

Lecture Notes in Artificial Intelligence 7302

Subseries of Lecture Notes in Computer Science

LNAI Series Editors

Randy Goebel

University of Alberta, Edmonton, Canada

Yuzuru Tanaka

Hokkaido University, Sapporo, Japan

Wolfgang Wahlster

DFKI and Saarland University, Saarbrücken, Germany

LNAI Founding Series Editor

Joerg Siekmann

DFKI and Saarland University, Saarbrücken, Germany

Pang-Ning Tan Sanjay Chawla
Chin Kuan Ho James Bailey (Eds.)

Advances in Knowledge Discovery and Data Mining

16th Pacific-Asia Conference, PAKDD 2012
Kuala Lumpur, Malaysia, May 29 – June 1, 2012
Proceedings, Part II

 Springer

Series Editors

Randy Goebel, University of Alberta, Edmonton, Canada
Jörg Siekmann, University of Saarland, Saarbrücken, Germany
Wolfgang Wahlster, DFKI and University of Saarland, Saarbrücken, Germany

Volume Editors

Pang-Ning Tan
Michigan State University, Department of Computer Science and Engineering
428 S. Shaw Lane, 48824-1226 East Lansing, MI, USA
E-mail: ptn@cse.msu.edu

Sanjay Chawla
University of Sydney, School of Information Technologies
1 Cleveland St., 2006 Sydney, NSW, Australia
E-mail: sanjay.chawla@sydney.edu.au

Chin Kuan Ho
Multimedia University, Faculty of Computing and Informatics
Jalan Multimedia, 63100 Cyberjaya, Selangor, Malaysia
E-mail: ckho@mmu.edu.my

James Bailey
The University of Melbourne, Department of Computing and Information Systems
111 Barry Street, 3053 Melbourne, VIC, Australia
E-mail: baileyj@unimelb.edu.au

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-30219-0 e-ISBN 978-3-642-30220-6
DOI 10.1007/978-3-642-30220-6
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012937031

CR Subject Classification (1998): I.2, H.3, H.4, H.2.8, C.2, J.1

LNCS Sublibrary: SL 7 – Artificial Intelligence

© 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)

Preface

PAKDD 2012 was the 16th conference of the Pacific Asia Conference series on Knowledge Discovery and Data Mining. For the first time, the conference was held in Malaysia, which has a vibrant economy and an aspiration to transform itself into a knowledge-based society. Malaysians are also known to be very active in social media such as Facebook and Twitter. Many private companies and government agencies in Malaysia are already adopting database and data warehousing systems, which over time will accumulate massive amounts of data waiting to be mined. Having PAKDD 2012 organized in Malaysia was therefore very timely as it created a good opportunity for the local data professionals to acquire cutting-edge knowledge in the field through the conference talks, tutorials and workshops.

The PAKDD conference series is a meeting place for both university researchers and data professionals to share the latest research results. The PAKDD 2012 call for papers attracted a total of 241 submissions from 32 countries in all six continents (Asia, Europe, Africa, North America, South America, and Australasia), of which 20 (8.3%) were accepted for full presentation and 66 (27.4%) were accepted for short presentation. Each submitted paper underwent a rigorous double-blind review process and was assigned to at least four Program Committee (PC) members. Every paper was reviewed by at least three PC members, with nearly two-thirds of them receiving four reviews or more. One of the changes in the review process this year was the adoption of a two-tier approach, in which a senior PC member was appointed to oversee the reviews for each paper. In the case where there was significant divergence in the review ratings, the senior PC members also initiated a discussion phase before providing the Program Co-chairs with their final recommendation. The Program Co-chairs went through each of the senior PC members' recommendations, as well as the submitted papers and reviews, to come up with the final selection. We thank all reviewers (Senior PC, PC and external invitees) for their efforts in reviewing the papers in a timely fashion (altogether, more than 94% of the reviews were completed by the time the notification was sent). Without their hard work, we would not have been able to see such a high-quality program.

The three-day conference program included three keynote talks by world-renowned data mining experts, namely, Chandrakant D. Patel from HP Labs (*Joules of Available Energy as the Global Currency: The Role of Knowledge Discovery and Data Mining*); Charles Elkan from the University of California at San Diego (*Learning to Make Predictions in Networks*); and Ian Witten from the University of Waikato (*Semantic Document Representation: Do It with Wikification*). The program also included four workshops, three tutorials, a doctoral symposium, and several paper sessions. Other than these intellectually inspiring events, participants of PAKDD 2012 were able to enjoy several social events

throughout the conference. These included a welcome reception on day one, a banquet on day two and a free city tour on day three. Finally, PAKDD 2012 organized a data mining competition for those who wanted to lay their hands on mining some real-world datasets.

