



Proceedings of the VLDB Endowment

Volume 12, No. 13 – September 2019

Editors in Chief:

Lei Chen and Fatma Özcan

Associate Editors:

Azza Abouzied, Selcuk Candan, Surajit Chaudhuri, Amol Desphande, Johann-Christoph Freytag, Rainer Gemulla, Nick Koudas, Georgia Koutrika, Yunyao Li, Alexandra Meliou, Arnab Nandi, M. Tamer Özsu, Themis Palpanas, Alkis Polyzotis, Kyuseok Shim, Xiaokui Xiao, Meihui Zhang

Publication Editors:

Abdul Quamar, Yongxin Tong

PVLDB – Proceedings of the VLDB Endowment

Volume 12, No. 13, September 2019.

All papers published in this issue will be presented at the 46th International Conference on Very Large Data Bases, Tokyo, Japan, 2020.

Copyright 2019 VLDB Endowment

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>. For any use beyond those covered by this license, obtain permission by emailing info@vldb.org.

Volume 12, Number 13, September 2019

Pages i – viii and 2325 - 2421

ISSN 2150-8097

Available at: <http://www.pvldb.org> and <https://dl.acm.org>.

TABLE OF CONTENTS

Front Matter

Copyright Notice	i
Table of Contents	ii
PVLDB Organization and Review Board – Vol. 12	iii
External Reviewers - Vol. 12	vi

Research Papers

Strong consistency is not hard to get: Two-Phase Locking and Two-Phase Commit on Thousands of Cores.....	2325
<i>Claude Barthels, Ingo Müller, Konstantin Taranov, Gustavo Alonso, Torsten Hoefler</i>	
Discovery and Ranking of Embedded Uniqueness Constraints.....	2339
<i>Ziheng Wei, Uwe Leck, Sebastian Link</i>	
Online Density Bursting Subgraph Detection from Temporal Graphs	2353
<i>Lingyang Chu, Yanyan Zhang, Yu Yang, Lanjun Wang, Jian Pei</i>	
Progressive Indexes: Indexing for Interactive Data Analysis	2366
<i>Pedro Holanda, Stefan Manegold, Hannes Mühleisen, Mark Raasveldt</i>	
Distributed Edge Partitioning for Trillion-edge Graphs	2379
<i>Masatoshi Hanai, Toyotaro Suzumura, Wen Jun Tan, Elvis Liu, Georgios Theodoropoulos, Wentong Cai</i>	
Optimal Column Layout for Hybrid Workloads.....	2393
<i>Manos Athanassoulis, Kenneth Bøgh, Stratos Idreos</i>	
Selecting Data to Clean for Fact Checking: Minimizing Uncertainty vs. Maximizing Surprise	2408
<i>Stavros Sintos, Pankaj Agarwal, Jun Yang</i>	

PVLDB ORGANIZATION AND REVIEW BOARD - Vol. 12

Editors in Chief of PVLDB

Lei Chen, HKUST
Fatma Özcan, IBM Research - Almaden

Associate Editors of PVLDB

Azza Abouzied, NYU Abu Dhabi
Selcuk Candan, Arizona State University
Surajit Chaudhuri, Microsoft Research
Amol Desphande, University of Maryland
Johann-Christoph Freytag, HU Berlin
Rainer Gemulla, University of Mannheim
Nick Koudas, University of Toronto
Georgia Koutrika, Athena Research Center
Yunyao Li, IBM Research - Almaden
Alexandra Meliou, University of Massachusetts
Arnab Nandi, Ohio State University
M. Tamer Özsu, University of Waterloo
Themis Palpanas, French University Institute
Alkis Polyzotis, Google
Kyuseok Shim, Seoul National University

Xiaokui Xiao, National University of Singapore
Meihui Zhang, Beijing Institute of Technology

Publication Editors

Abdul Quamar, IBM Research - Almaden
Yongxin Tong, Beihang University

PVLDB Managing Editor

Wolfgang Lehner, TU Dresden

PVLDB Advisory Committee

Sihem Amer-Yahia, CNRS
Peter Boncz, Centrum Wiskunde & Informatica
Xin Luna Dong, Amazon
Juliana Freire, New York University
Wolfgang Lehner, TU Dresden
Renée J. Miller, University of Toronto
Tova Milo, Tel Aviv University
M. Tamer Özsu, University of Waterloo
Divesh Srivastava, AT&T Labs-Research

