Communications in Computer and Information Science 1281

Editorial Board Members

Joaquim Filipe
Polytechnic Institute of Setúbal, Setúbal, Portugal

Ashish Ghosh
Indian Statistical Institute, Kolkata, India

Raquel Oliveira Prates
Federal University of Minas Gerais (UFMG), Belo Horizonte, Brazil

Lizhu Zhou
Tsinghua University, Beijing, China

More information about this series at http://www.springer.com/series/7899

Lu Qin · Wenjie Zhang ·
Ying Zhang · You Peng ·
Hiroyuki Kato · Wei Wang ·
Chuan Xiao (Eds.)

Software Foundations for Data Interoperability and Large Scale Graph Data Analytics

4th International Workshop, SFDI 2020
and 2nd International Workshop, LSGDA 2020
held in Conjunction with VLDB 2020
Tokyo, Japan, September 4, 2020
Proceedings

Springer

Editors
Lu Qin
University of Technology Sydney
Sydney, NSW, Australia

Ying Zhang
University of Technology Sydney
Sydney, NSW, Australia

Hiroyuki Kato
National Institute of Informatics
Tokyo, Japan

Chuan Xiao
Osaka University
Osaka, Japan

Wenjie Zhang
The University of New South Wales
Sydney, NSW, Australia

You Peng
The University of New South Wales
Sydney, NSW, Australia

Wei Wang
The University of New South Wales
Sydney, NSW, Australia

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-030-61132-3 ISBN 978-3-030-61133-0 (eBook)
https://doi.org/10.1007/978-3-030-61133-0

© Springer Nature Switzerland AG 2020
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, expressed 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.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume gathers the papers presented at the Second International Workshop on Large Scale Graph Data Analytics (LSGDA 2020) and the 4th Workshop on Software Foundations for Data Interoperability (SDFI 2020), held in Tokyo, Japan, on September 4, 2020. The LSGDA and the SFDI workshops continued the series of annual workshops which have previously been held in Macau, China (2019), Kyoto, Japan (2018 and 2019), and Fukuoka, Japan (2019), respectively. The Second LSGDA workshop was organized by the University of New South Wales, Australia, in cooperation with the University of Technology Sydney, Australia. The 4th SFDI workshop was organized by the National Institute of Informatics, Japan, the University of New South Wales, Australia, and Osaka University, Japan.

Unfortunately, due to the escalation of the COVID-19 pandemic, and the anti-COVID-19 regulations in Japan and the rest of the world, the planned onsite conferences in Tokyo had to be reformatted. The workshops were held online as virtual conferences featuring live and semi-live presentations during the same period of time.

The LSGDA and the SFDI series of workshops have served as international forums for researchers, practitioners, and PhD students to exchange research findings and ideas on the crucial matters on large-scale graph data analytics and data interoperability, respectively. The LSGDA 2020 and the SDFI 2020 workshops continued this tradition and featured original research and application papers on the development of novel graph analytics models, scalable graph analytics techniques and systems, data integration, and data exchange.

The program of the LSGDA 2020 included three invited keynote talks, given by Prof. Chengfei Liu (Swinburne University of Technology, Australia), Prof. Da Yan (University of Alabama at Birmingham, USA), and Prof. Weiren Yu (University of Warwick, UK). The program of SFDI 2020 included three invited keynote talks, given by Prof. Koiti Hasida (The University of Tokyo and Riken, Japan), Prof. Rui Zhang (The University of Melbourne, Australia), and Prof. Kazutaka Matsuda (Tohoku University, Japan). The call for papers for LSGDA 2020 welcomed original unpublished research and application experience papers on graph data model, storage, indexing and query processing techniques, graph mining techniques, techniques for distributed graph analytics, graph visualization techniques and system interfaces, dynamic and streaming graph data analytics, spatial-temporal graph analytics, AI techniques for graphs, machine learning techniques for graphs, and vision papers to survey the area of graph data analytics as well as describe the future research directions. The call for papers for SFDI 2020 welcomed original unpublished research and application experience papers on software foundations for data interoperability, including data integration, data exchange, distributed collaborative systems, and applications in real-world systems such as data markets. The two workshops collectively received 38 submissions with authors coming from 10 countries. Each paper was evaluated through single-blind review by at least three members of the Program

Committee (PC). The reviewers were assigned after careful consideration of all potential conflicts. Papers were not assigned to PC members originating from the same affiliation or having any known conflicting interests. After the review process, papers with consistent negative evaluations were rejected, whereas papers with mixed ratings (positive and negative) were additionally evaluated by program chairs prior to the meeting, in which all the papers and final decisions regarding them were thoroughly discussed. The evaluation process resulted in the selection of 15 papers (acceptance rate of 39%), which were accepted for presentation at the conferences and publication in this joint proceedings.

The original research results presented in this volume concern well-established fields such as graph data model, storage, indexing and query processing techniques, graph mining techniques, techniques for distributed graph analytics, graph visualization techniques and system interfaces, dynamic and streaming graph data analytics, spatial-temporal graph analytics, AI techniques for graphs, machine learning techniques for graphs, similarity query processing techniques, solutions to data exchange and data integration, heterogeneous data management, and distributed data management. The research results feature vision papers to survey the areas of graph data analytics and data interoperability as well as describe future research directions. The volume also includes three papers for the keynote talks of LSGDA 2020.

Finally, we express our deep gratitude to the members of the Program Committees of the two workshops for their time, comments, and constructive evaluations. We would like to thank everyone from the Organizing/Steering Committees for their time and dedication, which helped make the conferences successful. We are also grateful to the authors and all the participants who truly made the conferences successful, even within the short time frame we had to reorganize the workshops due to the COVID-19 outbreak, we managed to face the new reality and hold the conferences virtually from a safe distance.