Putting a conference together with a scale like PAKDD 2012 requires tremendous efforts from the organizing team as well as financial support from the sponsors. We thank Takashi Washio, Jun Luo and Hui Xiong for organizing the workshops and tutorials, and coordinating with the workshop/tutorial organizers/speakers. We also owe James Bailey a big thank you for preparing the conference proceedings. Finally, we had a great team of Publicity Co-chairs, Local Organization Co-chairs, and helpers. They ensured the conference attracted many local and international participants, and the conference program proceeded smoothly.

We would like to express our gratitude to SAS, AFOSR/AOARD (Air Force Office of Scientific Research/Asian Office of Aerospace Research and Development), MDeC (Multimedia Development Corporation), PIKOM (Computer Industry Association of Malaysia) and other organizations for their generous sponsorship and support. We also wish to thank the PAKDD Steering Committee for offering the student travel support grant and the grant for the best student paper award(s), and UTAR and MMU for providing the administrative support.

Philip Yu
Ee-Peng Lim
Hong-Tat Ewe
Pang-Ning Tan
Sanjay Chawla
Chin-Kuan Ho

Organization

Organizing Committee

Conference Co-chairs

Philip Yu	University of Illinois at Chicago, USA
Hong-Tat Ewe	Universiti Tunku Abdul Rahman, Malaysia
Ee-Peng Lim	Singapore Management University, Singapore

Program Co-chairs

Pang-Ning Tan	Michigan State University, USA
Sanjay Chawla	The University of Sydney, Australia
Chin-Kuan Ho	Multimedia University, Malaysia

Workshop Co-chairs

Takashi Washio	Osaka University, Japan
Jun Luo	Shenzhen Institute of Advanced Technology, China

Tutorial Co-chair

Hui Xiong	Rutgers University, USA
-----------	-------------------------

Local Organization Co-chairs

Victor Tan	Universiti Tunku Abdul Rahman, Malaysia
Wen-Cheong Chin	Multimedia University, Malaysia
Soung-Yue Liew	Universiti Tunku Abdul Rahman, Malaysia

Publicity Co-chairs

Rui Kuang	University of Minnesota, USA
Ming Li	Nanjing University, China
Myra Spiliopoulou	University of Magdeburg, Germany

Publication Chair

James Bailey	University of Melbourne, Australia
--------------	------------------------------------

Local Arrangements Committee

Soung Yue Liew (Co-chair)	Victor Tan (Co-chair)
Wen Cheong Chin (Co-chair)	Kok Why Ng (Co-chair)
Nadim Jahangir	Choo Yee Ting
Chee Onn Wong	Chiung Ching Ho
Chong Pei Fen	Hau Lee Tong
Timothy Yap	James Ooi
Kok Leong Chan	Yong Haur Tay
Azurawati	Chian Wen Too
Khong Leng Lim	Mariam
Michelle	Meei Hao Hoo
Kean Vee Sor	Priya
Madhavan	Simon Lau
Chin Chwee Wong	Swee Ling Chean

Steering Committee

Co-chairs

Graham Williams	Australian National University, Australia
Tu Bao Ho	Japan Advanced Institute of Science and Technology, Japan

Life Members

Hiroshi Motoda	AFOSR/AOARD and Osaka University, Japan
Rao Kotagiri	University of Melbourne, Australia
Ning Zhong	Maebashi Institute of Technology, Japan
Masaru Kitsuregawa	Tokyo University, Japan
David Cheung	University of Hong Kong, China
Graham Williams	Australian National University, Australia
Ming-Syan Chen	National Taiwan University, Taiwan, ROC

Members

Huan Liu	Arizona State University, USA
Kyu-Young Whang	Korea Advanced Institute of Science and Technology, Korea
Chengqi Zhang	University of Technology Sydney, Australia
Tu Bao Ho	Japan Advanced Institute of Science and Technology, Japan
Ee-Peng Lim	Singapore Management University, Singapore
Jaideep Srivastava	University of Minnesota, USA
Zhi-Hua Zhou	Nanjing University, China
Takashi Washio	Institute of Scientific and Industrial Research, Osaka University
Thanaruk Theeramunkong	Thammasat University, Thailand

P. Krishna Reddy	International Institute of Information Technology, Hyderabad (IIIT-H), India
Joshua Z. Huang	Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China

Senior Program Committee

Anirban Dasgupta	Yahoo! Research Silicon Valley, USA
Arno Siebes	Universiteit Utrecht, The Netherlands
Bart Goethals	University of Antwerp, Belgium
Bernhard Pfahringer	The University of Waikato, New Zealand
Dacheng Tao	Nanyang Technological University, Singapore
Ee-Peng Lim	Singapore Management University, Singapore
Haixun Wang	Microsoft Research Asia, China
Hisashi Kashima	University of Tokyo, Japan
Jeffrey Xu Yu	The Chinese University of Hong Kong, Hong Kong
Jian Pei	Simon Fraser University, Canada
Jianyong Wang	Tsinghua University, China
Jiuyong Li	University of South Australia, Australia
Kyuseok Shim	Seoul National University, Korea
Masashi Sugiyama	Tokyo Institute of Technology, Japan
Ng Wee Keong	Nanyang Technological University, Singapore
Nitesh V. Chawla	University of Notre Dame, USA
Osmar R. Zaiane	University of Alberta, Canada
Panagiotis Karras	Rutgers University, USA
Peter Christen	The Australian National University, Australia
Sameep Mehta	IBM Research, India
Sanjay Ranka	University of Florida, USA
Shivani Agarwal	Indian Institute of Science, India
Wei Wang	University of North Carolina at Chapel Hill, USA
Yu Zheng	Microsoft Research Asia, China

