Guest Editorial: Special Issue on Methods and Infrastructures for Data Mining at the Edge of Internet of Things

THE Internet of Things (IoT) enables the interconnection of new cyber–physical devices that generate significant traffic of distributed, heterogeneous, and dynamic data at the network edge. Since several IoT applications demand for short response times (e.g., industrial applications, emergency management, real-time systems, and healthcare systems) and, at the same time, rely on resource-constrained devices, the adoption of traditional data mining techniques is neither effective nor efficient. Therefore, conventional data mining techniques need to be adjusted for optimizing response times, energy consumption, and data traffic while still providing adequate accuracy as required by the IoT applications.

In this special issue, new data mining approaches particularly tailored for the IoT scenario were investigated, in particular with respect to the promising and emerging novel distributed computing paradigm of edge computing. The focus of the special issue was therefore on high-quality original papers aiming at demonstrating effective and efficient data mining approaches at the IoT edge, which at the same time considers data, device, and infrastructure perspectives and related issues. Selected topics included: edge-based platforms for IoT data mining, adapting data mining algorithms and methods to edge computing, novel edge-based data mining methods and algorithms for IoT, novel edge-based machine/deep learning methods and algorithms for IoT, simulation of edge IoT mining, methodologies for driving data mining-based IoT edge systems development, security, privacy, and trust for data mining at the IoT edge, and real/industry applications and systems of data mining at the edge.

The response to our call for this special issue was overwhelming, as we received in total 65 submissions from around the world. During the review process, each article was assigned to and reviewed by at least three experts in the field, with a rigorous multiround review process. Thanks to the great support from the Editor-in-Chief, Prof. Honggang Wang, and the dedicated work of numerous reviewers, we were able to accept 14 excellent articles covering various topics in methods and infrastructures for data mining at the edge of IoT. In the following, we will introduce these articles and highlight their main contributions.

In their article "CLONE: Collaborative learning on the edges," Lu *et al.* propose a collaborative learning framework on the edges (CLONE), including CLONE training and CLONE inference. It demonstrates the effectiveness of privacy serving and latency reduction through emblematic use cases concerning failure prediction of electric vehicle battery and multitarget multicamera tracking in a grocery store.

The article "Communication-efficient offloading for mobile-edge computing in 5G heterogeneous networks," by Zhou *et al.*, proposes a service deployment architecture 5G heterogeneous networks in which three cognitive engines of data, resource, and user QoE are key components. In a multiuser service scenario, experiments show that the defined adaptive task offloading scheme has the lowest total response time compared to the scenarios of local computing, edge offloading computing, and cloud offloading computing.

Chen *et al.*, in the article "A security awareness and protection system for 5G smart healthcare based on zero-trust architecture," propose a security awareness and protection system that leverages zero-trust architecture for a 5G-based smart medical platform. The system, implemented and tested thoroughly at industrial grade, constructs trustable dynamic access control models and achieves real-time network security situational awareness, continuous identity authentication, analysis of access behavior, and fine-grained access control.

In the article "Distributed learning on mobile devices: A new approach to data mining in the Internet of Things," Zhang *et al.* focus on training deep learning models in multiple mobile devices that learn a shared model collaboratively. Specifically, the novel architecture GREAT is proposed and implemented using Alpha-GossipSGD, a dynamic control algorithm. Performance evaluation shows that Alpha-GossipSGD can realize stable learning effectiveness over unreliable networks with constrained resources.

The article "Detection of anomalies in Industrial IoT systems by data mining: Study of CHRIST Osmotron water purification system," by Sadeghi Garmaroodi *et al.*, uses two machine-learning-based approaches for detecting anomalies or faults at the edge in a real-world industrial automation system, i.e., CHRIST Osmotron water purifier. Extensive experiments show the accuracy of the implemented data-driven and model-based anomaly detection methods.

In their article "PPCS: An intelligent privacy-preserving mobile-edge crowdsensing strategy for Industrial IoT," Wang *et al.* define the novel RL-based PPCS strategy for IIoT. PPCS provides a Kullback–Leibler privacy-preserving-based data aggregation utilizing an incentive mechanism to solve the problem of noise selection. Theoretical analysis and validation experiments indicate that the weighted average of aggregated data established by PPCS has a better aggregation accuracy than contemporary strategies.

Digital Object Identifier 10.1109/JIOT.2021.3075304

The article "Multipath TCP meets transfer learning: A novel edge-based learning for Industrial IoT," by Pokhrel *et al.*, presents the design of a novel distributed transfer learning framework to maximize multipath communication networking performance for Industry 4.0 environment. The framework is evaluated using mathematical-based models and NS-3 simulator. Results show that it can potentially satisfy Industry 4.0 time-evolving network scenarios, low delays, and reliability constraints.

