Advances in Web-Age Information Management

5th International Conference, WAIM 2004
Dalian, China, July 15-17, 2004
Proceedings

Springer
Preface

WAIM 2004, Dalian (China), 15–17 July, was hosted by Northeastern University, China and Dalian University of Technology, China in cooperation with the Database Society of CCF. It was the International Conference on Web-Age Information Management. WAIM 2004 brought together participants from universities, industry, and government to present, discuss, and address both current issues and novel approaches in the practice, deployment, theory, and methodology of information management in the context of the World Wide Web. The conference provided an international forum for the sharing of original research results and practical development experiences among researchers and application developers from the many Web-related areas. The previous WAIM conferences were held in Shanghai (WAIM 2000), Xi’an (WAIM 2001), Beijing (WAIM 2002), and Chengdu, China (WAIM 2003). These conferences turned out to be very successful and each attracted more than 100 participants from all over the world.

The WAIM 2004 conference included technical sessions organized around important subtopics such as data stream processing, time series data processing, mobile computing, cache management, security, XML, Web services, Web search engines, query evaluation, and data mining, etc. Following careful review of the 291 submissions by members of the international program committee, 57 regular papers and 23 short/industrial papers were accepted for presentation at the conference and for publication in this volume. The submissions came from a total of 22 countries and regions, and the accepted ones were from 13 countries and regions, including Australia, Canada, Chile, China, Hong Kong SAR, Korea, Norway, Singapore, Spain, Taiwan, UAE, UK, and USA. The conference also consisted of a poster and demonstration track, the publication of which was done in CD-ROM form.

A conference such as this can only succeed as a team effort. We would like to thank program committee members and reviewers for their efforts, and the WAIM steering committee members for their invaluable input and advice. We thank the conference chairs Prof. Wang Shan and Tok Wang Ling whose involvement and support added greatly to the quality of the conference. Our sincere gratitude goes to all of the authors who submitted papers. We are grateful to our sponsors for their generous support. Special thanks go to Prof. Yu Ge who served as both Local Arrangements Co-chair and Finance and Registration Chair, Prof. Katsumi Tanaka who selected and organized high-quality tutorials, Profs. Zhiyong Peng, Arne Solvberg, and Mengchi Liu for their effective services as Publicity Co-chairs, and Drs. Mukesh Mahania and Ji-Rong Wen who organized the industrial track and the poster and demonstration track, respectively.

On behalf of the Organizing and Program Committees of WAIM 2004, we trust the attendees found the conference a fruitful experience and that they had an enjoyable stay in the beautiful city of Dalian, China.

July 2004

Qing Li, Guoren Wang, Ling Feng
Organization

WAIM 2004 was organized by Northeastern University, China and Dalian University of Technology, China in cooperation with the China Computer Federation, Database Society and ACM SIGMOD. It was sponsored by Fujitsu Ltd., Japan, Oracle Corporation, China, and SicNoah Software Technology Co., Ltd., and the National Science Foundation, China

Executive Committee

Conference Co-chairs
Wang Shan (Renmin University, China)
Tok Wang Ling (NUS, Singapore)

Program Co-chairs
Qing Li (City University, HK SAR, China)
Guoren Wang (Northeastern University, China)

Local Organization Co-chairs
Yu Ge (Northeastern University, China)
Xiukun Wang (Dalian University Tech, China)

Publication Chair
Feng Ling (University of Twente, Netherlands)

Finance and Registration Chair
Yu Ge (Northeastern University, China)

Publicity Co-chairs
Peng Zhiyong (Asia-Pacific)
Arne Solvberg (Europe)
Mengchi Liu (America)

Tutorial Chair
Katsumi Tanaka (Kyoto University, Japan)

Panel Chair
Ee-Peng Lim (NTU, Singapore)

Industrial Track Chair
Mukesh K. Mohania (IBM India)

Exhibition and Demo Chair
Ji-Rong Wen (Microsoft Research Asia)

