

---

# Advances in Database Technology — EDBT 2018

21st International Conference  
on Extending Database Technology  
Vienna, Austria, March 26–29, 2018  
Proceedings

*Editors*

Michael Böhlen  
Reinhard Pichler  
Norman May  
Erhard Rahm  
Shan-Hung Wu  
Katja Hose



Advances in Database Technology – EDBT 2018  
Proceedings of the 21st International Conference  
on Extending Database Technology  
Vienna, Austria, March 26–29, 2018

Series ISSN: 2367-2005

*Editors*

Michael Böhlen, University of Zürich, Switzerland  
Reinhard Pichler, TU Wien, Austria  
Norman May, SAP Research, Germany  
Erhard Rahm, University of Leipzig, Germany  
Shan-Hung Wu, National Tsing Hua University, Taiwan  
Katja Hose, Aalborg University, Denmark



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

COPYRIGHT NOTICE: Copyright © 2018 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: 978-3-89318-078-3

DOI of this front matter: 10.5441/002/edbt.2018.01

# Foreword

The International Conference on Extending Database Technology (EDBT) is a leading international forum for database researchers, developers, and users to present and discuss cutting-edge ideas, and to exchange techniques, tools, and experiences related to data management. Data management is an essential enabling technology for scientific, business, and social communities. It is driven by the requirements of applications across many scientific, business and social 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 that enable 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 2018 solicited submissions of original research contributions, descriptions of industrial solutions and applications, and proposals for tutorials and software demonstrations. We encouraged submissions of research papers related to all aspects of data management. In addition to regular research paper submissions, EDBT 2018 again encouraged the submission of short research papers, which includes visionary papers that provide a forum for the identification and discussion of new or emerging areas, innovative or risky approaches, or emerging applications that require extensions of established techniques. Short papers are presented as posters at plenary poster sessions of the conference. This provides an excellent opportunity to describe significant work in progress or research that is best communicated interactively and fosters discussions.

The program committees of EDBT accepted 35 out of 142 submitted regular research papers, resulting in an acceptance rate of 24.6% for the research track; 23 out of 84 submitted short research papers, resulting in an acceptance rate of 27.4% for short research papers; 16 out of 37 demos, resulting in an acceptance rate of 43.2% for the demonstration track; and 10 out of 28 industrial and application papers, resulting in an acceptance rate of 35.7% for industrial and application papers. The papers will be presented in twelve research paper sessions, three industrial and application sessions, as well as two plenary poster and demonstration sessions. The program additionally features four workshops, one of which is the well-established DOLAP workshop that has successfully been co-located with EDBT since many years. Finally, the conference program includes four tutorials and an EDBT and ICDT joint session on research challenges.

I would like to thank all authors for their contributions: a successful conference crucially depends on high-quality submissions. I also would like to thank all reviewers for serving on the EDBT 2018 program committee, in particular for the high quality and timely handling of all reviews and discussions. This community service requires a lot of work on a tight schedule, and is what makes our research community function and ensures the sustained impact of our research. Thanks to this effort we can look forward to an exciting program and attractive EDBT conference in Vienna from March 26-29, 2018. A warm thanks to Norman Paton and Divesh Srivastava for serving on the Test-of-Time Award committee to select the paper from EDBT 2008 that has had the most lasting influence. The committee selected the paper Social ties and their relevance to churn in mobile telecom networks by Koustuv Dasgupta, Rahul Singh, Balaji Viswanathan, Dipanjan Chakraborty, Sougata Mukherjea, Amit A. Nanavati, and Anupam Josi for the test-of-time award.

Lei Chen, Wolfgang Lehner and Kian-Lee Tan generously accepted to serve on the Best Paper committee. As best paper, the committee selected the paper Temporally-Biased Sampling for Online Model Management by Brian Hentschel from Harvard University, Peter Haas from the University of Massachusetts Amherst, and Yuanyuan Tian from IBM Almaden Research Center. The EDBT best paper runner-up was awarded to the paper GeoAlign: Interpolating Aggregates over Unaligned Partitions by Jie Song from the University of Michigan, Danai Koutra from the University of Michigan, Murali Mani from the University of Michigan, Flint, and H. Jagadish from the University of Michigan. Congratulations to the awardees and a warm thanks to the committee members for their work.

