12th International Conference on Extending Database Technology 2009

Advances in Database Technology

(EDBT 2009)

ACM International Conference Proceedings Series Volume 360

Saint Petersburg, Russia
24 – 26 March 2009

Volume 1 of 2

ISBN: 978-1-61567-503-6
Table of Contents

Foreword
Program Committee Members

Invited Papers

Data Integration Flows for Business Intelligence
  Umeshwar Dayal, Malu Castellanos, Alkis Simitis, and Kevin Wilkinson .......... 1-11

Research Sessions

System Architectures
  Optimized Union of Non-disjoint Distributed Data Sets
    Itay Dar, Tova Milo, and Elad Verbin ................................. 12–23
  Shore-MT: A Scalable Storage Manager for the Multicore Era
    Ryan Johnson, Ippokratis Pandis, Nikos Hardavellas, Anastasia Ailamaki, and
    Babak Falsafi ........................................................... 24–35

Workload-Aware Data Partitioning in Community-Driven Data Grids
  Tobias Scholl, Bernhard K. Bauer, Jessica Müller, Benjamin Gufler, Angelika
  Reiser, and Alfons Kemper ............................................. 36–47

Spatio-Temporal
  Sequenced, Spatio-Temporal Aggregation in Road Networks
    Igor Timko, Michael Boehlen, and Johann Gamper ........................ 48–59
  Processing Probabilistic Spatio-Temporal Range Queries over Moving Objects with
  Uncertainty
    Bruce Chung, Wang-Chien Lee, and Arbee L.P. Chen ........................ 60–71
  Anonymizing Moving Objects: How to Hide a MOB in a Crowd?
    Roman Yarovoy, Francesco Bonchi, Laks V.S. Lakshmanan, and Hui Wang ...... 72–83

Database Summarization
  Type-Based Categorization of Relational Attributes
    Babak Ahmadi, Marios Hadjieleftheriou, Thomas Seidl, Divesh Srivastava, and
    Suresh Venkatasubramanian ............................................. 84–95
  AlphaSum: Size-Constrained Table Summarization using Value Lattices
    K. Selcuk Candan, Huiping Cao, Yan Qi, and Maria Luisa Sapino ............ 96–107
  Answering Aggregate Keyword Queries on Relational Databases Using Minimal
  Group-bys
    Bin Zhou and Jian Pei .................................................. 108–119
Query Processing

Rule-Based Multi-Query Optimization

Mingsheng Hong, Mirek Riedewald, Christoph Koch, Johannes Gehrke, and Alan Demers 120-131

Managing Long-Running Queries

Stefan Krompaß, Harumi Kuno, Janet Wiener, Kevin Wilkinson, Umeshwar Dayal, and Alfons Kemper 132-143

Continuous Visible Nearest Neighbor Queries

Yunjun Gao, Baidhua Zheng, Wang-Chien Lee, and Gencai Chen 144-155

XML, XPath, XQuery

Query Ranking in Probabilistic XML Data

Lijun Chang, Jeffrey Xu Yu, and Lu Qin 156-167

On Rewriting XPath Queries Using Views

Foto Afrati, Rada Chirkova, Manolis Gergatsoulis, Vassia Pavlaki, Benny Kimelfeld, and Yehoshua Sagiv 168-179

Parallelization of XPath Queries using Multi-core Processors: Challenges and Experiences

Lipyeow Lim, Rajesh Bordawekar, and Oded Shmueli 180-191

Graph Techniques

GADDI: Distance Index based Subgraph Matching in Biological Networks

Shijie Zhang, Shirong Li, and Jiong Yang 192-203

A Novel Approach for Efficient Supergraph Query Processing on Graph Databases

Shuo Zhang, Jianzong Li, Hong Gao, and Zhaonian zou 204-215

Flexible Query Answering on Graph-modeled Data

Giorgio Villani, Federico Mandreoli, Riccardo Martoglia, and Wilma Penzo 216-227

Privacy & Security

Privacy-Preserving Data Mashup

Noman Mohammed, Benjamin C. M. Fung, Ke Wang, and Patrick C. K. Hung 228-239

On the Comparison of Microdata Disclosure Control Algorithms

Rinku Dewri, Indrajit Ray, Indrakshi Ray, and Darrell Whitley 240-251

Detecting Privacy Violations in Database Publishing using Disjoint Queries

Millist Vincent, Mukesh Mohania, and Mizuho Iwaihara 252-262

Data Models

On Keys, Foreign Keys and Nullable Attributes in Relational Mapping Systems

Luca Cabibbo 263-274

A runtime approach to model-independent schema and data translation

Paolo Atzeni, Luigi Bellomarini, Francesco Bugiotti, and Giorgio Gianforme 275-286

