Peer-to-Peer Networking and Applications (2021) 14:3891–3894
https://doi.org/10.1007/s12083-021-01239-7

Convergence of Edge Computing and Next Generation Networking

Deze Zeng[1] · Geyong Min[2] · Qiang He[3] · Song Guo[4]

Accepted: 9 August 2021 / Published online: 1 September 2021

The fast development of cloud computing has spurred a broad range of applications and services to facilitate our lives. However, due to the unexpected connectivity, delay and jitter between the end users and the cloud, many applications/services, especially those delay-sensitive ones (e.g., IoT applications), may occasionally suffer from Quality-of-Service (QoS) degradation. To address such problem, pioneering researchers introduce a new concept, i.e., edge computing, to facilitate these applications/services via the available resources at the network edge, i.e., pushing the cloud computing and services to the network edge. By exploring the proximity feature of the edge computing, it has been widely agreed that mobile edge computing is a promising alternative to provide a diversity of cloud services to the users with high QoS (service availability, latency, robustness, etc.).

On the other hand, with the fast development of information technology, the network is experiencing unprecedented transformation as both the number of connections and the volume of data are severely increased. Besides, a diversity of applications are with different communication requirements. To make the networks fitting for such ever-growing diverse needs, many new future networking technologies have been proposed, covering both the access network and the backbone network. For example, software-defined networking (SDN) decouples the control plane from the data plane to enable more flexible and customized network flow control. Network Function Virtualization (NFV) softwarizes the traditional hardware based network functions and make them be able to run on standard servers (e.g.,×86) servers, significantly improving the network flexibility. Information Centric Network (ICN) evolves current Internet infrastructure from host-centric paradigm to the data or service name centric paradigm. These newly emerging networking technologies have already been widely regarded as the key enabling technologies in future networks.

However, at the early stage of edge computing and the above mentioned newly emerging networking technologies, there are still many challenging issues to be addressed for the integration of them. Therefore, it is significant to investigate how to integrate these new networking technologies into edge computing for more efficient resource usage and better service provision. To this end, we organize this SI to collect the-state-of-the-art representative studies in edge computing and emerging networking technologies from different aspects. Our SI successfully attracted 27 submissions, each of which was rigorously reviewed by at least 3 reviewers. Finally, 11 papers were accepted, achieving acceptance ratio 40.7%. Generally, the accepted papers mainly focus on two aspects, the end-devices (e.g., IoT devices) and edge computing infrastructure. We briefly discuss the main contributions of these 11 papers as follows.

The first contribution entitled "Towards Efficient and Energy-aware Query Processing for Industrial Internet of Things", Liu et al. advocate that the Industrial IoT (IIoT) devices are typically deployed in noisy environments, requiring flexibility and efficiency of network management. They propose a secure and efficient algorithm with trust and energy awareness, called Trust and Energy-aware based Holistic optimizing Algorithm for Spatial window query (TEHAS). It only requests part of nodes in the query area to forward query messages and return sensing data to sink, while sensing data of the query area can be returned directly through the geo-routing protocol instead of using the coordinator node. They also design a trust management and energy awareness mechanism during the cluster head selection, which can enhance network security and prolong the network lifetime.

This article is part of the Topical Collection: *Special Issue on Convergence of Edge Computing and Next Generation Networking*

Guest Editors: Deze Zeng, Geyong Min, Qiang He, and Song Guo

✉ Deze Zeng
deze@cug.edu.cn

[1] China University of Geosciences, Wuhan, China

[2] University of Exeter, Exeter, UK

[3] Swinburne University of Technology, Melbourne, Australia

[4] The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Experimental results demonstrate that TEHAS could outperform similar algorithms like SWIF and ESRT.

In the second paper "An Efficient Moving Object Tracking Framework for WSNs using Sequence-to- Sequence Learning Model", Cao et al. study the moving object tracking problem in wireless sensor networks. Toward the goals of (1) locating and tracking a moving target without using any GPS-based sensors, (2) delivering the wake-up and resetting message to the area around the target's path quickly and energy efficiently, and (3) preserving the location anonymity of the sensors tracking objects, the authors propose an energy efficient framework. The proposed framework predicts the trajectory of the moving object using a Sequence-to-Sequence learning (Seq2Seq) model and only wakes-up the sensors that fall within the predicted trajectory of the moving object with a specially designed control packet. Different from traditional moving object tracking methods, the Seq2Seq model reduces the computation time of encoding geospatial trajectory and the proposed framework preserves the location anonymity by only transmitting the hop's information instead of GPS values.

Besides dedicatedly deployed sensors, crowdsensing is also significant to data gathering. In the third paper "An Incentive Mechanism Based on Endowment Effect Facing Social Welfare in Crowdsensing", based on the theory of behavioral economics, the authors design an effective incentive mechanism from the point of view of reducing cost and improving social welfare. In the current research, the incentive purpose is mainly achieved by paying certain rewards to service providers. However, due to the demand for data quality and quantity, platforms often have high consumption. The incentive mechanism based on the traditional economic theory ignores the node's value evaluation strategy. Therefore, this paper constructs the mapping relationship between the theory of behavioral economics and MCS, designs the RCBEE. This paper analyzes the change of the node's value evaluation strategy, changes the node's return, and reconstructs the return matrix. The simulation results show that compared with the traditional incentive mechanism, RCBEE reduces the payment cost and improves the social welfare.