Review Board

Abdul Quamar, IBM Research - Almaden
Ada Waichee Fu, Chinese University of Hong Kong
Ahmet Erdem Sariyuca, University at Buffalo
Alan Fekete, University of Sydney
Alkis Simitsis, Hewlett Packard Labs
Ambuj Singh, UC Santa Barbara
Andrew Pavlo, Carnegie Mellon University
Angela Bonifati, University of Lyon
Arijit Khan, Nanyang Technological University
Arnab Bhattacharya, IIT Kanpur
Arun Kumar, University of California, San Diego
Arvind Arasu, Microsoft Research
Ashraf Aboulnaga, QCRI
Ashwin Machanavajhala, Duke University
Avrilia Floratou, Microsoft Research
Azade Nazi, Microsoft Research
Badrish Chandramouli, Microsoft Research
Barzan Mozafari, University of Michigan
Beng Chin OOI, National University of Singapore
Berthold Reinwald, IBM Research - Almaden
Bin Cui, Peiking University
Bobbie Cochrane, IBM
Bolin Ding, Alibaba
Boris Glavic, Illinois Institute of Technology
Bugra Gedik, Bilkent University
Byron Choi, Hong Kong Baptist University
Carlo Curino, Microsoft Research
Chee-Yong Chan, National University of Singapore
Chen Li, University of California, Irvine
Chengkai Li, UT Arlington
Chuan Lei, IBM Research - Almaden
Cong Yu, Google
Curtis Dyreson, Utah State University
Danica Probic, Oracle
Daniel Kifer, Penn State University
Davide Mottin, Hasso-Plattner Institute
Demetrios Zeinalipour-Yazti, University of Cyprus
Dimitris Papadias, HKUST
Diptikalyan Saha, IBM Research - India
Divyakant Agrawal, UC Santa Barbara
Donald Kossmann, Microsoft Research
Egemen Tanin, University of Melbourne
Eser Kandogan, IBM Research - Almaden
Essam M. Mansour, QCRI
Fabio Porto, LNCC
Fei Chiang, McMaster University
Feifei Li, University of Utah
Florin Rusu, University of California, Merced
Floris Geerts, University of Antwerp
George Papadakis, University of Athens
Goetz Graefe, Google
Guoliang Li, Tsinghua University
H. V. Jagadish, University of Michigan
Hakan Ferhatosmanoglu, Bilkent University
Hakan Hacigumus, Google
Hanghang Tong, Arizona State University
Helen Huang, University of Queensland

Heng Tao Shen, UESTC
Hong Cheng, Chinese University of Hong Kong
Hongzhi Yin, University of Queensland
Hua Lu, Aalborg University
Huiping Cao, New Mexico State University
Ilaria Bartolini, University of Bologna
Ilkay Altintas, San Diego Supercomputing Center
Immanuel Trummer, Cornell University
Ioana Manulescu, INRIA
Ismail Sengor Altingovde, METU
James Cheng, Chinese University of Hong Kong
Jens Dittrich, University of Saarland
Jens Teubner, TU Dortmund
Jianliang Xu, Hong Kong Baptist University
Jignesh Patel, University of Wisconsin - Madison
Jinyang Gao, National University of Singapore
Johann Gamper, Free University of Bozen-Bolzano
Jun Yang, Duke University
Junjie Yao, East China Normal University
Kai Zheng, UESTC
Karthik Sankaranarayanan, IBM Research - India
Katja Hose, Aalborg University
Khuzaima Daudjee, University of Waterloo
Kostas Stefanidis, University of Tampere
Kostas Zoumpatianos, Harvard University
Letizia Tanca, Polytechnic University of Milan
Lucian Popa, IBM Research - Almaden
Luna Dong, Amazon
Manos Karpathiotakis, Facebook London
Maria Luisa Sapino, University of Torino
Mario Nascimento, University of Alberta
Martin Theobald, University of Luxemburg
Mary Roth, IBM Research - Almaden
Matthias Boehm, Graz University of Technology
Matthias Renz, George Mason University
Maya Ramanath, Indian Institute of Technology Delhi
Melanie Herschel, University of Stuttgart
Michael Böhlen, University of Zurich
Michael Hay, Colgate University
Michael Mathioudakis, University of Helsinki
Min Li, JD.com
Mirek Riedewald, Northeastern University
Mirella Moro, Universidade Federal de Minas Gerais
Mohamed Eltabakh, Worcester Polytechnic Institute
Mohamed Mokbel, QCRI
Mohamed Sarwat, Arizona State University
Murat Kantarcioglu, University of Texas at Dallas
Nan Tang, QCRI
Nicolas Ancaux, INRIA
Nikolaus Augsten, University of Salzburg
Oktie Hassanzadeh, IBM Research - Yorktown
Olga Papaemmanouil, Brandeis University
Paolo Papotti, EURECOM
Parth Nagarkar, New Mexico State University
Pelin Angin, Middle East Technical University
Philip Bernstein, Microsoft Research
Philippe Bonnet, IT University of Copenhagen
Pinar Karagoz, Middle East Technical University
Pinar Tozun, IT University of Copenhagen

