Lecture Notes in Computer Science

7235

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

Alfred Kobsa

University of California, Irvine, CA, 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

TU Dortmund University, Germany

Madhu Sudan

Microsoft Research, Cambridge, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbruecken, Germany

Quan Z. Sheng Guoren Wang Christian S. Jensen Guandong Xu (Eds.)

Web Technologies and Applications

14th Asia-Pacific Web Conference, APWeb 2012 Kunming, China, April 11-13, 2012 Proceedings



Volume Editors

Quan Z. Sheng

The University of Adelaide, School of Computer Science

Adelaide, SA 5005, Australia

E-mail: qsheng@cs.adelaide.edu.au

Guoren Wang

Northeastern University, College of Information Science and Engineering

Shenyang 110819, China

E-mail: wanggr@mail.neu.edu.cn

Christian S. Jensen

Aarhus University, Department of Computer Science

8200 Aarhus N, Denmark

E-mail: csj@cs.au.dk

Guandong Xu

Victoria University, Centre for Applied Informatics

Melbourne City, VIC 8001, Australia

E-mail: guandong.xu@vu.edu.au

ISSN 0302-9743 ISBN 978-3-642-29252-1

e-ISSN 1611-3349 e-ISBN 978-3-642-29253-8

DOI 10.1007/978-3-642-29253-8

Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012934199

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

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web

and HCI

© Springer-Verlag Berlin Heidelberg 2012

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.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Message from the General Chairs

Welcome to APWeb 2012, the 14th Edition of Asia Pacific Web Technology Conference. APWeb is a leading international conference on research, development, and applications of Web technologies, database systems, information management, and software engineering, with a focus on the Asia-Pacific region. Previous APWeb conferences were held in Beijing (2011), Busan (2010), Suzhou (2009), Shenyang (2008), Huangshan (2007), Harbin (2006), Shanghai (2005), Hangzhou (2004), Xi'an (2003), Changsha (2001), Xi'an (2000), Hong Kong (1999), and Beijing (1998).

APWeb conferences cover contemporary topics in the fields of Web management and World Wide Web-related research and applications, such as advanced application of databases, cloud computing, content management, data mining and knowledge discovery, distributed and parallel processing, grid computing, Internet of Things, Semantic Web and Web ontology, security, privacy and trust, sensor networks, service-oriented computing, Web community analysis, Web mining and social networks.

The APWeb 2012 program featured a main conference and five satellite workshops. The main conference had two eminent keynote speakers, Patrick McDaniel from Pennsylvania State University, USA, and Vijay Varadharajan from Macquarie University, Australia, 39 full research papers, 34 short research papers, and 5 demonstration papers. The four workshops reported recent developments and advances in the contemporary topics in the related fields of:

- Information Extraction and Knowledge Base Building (IEKB)
- Intelligent Data Processing (IDP 2012)
- Sensor Networks and Data Engineering (SenDe 2012)
- Mobile Business Collaboration (MBC 2012)

Both main conference program and workshop sessions were of high quality owing to the strong support and commitment from their international Program Committee. We wish to take this opportunity to thank Program Co-chairs Michael Sheng, Guoren Wang, and Christian S. Jensen, for their dedication and effort in ensuring a high-quality program. We would also like to thank Workshop Co-chairs Hua Wang and Lei Zou, and each workshop organizers for their contributions to developing an interesting and attractive workshop program. Many colleagues helped toward the success of APWeb 2012. They are: Local Arrangements Co-chairs: Jianfeng He and Guangyan Huang; Financial Chair: Jing He; Publication Chair: Guandong Xu; Industry Co-chairs: Gary Morgan and Qingzhong Li; Demonstration Chair: Chaoyi Pang; Publicity Co-chairs: Haolan Zhang and Jiangang Ma, and Webmaster: Zhi Qiao and Zhangwei Jiang.

