

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Moshe Y. Vardi

Rice University, Houston, TX, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

Gianluca Moro Sonia Bergamaschi
Sam Joseph Jean-Henry Morin
Aris M. Ouksel (Eds.)

Databases, Information Systems, and Peer-to-Peer Computing

International Workshops, DBISP2P 2005/2006
Trondheim, Norway, August 28-29, 2005
Seoul, Korea, September 11, 2006
Revised Selected Papers

Volume Editors

Gianluca Moro

University of Bologna, Dept. of Electronics, Computer Science and Systems (DEIS)

Via Venezia, 52, 47023 Cesena (FC), Italy

E-mail: gmoro@deis.unibo.it

Sonia Bergamaschi

Università di Modena e Reggio Emilia, Dip. di Ingegneria dell'Informazione

Via Vignolese 905, 41100 Modena, Italy

E-mail: bergamaschi.sonia@unimo.it

Sam Joseph

University of Hawaii, Dept. of Information and Computer Science

1680 East-West Road, POST 309, Honolulu, HI 96822, USA

E-mail: srjoseph@hawaii.edu

Jean-Henry Morin

Korea University Business School

Anam-Dong, Seongbuk-Gu, Seoul 136-701, Korea

E-mail: jhmorin@korea.ac.kr

Aris M. Ouksel

The University of Illinois at Chicago, Dept. of Information and Decision Sciences

2411 University Hall, Chicago, IL, USA

E-mail: aris@uic.edu

Library of Congress Control Number: 2007924345

CR Subject Classification (1998): H.2, H.3, H.4, C.2, I.2.11, D.2.12, D.4.3, E.1

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743

ISBN-10 3-540-71660-2 Springer Berlin Heidelberg New York

ISBN-13 978-3-540-71660-0 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2007

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper SPIN: 12042593 06/3180 5 4 3 2 1 0

Preface

The aim of the International Workshop on Databases, Information Systems and P2P Computing was to explore the promise of P2P to offer exciting new possibilities in distributed information processing and database technologies. The realization of this promise lies fundamentally in the availability of enhanced services such as structured ways for classifying and registering shared information, verification and certification of information, content distributed schemes and quality of content, security features, information discovery and accessibility, interoperation and composition of active information services, and finally market-based mechanisms to allow cooperative and noncooperative information exchanges.

The P2P paradigm lends itself to constructing large-scale, complex, adaptive, autonomous and heterogeneous database and information systems, endowed with clearly specified and differential capabilities to negotiate, bargain, coordinate and self-organize the information exchanges in large-scale networks. This vision will have a radical impact on the structure of complex organizations (business, scientific or otherwise) and on the emergence and the formation of social communities, and on how the information is organized and processed. The P2P information paradigm naturally encompasses static and wireless connectivity and static and mobile architectures. Wireless connectivity combined with the increasingly small and powerful mobile devices and sensors poses new challenges as well as opportunities to the database community. Information becomes ubiquitous, highly distributed and accessible anywhere and at any time over highly dynamic, unstable networks with very severe constraints on the information management and processing capabilities. Which techniques and data models may be appropriate for this environment, and yet guarantee or approach the performance, versatility and capability that users and developers come to enjoy in a traditional static, centralized and distributed database environment? Is there a need to define new notions of consistency and durability, and completeness, for example?

The workshop concentrated on exploring the synergies between current database research and P2P computing. It is our belief that database research has much to contribute to the P2P grand challenge through its wealth of techniques for sophisticated semantics-based data models, new indexing algorithms and efficient data placement, query processing techniques and transaction processing. Database technologies in the new information age form the crucial components of the first generation of complex adaptive P2P information systems, which are characterized by their ability to continuously self-organize, adapt to new circumstances, promote emergence as an inherent property, optimize locally but not necessarily globally, deal with approximation and incompleteness. This workshop also concentrated on the impact of complex adaptive information

systems on current database technologies and their relation to emerging industrial technologies such as IBM's autonomic computing initiative.

