

VLDB2014

40th International Conference on Very Large Data Bases, Hangzhou, China



Proceedings of the VLDB Endowment

Volume 7, No. 10 – June 2014

**Proceedings of the 40th International Conference on
Very Large Data Bases, Hangzhou, China**

Program Chairs and Editors-in-Chief:

H. V. Jagadish, Aoying Zhou

Associate Editors – Research and Innovative Systems Tracks:

**Shivnath Babu, Lei Chen, Graham Cormode, Bin Cui, Wynne Hsu, Martin Kersten,
Donald Kossman, Elke Rundensteiner, Kyuseok Shim, Wang-Chiew Tan, Letizia Tanca, Jeffrey Yu**

Associate Editors – Experiments and Analysis Track:

Gao Cong, Jens Dittrich

Associate Editors – Vision Track:

Zachary Ives

Proceedings Chairs:

Li Xiong, Cong Yu

PVLDB – Proceedings of the VLDB Endowment

Volume 7, No. 10, June 2014.

The 40th International Conference on Very Large Data Bases, Hangzhou, China.

Copyright 2014 VLDB Endowment

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>. Obtain permission prior to any use beyond those covered by the license. Contact copyright holder by emailing info@vldb.org.

Volume 7, Number 10, June 2014: VLDB 2014

Pages ii - xi and 785 - 930

ISSN 2150-8097

Additional copies only online at: portal.acm.org, arxiv.org/corr, and www.vldb.org

TABLE OF CONTENTS

Front Matter

Copyright Notice	ii
Table of Contents	iii
VLDB 2014 Organization and Review Board	v

Letters

Letter from VLDB 2014 General Chair	<i>Sharad Mehrotra</i> x
---	--------------------------

Research Papers

Reverse k-Ranks Query.....	785
..... <i>Zhao Zhang, Cheqing Jin, Qiangqiang Kang</i>	
M4: A Visualization-Oriented Time Series Data Aggregation	797
..... <i>Uwe Jugel, Zbigniew Jerzak, Gregor Hackenbroich, Volker Markl</i>	
Continuous Matrix Approximation on Distributed Data	809
..... <i>Mina Ghashami, Jeff M. Phillips, Feifei Li</i>	
An Evaluation of the Advantages and Disadvantages of Deterministic Database Systems	821
..... <i>Kun Ren, Alexander Thomson, Daniel J. Abadi</i>	
Efficient In-memory Data Management: An Analysis.....	833
..... <i>Hao Zhang, Bogdan Marius Tudor, Gang Chen, Beng Chin Ooi</i>	
Workload Matters: Why RDF Databases Need a New Design	837
..... <i>Gunes Aluc, M. Tamer Ozsü, Khuzaima Daudjee</i>	
Storage Management in AsterixDB	841
..... <i>Sattam Alsubaiee, Alexander Behm, Vinayak Borkar, Zachary Heilbron, Young-Seok Kim</i>	
..... <i>Michael J. Carey, Markus Dreseler, Chen Li</i>	
Building Efficient Query Engines in a High-Level Language.....	853
..... <i>Yannis Klonatos, Christoph Koch, Tiark Rompf, Hassan Chafi</i>	
Scalable Logging through Emerging Non-Volatile Memory	865
..... <i>Tianzheng Wang, Ryan Johnson</i>	
When Data Management Systems Meet Approximate Hardware: Challenges and Opportunities...	877
..... <i>Bingsheng He</i>	
From Data Fusion to Knowledge Fusion	881
..... <i>Xin Luna Dong, Evgeniy Gabrilovich, Jeremy Heitz,</i>	
..... <i>Wilko Horn, Kevin Murphy, Shaohua Sun, Wei Zhang</i>	
On k-Path Covers and their Applications	893
..... <i>Stefan Funke, Andre Nusser, Sabine Storandt</i>	

The Case for Data Visualization Management Systems.....	903
..... <i>Eugene Wu, Leilani Battle, Samuel R. Madden</i>	
WideTable: An Accelerator for Analytical Data Processing	907
..... <i>Yinan Li, Jignesh M. Patel</i>	
A Framework for Protecting Worker Location Privacy in Spatial Crowdsourcing.....	919
..... <i>Hien To, Gabriel Ghinita, Cyrus Shahabi</i>	

