
Advances in Database Technology — EDBT 2017

20th International Conference
on Extending Database Technology
Venice, Italy, March 21–24, 2017
Proceedings

Editors

Volker Markl
Salvatore Orlando
Bernhard Mitschang
Periklis Andritsos
Kai-Uwe Sattler
Sebastian Breß



Advances in Database Technology — EDBT 2017
Proceedings of the 20th International Conference
on Extending Database Technology
Venice, Italy, March 21–24, 2017

Series ISSN: 2367-2005

Editors

Volker Markl, Technische Universität Berlin (TU Berlin), Germany
Salvatore Orlando, Ca' Foscari University of Venice (CFU), Italy
Bernhard Mitschang, Universität Stuttgart, Germany
Periklis Andritsos, University of Toronto, Canada
Kai-Uwe Sattler, TU Ilmenau, Germany
Sebastian Breß, German Research Center for Artificial Intelligence (DFKI), Germany



OpenProceedings.org
University of Konstanz
University Library
78457 Konstanz, Germany

COPYRIGHT NOTICE: Copyright © 2017 by the authors of the individual papers.

Distribution of all material contained in this volume is permitted under the terms of the Creative Commons license
CC-by-nc-nd 4.0

OpenProceedings ISBN: 978-3-89318-073-8

DOI of this front matter: 10.5441/002/edbt.2017.01

Foreword

The International Conference on Extending Database Technology (EDBT) is a leading international forum for database researchers, practitioners, developers, and users to discuss cutting-edge ideas, and to exchange techniques, tools, and experiences related to data management. Data management is an essential enabling technology for scientific, engineering, business, and social communities. Data management technology is driven by the requirements of applications across many scientific and business communities, and runs on diverse technical platforms associated with the web, enterprises, clouds and mobile devices. The database community has a continuing tradition of contributing with models, algorithms, and architectures, to the set of tools and applications enabling day-to-day functioning of our societies. Faced with the broad challenges of today's applications, data management technology constantly broadens its reach, exploiting new hardware and software to achieve innovative results.

EDBT 2017 solicited submissions of original research contributions, as well as descriptions of industrial and application achievements, and proposals for tutorials and software demonstrations. We encouraged submissions of research papers related to all aspects of data management defined broadly, and particularly encouraged work on topics of emerging interest in the research and development communities.

In addition to regular research paper submissions, EDBT 2017 solicited the submission of research papers that come within special topics of interest: "Vision", "Experiments and Analyses" and "Database Technology and Behavior, Security, Ethics, Rights and Duties of Citizens". These papers were reviewed by the same program committee as regular research papers. However, a dedicated co-chair for each special topic provided specific instructions to the reviewers of these papers and coordinated discussions, decisions, and meta-review formulation.

One innovation of EDBT 2017 is the solicitation of short papers, which are presented as posters at the plenary poster session of the conference. These short papers provide an opportunity to describe significant work in progress or research that is best communicated in an interactive or graphical format. In particular, these works contain smaller or more speculative ideas, controversial research topics, and new applications of old ideas or the reworking of previous studies. Short papers were reviewed by the research program committee in a second, independent call after the regular research paper submissions had been reviewed and decided. The program committees of EDBT accepted 37 out of 168 submitted regular research papers, resulting in an acceptance rate of 22% for the research track; 22 out of 93 submitted short papers, resulting in an acceptance rate of 23.6% for short research papers; 18 out of 45 demos, resulting in an acceptance rate of 40% for the demonstration track; 9 out of 24 industrial and application papers, resulting in an acceptance rate of 37.5%, as well as 3 out of 8 tutorials, again resulting in an acceptance rate of 37.5%.

The papers will be presented in nine research paper sessions, four industrial and application sessions (one invited), two plenary poster sessions, and two demo sessions. In addition, the program features six workshops, one of which is dedicated to European Research projects with a focus on the Horizon 2020 program, four joint keynotes with the ICDT conference, three tutorials, and one panel on the special topic "Database Technology and Behavior, Security, Ethics, Rights and Duties of Citizens". I would like to thank all authors for their contributions, as a successful conference crucially relies on high-quality submissions. The submission numbers indicate a healthy EDBT community. I also would like to thank all co-chairs and reviewers for serving on the EDBT program committee, in particular for the timely handling of all reviews and discussions with a high degree of professionalism and very high review and discussion quality. This enabled us to notify authors with no or only very little delay despite several reviewing cycles and only one month of reviewing and discussion time. Even though these community service contributions require a lot of work on a tight schedule, they are what make our research community function and ensure the overall impact of research in our field.