The workshop brought together key researchers from all over the world working on databases and P2P computing with the intention of strengthening this connection. In particular, the workshop series emphasizes discussions about methodologies, models, algorithms and technologies related to data management and P2P systems. Researchers from other related areas such as distributed systems, networks, multi-agent systems, artificial intelligence and complex systems are also invited. We seek high-quality and original contributions on the following non-exhaustive list of topics:

- Data models and query languages for P2P systems
- Data placement and query answering in P2P systems
- Indexing, caching and replication techniques for P2P systems
- Transaction management for P2P systems
- Metadata management in P2P systems
- Dynamic schema integration, interoperation
- Emergent semantics and evolution of ontologies in P2P systems
- P2P systems and the Semantic Web
- Wireless and mobile data dissemination, delivery, replication and synchronization in P2P systems
- Scalability, coordination, robustness and adaptability in P2P systems
- Information usage in P2P mobile ad-hoc networks
- Self-organization and emergent behavior in P2P data management systems
- E-commerce and P2P computing
- Participation and contract incentive mechanisms in P2P Systems
- Computational models of trust and reputation
- Community of interest building and regulation, and behavioral norms
- Intellectual property rights in P2P systems
- Resource allocation in P2P systems
- Scalable data structures for P2P systems
- Scalable infrastructure for discovery and composition of P2P services, service definition language and filtering
- Market-based mechanisms for the exchange of information services and resources allocation in P2P systems
- Knowledge discovery and P2P data mining
- P2P-oriented information systems
- Complex adaptive information systems
- Information ecosystems and P2P systems
- Security and privacy in ubiquitous P2P systems
- Grid computing infrastructure based on the P2P paradigm
- Multidisciplinary approaches to P2P systems
- Legal issues in P2P networks

This volume is the post-proceedings of DBISP2P 2005 and 2006, the 3rd and 4th International Workshop on Databases, Information Systems and P2P

Computing,¹. Both editions were held in conjunction with the International Conference on Very Large DataBases (VLDB), the 2005 edition was in Trondheim, Norway (August 28–29, 2005), while the last one was in Seoul, South Korea (September 11, 2006). The volume contains the papers presented at the workshop, fully revised to incorporate reviewers’ comments and discussions; moreover, it includes an invited paper.

The volume is organized according to the following sessions held in the two editions:

Third Edition

- Knowledge Discovery and Emergent Semantics
- Query Answering and Overlay Communities
- Indexing, Caching and Replication Techniques
- Complex Query Processing and Routing
- Semantic Overlay Networks
- Services, Agents and Communities of Interest

Fourth Edition

- Data Placement and Searching
- Semantic Search
- Querying Processing and Workload Balancing
- Continuous Queries and P2P Computing

We would like to thank the invited speakers of the third and fourth edition, respectively, Karl Aberer, full professor at EPFL-IC-IIF-LSIR, for the talk on “Managing trust in distributed environments” and Vana Kalogeraki, assistant professor at the University of California, for the presentation of the work on “Middleware for reliable real-time sensor data management”.

We express our deepest appreciation to the invited panelists of the third edition, Karl Aberer, Sonia Bergamaschi, Witold Litwin and Pavel Zezula, who conducted the panel on the theme “Semantic search: is it any new issue in P2P systems ?”

As far as the number of papers is concerned, in the third edition we received 39 contributions, out of which 12 were accepted as full papers and 11 as short works. In the fourth edition, which lasted only one day, among the 29 submissions, 11 works were selected as full papers and 4 as short contributions. All submissions were reviewed for scope, quality, originality and applicability by the Program Committee, to which we express our gratitude for preparing high-quality reviews in a short time. Finally, we would like to acknowledge the Steering Committee for its guidance and encouragement.

This workshop followed the successful first and second editions, which were both held in conjunction with VLDB, in Berlin (Germany) and in Toronto (Canada), in 2003 and 2004, respectively. In recognition of the interdisciplinary

¹ <http://dbisp2p.ingce.unibo.it/>

nature of P2P computing, a sister event called the International Workshop on Agents and Peer-to-Peer Computing (AP2PC)² was held in New York, USA and in Utrecht, The Netherlands, respectively in 2004 and 2005, both in conjunction with the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS).