Kasi *et al.* in "Heuristic edge server placement in Industrial Internet of Things and cellular networks" focus on the edge server placement problem. Specifically, the problem is addressed within an existing network infrastructure obtained from Shanghai Telecom's base station data set. A genetic algorithm is used to solve the edge placement problem formulated as a multiobjective constraint optimization problem. The experimental evaluation demonstrates that the genetic algorithm quickly reaches to a solution state such that it reduces access delay in the network and maintains the workload balance between edge servers with local information.

In "A deep-learning-based smart healthcare system for patient's discomfort detection at the edge of Internet of Things," Ahmed *et al.* present an IP camera-based noninvasive automated patient's discomfort monitoring/detection system that is implemented using a deep-learning-based algorithm, named Mask-RCNN. Through top-view cameras, the patient's body position and posture are continuously monitored, based on which comfort and discomfort level can be discriminated. The experimental results show that the system achieves a very good accuracy.

The article "Adversarial attacks against network intrusion detection in IoT systems," by Qiu *et al.*, deals with the design of a novel adversarial attack against deep-learning-based network intrusion detection systems (NIDS) in the IoT environment, with only black-box accesses to the deep-learning-based model. Adversarial examples using conventional methods are generated. The technique is applied to compromise one state-of-the-art NIDS, Kitsune.

In their article "TinyRadarNN: Combining spatial and temporal convolutional neural networks for embedded gesture recognition with short range radars," Scherer *et al.* propose a low-power high-accuracy embedded hand-gesture recognition algorithm targeting battery-operated wearable devices using low-power short-range RADAR sensors. The results show the effectiveness and potential of RADAR-based hand-gesture recognition for embedded devices, as well as the network design, using the temporal convolutional neural network approach.

The article "An IoT-based hedge system for solar power generation," authored by Syu *et al.*, proposes a hedging system to hedge the low-radiation risk for solar investors through the designed IoT-based data, edge-based models for predicting solar radiation as well as hedging options. The precise prediction model is proposed to predict solar radiation in high precision, and the light predictive model is proposed to meet the low latency and less computational cost on the edge system.

In the article "Ultra large-scale crowd monitoring system architecture and design issues" by Jiang *et al.*, the authors propose a novel ultra large-scale crowd monitoring system, namely the ULCM system. The presented ULCM system enables advanced sensing and networking technologies aimed at collecting and processing multimodal, multiperspective, and real-time crowding data relevant to the crowd management. To achieve a full comprehensive scene overview, the ULCM deployment utilizes multiplicity of UAV agents in different operational scenarios. The aerial deployment and control is realized by custom multiple UAVs network and airborne LiDAR sensors. The deployment and control on the ground sensory agents is based on multiple sub-networks including CCTV, infrared gas, and ultrasonic sensors networks. Eventually, ULCM employs SDN and edge cloud technologies to optimize the networking and data analytics performance from the infrastructure perspective.

The article "VID-WIN: Fast Video Event Matching With Query-Aware Windowing at the Edge for the Internet of Multimedia Things" by Yadav *et al.* presents a novel approach, named VID-WIN, to accelerate video event analytics, specifically CEP queries, in an edge-cloud paradigm in the context of Internet of Multimedia Things. The approach expands the concept of content-driven widowing from structured data streams to unstructured video data via the input transformation technique. Extensive experiments on real-world video data sets show the efficacy of VID-WIN-based windowing on edge resources and faster execution time with minimal effect on query accuracy.

We would like to express our sincere thanks to all the authors for submitting their papers and to the reviewers for their valuable comments and suggestions that significantly enhanced the quality of these articles. We are also grateful to Prof. H. Wang, the Editor-in-Chief of the IEEE INTERNET OF THINGS JOURNAL, for his great support throughout the whole review and publication process of this special issue, and, of course, all the editorial staff. We hope that this special issue will serve as a useful reference for researchers, scientists, engineers, and academics in the field of data mining and machine learning at the "edge" of the IoT.

GIANCARLO FORTINO
Department of Informatics, Modeling, Electronics and Systems
University of Calabria
87036 Rende, Italy

RAJKUMAR BUYYA
Cloud Computing and Distributed Systems (CLOUDS) Laboratory
School of Computing and Information Systems
University of Melbourne
Melbourne, VIC 3010, Australia

MIN CHEN
School of Computer Science and Technology
Huazhong University of Science and Technology
Wuhan 430074, China

FRANCISCO HERRERA
Computer Science and Artificial Intelligence Department
Universidad de Granada
18010 Granada, Spain