I firmly believe that we can look forward to an interesting program and exciting conference on March 21–24, 2017, in Venice.

Volker Markl
EDBT 2017 Program Chair

Program Committee Members

Research Program Committee

Ziawasch Abedjan, TU Berlin
Bernd Amann, U Pierre et Marie Curie
Walid G. Aref, Purdue U
Khalid Belhajjame, U Paris-Dauphine
Michael Böhlen, U Zürich
Klemens Böhm, Karlsruhe Inst. Techn. (KIT)
Francesco Bonchi, Yahoo! Labs
Angela Bonifati, Lille 1 U
Philippe Bonnet, IT U Copenhagen
Sebastian Breß, DFKI
Nieves Brisaboa, U La Coruna
Lei Chen, Hong Kong U of Sc. & Techn.
Sara Cohen, Hebrew U Jerusalem
Dario Colazzo, U Paris-Dauphine
Alfredo Cuzzocrea, U Trieste
Antonios Deligiannakis, TU Crete
Elena Ferrari, U Insubria
Peter M. Fischer, U Freiburg
Avigdor Gal, Technion
Helena Galhardas, U Lisbon
Johann Gamper, Free U Bozen-Bolzano
Minos Garofalakis, TU Crete
Tingjian Ge, U Mass. Lowell
François Goasdoué, U Rennes 1
Goetz Graefe, HP Labs
Michael Grossniklaus, U Konstanz
Moerkotte Guido, U Mannheim
Wook-Shin Han, Postech
Thomas Heinis, Imperial College London
Arantza Illarramendi, U País Vasco
Alekh Jindal, MIT
Alfons Kemper, TU Munich
Martin Kersten, CWI
Arijit Khan, Nanyang Technological U
Georg Lausen, U Freiburg
Wang-Chien Lee, Penn State U
Wolfgang Lehner, TU Dresden
Chuan Lei, NEC Labs America
Hong Va Leong, Hong Kong Polytechnic U
Ulf Leser, HU Berlin
Guoliang Li, Tsinghua U
Eric Lo, Polytecnic U Hong Kong
Sofian Maabout, U Bordeaux
Patrick Marcel, U Tours
Sebastian Michel, TU Kaiserslautern
Eltabakh Mohammed, Worcester PI
Mohamed Mokbel, U Minnesota
Kjetil Nørvåg, Norwegian U Sc. & Techn.
Esther Pacitti, U Montpellier and INRIA
Marta Patino-Martinez, Politecnica Madrid
Torben Bach Pedersen, Aalborg U
Jian Pei, Simon Fraser U
Jean-Marc Petit, Université Lyon, CNRS
Peter Pietzuch, Imperial College London
Evaggelia Pitoura, U Ioannina
Pascal Poncelet, U Montpellier
Nicoleta Preda, U Versailles
Philippe Pucheral, U Versailles
Jorge-Arnulfo Quiane-Ruiz, Qatar CRI
Maya Ramanath, IIT Delhi
Rodolfo F. Resende, U Fed. Minas Gerais
Tore Risch, Uppsala U
Elke A. Rundensteiner, Worcester PI
Pierangela Samarati, Università Milano
Mohamed Sarwat, Arizona State U
Sebastian Schelter, Amazon
Marc H. Scholl, U Konstanz
Heiko Schuldt, U Basel
Timos Sellis, Swinburne U Techn.
Divesh Srivastava AT&T Labs-Research
Jianwen Su, UC Santa Barbara
Nan Tang, Qatar Comptg. Res. Inst.
Martin Theobald, U Antwerp
Farouk Toumani, U Clermont-Ferrand
Peter Triantafillou, U Glasgow
Vassilis Tsotras, UC Riverside
Yannis Velegrakis, U Trento
Daisy Wang, U Florida
Wei Wang, U New South Wales
Yinghui Wu, Washington State U
Xiaochun Yang, Northeastern U
Demetris Zeinalipour, U Cyprus
Xiaofang Zhou, The U Queensland
Esteban Zimanyi, U Libre Bruxelles

