jacobson (UC Berkeley)

INDUSTRIAL AND APPLICATIONS SESSIONS

Industrial Session 1: Novel SQL Extensions

Returning Modified Rows - SELECT Statements with Side Effects 987
Andreas Behm, Serge Rielau, Richard Swagerman (IBM Toronto Lab)

PIVOT and UNPIVOT: Optimization and Execution Strategies in an RDBMS 998
Conor Cunningham, Cesar Galindo-Legaria, Goetz Graefe (Microsoft Corp.)

A Multi-Purpose Implementation of Mandatory Access Control in Relational Database Management Systems 1010
Walid Rjaibi, Paul Bird (IBM Toronto Lab)

Industrial Session 2: New DBMS Architectures and Performance

Hardware Acceleration in Commercial Databases: A Case Study of Spatial Operations 1021

Nagender Bandi (Univ. of California, Santa Barbara), Chengyu Sun (California State Univ., Los Angeles), Divyakant Agrawal, Amr El Abbadi (Univ. of California, Santa Barbara)

P*TIME: Highly Scalable OLTP DBMS for Managing Update-Intensive Stream Workload 1033

Sang K. Cha (Transact In Memory, Inc.), Changbin Song (Seoul National Univ.)

Generating Thousand Benchmark Queries in Seconds 1045

Meikel Poess (Oracle Corp.), John M. Stephens, Jr. (Gradient Systems)

Industrial Session 3: Semantic Query Approaches

Supporting Ontology-Based Semantic Matching in RDBMS 1054

Souripriya Das, Eugene Chong, George Eadon, Jagannathan Srinivasan (Oracle Corp.)

BioPatentMiner: An Information Retrieval System for BioMedical Patents 1066

Sougata Mukherjea, Bhuvan Bamba (IBM India Research Lab)

Flexible String Matching Against Large Databases in Practice 1078

Nick Koudas, Amit Marathe, Divesh Srivastava (AT&T Labs–Research)

Industrial Session 4: Automatic Tuning in Commercial DBMSs

DB2 Design Advisor: Integrated Automatic Physical Database Design 1087

Daniel Zilio (IBM Toronto Lab), Jun Rao (IBM Almaden), Sam Lightstone (IBM Toronto Lab), Guy Lohman (IBM Almaden), Adam Storm, Christian Garcia-Arellano (IBM Toronto Lab), Scott Fadden (IBM Portland)

Automatic SQL Tuning in Oracle 10g 1098

Benoit Dageville, Dinesh Das, Karl Dias, Khaled Yagoub, Mohamed Zait, Mohamed Ziauddin (Oracle Corp.)

Database Tuning Advisor for Microsoft SQL Server 2005 1110

Sanjay Agrawal, Surajit Chaudhuri, Lubor Kollar, Arun Marathe, Vivek Narasayya, Manoj Syamala (Microsoft Corp.)

Industrial Session 5: XML Implementations, Automatic Physical Design and Indexing

Query Rewrite for XML in Oracle XML DB 1122

Muralidhar Krishnaprasad, Zhen Hua Liu, Anand Manikutty, James W. Warner, Vikas Arora, Susan Kotsovolos (Oracle Corp.)

Indexing XML Data Stored in a Relational Database 1134

Shankar Pal, Istvan Cseri, Oliver Seeliger, Gideon Schaller, Leo Giakoumakis, Vasili Zolotov (Microsoft Corp.)

Automated Statistics Collection in DB2 UDB 1146
Ashraf Aboulnaga, Peter Haas (IBM Almaden), Mokhtar Kandil, Sam Lightstone (IBM Toronto Lab), Guy Lohman, Volker Markl (IBM Almaden), Ivan Popivanov (IBM Toronto Lab), Vijayshankar Raman (IBM Almaden)

High Performance Index Build Algorithms for Intranet Search Engines 1158
Marcus Fontoura, Eugene Shekita, Jason Zien, Sridhar Rajagopalan, Andreas Neumann (IBM Almaden)

Automated Design of Multi-Dimensional Clustering Tables in Relational Databases 1170
Sam Lightstone (IBM Toronto Laboratory), Bishwaranjan Bhattacharjee (IBM T.J. Watson Res. Ctr.)

Industrial Session 6: Data Management with RFIDs and Ease of Use

Integrating Automatic Data Acquisition with Business Processes - Experiences with SAP's Auto-ID Infrastructure (invited) 1182
Christof Bornhövd, Tao Lin, Stephan Haller, Joachim Schaper (SAP Research)

Managing RFID Data (invited)..... 1189
Sudarshan S. Chawathe (Univ. of Maryland), Venkat Krishnamurthy, Sridhar Ramachandran (OAT Systems), and Sanjay Sarma (MIT)

Production Database Systems: Making Them Easy is Hard Work (invited) 1196
David Campbell (Microsoft Corp.)

Industrial Session 7: Data Management Challenges in Life Sciences and Email Systems

Managing Data from High-Throughput Genomic Processing: A Case Study (invited) 1198
Toby Bloom, Ted Sharpe (Broad Institute of MIT and Harvard)

Database Challenges in the Integration of Biomedical Data Sets (invited)..... 1202
Rakesh Nagarajan (Washington Univ.), Mushtaq Ahmed, Aditya Phatak (Persistent Systems Pvt. Ltd.)

