IDEAS'09

IDEAS'09

International Database Engineering & Applications Symposium

2009 Proceedings

September 19-21, 2009
Cetraro – Calabria – Italy

Program Chairs
Domenico Saccà, Sergio Greco

Local Chairs
Alfredo Cuzzocra, Sergio Flesca

Editor, General Chair
Bipin C. Desai

Bipin C. Desai

ISBN 978-1-60558-402-7

CINDI

ConfSys.org

Proceedings of

IDEAS'09

Cetraro - Calabria - Italy
16-18 September 2009

Program Chairs:
Domenico Saccà (Università della Calabria, Rende, Italy)
Sergio Greco (Università della Calabria, Rende, Italy)

Local Chairs
Alfredo Cuzzocrea (ICAR-CNR, Rende, Italy)
Sergio Flesca (Università della Calabria, Rende, Italy)

Editor, General Chair:
Bipin C. Desai (Concordia University, Montréal, Canada)

Sponsors:

ACM
BytePress
Concordia University
ACM International Conference Proceeding Series
Università della Calabria, Rende(CS), Italy
ICAR-CNR, Rende (CS), Italy

The Association for Computing Machinery
2 Penn Plaza, Suite 701
New York New York 10121-0701

ACM COPYRIGHT NOTICE. Copyright © 2008 by the Association for Computing Machinery, Inc. Permission to make digital or hard copies of part or all 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. Copyrights for components of this work owned by others than ACM must be honoured. 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 permissions from Publications Dept., ACM, Inc., fax +1 (212) 869-0481, or permissions@acm.org.

For other copying of articles that carry a code at the bottom of the first or last page, copying is permitted provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923,
+1-978-750-8400, +1-978-750-4470 (fax).

Notice to Past Authors of ACM-Published Articles

ACM intends to create a complete electronic archive of all articles and/or other material previously published by ACM. If you have written a work that was previously published by ACM in any journal or conference proceedings prior to 1978, or any SIG Newsletter at any time, and you do NOT want this work to appear in the ACM Digital Library, please inform permissions@acm.org, stating the title of the work, the author(s), and where and when published.

ACM ISBN: 978-1-60558-402-7/09/09 $5.00

Editorial production by: BytePress
Cover artwork and photos © by (front) Lettie and (back) Bipin C. Desai

Table of Contents

Full Papers iii
Short Papers v
Poster papers vi
Preface vii
IDEAS'09 Program Committee viii
IDEAS'09 External Reviewers ix
Organization x

Full Papers

Efficient Discovery of Join Plans in Schemaless Data **1**

Aybar C. Acar (Bilkent University), Amihai Motro (George Mason University)

A Magic Approach to Optimizing Incremental Relational Expressions **12**

Andreas Behrend (Rheinische Friedrich-Wilhelms Universitat Bonn)

A Logical Framework for Detecting Anomalies in Drug Resistance Algorithms **23**

Luciano Caroprese (University of Calabria), Breanndán Ó Nualláin (University of Amsterdam), P.M.A. (Peter) Sloot (University of Amsterdam), Ester Zumpano (University of Calabria)

A Complete Translation from SPARQL into Efficient SQL **31**

Brenden Elliot(Case Western Reserve University), En Cheng (Case Western Reserve University), Chimezie Thomas-Ogbuji(Microsoft Corp.), Z. Meral Ozsoyoglu (Case Western Reserve University)

Efficiently Support Map-Reduce-like Computation Models Inside Parallel DBMS **43**

Qiming Chen (Hewlett-Packard Company), Andy Therber (Hewlett-Packard Company), Meichun Hsu (Hewlett-Packard Company), Hans Zeller (Hewlett-Packard Company), Bin Zhang (Hewlett-Packard Company), Ren Wu (Hewlett-Packard Company)

Depth First Algorithms and Inferencing for AFD mining **54**

Jeremy T. Engle (Indiana University at Bloomington), Edward L. Robertson (Indiana University at Bloomington)