The EDBT 2018 program is the result of the joint effort of many people who shared their experience and time to contribute to the EDBT 2018 program and make the conference a success. Norman May served as PC chair for industrial and application papers; Shan-Hung as PC chair for the demonstration track; Erhard Rahm as tutorial chair; Nikolaus Augsten as workshop chair; and Dan Olteanu as challenge session organizer. My warmest thank to all these people.

The general chair, Reinhard Pichler and the local organizers worked hard to make all necessary arrangements for a successful event. Special thanks to Katja Hose, the proceedings chair; Dimitris Sacharidis, the sponsorship chair; and Shqiponja Ahmetaj, Markus Kröll, and Wolfgang Fischl, the website chairs, for tirelessly finding solutions for all our needs and making things happen. Norman Paton was most helpful in advising and coordinating with the EDBT Executive Board.

I hope that you find EDBT 2018 inspiring, informative, and enjoyable and look forward to meeting you in Vienna.

Michael Böhlen  
EDBT 2018 Program Chair

# Program Committee Members

## Research Program Committee

Ahmed Eldawy, UC Riverside  
Angela Bonifati, U Lyon  
Anton Dignös, Free U Bozen-Bolzano  
Arash Termehchy, Oregon State U  
Arijit Khan, Nanyang Techn. U  
Arnab Bhattacharya, IIT Kanpur  
Bernhard Seeger, U Marburg  
Bin Cui, Peking U  
Boris Glavic, Illinois Inst. of Techn.  
Carsten Binning, Brown U  
Ce Zhang, ETH Zurich  
Curtis Dyreson, Utah State U  
Daniele Dell’Aglia, U Zürich  
Davide Martinenghi, Politecnico di Milano  
Denilson Barbosa, U Alberta  
Elena Ferrari, U Insubria  
Floris Geerts, U Antwerp  
Giansalvatore Mecca, U Basilicata  
Gottfried Vossen, U Münster  
Guoliang Li, Tsinghua U  
Heiko Schuldt, U Basel  
Hong Va Leong, Hong Kong Polytechnic U  
Jens Teubner, TU Dortmund  
K. Selçuk Candan, Arizona State U  
Karl Aberer, EPF Lausanne  
Kian-Lee Tan, National U Singapore  
Kjetil Norvag, NTNU  
Klaus Berberich, MPI Informatics  
Klemens Böhm, Karlsruhe Inst. Technology  
Kostas Stefanidis, U Tampere  
Kyriakos Mouratidis, Singapore Mgmt. U  
Lei Chen, Hong Kong U Sc. & Tech.  
Man Lung Yiu, Hong Kong Polytechn. U  
Marc Spaniol, U Caen Basse-Normandie  
Martin Theobald, U Luxembourg  
Matthias Renz, George Mason U  
Maurice van Keulen, UTwente  
Mohamed Eltabakh, Worcester PI  
Mohammad Sadoghi, Purdue U  
Mourad Ouzzani, Qatar Cmptg. Res. Inst.  
Nikos Pelekis, U Piraeus  
Nikos Mamoulis, U Ioannina  
Panagiotis Papapetrou, Stockholm U  
Panagiotis Bouros, Aarhus U  
Panagiotis Karras, Aalborg U  
Paolo Papotti, EURECOM  
Periklis Andritsos, U Toronto  
Peter Triantafillou, U Glasgow  
Philippe Cudre-Mauroux, U Fribourg  
Pierre Senellart, ENS  
Rainer Gemulla, U Mannheim  
Ralf Hartmut Güting, FU Hagen  
Reynold Cheng, U Hong Kong  
Riccardo Torlone, U Roma Tre  
Ryan Stutsman, U Utah  
Sabrina De Capitani di Vimercati, U Milano  
Sebastian Michel, TU Kaiserslautern  
Senjuti Roy, New Jersey Inst. Techn.  
Sherif Sakr, U New South Wales  
Sourav S Bhowmick, Nanyang Techn. U  
Stefan Manegold, CWI Amsterdam  
Stratis Viglas, U Edinburgh  
Sven Helmer, Free U Bozen-Bolzano  
Themis Palpanas, Paris Descartes U  
Theodore Johnson, AT&T Labs  
Thomas Seidl, LMU München  
Timos Sellis, Swinburne UT  
Tore Risch, Uppsalla U  
Torsten Grust, U Tübingen  
Ulf Leser, HU Berlin  
Verena Kantere, U Geneva  
Wai Kit Wong, Hang Seng Mgmt. College  
Walid Aref, Purdue U  
Wei-Shinn Ku, Auburn U  
Wolfgang Lehner, TU Dresden  
Wook-Shin Han, Postech  
Xin Huang, Hong Kong Baptist U  
Yang Cao, U Edinburgh  
Yannis Theodoridis, U Piraeus  
Yannis Velegarakis, U Trento  
Yaron Kanza, Technion  
Ying Yang, State U New York  
Ying Zhang, UT Sydney