October 2006

Sonia Bergamaschi
Sam Joseph
Jean-Henry Morin
Gianluca Moro
Aris M. Ouksel

² <http://p2p.ingce.unibo.it/>

Executive Committees

Organizers of the Third Edition

Program Co-chairs	Gianluca Moro (main contact) Dept. of Electronics, Computer Science and Systems University of Bologna, Italy
	Sonia Bergamaschi Dept. of Science Engineering University of Modena and Reggio-Emilia, Italy
	Aris M. Ouksel Dept. of Information and Decision Science University of Illinois at Chicago, USA
Invited Panelists	Karl Aberer EPFL, Lausanne, Switzerland
	Sonia Bergamaschi Dept. of Science Engineering University of Modena and Reggio-Emilia, Modena, Italy
	Witold Litwin Centre d'Études et de Recherches en Informatique Appliquée Université Paris Dauphine, Paris, France
	Pavel Zezula Dept. of Computer Systems and Communications Masaryk University, Brno, Czech Republic
Web Review System	Francesco Guerra and Mirko Orsini University of Modena and Reggio Emilia, Italy

Organizers of the Fourth Edition

Program Co-chairs	Gianluca Moro (main contact) Dept. of Electronics, Computer Science and Systems University of Bologna, Italy
-------------------	--------------------------------------------------------------------------------------------------------------------

Sonia Bergamaschi
Dept. of Science Engineering
University of Modena and Reggio-Emilia, Italy

Sam Joseph
Dept. of Information and Computer Science
University of Hawaii at Manoa, USA

Jean-Henry Morin
Korea University Business School, Seoul, Korea

Web Review System Francesco Guerra and Mirko Orsini
University of Modena and Reggio-Emilia, Italy

Steering Committee

Karl Aberer, EPFL, Lausanne, Switzerland

Sonia Bergamaschi, Dept. of Science Engineering
University of Modena and Reggio-Emilia, Italy

Manolis Koubarakis, Department of Informatics and Telecommunications
National and Kapodistrian University of Athens, Greece

Paul Marrow, Intelligent Systems Laboratory,
BTextact Technologies, UK

Gianluca Moro, Dept. of Electronics, Computer Science and Systems
University of Bologna, Italy

Aris M. Ouksel, Dept. of Information and Decision Science
University of Illinois at Chicago, USA

Claudio Sartori, IEIIT-BO-CNR, University of Bologna, Italy

Munindar P. Singh, Dept. of Computer Science
North Carolina State University, USA

Program Committee

Karl Aberer, EPFL, Switzerland

Alessandro Agostini, ITC-IRST Trento, Italy

Peter A. Boncz, CWI, The Netherlands

Silvana Castano, University of Milan, Italy

Isabel Cruz, University of Illinois, USA

Bin Cui, Singapore-MIT Alliance, Singapore

Alex Delis, Polytechnic University, NY, USA

Asuman Dogac, Middle East Technical University, Turkey

Fausto Giunchiglia, University of Trento, Italy

Francesco Guerra, University of Modena and Reggio-Emilia, Italy
 Mohand-Said Hacid, Lyon, France
 Manfred Hauswirth, EPFL, Switzerland
 Vana Kalogeraki, University of California, Riverside, USA
 Achilles D. Kameas, Computer Technology Institute, Greece
 Anastasios Kementsietsidis, University of Edinburgh, UK
 Manolis Koubarakis, University of Athens, Greece
 Matthias Klusch, DFKI, Saarbrücken, Germany
 Tan Kian Lee, National University of Singapore, Singapore
 Maurizio Lenzerini, University of Rome “La Sapienza”, Italy
 Witold Litwin, University Paris 9 Dauphine, France
 Pericles Loucopoulos, UMIST, Manchester, UK
 Jayant Madhavan, University of Washington, USA
 Alberto Montresor, University of Bologna, Italy
 Jean-Henry Morin, University of Geneva, Switzerland
 Gianluca Moro, University of Bologna, Italy
 Enrico Nardelli, University of Rome Tor Vergata, Italy
 Wolfgang Nejdl, Learning Lab Lower Saxony, Germany
 Wee Siong Ng, Singapore-MIT Alliance, Singapore
 Maria S. Pérez-Hernandez, Universidad Politécnica de Madrid, Spain
 Jean Marc Pierson, INSA de Lyon, France
 Evagelia Pitoura, University of Ioannina, Greece
 Dimitris Plexousakis, Institute of Computer Science FORTH, Greece
 Rachel Pottinger, The University of British Columbia, Canada
 Wolf Siberski, University of Hannover, Germany
 Steffen Staab, University of Koblenz-Landau, Germany
 Peter Triantafillou, RA Computer Technology Institute and University of
 Patras, Greece
 Ouri Wolfson, University of Illinois, Chicago USA
 Martin Wolpers, Learning Lab Lower Saxony, Germany
 Pinar Yolum, Bogazici University, Turkey
 Pavel Zezula, University of Brno, Czech Republic