Speeding-up Data-driven Applications with Program Summaries **66**

Sonia Guehis (Universite Paris-Dauphine (Paris IX)), Virginie Goasdoué-Thion (Universite Paris-Dauphine (Paris IX)), Philippe Rigaux (Universite Paris-Dauphine (Paris IX))

Concurrent Updating Transactions on Versioned Data **77**

Tuukka Haapasalo (Helsinki University of Technology), Ibrahim Jaluta (Helsinki University of Technology) Seppo Sippu (University of Helsinki), Eljas Soisalon-Soininen (Helsinki University of Technology)

Semantics of a Runtime Adaptable Transaction Manager **88**

Florian Irmert (Friedrich-Alexander Universitat Erlangen-Nurnberg), Frank Lauterwald (Friedrich-Alexander Universitat Erlangen-Nurnberg), Christoph P. Neumann (Friedrich-Alexander Universitat Erlangen-Nurnberg), Michael Daum (Friedrich-Alexander Universitat Erlangen-Nurnberg), Richard Lenz (Friedrich-Alexander Universitat Erlangen-Nurnberg), Klaus Meyer-Wegener (Friedrich-Alexander Universitat Erlangen-Nurnberg)

Mining frequent conjunctive queries in star schemas **97**

Tao-Yuan Jen (Universite de Cergy-Pontoise),Dominique Laurent (Universite de Cergy-Pontoise), Nicolas Spyratos (Universite Paris Sud (Paris XI))

Mining Uncertain Data for Constrained Frequent Sets **109**

Carson K. Leung (University of Manitoba), Dale A. Brajczuk (University of Manitoba)

XKMis: Effective and Efficient Keyword Search on XML Databases **121**

Jiang Li (Griffith University), Junhu Wang (Griffith University), Maolin Huang (University of Technology Sydney)

Managing the Mappings between Domain Ontologies and Database Schemas when Formulating Relational Queries **131**

Kamran Munir (University of the West of England, Bristol), Mohammed Odeh (University of the West of England, Bristol), Richard McClatchey (University of the West of England, Bristol)

Near-Duplicate Detection for Web-Forums **142**

Klemens Muthmann (Technische Universitat Dresden), Wojciech Barczynski (SAP), Alexander Löser (Technische Universitat Berlin), Falk Brauer (SAP)

NexusDS: A Flexible and Extensible Middleware for Distributed Stream Processing **152**

Nazario Cipriani (Universitat Stuttgart), Mike Eissele (Universitat Stuttgart), Andreas Brodt (Universitat Stuttgart), Matthias Grossmann (Universitat Stuttgart), Bernhard Mitschang (Universitat Stuttgart)

Discovering Expressive Process Models from Noised Log Data **162**

Francesco Folino (Consiglio Nazionale delle Ricerche), Gianluigi Greco (University of Calabria), Antonella Guzzo (University of Calabria), Luigi Pontieri (Consiglio Nazionale delle Ricerche)

Finding reliable users and social networks in a social internetworking system **173**

Pasquale De Meo (University of Reggio Calabria), Antonino Nocera (University of Reggio Calabria), Giovanni Quattrone (University of Reggio Calabria), Domenico Rosaci(University of Reggio Calabria), Domenico Ursino (University of Reggio Calabria)

A Cluster-Based Approach to XML Similarity Joins **182**

Leonardo A. Ribeiro (Universitat Kaiserslautern), Theo Härder (Universitat Kaiserslautern), Fernanda S Pimenta (Universidade Federal do Rio Grande do Sul)

Incremental Aggregation of RFID Data **194**

Roberto De Virgilio (Roma Tre University), Pierpaolo Sugamiele (Roma Tre University) , Riccardo Torlone (Roma Tre University)

Efficient Keyword Proximity Search using A Frontier-Reduce Strategy based on d-Distance Graph Index **206**

Ming Zhong (Wuhan University), Mengchi Liu (Carleton University)

Short Papers

An Online Spatio-Temporal Association Rule Mining Framework for Analyzing and Estimating Sensor Data **217**

