

Springer

Berlin

Heidelberg

New York

Barcelona

Hong Kong

London

Milan

Paris

Singapore

Tokyo

X. Sean Wang Ge Yu Hongjun Lu (Eds.)

Advances in Web-Age Information Management

Second International Conference, WAIM 2001
Xi'an, China, July 9-11, 2001
Proceedings



Springer

Series Editors

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

Volume Editors

X. Sean Wang
George Mason University, Department of Information and Software Engineering
Fairfax, VA 22030-4444, USA
E-mail: xywang@gmu.edu

Ge Yu
Northeastern University, Department of Computer Science and Engineering
Shenyang, 110004, China
E-mail: yuge@mail.neu.edu.cn

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

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Advances in web age information management : second international conference ;
proceedings / WAIM 2001, Xi'an, China, July 9 - 11, 2001. X. Sean Wang ...
(ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ;
Milan ; Paris ; Singapore ; Tokyo : Springer, 2001
(Lecture notes in computer science ; Vol. 2118)
ISBN 3-540-42298-6

CR Subject Classification (1998): H.2, H.3, H.4, I.2, H.5, J.1, C.2

ISSN 0302-9743

ISBN 3-540-42298-6 Springer-Verlag 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-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

<http://www.springer.de>

© Springer-Verlag Berlin Heidelberg 2001
Printed in Germany

Typesetting: Camera-ready by author, data conversion by PTP-Berlin, Stefan Sossna
Printed on acid-free paper SPIN: 10839875 06/3142 5 4 3 2 1 0

Preface

In this Web-age, international collaboration and exchange are even more important and interesting than before for information management research and development. In June 2000, the first WAIM conference was held in Shanghai, China, to promote such collaboration and exchange. It was a success and attracted more than 100 participants from 16 countries. WAIM 2001 continued this tradition, providing 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 and to identify new issues and directions for research and development work in the new millennium.

Like the inaugural conference in the year 2000, WAIM 2001 attracted submissions from researchers of many countries and regions, including Australia, Austria, Canada, China, France, Germany, Greece, Hong Kong, India, Italy, Japan, Jordan, Korea, Malaysia, Norway, Poland, Singapore, Spain, Switzerland, Taiwan, The Netherlands, UK, and USA. Twenty-one regular research papers and twelve short research papers, contained in these proceedings, were presented at the conference. These papers cover a wide range of topics, including multimedia databases, information retrieval, data warehousing, information management in e-commerce, semistructured data management, data mining, data integration and filtering, and workflow. Also presented at the conference, and contained in these proceedings, were four research experience papers. This reflects our efforts to extend the conference to also become a forum for promoting and exchanging good information management practices.

The preparation of a successful conference requires a lot of efforts from many people. In addition to the standard program committee work like reviewing and selecting papers, efforts are required that range from publicity to conference logistics. We appreciate the work of the people in our WAIM 2001 organization committees. We would also like to thank Northwestern Polytechnic University, China, for organizing the conference, and to thank Mr. Like Gao for maintaining the conference paper submission and reviewing website.

July 2001

X. Sean Wang, Ge Yu
WAIM 2001 Program Co-Chairs

Organization

General Chair:

Deyuan Gao Northwestern Polytechnic University, China

Conference Chairs:

Hongjun Lu Hong Kong University of Science & Technology,
China

Akifumi Makinouchi Kyushu University, Japan

Program Chairs:

X. Sean Wang George Mason University, USA

Ge Yu Northeastern University, China

Tutorial&Panel Chair:

Min Wang IBM T. J. Watson Research Center, USA

Publicity Chair:

Qiang Zhu The University of Michigan, USA

Organization Chairs:

Zhanhui Li Northwestern Polytechnic University, China

Kaiming Shi Northwestern Polytechnic University, China

WAIM Steering Committee:

Jianzhong Li Harbing Institute of Technology, China

Hongjun Lu (**Chair**) Hong Kong University of Science & Technology,
China

Baile Shi Fudan University, China

Jianwen Su University of California, Santa Barabara, USA

Shan Wang Renmin University, China

X. Sean Wang George Mason University, USA

Ge Yu Northeastern University, China

Aoying Zhou Fudan University, China

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

Northwestern Polytechnic University, China

in cooperation with

China Computer Federation, Database Society

and sponsored by

Fujitsu Limited, Japan,

TechnoProject Limited, Japan, and

National Science Foundation, China

Program Committee

Vijay Atluri	Rutgers University, USA
Daniel Barab�	George Mason University, USA
Claudio Bettini	University of Milan, Italy
Sang K. Cha	Seoul National University, Korea
Arbee L.P. Chen	National Tshing Hua University, Taiwan
Wan-Sup Cho	Chungbuk National University, Korea
Carlo Combi	University of Udine, Italy
Alex Delis	Polytechnic University, USA
Guozhu Dong	University of South Australia, Australia
Susan Gauch	University of Kansas, USA
Stephane Grumbach	INRIA, France
Jiawei Han	Simon Fraser University, Canada
Xingui He	Beijing Institute of System Eng., China
Larry Kerschberg	George Mason University, USA
Hiroyuki Kitagawa	University of Tsukuba, Japan
Yasushi Kiyoki	Keio University, Japan
Jiajin Le	Donghua University, China
Hans-J. Lenz	Free University, Berlin, Germany
Jianzhong Li	Harbin Institute of Technology, China
Tianzhu Li	Hebei University, China
Zhanhuai Li	Northwestern Polytechnic University, China
Huan Liu	Arizona State University, USA
Ling Liu	Georgia Tech, USA
Weiyi Meng	SUNY at Binghamton, USA
Xiaofeng Meng	Renmin University, China
William Perrizo	North Dakota State University, USA
Niki Pissinou	University of Louisiana at Lafayette, USA
Peter Revesz	University of Nebraska-Lincoln, USA
Baile Shi	Fudan University, China
Kyuseok Shim	KAIST, Korea
Dan Suciu	University of Washington, USA
Chanjie Tang	Sihuan University, China
Shiwei Tang	Peking University, China
Fu Tong	Shanghai University, China
Vassilis Tsotras	University of California, Riverside, USA
Shunsuke Uemura	NAIST, Japan
Victor Vianu	University of California, San Diego, USA
Guoren Wang	Northeastern University, China
Haiyang Wang	Shandong University, China
Jason Wang	New Jersey Institute of Technology, USA
Min Wang	IBM T. J. Watson Research Center, USA
Nengbin Wang	Southeast University, China
Shan Wang	Renmin University, China

(Program committee continued)

Kyu-Young Whang	KAIST, Korea
Jef Wijsen	University of Mons-Hainaut, Belgium
Jiepan Xu	Nanjing University, China
Dongqing Yang	Peking University, China
Masatoshi Yoshikawa	NAIST, Japan
Jeffrey X. Yu	Chinese University of Hong Kong, China
Aidong Zhang	SUNY Buffalo, USA
Aoying Zhou	Fudan University, China
Lizhu Zhou	Tsinghua University, China
Longxiang Zhou	Institute of Mathematics, China
Qiang Zhu	University of Michigan, USA

External Referees

Maggie Y.C. Chang	Xiuli Ma
George Chang	Atsuyuki Morishima
Pei-Yi Chen	Pratik Mukhopadhyay
Jie Chen	Chen Ning
Zhihong Deng	Jian Pei
Qin Ding	Weining Qian
Kai Essig	Huiyuan Shan
Youlin Fang	Xiaochun Tang
Jun Gao	Chun Tang
Zhiguo Gong	Anthony Tung
Katherine G. Herbert	Tengjiao Wang
Wei Huang	Dalin Wang
Yoshiharu Ishikawa	Xiong Wang
William Jockheck	Wei Wang
Maleq Khan	Lianghuai Yang
Yong-Sik Kwon	Michael M. Yin
Keunjoo Kwon	Liang Zhang
Shengen Li	Sen Zhang
Wang Lian	Shihui Zheng

Table of Contents

Multimedia Databases & High Dimensional Indexing (Session 1A)