We would like to sincerely thank our financial supporters and sponsors. The following organizations generously supported and sponsored APWeb 2012:

Message from the General Chairs

Hebei University of Engineering, Nanjing University of Finance and Economics, National Science Foundation of China, Kunming University of Science and Technology, Graduate University of Chinese Academy of Science, and Victoria University.

Finally, we wish to thank the APWeb Steering Committee, led by Xuemin Lin, for offering the opportunity to organize APWeb 2012 in Kunming. We also wish to thank the host organization Kunming University of Science and Technology and local arrangements committee and volunteers for the assistance in organizing this conference. The following members helped with the registration, accommodation, and various logistics: Jing Yang, Xun Zhou, Fenhua Li, and Shang Hu.

February 2012

VI

Yaoxue Zhang Yanchun Zhang Masaru Kitsuregawa

Message from the Program Chairs

The relentless growth in Internet functionality and broadband access has enabled a new wave of innovations that is transforming the way people and organizations interact, communicate, and collaborate. APWeb 2012 built on the tradition of 13 successful previous editions of the conference and aimed at presenting original contributions related to all fields of World Wide Web-related research, development, and applications.

The call for papers generated substantial interest. A total of 167 paper submissions were received, which represents an increase of over 60% compared to APWeb 2011. The paper submissions, from 17 countries and regions, were reviewed carefully by members of the Program Committee, which comprised 104 experts from 20 countries. Based on the reviews, 39 submissions were selected as full papers, yielding an acceptance rate of 23%. In addition, 34 submissions were selected as short papers. The selected papers covered a wide variety of important topics related to the World Wide Web, including Web mining, service-oriented computing, XML and semi-structured query processing, social networks, Web information extraction, and cloud computing. Authors of a few selected papers will be invited to submit extended versions to a special issue of the World Wide Web journal, published by Springer.

The conference program featured two keynote addresses given by Patrick Mc-Daniel, Editor-in-Chief of *ACM Transactions on Internet Technology*, and Vijay Varadharajan, Microsoft Chair in Innovation in Computing at Macquarie University, Australia. The program encompassed five demonstrations, which were selected by a separate committee chaired by Chaoyi Pang.

The program also contained a collection of papers that were presented in affiliated workshops and that were published in a separate LNCS volume, including the International Workshop on Intelligent Data Processing (IDP 2012) organized by Chaoyi Pang, Junhu Wang, and Haolan Zhang, the Workshop on Information Extraction and Knowledge Base Building (IEKB 2012) organized by Dongyan Zhao, the Workshop on Sensor Networks and Data Engineering (SenDe 2012) organized by Jing He, Guangyan Huang, and Xiedong Cao, and the Third International Workshop on Mobile Business Collaboration (MBC 2012) organized by Dickson Chiu, Yi Zhuang, and Zhiang Wu.

We thank all authors for their submissions and the Program Committee members and external reviewers for their excellent work. We hope that you find the papers in the proceedings interesting and stimulating.

February 2012

Quan Z. Sheng Guoren Wang Christian S. Jensen

Organization

Honorary General Chair

Yaoxue Zhang Central South University, China

General Co-chairs

Yanchun Zhang Victoria University, Australia Masaru Kitsuregawa The University of Tokyo, Japan

Program Co-chairs

Quan Z. Sheng The University of Adelaide, Australia Guoren Wang Northeastern University, China Christian S. Jensen Aarhus University, Denmark

Industry Co-chairs

Gary Morgan CSIRO, Australia

Qingzhong Li Shandong University, China

Workshop Co-chairs

Hua Wang University of Southern Queensland, Australia

Lei Zou Peking University, China

Local Arrangements Chairs

Jianfeng He Kunming University of Science and Technology,

China

Guangyan Huang Victoria University, Australia

Financial Chair

Jing He Victoria University, Australia

Publication Chair

Guandong Xu Victoria University, Australia

Demo Co-chair

Chaoyi Pang Australia e-Health Research Centre, Australia

Publicity Co-chairs

Haolan Zhang NIT, Zhejiang University, China Jiangang Ma The University of Adelaide, Australia

