Journal of Systems Architecture 118 (2021) 102186

Contents lists available at ScienceDirect

Journal of Systems Architecture

journal homepage: www.elsevier.com/locate/sysarc

Editorial

Editorial to special issue on resource management for edge intelligence

Due to rapid growth of Internet of Things (IoT) applications that demand large amounts of computing power, resource-constrained IoT devices typically do not have enough computing capability to satisfy the computational requirements. For example, inference or training of large machine learning models, e.g., Deep Neural Networks (DNNs), are typically quite resource-intensive. Computational offloading for edge intelligence can help overcome resource constraints of IoT devices, reduce computing load and improve task processing efficiency. Edge Computing and Fog Computing aim to harness computing resources in the proximity of the IoT devices to provide services in a timely manner with reduced accesses to the cloud, which may be time-consuming. For example, a DNN can be partitioned and processed collaboratively between the device, edge server and the cloud to optimize objectives such as performance and power consumption, based on factors such as workload, hardware platform capabilities and wireless bandwidth. This special issue includes twelve papers that address a variety of topics surrounding resource management for edge intelligence.

Chen et al. [1], propose a lightweight and robust learning algorithm in a dynamic network allowing topology changes. In their model, each agent is assumed to have only limited memory and communicate with each other asynchronously. Rigorous analysis shows that despite these harsh constraints, the best arm/option can be identified collaboratively by the agents and the algorithm converges efficiently. Li et al. [2], present theoretical analysis to the overall throughput of establishing multiple subflows with Multipath TCP in mobile edge networks, based on which they propose an adaptive subflow allocation algorithm to determine the proper number of subflows for each network interface to optimize the performance gain. Liu et al. [3] propose a new task offloading scheduler, Horae, to not only improve the resource utilization of MEC environment but also guarantee to select the edge server which could satisfy placement constraints for each offloaded task. Concretely, considering the fact that each job would experience slack time as a result of competing for limited resource with other jobs in MEC, Horae minimizes the sum of all slack time values of all the jobs while guaranteeing placement constraints, and therefore improve the resource utilization of the system.

Wu et al. [4] consider the online edge user allocation (EUA) problem where edge users' resource demands arrive and depart dynamically. They consider the long-term edge user allocation rate, edge server hiring cost, and edge server energy consumption as allocation targets from the mobile application vendor perspective, and propose a decentralized reactive approach by employing a fuzzy control mechanism to yield the real-time allocation decisions. Wei et al. [5] present a new caching algorithm, called Similarity-Aware Popularity-based Caching (SAPoC), to promote the performance of wireless edge caching in dynamic scenarios through utilizing the similarity among contents. In SAPoC algorithm, a content's popularity is determined by not only its requests history but also its similarity with existing popular ones to enable a quick-start of newly arrived contents. Wang et al. [6] formulate an optimization problem for dynamic resource allocation and automatically select the best partition point to minimize the overall latency of all vehicles, which is NP-hard. Then they design a chemical reaction optimization based algorithm for low complexity to solve the problem. Wang et al. [7] present an edge simulator called SimEdgeIntel for resource management that opens up detailed configuration options, enabling researchers quickly deploy mobile with edge intelligence. It supports researchers to customize the development of mobility models, caching algorithms and switching strategies. Kuang et al. [8] propose a joint iterative algorithm based on the Lagrangian dual decomposition, ShengJin Formula method, and monotonic optimization method. The CPU cycle frequency allocation is handled by the ShengJin Formula method due to the cubic equation of one variable about the CPU frequency allocation. The transmission power assignment is handled by the monotonic optimization method.