Hamed Chok (University of Oklahoma), Le Gruenwald (University of Oklahoma)

Semantic caching for pervasive grids **227**

Laurent d'Orazio (Universite Blaise Pascal (Clermont-Ferrand II)), Mamadou Kaba Traoré (Universite Blaise Pascal (Clermont-Ferrand II))

Mining Very Long Sequences in Large Databases with PLWAPLong **234**

C. I. Ezeife (University of Windsor), Kashif Saeed (University of Windsor), Dan Zhang (University of Windsor)

A Model for Matching and Integrating Heterogeneous Relational Biomedical Databases Schemas **242**

Yaser Karasneh (Universiti Putra Malaysia), Hamidah Ibrahim (Universiti Putra Malaysia) , Mohaed Othmanm(Universiti Putra Malaysia), Razali Yaakob(Universiti Putra Malaysia)

Efficient and Effective RFID Data Warehousing **251**

Bettina Fazzinga (University of Calabria), Sergio Flesca (University of Calabria), Filippo Furfaro (University of Calabria), Elio Masciari (Consiglio Nazionale delle Ricerche)

RDFKB: Efficient Support For RDF Inference Queries and Knowledge Management **259**

James P. McGlothlin (University of Texas at Dallas), Latifur R Khan (University of Texas at Dallas)

Maintaining Integrated XML Schema **267**

Kalpdrum Passi (Laurentian University of Sudbury), Don Morgan (McMaster University), Sanjay Madria(Univ. Of Missouri,-Rolla Rolla)

Formal and Conceptual Modeling of Spatio-Temporal Granularities **275**

Alberto Belussi (University of Verona), Carlo Combi (University of Verona), Gabriele Pozzani (University of Verona)

Information Integration with Uncertainty **284**

Ala A. Eshmawi (University of North Carolina at Greensboro), Fereidoon Sadri (University of North Carolina at Greensboro)

Optimizing Data Warehouse Loading Procedures for Enabling Useful-Time Data Warehousing **292**

Ricardo Jorge Santos (Universidade de Coimbra), Jorge Bernardino (Instituto Politecnico de Coimbra)

Poster Papers

Algebraic Foundation of a Data Model for an Extensible Space-Based Collaboration Protocol **301**

Stefan Craß (Technische Universitat Wien), Eva Kühn (Technische Universitat Wien), Gernot Salzer (Technische Universitat Wien)

Non-Functional Properties to Services Coordination Using Contracts **307**

Javier Alfonso Espinosa-Oviedo (Institut National Polytechnique de Grenoble), Genoveva Vargas-Solar (Institut National Polytechnique de Grenoble), José-Luis Zechinelli-Martini (Universidad de las Americas - Puebla) , Christine Collet (Institut National Polytechnique de Grenoble)

Mining Product Intention Rules from Transaction Logs of an Ecommerce Portal **311**

Ravi Chandra Jammalamadaka (eBay, Inc.), Naren Chittar (eBay, Inc.), Sanjay Ghatare (eBay, Inc.)

Uniform Query Processing in a Federation of RDFS and Relational Resources **315**

Merja Karjalainen (Chalmers University of Technology), Graham J. L. Kemp (Chalmers University of Technology)

An Organizational File Permission Management System Using the Cellular Data System **321**

Toshio Kodama (Advance Computer Systems, Inc.), Tosiyasu L. Kunii (Tokyo University), Yoichi Seki (Advance Computer Systems, Inc.)

Semantic Reconciliation in Peer Multi-Data source Management System **326**

Gilles Nachouki (Universite de Nantes), Mirna Nachouki (Ajman University of Science & Technology), Marie-pierre Chastang (Universite de Nantes)

Low-voltage Electricity Customer Profilling based on Load Data Clustering **330**

Francesco Gullo (University of Calabria), Giovanni Ponti (University of Calabria), Andrea Tagarelli (University of Calabria), Salvatore Iiritano (Exeura s.r.l.), Massimiliano Ruffolo (Exeura s.r.l.), Diego Labate (Enel s.p.a.)