APWeb Steering Committee Liaison

Xuemin Lin The University of New South Wales, Australia

WISE Steering Committee Liaison

Qing Li City University of Hong Kong, China

Steering Committee

Xiaoming Li Peking University, China

Xuemin Lin University of New South Wales, Australia Maria Orlowska Ministry of Science and Higher Education,

Poland

Kyu-Young Whang Korea Advanced Institute of Science and

Technology, Korea

Jeffrey Yu Chinese University of Hong Kong, China

Yanchun Zhang Victoria University, Australia
Xiaofang Zhou University of Queensland, Australia

Program Committee

Toshiyuki Amagasa University of Tsukuba, Japan Denilson Barbosa University of Alberta, Canada Pablo Barcelo Universidad de Chile, Chile Djamal Benslimane University of Lyon, France Geert Jan Bex Hasselt University, Belgium

Stephane Bressan Singapore National University, Singapore Jae-Woo Chang Chonbuk National University, Korea

Wenguang Chen Peking University, China Mauro Caporuscio Politecnico di Milano, Italy

Haiming Chen
Chinese Academy of Sciences, China
Hong Chen
Renmin University of China, China
Hanxiong Chen
University of Tsukuba, Japan
Jian Chen
Shaanxi Normal University, China
Renmin University of China, China

Jin Chen Michigan State University, USA
Reynold Cheng The University of Hong Kong, China
David Cheung The University of Hong Kong, China
Dishard Care and Chemical University of Startholds LW

Richard Connor University of Strathclyde, UK

Alfredo Cuzzocrea ICAR-CNR and University of Calabria, Italy

Ke Deng University of Queensland, Australia

Ting Deng Beihang University, China Paolo Falcarin University of East London, UK Ling Feng Tsinghua University, China

Li Guo Chinese Academy of Science, China Xueqing Gong East China Normal University, China

Rajeev Gupta IBM Research, India Qi He IBM Research, USA Sven Helmer University of London, UK

Zi Huang University of Queensland, Australia

Bin Hu Lanzhou University, China Yoshiharu Ishikawa Nagoya University, Japan

Jin-Ho Kim Kangwon National University, Korea

Kyong-Ho Lee Yonsei University, Korea SangKeun Lee Korea University, Korea Sang Ho Lee Soongsil University, Korea

Qing Li Hong Kong City University, China Yingshu Li Georgia State University, USA Yinsheng Li Fudan University, China

Xitong Li MIT, USA

Xue Li University of Queensland, Australia

Chengfei Liu Swinburne University of Technology, Australia

Mengchi Liu Carleton University, Canada
Jiaheng Lu Renmin University, China
Hua Lu Aalborg University, Demark
Qiang Ma Kyoto University, Japan
Shuai Ma Beihang University, China
Jiangang Ma University of Adelaide, Australia
Zaki Malik Wayne State University, USA

Zakaria Maamar Zayed University, UAE

Hamid Motahari HP Lab, USA

Weiyi Meng State University of New York at Binghamton,

USA

Azzam Mourad Lebanese American University, Lebanon Michael Mrissa University of Lyon, France Shinsuke Nakajima Kyoto Sangyo University, Japan Miyuki Nakano University of Tokyo, Japan

Joachim Niehren INRIA Lille Nord Europe, France

Yuan Ni IBM Research, China

Werner Nutt Free University of Bozen-Bolzano, Italy

Satoshi Oyama Hokkaido University, Japan

XII Organization

Helen Paik University of New South Wales, Australia

Chaoyi Pang CSIRO. Australia Sanghyun Park Yonsei University, Korea Zhivong Peng Wuhan University, China Alex Poulovassilis University of London, UK Florian Rosenberg IBM Research, USA

Caspar Ryan RMIT University, Australia

KeunHo Rvu Chungbuk National University, Korea Marc Scholl Universitaet Konstanz, Germany