VLDB 2014 ORGANIZATION AND REVIEW BOARD

Honorary Chair

Yunhe Pan, Chinese Academy of Engineering

General Chairs

Chun Chen, Zhejiang University

Sharad Mehrotra, University of California, Irvine

Program Chairs and Editors-in-Chief of PVLDB 7

H. V. Jagadish, University of Michigan

Aoying Zhou, East Normal University, China

Research and Innovative Systems Tracks Associate Editors

Shivnath Babu, Duke University

Lei Chen, Hong Kong University of Science and Technology

Graham Cormode, University of Warwick

Bin Cui, Peking University, China

Wynne Hsu, NUS

Martin Kersten, CWI

Donald Kossman, ETH

Elke Rundensteiner, WPI

Kyuseok Shim, Seoul National University

Wang-Chiew Tan, University of California, Santa Cruz

Letizia Tanca, Poli Milano

Jeffrey Yu, Chinese University of Hong Kong

Experiments and Analysis Track Associate Editors

Gao Cong, Nanyang Technology University

Jens Dittrich, Saarland

Visions Track Associate Editor

Zachary Ives, University of Pennsylvania

Industrial and Applications Track Associate Editors

Umeshwar Dayal, HP

C. Mohan, IBM

Ge Yu, Northeastern University, China

Demonstration Chairs

Mong-Li Lee, NUS
Feifei Li, University of Utah
Sunil Prabhakar, Purdue

Tutorial Chairs

Xiaoyong Du, Renmin University
Murat Kantarcioglu, University of Texas, Dallas
Divesh Srivastava, AT&T Labs

Research Track Review Board

Sibel Adali, Rensselaer Polytechnic Institute
Foto Afrati, NTU Athens
Yanif Ahmad, JHU
Jose Luis Ambite, ISI - USC
Walid Aref, Purdue University
Claudia Bauzer Medeiros, University of Campinas
Srikanta Bedathur, IIIT Delhi
Michael Benedikt, Oxford University
Sonia Bergamaschi, Universita Modena
Laure Berti-Equille, IRD, France
Leopoldo Bertossi, Carleton University, Ottawa
Subhash Bhalla, University of Aizu, Japan
Peter Boncz, CWI
Angela Bonifati, University of Lille 1
Rajesh Bordawekar, IBM Watson Research Center
Omar Boucelma, Aix-Marseille University
Nico Bruno, Microsoft Research
Andrea Cali, University of London, Birkbeck College
Malu Castellanos, HP Labs
Badrish Chandramouli, Microsoft Research
Adriane Chapman, Mitre
Gang Chen, Zhejiang University
Yi Chen, New Jersey Institute of Technology
James Cheng, CUHK
Reynold Cheng, University of Hong Kong

Workshop Chairs

Anastasia Ailamaki, EPFL
Kaushik Chakrabarti, Microsoft

Panel Chairs

Hakan Hacigumus, NEC Labs
Jignesh Patel, University of Wisconsin
Xiaoyang Sean Wang, Fudan University

Brian Cooper, Google, USA
Philippe Cudré-Mauroux, University of Fribourg
Carlo Curino, MIT
Gautam Das, UT Arlington and QCRI
Sudipto Das, Microsoft Research
Anish Das Sarma
Atish Das Sarma, eBay Research Labs
Khuzaima Daudjee, University of Waterloo
Antonios Deligiannakis, Technical University of Crete
Daniel Deutch, Ben Gurion University
Yanlei Diao, University of Massachusetts Amherst
Xin (Luna) Dong, Google, USA
Sameh Elnikety, Microsoft Research
Mohamed Eltabakh, Worcester Polytechnic Institute
Ihab F. Ilyas, QCRI
Hakan Ferhatosmanoglu, Bilkent University
Ada Wai-Chee Fu, Chinese University of Hong Kong
Minos Garofalakis, Technical University of Crete
Wolfgang Gatterbauer, Carnegie Mellon University
Tingjian Ge, University of Massachusetts, Lowell
Buğra Gedik, Bilkent University
Rainer Gemulla, Max-Planck-Institut Saarbrücken
Gabriel Ghinita, University of Massachusetts Boston
Parke Godfrey, York University
Lukasz Golab, University of Waterloo