Raymond Ng, University of British Columbia
Sai Wu, Zhejiang University
Sang Kyun Cha, Seoul National University
Sebastian Breß, DFKI and TU Berlin
Semih Salihoglu, University of Waterloo
Senjuti Basu Roy, New Jersey Institute of Technology
Seung-Won Hwang, Yonsei University
Shaoxu Song, Tsinghua University
Shuo Shang, IIAI
Spyros Blanas, Ohio State University
Stefan Mangeold, Centrum Wiskunde & Informatica
Stefano Paraboschi, University of Bergamo
Steffen Zeuch, DFKI and TU Berlin
Stratis Viglas, University of Edinburgh
Sudip Roy, Google
Tingjian Ge, University of Massachusetts Lowell
Tyson Condie, University of California, Los Angeles
Umar Farooq Minhas, Microsoft Research
Vijayshankar Raman, Google
Viktor Leis, University of Jena
Vincent Oria, New Jersey Institute of Technology
Vivek Narasayya, Microsoft Research
Wenjie Zhang, University of New South Wales
Wook-Shin Han, POSTECH

Xiang Lian, Kent State University
Xiangmin Zhou, RMIT
Xiaochun Yang, Northeastern University
Xiaofang Zhou, University of Queensland
Li Xiong, Emory University
Xu Chu, Georgia Institute of Technology
Xuemin Lin, University of New Southwales
Yael Amsterdamer, Bar-Ilan University
Yannis Velegarakis, Utrecht University
Yanyan Shen, Shanghai Jiao Tong University
Yi Chen, New Jersey Institute of Technology
Ying Zhang, University of Technology Sydney
Yinghui Wu, Washington State University
Yingjun Wu, IBM Research - Almaden
Yingxia Shao, Peking University
Yongxin Tong, Beihang University
Yoshiharu Ishikawa, Nagoay University
Ye Yuan, Northeastern University
Yuanyuan Tian, IBM Research - Almaden
Yucel Saygin, Sabanci University
Yunjun Gao, Zhejiang University
Zhiguo Gong, University of Macau

