Load-Aware Computation Offloading with Privacy Preservation for 5G Networks in Edge Computing

Xiaolong Xu[1,2], Xihua Liu[1,2], Xuyun Zhang[3], Lianyong Qi[4], and Yuan Yuan[5(✉)]

[1] School of Computer and Software,
Nanjing University of Information Science and Technology, Nanjing, China
`njuxlxu@gmail.com, liuxihua710@gmail.com`
[2] Jiangsu Collaborative Innovation Center of Atmospheric Environment
and Equipment Technology (CICAEET),
Nanjing University of Information Science and Technology, Nanjing, China
[3] Department of Electrical and Computer Engineering,
University of Auckland, Auckland, New Zealand
`xuyun.zhang@auckland.ac.nz`
[4] School of Information Science and Engineering,
Qufu Normal University, Qufu, China
`lianyongqi@gmail.com`
[5] Department of Computer Science and Engineering,
Michigan State University, East Lansing, MI, USA
`yyuan@msu.edu`

Abstract. Nowadays, with the advances in wireless communication, the mobile devices are becoming important due to various applications which provide mobile users with plentiful services in the devices. The mobile devices can hardly complete all the computing tasks as they have limitations on the battery capacity, physical size, etc. In order to release these limitations, in the fifth generation (5G), the computing tasks can be offloaded from the mobile devices to the central units (CUs) which are enhanced into edge nodes (ENs) for processing. However, it is still a problem to select the appropriate offloading destination, aiming to improve the load balance for all the ENs. In this paper, we first formulate an optimization problem to improve the load balance of all the ENs for 5G networks in edge computing, considering the time consumption and the privacy conflicts. Then, a load-aware computation offloading method with privacy preservation, named LCOP, is designed. Finally, experimental results and evaluations validate our proposed method is both effective and feasible.

Keywords: Mobile devices · 5G · Edge nodes · Load balance · Privacy conflicts

1   Introduction

In recent years, along with the advances in the mobile communication technologies, a increasing number of mobile users are attracted to enjoy the services supported by the mobile devices, which leads to an unprecedented growth in the mobile data traffic. Based on the analysis of Cisco's networking visual index report, it is expected that the data traffic will grow at the rate of 57%, a tenfold increase over 2014, by 2019 [1]. In order to cope with this condition, fifth generation (5G) is developed [2]. In 5G networks, the theoretical peak transmission speed can reach more than $10\,\mathrm{Gb/s}$, which is hundreds of times faster than that of 4G networks [3].

Generally, the computing resources of mobile devices are limited due to the restrictions like physical size, battery capacity, etc. Therefore, the computing tasks from the mobile devices need to be offloaded to the cloud via the distributed units (DUs) and central units (CUs) in 5G networks [4]. However, on condition that the tasks are executed on the remote cloud, the efficiency of processing the computing tasks will be affected. However, driven by the edge computing technology, the CUs can be enhanced into edge nodes (ENs) to provide storage and process ability [5]. The computing tasks which do not need to be dealt with immediately are offloaded to the remote cloud for processing, and the other tasks which have high priority are executed in the ENs. In this way, the experience of the mobile users is greatly improved.

However, the offloading process among ENs in 5G networks is imposed with several weaknesses, especially the security and privacy issues [6,7]. Some important security information, including current location, remote video, voice chat, etc., may also need to be offloaded from the mobile devices for processing [8]. On condition that the privacy information is offloaded from the mobile devices, the networks can obtain the contents of these information, which may lead to the risk of privacy leakage. The disclosure of private information will bring out many terrible problems. Therefore, it is of utmost significance to avoid privacy leakage during the computation offloading in 5G networks [9].

On the other hand, while the CUs have been enhanced into ENs which have data-processing ability in 5G networks, some computing tasks offloaded from the DUs need to be migrated to the other ENs for resource response due to the limited ENs computing power, which may lead to the number of the tasks hosted on ENs are not nearly equal. Hence, from the perspective of ENs, the load balance should be taken into consideration. The load balance makes differences in the throughput, data-processing capability, flexibility and availability of the networks in 5G. Hence, it is important and necessary to make an appropriate strategy for the computing tasks to find a better route to be offloaded across ENs [10].

