Web-Age
Information Management

First International Conference, WAIM 2000
Shanghai, China, June 21–23, 2000
Proceedings
Series Editors

Gerhard Goos, Karlsruhe University, Germany
Juris Hartmanis, Cornell University, NY, USA
Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editors

Hongjun Lu
Hong Kong University of Science and Technology
Department of Computer Science
Clear Water Bay, Kowloon, Hong Kong, China
E-mail: luhj@cs.ust.hk

Aoying Zhou
Fudan University
Department of Computer Science
220 Handan Road, Shanghai, China
E-mail: ayzhou@fudan.edu.cn

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Web age information management : first international conference ; proceedings / WAIM 2000, Shanghai, China, June 21 - 23, 2000. Hongjun Lu ; Aoying Zhou (ed.) - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer, 2000
(Lecture notes in computer science ; Vol. 1846)
ISBN 3-540-67627-9

CR Subject Classification (1991): H.2, H.3, H.4, I.2, H.5, J.1

ISSN 0302-9743

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-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag is a company in the BertelsmannSpringer publishing group.
© Springer-Verlag Berlin Heidelberg 2000
Printed in Germany

Typesetting: Camera-ready by author
Printed on acid-free paper
SPIN 10722078 06/3142 5 4 3 2 1 0
Database research and development has been remarkably successful over the past three decades. Now the field is facing new challenges posted by the rapid advances of technology, especially the penetration of the Web and Internet into everyone's daily life. The economical and financial environment where database systems are used has been changing dramatically. In addition to being able to efficiently manage a large volume of operational data generated internally, the ability to manage data in cyberspace, extract relevant information, and discover knowledge to support decision making is critical to the success of any organization. In order to provide researchers and practitioners with a forum to share their experiences in tackling problems in managing and using data, information, and knowledge in the age of the Internet and Web, the First International Conference on Web-Age Information Management (WAIM 2000) was held in Shanghai, China, June 21-23.

The inaugural conference in its series was well received. Researchers from 17 countries and regions, including Austria, Australia, Bahrain, Canada, China, France, Germany, Japan, Korea, Malaysia, The Netherlands, Poland, Singapore, Spain, Taiwan, UK, and USA submitted their recent work. Twenty-seven regular and 14 short papers contained in these proceedings were presented during the two-day conference. These papers cover a large spectrum of issues, from classical data management such as object-oriented modeling, spatial and temporal databases to recent hits like data mining, data warehousing, semi-structured data, and XML. More importantly, there are also sections of papers devoted to related research areas, such as information retrieval, artificial intelligent agents, and electronic commerce, which broaden the scope of the conference to cover the entire process of data management, information retrieval and dissemination, knowledge discovery, and applications with the Internet as the media.

Any successful conference requires tremendous efforts from numerous people. In addition to the people in various conference organization committees, we would like to thank ACM SIGMOD who supported the conference from its planning stage and granted it the in-cooperation status. We are grateful to our sponsors, in particular the K. C. Wong Education Foundation, Hong Kong, for their generous support.

June 2000

Hongjun Lu, Aoying Zhou
Conference Organization

General Chairs:

Rakesh Agrawal  IBM Almaden Research Center, USA
Baile Shi  Fudan University, China

Program Committee Chairs:

Hongjun Lu  Hong Kong University of Science and Technology, China
Aoying Zhou  Fudan University, China

Organizing Committee Chairs:

Shoucai He  Shanghai No.2 University of Technology, China
Jinhai Chen  Fudan University, China

Publicity Chairs:

X. Sean Wang  George Mason University, USA
Yujun Wang  Fudan University, China

Local Arrangement:

Jiajin Le  China Textile University
Yangyong Zhu  Shanghai Computer Society, China

Program Committee:

Arbee Chen  National Tsing Hua University, Taiwan
Chen Guoqing  Tsinghua University, China
Ming-Syan Chen  National Taiwan University, Taiwan
David W. Cheung  Hong Kong University, China
Umeshwar Dayal  Hewlett-Packard Laboratories, USA
Guozhu Dong  Wright State University, USA
Martin Ester  University of Munich, Germany
Feng Yucai  Huazhong University of Science and Technology, China
Johannes Gehrke  Cornell University, USA
Stephane Grumbach  INRIA, France
Jiawei Han  Simon Fraser University, Canada
He Xingui  Beijing Institute of System Engineering, China
Kamal Karlapalem  Hong Kong University of Science and Technology, China
Masaru Kitsuregawa  University of Tokyo, Japan
Chiang Lee  National Cheng-Kung University, Taiwan
Qing Li  City University of Hong Kong, China
Jianzhong Li  Harbin Institute of Technology, China
Li Zhanhuai  Northwestern Polytechnical University, China
Tok Wang Ling  National University of Singapore, Singapore
Alberto Mendelzon  University of Toronto, Canada
Mukesh Mohania  University of South Australia, Australia
Shojiro Nishio    Osaka University, Japan
Ooi Beng Chin    National University of Singapore, Singapore
Maria E. Orlowska    University of Queensland, Australia
Michael Papazoglou    Tilburg University, The Netherlands
Qu Zhaorong    East China Institute of Computer Technology, China
Ming-Chien Shan    Hewlett-Packard Laboratories, USA
Jianwen Su    University of California at Santa Barbara, USA
Zhao-Hui Tang    Microsoft, USA
Tang Changjie    Sichuan University, China
Tang Shiwei    Peking University, China
Tian Zengping    Fudan University, China
Tong Fu    Shanghai University, China
X. Sean Wang    George Mason University, USA
Wang Shan    Renmin University of China, China
Kyu-Young Whang    KAIST, Korea
Wong Lim Soon    Kent Ridge Digital Laboratories, Singapore
Xu Jiepan    Nanjing University, China
Jian Yang    CSIRO, Australia
Jeffrey X. Yu    Chinese University of Hong Kong, China
Philip S. Yu    IBM T.J. Watson Research Center, USA
Yu Ge    Northeastern University, China
Yanchun Zhang    University of Southern Queensland, Australia
J. Leon Zhao    University of Arizona, USA
Zhou Lizhu    Tsinghua University, China
Zhou Longxiang    China Academy of Science, China

The First International Conference on Web-Age Information Management was organized by

Fudan University, China

in cooperation with

ACM SIGMOD
Shanghai Computer Society

and sponsored by

K.C. Wong Education Foundation, Hong Kong
Table of Contents

Invited Talks

Visual Data Mining for Business Intelligence Applications ................................. 3
Ming Hao, Umeshwar Dayal, Meichun Hsu
(Hewlett-Packard Laboratories, USA)

Parallel Data Mining on Large Scale PC Cluster............................................ 15
Masaru Kitsuregawa (University of Tokyo, Japan)

XML: DTD and Queries

XML Query Languages in Practice : An Evaluation ........................................ 29
Zachary G. Ives (University of Washington, USA),
Ying Lu (University of Wisconsin, USA)

A Two-Level Method for Clustering DTDs.................................................... 41
Weining Qian, Long Zhang, Yuqi Liang, Hailei Qian, Wen Jin
(Fudan University, China)

XML Queries via SQL.................................................................................... 53
Cindy X. Chen (University of California Los Angeles, USA),
Ashok Malhotra (IBM T. J. Watson Research Center, USA)

Inferring DTD to Facilitate User-Oriented XML Document Query .................. 61
Shihui Zheng (Fudan University, China)

O-O: Modeling and Implementation

An Object Oriented Multidimensional Data Model for OLAP ....................... 69
T. B. Nguyen, A. M. Tjoa (Vienna University of Technology, Austria),
R. R. Wagner (University of Linz, Austria)

Applying Object-Oriented Conceptual Modeling Techniques to the Design of Multidimensional Databases and OLAP Applications ......................... 83
Juan Carlos Trujillo, Manuel Palomar, Jaime Gómez
(University of Alicante, Spain)

An Optimal Locking Scheme in Object-Oriented Database Systems ............ 95
Woochun Jun (Seoul National University of Education, Korea),
Le Gruenwald (University of Oklahoma, USA)
Association Rules: Mining and Application
An Efficient Distributed Algorithm for Computing Association Rules....... 109
  Yijun Li, Xuemin Lin (University of New South Wales, Australia),
  Chi-Ping Tsang (University of Western Australia, Australia)
Mining Association Rules with Negative Items Using Interest Measure...... 121
  Haofeng Zhou, Pan Gao, Yangyong Zhu (Fudan University, China)
Probabilistic Approach to Association Rules in Incomplete Databases....... 133
  Marzena Kryszkiewicz (Warsaw University of Technology, Poland)
Discovering Sequential Patterns from Non-Uniform Databases............... 139
  Du Dang, X. Sean Wang (George Mason University, USA)
An Effective Approach to Mining Exception Class Association Rules ....... 145
  Fang Yu, Wen Jin (Fudan University, China)

