

# **VLDB**2012

38<sup>th</sup> International Conference on Very Large Data Bases, Istanbul, Turkey



# Proceedings of the VLDB Endowment

Volume 5, No. 8 – April 2012

**Proceedings of the 38th International Conference on  
Very Large Data Bases, Istanbul, Turkey**

Editor-in-Chief:

**Z. Meral Özsoyoğlu**

Associate Editors - Research Track:

**Uğur Çetintemel, Nilesh Dalvi, Hank Korth, Anthony Tung**

Associate Editors - Experiments and Analysis Track:

**Gustavo Alonso, Juliana Freire**

Proceedings Editors:

**Ahmet Saçan, Nesime Tatbul**

PVLDB – Proceedings of the VLDB Endowment

Volume 5, No. 8, April 2012.

The 38th International Conference on Very Large Data Bases, Istanbul, Turkey.

## **Copyright 2012 VLDB Endowment**

Permission to make digital or hard copies of portions of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright for components of this work owned by others than VLDB Endowment must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists requires prior specific permission and/or a fee. Request permission to republish from PVLDB under email: [info@vldb.org](mailto:info@vldb.org).

Volume 5, Number 8: VLDB 2012

Pages i - vii and 692 - 787

ISSN 2150-8097, April 2012.

Additional copies only online at: [portal.acm.org](http://portal.acm.org), [arxiv.org/corr](http://arxiv.org/corr), and [www.vldb.org](http://www.vldb.org)

## TABLE OF CONTENTS

### Front Matter

Copyright Notice .....	ii
Table of Contents .....	iii
PVLDB Review Board .....	iv

### Letters

Letter from the Associate Editors .....	vii
..... <i>Uğur Çetintemel, Nilesh Dalvi</i>	

### Research Track Papers

Shortest Path Computation with No Information Leakage .....	692
..... <i>Kyriakos Mouratidis, Man Lung Yiu</i>	
V-SMART-Join: A Scalable MapReduce Framework for All-Pair Similarity Joins of Multisets and Vectors .....	704
..... <i>Ahmed Metwally, Christos Faloutsos</i>	
Distributed GraphLab: A Framework for Machine Learning in the Cloud .....	716
..... <i>Yucheng Low, Joseph Gonzalez, Aapo Kyrola, Danny Bickson, Carlos Guestrin, Joseph M. Hellerstein</i>	
Adding Logical Operators to Tree Pattern Queries on Graph-Structured Data .....	728
..... <i>Qiang Zeng, Xiaorui Jiang, Hai Zhuge</i>	
Learning Semantic String Transformations from Examples .....	740
..... <i>Rishabh Singh, Sumit Gulwani</i>	
Cologne: A Declarative Distributed Constraint Optimization Platform .....	752
..... <i>Changbin Liu, Lu Ren, Boon Thau Loo, Yun Mao, Prithwish Basu</i>	
Optimizing I/O for Big Array Analytics .....	764
..... <i>Yi Zhang, Jun Yang</i>	
Probabilistically Bounded Staleness for Practical Partial Quorums .....	776
..... <i>Peter Bailis, Shivaram Venkataraman, Michael J. Franklin, Joseph M. Hellerstein, Ion Stoica</i>	

