

*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

*New York University, NY, 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*

Wee Siong Ng Beng Chin Ooi  
Aris Ouksel Claudio Sartori (Eds.)

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

Second International Workshop, DBISP2P 2004  
Toronto, Canada, August 29-30, 2004  
Revised Selected Papers

## Volume Editors

Wee Siong Ng  
National University of Singapore  
Singapore-MIT Alliance  
4 Engineering Drive 3, Singapore, Malaysia  
E-mail: ngws@comp.nus.edu.sg

Beng Chin Ooi  
National University of Singapore  
Department of Computer Science  
School of Computing  
Kent Ridge, Singapore 117543, Malaysia  
E-mail: ooibc@comp.nus.edu.sg

Aris Ouksel  
University of Illinois at Chicago  
Department of Information and Decision Sciences  
601 South Morgan Street, Chicago, IL 60607, USA  
E-mail: aris@uic.edu

Claudio Sartori  
University of Bologna  
Department of Electronics, Computer Science and Systems  
Viale Risorgimento, 2, 40136 Bologna, Italy  
E-mail: claudio.sartori@unibo.it

Library of Congress Control Number: 2005921896

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

ISSN 0302-9743  
ISBN 3-540-25233-9 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  
springeronline.com

© Springer-Verlag Berlin Heidelberg 2005  
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India  
Printed on acid-free paper SPIN: 11404033 06/3142 5 4 3 2 1 0

# Preface

Peer-to-peer (P2P) computing promises 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 non-cooperative 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 pose new challenges to as well as opportunities for 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. What techniques and data models may be appropriate for this environment, and yet guarantee or approach the performance, versatility, and capability that users and developers have come to enjoy in traditional static, centralized, and distributed database environments? Is there a need to define new notions of consistency and durability, and completeness, for example?

This 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 will form the crucial components of the first generation of complex adaptive P2P information systems, which will be characterized by their ability to continuously self-organize, adapt to new circumstances, promote emergence as an inherent property, optimize locally but not necessarily globally, and deal with approximation and incompleteness. This workshop examined 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 was collocated with VLDB, the major international database and information systems conference. It offered the opportunity for experts from all over the world working on databases and P2P computing to exchange ideas on the more recent developments in the field. The goal was not only to present these new ideas, but also to explore new challenges as the technology matures. The workshop provided also a forum to interact with researchers in related disciplines. Researchers from other related areas such as distributed systems, networks, multiagent systems, and complex systems were invited.

Broadly, the workshop participants were asked to address the following general questions:

- What are the synergies as well as the dissonances between the P2P computing and current database technologies?
- What are the principles characterizing complex adaptive P2P information systems?
- What specific techniques and models can database research bring to bear on the vision of P2P information systems? How are these techniques and models constrained or enhanced by new wireless, mobile, and sensor technologies?

After undergoing a rigorous review by an international Program Committee of experts, including online discussions to clarify the comments, 14 papers were finally selected. The organizers are grateful for the excellent professional work performed by all the members of the Program Committee. The keynote address was delivered by Ouri Wolfson from the University of Illinois at Chicago. It was entitled “DRIVE: Disseminating Resource Information in Vehicular and Other Mobile Peer-to-Peer Networks.” A panel, chaired by Karl Aberer from EPFL (Ecole Polytechnique Fédérale de Lausanne) in Switzerland, addressed issues on next-generation search engines in a P2P environment. The title of the panel was “Will Google2Google Be the Next-Generation Web Search Engine?”

The organizers would particularly like to thank Wee Siong Ng from the University of Singapore for his excellent work in taking care of the review system and the website. We also thank the VLDB organization for their valuable support and the Steering Committee for their encouragement in setting up this series of workshops and for their continuing support.