Enterprise Information Systems and E-Commerce
Integrated Data Management and Enterprise Models ............................. 153
  Alain Bazan, Florida Estrellam, Zsolt Kovacs, Paul Lecoq, Jean-Marie Le
  Goff, Richard McClatchey, Steve Murray, Tony Solomonides,
  Jean-Pierre Vialle (University of Western England, UK)
A Virtual Private Network for Virtual Enterprise Information Systems....... 165
  Lina Wang, Ge Yu, Guoren Wang, Xiaochun Yang, Dan Wang, Xiaomei
  Dong, Daling Wang, Zhe Mei (Northeastern University, China)
Adaptive Online Retail Web Site Based on Hidden Markov Model............ 177
  Shi Wang, Wen Gao, Tiejun Huang, Jiyong Ma, Jintao Li, Hui Xie
  (Chinese Academic of Science, China)
A Protocol for Untraceable Electronic Cash ....................................... 189
  Hua Wang (Information Engineering University, China),
  Yanchun Zhang (University of Southern Queensland, Australia)

The Web: Document Management and Retrieval
Materializing Web Data for OLAP and DSS ....................................... 201
  Yan Zhu, Christof Bornhövd, Doris Sautner, Alejandro P. Buchmann
  (Darmstadt University of Technology, Germany)
Hierarchically Classifying Chinese Web Documents without Dictionary
  Support and Segmentation Procedure .............................................. 215
  Shuigeng Zhou, Ye Fan, Jiangtao Hu, Fang Yu, Yunfa Hu
  (Fudan University, China)
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web Clustering and Association Rule Discovery for Web Broadcast</td>
<td>227</td>
</tr>
<tr>
<td>Shi Wang, Wen Gao, Jintao Li, Tiejun Huang, Hui Xie</td>
<td></td>
</tr>
<tr>
<td>(Chinese Academic of Science, China)</td>
<td></td>
</tr>
<tr>
<td>A Non-Euclidean Model for Web Retrieval</td>
<td>233</td>
</tr>
<tr>
<td>Z. W. Wang, R. B. Maguire, Y. Y. Yao (University of Regina, Canada)</td>
<td></td>
</tr>
<tr>
<td>Advanced Replacement Policies for WWW Caching</td>
<td>239</td>
</tr>
<tr>
<td>Kai Cheng, Yahiko Kambayashi (Kyoto University, Japan)</td>
<td></td>
</tr>
<tr>
<td><strong>Spatial and Temporal Databases</strong></td>
<td></td>
</tr>
<tr>
<td>Extending Rectangle Join Algorithms for Rectilinear Polygons</td>
<td>247</td>
</tr>
<tr>
<td>Hongjun Zhu, Jianwen Su, Oscar H. Ibarra</td>
<td></td>
</tr>
<tr>
<td>(University of California at Santa Barbara, USA)</td>
<td></td>
</tr>
<tr>
<td>HMT: Modeling Temporal Aspects in Hypermedia Applications</td>
<td>259</td>
</tr>
<tr>
<td>Günther Specht (Technische Universität München, Germany), Peter Zoller</td>
<td></td>
</tr>
<tr>
<td>(Bavarian Research Centre for Knowledge-Based Systems, Germany)</td>
<td></td>
</tr>
<tr>
<td>Hana Tree: A Dynamic and Robust Access Method for Spatial Data Handling</td>
<td>271</td>
</tr>
<tr>
<td>Yongwon Kwon, Chang Sung Jeong (Korea University, Korea)</td>
<td></td>
</tr>
<tr>
<td>MADGIS: A New Architecture for Distributed GIS under Internet Environment</td>
<td>283</td>
</tr>
<tr>
<td>Jihong Guan, Shuigeng Zhou, Fuling Bian, Yunfa Hu</td>
<td></td>
</tr>
<tr>
<td>(Fudan University, China)</td>
<td></td>
</tr>
<tr>
<td><strong>Data Warehousing</strong></td>
<td></td>
</tr>
<tr>
<td>Global View Maintenance by Using Inference Relationship among Views</td>
<td>291</td>
</tr>
<tr>
<td>Haifeng Liu, Wee-Keong Ng, Chaohui Li</td>
<td></td>
</tr>
<tr>
<td>(Nanyang Technological University, Singapore)</td>
<td></td>
</tr>
<tr>
<td>Maintaining Materialized Views for Data Warehouses with Multiple Remote Sources</td>
<td>299</td>
</tr>
<tr>
<td>Weifa Liang, Chris Johnson (The Australian National University, Australia), Jeffrey X. Yu, (Chinese University of Hong Kong)</td>
<td></td>
</tr>
<tr>
<td>Using Loglinear Models to Compress Datacube</td>
<td>311</td>
</tr>
<tr>
<td>Daniel Barbará, Xintao Wu (George Mason University, USA)</td>
<td></td>
</tr>
<tr>
<td>Materialized View Selection in a Data Warehouse</td>
<td>323</td>
</tr>
<tr>
<td>Feng Yu, Du Xiaoyong, Wang Shan (Renmin University of China, China)</td>
<td></td>
</tr>
</tbody>
</table>
# Table of Contents