## **PVLDB REVIEW BOARD**

### **VLDB 2012 General PC Chair**

Z. Meral Özsoyoğlu, Case Western Reserve University, USA

### **VLDB 2012 Research Track PC Co-Chairs**

Uğur Çetintemel, Brown University, USA

Nilesh Dalvi, Yahoo Research

Hank Korth, Lehigh University, USA

Anthony Tung, National Univ. of Singapore, Singapore

### **VLDB 2012 Experiments and Analysis Track PC Co-Chairs**

Gustavo Alonso, ETH Zurich, Switzerland

Juliana Freire, NYU Poly, USA

### **VLDB 2012 Research Track Reviewers**

Karl Aberer, EPFL, Switzerland

Ashraf Aboulnaga, University of Waterloo

Yanif Ahmad, Johns Hopkins University

Shivnath Babu, Duke University

Magdalena Balazinska, University of Washington

Elisa Bertino, Purdue University

Claudio Bettini, University of Milano

Sourav Bhowmick, Nayang Technological University

Angela Bonifati, Icar-CNR

Randal Burns, Johns Hopkins University

Michael Cafarella, Michigan

Kevin Chang, Univ. of Illinois at Urbana-Champaign

Gang Chen, Zhejiang University, China

Lei Chen, Hong Kong University of SaT, China

Reynold Cheng, University of Hong Kong

Jan Chomicki, University of Buffalo

Chris Clifton, Purdue University

Gao Cong, Nayang Technological University

Brian Cooper, Google, USA

Graham Cormode, AT&T

Bin Cui, Peking University

Carlo Curino, MIT

Anish Das Sharma, Yahoo

Antonios Deligiannakis, Technical University of Crete

Amr El Abbadi, UC Santa Barbara

Alan Fekete, University of Sydney

Hakan Ferhatosmanoğlu, Ohio State University

Alvaro Fernandes, University of Manchester

Johann-Christoph Freytag, Humboldt-Universität zu Berlin

Minos Garofalakis, TU Creta, Greece

Bugra Gedik, IBM Watson  
 Gabriel Ghinita, Purdue University  
 Goetz Graefe, HP  
 Sergio Greco, University of Calabria  
 Todd Green, University of California, Davis  
 Le Gruenwald, University of Oklahoma  
 Haryadi Gunawi, University of California, Berkeley  
 Rahul Gupta, IIT Bombay  
 Stavros Harizopoulos, HP  
 Howard Ho, IBM  
 Vagelis Hristidis, Florida International University  
 Meichun Hsu, HP  
 Jeong-Hyon Hwang, SUNY, Albany  
 Ihab Ilyas, University of Waterloo  
 Zachary Ives, University of Pennsylvania  
 Dean Jacobs, SAP  
 Hans-Arno Jacobsen, University of Toronto  
 Christian Jensen, Aalborg University  
 Chris Jermaine, Rice University  
 Ricardo Jumenez-Peris, Technical University of Madrid  
 Theodore Johnson, AT&T  
 Vana Kalogeraki, Athens U of Econ. and Business  
 Raghav Kaushik, Microsoft Research  
 Eamonn Keogh, University of California, Riverside  
 Martin Kersten, CWI  
 Benny Kimelfeld, IBM  
 George Kollios, Boston University  
 Flip Korn, AT&T Labs Research  
 Donald Kossmann, ETH Zurich  
 Laks V.S. Lakshmanan, University of British Columbia  
 Mong-Li Lee, National University of Singapore  
 Wang-Chien Lee, Penn State University  
 Cuiping Li, Renmin University of China  
 Guoliang Li, Tsinghua University, China  
 Lipyeow Lim, University of Hawaii at Manoa  
 David Lomet, Microsoft Research  
 Boon Thau Loo, University of Pennsylvania  
 Qiong Luo, Hong Kong University of SaT, China  
 Ashwin Machanavajjhala, Yahoo  
 Sanjay Madria, University of Missouri-Rolla  
 Stefan Manegold, CWI, Netherlands  
 Sharad Mehrotra, University of California, Irvine  
 Gerome Miklau, Univ. of Massachusetts, Amherst  
 Bongki Moon, University of Arizona  
 Kyriakos Mouratidis, Singapore Management Univ.  
 Gero Muhl, University of Rostock  
 Suman Nath, MSR  
 Sylvia Nittel, University of Maine  
 Dan Olteanu, Oxford  
 M. Tamer Özsu, University of Waterloo  
 Esther Pacitti, University of Montpellier  
 Olga Papaemmanouil, Brandeis University  
 Srinivasan Parthasarathy, The Ohio State University  
 Jian Pei, Simon Fraser University  
 Peter Pietzuch, Imperial College  
 Lucian Popa, IBM  
 Rachel Pottinger, University of British Columbia  
 Sunil Prabhakar, Purdue University  
 Vijayshankar Raman, IBM, USA  
 Vibhar Rastogi, Yahoo  
 Christopher Re, University of Wisconsin, Madison  
 Ken Ross, Columbia University  
 Elke Rundensteiner, Worcester Polytechnic Institute  
 Maria Sapino, University of Torino  
 Yücel Saygın, Sabanci University  
 Kyuseok Shim, Seoul National University  
 Lidan Shou, Zhejiang University, China  
 Adam Silberstein, Yahoo!, USA  
 Yannis Sismanis, IBM, USA  
 Richard Snodgrass, University of Arizona  
 Divesh Srivastava, AT&T  
 Dan Suciu, University of Washington

