Guest Editorial: Special Section on the International Conference on Data Engineering

Christian S. Jensen, *Fellow, IEEE*, Christopher Jermaine, and Xiaofang Zhou

THE 29th International Conference on Data Engineering was held in Brisbane, QLD, Australia, on April 8-11, 2013. ICDE 2013 attracted 443 submissions in the research track, 20 submissions in the industrial track, and 69 demo proposals. Each submission was assigned to three reviewers. The evaluation process had several phases: assignment of papers to reviewers, reviewing, discussions among reviewers, decision making by area chairs, consolidation of decisions, and handling of papers assigned for shepherding. As a result of these efforts, 95 research papers, eight industrial papers, and 27 demos were selected for inclusion in the conference program.

This special section consists of journal versions of seven outstanding papers selected among the 95 accepted research papers. All papers were revised and substantially extended over their conference versions and went through a rigorous review process to ensure that the high quality standards of the *IEEE Transactions on Knowledge and Data Engineering* were met. The seven papers cover a broad range of topics that attest to the scope of ongoing research in data engineering.

The paper "Breaking the Barrier to Transferring Link Information across Networks" by Guo-Jun Qi, Charu Aggarwal, and Thomas S. Huang proposes a transfer learning based method for cross-network link prediction. This link prediction model can transfer linkage information from a "mature" source social network to a "young" target network by a bias-correction sampling technique.

The paper "Main-Memory Hash Joins on Modern Processor Architectures" by Cagri Balkesan, Jens Teubner, Gustavo Alonso, and M. Tamer Ozsu compares hardware-conscious and hardware-oblivious hash join approaches in empirical studies that consider a large number of workloads and configurations. They find that hardware-conscious algorithms generally outperform hardware-oblivious algorithms, although hardware-oblivious algorithms are competitive under certain circumstances that involve aggressive simultaneous multi-threading.

The paper "Efficient Notification of Meeting Points for Moving Groups via Independent Safe Regions" by Jing Li, Jeppe Rishede Thomsen, Man Lung Yiu, and Nikos Mamoulis studies the problem of continuously notifying a group of moving users of their best meeting point. They propose and evaluate new kinds of approximate safe regions that enable the delivery of this functionality in a manner that reduces communication and computational costs.

The paper "Real-Time City-Scale Taxi Ridesharing" by Shuo Ma, Yu Zheng, and Ouri Wolfson reports on the design and evaluation of a cloud-based system that enables ridesharing by assigning real-time ride requests to appropriate taxis while offering passengers and drivers monetary incentives for ridesharing. Empirical studies based on real taxi trajectories suggest that the system is effective, e.g., allowing the same fleet of taxis to serve more passengers.

The paper "Crowdsourcing Enumeration Queries: Estimators and Interfaces" by Beth Trushkowsky, Tim Kraska, Michael Franklin, Purnamrita Sarkar, and Venketaram Ramachandran studies an interesting and important aspect of crowdsourcing, which is a promising computing paradigm that involves humans to solve problems that are challenging to machines. They propose effective techniques for query result size estimation in crowdsourcing databases. A novel interface with "negative suggest" is proposed to control the amount of duplicate answers, with a partitioning technique that can improve the diversity of answers.

The paper "Fast All-Pairs SimRank Assessment on Large Graphs and Bipartite Domains" by Weiren Yu, Xuemin Lin, Wenjie Zhang, and Julie McCann considers how to speed up SimRank computation on large graphs; SimRank is a similarity measure that considers two vertices in a graph to be similar if they are referenced by similar vertices. The primary problem with SimRank is that it is very expensive to compute. The authors consider how it can be computed very quickly, proposing a series of algorithms that outperform all of the obvious alternatives.

Finally, the paper "SociaLite: An Efficient Graph Query Language Based on Datalog" by Jiwon Seo, Stephen Guo, and Monica Lam presents the SocialLite programming language and engine. Not only are common, recursive graph computations exceedingly succinct when coded in SocialLite, but they have excellent performance due to a few key features of SocialLite, such as the ability of the programmer to give hints to the execution engine. By implementing and evaluating nine different graph algorithms, the authors show that SocialLite implementations have speeds comparable to those of hand-coded, heavily optimized Java implementations.

- *C.S. Jensen is with the Department of Computer Science, Aalborg University, Selma Lagerlöfs Vej 300, DK-9220 Aalborg Øst, Denmark. E-mail: csj@cs.aau.dk.*
- *C. Jermaine is with the Department of Computer Science, Rice University, P.O. Box 1892, MS-132, Houston, TX 77251-1892. E-mail: cmj4@rice.edu.*
- *X. Zhou is with the School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane QLD 4072 Australia. E-mail: zxf@itee.uq.edu.au.*

For information on obtaining reprints of this article, please send e-mail to: reprints.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TKDE.2015.2419315

We thank all of the authors and reviewers of this special section for their hard work. We also extend our thanks to the program committee, the area chairs, the organization committee, and the authors and participants of ICDE 2013 for contributing to achieving a rewarding and successful conference. We hope that you find the selected papers in this special section enjoyable, informative, and useful.

Christian S. Jensen
Christopher Jermaine
Xiaofang Zhou
Guest Editors
ICDE 2013 Program Committee Chairs

Christian S. Jensen is an Obel professor of computer science at Aalborg University, Denmark. He was a professor at Aarhus University from 2010 to 2013, and he was previously at Aalborg University for two decades. He recently spent a one-year sabbatical at Google Inc., Mountain View, CA. His research concerns data management and data-intensive systems, and its focus is on temporal and spatio-temporal data management. He has received several national and international awards for his research. He is an editor-in-chief of *ACM TODS* and was an editor-in-chief of *The VLDB Journal* from 2008 to 2014. He is a fellow of the ACM and IEEE, and he is a member of the Academia Europaea, the Royal Danish Academy of Sciences and Letters, and the Danish Academy of Technical Sciences.

Christopher Jermaine is an associate professor of computer science at Rice University in Houston, TX. He focuses on building software systems that allow the application of statistical methods to large and complex data sets. He is currently on the editorial boards of the *ACM Transactions on Database Systems* and the *IEEE Data Engineering Bulletin*; he was previously on the editorial boards of the *IEEE Transactions on Knowledge and Data Engineering* and *The Very Large Database Journal*. He received of an Alfred P. Sloan Foundation Research Fellowship, a US National Science Foundation CAREER award, and an ACM SIGMOD Best Paper Award.

Xiaofang Zhou is a professor of computer science and the head in the Data and Knowledge Engineering Research Division, University of Queensland, Australia. His research focus is to find effective and efficient solutions for managing, integrating, and analyzing very large amount of complex data for business, scientific, and personal applications. He has been working in the area of spatial and multimedia databases, data quality, high-performance query processing, and web information systems. He is the chair of the IEEE Technical Committee on Data Engineering and a current member of the editorial boards of *The VLDB Journal*, the *IEEE Transactions on Cloud Computing*, the *World Wide Web Journal*, *Distributed and Parallel Databases*, and the *IEEE Data Engineering Bulletin*. He was an associate editor of the *IEEE Transactions on Knowledge and Data Engineering* (2009-2013). He is also a specially appointed adjunct professor in Soochow University, China, where he leads the Research Center on Advanced Data Analytics (ADA).