Industrial Program Committee

Nipun Agarwal, Oracle
Divy Agrawal, UC Santa Barbara
Stefan Dessloch, TU Kaiserslautern
Michael Gesmann, Software AG
Christoph Groeger, Bosch
Martin Grund, Amazon
Thomas Jörg, Google
Frank Leymann, U Stuttgart
Kostas Magoutis, U Ioannina
Albert Maier, IBM
Christian Mathis, SAP
Bernhard Mitschang, U Stuttgart
Mario Nascimento, U Alberta
Berthold Reinwald, IBM
Fabrizio Silvestri, Facebook
Eric Simon, SAP

Demonstration Program Committee

Ira Assent, Aarhus U
Carsten Binnig, Brown U
Angela Bonifati, Lille 1 Un
Isabel Cruz, U Illinois Chicago
Michael Gertz, Heidelberg U
Dirk Habich, TU Dresden
Sven Helmer, Free U Bolzano
Katja Hose, Aalborg U
Stefan Manegold, CWI Amsterdam

Anisoara Nica, SAP SE
Norman Paton, U Manchester
Heiko Schuldt, U Basel
Letizia Tanca, Politecnico Milano
Kian-Lee Tan, National U Singapore
Vassalos Vasilis, Athens U
Demetris Zeinalipour, MPI Informatics & U
Cyprus

External Reviewers

Julien Aligon, Toulouse U
Mohand Boughanem, IRIT U Toulouse
Marc Bux, HU Berlin
Sean Chester, Norwegian U Sc. & Techn.
Dong-Wan Choi, Simon Fraser U
Lingyang Chu, Simon Fraser U
Mahfoud Djedaini, U Tours
Miguel Forte, Workable
Olga Gkountouna, IMIS/RC Athens
Aris Gloulalas-Divanis, IBM
Nicolas Labroche, U Tours
Mirjana Mazuran, Politecnico Milano
Xinghao Pan, UC Berkeley
Miguel Pardal, IST, U Lisbon
Emanuele Rabosio, Politecnico Milano
Patrick Schäfer, HU Berlin
George Stoilos, AUEB
Yu Yang, Simon Fraser U

Test-of-Time Award

In 2014, the Extended Database Technology conference (EDBT) began awarding the EDBT test-of-time (ToT) award, with the goal of recognising papers presented at EDBT Conferences that have had the most impact in terms of research, methodology, conceptual contribution, or transfer to practice.

This year, covering the conferences from 1996 to 2002, the award has been given to:

Mining Sequential Patterns: Generalizations and Performance Improvements.

by Ramakrishnan Srikant, Rakesh Agrawal

published in the EDBT 1996 proceedings, 3–17.

This paper has made substantial contributions to data mining, and has had great influence on the work of others, as reflected by over 2900 citations on Google Scholar.

The paper formalizes a new variant of the problem of mining *sequential* patterns and develops and implements GSP, an algorithm to solve this problem. This paper extends the definition of sequence mining that was introduced by the same authors in a previous publication: Mining Sequential Patterns. ICDE 1995. The goal is to discover all sequential patterns with a user-specified minimum support from a database of sequences, where each sequence is a list of transactions ordered by transaction-time, and each transaction is a set of items. The proposed extensions are:

1. Time constraints: the authors generalised their previous definition of sequential patterns to admit max-gap and min-gap time constraints between adjacent elements of a sequential pattern.
2. Sliding windows: the authors relaxed the restriction that all the items in an element of a sequential pattern must come from the same transaction, and allowed a user-specified window-size within which the items can be present.
3. Taxonomies: the sequential patterns may include items across different levels of a taxonomy.

GSP guarantees that all rules that have a user-specified minimum support. It is shown to be much faster than the AprioriAll algorithm in the previous publication (on both synthetic and real data). GSP has been implemented as part of the Quest data mining prototype at IBM Research, and is incorporated in the IBM data mining product.

