

Proceedings

15th International Conference on Data Engineering

March 23 – 26, 1999

Sydney, Australia

Edited by

Masaru Kitsuregawa, Leszek Maciaszek,
Mike Papazoglou and Calton Pu

Sponsored by

IEEE Computer Society Technical Committee on
Data Engineering



Los Alamitos, California

Washington • Brussels • Tokyo

Copyright © 1999 by The Institute of Electrical and Electronics Engineers, Inc.
All rights reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries may photocopy beyond the limits of US copyright law, for private use of patrons, those articles in this volume that carry a code at the bottom of the first page, provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

Other copying, reprint, or republication requests should be addressed to: IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, P.O. Box 133, Piscataway, NJ 08855-1331.

The papers in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and, in the interests of timely dissemination, are published as presented and without change. Their inclusion in this publication does not necessarily constitute endorsement by the editors, the IEEE Computer Society, or the Institute of Electrical and Electronics Engineers, Inc.

IEEE Catalog Number 99CB36337
ISBN: 0-7695-0071-4 (Softbound)
ISBN: 0-7803-5581-4 (Casebound)
ISBN: 0-7695-0073-0 (Microfiche)
ISSN: 1063-6382
PR0071

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: + 1-714-821-8380
Fax: + 1-714-821-4641
E-mail: cs.books@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1-732-981-1393
Fax: + 1-732-981-9667
mis.custserv@computer.org

IEEE Computer Society
Ooshima Building
1-4-2 Minami-Aoyama
Minato-ku, Tokyo 107-0062
JAPAN
Tel: + 81-3-3408-3118
Fax: + 81-3-3408-3553
Tokyo.ofc@computer.org

Editorial production by Regina Spencer Sipple and Danielle Martin

Cover art design by Joseph Daigle/Studio Productions

Printed in the United States of America by The Printing House

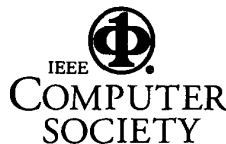


Table of Contents

15th International Conference on Data Engineering (ICDE '99)	
Message from the General Co-Chairs	xiii
Message from the Program Co-Chairs	xv
Organizing Committee	xvi
Program Committee Chairs	xviii
Program Committee Members	xix
External Reviewers.....	xxii
Keynote Address I	
Que Sera, Sera: The Coincidental Confluence of Economics, Business, and Collaborative Computing.....	2
<i>M.L. Brodie, GTE Laboratories, Inc.</i>	
Session 1: Multimedia Databases	
A Database Approach for Modeling and Querying Video Data	6
<i>C. Declair, M.-S. Hacid, and J. Kouloumdjian</i>	
Efficient Theme and Non-Trivial Repeating Pattern Discovering in Music Databases.....	14
<i>C.-C. Liu, J.-L. Hsu, and A.L.P. Chen</i>	
Query Processing Issues in Image (Multimedia) Databases.....	22
<i>S. Nepal and M.V. Ramakrishna</i>	
Session 2: Internet Query Processing	
On Getting Some Answers Quickly, and Perhaps More Later	32
<i>K.-L. Tan, C.H. Goh, and B.C. Ooi</i>	
Systematic Multiresolution and Its Application to the World Wide Web	40
<i>S. Acharya, H.F. Korth, and V. Poosala</i>	
Capability-Sensitive Query Processing on Internet Sources	50
<i>H. Garcia-Molina, W. Labio, and R. Yerneni</i>	
Industry Session 1: Application Integration	
A Messaging-Based Architecture for Enterprise Application Integration.....	62
<i>T. Joseph</i>	
Database as an Application Integration Platform.....	64
<i>A.R. Saxena</i>	

Business Objects and Application Integration.....	65
<i>S. Zeldin</i>	

Panel Session 1

Multimedia Information Management: Australian Industry Perspective

Chair: *Uma Srinivasan, CSIRO*

Session 3: Memory Management

Cooperative Caching in Append-only Databases with Hot Spots.....	70
<i>A. Sinha, C. Chase, and M. Cochinwala</i>	
Improving RAID Performance Using a Multibuffer Technique.....	79
<i>K.A. Hua, K. Vu, and T.-S. Hu</i>	
Managing Distributed Memory to Meet Multiclass Workload Response Time Goals.....	87
<i>M. Sinnwell and A.C. König</i>	

