13th International Conference on Extending Database Technology 2010

(EDBT 2010)

ACM International Conference Proceedings Series Volume 426

Lausanne, Switzerland
22-26 March 2010

Editors:

Ioana Manolescu  Alain Leger
Stefano Spaccapietra  Felix Naumann
Jens Teubner  AnastasiaAilamaki
Masaru Kitsuregawa  Fatma Ozcan

Table of Contents

Foreword ................................................................. ix
Program Committee Members ....................................... x–xi

Invited Papers

Provenance for Database Transformations
  Val Tannen ............................................................. 1
Scalable Ontology-Based Information Systems
  Ian Horrocks .......................................................... 2

Research Sessions

XML and Semi-Structured Data
Feedback-driven Result Ranking and Query Refinement for Exploring Semi-structured Data Collections
  Huiping Cao, Yan Qi, K. Selcuk Candan, and Maria Luisa Sapino .................. 3–14
Beyond Pages: Supporting Efficient, Scalable Entity Search with Dual-Inversion Index
  Tao Cheng and Kevin Chang .......................................... 15–26
Processing XPath queries with forward and downward axes over XML Streams
  Makoto Onizuka ....................................................... 27–38

OLAP and Decision Support

Correlation Aware Synchronization for Near Real Time Decision Support Systems
  Ying Yan and Wen-Syan Li ............................................ 39–50
Turbo-Charging Hidden Database Samplers with Overflowing Queries and Skew Reduction
  Arjun Dasgupta, Nan Zhang, and Gautam Das .................................. 51–62
Region-based Online Promotion Analysis
  Tianyi Wu, Yizhou Sun, Cuiping Li, and Jiawei Han .............................. 63–74

Distributed Databases

The Data Cyclotron Query Processing Scheme
  Romulo Goncalves and Martin Kersten .................................... 75–86
Gossiping Personalized Queries
  Xiao Bai, Marin Bertier, Rachid Guerraoui, Anne-Marie Kermarrec, and Vincent Leroy .................................................................................. 87–98
Optimizing Joins in a Map-Reduce Environment
  Foto Afrati and Jeffrey Ullman .............................................. 99–110
Security and Privacy 1

Private Record Matching Using Differential Privacy
Ali Inan, Murat Kantarcioglu, Gabriel Ghinita, and Elisa Bertino ......................... 111–122

The Hardness and Approximation Algorithms for L-Diversity
Xiaokui Xiao, Ke Yi, and Yufei Tao .................................................. 123–134

K-Symmetry Model for Identity Anonymization in Social Networks
Wentao Wu, Yanghua Xiao, Wei Wang, Zhenying He, and Zhihui Wang ............ 135–146

XPath and XQuery Optimization

Let SQL Drive the XQuery Workhorse
Torsten Grust, Manuel Mayr, and Jan Rittinger ........................................... 147–158

Efficient Physical Operators for cost-based XPath Execution
Harris Georgiadis, Minas Charalambides, and Vasilis Vassalos ......................... 159–170

Statistics-based Parallelization of XPath Queries in Shared Memory Systems
Rajesh Bordawekar, Lipyeow Lim, Anastasios Kementsietsidis, and Bryant Wei-Lun Kok .................................................. 171–182

Query Processing and Optimization 1

Adaptive Join Processing in Pipelined Plans
Kwanchai Eurviriyankul, Norman W. Paton, Alvaro A. A. Fernandes, and Steven Lynden .................................................. 183–194

BSkyTree: Scalable Skyline Computation Using A Balanced Pivot Selection
Jongwuk Lee and Seung-won Hwang .................................................. 195–206

Stream Schema: Providing and Exploiting Static Metadata for Data Stream Processing
Peter M. Fischer, Kyumars Sheykh Esmaili, and Renée J. Miller .................... 207–218

Optimization for Modern Hardware

Warm Cache Costing – A Feedback Optimization Technique for Buffer Pool Aware Costing
Ramanujam Halasipuram S and Edwin Seputis ........................................... 219–227

Position List Word Aligned Hybrid: Optimizing Space and Performance for Compressed Bitmaps
François Deliège and Torben Bach Pedersen ........................................... 228–239

Suffix Tree Construction Algorithms on Modern Hardware
Dimistris Tsirogiannis and Nick Koudas ........................................... 240–251