Steering Committee Liaison
Hongjun Lu (HKUST, HK SAR, China)
Program Committee

Stephane Bressan  National University of Singapore, Singapore
Ying Cai  Iowa State University, USA
Keith Chan  Hong Kong Polytechnic University, Hong Kong SAR
Zheng Chen  Microsoft Research Asia, Beijing, China
David Cheung  University of Hong Kong, Hong Kong SAR
Dickson K.W. Chiu  Dickson Computer Systems, Hong Kong SAR
Gill Dobbie  University of Auckland, New Zealand
Xiaoyong Du  Renmin University of China, China
David Dagan Feng  University of Sydney, Australia
Jianhua Feng  Tsinghua University, China
Vivekanand Gopalkrishnan  NTU, Singapore
Hong Gao  Harbin Institute of Technology, China
Jingfeng Guo  Yanshan University, China
Yanbo Han  Institute of Computing Technology, Chinese Academy of Sciences, China
Yanxiang He  Wuhan University, China
Kien A. Hua  University of Central Florida, USA
Richard Hull  Bell Labs Research, Lucent, USA
Patrick Hung  CSIRO, Australia
Xiao Ji  Shanghai Baosight Ltd. Company, China
Yan Jia  National University of Defence Technology, China
Qingshan Jiang  Xiamen University, China
Kamal Karlapalem  International Institute of Information Technology, India
Irwin King  Chinese University of Hong Kong, Hong Kong SAR
Jiajin Le  Donghua University, China
Dik Lun Lee  HKUST, Hong Kong SAR
YoonJoon Lee  KAIST, Korea
Hong Va Leong  Hong Kong Polytechnic University, Hong Kong SAR
Jianzhong Li  Harbin Institute of Technology, China
Tianzhu Li  Hebei University, China
Zhanhuai Li  Northwestern Polytechnic University, China
Yaping Lin  Huana University, China
Huan Liu  Arizona State University, USA
Mengchi Liu  Carleton University, Canada
Wenyin Liu  City Univ. of Hong Kong, Hong Kong SAR
Weiyi Liu  YunNan University, China
Frederick H. Lochovsky  HKUST, Hong Kong SAR
Akifumi Makinouchi  Kyushu University, Japan
Dennis McLeod  University of Southern California, USA
<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xiaofeng Meng</td>
<td>Renmin University of China, China</td>
</tr>
<tr>
<td>Maria Orlowska</td>
<td>University of Queensland, Australia</td>
</tr>
<tr>
<td>Zhiyong Peng</td>
<td>Wuhan University, China</td>
</tr>
<tr>
<td>Prasan Roy</td>
<td>IBM India Research Lab, India</td>
</tr>
<tr>
<td>Junyi Shen</td>
<td>Xi’an Jiaotong University, China</td>
</tr>
<tr>
<td>Baile Shi</td>
<td>Fudan University, China</td>
</tr>
<tr>
<td>John R. Smith</td>
<td>IBM T.J. Watson Research Center, USA</td>
</tr>
<tr>
<td>Jianwen Su</td>
<td>University of California, Santa Barbara, USA</td>
</tr>
<tr>
<td>Changjie Tang</td>
<td>Sichuan University, China</td>
</tr>
<tr>
<td>Shiwei Tang</td>
<td>Peking University, China</td>
</tr>
<tr>
<td>Yufei Tao</td>
<td>City Univ. of Hong Kong, Hong Kong SAR</td>
</tr>
<tr>
<td>Wallapak Tavanapong</td>
<td>Iowa State University, USA</td>
</tr>
<tr>
<td>Duc A. Tran</td>
<td>University of Dayton, USA</td>
</tr>
<tr>
<td>Christelle Vangenot</td>
<td>Swiss Federal Institute of Technology, Switzerland</td>
</tr>
<tr>
<td>Daling Wang</td>
<td>Northeastern University, China</td>
</tr>
<tr>
<td>Min Wang</td>
<td>IBM T.J. Watson Research Center, USA</td>
</tr>
<tr>
<td>Tengjiao Wang</td>
<td>Peking University, China</td>
</tr>
<tr>
<td>Xiukun Wang</td>
<td>Dalian University of Technology, China</td>
</tr>
<tr>
<td>X. Sean Wang</td>
<td>University of Vermont, USA</td>
</tr>
<tr>
<td>Raymond Wong</td>
<td>University of New South Wales, Australia</td>
</tr>
<tr>
<td>Chunxiao Xing</td>
<td>Tsinghua University, China</td>
</tr>
<tr>
<td>Christopher C. Yang</td>
<td>Chinese University of Hong Kong, Hong Kong SAR</td>
</tr>
<tr>
<td>Dongqing Yang</td>
<td>Peking University, China</td>
</tr>
<tr>
<td>Jian Yang</td>
<td>University of Tilburg, The Netherlands</td>
</tr>
<tr>
<td>Masatoshi Yoshikawa</td>
<td>Nagoya University, Japan</td>
</tr>
<tr>
<td>Ge Yu</td>
<td>Northeastern University, China</td>
</tr>
<tr>
<td>Jeffrey X. Yu</td>
<td>Chinese University of Hong Kong, Hong Kong SAR</td>
</tr>
<tr>
<td>Lihua Yue</td>
<td>University of Science and Technology, China</td>
</tr>
<tr>
<td>Yanchun Zhang</td>
<td>Victoria University, Australia</td>
</tr>
<tr>
<td>Aoying Zhou</td>
<td>Fudan University, China</td>
</tr>
<tr>
<td>Lizhu Zhou</td>
<td>Tsinghua University, China</td>
</tr>
<tr>
<td>Shuigeng Zhou</td>
<td>Fudan University, China</td>
</tr>
<tr>
<td>Xiaofang Zhou</td>
<td>University of Queensland, Australia</td>
</tr>
<tr>
<td>Yueting Zhuang</td>
<td>Zhejiang University, China</td>
</tr>
<tr>
<td>Hai Zhuge</td>
<td>Institute of Computing Technology, Chinese Academy of Sciences, China</td>
</tr>
</tbody>
</table>
External Referees