Poster Session 1: WWW, Integration, Workflow

Scalable Web Server Design for Distributed Data Management	96
<i>S.M. Baker and B. Moon</i>	
A Transparent Replication of HTTP Service.....	97
<i>R. Vingralek, Y. Breitbart, M. Sayal, and P. Scheuermann</i>	
Atomic Commitment in Database Systems over Active Networks.....	98
<i>Z. Zhang, W. Perrizo, and V.T.-S. Shi</i>	
A Hypertext Database for Advanced Sharing of Distributed Web Pages	99
<i>T. Yamakita and T. Fuji</i>	
Similarity Searching in Text Databases with Multiple Field Types	100
<i>K. Tzeras and E.G.M. Petrakis</i>	
Integrating Heterogeneous OO Schemas	101
<i>Y. Chen and W. Benn</i>	
Confirmation: A Solution for Non-Compensatability in Workflow Applications.....	102
<i>C. Liu, M. Orlowska, X. Zhou, and X. Lin</i>	
Policies in a Resource Manager of Workflow Systems: Modeling, Enforcement and Management.....	103
<i>Y.-N. Huang and M.-C. Shan</i>	
TP-Monitor-based Workflow Management System Architecture	104
<i>C. Bussler</i>	

Session 4: Data Mining I

Efficient Mining of Partial Periodic Patterns in Time Series Database	106
<i>J. Han, G. Dong, and Y. Yin</i>	
STING+: An Approach to Active Spatial Data Mining	116
<i>W. Wang, J. Yang, and R. Muntz</i>	
Efficient Time Series Matching by Wavelets	126
<i>K.-P. Chan and A.W.-C. Fu</i>	

Session 5: WWW & Digital Libraries

Enhancing Semistructured Data Mediators with Document Type Definitions	136
<i>Y. Papakonstantinou and P. Velikhov</i>	
Estimating the Usefulness of Search Engines	146
<i>W. Meng, K.-L. Liu, C. Yu, W. Wu, and N. Rishe</i>	
Query Routing in Large-Scale Digital Library Systems	154
<i>L. Liu</i>	

Industry Session 2: Document and WWW Information Service

Database Extensions for Complex Forms of Data	166
<i>S. DeFazio</i>	
Using XML in Relational Database Applications	167
<i>S. Malaika</i>	
Document Warehousing Based on a Multimedia Database System	168
<i>H. Ishikawa, K. Kubota, Y. Noguchi, K. Kato, M. Ono, N. Yoshizawa, and Y. Kanemasa</i>	
The ECHO Method: Concurrency Control Method for a Large-Scale Distributed Database	174
<i>Y. Shirota, A. Iizawa, H. Mano, and T. Yano</i>	

Keynote Address II

Data Mining from an AI Perspective	186
<i>Ross Quinlan</i>	

Session 6: Data Mining II

Constraint-Based Rule Mining in Large, Dense Databases	188
<i>R.J. Bayardo Jr., R. Agrawal, and D. Gunopulos</i>	
Parallel Classification for Data Mining on Shared-Memory Multiprocessors	198
<i>M.J. Zaki, C.-T. Ho, and R. Agrawal</i>	
Mining Optimized Support Rules for Numeric Attributes	206
<i>R. Rastogi and K. Shim</i>	

Session 7: Distributed & Mobile Objects

Using Java and CORBA for Implementing Internet Databases	218
<i>A. Bouguettaya, B. Benatallah, M. Ouzzani, and L. Hendra</i>	
Mobile Agents for WWW Distributed Database Access	228
<i>S. Papastavrou, G. Samaras, and E. Pitoura</i>	
Working Together in Harmony — An Implementation of the CORBA Object Query Service and Its Evaluation	238
<i>U. Röhm and K. Böhm</i>	

