EDITORIAL



Special issue on big graph data management and processing

Angela Bonifati¹ · Hannes Voigt²

Published online: 7 March 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

The aim of this special issue is to focus on recent advances in research and development in big graph data management and processing.

In our world, data are not just getting bigger, it is also getting more connected. Exploring, describing, predicting, and explaining phenomena connected to the interconnected world require the use of adequate data abstractions. Graphs are recognized as a general, natural, and flexible data abstraction that can model complex relationships, interactions, and interdependencies between objects. Graphs have been widely used to represent datasets and encode problems across an already extensive range of application domains. The everincreasing size of graph-structured data for these applications creates a critical need for scalable and even elastic systems that can process large amounts of it efficiently. Additionally, the complexity of using multiple datasets simultaneously in complex analysis raises numerous challenges for graph processing, from new requirements to new capabilities. A recent vision paper on "The Future of Big Graphs" [1] highlighted the need of graph ecosystems, encompassing the above challenges.

The special issue received 16 submissions, out of which 8 submissions have been accepted after rigorous review and at least one revision cycle. The accepted articles cover various areas within this theme, ranging from novel algorithms for density decomposition and influence maximization to frameworks and interfaces that provide technical abstraction of common graph analytics tasks. Concretely, this issue consists of eight papers that are briefly discussed as follows.

In the first article "Towards Efficient Solutions of Bitruss Decomposition for Large-scale Bipartite Graphs", K. Wang et al. study bitruss decomposition for bipartite graphs. Various measures, subgraph models, and algorithms exist for

 Angela Bonifati angela.bonifati@univ-lyon1.fr
Hannes Voigt hannes.voigt@neo4j.com

¹ Lyon 1 University, Nautibus Building, 23-25 Avenue Pierre de Coubertin, 69100 Villeurbanne Cedex, France

² Neo4j, Erich-Zeigner-Allee 69-73, 04229 Leipzig, Germany

decomposing a graph into a density hierarchy. Efficient density decomposition is still an ongoing research topic. K. Wang et al. propose a novel online index and two accompanying bitruss decomposition for bipartite graphs.

In the second article "Anchored Coreness: Efficient Reinforcement of Social Networks", Q. Linghu et al. study network reinforcement based on k-core decomposition. Density decomposition can also be used to stabilize social networks by anchoring individual nodes to not leave the network. Given a budget for anchoring a small number of nodes, this poses an optimization problem of maximizing the reinforcement of the network. Q. Linghu et al. propose to do this in a global manner for the whole network rather than locally for individual k-cores.

The third article "PrefixFPM: A Parallel Framework for General-Purpose Mining of Frequent and Closed Patterns" by D. Yan et al. proposes a framework for frequent and close pattern mining. The framework generalizes state-of-the-art pattern mining algorithms and others the user a unify API to customize the framework for mining their desired patterns.

In the fourth article "G-thinker: A General Distributed Framework for Finding Qualified Subgraphs in a Big Graph with Load Balancing", D. Yan et al. propose a similar framework for subgraph finding. G-thinker executes subgraph finding distributively in a CPU-bound manner, handling all the scheduling, communication, caching, CPU idle time minimization for the user.

The fifth article by M. Aisha et al. on "RDFFrames: Knowledge Graph Access for Machine Learning Tools" proposes an interface between RDF databases and Machine Learning tools in order to fill the gap of the data model and programming paradigm. It shows how to combine the usability of PyData with the performance and efficiency of RDF databases.

In the sixth article on "A Design Space for RDF Data Representations", T. Sagi et al. provide a three-dimension design space for RDF storage systems along the subdivision, compression, and redundancy. This design space can help existing systems to identify storage dimensions on which such systems are not optimized thus leading to their improvement. The seventh article by C. Rost et al. on "Distributed Temporal Graph Analytics with GRADOOP" is a complete overview of the GRADOOP system, a graph dataflow system for scalable, distributed analytics of temporal property graphs. The system leverages a bitemporal graph model and suitable analytical operators that guarantee both scalability and efficiency.

In the eight article on "A Fractional Memory Efficient Approach for Online Continuous-time Influence Maximization", G. S. Bevilacqua and L. V. S. Lakshmanan study influence maximization, which is an intractable problem for which approximation algorithms have been proposed. The paper focuses on the design of a new algorithm that allows the influence samples to be processed in a streaming manner, avoiding the need to store large collections of them and the associated prohibitive cost.

In summary, this special issue embraces several fundamental and challenging topics in the areas of graph algorithms, graph data management, graph analytics, and graph learning. We hope that you will find this special issue interesting and enjoyable.

Angela Bonifati and Hannes Voigt Guest Editors of the Special Issue

Reference

 Sakr, S., Bonifati, A., Voigt, H., Iosup, A., Ammar, K., Angles, R., Aref, W.G., Arenas, M., Besta, M., Boncz, P.A., Daudjee, K., Valle, E.D., Dumbrava, S., Hartig, O., Haslhofer, B., Hegeman, T., Hidders, J., Hose, K., Iamnitchi, A., Kalavri, V., Kapp, H., Martens, W., Özsu, M.T., Peukert, E., Plantikow, S., Ragab, M., Ripeanu, M., Salihoglu, S., Schulz, C., Selmer, P., Sequeda, J.F., Shinavier, J., Szárnyas, G., Tommasini, R., Tumeo, A., Uta, A., Varbanescu, A.L., Wu, H.-Y., Yakovets, N., Yan, D., Yoneki, E.: The future is big graphs: a community view on graph processing systems. Commun. ACM, 64(9), 62–71 (2021)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.