To improve the response time and efficiency for executing the computing tasks in 5G networks, the ability of CUs has been expanded into ENs to provide storage and computation power for the tasks. However, due to the privacy conflicts of the datasets for running the tasks, some tasks cannot be offloaded to the same EN for execution to avoid the privacy leakage. Furthermore, the

resource limitations of ENs need to be taken into consideration when the ENs are employed to execute the computing tasks. On condition that the computing task hosted on the first EN need to be offloaded to the other ENs, the load balance of all the ENs in 5G networks is also a significant problem waiting to be solved.

The main contributions of this paper are as follows:

(1) Analyze the load balance during the offloading of all the CUs, and the computation offloading problem with privacy preservation for 5G in edge computing is defined as a standard simple objective optimization problem.
(2) A load-aware computation offloading method with privacy preservation (LCOP) is adopted to realize the optimization to improve the load balance of all the CUs while guarding against privacy conflicts of the mobile applications.
(3) Conduct comprehensive experiments and evaluations to demonstrate the effectiveness and efficiency of our proposed method.

The reminder of this paper is organized as follows. Section 2 describes the mathematical modeling and the formulation. Section 3 develops a load-aware computation offloading with privacy preservation method for 5G networks in edge computing. In Sect. 4, simulation experiments and comparison analysis are presented. In Sect. 5, the related work is summarized. Finally, conclusions and future work are outlined in Sect. 6.

2 System Model and Problem Formulation

In this section, we establish the offloading task in 5G infrastructure. Besides, the resource usage and the load balance of EN are also analyzed.

Table 1. Key notations and descriptions.

Terms	Descriptions
MD	The set of mobile devices, $MD = \{md_1, md_2, md_3, \dots, md_k\}$
K	The number of data-processing tasks and mobile devices
DA	The set of data-processing tasks, $DA = \{da_1, da_2, da_3, \dots, da_k\}$
Q	The number of DUs
SD	The set of DUs, $SD = \{sd_1, sd_2, sd_3, \dots, sd_q\}$
N	The number of ENs
SC	The set of ENs, $SC = \{sc_1, sc_2, sc_3, \dots, sc_n\}$
RV	The set of running VM, $RV = \{rv_1, rv_2, rv_3, \dots, rv_m\}$
RU_n	The resource utilization of n-th computing nodes
ALB	The overall load balance variance
UT	The urgency of the task in mobile devices
AT	The time consumed by the task migration

2.1   Resouce Model

We assume that there are k mobile devices, k data-processing tasks, Q DUs and N ENs in 5G, denoted as $MD = \{md_1, md_2, \ldots, md_k\}$, $DA = \{da_1, da_2, \ldots, da_k\}$, $SD = \{sd_1, sd_2, sd_3, \ldots, sd_q\}$ and $SC = \{sc_1, sc_2, \ldots, sc_n\}$, respectively. Consider a scenario, only one server is deployed in each EN. Besides, we assume that there are M VMs running in ENs, denoted as $RV = \{rv_1, rv_2, \ldots, rv_m\}$ (Fig. 1).
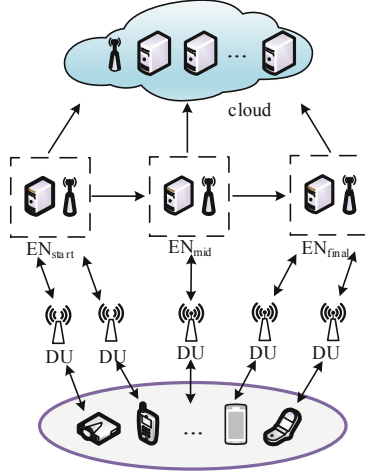
Fig. 1. The framework of migrating tasks in 5G.

2.2   Load Balance Analysis of ENs

VMs are rented for resource allocation in the ENs. The data-processing tasks from mobile devices are hosted by VMs. Note that the resource requirement of datasets and the capacity of servers in this paper are weighed by the number of VMs.