Program Committee

Aditya Krishna Menon	University of California, USA
Aixin Sun	Nanyang Technological University, Singapore
Akihiro Inokuchi	Osaka University, Japan
Albrecht Zimmerman	Katholieke Universiteit Leuven, Belgium
Alexandre Termier	Université Joseph Fourier, France
Alfredo Cuzzocrea	ICAR-CNR and University of Calabria, Italy
Amol Ghoting	IBM T.J. Watson Research Center, USA
Andreas Hotho	University of Kassel, Germany
Andrzej Skowron	University of Warsaw, Poland
Annalisa Appice	Università degli Studi di Bari, Italy

Anne Denton	North Dakota State University, USA
Anne Laurent	Montpellier 2 University, France
Aoying Zhou	East China Normal University, Shanghai, China
Arbee Chen	National Chengchi University, Taiwan
Aristides Gionis	Yahoo! Research, Spain
Aryya Gangopadhyay	University of Maryland, USA
Atsuhiko Takasu	National Institute of Informatics, Japan
Atsuyoshi Nakamura	Hokkaido University, Japan
Benjamin C.M. Fung	Concordia University, Canada
Bettina Berendt	Katholieke Universiteit Leuven, Belgium
Bo Zhang	Tsinghua University, China
Bradley Malin	Vanderbilt University, USA
Bruno Cremilleux	Université de Caen, France
Chandan Reddy	Wayne State University, USA
Chang-Tien Lu	Virginia Polytechnic Institute and State University, USA
Charles Ling	The University of Western Ontario, Canada
Chengkai Li	The University of Texas at Arlington, USA
Chengqi Zhang	University of Technology, Australia
Chiranjib Bhattachar	Indian Institute of Science, India
Choochart Haruechaiy	National Electronics and Computer Technology Center (NECTEC), Thailand
Chotirat Ratanamatan	Chulalongkorn University, Thailand
Chunsheng Yang	Institute for Information Technology, Canada
Clement Yu	University of Illinois at Chicago, USA
Daisuke Ikeda	Kyushu University, Japan
Dan Simovici	University of Massachusetts Boston, USA
Dao-Qing Dai	Sun Yat-Sen University, China
Daoqiang Zhang	Nanjing University of Aeronautics and Astronautics, China
David Albrecht	Monash University, Australia
David Taniar	Monash University, Australia
David Lo	Singapore Management University, Singapore
David F. Gleich	Purdue University, USA
Davood Rafiei	University of Alberta, Canada
Deept Kumar	Virginia Polytechnic Institute and State University, USA
Dejing Dou	University of Oregon, USA
Di Wu	Polytechnic Institute of NYU, USA
Diane Cook	Washington State University, USA
Diansheng Guo	University of South Carolina, USA
Dragan Gamberger	Rudjer Boskovic Institute, Croatia
Du Zhang	California State University, USA
Efstratios Gallopoulos	University of Patras, Greece
Elena Baralis	Politecnico di Torino, Italy

Eyke Huellermeier	University of Marburg, Germany
Fabrizio Silvestri	Istituto di Scienza e Tecnologie dell'Informazione (ISTI), Italy
Feifei Li	Florida State University, USA
Florent Massegia	INRIA, France
Fosca Giannotti	Università di Pisa, Italy
Francesco Bonchi	Yahoo! Research, Spain
Frans Coenen	University of Liverpool, UK
Gang Li	Deakin University, Australia
Gao Cong	Nanyang Technological University, Singapore
George Karypis	University of Minnesota, USA
Giuseppe Manco	Università della Calabria, Italy
Graham Williams	Australian Taxation Office, Australia
Hady Lauw	Institute for Infocomm Research, Singapore
Haibin Cheng	Yahoo! Labs, USA
Haimonti Dutta	Columbia University, USA
Hanghang Tong	IBM T.J. Watson Research Center, USA
Harry Zhang	University of New Brunswick, Canada
Hassab Elgawi Osman	University of Tokyo, Japan
Hideo Bannai	Kyushu University, Japan
Hiroyuki Kawano	Nanzan University, Japan
Hong Cheng	The Chinese University of Hong Kong, Hong Kong
Hua Lu	Aalborg University, Denmark
Huan Liu	Arizona State University, USA
Hui Wang	University of Ulster, UK
Huidong Jin	Chinese University of Hong Kong, Hong Kong
Ioannis Androulakis	Rutgers University, USA
Irena Koprinska	University of Sydney, Australia
Ivor Tsang	The Hong Kong University of Science and Technology, Hong Kong
Jaakko Hollmen	Aalto University, Finland
James Caverlee	Texas A&M University, USA
Jason Wang	New Jersey's Science and Technology University, USA
Jean-Francois Boulicaut	Université de Lyon, France
Jean-Marc Petit	Université de Lyon, France
Jeffrey Ullman	Stanford University, USA
Jialie Shen	Singapore Management University, Singapore
Jian Yin	Sun Yat-Sen University, China
Jieping Ye	Arizona State University, USA
Jinze Liu	University of Kentucky, USA
John Keane	The University of Manchester, UK
Josep Domingo-Ferrer	Universitat Rovira i Virgili, Spain
Junbin Gao	Charles Sturt University, Australia
Junping Zhang	Fudan University, China