Scientific Databases and OLAP

Reducing Metadata Complexity for Faster Table Summarization
K. Selcuk Candan, Mario Cataldi, and Maria Luisa Sapino ........................................... 252–263

Splash: Ad-Hoc Querying of Data and Statistical Models
Lujun Fang and Kristen LeFevre .................................................. 264–275

Anchoring Millions of Distinct Reads on the Human Genome within Seconds
Tien Huynh, Michail Vlachos, and Isidore Rigoutsos ........................................... 276–286
Data Provenance
Techniques for Efficiently Querying Scientific Workflow Provenance Graphs
  Manish Anand, Shawn Bowers, and Bertram Ludaescher ...................... 287-298
Fine-grained and efficient lineage querying of collection-based workflow provenance
  Paolo Missier, Norman W. Paton, and Khalid Belhajjame .................... 299-310
Lost Source Provenance
  Jing Zhang and H. V. Jagadish ............................................. 311-322

Probabilistic and Spatial Databases
Bridging the Gap Between Intensional and Extensional Query Evaluation in Probabilistic Databases
  Abhay Jha, Dan Olteanu, and Dan Suciu .................................... 323-334
Probabilistic Path Queries in Road Networks: Traffic Uncertainty Aware Path Selection
  Ming Hua and Jian Pei .......................................................... 335-346
Probabilistic Threshold k Nearest Neighbor Queries over Moving Objects in Symbolic Indoor Space
  Bin Yang, Hua Lu, and Christian S. Jensen ............................... 347-358

Query Processing and Optimization 2
A Simple (yet Powerful) Algebra for Pervasive Environments
  Yann Gripay, Frédérique Laforest, and Jean-Marc Petit .................... 359-370
Self-selecting, self-tuning, incrementally optimized indexes
  Goetz Graefe and Harumi Kuno ............................................. 371-381
Minimizing Database Repros using Language Grammars
  Nicolas Bruno ................................................................. 382-393

Spatial Databases
Efficient and Scalable Multi-Geography Route Planning
  Vidhya Balasubramanian, Dmitri Kalashnikov, Sharad Mehrotra, and Nalini Venkatasubramanian ................................. 394-405
Querying Trajectories Using Flexible Patterns
Querying Spatial Patterns
  Vishwakarma Singh, Arnab Bhattacharya, and Ambuj K. Singh ............. 418-429

Technologies for the Web
Indexing Relations on the Web
  Sergio Mergen, Juliana Freire, and Carlos Heuser ......................... 430-440
An Execution Environment for C-SPARQL Queries
  Davide Francesco Barbieri, Daniele Braga, Stefano Ceri, and Michael Grossniklaus ...................................... 441-452
Rewrite Techniques for Performance Optimization of Schema Matching Processes
  Eric Peukert, Henrik Berthold, and Erhard Rahm .......................... 453-464
Ranking and Nearest Neighbor

Fast Computation of SimRank for Static and Dynamic Information Networks
Cuiping Li, Jiawei Han, Guoming He, Xin Jin, Yizhou Sun, Yintao Yu, and Tianyi Wu ................................................................. 465–476

Probabilistic Ranking over Relations
Lijun Chang, Jeffrey Xu Yu, Lu Qin, and Xuemin Lin ......................... 477–488

Privacy Preserving Group Nearest Neighbor Queries
Tanzima Hashem, Lars Kulik, and Rui Zhang ................................. 489–500

Data Cleaning and Curation

Subsumption and Complementation as Data Fusion Operators
Jens Bleiholder, Sascha Szott, Melanie Herschel, Frank Kaufer, and Felix Naumann ................................................................. 501–512

HARRA: Fast Iterative Hashed Record Linkage for Large-Scale Data Collections
Hung-sik Kim and Dongwon Lee .................................................... 513–524

Finding Misplaced Items in Retail by Clustering RFID Data
Leonardo Weiss Ferreira Chaves, Erik Buchmann, and Klemens Böhm ....... 525–536

XML Keyword Search

Keyword Search for Data-Centric XML Collections with Long Text Fields
Arash Termehchy and Marianne Winslett ........................................ 537–548

Fast ELCA Computation for Keyword Queries on XML Data
Rui Zhou, Chengfei Liu, and Jianxin Li ........................................... 549–560

Suggestion of Promising Result Types for XML Keyword Search
Jianxin Li, Chengfei Liu, Rui Zhou, and Wei Wang .......................... 561–572

