

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, Lancaster, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Zurich, Switzerland

John C. Mitchell

Stanford University, Stanford, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

TU Dortmund University, Dortmund, Germany

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbrücken, Germany

More information about this series at <http://www.springer.com/series/7409>

Zhifeng Bao · Goce Trajcevski
Lijun Chang · Wen Hua (Eds.)

Database Systems for Advanced Applications

DASFAA 2017 International Workshops:
BDMS, BDQM, SeCoP, and DMMOOC
Suzhou, China, March 27–30, 2017
Proceedings

Editors

Zhifeng Bao
Royal Melbourne Institute of Technology
Melbourne
Australia

Goce Trajcevski
Northwestern University
Evanston, IL
USA

Lijun Chang
University of New South Wales
Sydney, NSW
Australia

Wen Hua
The University of Queensland
Brisbane, QLD
Australia

ISSN 0302-9743

ISSN 1611-3349 (electronic)

Lecture Notes in Computer Science

ISBN 978-3-319-55704-5

ISBN 978-3-319-55705-2 (eBook)

DOI 10.1007/978-3-319-55705-2

Library of Congress Control Number: 2017934640

LNCS Sublibrary: SL3 – Information Systems and Applications, incl. Internet/Web, and HCI

© Springer International Publishing AG 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, 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.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Along with the main conference, the DASFAA 2017 workshops provided an international forum for researchers and practitioners to gather and discuss research results and open problems, aiming at more focused problem domains and settings. This year there were four workshops held in conjunction with DASFAA 2017:

- The 4th International Workshop on Big Data Management and Service (BDMS 2017)
- The Second Workshop on Big Data Quality Management (BDQM 2017)
- The 4th International Workshop on Semantic Computing and Personalization (SeCoP 2017)
- The First International Workshop on Data Management and Mining on MOOCs (DMMOOC 2017)

All the workshops were selected after a public call-for-proposals process, and each of them focused on a specific area that contributes to, and complements, the main themes of DASFAA 2017. Each workshop proposal, in addition to the main topics of interest, provided a list of the Organizing Committee members and (a tentative) Program Committee. Once the selected proposals were accepted, each of the workshops proceeded with their own call for papers and reviews of the submissions.

In total, 37 papers were accepted, including 13 papers for BDMS 2017, five papers for BDQM 2017, five papers for SeCoP 2017, and 14 papers (nine full and five short) for DMMOOC 2017.

We would like to thank all of the members of the Organizing Committees of the respective workshops, along with their Program Committee members, for their tremendous effort in making the DASFAA 2017 workshops a success. In addition, we are grateful to the main conference organizers for their generous support as well as the efforts in including the papers from the workshops in the proceedings series.

March 2017

Zhifeng Bao
Goce Trajcevski

Organization

The 4th International Workshop on Big Data Management and Service (BDMS 2017)

Workshop Chairs

Xiaoling Wang	East China Normal University, China
Kai Zheng	Soochow University, China
An Liu	Soochow University, China

Program Committee

Jialong Han	Nanyang Technological University, Singapore
Tieke He	Nanjing University, China
Yuwei Peng	Wuhan University, China
Han Su	University of Electric Science and Technology of China, China
Da Yan	University of Alabama at Birmingham, USA
Kun Yue	Yunnan University, China
Wei Zhang	East China Normal University, China
Xin Zhao	Renmin University of China, China
Bolong Zheng	University of Queensland, Australia
Yaqian Zhou	Fudan University, China
Zhixu Li	Soochow University, China
Ke Sun	King Abdullah University of Science and Technology, Saudi Arabia
Haoran Xie	Education University of Hong Kong, SAR China

The Second Workshop on Big Data Quality Management (BDQM 2017)

Honorable Chair

Jianzhong Li	Harbin Institute of Technology, China
--------------	---------------------------------------

Workshop Chairs

Hongzhi Wang	Harbin Institute of Technology, China
Jing Gao	University at Buffalo, USA/State University of New York, USA

Program Committee

Cheqing Jin	East China Normal University, China
Jiannan Wang	Simon Fraser University, Canada

Lingli Li	Heilongjiang University, China
Rihan Hai	Lehrstuhl Informatik 5, RWTH Aachen University, Germany
Wenjie Zhang	University of New South Wales, Australia
Yingyi Bu	Couchbase, USA
Yueguo Chen	Renmin University of China, China
Zhaonian Zou	Harbin Institute of Technology, China
Zhijing Qin	Pinterest, USA

The 4th International Workshop on Semantic Computing and Personalization (SeCoP 2017)

General Chairs

Haoran Xie	The Education University of Hong Kong, SAR Hong Kong
Fu Lee Wang	Caritas Institute of Higher Education, SAR Hong Kong
Tak-Lam Wong	The Education University of Hong Kong, SAR Hong Kong
Yi Cai	South China University of Technology, China