Abhishek Verma  
Amit Mandvikar  
Bin Zhou  
Cagdas Gerede  
Cheong Fung  
Chung-yi Chang  
Daoshun Wang  
Deng Ke  
Edward Ho  
Ehtesham Haque  
Fang Ying  
Gabriel Pui  
Hoda M.O. Mokhtar  
Indra Budi  
J. Ravindranath  
James Liu  
Jhansi Rani V  
Jian Zhai  
Jianpo Ou  
Jigar Mody  
Jing Sun  
Jingyu Hou  
Jiying Wang  
Jong-eun Jun  
Joshua Pun  
Juanzi Li  
Jun He  
Ken C.K. Lee  
Krishna Prasad  
Chitrapur  
Kun Yue  
Lance R. Parsons  
Laurent Mignet  
Lei Yi  
Lei Yu  
Leila Kaghazian  
Liang Zhu  
Luo Yi  
Manish Bhide  
Mohamed Rafi  
Nadine Cullot  
Omar Boucelma  
Panagiotis Kalnis  
Qi Guo  
Qingqing Yuan  
Sai Sun  
Sang-Soo Sung  
V.R. Satyanarayana  
Seoungwook Youn  
Shan Prasher  
Shijun Li  
V. Soujanya  
Srihari Venkatesan  
Sufang Hou  
Tok Wee Hyong  
Wai-Shing Ho  
Xiaoling Wang  
WeiHong Han  
Weining Qian  
Weixia Wei  
Xiang Fu  
Xuan Tian  
Xue Li  
Xueqing Gong  
Xuhui Li  
Yang Cao  
Yang Xiaochun  
Yanyan Guo  
Yifeng Gao  
Ying Cai  
Yu Li  
Yu Wang  
Yuan Yuan  
Zeng Cheng  
Zhimao Guo  
Zhiqiang Zhang  
M. Zimmer
# Table of Contents

## Invited Talks

A Framework for a Trusted Environment for Virtual Collaboration  
*Tharam S. Dillon, Elizabeth Chang, and Farookh Hussain*

Future Interconnection Environment –
Dream, Principle, Challenge and Practice  
*Hai Zhuge*

Recent Advances in Histogram Construction Algorithms  
*Kyuseok Shim*

## Data Stream Processing

Dynamic Adjustment of Sliding Windows over Data Streams  
*Dongdong Zhang, Jianzhong Li, Zhaogong Zhang, Weiping Wang, and Longjiang Guo*