### **Industrial Program Committee**

Berthold Reinwald, IBM Research  
Carl-Christian Kanne, Workday  
Christian Mathis, SAP  
Danica Porobic, Oracle  
Eric Simon, SAP-BO  
Florian Funke, Snowflake  
Jörg Schad, Mesosphere  
Manuel Then, Tableau  
Martin Grund, Amazon  
Matthias Brantner, Oracle  
Matthias Boehm, IBM  
Philipp Unterbrunner, Facebook  
Pinar Tozun, IBM  
Stefan Mandl, Exasol  
Tobias Muehlbauer, Muehlbauer Tableau

### **Demonstration Program Committee**

Alessandro Campi, Politecnico di Milano  
Alfredo Cuzzocrea, U Trieste & ICAR-CNR  
Anisoara Nica, SAP SE Waterloo  
Berthold Reinwald, IBM Research  
Danai Symeonidou, French Nat.Inst. Agricultural Res.  
Demetris Zeinalipour, MPI Informatics & U Cyprus  
Dirk Habich, TU Dresden  
Elisa Quintarelli, Politecnico di Milano  
Eric Lo, Chinese U Hong Kong  
Ernest Teniente, UP Catalunya  
George Fletcher, TU Eindhoven  
Guoliang Li, Tsinghua U  
Haruo Yokota, Tokyo Inst. Technology  
Kai-Uwe Sattler, TU Ilmenau  
Katja Hose, Aalborg U  
Leopoldo Bertossi, U Carleton  
Letizia Tanca, Politecnico di Milano  
Michael Gertz, Heidelberg U  
Neil Conway, Mesosphere  
Norman Paton, U Manchester  
Tony Tan, Nat. Taiwan U  
Vasilis Vassalos, Athens U Economics & Business  
Yi-Leh Wu, Nat. Taiwan U Sci. & Techn.  
Yingyi Bu, Couchbase  
Zhifeng Bao, RMIT U

### **External Reviewers**

Abdallah Arioua, Lyon 1 U  
Alessio Conte, Nat.Inst. Informatics Tokyo  
Amit Gupta, EPF Lausanne  
Andreas Spitz, Heidelberg U  
Anna Beer, LMU Munich  
Ben McCamish, Oregon State U  
Bo Tang, Southern U Sc. & Techn. China  
Chenhao Ma, U Hong Kong  
Christian Beilschmidt, U Marburg  
Danhao Ding, U Hong Kong  
Daniyal Kazempour, LMU Munich  
Denis Martins, U Münster  
Dimitris Sacharidis, TU Wien  
Donatella Firmani, U Roma Tre

Donatello Santoro, U Basilicata  
Duong Chi Thang, EPF Lausanne  
Evgeniy Faerman, LMU Munich  
Fabio Valdés, FU Hagen  
Fan Zhang, U New South Wales  
Florian Funke, Snowflake Computing  
Florian Richter, LMU Munich  
Giulia Pretti, U Trento  
Hamza Harkous, EPF Lausanne  
Hanchen Wang, UT Sydney  
Harris Georgiou, U Piraeus  
Jan-Kristof Nidzwetzki, FU Hagen  
Janina Sontheim, LMU Munich  
Jens Lechtenbörger, U Münster  
Jiafeng Hu, U Hong Kong  
Jilian Zhang, Jinan U  
Johannes Droenner, U Marburg  
Jose Picado, Oregon State U  
Julian Busch, LMU Munich  
Kewen Liao, Swinburne UT  
Kiril Panev, TU Kaiserslautern  
Koninika Pal, TU Kaiserslautern  
Kostas Patroumpas, IMIS Athena Res. Ctr.  
Kostas Zoumpatianos, Harvard U  
Leschek Homann, U Münster  
Loc Do, U Hong Kong  
Manuel Hoffmann, TU Kaiserslautern  
Matteo Lissandrini, U Trento  
Max Berrendorf, LMU Munich  
Michael Körber, U Marburg  
Michael Mattig, U Marburg  
Michele Linardi, Paris Descartes U  
Mingjie Li, UT Sydney  
Nicolas Pflanzl, U Münster  
Nikolaus Glombiewski, U Marburg  
Panayiotis Smeros, EPF Lausanne  
Parisa Ataie, Oregon State U  
Patrick Schaefer, HU Berlin  
Pavlos Paraskevopoulos, George Mason U  
Peipei Yi, Hong Kong Baptist U  
Prithvi Sen, IBM Almaden  
Ruey-Cheng Chen, RMIT U  
Spencer Pearson, Purdue U  
Suyash Gupta, UC Davis  
Tam Nguyen Thanh, EPF Lausanne  
Thamir Qadah, Purdue U  
Theodoros Chondrogiannis, U Konstanz  
Thomas Fober, U Marburg  
Thomas Krause, HU Berlin  
Ugo Comignani, Lyon 1 U  
Vahid Ghadakchi, Oregon State U  
Yifeng Lu, LMU Munich  
Yixiang Fang, U Hong Kong  
Yodsawalai Chodpathumwan, Oregon State U  
Yongrui Qin, U Huddersfield  
You Peng, U New South Wales  
Yuli Jiang, Chinese U Hong Kong  
Zhipeng Huang, U Hong Kong  
Zhipeng Zhang, Peking U  
Zichen Zhu, U Hong Kong

