

Resource Allocation for Heterogeneous Service in Green Mobile Edge Networks Using Deep Reinforcement Learning

Si-yuan Sun^{1*}, Ying Zheng^{1*}, Jun-hua Zhou², Jiu-xing Weng³, Yi-fei Wei¹, Xiao-jun Wang⁴

¹ School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, 100876, China

[e-mail: sunsy@bupt.edu.cn; weiyifei@bupt.edu.cn; yingzheng@bupt.edu.cn]

² State Key Laboratory of Intelligent Manufacturing System Technology, Beijing Institute of Electronic System Engineering, Beijing, 100854, China

[e-mail: zhoujunhua_zjh@163.com]

³Ningbo Sunny Intelligent Technology Co., LTD, Yuyao, 315400, China

[e-mail: jiuxing@sunnyoptical.com]

⁴Dublin City University, Dublin 9, Ireland

[e-mail: xiaojun.wang@dcu.ie]

*Corresponding author: Sun Si-yuan and Zheng Ying.

*Received February 5, 2021; revised May 27, 2021; accepted June 26, 2021;
published July 31, 2021*

Abstract

The requirements for powerful computing capability, high capacity, low latency and low energy consumption of emerging services, pose severe challenges to the fifth-generation (5G) network. As a promising paradigm, mobile edge networks can provide services in proximity to users by deploying computing components and cache at the edge, which can effectively decrease service delay. However, the coexistence of heterogeneous services and the sharing of limited resources lead to the competition between various services for multiple resources. This paper considers two typical heterogeneous services: computing services and content delivery services, in order to properly configure resources, it is crucial to develop an effective offloading and caching strategies. Considering the high energy consumption of 5G base stations, this paper considers the hybrid energy supply model of traditional power grid and green energy. Therefore, it is necessary to design a reasonable association mechanism which can allocate more service load to base stations rich in green energy to improve the utilization of green energy. This paper formed the joint optimization problem of computing offloading, caching and resource allocation for heterogeneous services with the objective of minimizing the on-grid power consumption under the constraints of limited resources and QoS guarantee. Since the joint optimization problem is a mixed integer nonlinear programming problem that is impossible to solve, this paper uses deep reinforcement learning method to learn the optimal strategy through a lot of training. Extensive simulation experiments show that compared with other schemes, the proposed scheme can allocate resources to heterogeneous service according to the green energy distribution which can effectively reduce the traditional energy consumption.

Keywords: Deep reinforcement learning, green energy, heterogeneous service, mobile edge computing, power consumption, resource allocation

1. Introduction

Nowadays, the advancements in wireless technologies and the popularity of smart devices have spawned various emerging applications and heterogeneous services [1–3]. Those new applications and services put forward higher requirements on data rate and computational capabilities, and at the same time bring explosive mobile traffic, which pose a severe challenge to the construction of the next generation mobile network [4–5]. To meet strict requirements of heterogeneous services and improve user quality of experience (QoE), it is necessary to develop various enabling technologies to improve the utilization of network resources.

Mobile edge computing/cache (MEC) reduces network transmission redundancy and delay by developing computing offload and intelligent content cache at mobile network edge, which further improving network content distribution efficiency and computing processing capacity [6]. As a promising paradigm, Mobile edge computing can enable computation-intensive tasks with stringent low-latency requirement to be processed with quality of service (QoS) guarantee by provide considerable computing resources in proximity to users [7–8]. There are many related works, mainly focusing on the joint optimization of offloading strategy and resource allocation. For multi-unmanned aerial vehicle (UAV) aided MEC system, the authors of [9] propose a reinforcement learning based algorithm to solve the user association and resource allocation problem between UAVs and user equipments (UEs), while minimizing the energy consumption of all UEs. [10] considers the problem of access points (APs) assignment and resource allocation in dynamic scenario, and develops a low-complexity online scheduling algorithm combining stochastic optimization tools and matching theory. Similarly, in order to solve the problem of user association and resource allocation between APs and UEs, [11] combines the successive convex approximation techniques and matching theory to minimize the total transmit power of UEs under delay constraints. [12] decomposes the original problem into task offloading problem and resource allocation problem to simplify the solution, and uses the convex optimization theory and heuristic algorithm to solve the subproblems respectively. [13] integrate non-orthogonal multiple access and introduce quantum behaved particle swarm optimization algorithm to solve APs selection and resource allocation, so as to improve system energy efficiency. [14] uses the deep reinforcement learning framework to solve the computation offloading problem of MEC network, the offloading strategy is optimized to reduce the overhead of mobile users.