Spatial-Match Iconic Image Retrieval with Ranking in Multimedia Databases	3
<i>Jae-Woo Chang and Yeon-Jung Kim (Chonbuk National University, Korea)</i>	
Similarity-Based Operators in Image Database Systems	14
<i>Solomon Atnafu, Lionel Brunie (LISI, INSA de Lyon, France), and Harald Kosch (University Klagenfurt, Austria)</i>	
An Efficient High-Dimensional Index Structure Using Cell Signatures for Similarity Search	26
<i>Jae-Woo Chang (Chonbuk National University, Korea) and Kwang-Taek Song (WOORIN R&D Laboratory, Korea)</i>	
Using Object-Relational Database Systems and XML in the Context of an Wireless-Oriented Multimedia Information System	34
<i>Christoph Dittberner, Thomas Courvoisier, and Guntram Flach (Computer Graphics Center Rostock, Germany)</i>	

Information Retrieval & Text Indexing (Session 1B)

Utilizing the Correlation between Query Keywords for Information Retrieval	49
<i>Tetsuya Yoshida, Daiki Shinkai, and Shogo Nishida (Osaka University, Japan)</i>	
Adjacency Matrix Based Full-Text Indexing Models	60
<i>Shuigeng Zhou, Jihong Guan (Wuhan University, China), Yunfa Hu, Jiangtao Hu, and Aoying Zhou (Fudan University, China)</i>	
Goal-Oriented Information Retrieval Using Feedback from Users	72
<i>Hiroiyuki Toda, Toshifumi Enomoto, and Tetsuji Satoh (NTT Cyber Space Laboratories, Japan)</i>	
Classify Web Document by Key Phrase Understanding	80
<i>Changjie Tang, Tong Li, Changyu Liu, and Yin Ge (Sichuan University, China)</i>	

Data Mining (Session 2A)

Deriving High Confidence Rules from Spatial Data Using
Peano Count Trees 91
*William Perrizo, Qin Ding, Qiang Ding, and Amalendu Roy (North
Dakota State University, USA)*

A Distributed Hierarchical Clustering System for Web Mining 103
*Catherine W. Wen, Huan Liu (Arizona State University, USA),
Wilson X. Wen (GoConnect, Australia), and
Jeffery Zheng (CMST CSIRO, Australia)*

ARMiner: A Data Mining Tool Based on Association Rules 114
*Haofeng Zhou, Beijun Ruan, Jianqiu Zhu, Yangyong Zhu, and
Baile Shi (Fudan University, China)*

An Integrated Classification Rule Management System for Data Mining .. 122
*Daling Wang , Yubin Bao, Xiao Ji, Guoren Wang, and
Baoyan Song (Northeastern University, China)*

Semistructured Data Management (Session 2B)

Characterizing Web Document Change 133
*Lipyeow Lim (Duke University, USA), Min Wang,
Sriram Padmanabhan (IBM T.J. Watson Research Center, USA),
Jeffrey S. Vitter (Duke University, USA), and
Ramesh Agarwal (IBM T.J. Watson Research Center, USA)*

Identification of Syntactically Similar DTD Elements for
Schema Matching 145
*Hong Su (Worcester Polytechnic Institute, USA),
Sriram Padmanabhan, and Ming-Ling Lo (IBM T.J. Watson
Research Center, USA)*

An XML Based Electronic Medical Record Integration System 160
*Hongyan Li, Shiwei Tang, Dongqing Yang (Peking University,
China), and Yi Yu (PeridotHealth Systems Pte Ltd, Singapore)*

Enhancing XML Data Processing in Relational System with Indices 168
*Yuqi Liang, Aoying Zhou, Shihui Zheng, Long Zhang, and
Wenyun Ji (Fudan University, China)*

Data Warehousing & Federated Databases (Session 3A)

A Multidimensional Data Model Supporting Complicated Aggregations ..	179
<i>Hongxing Tan (Henan University, Chinese Academy of Sciences, China), Wenwen Qi (Henan University, China), and Longxiang Zhou (Chinese Academy of Sciences, China)</i>	
Comparing String Similarity Measures for Reducing Inconsistency in Integrating Data from Different Sources	191
<i>Sergio Luján-Mora and Manuel Palomar (Universidad de Alicante, Spain)</i>	
Revisit on View Maintenance in Data Warehouses	203
<i>Weifa Liang (The Australian National University, Australia) and Jeffrey X. Yu (Chinese University of Hong Kong, China)</i>	
Experiences with a Hybrid Implementation of a Globally Distributed Federated Database System	212
<i>Xuebiao Xu, Andrew C. Jones, Nick Pittas, W. Alex Gray, Nick J. Fiddian (Cardiff University, UK), Richard J. White, John Robinson (University of Southampton, UK), Frank A. Bisby, and Sue M. Brandt (University of Reading, UK)</i>	