Aviv Segev KAIST, Korea

Mohamed Sharaf University of Queensland, Australia Claudia Szabo University of Adelaide, Australia

Kazutoshi Sumiya University of Hyogo, Japan

Aixin Sun Nanyang Technological University, Singapore

Alex Thomo University of Victoria, Canada Tsinghua University, China Chaokun Wang Daling Wang Northeastern University, China Hongzhi Wang Harbin Institute of Technology, China

Hua Wang University of Southern Queensland, Australia

Fudan University, China Wei Wang

East China Normal University, China Xiaoling Wang Yinglin Wang Shanghai Jiaotong University, China Jef Wijsen University of Mons-Hainaut, Belgium

Peter Wood University of London, UK

Yanbo Wu University of Adelaide, Australia

Yuqing Wu Indiana University, USA

Jianliang Xu Hong Kong Baptist University, China

Linhao Xu IBM Research, China

Bin Yang Max Planck Institute, Germany Xiaochun Yang Northeastern University, China Hamdi Yahyaoui Kuwait University, Kuwait Sun Yat-Sen University, China Jian Yin Muhammad Younas Oxford Brookes University, UK

Haruo Yokota Tokyo Institute of Technology, Japan

Jian Yu Swinburne University of Technology, Australia

Philip Yu University of Illinois at Chicago, USA

Ming Zhang Beijing University, China

Rui Zhang The University of Melbourne, Australia University of New South Wales, Australia Ying Zhang University of New South Wales, Australia Wenjie Zhang Renmin University of China, China

Xiao Zhang

Zhenjie Zhang ADSC, Singapore Shuigeng Zhou Fudan University, China

Xiangmin Zhou CSIRO, Australia

Xuan Zhou Renmin University, China

Zhangbing Zhou Institute TELECOM SudParis, France

Table of Contents

Keynotes	
----------	--

Scalable Integrity-Guaranteed AJAX	1
Security and Trust in the Web	20
Full Papers	
Estimate Unlabeled-Data-Distribution for Semi-supervised PU Learning	22
Disputed Sentence Suggestion towards Credibility-Oriented Web Search	34
Exploration and Visualization of Administrator Network in Wikipedia	46
An Efficient Algorithm for Arbitrary Reverse Furthest Neighbor Queries	60
Memory-Aware BWT by Segmenting Sequences to Support Subsequence Search	73
Mining Frequent Association Tag Sequences for Clustering XML Documents	85
Context-Aware Personalized Search Based on User and Resource Profiles in Folksonomies	97
Distance-Based Outlier Detection on Uncertain Data of Gaussian Distribution	109

Extracting Keyphrase Set with High Diversity and Coverage Using Structural SVM	122
Discovering the Most Potential Stars in Social Networks with Infra-skyline Queries	134
Feature Based Informative Model for Discriminating Favorite Items from Unrated Ones	146
Search for Minority Information from Wikipedia Based on Similarity of Majority Information	158
An Incremental Approach to Analyzing Temporal Constraints of Workflow Processes	170
Keywords Filtering over Probabilistic XML Data	183
H-Tree: A Hybrid Structure for Confidence Computation in Probabilistic Databases	195
Predicting Online Auction Final Prices Using Time Series Splitting and Clustering	207
Taxonomy-Oriented Recommendation towards Recommendation with Stage	219
A GPU-Based Accelerator for Chinese Word Segmentation	231
The Equi-Join Processing and Optimization on Ring Architecture Key/Value Database	243
Efficient Probabilistic Image Retrieval Based on a Mixed Feature Model	255