Organizing Chairs

Wei Chen	Agricultural Information Institute of CAAS, China
Tianyong Hao	Guangdong University of Foreign Studies, China
Zhaoqing Pan	Nanjing University of Information Science and Technology, China

Publicity Chairs

Xiaohui Tao	Southern Queensland University, Australia
Di Zou	The Hong Kong Polytechnic University, SAR China
Xudong Mao	City University of Hong Kong, SAR China
Yanghui Rao	Sun-Yet San University, China
Yunhui Zhuang	City University of Hong Kong, SAR China
Zhenguo Yang	Guangdong University of Technology, China

Program Committee

Zhiwen Yu	South China University of Technology, China
Jian Chen	South Chia University of Technology, China
Raymong Y.K. Lau	City University of Hong Kong, SAR China
Rong Pan	Sun Yat-Sen University, China
Yunjun Gao	Zhejiang University, China
Shaojie Qiao	Southwest Jiaotong University, China
Jianke Zhu	Zhejiang University, China
Neil Y. Yen	University of Aizu, Japan
Derong Shen	Northeastern University, China

Jing Yang	Research Center on Fictitious Economy and Data Science CAS, China
Wen Wu	Hong Kong Baptist University, SAR China
Raymong Wong	Hong Kong University of Science and Technology, SAR China
Wenjuan Cui	China Academy of Sciences, China
Xiaodong Li	Hohai University, China
Xiangping Zhai	Nanjing University of Aeronautics and Astronautics, China
Xu Wang	Shenzhen University, China
Ran Wang	Shenzhen University, China
Debby Dan Wang	Caritas Institute of Higher Education, SAR China
Jianming Lv	South China University of Technology, China
Tao Wang	The University of Southampton, UK
Guangliang Chen	TU Delft, The Netherlands
Wenji Ma	Columbia University, USA
Kai Yang	South China University of Technology, China
Yun Ma	City University of Hong Kong, SAR China

The First International Workshop on Data Management and Mining on MOOCs (DMMOOC 2017)

Workshop Chairs

Wenjun Wu	Beihang University, China
Yan Zhang	Peking University, China
Yongxin Tong	Beihang University, China

Program Committee

Yurong Cheng	Northeastern University, China
Dawei Gao	Beihang University, China
Xiaonan Guo	Stevens Institute of Technology, USA
Di Jiang	Baidu Inc., China
Jun Liu	Xi'an Jiaotong University, China
Xinjun Mao	National University of Defense Technology, China
Rui Meng	Hong Kong University of Science and Technology, SAR China
Longfei Shangguan	Princeton University, USA
Jieying She	Hong Kong University of Science and Technology, SAR China
Tianshu Song	Beihang University, China
Zhiyang Su	Microsoft, China
Jie Tang	Tsinghua University, China
Qian Tao	Beihang University, China
Libin Wang	Beihang University, China
Qiong Wang	Peking University, China

Ting Wang	National University of Defense Technology, China
Wei Xu	Tsinghua University, China
Gang Yin	National University of Defense Technology, China
Xiaolong Zheng	Tsinghua University, China
Zimu Zhou	ETH Zurich, Switzerland

Contents

BDMS

Automatically Classify Chinese Judgment Documents Utilizing Machine Learning Algorithms	3
<i>Miaomiao Lei, Jidong Ge, Zhongjin Li, Chuanyi Li, Yemao Zhou, Xiaoyu Zhou, and Bin Luo</i>	
A Partitioning Scheme for Big Dynamic Trees	18
<i>Atsushi Sudoh, Tatsuo Tsuji, and Ken Higuchi</i>	
Optimization Factor Analysis of Large-Scale Join Queries on Different Platforms	35
<i>Chao Yang, Qian Wang, Qing Yang, Huibing Zhang, Jingwei Zhang, and Ya Zhou</i>	
Which Mapping Service Should We Select in China?	47
<i>Detian Zhang, Jia-ao Wang, and Fei Chen</i>	
An Online Prediction Framework for Dynamic Service-Generated QoS Big Data.	60
<i>Jianlong Xu, Changsheng Zhu, and Qi Xie</i>	
Discovering Interesting Co-location Patterns Interactively Using Ontologies . . .	75
<i>Xuguang Bao and Lizhen Wang</i>	
LFLogging: A Latch-Free Logging Scheme for PCM-Based Big Data Management Systems	90
<i>Wenqiang Wang, Peiquan Jin, Shouhong Wan, and Lihua Yue</i>	
RTMatch: Real-Time Location Prediction Based on Trajectory Pattern Matching.	103
<i>Dong Zhenjiang, Deng Jia, Jiang Xiaohui, and Wang Yongli</i>	
Online Formation of Large Tree-Structured Team	118
<i>Cheng Ding, Fan Xia, Gopakumar, Weining Qian, and Aoying Zhou</i>	
Cell-Based DBSCAN Algorithm Using Minimum Bounding Rectangle Criteria	133
<i>Tatsuhiro Sakai, Keiichi Tamura, and Hajime Kitakami</i>	
Time-Aware and Topic-Based Reviewer Assignment	145
<i>Hongwei Peng, Haojie Hu, Keqiang Wang, and Xiaoling Wang</i>	