Additional Reviewers and Helpers

Christos Goumopoulos
 Riad Mokadem
 Gabriele Monti
 Mirko Orsini
 Sahri Soror

Preceding Editions of DBISP2P

The references to the preceding editions of DBISP2P, including the volumes of revised and invited papers, are:

- DBISP2P 2004 was held in Toronto, Canada, August 29–30, 2004. The Web site can be found at <http://dbisp2p.ingce.unibo.it/2004/>; the proceedings were published by Springer as LNCS volume no. 3367 and are available online at: <http://springerlink.metapress.com/content/hhe7htl85kw7/>
- DBISP2P 2003 was held in Berlin, Germany, September 7–8, 2003. The Web site can be found at <http://dbisp2p.ingce.unibo.it/2003/>; the proceedings were published by Springer as LNCS volume no. 2944 and are available online at: <http://springerlink.metapress.com/content/v9fpfwe6c2t9/>

Table of Contents

Third Edition

Knowledge Discovery and Emergent Semantics

Galois Connections, T-CUBES, and P2P Data Mining	1
<i>Witold Litwin</i>	
Querying a Super-Peer in a Schema-Based Super-Peer Network	13
<i>Domenico Beneventano, Sonia Bergamaschi, Francesco Guerra, and Maurizio Vincini</i>	

Query Answering and Overlay Communities

Database Selection and Result Merging in P2P Web Search	26
<i>Sergey Chernov, Pavel Serdyukov, Matthias Bender, Sebastian Michel, Gerhard Weikum, and Christian Zimmer</i>	
Multiple Dynamic Overlay Communities and Inter-space Routing	38
<i>Pedro Furtado</i>	
Benefit and Cost of Query Answering in PDMS	50
<i>Armin Roth and Felix Naumann</i>	

Indexing, Caching and Replication Techniques

Cooperative Prefetching Strategies for Mobile Peers in a Broadcast Environment	62
<i>Wei Wu and Kian-Lee Tan</i>	
Symmetric Replication for Structured Peer-to-Peer Systems	74
<i>Ali Ghodsi, Luc Onana Alima, and Seif Haridi</i>	
A Gradient Topology for Master-Slave Replication in Peer-to-Peer Environments	86
<i>Jan Sacha and Jim Dowling</i>	

Complex Query Processing and Routing

A Content-Addressable Network for Similarity Search in Metric Spaces	98
<i>Fabrizio Falchi, Claudio Gennaro, and Pavel Zezula</i>	
Range Query Optimization Leveraging Peer Heterogeneity in DHT Data Networks	111
<i>Nikos Ntarmos, Theoni Pitoura, and Peter Triantafillou</i>	

Guaranteeing Correctness of Lock-Free Range Queries over P2P Data	123
<i>Stacy Patterson, Divyakant Agrawal, and Amr El Abbadi</i>	
Publish/Subscribe with RDF Data over Large Structured Overlay Networks	135
<i>Erietta Liarou, Stratos Idreos, and Manolis Koubarakis</i>	

Semantic Overlay Networks

A Semantic Information Retrieval Advertisement and Policy Based System for a P2P Network	147
<i>Giovanna Guerrini, Viviana Mascardi, and Marco Mesiti</i>	
Cumulative Algebraic Signatures for Fast String Search, Protection Against Incidental Viewing and Corruption of Data in an SDDS	155
<i>Witold Litwin, Riad Mokadem, and Thomas Schwarz</i>	
PARIS: A Peer-to-Peer Architecture for Large-Scale Semantic Data Integration	163
<i>Carmela Comito, Simon Patarin, and Domenico Talia</i>	
Processing Rank-Aware Queries in P2P Systems	171
<i>Katja Hose, Marcel Karnstedt, Anke Koch, Kai-Uwe Sattler, and Daniel Zinn</i>	
Semantic Caching in Schema-Based P2P-Networks	179
<i>Ingo Brunkhorst and Hadhami Dhraief</i>	
Aggregation of a Term Vocabulary for P2P-IR: A DHT Stress Test	187
<i>Fabius Klemm and Karl Aberer</i>	