Kamalika Das	NASA Ames Research Center, USA
Kanishka Bhaduri	NASA, USA
Keith Marsolo	Cincinnati Children's Hospital Medical Center, USA
Keith Chan	The Hong Kong Polytechnic University, Hong Kong
Kennichi Yoshida	University of Tsukuba, Japan
Kitsana Waiyamai	Kasetsart University, Thailand
Konstantinos Kalpakis	University of Maryland Baltimore County, USA
Kouzou Ohara	Aoyama-Gakuin University, Japan
Krishnamoorthy Sivakumar	Washington State University, USA
Kun Liu	Yahoo! Labs, USA
Kuo-Wei Hsu	National Chengchi University, Taiwan
Larry Hall	University of South Florida, USA
Larry Holder	Washington State University, USA
Latifur Khan	University of Texas at Dallas, USA
Liang Wang	NLPR, Institute of Automation Chinese Academy of Science, China
Lim Chee Peng	Universiti Sains Malaysia, Malaysia
Lisa Singh	Georgetown University, USA
Maguelonne Teisseire	Maison de la Teledetection, France
Manabu Okumura	Japan Advanced Institute of Science and Technology, Japan
Marco Maggini	Università degli Studi di Siena, Italy
Marian Vajtersic	University of Salzburg, Austria
Marut Buranarach	National Electronics and Computer Technology Center, Thailand
Mary Elaine Califf	Illinois State University, USA
Marzena Kryszkiewicz	Warsaw University of Technology, Poland
Masayuki Numao	Osaka University, Japan
Masoud Makrehchi	University of Waterloo, Canada
Matjaz Gams	J. Stefan Institute, Slovenia
Mengjie Zhang	Victoria University of Wellington, New Zealand
Michael Hahsler	Southern Methodist University, USA
Michael Bruckner	University of Potsdam, Germany
Michalis Vazirgianni	INRIA/FUTURS, France
Min Song	New Jersey Institute of Technology, USA
Min Yao	Zhejiang University, China
Ming-Syan Chen	National Taiwan University, Taiwan
Mingli Song	Zhejiang University, China
Mirco Nanni	Istituto di Scienza e Tecnologie dell'Informazione (ISTI), Italy
Murali Mani	Worcester Polytechnic Institute, USA
Murat Kantarcioglu	University of Texas at Dallas, USA
Nagaraj Kota	Yahoo! Labs, India
Ngoc-Thanh Nguyen	Wroclaw University of Technology, Poland

Olivia Sheng	University of Utah, USA
Pabitra Mitra	Indian Institute of Technology Kharagpur, India
Panagiotis Papadimitriou	Stanford University, USA
Philippe Lenca	Telecom Bretagne, France
Ping Li	Cornell University, USA
Qi Li	Western Kentucky University, USA
Qi He	IBM Research, USA
Qingshan Liu	NLPR, Institute of Automation Chinese Academy of Science, China
Richi Nayak	Queensland University of Technologies, Australia
Robert Hilderman	University of Regina, Canada
Roberto Bayardo	Google, Inc, USA
Rohan Baxter	Australian Taxation Office, Australia
Rui Camacho	Universidade do Porto, Portugal
Ruoming Jin	Kent State University, USA
Sachindra Joshi	IBM Research, India
Sanjay Jain	National University of Singapore, Singapore
Scott Sanner	Australian National University, Australia
See-Kiong Ng	Singapore University of Technology and Design, Singapore
Selcuk Candan	Arizona State University, USA
Shashi Shekhar	University of Minnesota, USA
Shen-Shyang Ho	California Institute of Technology, USA
Sheng Zhong	State University of New York at Buffalo, USA
Shichao Zhang	University of Technology, Australia
Shiguang Shan	Institute of Computing Technology Chinese Academy of Sciences, China
Shoji Hirano	Shimane University, Japan
Shu-Ching Chen	Florida International University, USA
Shuigeng Zhou	Fudan University, China
Shusaku Tsumoto	Shimane University, Japan
Shyam-Kumar Gupta	Indian Institute of Technology, India
Silvia Chiusano	Politecnico di Torino, Italy
Songcan Chen	Nanjing University of Aeronautics and Astronautics, China
Sourav S. Bhowmick	Nanyang Technological University, Singapore
Srikanta Tirthapura	Iowa State University, USA
Srivatsan Laxman	Microsoft Research, India
Stefan Rueping	Fraunhofer IAIS, Germany
Sung-Ho Ha	Kyungpook National University, Korea
Szymon Jaroszewicz	University of Massachusetts Boston, USA
Tadashi Nomoto	National Institute of Japanese Literature, Japan
Takehisa Yairi	University of Tokyo, Japan