# Organization

## Program Chair

Beng Chin Ooi  
Aris M. Ouksel  
Claudio Sartori

National University of Singapore, Singapore  
University of Illinois, Chicago, USA  
University of Bologna, Italy

## Steering Committee

Karl Aberer  
Sonia Bergamaschi  
Manolis Koubarakis  
Paul Marrow

EPFL, Lausanne, Switzerland  
University of Modena and Reggio-Emilia, Italy  
Technical University of Crete, Crete  
Intelligent Systems Laboratory,  
BTextact Technologies, UK  
University of Bologna, Cesena, Italy  
University of Illinois, Chicago, USA  
North Carolina State University, USA  
University of Bologna, Italy

Gianluca Moro  
Aris M. Ouksel  
Munindar P. Singh  
Claudio Sartori

## Program Committee

Divyakant Agrawal  
Boualem Benattallah  
Peter A. Boncz

University of California, Santa Barbara, USA  
University of New South Wales, Australia  
CWI, Netherlands and Alex Delis Polytechnic  
University, New York, USA

Fausto Giunchiglia  
Manfred Hauswirth  
Vana Kalogeraki  
Achilles D. Kameas  
Peri Loucopoulos  
Alberto Montresor  
Jean-Henry Morin  
Gianluca Moro  
Wolfgang Nejdl  
Wee Siong Ng  
Thu D. Nguyen  
Evaggelia Pitoura

University of Trento, Italy  
EPFL, Switzerland  
University of California, Riverside, USA  
Computer Technology Institute, Greece  
UMIST Manchester, UK  
University of Bologna, Italy  
University of Geneva, Switzerland  
University of Bologna, Italy  
Learning Lab Lower Saxony, Germany  
Singapore-MIT Alliance, Singapore  
Rutgers University, USA  
University of Ioannina, Greece

VIII Table of Contents

Dimitris Plexousakis	Institute of Computer Science, FORTH, Greece
Krithi Ramamritham	IIT, Bombay, India and Wolf Siberski University of Hannover, 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
Aoying Zhou	Fudan University, China

**Sponsoring Institutions**

Microsoft Corporation, USA  
Springer

# Table of Contents

## Keynote Address

Data Management in Mobile Peer-to-Peer Networks <i>Bo Xu, Ouri Wolfson</i> .....	1
---	---

## Query Routing and Processing

On Using Histograms as Routing Indexes in Peer-to-Peer Systems <i>Yannis Petrakis, Georgia Koloniari, Evaggelia Pitoura</i> .....	16
Processing and Optimization of Complex Queries in Schema-Based P2P-Networks <i>Hadhami Dhraief, Alfons Kemper, Wolfgang Nejdl, Christian Wiesner</i> .....	31
Using Information Retrieval Techniques to Route Queries in an InfoBeacons Network <i>Brian F. Cooper</i> .....	46

## Similarity Search in P2P Networks

Content-Based Similarity Search over Peer-to-Peer Systems <i>Ozgur D. Sahin, Fatih Emekci, Divyakant Agrawal, Amr El Abbadi</i> .....	61
A Scalable Nearest Neighbor Search in P2P Systems <i>Michal Batko, Claudio Gennaro, Pavel Zezula</i> .....	79
Efficient Range Queries and Fast Lookup Services for Scalable P2P Networks <i>Chu Yee Liau, Wee Siong Ng, Yanfeng Shu, Kian-Lee Tan, Stéphane Bressan</i> .....	93
The Design of PIRS, a Peer-to-Peer Information Retrieval System <i>Wai Gen Yee, Ophir Frieder</i> .....	107

## Adaptive P2P Networks

Adapting the Content Native Space for Load Balanced Indexing <i>Yanfeng Shu, Kian-Lee Tan, Aoying Zhou</i> .....	122
---	-----

On Constructing Internet-Scale P2P Information Retrieval Systems  
*Demetrios Zeinalipour-Yazti, Vana Kalogeraki,  
Dimitrios Gunopoulos* ..... 136

AESOP: Altruism-Endowed Self-organizing Peers  
*Nikos Ntarmos, Peter Triantafillou* ..... 151

**Information Sharing and Optimization**

Search Tree Patterns for Mobile and Distributed XML Processing  
*Adelhard Türling, Stefan Böttcher* ..... 166

Dissemination of Spatial-Temporal Information in Mobile Networks  
with Hotspots  
*Ouri Wolfson, Bo Xu, Huabei Yin* ..... 185

Wayfinder: Navigating and Sharing Information in a Decentralized  
World  
*Christopher Peery, Francisco Matias Cuenca-Acuna,  
Richard P. Martin, Thu D. Nguyen* ..... 200

CISS: An Efficient Object Clustering Framework for DHT-Based  
Peer-to-Peer Applications  
*Jinwon Lee, Hyonik Lee, Seungwoo Kang, Sungwon Choe,  
Junehwa Song* ..... 215

**Author Index** ..... 231