# Test-of-Time Award

Since 2014, the Extended Database Technology (EDBT) Conference awards the EDBT test-of-time award, with the goal of recognizing papers presented at EDBT Conferences that have had the most impact in terms of research, methodology, conceptual contribution, or transfer to practice.

This year the award has been given to a paper from the EDBT 2008 Conference in Nantes, France. The award was bestowed upon the paper:

## **Social ties and their relevance to churn in mobile telecom networks**

by Koustuv Dasgupta, Rahul Singh, Balaji Viswanathan, Dipanjan Chakraborty, Sougata Mukherjea, Amit A. Nanavati and Anupam Josi

published in the EDBT 2008 Proceedings, pp. 668–677, DOI: 10.1145/1353343.1353424.

This industry track paper reports on an analysis of mobile telecoms data, with a view to predicting which customers will leave. The analysis involves commercial mobile telephony data, in which nodes are customers and edges represent calls. The hypothesis tested is that it is possible to predict who will leave a network based on earlier departures among their connections. The main technique investigated is the use of spreading activation, to predict the heat of nodes based on the heat of connected nodes. It is shown how the approach based on connections is more effective than classification techniques based on other properties of the nodes. As a result, the paper provides early and compelling experience on the combination an important real problem (churn in mobile telecom networks) with a powerful technique (social ties) and applies it on large real data (telecom operator network over 4 months). The approach has influenced many subsequent studies, for the same problem, but also for analyses involving different types of network and different hypotheses. Social network analysis continues as an important and active area ten years later, and this paper continues to be widely cited.

The EDBT 2018 Test-of-Time Award Committee consisted of Michael Böhlen, Divesh Srivastava and Norman Paton. The EDBT Test-of-Time award for 2018 will be presented during the EDBT/ICDT 2018 Conference, March 26–29, in Vienna, Austria (<http://edbticdt2018.at>).

# Best Paper Award

The best paper award was bestowed upon the paper:

## **Temporally-Biased Sampling for Online Model Management**

by Brian Hentschel from Harvard University, Peter Haas from the University of Massachusetts Amherst, and Yuanyuan Tian from the IBM Almaden Research Center. DOI: 10.5441/002/edbt.2018.11

The paper proposes a temporally-biased sampling method for a stream of batches that weighs recent data items more heavily. The inclusion probabilities of data items decay exponentially over time. The authors introduce a reservoir-based temporally-biased sampling method that asserts an upper bound on the sample size while keeping the decay of the sample predictable. The problem is well motivated and described, and the paper offers an excellent solution that is formalized precisely, is robust in the presence of evolving data, and has been implemented and evaluated for a distributed setting.

The best paper runner-up award was bestowed upon the paper:

## **GeoAlign: Interpolating Aggregates over Unaligned Partitions**

by Jie Song from the University of Michigan, Danai Koutra from the University of Michigan, Murali Mani from the University of Michigan, Flint, and H. Jagadish from the University of Michigan. DOI: 10.5441/002/edbt.2018.32

This paper introduces a novel technique to integrate geographical summaries over unaligned geographical regions, e.g., counties and ZIP codes. While traditional techniques assume that the data in each region is uniformly distributed, the proposed approach infers the distribution based on other datasets. The proposed idea is novel, refreshing, and nicely motivated. The described solutions are practical, have been implemented and evaluated, and there is good potential for follow-up work.