The Bloomba Personal Content Database (invited)..... 1214
Raymie Stata (Stata Labs and UC Santa Cruz), Patrick Hunt (Stata Labs), Thiruvalluvan M. G. (iSoftTech Ltd.)

Industrial Session 8: Issues in Data Warehousing

Trends in Data Warehousing: A Practitioner's View (invited)..... 1224
William O'Connell (IBM Toronto Lab)

Technology Challenges in a Data Warehouse (invited) 1225
Ramesh Bhashyam (NCR Corporation)

Sybase IQ Multiplex - Designed For Analytics (invited) 1227
Roger MacNicol, Blaine French (Sybase Inc.)

PANEL SESSIONS

Biological Data Management: Research, Practice and Opportunities 1233

Moderator: Thodoros Topaloglou (MDS Proteomics). Panelists: Susan B. Davidson (Univ. of Pennsylvania), H. V. Jagadish (Univ. of Michigan), Victor M. Markowitz (Lawrence Berkeley National Laboratory), and Evan W. Steeg (Consultant), Mike Tyers (Univ. of Toronto).

Where is Business Intelligence taking today's Database Systems? 1237

Moderator: William O'Connell (IBM Toronto Lab). Panelists: Andy Witkowski (Oracle Corp.), Ramesh Bhashyam (NCR – Teradata), Surajit Chauduri (Microsoft Corp.)

TUTORIAL SESSIONS

Database Architecture for New Hardware 1241

Anastassia Ailamaki (Carnegie Mellon Univ.)

Security of Shared Data in Large Systems: State of the Art and Research Directions 1242

Arnon Rosenthal (MITRE Corp.), Marianne Winslett (Univ. of Illinois at Urbana-Champaign)

Self-Managing Technology in Database Management Systems 1243

Surajit Chaudhuri (Microsoft Research), Benoit Dageville (Oracle Corp.), Guy Lohman (IBM Almaden)

Architectures and Algorithms for Internet-Scale (P2P) Data Management 1244

Joseph M. Hellerstein (UC Berkeley and Intel Research Berkeley)

The Continued Saga of DB-IR Integration 1245

Ricardo Baeza-Yates (Univ. of Chile) and Mariano Consens (Univ. of Toronto)

DEMONSTRATIONS

GPX: Interactive Mining of Gene Expression Data 1249

Daxin Jiang, Jian Pei, Aidong Zhang (State Univ. of New York at Buffalo)

Computing Frequent Itemsets Inside Oracle 10G 1253

Wei Li, Ari Mozes (Oracle Corp)

StreamMiner: A Classifier Ensemble-based Engine to Mine Concept-drifting Data Streams 1257

Wei Fan (IBM T.J. Watson Res. Ctr.)

Semantic Mining and Analysis of Gene Expression Data 1261

Xin Xu, Gao Cong, Beng Chin Ooi, Kian-Lee Tan, Anthony K.H.Tung (National Univ. of Singapore)

HOS-Miner: A System for Detecting Outlying Subspaces of High-dimensional Data 1265

Ji Zhang, Meng Lou (Univ. of Toronto), Tok Wang Ling (National Univ. of Singapore), Hai Wang (Univ. of Toronto)

VizTree: a Tool for Visually Mining and Monitoring Massive Time Series Databases 1269
Jessica Lin, Eamonn Keogh, Stefano Lonardi (Univ. of California, Riverside), Jeffrey P. Lankford, Daonna M. Nystrom (The Aerospace Corporation)

An Electronic Patient Record “on Steroids”: Distributed, Peer-to-Peer, Secure and Privacy-conscious 1273
Serge Abiteboul (INRIA), Bogdan Alexe (Bell Labs, Lucent Technologies), Omar Benjelloun, Bogdan Cautis (INRIA), Irini Fundulaki (Bell Labs, Lucent Technologies), Tova Milo (INRIA, France, and Tel-Aviv Univ.), Arnaud Sahuguet (Tel-Aviv Univ.)

Queries and Updates in the coDB Peer to Peer Database System 1277
Enrico Franconi (Free Univ. of Bozen-Bolzano), Gabriel Kuper (Univ. of Trento), Andrei Lopatenko (Free Univ. of Bozen-Bolzano and Univ. of Manchester), Ilya Zaihrayeu (Univ. of Trento)

A-ToPSS: A Publish/Subscribe System Supporting Imperfect Information Processing 1281
Haifeng Liu, Hans-Arno Jacobsen (Univ. of Toronto)

Efficient Constraint Processing for Highly Personalized Location Based Services 1285
Zhengdao Xu, Hans-Arno Jacobsen (Univ. of Toronto)

LH*RS: A Highly Available Distributed Data Storage 1289
Witold Litwin, Rim Moussa (Université Paris Dauphine), Thomas Schwarz (Santa Clara Univ.)