The resource usage is a pivotal index to measure the performance of the ENs. Note O_n as the variable to estimate whether the sc_n $(1 \leq n \leq N)$ is occupied, which is determined by

$$O_n = \begin{cases} 1, \text{if } sc_n \text{ isoccupied}, \\ 0, \text{otherwise}. \end{cases} \tag{1}$$

Then the OS_c^n is a binary variable to judge whether the dataset DA_c $(1 \leq c \leq C)$ is deployed on sc_n, which is defined by

$$OS_c^n = \begin{cases} 1, \text{if } DA_c \text{ is placed on } sc_n, \\ 0, \text{otherwise}. \end{cases} \tag{2}$$

The average resource usage of the EN is calculated according to the number of servers which are occupied. The number of occupied servers, denoted as MS, is determined by

$$MS = \sum_{n=1}^{N} OS_c^n. \tag{3}$$

The resource utilization represents the usage of the VM instances. The resource utilization of sc_n is calculated by

$$RU_n = \frac{1}{\alpha_n} \sum_{c=1}^{C} v_c \cdot O_n, \tag{4}$$

where α_n represents the capacity of the n-th EN and v_c represents the consumed number of VMs for dataset DA_c.

Finally, the average resource usage of the EN is calculated by

$$AU = \frac{1}{MS} \sum_{n=1}^{N} RU_n \tag{5}$$

Then, we consider the load balance of ENs in 5G. Through calculating the variance of resource utilization, the load balance variance of sc_n is calculated by

$$LB = (AU - RU_n)^2. \tag{6}$$

For all ENs, the average load balance variance is calculated by

$$ALB = \frac{1}{MS} \sum_{n=1}^{N} LB \cdot OS_c^n. \tag{7}$$

2.3   Privacy Model of Computing Tasks

In 5G networks, the computing tasks from different mobile devices combine privacy conflicts. These tasks need different datasets to accomplish their targets. However, the datasets may have requirements of different privacy preservation. Therefore, some computing tasks cannot be deployed in the same EN for further process.

A graph $\gamma = (A, R)$, where A represents the set of computing tasks and R represents the set of conflicting relations, is used to model the privacy conflicts of computing tasks. In order to make sure the privacy information of the mobile devices, a pair of conflict relations $(a_k, a_{k'})\,(a_k, a_{k'} \in A)$ is incapable of the deployment in the same EN. The conflict computing tasks for a_n are obtained based on $ca_k = \{a_{k'} | (a_k, a_{k'}) \in R, k' = \{1, 2, 3, ..., K\}\}$.

The load-aware computation offloading strategy for the computing tasks is denoted as $OS = \{os_1, os_2, os_3, ..., os_N\}\,(os_n \in N)$, where os_n is the destination EN for hosting a_n.

Based on the acquired conflicting tasks set for hosting a_n, the deployed location a_n has the conflicting EN set, which is acquired by $cc_k = \{os_j | os_j \in ca_k, j = \{1, 2, 3, ..., |ca_k|\}\}$.

2.4 Time Consumption of Migration

On condition that the server of the first EN is full of tasks, we shall migrate the VMs from hosted EN to other ENs or the cloud. However, the process of migration consumes a quantities of transmission time.

$VF_m^{n,n'}$ is a binary variable that indicates whether the vm_m is migrated from sc_n to $sc_{n'}$ and the cloud, which is calculated by

$$VF_m^{n,n'} = \begin{cases} 1, \text{if } vm_m \text{ is migrated from } sc_n \text{ to } sc_{n'} \text{ and the cloud,} \\ 0, \text{otherwise.} \end{cases} \tag{8}$$

When the EN sc_n needs to be transferred between the EN, the time consumed by the AP and the VM is calculated by

$$TC = \sum_{m=1}^{M} \sum_{n'=1}^{N} OS_c^n \cdot VF_m^{n,n'} \cdot \frac{DS_m}{TE_1}, \tag{9}$$

where DS_m is the data size of the vm_m, and TE_1 is the transmission efficiency between the ENs.

Let $NU_{n,n'}$ be the number of ENs between the EN_{start} and the EN_{final}. The time consumed by migration from EN_{start} and the EN_{final} is calculated by

$$TM = \sum_{m=1}^{M} \sum_{n'=1}^{N} OS_c^n \cdot VF_m^{n,n'} \cdot \frac{DS_m}{TE_2} \cdot (NU_{n,n'} - 1), \tag{10}$$

where TE_2 represents the transmission efficiency between ENs.

If the VMs are migrated from the EN_{start} to the cloud, the time consumed by migration from EN_{start} and the cloud is calculated by

$$UC = \sum_{m=1}^{M} \sum_{n'=1}^{N} OS_c^n \cdot VF_m^{n,n'} \cdot \frac{DS_m}{TE_3}, \tag{11}$$

where TE_3 represents the transmission efficiency between EN and the cloud.