The EDBT 2018 Best Paper Award Committee consisted of Michael Böhlen, Lei Chen, Wolfgang Lehner, and Kian-Lee Tan. The EDBT Best Paper Awards for 2018 will be presented during the EDBT/ICDT 2018 Conference, March 26-29, in Vienna, Austria (<http://edbticdt2018.at>).

# Table of Contents

Foreword	i
Program Committee Members	ii
Test-of-Time Award	iv
Best Paper Award	v
Table of Contents	vi
<b>Research Papers</b>	
ID Repair for Trajectories with Transition Graphs <i>Xingcan Cui, Xiaohui Yu, Xiaofang Zhou, Jiong Guo</i>	1
MTBase: Optimizing Cross-Tenant Database Queries <i>Lucas Braun, Renato Marroquin, Ken Tsay, Donald Kossmann</i>	13
Extending In-Memory Relational Database Engines with Native Graph Support <i>Mohamed Hassan, Tatiana Kuznetsova, Hyun Chai Jeong, Walid Aref, Mohammad Sadoghi</i>	25
Sequenced Route Query with Semantic Hierarchy <i>Yuya Sasaki, Yoshiharu Ishikawa, Yasuhiro Fujiwara, Makoto Onizuka</i>	37
On Complexity and Efficiency of Mutual Information Estimation on Static and Dynamic Data <i>Michael Vollmer, Ignaz Rutter, Klemens Böhm</i>	49
Finding All Maximal Connected s-Cliques in Social Networks <i>Rachel Behar, Sara Cohen</i>	61
Summarization Algorithms for Record Linkage <i>Dimitrios Karapiperis, Aris Gkoulalas-Divanis, Vassilios S. Verykios</i>	73
Continuous Monitoring of Pareto Frontiers on Partially Ordered Attributes for Many Users <i>Afroza Sultana, Chengkai Li</i>	85
Optimizing Selection Processing for Encrypted Database using Past Result Knowledge Base <i>Wai Kit Wong, Kwok Wai Wong, Ho-Yin Yue</i>	97
Temporally-Biased Sampling for Online Model Management <i>Brian Hentschel, Peter J. Haas, Yuanyuan Tian</i>	109
Detecting Database File Tampering through Page Carving <i>James Wagner, Alexander Rasin, Tanu Malik, Karen Heart, Jacob Furst, Jonathan Grier</i>	121
User-guided Repairing of Inconsistent Knowledge Bases <i>Abdallah Arioua, Angela Bonifati</i>	133
Synchronous Multi-GPU Training for Deep Learning with Low-Precision Communications: An Empirical Study <i>Demjan Grubic, Leo Tam, Dan Alistarh, Ce Zhang</i>	145
EasyCommit: A Non-blocking Two-phase Commit Protocol <i>Suyash Gupta, Mohammad Sadoghi</i>	157
Beyond Frequencies: Graph Pattern Mining in Multi-weighted Graphs <i>Giulia Preti, Matteo Lissandrini, Davide Mottin, Yannis Velegrakis</i>	169
Scalable Evaluation of k-NN Queries on Large Uncertain Graphs <i>Xiaodong Li, Reynold Cheng, Yixiang Fang, Jiafeng Hu, Silviu Maniu</i>	181