September 2020

Lu Qin
Wenjie Zhang
Ying Zhang
You Peng
Hiroyuki Kato
Wei Wang
Chuan Xiao

LSGDA 2020 Organization

General Chair

Xuemin Lin	University of New South Wales, Australia

Program Committee Co-chairs

Lu Qin	University of Technology Sydney, Australia
Wenjie Zhang	University of New South Wales, Australia
Ying Zhang	University of Technology Sydney, Australia

Program Committee

Witold Abramowicz	Poznań University of Economics and Business, Poland
Anil Pacaci	University of Waterloo, Canada
Bolin Ding	Alibaba Group, USA
Chuan Xiao	Osaka University, Japan
Chunbin Lin	Amazon Web Services, USA
Donatella Firmani	Roma Tre University, Italy
Huasong Shan	JD.COM, USA
Jiafeng Hu	Google, China
Jianye Yang	Hunan University, China
Lijun Chang	The University of Sydney, Australia
Matteo Lissandrini	Aalborg University, Denmark
Rong-Hua Li	Beijing Institute of Technology, China
Sergey Pupyrev	Facebook, USA
Stefano Leucci	University of L'Aquila, Italy
Verena Kantere	National Technical University of Athens, Greece
Vijil Chenthamarakshan	IBM AI Research, USA
Weiren Yu	University of Warwick, UK
Xiang Zhao	National University of Defense Technology, China
Xin Cao	University of New South Wales, Australia
Yuanyuan Zhu	Wuhan University, China
Zhaonian Zou	Harbin Institute of Technology, China

Organizing Committee

Dong Wen	University of Technology Sydney, Australia
You Peng	University of New South Wales, Australia
Kai Wang	University of New South Wales, Australia

SFDI 2020 Organization

Steering Committee

Zhenjiang Hu	Peking University, China, and NII, Japan
Makoto Onizuka	Osaka University, Japan
Masatoshi Yoshikawa	Kyoto University, Japan

Program Committee Co-chairs

Hiroyuki Kato	NII, Japan
Wei Wang	University of New South Wales, Australia
Chuan Xiao	Osaka University, Japan

Program Committee

Yang Cao	Kyoto University, Japan
Muhammad Aamir Cheema	Monash University, Australia
Yuyang Dong	NEC, Japan
Raul Castro Fernandez	University of Chicago, USA
Torsten Grust	University of Tübingen, Germany
Sheng Hu	Hokkaido University, Japan
Verena Kantere	National Technical University of Athens, Greece
Sebastian Maneth	University of Bremen, Germany
Parth Nagarkar	New Mexico State University, USA
George Papadakis	University of Athens, Greece
Jianbin Qin	Shenzhen University, China
Lu Qin	University of Technology Sydney, Australia
Yuya Sasaki	Osaka University, Japan
Toshiyuki Shimizu	Kyoto University, Japan
Massimo Tisi	IMT Atlantique, France
Jiannan Wang	Simon Fraser University, Canada
Xiang Zhao	National University of Defense Technology, China
Wenjie Zhang	University of New South Wales, Australia
Erkang Zhu	Microsoft, USA
Kostas Zoumpatianos	Harvard University, USA

Contents

Large Scale Graph Data Analytics

Attribute Diversified Community Search 3
Chengfei Liu, Lu Chen, Rui Zhou, and Afzal Azeem Chowdhary

Parallel Mining of Frequent Subtree Patterns 18
Wenwen Qu, Da Yan, Guimu Guo, Xiaoling Wang, Lei Zou, and Yang Zhou

An Axiomatic Role Similarity Measure Based on Graph Topology 33
Weiren Yu, Sima Iranmanesh, Aparajita Haldar, Maoyin Zhang, and Hakan Ferhatosmanoglu

Scalable In-Memory Graph Pattern Matching on Symmetric Multiprocessor Systems 49
Alexander Krause, Dirk Habich, and Wolfgang Lehner

A Graph-Based Approach Towards Risk Alerting for COVID-19 Spread 63
Aibo Guo, QianZhen Zhang, and Xiang Zhao

Distributed Graph Analytics with Datalog Queries in Flink. 70
Muhammad Imran, Gábor E. Gévay, and Volker Markl

Explaining Results of Path Queries on Graphs: Single-Path Results for Context-Free Path Queries. 84
Jelle Hellings

Software Foundations for Data Interoperability

NGNC: A Flexible and Efficient Framework for Error-Tolerant Query Autocompletion. 101
Yukai Miao, Jianbin Qin, Sheng Hu, Yuyang Dong, Yoshiharu Ishikawa, and Makoto Onizuka

A Cheap Implementation of Resugaring in BIRDS Based on Bidirectional Transformation 116
Xing Zhang, Van-Dang Tran, and Zhenjiang Hu

Toward Appropriate Data Publishing in Relational Data Exchange Framework 131
Yasunori Ishihara

Toward Programmable Strategy for Co-existence of Relational Schemes 138
Jumpei Tanaka, Van-Dang Tran, and Zhenjiang Hu

Data Integration Models and Architectures for Service Alliances 152
Yasuhito Asano, Zhenjiang Hu, Yasunori Ishihara, Makoto Onizuka, Masato Takeichi, and Masatoshi Yoshikawa

Towards Smart Data Sharing by Updatable Views. 165
Makoto Onizuka, Yasunori Ishihara, and Masato Takeichi

Integration of Fast-Evolving Data Sources Using a Deep Learning Approach 172
Zijie Wang, Lixi Zhou, and Jia Zou

Towards Guaranteeing Global Consistency for Peer-Based Data Integration Architecture 187
Kota Miyake, Yusuke Wakuta, Yuya Sasaki, and Makoto Onizuka

Author Index 195