Space Efficient Quantile Summary for Constrained Sliding Windows on a Data Stream  
*Jian Xu, Xuemin Lin, and Xiaofang Zhou*

## Time Series Data Processing

The Dimension-Reduced Multi-scale Histogram: A New Method for Fast and Efficient Retrieval of Similar Time Sequences and Subsequence  
*Jian-Wei Liu, Shou-Jian Yu, and Jia-Jin Le*

Time Series Prediction Based on Gene Expression Programming  
*Jie Zuo, Chang-jie Tang, Chuan Li, Chang-an Yuan, and An-long Chen*

A Grid-Based Index Method for Time Warping Distance  
*Jiyuan An, Yi-Ping Phoebe Chen, and Eamonn Keogh*

## Security

Fail-Stop Authentication Protocol for Digital Rights Management in Adaptive Information Retrieval System  
*Zhaofeng Ma and Boqin Feng*

Effective Web-Related Resource Security Using Distributed Role Hierarchy  
*Gunhee Lee, Wonil Kim, Dong-kyoo Kim, and Hongjin Yeh*
Mobile Computing

A Low-Cost Checkpointing Scheme for Mobile Computing Systems
Guohui Li, Hongya Wang, and Jixiong Chen

Semantic and Adaptive Middleware for Data Management
in Smart Vehicle Space
Qing Wu, Zhaohui Wu, Bin Wu, and Zhou Jiang

Dynamic Resource Reservation Using the Differentiated Handoff Estimation Model for Mobile Networks
Si-Yong Park and Ki-Dong Chung

Adaptive Generic Communications for Integrated Mobile
and Internet Web-Services
Man-Ching Yuen, Leung Cheng, Pui-On Au, and Weijia Jia

Spatio-temporal Database with Multi-granularities
Sheng-sheng Wang and Da-you Liu

Cache Management

A Simple but Effective Dynamic Materialized View Caching
Chi-Hon Choi, Jeffrey Xu Yu, and Hongjun Lu

A Cache Replacement Algorithm in Hierarchical Storage
of Continuous Media Object
Yaoqiang Xu, Chunxiao Xing, and Lizhu Zhou

CCAN: Cache-Based CAN Using the Small World Model
Futai Zou, Yin Li, Liang Zhang, and Fanyuan Ma

Performance Evaluations of Replacement Algorithms
in Hierarchical Web Caching
Haohuan Fu, Pui-On Au, and Weijia Jia

Component-Based WebGIS and Its Spatial Cache Framework
Yingwei Luo, Xiaolin Wang, and Zhuoqun Xu

Query Evaluation

Learning-Based Top-N Selection Query Evaluation
over Relational Databases
Liang Zhu and Weiyi Meng

DHT Based Searching Improved by Sliding Window
Shen Huang, Gui-Rong Xue, Xing Zhu, Yan-Feng Ge, and Yong Yu
A Scalable and I/O Optimal Skyline Processing Algorithm ............... 218  
Yi Luo, Hai-Xin Lu, and Xuemin Lin

Linearization Approach for Efficient KNN Search of High-Dimensional Data ........................................... 229  
Zaheer Al Aghbari and Akifumi Makinouchi

SQL/MDR: Query Language for Consistent Access of Metadata Registries Based on SQL3 ................................. 239  
Dongwon Jeong, Young-Gab Kim, and Doo-Kwon Baik

Solving Queries over Semantically Integrated Biological Data Sources . 249  
José Francisco Aldana-Montes, Ismael Navas-Delgado,  
and María del Mar Roldán-García

Fast Approximate Search in Text Databases .......................... 259  
Fei Shi

Web Search Engine

Apply Feedback Control Theory to Design Soft Real-Time Search Engine System ..................................................... 268  
Huayong Wang and Yiqi Dai

User Interest Detection on Web Pages for Building Personalized Information Agent ................................................. 280  
Yin Liu, Wenyin Liu, and Changjun Jiang

Improved Link-Based Algorithms for Ranking Web Pages .......... 291  
Ziyang Wang

Effectiveness Evaluation and Comparison of Web Search Engines and Meta-search Engines ........................................ 303  
Shengli Wu and Jieyu Li

Semantic Search in the WWW Supported by a Cognitive Model ........ 315  
Katia Wechsler, Jorge Baier, Miguel Nussbaum,  
and Ricardo Baeza-Yates