Semantic Query Optimization in an Automata-Algebra Combined XQuery Engine over XML Streams 1293
Hong Su, Elke A. Rundensteiner, Murali Mani (Worcester Polytechnic Institute)

ShreX: Managing XML Documents in Relational Databases 1297
Fang Du (Oregon Health & Science Univ), Sihem Amer-Yahia (AT&T Labs-Research), Juliana Freire (Oregon Health & Science Univ)

A Uniform System for Publishing and Maintaining XML Data 1301
Byron Choi (Univ. of Pennsylvania), Wenfei Fan (Univ. of Edinburgh and Bell Laboratories), Xibei Jia (Univ. of Edinburgh), Arek Kasprzyk (European Bioinformatics Institute)

An Injection of Tree Awareness: Adding Staircase Join to PostgreSQL 1305
Sabine Mayer, Torsten Grust (Univ. of Konstanz), Maurice van Keulen (Univ. of Twente), Jens Teubner (Univ. of Konstanz)

FluXQuery: An Optimizing XQuery Processor for Streaming XML Data 1309
Christoph Koch, Stephanie Scherzinger (Technische Univ. Wien), Nicole Schweikardt (Humboldt Univ. Berlin), Bernhard Stegmaier (Technische Univ. München)

COMPASS: A Concept-based Web Search Engine for HTML, XML, and Deep Web Data 1313
Jens Graupmann, Michael Biwer, Christian Zimmer, Patrick Zimmer, Matthias Bender, Martin Theobald, Gerhard Weikum (Max-Planck Institute for Computer Science)

Discovering and Ranking Semantic Associations over a Large RDF Metabase 1317
Chris Halaschek, Boanerges Aleman-Meza, I. Budak Arpinar, Amit P. Sheth (Univ. of Georgia)

An Automatic Data Grabber for Large Web Sites 1321
Valter Crescenzi (Univ. of Roma), Giansalvatore Mecca (Univ. of Potenza), Paolo Merialdo , Paolo Missier (Univ. of Roma)

WS-CatalogNet: An Infrastructure for Creating, Peering, and Querying e-Catalog Communities 1325
Karim Baina, Boualem Benatallah, Hye-young Paik (Univ. of New South Wales), Farouk Toumani, Christophe Rey (ISIMA), Agnieszka Rutkowska, Bryan Harianto (Univ. of New South Wales)

Trust-Serv: A Lightweight Trust Negotiation Service 1329
Halvard Skogsrud, Boualem Benatallah (Univ. of New South Wales), Fabio Casati (Hewlett-Packard Labs), Manh Q. Dinh (Univ. of New South Wales)

Green Query Optimization: Taming Query Optimization Overheads through Plan Recycling 1333
Parag Sarda, Jayant R. Haritsa (Indian Institute of Science)

Progressive Optimization in Action 1337
Vijayshankar Raman, Volker Markl, David Simmen, Guy Lohman, Hamid Pirahesh (IBM Almaden)

CORDS: Automatic Generation of Correlation Statistics in DB2 1341
Ihab F. Ilyas (Univ. of Waterloo), Volker Markl, Peter J. Haas, Paul G. Brown, Ashraf Aboulnaga (IBM Almaden)

CHICAGO: A Test and Evaluation Environment for Coarse-Grained Optimization 1345
Tobias Kraft, Holger Schwarz (Univ. of Stuttgart)

SVT: Schema Validation Tool for Microsoft SQL-Server 1349
Ernest Teniente, Carles Farré, Toni Urpí, Carlos Beltrán, David Gañán (Univ. of Catalunya)

CAPE: Continuous Query Engine with Heterogeneous-Grained Adaptivity 1353
Elke A. Rundensteiner, Luping Ding, Timothy Sutherland, Yali Zhu, Brad Pielech, Nishant Mehta (Worcester Institute of Technology)

HiFi: A Unified Architecture for High Fan-in Systems 1357
Owen Cooper, Anil Edakkunni, Michael J. Franklin (UC Berkeley), Wei Hong (Intel Research Berkeley), Shawn R. Jeffery, Sailesh Krishnamurthy, Fredrick Reiss, Shariq Rizvi, Eugene Wu (UC Berkeley)

An Integration Framework for Sensor Networks and Data Stream Management Systems 1361
Daniel J. Abadi, Wolfgang Lindner, Samuel Madden (MIT), Jörg Schuler (Tufts Univ.)

QStream: Deterministic Querying of Data Streams 1365
Sven Schmidt, Henrike Berthold, Wolfgang Lehner (Dresden Univ. of Technology)

AIDA: an Adaptive Immersive Data Analyzer 1369
Mehdi Sharifzadeh, Cyrus Shahabi, Bahareh Navai, Farid Parvini, Albert A. Rizzo (Univ. of Southern California)

BilVideo Video Database Management System 1373
Özgür Ulusoy, Uğur Güdükbay, Mehmet Emin Dönderler, Ediz Şaykol, Cemil Alper (Bilkent Univ.)

PLACE: A Query Processor for Handling Real-time Spatio-temporal Data Streams 1377
Mohamed F. Mokbel, Xiaopeng Xiong, Walid G. Aref, Susanne E. Hambrusch, Sunil Prabhakar, Moustafa A. Hammad (Purdue Univ.)