Sergio Greco, University of Calabria
Le Gruenwald, University of Oklahoma
Giovanna Guerrini, Università Genova
Krishna Gummadi, MPI-SWS
Rahul Gupta, Google Research
Rajeev Gupta, IBM Research
Shyam Gupta, IIT Delhi
Marios Hadjieleftheriou, AT&T labs
Wook-Shin Han, KNU, Korea
Kuno Harumi, HP Labs
Bingsheng He, NTU Singapore
Sven Helmer, Free University of Bozen-Bolzano
Jan Hidders, TUDelft
Wei Hong, Cisco System Inc.
Katja Hose, Aalborg University
Zi Huang, University of Queensland
Jeong-Hyon Hwang, SUNY - Albany
Seung-won Hwang, POSTECH, Korea
Stratos Idreos, CWI
Yoshiharu Ishikawa, Nagoya University
Zachary Ives, University of Pennsylvania
Ricardo Jimenez-Peris, Technical University of Madrid
Cheqing Jin, East China Normal University
Ruoming Jin, Kent State University
Alekh Jindal, Saarland University/MIT
Ryan Johnson, University of Toronto
Dmitri V Kalashnikov, UC Irvine
Panos Kalnis, KAUST, Saudi Arabia
Ben Kao, Hong Kong University
Panagiotis Karras, Rutgers University
Yiping Ke, Institute of High Performance Computing
Bettina Kemme, McGill University
Daniel Kifer, PSU
Benny Kimelfeld, IBM
Hideaki Kimura, Microsoft Jim Gray Systems Lab
George Kollios, Boston University
Christian König, Microsoft Research
Tim Kraska, Brown University

Laks V. S. Lakshmanan, University of British Columbia
Mounia Lalmas, Yahoo Inc.
Mong-Li Lee, National University of Singapore
Wolfgang Lehner, Technische University Dresden
Justin Levandoski, Microsoft Research
Chengkai Li, The University of Texas at Arlington
Cuiping Li, Renmin University of China
Feifei Li, University of Utah
Guoliang Li, Tsinghua University
Jianzhong Li, Harbin Institute of Technology
Yunyao Li, IBM Almaden
Zhanhuai Li, Northwestern Polytechnical University
Dan Lin, Missouri S&T, USA
Xuemin Lin, University of New South Wales
Bin Liu, NEC Labs America
Ziyang Liu, NEC Labs America
Eric Lo, The Hong Kong Polytechnic University
Qiong Luo, HKUST
Shuai Ma, Beihang University
Ashwin Machanavajjhala, Duke University
Brad Malin, Duke University
Nikos Mamoulis, University of Hong Kong
Stefan Manegold, CWI
Murali Mani, University of Michigan
Ioana Manolescu, INRIA, France
Amélie Marian, Rutgers University
Volker Markl, TU Berlin
Marta Mattoso, Federal University of Rio de Janeiro
Frank McSherry, Microsoft
Alexandra Meliou, Umass Amherst
Marco Mesiti, University of Milano
Dan Miranker, The University of Texas at Austin
Mohamed Mokbel, University of Minnesota
Bongki Moon, Seoul National University
Yasuhiko Morimoto, Hiroshima University
Mirella Moro, Universidade Federal de Minas Gerais
Kyriakos Mouratidis, SMU, Singapore
Karin Murthy, IBM India