MatchCatcher: A Debugger for Blocking in Entity Matching <i>Han Li, Pradap Konda, Paul Suganthan G C, Anhai Doan, Benjamin Snyder, Youngchoon Park, Ganesh Krishnan, Rohit Deep, Vijay Raghavendra</i> . . . . .	193
Extracting Statistical Graph Features for Accurate and Efficient Time Series Classification <i>Daoyuan Li, Jessica Lin, Tegawendé Bissyandé, Jacques Klein, Yves Le Traon</i> . . . . .	205
Counting Edges with Target Labels in Online Social Networks via Random Walk <i>Yang Wu, Cheng Long, Ada Fu, Zitong Chen</i> . . . . .	217
An Homophily-based Approach for Fast Post Recommendation on Twitter <i>Quentin Grossetti, Camelia Constantin, Cedric du Mouza, Nicolas Travers</i> . . . . .	229
Online Set Selection with Fairness and Diversity Constraints <i>Julia Stoyanovich, Ke Yang, H. Jagadish</i> . . . . .	241
Apollo: Learning Query Correlations for Predictive Caching in Geo-Distributed Systems <i>Brad Glasbergen, Michael Abebe, Khuzaima Daudjee, Scott Foggo, Anil Pacaci</i> . . . . .	253
Interactive Rule Refinement for Fraud Detection <i>Tova Milo, Slava Novgorodov, Wang-Chiew Tan</i> . . . . .	265
Privacy Preserving Group Nearest Neighbor Search <i>Yuncheng Wu, Ke Wang, Zhilin Zhang, weipeng lin, Hong Chen, Cuiping Li</i> . . . . .	277
Pattern Search in Temporal Social Networks <i>Andreas Züfle, Matthias Renz, Tobias Emrich, Maximilian Franzke</i> . . . . .	289
Scalable and Dynamic Regeneration of Big Data Volumes <i>Anupam Sanghi, Raghav Sood, Jayant Haritsa, Srikanta Tirthapura</i> . . . . .	301
TPStream: Low-Latency Temporal Pattern Matching on Event Streams <i>Michael Körber, Nikolaus Glombiewski, Bernhard Seeger</i> . . . . .	313
QUASII: QUery-Aware Spatial Incremental Index <i>Mirjana Pavlovic, Darius Sidlauskas, Thomas Heinis, Anastasia Ailamaki</i> . . . . .	325
Loom: Query-aware Partitioning of Online Graphs <i>Hugo Firth, Paolo Missier, Jack Aiston</i> . . . . .	337
Kernel-Based Cardinality Estimation on Metric Data <i>Michael Mattig, Thomas Fober, Christian Beilschmidt, Bernhard Seeger</i> . . . . .	349
GeoAlign: Interpolating Aggregates over Unaligned Partitions <i>Jie Song, Danai Koutra, Murali Mani, H. Jagadish</i> . . . . .	361
Distributed query-aware quantization for high-dimensional similarity searches <i>Gheorghii Guzun, Guadalupe Canahuate</i> . . . . .	373
Global-Scale Placement of Transactional Data Stores <i>Victor Zakhary, Faisal Nawab, Divy Agrawal, Amr El Abbadi</i> . . . . .	385
SlickDeque: High Throughput and Low Latency Incremental Sliding-Window Aggregation <i>Anatoli Shein, Panos Chrysanthis, Alexandros Labrinidis</i> . . . . .	397
Modeling and Exploiting Goal and Action Associations for Recommendations <i>Dimitra Papadimitriou, Yannis Velegrakis, Georgia Koutrika</i> . . . . .	409
<b>Short Papers</b>	
Very-Low Random Projection Maps <i>Anastasios Zouzias, Michail Vlachos</i> . . . . .	421