$flag$ is a binary variable that indicates where the vm_m is migrated, which is calculated by

$$flag = \begin{cases} 1, \text{if VMs in the EN}_{final}, \\ 0, \text{if VMs in the cloud.} \end{cases} \tag{12}$$

During the entire process of migration, the time consumption is calculated by

$$AT = \begin{cases} \sum_{n=1}^{N} 2TC + TM, \text{if } flag = 1, \\ \sum_{n=1}^{N} 2TC + UC, \text{if } flag = 0. \end{cases} \tag{13}$$

2.5   Problem Formulation

From the foregoing, the load balance of ENs is analyzed and quantified. In this paper, we aim to achieve the target of minimizing the load balance variance presented in (7). The formalized problem is formulated by

$$minALB. \tag{14}$$

$$\text{s.t.AT} \leq \varphi(t). \tag{15}$$

$$os_k \notin cc_k. \tag{16}$$

Formula (15) means that the time consumption should meet the time constraint which is defined as $\varphi(t)$ according to the data size, delay and distance. Besides, formula (16) means that the privacy constraints must be satisfied.

3   Load-Aware Computation Offloading Method with Privacy Preservation

In this paper, our goal is to make optimization of the load balance under several constraints. Compared with the traditional algorithms such as weighted coefficient method, GA has been widely used in the optimization problems because of its good robustness, parallel processing mechanism and global optimization.

3.1   Encoding

Firstly, the computing tasks should be encoded in this operation. The computing tasks which are offloaded from the mobile devices should be represented as a gene. All of the genes consist of the chromosome which represents the efficient offloading strategy for the computing tasks. In this paper, the chromosome is encoded in integer. In GA, the fitness functions are used to evaluate the pros and cons of each individual. Then, the inheritance opportunity is determined. In this paper, the fitness functions include one category: the average load balance variance (7). As is shown in (14), this method aims to make optimization of the load balance. The constraints are given by (15) (16), representing the time consumption should meet the time limitation, and some computing tasks cannot be deployed in the same EN for execution to meet the privacy constraint.

3.2   Initialization

In the operation of initialization, the related paraments including the size of population SP, the probability of crossover PC, the probability of mutation PM, the number of iterations NI and the size of archive SA need to be determined at first. For each offloading strategy, $OS^j = \{os_1, os_2, \ldots, os_K\}$, where OS^j represents the j-th chromosome in the population.

3.3   Selection

The selection operation selects those individuals which have better fitness from the current evolutionary group into the mating pool. The crossover operation and the mutation operation can only select individuals from the mating pool to generate a better population.

3.4   Crossover and Mutation

In the traditional single-point crossover operation, two parental chromosomes are combined to generate two new chromosomes. Firstly, one crossover point is selected in the crossover operation. Afterwards, two chromosomes are changed.

On condition that the offspring chromosomes perform no longer better than the parental chromosomes of them but do not reach global optimal solution, the premature convergence will take place. The mutation operation is selected to keep the individual diversity in the population. In the mutation operation, the probability of each gene which will mutate is exactly equal.

3.5   Method Review

In this paper, we aim at optimizing the average load balance for all ENs. GA-based algorithm is selected to solve this problem because of its good performance in the optimization problem. Firstly, all of the computing tasks are encoded. Both the fitness functions and the constraints are given for the load-aware computation offloading problem. Secondly, after the fine-grained fitness assignment strategy, better chromosomes are selected from the population, the environmental selection and mating selection. Besides, the crossover operation and mutation operation are leveraged to avoid the premature convergence and produce new better offspring chromosomes.

4   Experimental Evaluation

In this section, comprehensive experiments and simulations are conducted to evaluate the performance of our proposed load-aware computation offloading with privacy preservation method for 5G networks in edge computing. In this comparison, we denote the running state without migration as benchmark, first fit decreasing-based computation offloading is marked as FFD, and the abbreviation of our proposed method is marked as LCOP.

4.1   Experimental Context

In this experiment, LENOVO TS250 is configured as the server. The basic configurations of this server are Intel Xeon-E3-1225V6, Quad-Processor clocked at 3.4 GHz and 4 GB of RAM. In Table 2, there are five basic parameters and the range of the values in our experiment. For the effectiveness of this experiment, we set five different numbers of the computing tasks to generate five different scale datasets. The numbers of the computing tasks are set as 50, 100, 150, 200 and 250, respectively.