S Sudarshan, IIT Bombay  
Egemen Tanin, University of Melbourne  
Yufei Tao, Chinese University of Hong Kong, China  
Martin Theobald, Max Planck Institute, Germany  
Frank Tompa, University of Waterloo  
Stratis Viglas, University of Edinburgh  
Wei Wang, University of North Carolina

Gerhard Weikum, Max Planck Institute, Germany  
Dong Xin, Google  
Jiong Yang, Case Western Reserve University  
Jun Yang, Duke  
Ge Yu, Northeastern University, China  
Aoying Zhou, East China Normal University, China

### **VLDB 2012 Experiments and Analysis Track Reviewers**

Cristiana Amza, U. Toronto, Canada  
Phil Bernstein, Microsoft Research, USA  
Philippe Bonnet, IT U. of Copenhagen, Denmark  
Michael Carey, UC Irvine, USA  
Jens Dittrich, Uni. Saarland, Germany  
Mary Fernandez, AT&T Research, USA  
Bettina Kemme, McGill Uni., Canada  
Whang Kyu-Young, KAIST, Korea  
Samuel Madden, MIT, USA

Ioana Manolescu, INRIA, France  
Dimitris Papadias, Hong Kong University of SaT, China  
Hamid Pirahesh, IBM Almaden, USA  
Maya Ramanath, IIT Delhi, India  
Uwe Roehm, University of Sydney, Australia  
Jens Teubner, ETH Zurich, Switzerland  
Peter Triantafillou, U. Patras, Greece  
Cong Yu, Google, USA

### **PVLDB Information Director**

Gerald Weber, University of Auckland, New Zealand

### **VLDB 2012 Proceedings Co-Chairs**

Ahmet Saçan, Drexel University, USA  
Nesime Tatbul, ETH Zurich, Switzerland

### **Advisory Committee**

Philip Bernstein, Michael Boehlen, Peter Buneman, Susan Davidson, Z. Meral Özsoyoğlu, S. Sudarshan, Gerhard Weikum

## LETTER FROM THE ASSOCIATE EDITORS

We are delighted to introduce the eighth issue of Proceedings of the VLDB Endowment (PVLDB), Vol. 5. PVLDB publishes papers accepted as a result of a journal-style reviewing process from papers submitted via a monthly, year-around submission cycle. All PVLDB papers published throughout the year will be presented at the VLDB 2012 Conference, to be held in Istanbul, Turkey, August 27-31, 2012.

This issue includes 8 papers, which cover a variety of important and current topics that range from distributed machine learning to declarative constraint optimization, and from secure computation to probabilistic quorum systems. We believe that the innovative technique and insightful experimental results reported in these papers will influence data management theory and practice for years to come.

We would like to thank the authors for making high-quality submissions and working with the PC to improve them. We also would like to thank our PC members, the PC co-chairs, the general PC chair, and last, but certainly not least, the proceedings co-chairs for their hard work in continuing to execute this innovative and ambitious publication model smoothly.

We hope that you also find the papers in this issue noteworthy and educational. We look forward to seeing you all in Istanbul to experience what promises to be a great technical program, and a world city that you will immensely enjoy!

---

Uğur Çetintemel, Brown University  
Nilesh Dalvi, Yahoo Research  
Associate Editor, Proceedings of the VLDB Endowment, Volume 5  
Research Track PC Co-Chair, VLDB 2012