Takeshi Fukuda	IBM, Japan
Tamir Tassa	The Open University, Israel
Tao Li	Florida International University, USA
Tapio Elomaa	Tampere University of Technology, Finland
Tetsuya Yoshida	Hokkaido University, Japan
Thepchai Supnithi	National Electronics and Computer Technology Center, Thailand
Thomas Seidl	RWTH Aachen University, Germany
Tom Croonenborghs	Katholieke Hogeschool Kempen, Belgium
Toon Calders	Eindhoven University of Technology, The Netherlands
Toshihiro Kamishima	National Institute of Advanced Industrial Science and Technology, Japan
Toshiro Minami	Kyushu University Library, Japan
Tru Cao	Ho Chi Minh City University of Technology, Vietnam
Tsuyoshi Murata	Tokyo Institute of Technology, Japan
Tu-Bao Ho	Japan Advanced Institute of Science and Technology, Japan
Varun Chandola	Oak Ridge National Laboratory, USA
Vincent S. Tseng	National Cheng Kung University, Taiwan
Vincenzo Piuri	Università degli Studi di Milano, Italy
Vladimir Estivill-Castro	Griffith University, Australia
Wagner Meira	Universidade Federal de Minas Gerais, Brazil
Wai Lam	The Chinese University of Hong Kong, Hong Kong
Walter Kusters	Universiteit Leiden, The Netherlands
Wanpracha Chaovalitw	The State University of New Jersey Rutgers, USA
Wei Fan	IBM T.J. Watson Research Center, USA
Weining Qian	East China Normal University, China
Wen-Chih Peng	National Chiao Tung University, Taiwan
Wilfred Ng	Hong Kong University of Science and Technology, Hong Kong
Woong-Kee Loh	Sungkyul University, South Korea
Xiaofang Zhou	The University of Queensland, Australia
Xiaohua Hu	Drexel University, USA
Xiaohui Liu	Brunel University, UK
Xiaoli Li	Institute for Infocomm Research, Singapore
Xin Wang	University of Calgary, Canada
Xindong Wu	University of Vermont, USA
Xingquan Zhu	Florida Atlantic University, USA
Xintao Wu	University of North Carolina at Charlotte, USA
Xu Sun	Cornell University, USA
Xuan Vinh Nguyen	Monash University, Australia
Xue Li	The University of Queensland, Australia

Xuelong Li	University of London, UK
Xuemin Lin	The University of New South Wales, Australia
Xueyi Wang	Northwest Nazarene University, USA
Yan Liu	IBM Research, USA
Yan Jia	National University of Defense Technology, China
Yang Zhou	Yahoo!, USA
Yang-Sae Moon	Kangwon National University, Korea
Yasuhiko Morimoto	Hiroshima University, Japan
Yi-Dong Shen	Institute of Software, Chinese Academy of Sciences, China
Yi-Ping Chen	La Trobe University, Australia
Yifeng Zeng	Aalborg University, Denmark
Yiu-ming Cheung	Hong Kong Baptist University, Hong Kong
Yong Guan	Iowa State University, USA
Yonghong Peng	University of Bradford, UK
Yue Lu	University of Illinois at Urbana-Champaign, USA
Yun Chi	NEC Laboratories America, Inc., USA
Yunhua Hu	Microsoft Research Asia, China
Zheng Chen	Microsoft Research Asia, China
Zhi-Hua Zhou	Nanjing University, China
Zhiyuan Chen	University of Maryland Baltimore County, USA
Zhongfei Zhang	Binghamton University, USA
Zili Zhang	Deakin University, Australia

External and Invited Reviewers

Aur�lie Bertaux	Antonio Bruno
Tianyu Cao	Rui Chen
Zhiyong Cheng	Patricia Lopez Cueva
Jeremiah Deng	Stephen Guo
Raymond Heatherly	Lam Hoang
Peter Karsmakers	Sofiane Lagraa
St�phane Lallich	Ivan Lee
Peipei Li	Zhao Li
Lin Liu	Corrado Loglisci
Zhenyu Lu	Marc Mertens
Benjamin N�grevergne	Marc Plantevit
Jing Ren	Yelong Sheng
Arnaud Soulet	Vassilios Verykios
Petros Venetis	Guan Wang
Lexing Xie	Yintao Yu
Yan Zhang	