Table 2. Parameter settings.

Parameter description	Value
The number of computing tasks	50,100,150,200,250
The number of VMs on each EN	7
The number of running VMs on each cloudlet	[1,6]
The transmission rate between AP and the ENs	1200 Mb/s
The transmission rate between APs	540 Mb/s
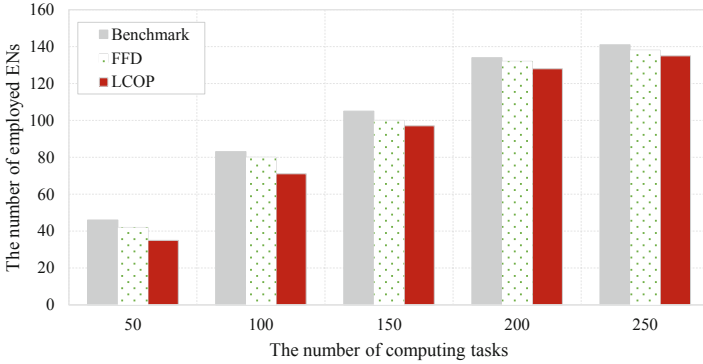
Fig. 2. The number of employed ENs.

Comparison of Employed ENs. In this part, the comparison of the whole employed ENs with Benchmark and FFD with the same experimental context is analyzed in detail. More ENs employed, lower resource utilization it has. Therefore, the number of the employed ENs is an important parameter which needs to be taken into consideration. Based on the outcome of this experiment, if the number of the computing tasks increases, the number of employed ENs will increase too. Besides, with the increase of the computing tasks, the increase of the number of employed ENs will be slow. In our proposed method, the number of computing tasks of 50, 100, 150, 200, 250 corresponds with the employed ENs of 35, 71, 97, 128 and 135, respectively. The number of the employed ENs is shown in Fig. 2.

Comparison of Resource Utilization. In this part, the comparison of the resource utilization with Benchmark and FFD with the same experimental context is analyzed. The resource utilization represents the percentage of the used VMs in ENs. If the resource utilization is low, the number of the employed ENs will increase. On this condition, the energy consumption will increase. Analyzed from our experiment, with the increase of the computing tasks, the resource utilization of LCOP will increase. In our proposed method, the number of computing tasks of 50, 100, 150, 200, 250 corresponds with the resource utilization of 59%, 63%, 67%, 70% and 75%, respectively. The resource utilization of ENs is shown in Fig. 3.
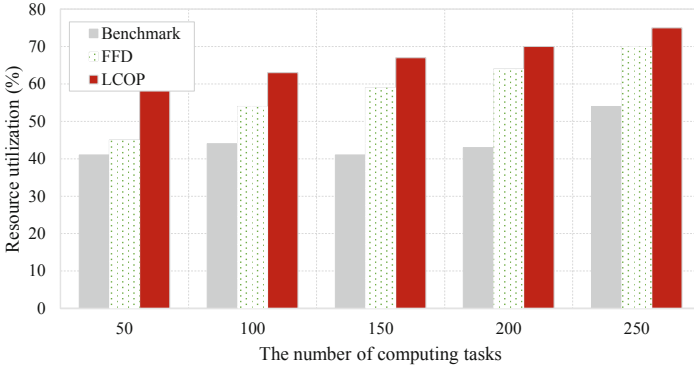
Fig. 3. The resource utilization of ENs.

Comparison of Load Balance. In this part, the comparison of the load balance with Benchmark and FFD with the same experimental context is conducted. The load balance represents the degree of the difference between the employed ENs. The load balance can increase the throughput, enhance the data-processing ability of the network, and increase the flexibility and availability of the network. Analyzed from our experiment, with the increase of the computing tasks, the value of the load balance in FFD and LCOP changes just a bit, and the value of the load balance in Benchmark changed a lot. In LCOP, the number of computing tasks of 50, 100, 150, 200, 250 corresponds with the load balance of 3%, 3%, 1%, 1% and 1%, respectively. The load balance for all the ENs is shown in Fig. 4.
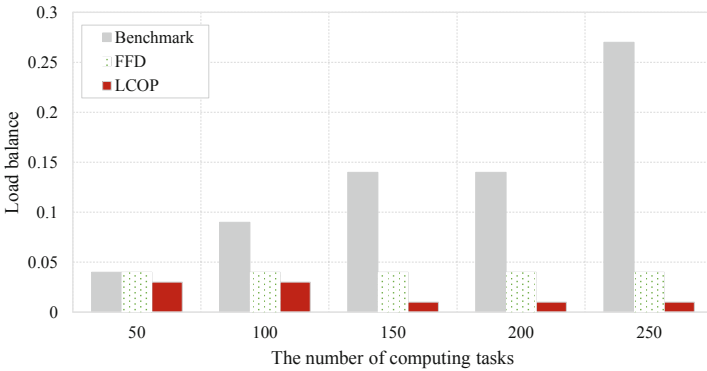