The EDBT 2017 Test of Time Award Committee consisted of Peter Triantafillou, Gustavo Alonso, Sihem Amer-Yahia, Ralf Hartmut Güting and Volker Markl.

The EDBT ToT award for 2017 will be presented during the EDBT/ICDT 2017 Joint Conference, March 21–24, in Venice, Italy (<http://edbticdt2017.unive.it/>).

Table of Contents

Foreword	i
Program Committee Members	ii
Test-of-Time Award	iv
Table of Contents	v
Research Papers	
Parallel Array-Based Single- and Multi-Source Breadth First Searches on Large Dense Graphs <i>Moritz Kaufmann, Manuel Then, Alfons Kemper, Thomas Neumann</i>	1
GraphCache: A Caching System for Graph Queries <i>Jing Wang, Nikos Ntarmos, Peter Triantafillou</i>	13
Subgraph Querying with Parallel Use of Query Rewritings and Alternative Algorithms <i>Foteini Katsarou, Nikos Ntarmos, Peter Triantafillou</i>	25
Reverse k-Ranks Queries on Large Graphs <i>Yuqiu Qian, Hui Li, Nikos Mamoulis, Yu Liu, David Cheung</i>	37
Analytics on Fast Data: Main-Memory Database Systems versus Modern Streaming Systems <i>Andreas Kipf, Varun Pandey, Jan Böttcher, Lucas Braun, Thomas Neumann, Alfons Kemper</i>	49
Self-managed collections: Off-heap memory management for scalable query-dominated collections <i>Fabian Nagel, Gavin Bierman, Aleksandar Dragojevic, Stratis Viglas</i>	61
Lightweight Data Compression Algorithms: An Experimental Survey (Experiments and Analyses) <i>Patrick Damme, Dirk Habich, Juliana Hildebrandt, Wolfgang Lehner</i>	72
SQL- and Operator-centric Data Analytics in Relational Main-Memory Databases <i>Linnea Passing, Manuel Then, Nina Hubig, Harald Lang, Michael Schreier, Stephan Günemann, Alfons Kemper, Thomas Neumann</i>	84
Data Exploration with SQL using Machine Learning Techniques <i>Julien Cumin, Jean-Marc Petit, Vasile-Marian Scuturici, Sabina Surdu</i>	96
Task-Optimized Group Search for Social Internet of Things <i>Chih-Ya Shen, Hong-Han Shuai, Kuo-Feng Hsu, Ming-Syan Chen</i>	108
Finding Socio-Textual Associations Among Locations <i>Paras Mehta, Dimitris Sacharidis, Dimitrios Skoutas, Agnes Voisard</i>	120
COP: Planning Conflicts for Faster Parallel Transactional Machine Learning <i>Faisal Nawab, Divy Agrawal, Amr El Abbadi, Sanjay Chawla</i>	132
ChronicleDB: A High-Performance Event Store <i>Marc Seidemann, Bernhard Seeger</i>	144
EXstream: Explaining Anomalies in Event Stream Monitoring <i>Haopeng Zhang, Yanlei Diao, Alexandra Meliou</i>	156
Real Time Contextual Summarization of Highly Dynamic Data Streams <i>Manoj Agarwal, Krithi Ramamritham</i>	168
An Effective and Efficient Truth Discovery Framework over Data Streams <i>Tianyi Li, Yu Gu, Xiangmin Zhou, Qian Ma, Ge Yu</i>	180