Personalization and Preferences

Feedback-Based Annotation, Selection and Refinement of Schema Mappings for Dataspaces
Khalid Belghaizeme, Norman W. Paton, Suzanne Embury, Alvaro A. A. Fernandes, and Cornelia Hedeler ................................. 573–584

PerK: Personalized Keyword Search in Relational Databases through Preferences
Kostas Stefanidis, Marina Drosou, and Evaggelia Pitoura .................. 585–596

Efficient Computation of Trade-Off Skylines
Christoph Lofi, Ulrich Güntzer, and Wolf-Tilo Balke ........................ 597–608

Security and Privacy 2

How to Authenticate Graphs Without Leaking
Ashish Kundu and Elisa Bertino .................................................... 609–620

Trustworthy Vacuuming and Litigation Holds in Long-term High-integrity Records Retention
Ragib Hasan and Marianne Winslett ............................................. 621–632

Algorithm-safe Privacy-Preserving Data Publishing
Xin Jin, Nan Zhang, and Gautam Das ............................................. 633–644
Industrial Sessions

Transactions and Distribution

BronzeGate: Real-time Transactional Data Obfuscation for GoldenGate

Shenoda Guirguis and Alok Pareek ........................................ 645–650

Logging Last Resource Optimization for Distributed Transactions in Oracle WebLogic Server

Tom Barnes, Adam Messinger, Paul Parkinson, Amit Ganesh, German Shegalov, Saraswathy Narayanan, and Srinivas Kareenhalli ........................................ 651–656

DEDUCE: At the Intersection of MapReduce and Stream Processing

Vibhore Kumar, Henrique Andrade, Bugra Gedik, and Kun-Lung Wu ............ 657–662

New Applications

An Experimental Study of Time-Constrained Aggregate Queries

Ying Hu, Wen-Chi Hou, Seema Sundara, and Jagannathan Srinivasan ............ 663–668

Xbase: Cloud-enabled Information Appliance for Healthcare

Wen-Syan Li, Jianfeng Yan, Ying Yan, and Jin Zhang ......................... 669–674

Data Warehousing and Analytics

Aggregation of asynchronous electric power consumption time series knowing the integral

Raja Chiky, Laurent Decreusefond, and Georges Hebrail ......................... 675–680

A Plan for OLAP

Bernhard Jaecksch, Franz Faerber, and Wolfgang Lehner ....................... 681–686

Augmenting OLAP Exploration with Dynamic Advanced Analytics

Benjamin Leonhardi, Bernhard Mitschang, Ruben Pulido de los Reyes, Christoph Sieb, and Michael Wurst .................. 687–692

Demonstrations

Demonstrations

Advanced Knowledge Discovery on Movement Data with the GeoPKDD system

Mirco Nanni, Roberto Trasarti, Chiara Renso, Fosca Giannotti, and Dino Pedreschi ........................................ 693–696

Timely YAGO: Harvesting, Querying, and Visualizing Temporal Knowledge from Wikipedia

Yafang Wang, mingjie zhu, Lizhen Qu, Marc Spaniol, and Gerhard Weikum .... 697–700

PARINDA: An Interactive Physical Designer for PostgreSQL

Cristina Maier, Debabrata Dash, Ioannis Alagiannis, Anastasia Ailamaki, and Thomas Heinis ........................................ 701–704

BIAEEditor - Matching Process and Operational Data for a Business Impact Analysis

Sylvia Radeschütz, Florian Niedermann, and Wolfgang Bischoff .................. 705–708

Pattern Detector: Fast Detection of Suspicious Stream Patterns for Immediate Reaction

Ira Assent, Hardy Kremer, Stephan Günemann, and Thomas Seidl .............. 709–712

BP-Ex – A uniform query engine for Business Process Execution traces

Eran Balan, Tova Milo, and Tal Sterenzky ...................................... 713–716
B-Fabric: The Swiss Army Knife for Life Sciences

Can Türker, Fuat Akal, Dieter Joho, Ralph Schlapbach, Christian Panse, Simon Barkow-Oesterreicher, and Hubert Rehrauer ........................................ 717–720

Tutorials

FPGAs: A New Point in the Database Design Space

Rene Mueller and Jens Teubner ........................................ 721–723

Querying the Deep Web

Andrea Cali and Davide Martinenghi ........................................ 724–727

Author Index