Arnab Nandi, Ohio State University
 Wolfgang Nejdl, University of Hannover
 Thomas Neumann, Technology University Munchen
 Boris Novikov, St Petersburg University
 Dan Olteanu, Oxford University
 Gultekin Ozsoyoglu, Case Western Reserve University
 Tamer Ozsü, University of Waterloo
 Esther Pacitti, University of Montpellier
 Themis Palpanas, University of Trento
 Ippokratis Pandis, IBM Almaden
 Stelios Paparizos, Microsoft Research
 Aditya Parameswaran, Stanford University
 Srinivasan Parthasarathy, The Ohio State University
 Jignesh Patel, University of Wisconsin
 Andrew Pavlo, Brown University
 Peter Pietzuch, Imperial College London
 Neoklis Polyzotis, University of California - Santa Cruz
 Cecilia M. Procopiuc, AT&T Labs
 Li Qian, University of Michigan
 Jorge Quiané-Ruiz, QCRI
 Elisa Quintarelli, Politecnico di Milano
 Maya Ramanath, IIT Delhi
 Louiqa Raschid, University of Maryland
 Vibhar Rastogi, Yahoo
 Matthias Renz, University of Munich
 Kenneth Ross, Columbia University
 Sourav S Bhowmick, NTU, Singapore
 Dimitris Sacharidis, IMIS Athena, Greece
 Kenneth Salem, University of Waterloo
 Maria Sapino, University of Torino
 Kai-Uwe Sattler, TU Ilmenau
 Monica Scannapieco, ISTAT
 Bernhard Seeger, University of Marburg
 Lidan Shou, Zhejiang University
 Adam Silberstein, Trifacta
 Lisa Singh, Georgetown University
 Radu Sion, Stony Brook University
 Yufei Tao, Chinese University of Hong Kong
 Nesime Tatbul, ETH Zurich
 Arash Termehchy, Oregon State University
 Evimaria Terzi, University of Boston
 Martin Theobald, Max Planck Institute, Germany
 Srikanta Tirthapura, Iowa State University
 Riccardo Torlone, Roma Tre University
 Anthony Tung, National University of Singapore
 Kostas Tzoumas, Technical University of Berlin
 Sergei Vassilvitskii, Google Research
 Marcos Vaz Salles, University of Copenhagen (DIKU)
 Stratis Viglas, University of Edinburgh
 Hoang Tam Vo, National University of Singapore
 Daisy Zhe Wang, University of Florida
 Haixun Wang, Microsoft Research Asia
 Ke Wang, Simon Fraser University
 Wei Wang, University of New South Wales
 Xiaoling Wang, East China Normal University
 Ingmar Weber, Yahoo
 Raymond Chi Wing Wong, HKUST
 Sai Wu, Zhejiang University
 Yuqing Wu, Indiana University
 Xiaokui Xiao, NTU
 Dong Xin, Google
 Jianliang Xu, Hong Kong Baptist University
 Jun (Jim) Xu, Georgia Institute of Technology
 Xifeng Yan, University of Santa Barbara
 Xiaoyan Yang, Advanced Digital Science Center
 Ke Yi, HKUST
 Ge Yu, Northeastern University, China
 Hwanjo Yu, POSTECH, Korea
 Meihui Zhang, National University of Singapore
 Wenjie Zhang, The University of New South Wales
 Ying Zhang, The University of New South Wales
 Zhenjie Zhang, Advanced Digital Science Center
 Wenzhao Zhou, Georgetown University
 Xiaofang Zhou, University of Queensland

PhD Workshop Chairs

Erich Neuhold, University of Vienna

Yunyao Li, IBM

Sponsorship Chairs

Mike Carey, University of California, Irvine

Lizhu Zhou, Tsinghua University

Local Organization Chair

Lidan Shou, Zhejiang University

Web Management Chair

Sai Wu, Zhejiang University

Conference and Registration Chairs

Ke Chen, Zhejiang University

Cuiping Li, Renmin University

Publicity Chairs

Vasilis Vassalos, AUEB, Greece

Dunlu Peng, USST, China

Proceedings Chairs

Li Xiong, Emory University

Cong Yu, Google Research

Treasury Chair

Li (Eric) Qian, University of Michigan

VLDB Endowment Liaison

Kyu-Young Whang, KAIST

PVLDB Managing Editor

Divesh Srivastava, AT&T Labs

PVLDB Information Director

Gerald Weber, University of Auckland

PVLDB Advisory Committee

Philip Bernstein, Michael Böhlen, Peter Buneman,
Susan Davidson, Z. Meral Ozsoyoglu, S. Sudarshan,
Gerhard Weikum

Logo Design

Guanmin Guo

LETTER FROM VLDB 2014 GENERAL CHAIR

On behalf of the conference committee for VLDB 2014, it gives me great pleasure to welcome you to Hangzhou, China for the 40th International Conference on Very Large Data Bases. VLDB is a premier annual conference for data management that attracts a large number of database researchers, practitioners, developers, and users from all over the world. As has been the practice, the conference will feature a combination of research and industrial talks, keynotes, tutorials, panels, workshops (including a Ph.D. workshop), and demonstrations. It will cover visionary ideas, new techniques and algorithms, technology breakthroughs, major advances in data management principles along with innovative exciting applications. The plenary sessions, keynotes, and other forums including workshops will highlight emerging challenges for the data management research community. We are fortunate to have two leading experts from industry -- Hasso Plattner from SAP and Shivakumar Venkataraman from Google -- give keynote addresses. To the best of my knowledge, for the first time, we will experiment with a new model for keynote sessions where the session in addition to the keynote talk also includes a parallel talk given in response to the keynote from the academic perspectives. Volker Markl and Divy Agrawal will provide such an academic response to talk by Plattner and Venkataraman respectively. We hope that the new format helps provide both a more entertaining as well as a broader perspectives representing both the academic and industrial views.