Maritime data integration and analysis: recent progress and research challenges <i>Christophe Claramunt, Cyril Ray, Elena Camossi, Anne-Laure Joussemme, Melita Hadzagic, Andrienko Gennady, Natalia Andrienko, Yannis Theodoridis, George Vouros, Loïc Salmon</i>	192
DeepSea: Progressive Workload-Aware Partitioning of Materialized Views in Scalable Data Analytics <i>Jiang Du, Renée Miller, Boris Glavic, Wei Tan</i>	198
Matching Web Tables To DBpedia - A Feature Utility Study <i>Dominique Ritze, Christian Bizer</i>	210
Schema Inference for Massive JSON Datasets <i>Mohamed-Amine Baazizi, Housseem Ben Lahmar, Dario Colazzo, Giorgio Ghelli, Carlo Sartiani</i>	222
Distributed in-memory SPARQL Processing via DOF Analysis <i>Roberto De Virgilio</i>	234
Motivation-Aware Task Assignment in Crowdsourcing <i>Julien Pilourdault, Sihem Amer-Yahia, Dongwon Lee, Senjuti Basu Roy</i>	246
A Probabilistic Framework for Estimating Pairwise Distances Through Crowdsourcing <i>Habibur Rahman, Senjuti Basu Roy, Gautam Das</i>	258
Information Propagation in Interaction Networks <i>Rohit Kumar, Toon Calders</i>	270
Flexible Caching in Trie Joins <i>Oren Kalinsky, Yoav Etsion, Benny Kimelfeld</i>	282
Querying Improvement Strategies <i>Guolei Yang, Ying Cai</i>	294
Grid-Index Algorithm for Reverse Rank Queries <i>Yuyang Dong, Hanxiong Chen, Jeffrey Xu Yu, Kazutaka Furuse, Hiroyuki Kitagawa</i>	306
Parallel and Distributed Processing of Spatial Preference Queries using Keywords <i>Christos Doulkeridis, Akrivi Vlachou, Dimitris Mpeatas, Nikos Mamoulis</i>	318
Continuous Imputation of Missing Values in Streams of Pattern-Determining Time Series <i>Kevin Wellenzohn, Michael Böhlen, Anton Dignös, Johann Gamper, Hannes Mitterer</i>	330
Data-driven Schema Normalization <i>Thorsten Papenbrock, Felix Naumann</i>	342
Towards Interactive Debugging of Rule-based Entity Matching <i>Fatemah Panahi, Wentao Wu, AnHai Doan, Jeffrey Naughton</i>	354
Convergent Interactive Inference with Leaky Joins <i>Ying Yang, Oliver Kennedy</i>	366
Efficient Motif Discovery in Spatial Trajectories Using Discrete Fréchet Distance <i>Bo Tang, Man Lung Yiu, Kyriakos Mouratidis, Kai Wang</i>	378
Group Trip Scheduling (GTS) Queries in Spatial Databases <i>Roksana Jahan, Tanzima Hashem, Sukarna Barua</i>	390
Towards Efficient Maintenance of Continuous MaxRS Query for Trajectories <i>Muhammed Mas-ud Hussain, Kazi Ashik Islam, Goce Trajcevski, Mohammed Eunus Ali</i>	402
Exact and Approximate Algorithms for Finding k-Shortest Paths with Limited Overlap <i>Theodoros Chondrogiannis, Panagiotis Bouros, Johann Gamper, Ulf Leser</i>	414
Efficient Parallel Spatial Skyline Evaluation Using MapReduce <i>Wenlu Wang, Ji Zhang, MinTe Sun, Wei-Shinn Ku</i>	426