EXTERNAL REVIEWERS - Vol. 12

Jianwen Zhao, The Chinese Univeristy of Hong Kong
Georgia Koutrika, Athena Research Institute
Sourav Medya, UC, Santa Barbara
Arlei Silva, UC, Santa Barbara
Wei Ye, UC, Santa Barbara
Alex Jones, UC, Santa Barbara
Ios Kotsogiannis, Duke University
Nisarg Raval, Duke University
Navneet Potti, UW-Madison
Sheng Wang, National University of Singapore
Wei Wang, National University of Singapore
Qian Lin, National University of Singapore
Guoliang Li, Tsinghua University
Anh Dinh, National University of Singapore
Daniel Ford, IBM
Rathijit Sen, Microsoft
Li Meiyang, National University of Singapore
TaiNing Wang, National University of Singapore
Zuozhi Wang, UC, Irvine
Avinash Kumar, UC, Irvine
Taewoo Kim, UC, Irvine
Jianfeng Jia, Google
Sadeem Saleh, UC, Irvine
Qiushi Bai, UC, Irvine
Sami Saeef, University of Texas at Arlington
Kai Yao, HKUST
Dimitris Tsaras, HKUST
Dimitris Tsakalidis, HKUST
Kai Yao, HKUST
Panos Simatis, HKUST
Mohammad Javad Amiri, UC, Santa Barbara
Sujaya Maiya, UC, Santa Barbara
Mohammad Javad Amiri, UC, Santa Barbara
Xianghai Sheng, UC, Merced
Weijie Zhao, UC, Merced
Stratis Viglas, Google
Fuat Basik, Bilkent University
Chris Conlan, University of Warwick
Teddy Cunningham, University of Warwick
Su Jiao, The Chinese University of Hong Kong
Lu Can, The Chinese University of Hong Kong
Marco Patella, DISI, University of Bologna
Zhi Liu, The Chinese University of Hong Kong
Yidi Wu, The Chinese University of Hong Kong
Xiao Yan, The Chinese University of Hong Kong
Xinyan Dai, The Chinese University of Hong Kong
Jian Zhang, The Chinese University of Hong Kong
Yunjian Zhao, The Chinese University of Hong Kong
Felix Martin Schuhknecht, Saarland University
Ankur Sharma, Saarland University
Cheng Xu, Hong Kong Baptist University
Zuyu Zhang, UW-Madison
Stavros Sintos, Duke University
Michael Abebe, University of Waterloo
Brad Glasbergen, University of Waterloo
Fabio Azzalini, Politecnico di Milano
Davide Azzalini, Politecnico di Milano
Daniel Garcia Ulloa, Keysight Technologies
Yang Cao, Kyoto University
Mark Ma, Emory University
Camila Costa, University of Alberta
Vinu Venugopal, University of Luxemburg
Neha Sengupta, IIT-Delhi
Srikanta Bedathur, IIT-Delhi
Sarah Oppold, University of Stuttgart
Kevin Wellenzohn, University of Zurich
Huasong Shan, JD.com American Technologies Corporation
Cuney Akcora, University of Texas at Dallas
Imrul Anindya, University of Texas at Dallas
Mustafa Ozdayi, University of Texas at Dallas
Aref Asvashirehjini, University of Texas at Dallas
Shivaram Venkataraman, Microsoft Research
Todd Porter, Microsoft
Bailu Ding, Microsoft Research
Bonaventura Del Monte, DFKI GmbH
Jonas Traub, TU Berlin
Pedro Holanda, CWI
Diego Tome, CWI
Tatiana Jin, The Chinese University of Hong Kong
David Lomet, Microsoft Research
Anshuman Dutt, Microsoft Research
Weilong Ren, Kent State University
Ahmed Al-Baghdadi, Kent State University
Niranjan Rai, Kent State University
Wen Hua, University of Queensland
Junhao Gan, University of Queensland
Lei Li, University of Queensland
Fengmei Jin, Renmin University of China
Pingfu Chao, The University of Queensland
Sen Wang, The University of Queensland
Longbin Lai, The University of New South Wales
Zhengyi Yang, The University of New South Wales
Boge Liu, The University of New South Wales
Fan Zhang, The University of New South Wales
Dong Wen, The University of Technology Sydney
Hanchen Wang, The University of New South Wales
Yixiang Fang, The University of New South Wales
Jingzhi Fang, Beihang University
Yingjun Wu, IBM Research - Almaden
Mehmet Ercan Nergiz, Sabanci University
Jingjing Li, UESTC
Wei Wang, The University of New South Wales
Bingsheng He, National University of Singapore
Prithviraj Sen, IBM Research - Almaden
Kostas Kollias, Google Research
Peter Bailis, Stanford University
Lingfei Wu, IBM Research - Yorktown
Niketani Pansare, Apple
Sergey Gorbunov, University of Waterloo

LETTER FROM THE EDITORS IN CHIEF

The Proceedings of the VLDB Endowment (PVLDB) provides a high-quality journal publication service to the data management research community. Each volume offers twelve monthly submission deadlines on the first day of each month and a quick, six weeks, reviewing cycle. This publication model was pioneered by PVLDB and combines a journal-style reviewing process, which includes a three-month revision cycle, with the agility and visibility provided by rapid on-line publication, and presentation at the annual VLDB conference.

PVLDB attracts many submissions spanning diverse data management topics, and the PVLDB reviewing process is implemented by a large team of dedicated researchers. The Review Board of PVLDB Volume 12 consists of 166 expert researchers, and reviewing is coordinated by 17 Associate Editors. Review Board members provide timely (within a 4-week deadline) high-quality reviews, and participate actively in online discussions led by the Associate Editors for each paper.

This is the thirteenth and the final issue of the twelfth volume of the PVLDB. There are seven papers accepted in this issue that will be presented at the 46th International Conference on Very Large Data Bases (VLDB 2020), to be held in Tokyo, Japan during August 31 to September 4, 2020.

For the thirteenth issue of PVLDB Volume 12, the review board has selected contributions on advanced analytics topics, including edge partitioning of large graphs, density bursting subgraph detection on temporal graphs, data selection for fact checking and discovery of uniqueness constraints in data profiling. This issue also contains papers on advanced data management such as strong consistency in distributed transactions using 2-phase commit, new indexes for interactive data exploration and data layouts for hybrid transactional and analytical workloads. We hope that the readers will find the selected papers engaging, and though provoking. We also hope that the selected papers will provide valuable insights and inspire novel systems and algorithmic contributions and follow-up research.

In this last issue of volume 12, we would like to thank all the reviewers and the associate editors for their diligent work in selecting high quality publications for PVLDB.

Lei Chen and Fatma Özcan

PVLDB Volume 12 Editors in Chief