Evaluation of Skyline Algorithms in PostgreSQL **334**

Hannes Eder (Technische Universitat Wien), Fang Wei (Albert-Ludwigs-Universitat Freiburg)

Preface

IDEAS'09 is the fourteenth meeting of a series of annual meetings bringing together academics and practitioners in the area of database engineering. The previous editions of IDEAS were held in Asia, Europe and North America.

IDEAS aims to address the needs of the academic community in database engineering. The convening of academics, government and industry professionals facilitates the discussion of recent progress and challenges in data mining, databases, information retrieval, applications of the Web in commerce and governance as well as database engineering. IDEAS also serves as a platform for theoreticians and practitioners to exchange their original research ideas on academic or application aspects of database engineering and their application in real systems; share their experiences on integrating new technologies into products and applications; discuss their work to real-life situations; describe their development and operations of challenging computing related systems; and identify unsolved challenges.

The IDEAS Program Committee consists of academics and professionals from different parts of the world. This committee has done an excellent job in finishing the single-blind review and on-line double-blind debate processes in record time. For example, on average, about 88% of reviews were submitted by the due date and about 3.5 reviews were received for each paper. Second, we have attracted many high-quality papers submitted by authors globally. The paper selection process was thorough and competitive. On average, each paper was refereed by at least three reviewers, and 24.3% of submissions were accepted as full research papers. For IDEAS'09, the program chairs have designed a rich program that spans over three days.

I am grateful to a number of people without whom we would not have been able to put the program together. These include our program co:chairs Domenico Saccà and Sergio Greco, and the program committee which has done an excellent job. I would also like to thank the many external reviewers who have helped ``in the background," and who made sure that we could stick with our schedule. I am grateful to the high number of authors who have considered IDEAS as the target for their work. Even though we could not accommodate every submission, we hope that the extensive reviews can be of help to people.

Many thanks also go to the hard work of Alfredo Cuzzocrea and Sergio Flesca for making the local arrangements. I am also grateful to Bernard Rous and Craig Rodkin of ACM and to ACM and BytePress for publishing these proceedings. Last, not least, I am indebted to the staff of Concordia University, Università della Calabria and ICAR-CNR for making this event a reality.

Bipin C. Desai
General Chair

IDEAS'09 Program Committee

- Afrati, Foto (National Technical University of Athens)
- Appice, Annalisa (University of Bari)
- Barga, Roger (Microsoft Research)
- Bernardino, Jorge (ISEC - Instituto Politécnico de Coimbra)
- Bonchi, Francesco (Yahoo! Reseearch, Barcelona)
- Cali, Andrea (Oxford University Computing Lab)
- Chakravarthy, Sharma (University of Texas at Arlington)
- Christodoulakis, Stavros (Technical University of Crete)
- Consens, Mariano P. (University of Toronto)
- Cruz, Isabel (University of Illinois)
- Cuzzocrea, Alfredo (University of Calabria)
- de Arno, Sandra (Federal University of Uberlândia)
- Desai, Bipin C. (Concordia University)
- Elmagarmid, Ahmed (Purdue University)
- Flesca, Sergio (University of Calabria)
- Furtado, Pedro (University of Coimbra)
- Ganguly, Sumit (ITT, Kanpur)
- Greco, Sergio (University of Calabria)
- Grumbach, Stephane (LIAMA)
- Gunopulos, Dimitrios (UCR, NKUA)
- Gyssens, Marc (Hasselt University)
- Lakshmanan, Laks V.S. (University of British Columbia)
- Leung, Carson K. (University of Manitoba)
- Ludaescher, Bertram (University of California at Davis)
- Mohania, Mukesh (IBM India Research Laboratory)
- Novikov, Boris (St-Petersburg University)
- Ozsoyoðlu, Z. Meral (Case Western Reserve University)
- Rigaux, Philippe (Université Paris-Dauphine)
- Risch, Tore (Uppsala University)
- Sacca, Domenico (University of Calabria)
- Sattler, Kai-Uwe (Technische Universitat Ilmenau)
- Saygin, Yücel (Sabanci University)
- Spyratos, Nicolas (Université de Paris-Sud)
- Tagarelli, Andrea (University of Calabria)
- Tanca, Letizia (Politecnico di Milano)
- Terracina, Giorgio (University of Calabria)
- Thalheim, Bernhard (University of Kiel)
- Tiberio, Paolo (University of Modena and Reggio Emilia)
- Torlone, Riccardo (Roma Tre University)
- Ursino, Domenico (University Mediterranea of Reggio Calabria,)
- Vargas-Solar, Genoveva (LSR-IMAG)
- Vidyasankar, Krishnamurthy (Memorial University of Newfoundland)
- Wang, Wei (University of South Wales)
- Xu Yu, Jeffery (Chinese University of Hong Kong)
- Zaniolo, Carlo (University of California)
- Zimányi, Esteban (ULB)