Poster Papers

Vertex-Centric Graph Processing: Good, Bad, and the Ugly <i>Arijit Khan</i>	438
Top-k Skyline Groups Queries <i>Haoyang Zhu, Peidong Zhu, Xiaoyong Li, Qiang Liu</i>	442
Towards sequenced semantics for evolving graphs <i>Vera Zaychik Moffitt, Julia Stoyanovich</i>	446
Crowdsourcing Strategies for Text Creation Tasks <i>Ria Mae Borromeo, Maha Alsaysneh, Sihem Amer-Yahia, Vincent Leroy</i>	450
Hybrid LSH: Faster Near Neighbors Reporting in High-dimensional Space <i>Ninh Pham</i>	454
SPST-Index: A Self-Pruning Splay Tree Index for Caching Database Cracking <i>Pedro Holanda, Eduardo Cunha de Almeida</i>	458
In-Memory Spatial Join: The Data Matters! <i>Sadegh Nobari, Qiang Qu, Christian Jensen</i>	462
Fairness and Transparency in Crowdsourcing <i>Ria Mae Borromeo, Thomas Laurent, Motomichi Toyama, Sihem Amer-Yahia</i>	466
TASWEET: Optimizing Disjunctive Path Queries in Graph Databases <i>Zahid Abul-Basher, Nikolay Yakovets, Parke Godfrey, Shadi Ghajar-Khosrav, Mark Chignell</i>	470
A New Division Operator to Handle Complex Objects in Very Large Relational Datasets <i>Andre Gonzaga, Robson Leonardo Ferreira Cordeiro</i>	474
Protecting Location Privacy in Spatial Crowdsourcing using Encrypted Data <i>Bozhong Liu, Ling Chen, Xingquan Zhu, Ying Zhang, Chengqi Zhang, Weidong Qiu</i>	478
Break the Windows: Explicit State Management for Stream Processing Systems <i>Alessandro Margara, Daniele Dell'Aglio, Abraham Bernstein</i>	482
Stability notions in synthetic graph generation: a preliminary study <i>Wilco van Leeuwen, Angela Bonifati, George Fletcher, Nikolay Yakovets</i>	486
Big Spatial Data Processing Frameworks: Feature and Performance Evaluation <i>Stefan Hagedorn, Philipp Götze, Kai-Uwe Sattler</i>	490
Implementation and Evaluation of Genome Type Processing for Disease-Causal Gene Studies on DBMS <i>Yoshifumi Ujibashi, Lilian Harada</i>	494
Authority-based Team Discovery in Social Networks <i>Morteza Zihayat, Aijun An, Lukasz Golab, Mehdi Kargar, Jaroslaw Szlichta</i>	498
Correlation-Aware Distance Measures for Data Series <i>Katsiaryna Mirylenka, Michele Dallachiesa, Themis Palpanas</i>	502
Load balancing for Key Value Data Stores <i>Ainhoa Azqueta-Alzúaz, Ivan Brondino, Marta Patino-Martinez, Ricardo Jimenez-Peris</i>	506
Entity Matching on Web Tables: a Table Embeddings approach for Blocking <i>Anna Lisa Gentile, Petar Ristoski, Steffen Eckel, Dominique Ritze, Heiko Paulheim</i>	510
Fast Subsequence Search on Time Series Data <i>Yuhong Li, Bo Tang, Leong Hou U, Man Lung Yiu, Zhiguo Gong</i>	514
Progressive Recovery of Correlated Failures in Distributed Stream Processing Engines <i>Li Su, Yongluan Zhou</i>	518

Optimal Obstructed Sequenced Route Queries in Spatial Databases <i>Anika Anwar, Tanzima Hashem</i>	522
Demonstrations	
I ² : Interactive Real-Time Visualization for Streaming Data <i>Jonas Traub, Nikolaas Steenbergen, Philipp Grulich, Tilmann Rabl, Volker Markl</i>	526
HDM: Optimized Big Data Processing with Data Provenance <i>Dongyao Wu, Sherif Sakr, Liming Zhu</i>	530
ChaseFUN: a Data Exchange Engine for Functional Dependencies at Scale <i>Angela Bonifati, Ioana Ileana, Michele Linardi</i>	534
GnosisMiner: Reading Order Recommendations over Document Collections <i>Georgia Koutrika, Alkis Simitsis, Yannis Ioannidis</i>	538
MovieFinder: A Movie Search System via Graph Pattern Matching <i>Xin Wang, Chengye Yu, Enyang Zhang, Tong Du</i>	542
VAT: A System for Data-Driven Biodiversity Research <i>Christian Beilschmidt, Johannes Dröner, Michael Mattig, Bernhard Seeger</i>	546
SDOS: Using Trusted Platform Modules for Secure Cryptographic Deletion in the Swift Object Store <i>Tim Waizenegger, Frank Wagner, Cataldo Mega</i>	550
Come and crash our database! – Instant recovery in action <i>Caetano Sauer, Gilson Souza, Goetz Graefe, Theo Härder</i>	554
TOP: Spatio-Temporal Detection and Summarization of Locally Trending Topics in Microblog Posts <i>Paras Mehta, Manuel Kotlarski, Dimitrios Skoutas, Dimitris Sacharidis, Kostas Patroumpas, Agnes Voisard</i>	558
Insights into the Comparative Evaluation of Lightweight Data Compression Algorithms <i>Patrick Damme, Dirk Habich, Juliana Hildebrandt, Wolfgang Lehner</i>	562
Multi-workflow optimization in PAW <i>Maxim Filatov, Verena Kantere</i>	566
Efficient spatio-temporal event processing with STARK <i>Stefan Hagedorn, Timo Räth</i>	570
Context-Aware Proactive Personalization of Linear Audio Content <i>Paolo Casagrande, Maria Luisa Sapino, K. Selcuk Candan</i>	574
In Search for Relevant, Diverse and Crowd-screen Points of Interests <i>Xiaoyu Ge, Samanvay Panati, Konstantinos Pelechrinis, Panos Chrysanthis, Mohamed Sharaf</i>	578
SIMDMS: Data Management and Analysis to Support Decision Making through Large Simulation Ensembles <i>Silvestro Poccia, Maria Luisa Sapino, Sicong Liu, Xilun Chen, Yash Garg, Shengyu Huang, Jung Kim, Xinsheng Li, Parth Nagarkar, K. Selcuk Candan</i>	582
In-Place Appends for Real: DBMS Overwrites on Flash without Erase <i>Sergey Hardock, Ilia Petrov, Robert Gottstein, Alejandro Buchmann</i>	586
CAESAR: Context-Aware Event Stream Analytics for Urban Transportation Services <i>Olga Poppe, Chuan Lei, Elke Rundensteiner, Dan Dougherty, Goutham Deva, Nicholas Fajardo, James Owens, Thomas Schweich, MaryAnn VanValkenburg, Sarun Paisarnsrisomsuk, Pitchaya Wiratchotisanian, George Gettel, Robert Hollinger, Devin Roberts, Daniel Tocco</i>	590