Interval Count Semi-Joins <i>Panagiotis Boursos, Nikos Mamoulis</i> . . . . .	425
Notable Characteristics Search through Knowledge Graphs <i>Davide Mottin, Bastian Grasnck, Axel Kroschk, Patrick Siegler, Emmanuel Müller</i> . . . . .	429
EmbedS: Scalable, Ontology-aware Graph Embeddings <i>Gonzalo Diaz, Achille Fokoue, Mohammad Sadoghi</i> . . . . .	433
All that Incremental is not Efficient: Towards Recomputation Based Complex Event Processing for Expensive Queries <i>Abderrahmen Kammoun, Syed Gillani, Julien Subercaze, Stephane Frenot, Kamal Singh, Frederique Laforest, Jacques Fayolle</i> . . . . .	437
DeepEye: Visualizing Your Data by Keyword Search <i>xuedi qin, Yuyu Luo, Nan Tang, Guoliang Li</i> . . . . .	441
Research Directions in Blockchain Data Management and Analytics <i>Hoang Tam Vo, Ashish Kundu, Mukesh Mohania</i> . . . . .	445
Scalable Active Temporal Constrained Clustering <i>Son Mai, Sihem Amer-Yahia, Ahlame Douzal Chouakria</i> . . . . .	449
Global Range Encoding for Efficient Partition Elimination <i>Jeremy Chen, Reza Sherkat, Mihnea ANDREI, Heiko Gerwens</i> . . . . .	453
NoFTL-KV: Tackling Write-Amplification on KV-Stores with Native Storage Management <i>Tobias Vincon, Sergej Hardock, Christian Riegger, Julian Oppermann, Andreas Koch, Ilia Petrov</i> . . . . .	457
Towards Hypothetical Reasoning Using Distributed Provenance <i>Daniel Deutch, Yuval Moskovitch, Itay Polak, Noam Rinetzky</i> . . . . .	461
On Answering Why-Not Queries Against Scientific Workflow Provenance <i>Khalid Belhajjame</i> . . . . .	465
PRoST: Distributed Execution of SPARQL Queries Using Mixed Partitioning Strategies <i>Matteo Cossu, Michael Färber, Georg Lausen</i> . . . . .	469
Deep Integration of Machine Learning Into Column Stores <i>Mark Raasveldt, Pedro Holanda, Hannes Mühleisen, Stefan Manegold</i> . . . . .	473
Scalable Detection of Concept Drifts on Data Streams with Parallel Adaptive Windowing <i>Philipp Marian Grulich, Rene Saitenmacher, Jonas Traub, Sebastian Breß, Tilmann Rabl, Volker Markl</i> . . . . .	477
Point-of-Interest Recommendation Using Heterogeneous Link Prediction <i>Alireza Pourali, Fattane Zarrinkalam, Ebrahim Bagheri</i> . . . . .	481
MetisIDX - From Adaptive to Predictive Data Indexing <i>Elvis Teixeira, Paulo Amora, Javam Machado</i> . . . . .	485
Efficient SIMD Vectorization for Hashing in OpenCL <i>Tobias Behrens, Viktor Rosenfeld, Jonas Traub, Sebastian Breß, Volker Markl</i> . . . . .	489
Histogram Domain Ordering for Path Selectivity Estimation <i>Nikolay Yakovets, Li Wang, George Fletcher, Craig Taverner, Alexandra Poulouvassilis</i> . . . . .	493
Nomadic Datacenters at the Network Edge: Data Management Challenges for the Cloud with Mobile Infrastructure <i>Faisal Nawab, Divy Agrawal, Amr El Abbadi</i> . . . . .	497
Dynamic Resource Routing using Real-Time Information <i>Sebastian Schmoll, Matthias Schubert</i> . . . . .	501

Data Structures for Efficient Computation of Influence Maximization and Influence Estimation <i>Diana Popova, Akshay Khot, Alex Thomo</i> . . . . .	505
A Roadmap towards Declarative Similarity Queries <i>Nikolaus Augsten</i> . . . . .	509
<b>Tutorials</b>	
Interactive Exploration of Composite Items <i>Sihem Amer-Yahia, Senjuti Basu Roy</i> . . . . .	513
Recent Advances in Recommender Systems: Matrices, Bandits, and Blenders <i>Georgia Koutrika</i> . . . . .	517
openCypher: New Directions in Property Graph Querying <i>Alastair Green, Martin Junghanns, Max Kiessling, Tobias Lindaaker, Stefan Plantikow, Petra Selmer</i> . . . . .	520
Real-Time Data Management for Big Data <i>Wolfram Wingerath, Felix Gessert, Erik Witt, Steffen Friedrich, Norbert Ritter</i> . . . . .	524
<b>Industrial and Applications Papers</b>	
Supporting Similarity Queries in Apache AsterixDB <i>Taewoo Kim, Wenhai Li, Alexander Behm, Inci Cetindil, Rares Vernica, Vinayak Borkar, Michael Carey, Chen Li</i> . . . . .	528
L-Store: A Real-time OLTP and OLAP System <i>Mohammad Sadoghi, Souvik Bhattacharjee, Bishwaranjan Bhattacharjee, Mustafa Canim</i> . . . . .	540
A Hybrid Approach for Alarm Verification using Stream Processing, Machine Learning and Text Analytics <i>Ana Sima, Kurt Stockinger, Katrin Affolter, Martin Braschler, Peter Monte, Lukas Kaiser</i> . . . . .	552
Efficient Secure k-Nearest Neighbours over Encrypted Data <i>Manish Kesarwani, Akshar Kaul, Prasad Naldurg, Sikhar Patranabis, Gagandeep Singh, sameep mehta, Debdeep Mukhopadhyay</i> . . . . .	564
A Parallel and Scalable Processor for JSON Data <i>Christina Pavlopoulou, E. Preston Carman Jr, Till Westmann, Michael Carey, Vassilis Tsotras</i> . . . . .	576
An Automated System for Internet Pharmacy Verification <i>Alberto Cordioli, Themis Palpanas</i> . . . . .	588
RQL: Retrospective Computations over Snapshot Sets <i>Nikos Tsikoudis, Liuba Shrira, Sara Cohen</i> . . . . .	600
Big Data Analytics for Time Critical Mobility Forecasting: Recent Progress and Research Challenges <i>George Vouros, Akrivi Vlachou, Giorgos Santipantakis, Christos Doulkeridis, Nikos Pelekis, Harris Georgiou, Yannis Theodoridis, Kostas Patroumpas, Elias Alevizos, Alexander Artikis, Christophe Claramunt, Cyril Ray, David Scarlatti, Georg Fuchs, Gennady Andrienko, Natalia Andrienko, Michael Mock, Elena Camossi, Anne-Laure Jous-selme, Jose Manuel Cordero Garcia</i> . . . . .	612
Scouter: A Stream Processing Web Analyzer to Contextualize Singularities <i>Badre Belabess, Musab Bairat, Jeeremy Lhez, Zakaria Khattabi, Yufan Zheng, Olivier CURE</i> . . . . .	624
Finding Contrast Patterns for Mixed Streaming Data (Application track) <i>Rohan Khade, Jessica Lin, Nital Patel</i> . . . . .	632
<b>Demonstrations</b>	
Recalibration of Analytics Workflows <i>Verena Kantere, Maxim Filatov, Maxim Filatov, Vasiliki Kantere, Verena Kantere</i> . . . . .	642