Fig. 4. The load balance for all the ENs

5   Related Work

Recently, due to advancements in wireless and mobile communication technologies, mobile devices such as tablets, laptops and smartphones are gaining great popularity in our daily life. There are various applications in these mobile devices such as remote video and GPS, which leading to quantities of data waiting to be solved [11,12]. Currently, in order to release the above problem, the 5th generation mobile network (5G), which can provide high data rate and low latency, is gaining much attention among the society [13,14]. The mobile applications can be offloaded from the mobile devices to the CUs or the cloud via DUs for execution in 5G networks.

However, some important information, including personal data and current locations, also needs to be offloaded for further processing. The privacy information may be leaked and the privacy leakage will bring a lot of problems with the customers [15–17].

In [18], Eiza et al. proposed a novel system for vehicular network in 5G, aiming to provide customers with a real-time, secure, reliable and privacy-aware video reporting service. In [19], Ni et al. introduced a secure and efficient service-oriented authentication framework. This network supports fog computing and network slicing for 5G-enabled IoT services. In [20], Fang et al. introduced the current features of different technologies and proposed a new 5G wireless security architecture in 5G networks.

If all the mobile applications are offloaded to the CUs without management, the load balance of the CUs will be very low. In this way, the data-processing capabilities, availability and flexibility of the 5G networks will be worse. Therefore, it is highly necessary for us to find the optimal migration strategy.

In [4], by implementing a mobile DBA and a fixed delay function firstly, Keita et al. illustrate the next generation-PON2 (NG-PON2) with low latency and small delay variation for 5G MFH. In [8], Zhang et al. proposed a mechanism for MEC named energy-efficient computation offloading (EECO) in 5G network, aiming to minimizing the energy consumption during the offloading. In [21], in order to release the end-to-end delay in 5G network, Chen et al. introduced a novel network architecture using data engine and a resource cognitive engine. Then, an optimal caching strategy was introduced for the macro-cell cloud and the small-cell cloud. In [22], Ketyko et al. summarized the NP-hard methods and problems related to load balancing, resource sharing, fairness and deployment among multiple users in 5G mobile networks. In [23], a statistical QoS-driven power adaptation scheme is proposed for the distributed caching assisted offloading scheme by Zhang et al. in 5G wireless networks. They also set up the system models for the partial in-network transcoding and the D2D assisted caching.

Based on the above analysis, the previous researches and studies have limitations on computation offloading for 5G networks in edge computing, and few works took the multiple-objective optimization for improving the load balance and protecting the privacy into consideration. In view of this challenge, a load-aware computation offloading with privacy preservation method for 5G networks in edge computing is proposed in this paper.

6   Conclusion and Future Work

In recent years, 5G has emerged as an important technology to release the mobile data traffic. The MEC paradigm, which is an effective paradigm, plays an important role in processing computing tasks in 5G networks. In order to optimize the load balance of all the ENs for 5G networks in edge computing while meet the constraints of the time consumption and the privacy conflicts of the computing tasks, a load-aware computation offloading with privacy preservation method is proposed in this paper. First, where the computing tasks are offloaded to is based on the priority of the computing tasks. Then, GA is leveraged to achieve the goal of optimizing the load balance. Subsequent experimental evaluations are conducted to verify the effectiveness and efficiency of our proposed method.

For future work, we plan to adjust and extend our proposed method to a scenario in the real-world and make investigation for other specific applications.