On the other hand, deploying cache in the edge network can enhance content delivery networks (CDN) and reducing the data traffic caused by a large number of repeated content requests. By analyzing popularity and proactively caching popular content from the core network to the network edge in advance, the subsequent requests for the same content can be obtained in the edge nodes without duplicate transmissions from remote core network, which can reduce transmission delay and alleviate the pressure of backhaul link and core network [15]. Edge caching has been widely studied due to the advantages of QoE improvement and energy saving. [16] studies the optimization problem of mobile edge caching placement and develops a greedy algorithm to reduce the service load of base stations while maintaining high QoE for users. [17] proposes a new content-centric collaborative edge caching framework, and introduces a vehicle-aided computing and caching scheme to schedule the two resources at the same time, which can effectively reduce access latency and improve resource utilization. Similarly, [18] proposes a learning-based cooperative learning caching (LECC) strategy by incorporating mobile edge computing. Firstly, the content popularity is determined by transfer learning, and then the greedy algorithm is used to solve the problem of cache content placement optimization, so as to improve the content hit rate. The content caching strategy, computation offloading policy and resource allocation are optimized simultaneously in [19], and a solution based on actor-critic reinforcement learning is designed for this joint optimization problem to reduce the service delay.

In addition to intensive computing tasks and concentrated hot contents, 5G network is also concerned about the issue of energy consumption, the advocacy of green 5G makes the integration of green energy into the energy system become a trend. The hybrid energy supply model of traditional power grid combined with green energy and the complementary of various energy sources have become research focus [20-21]. Most of the related work in MEC minimizes system energy consumption by jointly optimizing resource allocation and offloading strategy or cache placement strategy. For example, [22] combines edge computing and caching, and applies reinforcement learning to the dynamic allocation of virtual network resources to reduce the energy consumption. However, this method has limited effect on improving energy efficiency. Significantly improving energy efficiency and accelerating the shift from fossil fuels to renewable energy have become the driving force for 5G's sustainable development. However, less work has been done to integrate green energy into mobile edge computing and caching.

In this paper, we consider hybrid energy supply pattern and consider two typical heterogeneous services in MEC network: content delivery service and computing service similar to work [2] and [23]. By sensing the available green energy status of base stations and distributing more load to base stations rich in green energy, our goal is to fully utilize renewable energy to minimize the on-grid power consumption of system with QoS guarantees. Due to the coexistence of heterogeneous services and the sharing of limited base stations' resources, we must consider the joint optimization of offloading decision, cache placement and resource allocation at the same time, meanwhile guarantee the latency constraint of each service. Since the formulated problem is a mixed integer nonlinear programming (MINLP) which is NP-hard, so we designed a deep reinforcement learning (DRL) based solution to solve the problem. A large number of simulations show the effectiveness of the scheme.

The rest of the paper is organized as follows. Section 2 describes the system model and problem statement. Section 3 presents the proposed scheme in detail. The simulation results and conclusions are discussed and given in Section 4 and Section 5.

2. System Model and Problem Formulation

As illustrated in Fig. 1, in our scenario we consider a heterogeneous MEC network composed of K small base stations (SBSs) and one macro station (MBS), the MBS is connected to the core network via wired links. Each of SBS is equipped with an edge server (ES) and has certain storage capacity to provide services for UEs, the set of SBSs can be denoted by $K = \{1, 2, \dots, K\}$. Since MEC network has communication, computing and caching functions at the same time, the multi-dimensional resources in the network need to be properly configured to expand network capacity, improve network computing capability and content delivery rates. the resources of the SBS side can be expressed as $R_k = \{B_k^{\max}, F_k^{\max}, S_k^{\max}\}$, B_k^{\max} is the available bandwidth including uplink bandwidth and downlink bandwidth of SBS k and F_k^{\max} is the maximum computing capability which is quantified in CPU cycles[24], S_k^{\max} denotes the limited storage capacity of SBS k . We assume UEs associated with the same SBS are allocated orthogonal spectrum using Orthogonal Frequency Division Multiple Access (OFDMA) thus there's no intra-cell interference between UEs in uplink and downlink transmission.