Poster Session 2: Interoperable Databases

Ad Hoc OLAP: Expression and Evaluation	250
<i>D. Chatziantoniou</i>	
Fast Approximate Search Algorithm for Nearest Neighbor Queries in High Dimensions	251
<i>S. Pramanik and J. Li</i>	
Fast Approximate Query Answering Using Precomputed Statistics	252
<i>V. Poosala and V. Ganti</i>	
Data Warehouse Maintenance under Concurrent Schema and Data Updates	253
<i>X. Zhang and E.A. Rundensteiner</i>	
Hash in Place with Memory Shifting: Datacube Computation Revisited	254
<i>J.X. Yu and H. Lu</i>	
Data Warehouse Evolution: Trade-Offs between Quality and Cost of Query Rewritings	255
<i>A.J. Lee, A. Koeller, A. Nica, and E.A. Rundensteiner</i>	
Exploiting Data Lineage for Parallel Optimization in Extensible DBMSs	256
<i>E.C. Shek and R.R. Muntz</i>	
The Bulk Index Join: A Generic Approach to Processing Non-Equijoins	257
<i>J. van den Bercken, B. Seeger, and P. Widmayer</i>	
Algorithms for Index-Assisted Selectivity Estimation	258
<i>P.M. Aoki</i>	
Multiple Index Structures for Efficient Retrieval of 2D Objects	259
<i>C. Shahabi, M. Safar, and H. Ai</i>	
Declarative and Procedural Object-Oriented Views	260
<i>R. Busse and P. Fankhauser</i>	
Speeding up Heterogeneous Data Access by Converting and Pushing Down String Comparisons	261
<i>W. Zhang and P.-Å. Larson</i>	

Panel Session 2

Privacy Issues in Data Mining

Chair: *Bhavani Thuraisingham, MITRE*

Panelists: *T.Y. Lin, San Jose State University*

Elisa Bertino, University of Milan

Vijay Varadharajan, University of Western Sydney

Session 8: Advanced Topics

Scalable Trigger Processing	266
<i>E.N. Hanson, C. Carnes, L. Huang, M. Konyala, L. Noronha, S. Parthasarathy, J.B. Park, and A. Vernon</i>	
Using Codewords to Protect Database Data from a Class of Software Errors	276
<i>P. Bohannon, R. Rastogi, S. Seshadri, A. Silberschatz, and S. Sudarshan</i>	

Integrating Light-Weight Workflow Management Systems within Existing Business Environments.....	286
<i>P. Muth, J. Weissenfels, M. Gillmann, and G. Weikum</i>	

Industry Session 3: Commercial Advance in QP and Access Methods

Index Merging	296
<i>S. Chaudhuri and V. Narasayya</i>	
Optimizer and Parallel Engine Extensions for Handling Expensive Methods Based on Large Objects	304
<i>W. O'Connell, F. Carino, and G. Linderman</i>	
Developing a DataBlade for a New Index	314
<i>R. Bliujute, S. Šaltenis, G. Slivinskas, and C.S. Jensen</i>	

Poster Session 3: Free Discussion

Session 9: OLAP & Multi-Dimensional Data

Relative Prefix Sums: An Efficient Approach for Querying Dynamic OLAP Data Cubes	328
<i>S. Geffner, D. Agrawal, A. El Abbadi, and T. Smith</i>	
Multidimensional Data Modeling for Complex Data.....	336
<i>T.B. Pedersen and C.S. Jensen</i>	
Maintaining Data Cubes under Dimension Updates	346
<i>C.A. Hurtado, A.O. Mendelzon, and A.A. Vaisman</i>	

Session 10: Interoperable Databases

Semantic Brokering over Dynamic Heterogeneous Data Sources in InfoSleuth™	358
<i>M. Nodine, W. Bohrer, and A.H.H. Ngu</i>	
Heterogeneous Query Processing through SQL Table Functions	366
<i>B. Reinwald, H. Pirahesh, G. Krishnamoorthy, G. Lapis, B. Tran, and S. Vora</i>	
Data Integration by Describing Sources with Constraint Databases	374
<i>X. Cheng, G. Dong, T. Lau, and J. Su</i>	

Session 11: Active Databases

An Agent-Based Approach to Extending the Native Active Capability of Relational Database Systems.....	384
<i>L. Li and S. Chakravarthy</i>	