Web Information Management & E-commerce (Session 3B)

Link Based Clustering of Web Search Results	225
<i>Yitong Wang and Masaru Kitsuregawa (The University of Tokyo, Japan)</i>	
Three-Tier Clustering: An Online Citation Clustering System	237
<i>Haifeng Jiang, Wenwu Lou, and Wei Wang (Hong Kong University of Science and Technology, China)</i>	
Introducing Navigation Graphs as a Technique for Improving WWW User Browsing	249
<i>Christos Bouras, Agisilaos Konidaris (Computer Technology Institute, Greece), Eleni Konidari (Child Development Centre, Greece), and Afrodite Sevasti (Computer Technology Institute, Greece)</i>	
The Evolution of the Contract Net Protocol	257
<i>Lai Xu and Hans Weigand (Tilburg University, The Netherlands)</i>	

Temporal, Spatial & High Dimensional Information Management (Session 4A)

Semantic Compression of Temporal Data	267
<i>Claudio Bettini (Università di Milano, Italy)</i>	
Efficiently Computing Weighted Proximity Relationships in Spatial Databases	279
<i>Xuemin Lin, Xiaomei Zhou (University of New South Wales, Australia), Chengfei Liu (University of South Australia, Australia), and Xiaofang Zhou (University of Queensland, Australia)</i>	
Compressing the Index – A Simple and yet Efficient Approximation Approach to High-Dimensional Indexing	291
<i>Shuguang Wang, Cui Yu, and Beng Chin Ooi (National University of Singapore)</i>	

Data Mining & Constraint Management (Session 4B)

Finding Dense Clusters in Hyperspace: An Approach Based on Row Shuffling	305
<i>Daniel Barbará and Xintao Wu (George Mason University, USA)</i>	
Modelling Classification Performance for Large Data Sets (An Empirical Study)	317
<i>Baohua Gu, Feifang Hu (National University of Singapore), and Huan Liu (Arizona State University, USA)</i>	
Mapping Referential Integrity Constraints from Relational Databases to XML	329
<i>Xiaochun Yang and Guoren Wang (Northeastern University, China)</i>	

Data Integration & Filtering (Session 5A)

Attribute Value Extraction and Standardization in Data Integration	343
<i>Hongjun Lu (Hong Kong University of Science and Technology, China) and Zengping Tian (Fudan University, China)</i>	
Indexing Attributes and Reordering Profiles for XML Document Filtering and Information Delivery	355
<i>Wang Lian, David W. Cheung, and S.M. Yiu (The University of Hong Kong, China)</i>	

An Approach for Data Filtering Based on Rough Set Theory	367
<i>Xuri Yin, Zhihua Zhou, Ning Li, and Shifu Chen (Nanjing University, China)</i>	

Workflow & Adaptive Systems (Session 5B)

Web-Based Workflow Evolution in ADOME-WFMS	377
<i>Dickson K.W. Chiu (Hong Kong University of Science and Technology, China), Qing Li (City University of Hong Kong, China), and Kamalakar Karlapalem (IIT, India)</i>	
Meta-data Objects as the Basis for System Evolution	390
<i>Florida Estrella (UWE, UK), Zsolt Kovacs, Jean-Marie Le Goff (CERN, Switzerland), Richard McClatchey, and Norbert Toth (UWE, UK)</i>	
Adaptive Web Meta-search Enhanced by Constraint-Based Query Constructing and Mapping	400
<i>Lieming Huang, Ulrich Thiel, Matthias Hemmje, and Erich J. Neuhold (GMD-IPSI, Germany)</i>	
A Mixed Data Dissemination Strategy for Mobile Computing Systems ...	408
<i>Guohong Cao, Yiqiong Wu (The Pennsylvania State University, USA), and Bo Li (Xi'an Jiaotong University, China)</i>	
Author Index	417