A methodology for preference-based personalization of contextual data

Antonio Miele, Elisa Quintarelli, and Letizia Tanca 287-298
Stream Processing
Scalable Stream Join Processing with Expensive Predicates: Workload Distribution and Adaptation by Time-Slicing
Song Wang and Elke Rundensteiner ........................................ 299–310

Indexing Density Models for Incremental Learning and Anytime Classification on Data Streams
Thomas Seidl, Ira Assent, Philipp Kränen, Ralph Krieger, and Jennifer Herrmann 311–322

Exploiting the Power of Relational Databases for Efficient Stream Processing
Erietta Liarou, Romulo Gonçalves, and Stratos Idreos ........................................ 323–334

XML, XPath, XQuery
A Sampling Approach for XML Query Selectivity Estimation
Cheng Luo, Zhewei Jiang, Wen-Chi Hou, Feng Yu, and Qiang Zhu .................. 335–344

Recursion in XQuery: Put Your Distributivity Safety Belt On
Loredana Afanasiev, Torsten Grust, Maarten Marx, Jan Rittinger, and Jens Teubner ........................................ 345–356

Expressive, yet Tractable XML Keys
Sven Hartmann and Sebastian Link ........................................ 357–367

Database Summarization
It Takes Variety to Make a World: Diversification in Recommender Systems
Cong Yu, Laks V.S. Lakshmanan, and Sihem Amer-Yahia .......................... 368–378

Supporting Annotations on Relations
Mohamed Eltabakh, Walid G. Aref, Ahmed Elmagarmid, Mourad Ouzzani, and Yasin Silva ........................................ 379–390

DataClouds: Summarizing Keyword Search Results over Structured Data
Georgia Koutrika, Zahra Mohammadi Zadeh, and Hector García-Molina ........... 391–402

Query Processing
Sample Synopses for Approximate Answering of Group-By Queries
Philipp Rösch and Wolfgang Lehner ........................................ 403–414

A Query Processor for Prediction-Based Monitoring of Data Streams
Sergio Ilarri, Ouri Wolfson, Eduardo Mena, Arantza Illarramendi, and Prasad Sistla ........................................ 415–426

Flower-CDN: A Hybrid P2P Overlay for Efficient Query Processing in CDN
Manal El Dick, Esther Pacitti, and Bettina Kemme ........................................ 427–438

Top-K Techniques
Zerber+R: Top-k Retrieval from a Confidential Index
Sergej Zerr, Daniel Olmedilla, Wolfgang Nejdl, and Wolf Siberski .............. 439–449

Efficient Top-K count queries over imprecise duplicates
Sunita Sarawagi, Vinay Deshpande, and Sourabh Kasliwal ...................... 450–461

The C-ND Tree: A Multidimensional Index for Hybrid Continuous and Non-ordered Discrete Data Spaces
Changqing Chen, Sakti Pramanik, Qiang Zhu, Alok Wawve, and Gang Qian ...... 462–471
Graph Techniques

G-Hash: Towards Fast Kernel-based Similarity Search in Large Graph Databases
Xiaohong Wang, Aaron Smaller, Jun Huan, and Gerald Lushington 472–480

On-line Exact Shortest Distance Query Processing
Cheng Jiefeng and Jeffrey Xu Yu 481–492

Efficiently Indexing Shortest Paths by Exploiting Symmetry in Graphs
Yanghua Xiao, Wentao Wu, Jian Pei, Wei Wang, and Zhenying He 493–504

Data Mining

Estimating the Number of Frequent Itemsets in a Large Database
Ruoming Jin, Scott McCallen, Yuri Breitbart, David Fuhry, and Dong Wang 505–516

FOGGER: An Algorithm for Graph Generator Discovery
Zhiping Zeng, Jianyong Wang, Jun Zhang, and Lizhu Zhou 517–528

Neighbor-Based Pattern Detection for Windows Over Streaming Data
Di Yang, Elke Rundensteiner, and Matthew Ward 529–540

Efficient Constraint Evaluation in Categorical Sequential Pattern Mining for Trajectory Databases
Leticia Gomez and Alejandro Vaisman 541–552

Heterogeneous & Distributed

Flexible and Efficient Querying and Ranking on Hyperlinked Data Sources
Ramakrishna Varadarajan, Hector Rodriguez-Drumond, Vagelis Hristidis, Louiqa Raschid, Maria-Esther Vidal, and Luis Daniel Ibáñez 553–564

RankClus: Integrating Clustering with Ranking for Heterogenous Information Network Analysis
Yizhou Sun, Jiawei Han, Peizhang Zhao, Zhiyun Yin, Hong Cheng, and Tianyi Wu 565–576