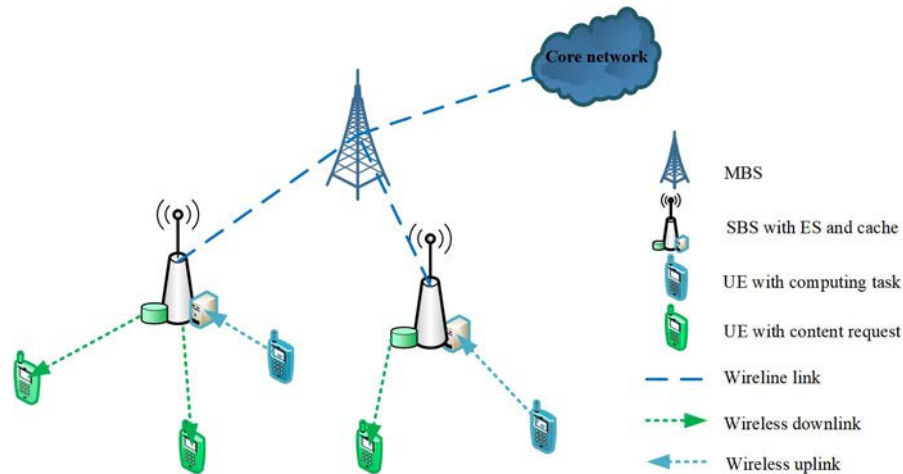


Fig. 1. System model

We assume that there are N UEs in the network and each UE has a service request that needs to be processed. Based on the heterogeneity of services, we divided UEs into two categories, denoted as $N_0 = \{1, 2, \dots, |N_0|\}$ and $N_1 = \{|N_0| + 1, |N_0| + 2, \dots, N\}$ respectively. The entire set of UEs can also be represented as $N = \{N_c | c = 0, 1\}$ in a uniform way, where the subscript c indicates the service category, if $c = 0$ represents computing service otherwise it represents content delivery service. For simplicity, we call them service 0 and service 1 respectively. For service 0, the computing tasks of UE n can be modeled as $task_n^0 = \{d_n^0, \omega_n^0, T_n^{0, \max}\}$, $\forall n \in N_0$, where d_n^0 represents the input data size for computing and ω_n^0 represents the CPU cycles required to complete the computing task, T_n^{\max} is the maximal tolerable delay which require UE's task to be completed within this time limit. Similarly, for service 1 the task of UE n for the content request can be characterized as $task_n^1 = \{d_n^1, T_n^{1, \max}\}$, $n \in N_1$, d_n^1 is the size of the requested content and $T_n^{1, \max}$ still denotes the latency constraint of task.

2.1 Computation Model

For computing service, UEs can choose to execute the task locally or offload to ESs by taking advantage of the ESs' rich computing resource. We use binary variable $\alpha_{n,k} = \{0,1\}, \forall n \in N_0, k \in K$ to represent the offloading decision, $\alpha_{n,k} = 1$ if UE n decide to associate with SBS k for offloading and $\alpha_{n,k} = 0$ otherwise. Assume that each UE can only

associate with at most one SBS, i.e. $\sum_{k=1}^K \alpha_{n,k} \leq 1, \forall n \in N_0$, when $\sum_{k=1}^K \alpha_{n,k} = 0, \forall n \in N_0$ it means

UE choose local execution instead of offloading. If UE process the task locally on the device, the local execution delay is the same as device computation delay which can be given as

$t_n^0 = \frac{\omega_n^0}{F_n^l}$, where F_n^l is the local computing capability. If UE's limited battery capacity or

computing power cannot support local computing, the UE will upload the task to available ESs. The offloading execution delay is composed of transmission delay and computing delay, note that since the computing result is usually very small, we only consider the uplink transmission delay. The transmission rate for uplink can be given by

$$r_{n,k}^u = B_{n,k}^u \log_2 \left(1 + \frac{h_{n,k} P_n^u}{N_0 B_{n,k}^u} \right) \quad (1)$$

where $B_{n,k}^u$ is uplink bandwidth allocated to UE n by SBS k , P_n^u is the transmission power of UE n , $h_{n,k}$ is the channel gain between SBS k and UE n . Denote $f_{n,k}$ as allocated computing resource, the total offloading execution delay of UE n can be calculated by