Sponsors



Table of Contents – Part II

Pattern Mining: Networks, Graphs, Time-Series and Outlier Detection

Heterogeneous Ensemble for Feature Drifts in Data Streams	1
<i>Hai-Long Nguyen, Yew-Kwong Woon, Wee-Keong Ng, and Li Wan</i>	
OMC-IDS: At the Cross-Roads of OLAP Mining and Intrusion Detection	13
<i>Hanen Brahmi, Imen Brahmi, and Sadok Ben Yahia</i>	
Towards Linear Time Overlapping Community Detection in Social Networks	25
<i>Jierui Xie and Boleslaw K. Szymanski</i>	
WeightTransmitter: Weighted Association Rule Mining Using Landmark Weights	37
<i>Yun Sing Koh, Russel Pears, and Gillian Dobbie</i>	
Co-occurring Cluster Mining for Damage Patterns Analysis of a Fuel Cell	49
<i>Daiki Inaba, Ken-ichi Fukui, Kazuhisa Sato, Junichirou Mizusaki, and Masayuki Numao</i>	
New Exact Concise Representation of Rare Correlated Patterns: Application to Intrusion Detection	61
<i>Souad Bouasker, Tarek Hamrouni, and Sadok Ben Yahia</i>	
Life Activity Modeling of News Event on Twitter Using Energy Function	73
<i>Rong Lu, Zhiheng Xu, Yang Zhang, and Qing Yang</i>	
Quantifying Reciprocity in Large Weighted Communication Networks	85
<i>Leman Akoglu, Pedro O.S. Vaz de Melo, and Christos Faloutsos</i>	
Hierarchical Graph Summarization: Leveraging Hybrid Information through Visible and Invisible Linkage	97
<i>Rui Yan, Zi Yuan, Xiaojun Wan, Yan Zhang, and Xiaoming Li</i>	
Mining Mobile Users' Activities Based on Search Query Text and Context	109
<i>Bingyue Peng, Yujing Wang, and Jian-Tao Sun</i>	
Spread of Information in a Social Network Using Influential Nodes	121
<i>Arpan Chaudhury, Partha Basuchowdhuri, and Subhashis Majumder</i>	

Discovering Coverage Patterns for Banner Advertisement Placement	133
<i>P. Gowtham Srinivas, P. Krishna Reddy, S. Bhargav, R. Uday Kiran, and D. Satheesh Kumar</i>	
Discovering Unknown But Interesting Items on Personal Social Network	145
<i>Juang-Lin Duan, Shashi Prasad, and Jen-Wei Huang</i>	
The Pattern Next Door: Towards Spatio-sequential Pattern Discovery	157
<i>Hugo Alatrística Salas, Sandra Bringay, Frédéric Flouvat, Nazha Selmaoui-Folcher, and Maguelonne Teisseire</i>	
Accelerating Outlier Detection with Uncertain Data Using Graphics Processors	169
<i>Takazumi Matsumoto and Edward Hung</i>	
Finding Collections of k -Clique Percolated Components in Attributed Graphs	181
<i>Pierre-Nicolas Mougél, Christophe Rigotti, and Olivier Gandrillon</i>	
Reciprocal and Heterogeneous Link Prediction in Social Networks	193
<i>Xiongcai Cai, Michael Bain, Alfred Krzywicki, Wayne Wobcke, Yang Sok Kim, Paul Compton, and Ashesh Mahidadia</i>	
Detecting Multiple Stochastic Network Motifs in Network Data	205
<i>Kai Liu, William K. Cheung, and Jiming Liu</i>	
Scalable Similarity Matching in Streaming Time Series	218
<i>Alice Marascu, Suleiman A. Khan, and Themis Palpanas</i>	
Scalable Mining of Frequent Tri-concepts from <i>Folksonomies</i>	231
<i>Chiraz Trabelsi, Nader Jelassi, and Sadok Ben Yahia</i>	
SHARD: A Framework for Sequential, Hierarchical Anomaly Ranking and Detection	243
<i>Jason Robinson, Margaret Lonergan, Lisa Singh, Allison Candido, and Mehmet Sayal</i>	
Instant Social Graph Search	256
<i>Sen Wu, Jie Tang, and Bo Gao</i>	
 Data Manipulation: Pre-processing and Dimension Reduction	
Peer Matrix Alignment: A New Algorithm	268
<i>Mohammed Kayed</i>	