VLDB 2014 is the first time the conference is being held in mainland China. Hangzhou city, the conference venue, is a major tourist attraction that offers large number of opportunities for sightseeing foremost of which is perhaps the very scenic and serene West Lake. Hangzhou is also famous for its Longjin green tea and for silk. I do hope you do not miss the opportunity to enjoy the wonderful sights, culture, and cuisine of Hangzhou. Hangzhou, besides being a major tourist destination, is also a major emerging technology hub often referred to as the "paradise silicon valley" in China. It is home to Zhejiang University, the local hosts for VLDB 2014 which, established in 1897, has long being considered as one of the top three universities of China.

The conference would not have been possible without the efforts of a large number of people who have worked very hard to make the conference a success. I would like to express my gratitude to Dr. Pan Yunhe for his role as honorary chair for VLDB 2014. I would specially like to acknowledge the contributions of my General Co-Chair Dr. Chen Chun and the local team headed by the local Organizations Chair Dr. Lidan Shou. I am very grateful to Lidan for his tireless attention to the minute details on almost every aspect of the conference organization without which it would have been difficult, if not impossible, for me to help organize the conference sitting on the other side of the Pacific Ocean. I would also like to thank the Technical Program Chairs Dr. H. V. Jagadish and Dr. Aoying Zhou and their team of Research track associate chairs Dr. Shivnath Babu, Dr. Lei Chen, Dr. Graham Cormode, Dr. Bin Cui, Dr. Wynne Hsu, Dr. Donald Kossman, Dr. Elke Rundensteiner, Dr. Kyuseok Shim, Dr. Wang-Chiew Tan, Dr. Letizia Tanca, and Dr. Jeffrey Yu for selecting a wonderful set of research papers to be presented at the conference. Thanks are also due to Dr. Gao Cong and Dr. Jens Dittrich for serving as Experiments and Analysis Track Associate Chairs; to Dr. Zachary Ivers for serving as the vision track associate editor; to Dr. Umesh Dayal, Dr. C Mohan, and Dr. Ge Yu for serving as the Industrial and Applications Track Associate Editors; to Dr. Mong-Li Lee, Dr. Fei Fei Li, and Dr. Sunil Prabhakar for serving as the demonstration chairs; Dr. Erich Neuhold and Dr. Yunyao Li, IBM for serving as Ph.D. workshop chairs; Dr. Anastasia Ailamaki and Dr. Kaushik Chakrabarti for their role as workshop chairs; Dr. Xiaoyong Du, Dr. Murat Kantarcioglu and Dr. Divesh Srivastava for serving as Tutorial chairs; Dr. Hakan Hacigumus, Dr. Jignesh Patel and Dr. Xiaoyang Sean Wang for serving as Panel Chair; Dr. Vasilis Vassalos and Dr. Dunlu Peng for serving as publicity chairs; Dr. Li Xiong and Dr. Cong Yu for their role as Proceedings Chairs; Dr. Sai Wu for his role as Web Management Chair; Dr. Eric Qian for serving as Treasury Chair; Ke Chen, Cuiping Li for conference and registration management; and Guanmim Guo for Logo Design. Special thanks to Dr. Mike Carey and Dr. Lizhu Zhou for helping raise sponsorship money in their role as sponsorship chairs. Last, but not the least, I would like to acknowledge Dr. Kyu-young Whang (VLDB endowment Liaison) and Dr. Beng-Chin Ooi for their continuous advise and guidance.

Thanks also to Dr. Themis Palpanas and Yannis Velegrakis (General Chairs, VLDB 2013); and to Dr. Ling Liu and Dr. Adnan Yazici (General Chairs, VLDB 2012) for sharing their experiences and conference preparation materials with us.

We are grateful to the generous support of our sponsors: Platinum Sponsors – Alibaba, Couchbase, CityCloud, Microsoft, Netease, Oracle, SAP; Gold Sponsors – Facebook, HTC, IBM, Google, Pivotal; Silver Sponsors – Baidu, Intel, Persistent Systems, Tencent; and Bronze Sponsors – Cisco, VMWare, TPC, NEC Labs America.

Finally, our sincere thanks to the authors of the papers, demos, panelists, speakers and, of course, to the hard work of the reviewers and to all who contributed to the program. Without your help we would not have a conference.

Welcome and enjoy the conference as well as the sights, culture, cuisine, and hospitality of the city of Hangzhou!

Sharad Mehrotra, University of California, Irvine, USA
General Chair, VLDB 2014