Services, Agents and Communities of Interest

Peer Group-Based Dependency Management in Service-Oriented Peer-to-Peer Architectures	195
<i>Sascha Alda</i>	
LEAP-DB: A Mobile-Agent-Based Distributed DBMS Not Only for PDAs	203
<i>Peter Ahlbrecht and Andreas Bothe</i>	
Models and Languages for Overlay Networks	211
<i>Stefan Behnel and Alejandro Buchmann</i>	
A Peer-to-Peer Membership Notification Service	219
<i>Roberto Baldoni and Sara Tucci Piergiovanni</i>	
Querying Communities of Interest in Peer Database Networks	227
<i>Md. Delwar Hossain and Iluju Kiringa</i>	

Fourth Edition

Invited Talk

Middleware for Reliable Real-Time Sensor Data Management	235
<i>Vana Kalogeraki</i>	

Data Placement and Searching

Oscar: Small-World Overlay for Realistic Key Distributions	247
<i>Sarunas Girdzijauskas, Anwitaman Datta, and Karl Aberer</i>	
Keyword Searching in Structured Overlays Via Content Distance Addressing	259
<i>Yu-En Lu, Steven Hand, and Pietro Lió</i>	

Semantic Search

XML Query Routing in Structured P2P Systems	273
<i>Leonidas Fegaras, Weimin He, Gautam Das, and David Levine</i>	
Reusing Classical Query Rewriting in P2P Databases	285
<i>Verena Kantere and Timos Sellis</i>	
Efficient Searching and Retrieval of Documents in PROSA	298
<i>Vincenza Carchiolo, Michele Malgeri, Giuseppe Mangioni, and Vincenzo Nicosia</i>	
P2P Query Reformulation over Both-As-View Data Transformation Rules	310
<i>Peter McBrien and Alexandra Poulovassilis</i>	
RDFCube: A P2P-Based Three-Dimensional Index for Structural Joins on Distributed Triple Stores	323
<i>Akiyoshi Matono, Said Mirza Pahlevi, and Isao Kojima</i>	

Query Processing and Workload Balancing

Optimal Caching for First-Order Query Load-Balancing in Decentralized Index Structures	331
<i>Anwitaman Datta, Wolfgang Nejdl, and Karl Aberer</i>	
On Triple Dissemination, Forward-Chaining, and Load Balancing in DHT Based RDF Stores	343
<i>Dominic Battré, Felix Heine, André Höing, and Odej Kao</i>	
Priority Based Load Balancing in a Self-interested P2P Network	355
<i>Xuan Zhou and Wolfgang Nejdl</i>	

A Self-organized P2P Network for an Efficient and Secure Content Location and Download	368
<i>Juan Pedro Muñoz-Gea, Josemaria Malgosa-Sanahuja, Pilar Manzanares-Lopez, Juan Carlos Sanchez-Aarnoutse, and Joan Garcia-Haro</i>	

Query Coordination for Distributed Data Sharing in P2P Networks	376
<i>Maybin Muyebe and M. Sulaiman Khan</i>	

Continuous Queries and P2P Computing

A Comparative Study of Pub/Sub Methods in Structured P2P Networks	385
<i>Matthias Bender, Sebastian Michel, Sebastian Parkitny, and Gerhard Weikum</i>	

Answering Constrained k-NN Queries in Unstructured P2P Systems	397
<i>Bin Wang, Xiaochun Yang, Guoren Wang, Lei Chen, Sean X. Wang, Xuemin Lin, and Ge Yu</i>	

Scalable IPv4/IPv6 Transition: A Peer-to-Peer Based Approach	406
<i>Jun Bi, Xiaoxiang Leng, and Jianping Wu</i>	

Author Index	417
-------------------------------	------------