Building Multi-Resolution Event-Enriched Maps From Social Data <i>Faizan Ur Rehman, Imad Afyouni, Ahmed Lbath, Sohaib Khan, Saleh Basalamah, Mohamed Mokbel</i>	594
Tutorials	
Declarative Graph Querying in Practice and Theory <i>George Fletcher, Hannes Voigt, Nikolay Yakovets</i>	598
Multi-model Data Management: What's New and What's Next? <i>Jiaheng Lu, Irena Holubová</i>	602
Data Security and Privacy for Outsourced Data In the Cloud <i>Cetin Sahin, Amr El Abbadi</i>	606
Industrial and Applications Papers	
Improving Company Recognition from Unstructured Text by using Dictionaries <i>Michael Loster, Zuo Zhe, Felix Naumann, Oliver Maspfuhl, Dirk Thomas</i>	610
Temporal group linkage and evolution analysis for census data <i>Victor Christen, Anika Groß, Jeffrey Fisher, Qing Wang, Peter Christen, Erhard Rahm</i>	620
In-DBMS Sampling-based Sub-trajectory Clustering <i>Nikos Pelekis, Panagiotis Tampakis, Marios Vondas, Costas Panagiotakis, Yannis Theodoridis</i>	632
Powering Archive Store Query Processing via Join Indices <i>Joseph D'silva, Bettina Kemme, Richard Grondin, Evgueni Fadeitchev</i>	644
RDF Keyword-based Query Technology Meets a Real-World Dataset <i>Grettel García, Yenier Izquierdo, Elisa Menendez, Frederic Dartayre, Marco Casanova</i>	656
Hi-WAY: Execution of Scientific Workflows on Hadoop YARN <i>Marc Bux, Jörgen Brandt, Carl Witt, Jim Dowling, Ulf Leser</i>	668
Buddy Instance - A Mechanism for Increasing Availability in Shared-Disk Clusters <i>Anjan Kumar Amirishetty, Yunrui Li, Tolga Yurek, Mahesh Girkar, Wilson Chan, Graham Ivey, Vsevolod Panteleen, Ken Wong</i>	680
DBaaS Cloud Capacity Planning - Accounting for Dynamic RDBMS System that Employ Clustering and Standby Architectures <i>Antony Higginson, Norman Paton, Suzanne Embury, Clive Bostock</i>	687
Herding the elephants: Workload-level optimization strategies for Hadoop <i>Sandeep Akinapelli, Ravi Shetye, Sangeeta T</i>	699