Improve Top-K Recommendation by Extending Review Analysis	435
CDDTA-JOIN: One-Pass OLAP Algorithm for Column-Oriented Databases	448
A Study of the Single Point Mutation Loci in the Hepatitis B Virus Sequences via Optimal Risk and Preventive Sets with Weights Qi Zhang, Junpeng Zhang, Jianmei Gao, Jianfeng He, Xinmin Yan, Lei Ma, Xianwen Zhang, and Jiuyong Li	460
Enforcing Interaction and Cooperation in Content-Based Web3.0	450
Applications	472
Adaptive Topic Community Tracking in Social Network Zheng Liang, Yan Jia, and Bin Zhou	484
Short Papers	
Extracting Difference Information from Multilingual Wikipedia Yuya Fujiwara, Yu Suzuki, Yukio Konishi, and Akiyo Nadamoto	496
Multi-strategic Approach of Fast Composition of Web Services Gang Wang, Li Zhang, and Kunming Nie	504
Collaborative Filtering via Temporal Euclidean Embedding Li'ang Yin, Yongqiang Wang, and Yong Yu	513
Transfer Learning with Local Smoothness Regularizer	521
Scalable Complex Event Processing on Top of MapReduce	529
Learning to Recommend Based on Slope One Strategy Yongqiang Wang, Li'ang Yin, Bing Cheng, and Yong Yu	537
Dataflow Optimization for Service-Oriented Applications	545
Group-Scope Query and Its Access Method	552
Introducing SaaS Capabilities to Existing Web-Based Applications	.
Automatically	560

Computing Resource Prediction for MapReduce Applications Using Decision Tree	570
Jing Tai Piao and Jun Yan	
Who Resemble You Better, Your Friends or Co-visited Users Jinjing Ma and Yan Zhang	578
P2P-Based Publication and Sharing of Axioms in OWL Ontologies for SPARQL Query Processing in Distributed Environment	586
When Sparsity Meets Noise in Collaborative Filtering	594
Data Sparsity: A Key Disadvantage of User-Based Collaborative Filtering?	602
Path Skyline for Moving Objects	610
A Graphical Audit Facility for Data Processing and Its Evaluation with Users	618
Efficient SPARQL Query Processing in MapReduce through Data Partitioning and Indexing	628
An Entity Class Model Based Correlated Query Path Selection Method in Multiple Domains	636
ℓ^1 -Graph Based Community Detection in Online Social Networks Liang Huang, Ruixuan Li, Yuhua Li, Xiwu Gu, Kunmei Wen, and Zhiyong Xu	644
Improving Recommendation Based on Features' Co-occurrence Effects in Collaborative Tagging Systems	652
Memory Performance Prediction of Web Server Applications Based on Grey System Theory	660
Performance Optimization of Analysis Rules in Real-Time Active Data Warehouses	669

Efficient Retrieval of Similar Workflow Models Based on Behavior Tao Jin, Jianmin Wang, and Lijie Wen	677
Scalable Subspace Logistic Regression Models for High Dimensional Data	685
Verifying Location-Based Services with Declassification Enforcement Cong Sun, Sheng Gao, and Jianfeng Ma	695
Single-Hop Friends Recommendation and Verification Based Incentive for BitTorrent	703
A Compact XML Storage Scheme Supporting Efficient Path Querying	711
Self-supervised Learning Approach for Extracting Citation Information on the Web	719
Complete-Thread Extraction from Web Forums	727
Characterizing Topic-Specific Hashtag Cascade in Twitter Based on Distributions of User Influence	735
A Study on Modeling of Lightweight Scientific Workflow Systems Using XML Schema	743
Hit Count Reliability: How Much Can We Trust Hit Counts?	751
A Topological Description Language for Agent Networks	759
Sentiment Analysis for Effective Detection of Cyber Bullying	767
Demo Papers	
Bizard: An Online Multi-dimensional Data Analysis Visualization Tool	775

Table of Contents	XIX
Jingwei+: A Distributed Large-Scale RDF Data Server	779
Baby Careware: A Online Secured Health Consultant	784
Mobile Applications towards Prevention and Management of Chronic Diseases	788
A Healthcare Information System with Augmented Access Controls Nagajyothi Gunti, Weiqing Sun, Mingzhe Xu, Zidong Liu, Mohammed Niamat, and Mansoor Alam	792
Author Index	797