Communication performance is critical to the overall system. The paper "Robust Multi-agent Reinforcement Learning for Noisy Environments" by Chen et al. focuses on the robustness of multi-agent learning in noisy environments where error observations may occur and lead to training failure. To improve the robustness of policy learning, the authors designed MAFTRL, a robust multi-agent reinforcement learning algorithm. MAFTRL proposes (i) an effective error detection mechanism based on an autoencoder to reduce the impact of environmental noise, and (ii) an attention communication medium to avoid the propagation of error observations. Experiments on multi-agent tasks show that MAFTRL has good performance and strong robustness in both reliable and noisy environments.

D2D communication is widely advocated to promote the communication efficiency in an infrastructureless way. The paper "Deep Reinforcement Learning-Based Incentive Mechanism Design for Short Video Sharing through D2D Communication" authored by Zhuo Li et al. studies how to dynamically motivate mobile user equipment(MUEs)to engage in short video sharing while ensuring the QoS. To address the three main challenges of (i) the selfishness of MUEs, (ii) the different levels of interest in shared content, and (iii) the constantly changing environment as the process proceeds, Dynamic Incentive Mechanism algorithm of D2D-based Short Video Sharing based on Asynchronous Advantage Actor-Critic (DIM-A3C) is proposed. DIM-A3C shows good performance in terms of convergence speed and user satisfaction. Simulation results show that DIM-A3C can increase the utility of mobile edge computing server by an average of 22% and 16%, compared with the existing proportional incentive mechanism (PIM) and scoring-based incentive mechanism (SIM). Meanwhile, DIM-A3C achieves a higher degree of satisfaction than PIM and SIM.

Edge computing, residing between the IoT and the cloud, is essential for IoT data gathering. In the paper "Energy-Efficient Sensory Data Gathering in IoT Networks with Mobile Edge Computing", Ren et al. study the energy-efficient sensory data gathering in mobile IoT networks to reduce energy consumption of resource-limited IoT devices, and prolong the network lifetime. To address the three challenges of (i) the large-scale and spatial–temporal evolutional characteristic of IoT networks, (ii) unstable data gathering network topology and uneven distribution of smart things in IoT networks, and (iii) the planning of the gathering data path in the edge network, an Energy-efficient sensory Data Gathering mechanism with Mobile edge computing (EDGM) is proposed. EDGM includes two stages: (i) an edge node level strategy, which evaluates the level value of edge nodes and divides the edge network; and (ii) a data gathering path of the mobile edge node that gathers sensory data from each single edge network, which can decrease energy consumption along data gathering path. Prototype implementation and experiments have conducted on EdgeCloudSim simulator. Experimental results show that EDGM can effectively reduce energy consumption, in comparison with benchmark techniques.

Vehicular applications are quite suitable to be deployed in edge computing infrastructure. In the paper "Joint Computation Offloading and Resource Allocation in Vehicular Edge Computing Based on An Economic Theory: Walrasian Equilibrium", Wang et al. study joint computation offloading and resource allocation in Vehicular Edge Computing (VEC). Limited computation resources lead to the existence of a mismatch between the resources requested by vehicles and the resources that can be allocated by servers, which

increases the computation delay. To address the problem, the authors model the computation offloading and resource allocation as a VEC network welfare maximization optimization problem based on Walrasian Equilibrium. An algorithm with a fast convergence rate is proposed to solve the optimization problem to find the Walrasian Equilibrium in VEC. Simulation results demonstrate that the proposed algorithm can eliminate the mismatch between the demand of offloading and computation resource, and can obtain the Walrasian equilibrium of VEC networks. Meanwhile, the total offloading delay is decreased, which is benefit for VEC servers and vehicles for saving energy.

Recommendation services can be deployed at the edge to satisfy the users nearby by exploiting the low latency advantage. The data quality is a critical issue to the overall recommendation quality. The paper "From edge data to recommendation: A double attention-based deformable convolutional network" by Li et al. effectively alleviates the problem of data sparsity in recommender systems by informative data obtained from applications of edge devices and deep learning technologies with powerful data processing ability. The proposed model DADCN not only applies the parallel deformable convolutional networks to flexibly capture the deep sematic features of users and items, but also adopts the word-level and review-level attention mechanisms to intensify the critical words and informative reviews by assigning relatively large attention weights to them. The results of extensive experiments conducted on four real-world datasets validate the effectiveness of DADCN.

Besides computing, data caching by edge computing infrastructure is also promising for QoS improvement. The paper "Video Placement and Delivery in Edge Caching Networks: Analytical Model and Optimization Scheme" by Wu et al. addresses the conflict between limited wireless resources of edge caching network and repeated requests from different users. With the help of stochastic geometry theory, the successful transmission probability, energy efficiency and system delay are first modeled under three caching policy, i.e., (i) caching based on SVC, (ii) caching based on quality version and (iii) caching based on maximum quality level. Based on the obtained models, a genetic-based video placement and delivery algorithm is studied to balance the energy efficiency and system average delay. Extensive simulations reveal the impacts of cache capacity, video file popularity distribution as well as weighted factor on the mean system delay and energy efficiency. Moreover, the effectiveness of the proposed scheme is also validated.