Adaptive Bayesian Network Structure Learning from Big Datasets 158
Yan Tang, Qidong Zhang, Huaxin Liu, and Wangsong Wang

A Novel Approach for Author Name Disambiguation
Using Ranking Confidence. 169
Xueqin Lin, Jia Zhu, Yong Tang, Fen Yang, Bo Peng, and Weiling Li

BDQM

Capture Missing Values with Inference on Knowledge Base. 185
Zhixin Qi, Hongzhi Wang, Fanshan Meng, Jianzhong Li, and Hong Gao

Weakly-Supervised Named Entity Extraction Using Word Representations. . . 195
Kejun Deng, Dongsheng Wang, and Junfei Liu

RDF Data Assessment Based on Metrics and Improved
PageRank Algorithm 204
Kai Wei, Pingfang Tian, Jinguang Gu, and Li Huang

Efficient Web-Based Data Imputation with Graph Model 213
*Yiwen Tang, Hongzhi Wang, Shiwei Zhang, Huijun Zhang,
and Ruoxi Shi*

A New Schema Design Method for Multi-tenant Database 227
Yaoqiang Xu and Jiakai Ni

SeCoP

Reader's Choice: A Recommendation Platform. 243
Sayar Kumar Dey and Günter Fahrnberger

Accelerating Convolutional Neural Networks Using Fine-Tuned
Backpropagation Progress 256
Yulong Li, Zhenhong Chen, Yi Cai, Dongping Huang, and Qing Li

A Personalized Learning Strategy Recommendation Approach for
Programming Learning 267
Peipei Gu, Junxia Ma, Wei Chen, Lujuan Deng, and Lan Jiang

Wikipedia Based Short Text Classification Method 275
Junze Li, Yi Cai, Zhiwei Cai, Hofung Leung, and Kai Yang

An Efficient Boolean Expression Index by Compression 287
Jin Tao, Chenxi Zhang, and Weixiong Rao

DMMOOC

MOOCon: A Framework for Semi-supervised Concept Extraction from MOOC Content	303
<i>Zhuoxuan Jiang, Yan Zhang, and Xiaoming Li</i>	
What Decides the Dropout in MOOCs?	316
<i>Xiaohang Lu, Shengqing Wang, Junjie Huang, Wenguang Chen, and Zengwang Yan</i>	
Exploring N-gram Features in Clickstream Data for MOOC Learning Achievement Prediction	328
<i>Xiao Li, Ting Wang, and Huaimin Wang</i>	
Predicting Student Examinee Rate in Massive Open Online Courses	340
<i>Wei Lu, Tongtong Wang, Min Jiao, Xiaoying Zhang, Shan Wang, Xiaoyong Du, and Hong Chen</i>	
Task Assignment of Peer Grading in MOOCs.	352
<i>Yong Han, Wenjun Wu, and Yanjun Pu</i>	
Predicting Honors Student Performance Using RBFNN and PCA Method . . .	364
<i>Moke Xu, Yu Liang, and Wenjun Wu</i>	
DKG: An Expanded Knowledge Base for Online Course	376
<i>Haimeng Duan, Yuanhao Zheng, Lei Shi, Changhong Jin, Hongwei Zeng, and Jun Liu</i>	
Towards Economic Models for MOOC Pricing Strategy Design	387
<i>Yongzheng Jia, Zhengyang Song, Xiaolan Bai, and Wei Xu</i>	
Using Pull-Based Collaborative Development Model in Software Engineering Courses: A Case Study	399
<i>Yao Lu, Xinjun Mao, Gang Yin, Tao Wang, and Yu Bai</i>	
A Method of Constructing the Mapping Knowledge Domains in Chinese Based on the MOOCs	411
<i>Zhengzhou Zhu, Yang Li, Youming Zhang, and Zhonghai Wu</i>	
Social Friendship-Aware Courses Arrangement on MOOCs	417
<i>Yuan Liang</i>	
Quality-Aware Crowdsourcing Curriculum Recommendation in MOOCs	423
<i>Yunpeng Gao</i>	

Crowdsourcing Based Teaching Assistant Arrangement for MOOC	429
<i>Dezhi Sun and Bo Liu</i>	
Quantitative Analysis of Learning Data in a Programming Course.	436
<i>Yu Bai, Liqian Chen, Gang Yin, Xinjun Mao, Ye Deng, Tao Wang,</i> <i>Yao Lu, and Huaimin Wang</i>	
Author Index	443