Semantic Board: A Bulletin Board System Supporting Dynamic Semantic Information ............................................ 325  
Eui-Hyun Jung

Efficient Evaluation of Sparse Data Cubes ................................ 336  
Lixin Fu
XML

Redundancy Free Mappings from Relations to XML .................. 346
Millist W. Vincent, Jixue Liu, and Chengfei Liu

Tree Multivalued Dependencies for XML Datasets ................. 357
Lawrence V. Saxton and Xiqun Tang

TROS: A Novel Model and Its Implementation
for XML Data Management ........................................... 368
Wenwen Qi and Aoying Zhou

Optimized Query Translation Strategy for XML Stored
in Relational Database ................................................. 378
Hongzhi Wang, Jianzhong Li, and Zhenying He

QReduction: Synopsizing XPath Query Set Efficiently
under Resource Constraint ........................................... 389
Jun Gao, Xiuli Ma, Dongqing Yang, Tengjiao Wang, and Shiwei Tang

Extracting Key Value and Checking Structural Constraints
for Validating XML Key Constraints ............................... 399
Yunfeng Liu, Dongqing Yang, Shiwei Tang, Tengjiao Wang,
and Jun Gao

Estimating the Selectivity of XML Path Expression
with Predicates by Histograms ...................................... 409
Yu Wang, Haixun Wang, Xiaofeng Meng, and Shan Wang

Wrapping Web Pages into XML Documents .......................... 419
Tao Fu

Web Services

Rapid Prototyping for Web Applications ............................... 429
Chee Chern Lim, Man Hing Yu, and Jesse J. Jin

Capacity Planning for Composite Web Services
Using Queueing Network-Based Models ............................... 439
Dunlu Peng, Yang Yuan, Kun Yue, Xiaoling Wang, and Aoying Zhou

A Web-Based Coordination Infrastructure for Grid Collective Services ... 449
Javier Jaén, Jose H. Canós, and Elena Navarro

Symbolic Agent Negotiation for Semantic Web Service Exploitation ...... 458
Peep Künagas, Jinghai Rao, and Mihhail Matskin
Object-Oriented Realization of Workflow Views for Web Services –
An Object Deputy Model Based Approach .......................... 468
    Zhe Shan, Zhiyi Long, Yi Luo, and Zhiyong Peng

Modeling QoS for Semantic Equivalent Web Services .............. 478
    Derong Shen, Ge Yu, Tiezheng Nie, Rui Li, and Xiaochun Yang

Classification

Improved Email Classification through Enriched Feature Space ........ 489
    Yunming Ye, Fanyuan Ma, Hongqiang Rong,
    and Joshua Zhexue Huang

PLD: A Distillation Algorithm for Misclassified Documents .......... 499
    Ding-Yi Chen and Xue Li

Sequential Classifiers Combination for Text Categorization:
An Experimental Study .............................................. 509
    Zheng Zhang, Shuigeng Zhou, and Aoying Zhou

Combining Active Learning and Boosting
for Naive Bayes Text Classifiers .................................... 519
    Han-joon Kim and Je-uk Kim

Data Mining

Using Interval Association Rules to Identify Dubious Data Values .... 528
    Ren Lu, Mong Li Lee, and Wynne Hsu

Mining Web Sequential Patterns Incrementally
with Revised PLWAP Tree ........................................... 539
    Christie I. Ezeife and Min Chen

Mining Frequent Items in Spatio-temporal Databases ................. 549
    Cheqing Jin, Fang Xiong, Joshua Zhexue Huang, Jeffrey Xu Yu,
    and Aoying Zhou

Discovery of Frequent XML Query Patterns
with DTD Cardinality Constraints .................................... 559
    Yunfeng Liu, Dongqing Yang, Shiwei Tang, Tengjiao Wang,
    and Jun Gao

Mining Inheritance Rules from Genealogical Data .................... 569
    Yen-Liang Chen and Jing-Tin Lu

Mining DAG Patterns from DAG Databases .......................... 579
    Yen-Liang Chen, Hung-Pin Kao, and Ming-Tat Ko
Mining Class Outliers: Concepts, Algorithms and Applications ........... 589  
Zengyou He, Joshua Zhexue Huang, Xiaofei Xu, and Shengchun Deng