## Information Retrieval

<table>
<thead>
<tr>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>ExSight: Highly Accurate Object Based Image Retrieval System Enhanced by Redundant Object Extraction</td>
<td>331</td>
</tr>
<tr>
<td>Kazuhiko Kushima, Hiroki Akama, Seiichi Kon’ya, Masashi Yamamura (NTT Cyber Space Systems Laboratories, Japan)</td>
<td></td>
</tr>
<tr>
<td>Applying Anaphora Resolution to Question Answering and Information Retrieval Systems</td>
<td>344</td>
</tr>
<tr>
<td>José L. Vicedo, Antonio Ferrández (Universidad de Alicante, Spain)</td>
<td></td>
</tr>
<tr>
<td>MB+Tree  A Dynamically Updatable Metric Index for Similarity Searches</td>
<td>356</td>
</tr>
<tr>
<td>Masahiro Ishikawa (University of Tsukuba, Japan), Hanxiong Chen (Tsukuba International University, Japan), Kazutaka Furuse (University of Tsukuba, Japan), Jeffrey X. Yu (Chinese University of Hong Kong, China), Nobuo Ohbo (University of Tsukuba, Japan)</td>
<td></td>
</tr>
<tr>
<td>An Information Store and Retrieval Facility on CORBA</td>
<td>374</td>
</tr>
<tr>
<td>Winston Lo (Tung Hai University, Taiwan), Yue-Shan Chang (National Chiao Tung University, Taiwan), Chii-Long Chou (Feng Chia University, Taiwan), Ruey-Kai Sheu, Shyan-Ming Yuan (National Chiao Tung University, Taiwan)</td>
<td></td>
</tr>
<tr>
<td>Retrieving Content Directly from Antique Book Images</td>
<td>380</td>
</tr>
<tr>
<td>Zhifeng Chen, Yong Wang, Liang Zhang, Baile Shi (Fudan University, China)</td>
<td></td>
</tr>
</tbody>
</table>

## Semi-Structured Data

<table>
<thead>
<tr>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schema Based Data Storage and Query Optimization for Semi-structured Data</td>
<td>389</td>
</tr>
<tr>
<td>Qinke Wang, Lizhu Zhou (Tsinghua University, China)</td>
<td></td>
</tr>
<tr>
<td>Automatic Wrapper Generation for Web Search Engines</td>
<td>399</td>
</tr>
<tr>
<td>Boris Chidlovskii, Jon Ragetli (Xerox Research Centre Europe, France), Maarten de Rijke (University of Amsterdam, The Netherlands)</td>
<td></td>
</tr>
<tr>
<td>Multi-level Schema Extraction for Heterogenous Semi-structured Data</td>
<td>411</td>
</tr>
<tr>
<td>Jong P. Yoon, Vijay Raghavan (University of Louisiana, USA)</td>
<td></td>
</tr>
</tbody>
</table>
# Table of Contents

**Agent, Rough Set and Learning**

<table>
<thead>
<tr>
<th>Rough Set Based WebCT Learning</th>
<th>425</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Aileen H. Liang, Brien Maguire, Julia Johnson</em></td>
<td></td>
</tr>
<tr>
<td><em>(University of Regina, Canada)</em></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MultiAgent Systems for Coalition Support</th>
<th>437</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Zakaria Maamar (Defence Research Establishment Valcartier, Canada), Nader Troudi (Laval Universisy, Canada)</em></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A Fast Globally Supervised Learning Algorithm for Gaussian Mixture Models</th>
<th>449</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Jiyong Ma, Wen Gao (Chinese Academic of Science , China)</em></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Optimizing Classifiers by Genetic Algorithms</th>
<th>455</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Wenyun Ji, Liang Zhang, Wen Jin (Fudan University, China)</em></td>
<td></td>
</tr>
</tbody>
</table>

**Author Index** | 461