Domain Transfer Dimensionality Reduction via Discriminant Kernel Learning	280
<i>Ming Zeng and Jiangtao Ren</i>	
Prioritizing Disease Genes by Bi-Random Walk	292
<i>Maoqiang Xie, Taehyun Hwang, and Rui Kuang</i>	
Selecting Feature Subset via Constraint Association Rules	304
<i>Guangtao Wang and Qinbao Song</i>	
RadialViz: An Orientation-Free Frequent Pattern Visualizer	322
<i>Carson Kai-Sang Leung and Fan Jiang</i>	
Feature Weighting by RELIEF Based on Local Hyperplane Approximation	335
<i>Hongmin Cai and Michael Ng</i>	
Towards Identity Disclosure Control in Private Hypergraph Publishing	347
<i>Yidong Li and Hong Shen</i>	
EWNI: Efficient Anonymization of Vulnerable Individuals in Social Networks	359
<i>Frank Nagle, Lisa Singh, and Aris Gkoulalas-Divanis</i>	
A Pruning-Based Approach for Searching Precise and Generalized Region for Synthetic Minority Over-Sampling	371
<i>Kamthorn Puntumapon and Kitsana Waiyamai</i>	
Towards More Efficient Multi-label Classification Using Dependent and Independent Dual Space Reduction	383
<i>Eakasit Pacharawongsakda and Thanaruk Theeramunkong</i>	
Automatic Identification of Protagonist in Fairy Tales Using Verb	395
<i>Hui-Ngo Goh, Lay-Ki Soon, and Su-Cheng Haw</i>	
CD: A Coupled Discretization Algorithm	407
<i>Can Wang, Mingchun Wang, Zhong She, and Longbing Cao</i>	
Co-embedding of Structurally Missing Data by Locally Linear Alignment	419
<i>Takehisa Yairi</i>	
Relevant Feature Selection from EEG Signal for Mental Task Classification	431
<i>Akshansh Gupta and R.K. Agrawal</i>	
Author Index	443

Table of Contents – Part I

Supervised Learning: Active, Ensemble, Rare-Class and Online

Time-Evolving Relational Classification and Ensemble Methods	1
<i>Ryan Rossi and Jennifer Neville</i>	
Active Learning for Hierarchical Text Classification	14
<i>Xiao Li, Da Kuang, and Charles X. Ling</i>	
TeamSkill Evolved: Mixed Classification Schemes for Team-Based Multi-player Games	26
<i>Colin DeLong and Jaideep Srivastava</i>	
A Novel Weighted Ensemble Technique for Time Series Forecasting	38
<i>Ratnadip Adhikari and R.K. Agrawal</i>	
Techniques for Efficient Learning without Search	50
<i>Houssam Salem, Pramuditha Suraweera, Geoffrey I. Webb, and Janice R. Boughton</i>	
An Aggressive Margin-Based Algorithm for Incremental Learning	62
<i>JuiHsi Fu and SingLing Lee</i>	
Two-View Online Learning	74
<i>Tam T. Nguyen, Kuiyu Chang, and Siu Cheung Hui</i>	
A Generic Classifier-Ensemble Approach for Biomedical Named Entity Recognition	86
<i>Zhihua Liao and Zili Zhang</i>	
Neighborhood Random Classification	98
<i>Djamel Abdelkader Zighed, Diala Ezzeddine, and Fabien Rico</i>	
SRF: A Framework for the Study of Classifier Behavior under Training Set Mislabeling Noise	109
<i>Katsiaryna Mirylenka, George Giannakopoulos, and Themis Palpanas</i>	
Building Decision Trees for the Multi-class Imbalance Problem	122
<i>T. Ryan Hoens, Qi Qian, Nitesh V. Chawla, and Zhi-Hua Zhou</i>	
Scalable Random Forests for Massive Data	135
<i>Bingguo Li, Xiaojun Chen, Mark Junjie Li, Joshua Zhexue Huang, and Shengzhong Feng</i>	

Hybrid Random Forests: Advantages of Mixed Trees in Classifying Text Data	147
<i>Baoxun Xu, Joshua Zhexue Huang, Graham Williams, Mark Junjie Li, and Yunming Ye</i>	
Learning Tree Structure of Label Dependency for Multi-label Learning	159
<i>Bin Fu, Zhihai Wang, Rong Pan, Guandong Xu, and Peter Dolog</i>	
Multiple Instance Learning for Group Record Linkage	171
<i>Zhichun Fu, Jun Zhou, Peter Christen, and Mac Boot</i>	
Incremental Set Recommendation Based on Class Differences	183
<i>Yasuyuki Shirai, Koji Tsuruma, Yuko Sakurai, Satoshi Oyama, and Shin-ichi Minato</i>	
Active Learning for Cross Language Text Categorization	195
<i>Yue Liu, Lin Dai, Weitao Zhou, and Heyan Huang</i>	
Evasion Attack of Multi-class Linear Classifiers	207
<i>Han Xiao, Thomas Stibor, and Claudia Eckert</i>	
Foundation of Mining Class-Imbalanced Data	219
<i>Da Kuang, Charles X. Ling, and Jun Du</i>	
Active Learning with ϵ -Certainty	231
<i>Eileen A. Ni and Charles X. Ling</i>	
A Term Association Translation Model for Naive Bayes Text Classification	243
<i>Meng-Sung Wu and Hsin-Min Wang</i>	
A Double-Ensemble Approach for Classifying Skewed Data Streams	254
<i>Chongsheng Zhang and Paolo Soda</i>	
Generating Balanced Classifier-Independent Training Samples from Unlabeled Data	266
<i>Youngja Park, Zijie Qi, Suresh N. Chari, and Ian M. Molloy</i>	
Nyström Approximate Model Selection for LSSVM	282
<i>Lizhong Ding and Shizhong Liao</i>	
Exploiting Label Dependency for Hierarchical Multi-label Classification	294
<i>Noor Alaydie, Chandan K. Reddy, and Farshad Fotouhi</i>	
Diversity Analysis on Boosting Nominal Concepts	306
<i>Nida Meddouri, Héla Khoufi, and Mondher Sadok Maddouri</i>	