IDEAS'09 External Reviewers

- Alessandra Martello (Universite della Calabria)
- Alirio Perez (Concordia University)
- Andrea Pugliese (University of Calabria)
- Anne Laurent (Univ Montpellier 2)
- Arnab Bhattacharya (Indian Institute of Technology, Kanpur)
- Barzan Mozafari (UCLA)
- Benjamin Fung (Concordia University)
- Bettina Fazzinga (University of Calabria)
- Carlo A. Curino (UCLA)
- Corrado Loglisci (University of Bary)
- Daniel Zinn (UC Davis)
- Davide Martinenghi (Politecnico di Milano)
- Dimitris Kotzinos (Technical Institution of Serres)
- Domenico Rosaci (University Mediterranea of Reggio Calabria)
- Dora Souliou (National Technical University of Athens)
- Elio Masciari (ICAR-CNR, Italy)
- Eugenio Cesario (ICAR-CNR)
- Fabio Fumarola (University of Bari)
- Filippo Furfaro (University of Calabria)
- Francesco Ricca (Universita' della Calabria)
- Francois GOASDOUE (Univ. Paris-Sud)
- Gianluca Caminiti (University Mediterranea of Reggio Calabria)
- Gianluca Lax (University Mediterranea of Reggio Calabria)
- Giovanni Ponti (University of Calabria)
- Giovanni Quattrone (University Mediterranea of Reggio Calabria)
- Hong Cheng (Chinese University of Hong Kong)
- Jean-Marc Petit (INSA Lyon)
- Kalpdrum Passi (Laurentian University)
- Michael Byrd (UC Davis)
- Michel de Rougemont (Universite Paris II)
- Nikos Kiourtis (National Technical University of Athens)
- Pasquale De Meo (University Mediterranea of Reggio Calabria)
- Renato Umeton (Universite della Calabria)
- Ricardo Santos (University of Coimbra)
- Rui Chen (Concordia University)
- Salvatore Garruzzo (Universite Mediterranea di Reggio Calabria)
- Shawn Bowers (University of California, Davis)
- Silvia Stefanova (Uppsala University)
- Solmaz Kolahi (UBC)
- TongYaun Wang (Concordia University)
- Victor Kyritsis (National Technical University of Athens)
- Viet Phan-Luong (LIF Universite de Provence)
- Wenwu Qu (CASIA)

Organized by

Concordia University, Montreal, Canada;
Università della Calabria, Rende(CS), Italy
ICAR-CNR, Rende (CS), Italy
with the cooperation of ACM and BytePress

Program Chairs:
Domenico Saccà (Università della Calabria, Rende, Italy)
Sergio Greco (Università della Calabria, Rende, Italy)

Local Chairs
Alfredo Cuzzocrea (ICAR-CNR, Rende, Italy)
Sergio Flesca (Università della Calabria, Rende, Italy)

General Chair:
Bipin C. Desai (Concordia University, Montréal, Canada)