Effective Quality Assurance for Data Labels through Crowdsourcing and Domain Expert Collaboration <i>Wei Lee, Chien-Wei Chang, Po-An Yang, Chi-Hsuan Huang, Ming-Kuang Wu, Chu-Cheng Hsieh, Kun-Ta Chuang</i>	646
Exploring Large Scholarly Networks with Hermes <i>Gabriel Campero Durand, Anusha Janardhana, Marcus Pinnecke, Yusra Shakeel, Jacob Krüger, Thomas Leich, Gunter Saake</i>	650
Don't write all data pages in one stream <i>Soyee Choi, Hyunwoo Park, Sang Won Lee</i>	654
eLinda: Explorer for Linked Data <i>Tal Yahav, Oren Kalinsky, Oren Mishali, Benny Kimelfeld</i>	658
FAIMUSS: Flexible Data Transformation to RDF from Multiple Streaming Sources <i>Giorgos Santipantakis, Apostolos Glenis, Nikolaos Kalaitzian, Akrivi Vlachou, Christos Doulkeridis, George Vouros</i>	662
SAMUEL: A Sharing-based Approach to processing Multiple SPARQL Queries with MapReduce <i>InA Kim, Kyong-Ha Lee, Kyuchul Lee</i>	666
GEDetector: Early Detection of Gathering Events Based on Cluster Containment Join in Trajectory Streams <i>Bin Zhao, Genlin Ji, Yu Yang, Zhaoyuan Yu, Xintao Liu, Ningfang Mi</i>	670
Reconciling Privacy and Data Sharing in a Smart and Connected Surrounding <i>Paul Tran-Van, Nicolas Ancaux, Philippe Pucheral</i>	674
Spatio-Temporal-Keyword Pattern Queries over Semantic Trajectories with Hermes@Neo4j <i>Fragkiskos Gryllakis, Nikos Pelekis, Christos Doulkeridis, Stylianos Sideridis, Yannis Theodoridis</i>	678
MDM: Governing Evolution in Big Data Ecosystems <i>Sergi Nadal, Alberto Abelló, Oscar Romero, Stijn Vansummeren, Panos Vassiliadis</i>	682
Provenance-Based Visual Data Exploration with EVLIN <i>Housseem BEN LAHMAR, Melanie Herschel, Michael Blumenschein, Daniel Keim</i>	686
Interactive Visualization of Large Similarity Graphs and Entity Resolution Clusters <i>M. Ali Rostami, Alieh Saeedi, Eric Peukert, Erhard Rahm</i>	690
FastOFD: Contextual Data Cleaning with Ontology Functional Dependencies <i>Zheng Zheng, Morteza Alipour, Zhi Qu, Ian Currie, Fei Chiang, Lukasz Golab, Jaroslaw Szlichta</i>	694
Analysis and Visualization of Urban Emission Measurements in Smart Cities <i>Dirk Ahlers, Frank Kraemer, Anders Braten, Xiufeng Liu, Fredrik Anthonisen, Patrick Driscoll, John Krogstie</i>	698
Pharos: Privacy Hazards of Replicating ORAM Stores <i>Victor Zakhary, Cetin Sahin, Amr El Abbadi, Huijia Lin, Stefano Tessaro</i>	702