CD-Trees: An Efficient Index Structure for Outlier Detection .......... 600  
Huanliang Sun, Yubin Bao, Faxin Zhao, Ge Yu, and Daling Wang

Industrial/Short Papers

Mobile Data Management with Sound ................................. 610  
Ook Lee

A Mixture Model of Classical Fuzzy Classifiers ....................... 616  
Yongchuan Tang and Shouqian Sun

An Empirical Study of Building Compact Ensembles ................... 622  
Huan Liu, Amit Mandvikar, and Jigar Mody

Merging Individually Learned Optimal Results to Accelerate Coordination ........................................ 628  
Huaxiang Zhang and Shangteng Huang

The Compression of Massive Offline Relations ....................... 634  
Jizhou Luo, Jianzhong Li, Hongzhi Wang, Yanqiu Zhang, and Kai Zhao

Finding Plagiarism Based on Common Semantic Sequence Model ....... 640  
Jun-Peng Bao, Jun-Yi Shen, Xiao-Dong Liu, Hai-Yan Liu,  
and Xiao-Di Zhang

Web Information Extraction Based on Similar Patterns ................ 646  
Na Ye, Xuejun Wu, Jingbo Zhu, Wenliang Chen, and Tianshun Yao

TS-Cache: A Novel Caching Strategy to Manage Data in MANET Database ........................................... 652  
Shengfei Shi, Jianzhong Li, Chaokun Wang, and Jinbao Li

Extended Chain-Conflicting Serializability for the Correct Schedule of Transactions in Hierarchical Multidatabase .................. 658  
Guoning Chen and Taoshen Li

An Efficient Hierarchical Failure Recovery Algorithm Ensuring Semantic Atomicity for Workflow Applications .................. 664  
Yi Ren, Quanyuan Wu, Yan Jia, and Jianbo Guan

A Macrocommittees Method of Combining Multistrategy Classifiers for Heterogeneous Ontology Matching ...................... 672  
Leyun Pan, Hui Song, and Fanyuan Ma
Hypertext Classification Algorithm  
Based on Co-weighting Multi-information ............................ 678  
Ya Peng, Ya-ping Lin, and Zhi-ping Chen

Verifying a MAC-Based Service Bundle Authentication Mechanism for the OSGi Service Platform ............................... 684  
Young-Gab Kim, Dongwon Jeong, and Doo-Kwon Baik

Optimal Data Dispatching Methods in Near-Line Tertiary Storage System ........................................... 690  
Baoliang Liu, Jianzhong Li, and Yanqi Zhang

A Real-Time Information Gathering Agent Based on Ontology ........ 696  
Jianqing Li, Shengping Liu, ZuoQuan Lin, and Cen Wu

Dynamical Schedule Algorithms Based on Markov Model in Tertiary Storage ........................................... 702  
Yanqiu Zhang, Jianzhong Li, Zhaogong Zhang, and Baoliang Liu

Image Retrieval Using Structured Logical Shape Feature .......... 708  
Nanhyo Bang and Kyhyun Um

Automatic HTML to XML Conversion ................................ 714  
Shijun Li, Mengchi Liu, Tok Wang Ling, and Zhiyong Peng

Quality Improvement Modeling and Practice in Baosteel Based on KIV-KOV Analysis ........................................... 720  
Xinyu Liu, Haidong Tang, Zhiping Fan, and Bing Deng

A Frequent Pattern Discovery Method for Outlier Detection ........ 726  
Zengyou He, Xiaofei Xu, Joshua Zhexue Huang, and Shengchun Deng

Adaptive Load Balancing and Fault Tolerance in Push-Based Information Delivery Middleware Service ........................................... 733  
Zhaofeng Ma and and Boqin Feng

Performance Optimization of Fractal Dimension Based Feature Selection Algorithm ........................................... 739  
Yubin Bao, Ge Yu, Huanliang Sun, and Daling Wang

Graph-Based Cyclic Multi-join Query Optimization Algorithm .......... 745  
Hong Yu, Xiu-Kun Wang, and Jian-Ying Zhang

Author Index .................................................... 751