References

1. C. V. N. Index: Global mobile data traffic forecast update, 2014–2019, White Paper, 1 February
2. Dai, L., Wang, B., Yuan, Y., Han, S., Chih-Lin, I., Wang, Z.: Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends. IEEE Commun. Mag. **53**(9), 74–81 (2015)
3. Gupta, A., Jha, R.K.: A survey of 5G network: architecture and emerging technologies. IEEE Access **3**, 1206–1232 (2015)
4. Takahashi, K., et al.: NG-PON2 demonstration with small delay variation and low latency for 5G mobile fronthaul. In: 2017 European Conference on Optical Communication (ECOC), pp. 1–3. IEEE (2017)
5. Wang, X., Yang, L.T., Xie, X., Jin, J., Deen, M.J.: A cloud-edge computing framework for cyber-physical-social services. IEEE Commun. Mag. **55**(11), 80–85 (2017)
6. Li, S., Da Xu, L., Zhao, S.: 5G Internet of Things: a survey. J. Ind. Inf. Integr. **10**, 1–9 (2018)
7. Wang, X., Yang, L.T., Kuang, L., Liu, X., Zhang, Q., Deen, M.J.: A tensor-based big-data-driven routing recommendation approach for heterogeneous networks. IEEE Netw. **33**(1), 64–69 (2018)
8. Zhang, K., et al.: Energy-efficient offloading for mobile edge computing in 5G heterogeneous networks. IEEE Access **4**, 5896–5907 (2016)
9. Alliance, N.: 5G white paper, Next generation mobile networks, white paper, pp. 1–125 (2015)
10. Ren, L., Cheng, X., Wang, X., Cui, J., Zhang, L.: Multi-scale dense gate recurrent unit networks for bearing remaining useful life prediction. Future Gen. Comput. Syst. **94**, 601–609 (2019)
11. Beyranvand, H., Lévesque, M., Maier, M., Salehi, J.A., Verikoukis, C., Tipper, D.: Toward 5G: FiWi enhanced LTE-A hetnets with reliable low-latency fiber backhaul sharing and wifi offloading. IEEE/ACM Trans. Network. **25**(2), 690–707 (2017)

12. Mumtaz, S., Huq, K.M.S., Ashraf, M.I., Rodriguez, J., Monteiro, V., Politis, C.: Cognitive vehicular communication for 5G. IEEE Commun. Mag. **53**(7), 109–117 (2015)
13. Wang, S., Zhou, A., Yang, M., et al.: Service composition in cyber-physical-social systems. IEEE Trans. Emerg. Top. Comput. (2017)
14. Wang, S., Zhou, A., Bao, R., et al.: Towards green service composition approach in the cloud. IEEE Trans. Serv. Comput. (2018)
15. Ferrag, M.A., Maglaras, L., Argyriou, A., Kosmanos, D., Janicke, H.: Security for 4G and 5G cellular networks: a survey of existing authentication and privacy-preserving schemes. J. Netw. Comput. Appl. **101**, 55–82 (2018)
16. Xu, X., et al.: An IoT-oriented data placement method with privacy preservation in cloud environment. J. Netw. Comput. Appl. **124**, 148–157 (2018)
17. Xu, X., et al.: An edge computing-enabled computation offloading method with privacy preservation for internet of connected vehicles. Future Gen. Comput. Syst. **96**, 89–100 (2019)
18. Eiza, M.H., Ni, Q., Shi, Q.: Secure and privacy-aware cloud-assisted video reporting service in 5G-enabled vehicular networks. IEEE Trans. Veh. Technol. **65**(10), 7868–7881 (2016)
19. Ni, J., Lin, X., Shen, X.S.: Efficient and secure service-oriented authentication supporting network slicing for 5G-enabled iot. IEEE J. Sel. Areas Commun. **36**(3), 644–657 (2018)
20. Fang, D., Qian, Y., Hu, R.Q.: Security for 5G mobile wireless networks. IEEE Access **6**, 4850–4874 (2018)
21. Chen, M., Qian, Y., Hao, Y., Li, Y., Song, J.: Data-driven computing and caching in 5G networks: architecture and delay analysis. IEEE Wirel. Commun. **25**(1), 70–75 (2018)
22. Ketykó, I., Kecskés, L., Nemes, C., Farkas, L.: Multi-user computation offloading as multiple knapsack problem for 5G mobile edge computing. In: 2016 European Conference on Networks and Communications (EuCNC), pp. 225–229. IEEE (2016)
23. Zhang, X., Wang, J.: Statistical QoS-driven power adaptation for distributed caching based mobile offloading over 5G wireless networks. In: IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 486–491. IEEE (2018)