Extreme Value Prediction for Zero-Inflated Data	318
<i>Fan Xin and Zubin Abraham</i>	
Learning to Diversify Expert Finding with Subtopics	330
<i>Hang Su, Jie Tang, and Wanling Hong</i>	
An Associative Classifier for Uncertain Datasets	342
<i>Metanat Hooshadat and Osmar R. Zaïane</i>	

Unsupervised Learning: Clustering, Probabilistic Modeling

Neighborhood-Based Smoothing of External Cluster Validity Measures	354
<i>Ken-ichi Fukui and Masayuki Numao</i>	
Sequential Entity Group Topic Model for Getting Topic Flows of Entity Groups within One Document	366
<i>Young-Seob Jeong and Ho-Jin Choi</i>	
Topological Comparisons of Proximity Measures	379
<i>Djamel Abdelkader Zighed, Rafik Abdesselam, and Asmelash Hadgu</i>	
Quad-tuple PLSA: Incorporating Entity and Its Rating in Aspect Identification	392
<i>Wenjuan Luo, Fuzhen Zhuang, Qing He, and Zhongzhi Shi</i>	
Clustering-Based k -Anonymity	405
<i>Xianmang He, HuaHui Chen, Yefang Chen, Yihong Dong, Peng Wang, and Zhenhua Huang</i>	
Unsupervised Ensemble Learning for Mining Top-n Outliers	418
<i>Jun Gao, Weiming Hu, Zhongfei(Mark) Zhang, and Ou Wu</i>	
Towards Personalized Context-Aware Recommendation by Mining Context Logs through Topic Models	431
<i>Kuifei Yu, Baoxian Zhang, Hengshu Zhu, Huanhuan Cao, and Jilei Tian</i>	
Mining of Temporal Coherent Subspace Clusters in Multivariate Time Series Databases	444
<i>Hardy Kremer, Stephan Günnemann, Arne Held, and Thomas Seidl</i>	
A Vertex Similarity Probability Model for Finding Network Community Structure	456
<i>Kan Li and Yin Pang</i>	
Hybrid- ε -greedy for Mobile Context-Aware Recommender System	468
<i>Djallel Bouneffouf, Amel Bouzeghoub, and Alda Lopes Gançarski</i>	

Unsupervised Multi-label Text Classification Using a World Knowledge Ontology	480
<i>Xiaohui Tao, Yuefeng Li, Raymond Y.K. Lau, and Hua Wang</i>	
Semantic Social Network Analysis with Text Corpora	493
<i>Dong-mei Yang, Hui Zheng, Ji-kun Yan, and Ye Jin</i>	
Visualizing Clusters in Parallel Coordinates for Visual Knowledge Discovery	505
<i>Yang Xiang, David Fuhry, Ruoming Jin, Ye Zhao, and Kun Huang</i>	
Feature Enriched Nonparametric Bayesian Co-clustering	517
<i>Pu Wang, Carlotta Domeniconi, Huzefa Rangwala, and Kathryn B. Laskey</i>	
Shape-Based Clustering for Time Series Data	530
<i>Warissara Meesrikamolkul, Vit Niennattrakul, and Chotirat Ann Ratanamahatana</i>	
Privacy-Preserving EM Algorithm for Clustering on Social Network	542
<i>Bin Yang, Issei Sato, and Hiroshi Nakagawa</i>	
Named Entity Recognition and Identification for Finding the Owner of a Home Page	554
<i>Vassilis Plachouras, Matthieu Rivière, and Michalis Vazirgiannis</i>	
Clustering and Understanding Documents via Discrimination Information Maximization	566
<i>Malik Tahir Hassan and Asim Karim</i>	
A Semi-supervised Incremental Clustering Algorithm for Streaming Data	578
<i>Maria Halkidi, Myra Spiliopoulou, and Aikaterini Pavlou</i>	
Unsupervised Sparse Matrix Co-clustering for Marketing and Sales Intelligence	591
<i>Anastasios Zouzias, Michail Vlachos, and Nikolaos M. Freris</i>	
Expectation-Maximization Collaborative Filtering with Explicit and Implicit Feedback	604
<i>Bin Wang, Mohammadreza Rahimi, Dequan Zhou, and Xin Wang</i>	
Author Index	617