ICN is a potential supporting technology to edge computing. In the paper "Comparative Analysis of Probabilistic Forwarding Strategies in ICN for Edge Computing", Zhang et al. use NDN forwarding strategy to solve the service invocation problem in edge computing. Due to the dynamic nature, how the requester knows and calls surrounding services is a key issue. The flexibility brought by the NDN forwarding strategy is inherently suitable for service discovery in highly dynamic scenarios. The service invocation problem in NDN edge computing is abstracted into a forwarding path selection problem, and a probabilistic forwarding framework based on the coefficient of variation method (CVPF) is proposed based on the multi-attribute decision model. The experiments conducted on ndnSIM show that CVPF improves network performance in comparison with benchmark techniques, and CVPF is more suitable for frequent and dynamic small flow requests in edge computing.

With the emergence and hotness of edge intelligence, the paper "A DNN Inference Acceleration Algorithm Combining Model Partition and Task Allocation in Heterogeneous Edge Computing System" by Shi et al. discusses how to partition a large DNN model and allocate the tasks in heterogeneous edge computing systems. Aiming to minimize the average task execution time, the authors first establish a mathematical model of adaptive DNN model partition and task offloading. Due to the involvement of a large number of binary variables, the optimization model is hard to solve to obtain the optimal solution. Hence, the authors then propose Partition-Points- Selection (PPS) algorithm and Greedy Strategy for Progressive Inference (GSPI) algorithm to pursue sub-optimal solutions. The authors use multiple DNN models and CIFAR-10 data set to practically run on the real systems to get raw data needed for experiments. Based on these raw data, experiment results verify the effectiveness and efficiency of the proposed algorithms.

Deze Zeng is currently a Full Professor in School of Computer Science, China University of Geosciences, Wuhan, China. His current research interests include: network function virtualization, software-defined networking, cloud computing and edge computing. He has authored 1 book and over 80 papers in refereed journals and conferences in these areas. He also received 5 best paper awards from IEEE/ACM conferences and journals. He serves in editorial boards of Journal of Network and Computer Applications and guest editors of many prestigious journals. He has been the in organization or program committees of many international conferences including ICPADS, ICA3PP, CollaberateCom, MobiQuitous, ICC, Globecom. He is a member of IEEE.

Geyong Min is a Professor of High Performance Computing and Networking in the Department of Computer Science within the College of Engineering, Mathematics and Physical Sciences at the University of Exeter, United Kingdom. He received the PhD degree in Computing Science from the University of Glasgow, United Kingdom, in 2003, and the B.Sc. degree in Computer Science from Huazhong University of Science and Technology, China, in 1995. His research interests include Future Internet, Computer Networks, Wireless Communications, Multimedia Systems, Information Security, High Performance Computing, Ubiquitous Computing, Modelling and Performance Engineering.

Qiang He is a Senior Lecturer at Swinburne University of Technology. He received his first PhD degree from the Swinburne University of Technology, Australia, and the second PhD degree as well his BSc degree from Huazhong University of Science and Technology, China. Dr. He' research interests mainly include edge computing, cloud computing, software engineering, and service computing. He has authored or coauthored papers published on leading conferences and journals, including IEEE TPDS, IEEE TSE, IEEE TCC, IEEE TSC, IEEE TBD, IEEE ICSE, WWW, ICDE, IJCAI, ICDM, ICSOC, ICWS and CLOUD.

Song Guo is a Full Professor and Associate Head at Department of Computing, The Hong Kong Polytechnic University. Before joining PolyU, he was a professor at the University of Aizu, Japan. He obtained his PhD degree in Computer Science from University of Ottawa, Master's degree in Computer Engineering from Beijing University of Posts and Telecommunications, and Bachelor degree in Computer Software from Huazhong University of Science and Technology. His research interests are mainly in the areas of big data, cloud computing, mobile computing, and distributed systems. Some of his Transactions papers were selected as Featured or Spotlight papers. He is the recipient of the 2019 IEEE TCBD Best Conference Paper Award, 2018 IEEE TCGCC Best Magazine Paper Award, 2019 & 2017 IEEE Systems Journal Annual Best Paper Award, and other 6 Best Paper Awards from IEEE/ACM conferences. His work was also recognized by the 2016 Annual Best of Computing: Notable Books and Articles in Computing in ACM Computing Reviews. Prof. Guo is an IEEE Fellow (Computer Society) and the Editor-in-Chief of IEEE Open Journal of the Computer Society. He was a Distinguished Lecturer of IEEE Communications Society (ComSoc) and served in the IEEE ComSoc Board of Governors. Prof. Guo has been named on editorial board of a number of prestigious international journals like IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Cloud Computing, IEEE Transactions on Emerging Topics in Computing, etc. He has also served as chairs of organizing and technical committees of many international conferences.