Xu et al. [9] utilize RL technique to design a privacy-preserving incentive mechanism for multiple providers and multiple IoT devices. Specifically, they model the pricing and demand problem of providers and IoT devices as a multi-leader multi-follower Stackelberg game, in which the providers work as leaders to determine their prices first, and then the IoT devices determine their demands as followers. They prove the existence and uniqueness of the Nash Equilibrium (NE) of this game. Liu et al. [10] propose a new privacy-preserving resource-trading scheme (PRTS), which leverages the concept of homomorphic cryptography and asymmetric searchable encryption, to simultaneously protect the privacy of the equipment factory and parts factories, while supporting best matching results in terms of parts parameters and price. Wang et al. [11] establish an audit model based on a designed binary tree assisted by edge computing, which provides computing capability for the resource-constrained terminals. The data pre-processing task is offloaded to the edge, which reduces computing load and improves the efficiency of task processing. They propose an improved correlation mechanism between data blocks and nodes on the binary tree so that all nodes on the binary tree can be fully utilized while existing methods use only leaf nodes and thus are required to establish multiple binary trees. Wang et al. [12] first propose a novel handover authentication model of ITS with multi-server edge computing architecture. Second, the proposed handover authentication scheme allows the authenticated server to assist users subsequently authenticate with other server, thereby achieving interactions with the server anytime and anywhere with low overhead. Finally, blockchain technology and strong anonymity

https://doi.org/10.1016/j.sysarc.2021.102186

Available online 27 May 2021

mechanism are introduced to protect users' privacy strictly.

References

[1] S. Chen, Y. Tao, D. Yu, F. Li, B. Gong, Distributed learning dynamics of multi-armed bandits for edge intelligence, J. Syst. Archit. 114 (2021) 101919.

[2] W. Li, L. Yu, C. Xue, H. Zhang, J. Lu, R. Cao, S. Lu, Adaptive subflow allocation for multipath data transmission in mobile edge networks, J. Syst. Archit. 116 (2021) 102012.

[3] B. Liu, X. Xu, L. Qi, Q. Ni, W. Dou, Task scheduling with precedence and placement constraints for resource utilization improvement in multi-user MEC environment, J. Syst. Archit. 114 (2021) 101970.

[4] C. Wu, Q. Peng, Y. Xia, Y. Ma, W. Zheng, H. Xie, S. Pang, F. Li, X. Fu, X. Li, et al., Online user allocation in mobile edge computing environments: a decentralized reactive approach, J. Syst. Archit. 113 (2021) 101904.

[5] X. Wei, J. Liu, Y. Wang, C. Tang, Y. Hu, Wireless edge caching based on content similarity in dynamic environments, J. Syst. Archit. 115 (2021) 102000.

[6] Q. Wang, Z. Li, K. Nai, Y. Chen, M. Wen, Dynamic resource allocation for jointing vehicle-edge deep neural network inference, J. Syst. Archit. (2021) 102133.

[7] C. Wang, R. Li, W. Li, C. Qiu, X. Wang, Simedgeintel: a open-source simulation platform for resource management in edge intelligence, J. Syst. Archit. 115 (2021) 102016.

[8] Z. Kuang, Z. Ma, Z. Li, X. Deng, Cooperative computation offloading and resource allocation for delay minimization in mobile edge computing, J. Syst. Archit. (2021).

[9] H. Xu, X. Qiu, W. Zhang, K. Liu, S. Liu, W. Chen, Privacy-preserving incentive mechanism for multi-leader multi-follower IoT-edge computing market: a reinforcement learning approach, J. Syst. Archit. 114 (2021) 101932.

[10] P. Liu, K. Liu, T. Fu, Y. Zhang, J. Hu, A privacy-preserving resource trading scheme for cloud manufacturing with edge-PLCs in IIot, J. Syst. Archit. 117 (2021) 102104.

[11] T. Wang, Y. Mei, X. Liu, J. Wang, H.-N. Dai, Z. Wang, Edge-based auditing method for data security in resource-constrained internet of things, J. Syst. Archit. 114 (2021) 101971.

[12] W. Wang, H. Huang, L. Xue, Q. Li, R. Malekian, Y. Zhang, Blockchain-assisted handover authentication for intelligent telehealth in multi-server edge computing environment, J. Syst. Archit. 115 (2021) 102024.

Shaohua Wan[*,a], Huaming Wu[b], Joarder Kamruzzaman[c], Sotirios Goudos[d]

[a] *Zhongnan University of Economics and Law, China*

[b] *Tianjin University, China*

[c] *Federation University Australia, Australia*

[d] *Aristotle University of Thessaloniki, Greece*

[*] Corresponding author.

E-mail address: shwanhust@gmail.com (S. Wan).