Big Data Research 25 (2021) 100215

Contents lists available at ScienceDirect

Big Data Research

www.elsevier.com/locate/bdr

Editorial for Special Issue of Journal of Big Data Research on "Big Data Meets Knowledge Graphs"

In recent years, an increasing number of large-scale knowledge graphs have been constructed and published, by both academic and industrial communities, such as DBpedia, YAGO, Freebase, Wikidata, Google Knowledge Graph, Microsoft Satori, Facebook Entity Graph, and others. In fact, large-scale knowledge graphs from various domains are essentially big data of graph structures consisting of entities, attributes, and semantic relationships with different expressive capabilities.

However, the graph-based ontology structure and the inherent complexity of these large knowledge graphs have been posing great challenges to the existing big data management and processing theories and methods. These challenges include big knowledge graph construction, big knowledge graph data management, data processing on big knowledge graphs, reasoning on big knowledge graphs, representation learning on big knowledge graphs, machine intelligence based on big knowledge graphs, big knowledge graph visualization, and applications of knowledge-graph-based big data.

This special issue aims at providing active researchers a platform to present recent advancements and address some of these challenges in the convergent research when big data meets knowledge graphs. After a rigorous peer review process, we have accepted ten original research papers out of sixteen, all of which report on state-of-the-art and recent achievements in research involving big knowledge graph data. These papers cover a wide range of topics as described below.

The paper "Richpedia: A Large-Scale, Comprehensive Multi-Modal Knowledge Graph" presents Richpedia, which aims at providing a comprehensive multi-modal knowledge graph by distributing sufficient and diverse images to textual entities in Wikidata. The Richpedia resource is accessible on the Web via a faceted query endpoint, which provides a pathway for knowledge graph and computer vision tasks, such as link prediction and visual relation detection.

The paper "WISE: Workload-Aware Partitioning for RDF Systems" proposes the WISE partitioning framework for RDF knowledge graphs, which can dynamically execute SPARQL query workloads and periodically move the RDF triples of a frequent query pattern to the same partition. The WISE framework significantly reduces the communication cost and improves the response time of queries, thus outperforming the state-of-the-art methods.

The paper "Ontology-Mediated SPARQL Query Answering over Knowledge Graphs" focuses on the extension of SPARQL queries over RDF knowledge graphs with ontology-mediated query answering. The authors formalize the requirements in terms of semantics to reconcile certain answers and SPARQL answers and give some important theoretical results on query answering under different semantics over knowledge graphs equipped with an ontology.

The paper "Iterative Visual Relationship Detection via Commonsense Knowledge Graph" proposes a novel model that is able to take advantage of commonsense knowledge in Scene Graph Generation. The model can predict visual relationships by using both visual features and semantic features, and meanwhile construct a specific commonsense knowledge graph for predicate prediction in an iterative process.

The paper "Retrofitting Soft Rules for Knowledge Representation Learning" proposes a retrofit framework that can iteratively enhance the knowledge representation and confidence of soft rules in knowledge graph completion using rule-enhanced learning. The soft rules guide the learning of knowledge representation, which, in turn, provides global semantics of the knowledge graph to optimize the confidence of soft rules.

The paper "JECI++: A Modified Joint Knowledge Graph Embedding Model for Concepts and Instances" proposes a novel knowledge graph embedding model that distinguishes between concepts and instances. In this model, called JECI++, an instance can be obtained by its neighbor instances and the concepts to which they belong. Then, circular convolution is used to locate an instance in the embedding space based on neighbor instances and simplified concepts. JECI++ can alleviate the problem of complex relations by incorporating neighbor information of instances.

The paper "Fine-Grained Evaluation of Knowledge Graph Embedding Model in Knowledge Enhancement Downstream Tasks" conducts comprehensive experiments on different knowledge graph embedding models in knowledge-graph-based question answering, recommendation, and NLP tasks. This work mines the characteristics of each knowledge graph embedding model in actual application scenarios, and provides guidance for the knowledge-enhanced downstream tasks.

The paper "SMR: Medical Knowledge Graph Embedding for Safe Medicine Recommendation" proposes a novel medicine recommendation system, which is based on constructing a high-quality heterogeneous graph by bridging electronic medical records and medical knowledge graphs. The proposed framework can jointly embed knowledge graphs of diseases, medicines, and patients into a shared lower dimensional space to decompose the medicine recommendation into a link prediction process while considering the patient's diagnoses and adverse drug reactions.

The paper "Knowledge Graph-Based Spatial-Aware User Community Preference Query Algorithm for LBSNs" proposes a query algorithm to effectively discover user's community preferences from Location-Based Social Networks, which considers both location-

https://doi.org/10.1016/j.bdr.2021.100215

based semantic information and preference weights of users' points of interest. An R-tree type spatial index is leveraged to improve query efficiency, and a community satisfaction degree model based on knowledge graphs is introduced to comprehensively evaluate whether the points of interest can best meet the preference requirements of a user community.

The paper "CSIP: Enhanced Link Prediction with Context of Social Influence Propagation" proposes to utilize dynamic information in social networks to improve the performance of link prediction in knowledge graphs with social attributes, i.e., social networks. The approach uses a new multi-output graph neural network framework to capture the information about influence propagation in social networks, and applies it to model user nodes to obtain potential relationships among users.

Finally, we would like to take this opportunity to thank the Editors-in-Chief of the Big Data Research journal, Professor H. Chen and Professor T. Palpanas, for their invaluable help and support throughout the preparation of this special issue, as well as the technical and production team of the journal. We express our sincere gratitude to all the reviewers who devoted much of their precious time reviewing all the papers submitted to this special issue. Their timely reviews greatly helped us in selecting the best papers to include in this issue. Also, we would like to thank all authors who contributed to this special issue and hope that the readers will find the selection of the above papers useful for their own research.

Guest Editors
Xin Wang
Tianjin University, China
E-mail address: wangx@tju.edu.cn

Diego Calvanese
Free University of Bozen-Bolzano, Italy

1 February 2021
Available online 19 February 2021