On the Semantics of Complex Events in Active Database Management Systems.....	392
<i>D. Zimmer and R. Unland</i>	
Formal Semantics of Composite Events for Distributed Environments	400
<i>S. Yang and S. Chakravarthy</i>	

Session 12: Temporal Databases

On Similarity-Based Queries for Time Series Data	410
<i>D. Rafiei</i>	
Parallel Algorithms for Computing Temporal Aggregates	418
<i>J.A.G. Gendrano, B.C. Huang, J.M. Rodrigue, B. Moon, and R.T. Snodgrass</i>	
Universal Temporal Extensions for Database Languages	428
<i>C.X. Chen and C. Zaniolo</i>	

Session 13: Index Schemas I

The Hybrid Tree: An Index Structure for High Dimensional Feature Spaces	440
<i>K. Chakrabarti and S. Mehrotra</i>	
Fat-Btree: An Update-Conscious Parallel Directory Structure	448
<i>H. Yokota, Y. Kanemasa, and J. Miyazaki</i>	
Real-Time Data Access Control on B-Tree Index Structures.....	458
<i>T.-W. Kuo, C.-H. Wei, and K.-Y. Lam</i>	

Session 14: Data Warehousing

Scalable Classification over SQL Databases.....	470
<i>S. Chaudhuri, U. Fayyad, and J. Bernhardt</i>	
Storage of Multidimensional Arrays Based on Arbitrary Tiling	480
<i>P. Furtado and P. Baumann</i>	
Complements for Data Warehouses	490
<i>D. Laurent, J. Lechtenbörger, N. Spyros, and G. Vossen</i>	

Session 15: Data Clustering

Clustering Large Datasets in Arbitrary Metric Spaces.....	502
<i>V. Ganti, R. Ramakrishnan, J. Gehrke, A. Powell, and J. French</i>	
ROCK: A Robust Clustering Algorithm for Categorical Attributes	512
<i>S. Guha, R. Rastogi, and K. Shim</i>	
Data Organization and Access for Efficient Data Mining.....	522
<i>B. Dunkel and N. Soparkar</i>	

Session 16: Tertiary Storage	
Scheduling and Data Replication to Improve Tape Jukebox Performance	532
<i>B.K. Hillyer, R. Rastogi, and A. Silberschatz</i>	
Improving the Access Time Performance of Serpentine Tape Drives	542
<i>O. Sandstå and R. Midtstraum</i>	
Tape-Disk Join Strategies under Disk Contention	552
<i>A. Kraiss, P. Muth, and M. Gillmann</i>	
Session 17: Query Processing & Reconciliation	
Processing Operations with Restrictions in RDBMS without External Sorting: The Tetris Algorithm	562
<i>V. Markl, M. Zirkel, and R. Bayer</i>	
A Graph Query Language and Its Query Processing	572
<i>L. Sheng, Z.M. Özsoyoglu, and G. Özsoyoglu</i>	
Multiversion Reconciliation for Mobile Databases	582
<i>S.H. Phatak and B.R. Badrinath</i>	
Session 18: Secondary Storage Organization	
Design and Evaluation of Disk Scheduling Policies for High-Demand Multimedia Servers	592
<i>M.F. Khan, A. Ghafoor, and M.N. Ayyaz</i>	
Raster-Spatial Data Declustering Revisited: An Interactive Navigation Perspective	600
<i>C.-M. Chen and R.K. Sinha</i>	
Concentric Hyperspaces and Disk Allocation for Fast Parallel Range Searching	608
<i>H. Ferhatosmanoglu, D. Agrawal, and A. El Abbadi</i>	
Session 19: Index Schemas II	
Indexing Constraint Databases by Using a Dual Representation	618
<i>E. Bertino, B. Catania, and B. Chidlovskii</i>	
I/O Complexity for Range Queries on Region Data Stored Using an R-tree	628
<i>G. Proietti and C. Faloutsos</i>	
An Index Structure for Spatial Joins in Linear Constraint Databases	636
<i>H. Zhu, J. Su, and O.H. Ibarra</i>	

Panel Session 3

Directions for Data Management Research

Chair: Surajit Chaudhuri, Microsoft Research

Author Index.....	646
--------------------------	------------