$t_n^0 = \frac{d_n^0}{r_{n,k}^u} + \frac{\omega_n^0}{f_{n,k}}$, A more concise way is used to represent the task execution delay of service

0 UEs like

$$t_n^0 = \sum_{k=1}^K \alpha_{n,k} \left(\frac{d_n^0}{r_{n,k}^u} + \frac{\omega_n^0}{f_{n,k}} \right) + \left(1 - \sum_{k=1}^K \alpha_{n,k} \right) \frac{\omega_n^0}{F_n^l} \quad (2)$$

Since SBSs allocate computing resources for computing UEs' tasks, which require to consume a certain amount of energy, we express the energy consumption of $task_n^0$ processed by ES k in one second as $\delta f_{n,k}$, namely the computing power consumption, in which δ denotes the energy consumption per CPU cycle of ES. Therefore, the total computing power consumption of the SBS k can be given as

$$P_k^{emp} = \delta \sum_{n \in N_0} \alpha_{n,k} f_{n,k} \quad (3)$$

2.2 Caching Model

The mobile edge cache takes SBSs as the intermediate node to cache the popular content in advance so as to avoid long-distance transmission and realize content reuse. In order to relieve network pressure, pre-caching usually occurs when network is idle. Suppose that there are L types of contents in the network and the content popularity (request probability of contents) follows Zipf distribution [25], thus the probability of l -th type of content being independently requested by each UE is

$$p_l = \frac{1/l^\delta}{\sum_{j=1}^L 1/j^\delta}, \forall l \quad (4)$$

where the component ϵ is usually set to be a positive number between 0.5 and 1 [26]. We still use $\alpha_{n,k}$ to represent the cache decision for service 1, $\alpha_{n,k}=1, n \in N_1$ indicate that the requested content by UE n has been pre-stored in SBS k which can be delivered directly and $\alpha_{n,k}=0$ otherwise. For each type of content, assume that only one SBS can be selected for caching, thus $\sum_{k=1}^K \alpha_{n,k}=0, \forall n \in N_1$ means that the l -th type of content requested by the UE n is not cached on SBSs and can only be retrieved from the remote core server, the requested content must be transferred to SBS via MBS before being delivered to UE. The downlink transmission rate from SBS to UE can be given as

$$r_{n,k}^d = B_{n,k}^d \log_2 \left(1 + \frac{h_{n,k} \rho_k}{N_0} \right) \quad (5)$$

in which $B_{n,k}^d$ is the downlink bandwidth towards UE n and ρ_k denotes transmit power-spectral density which is constant, that is, the transmitted power of each SBS on the unit bandwidth is fixed [27]. Note that we assume the content is transferred directly through the SBS closest to the UE if the requested content is not cached, therefore, for service 1 the content delivery delay can be expressed as

$$t_n^1 = \sum_{k=1}^K \alpha_{n,k} \frac{d_n^1}{r_{n,k}^d} + \left(1 - \sum_{k=1}^K \alpha_{n,k} \right) \left(\frac{d_n^1}{r^b} + \frac{d_n^1}{r_{n,k'}^d} \right) \quad (6)$$

where r^b is the average transmission rate of wireline link and $r_{n,k'}^d$ represents the transmission rate from the nearest SBS k' to the UE n [23]. The actual transmitting power from SBS k to UE n is $\rho_k B_{n,k}^d$, which is proportional to the allocated bandwidth $B_{n,k}^d$, so the total transmitting power of the SBS k is

$$P_k^d = \sum_{n \in \Theta_k} \rho_k B_{n,k}^d \quad (7)$$

where Θ_k includes UEs whose content can be delivered directly or need to be forwarded by SBS k . When contents are not cached, the power consumption of MBS should also be taken into account due to the transmission of data from MBS to SBS. We suppose the power allocated to all links with SBSs are the same, i.e. p^0 and denote κ as the number of UEs who need to get content from the core network, the total power consumption of MBS can be given

$$P_0^{\text{tra}} = \kappa p^0 \quad (8)$$

Usually, the cache placement occurs in the period of low network traffic [28], the power consumption of popular content pre-caching should also be considered and we set it to be proportional to the size of the requested data, that is $\sigma d^l, \forall l \in L$. Then the total power consumption for caching can be given as

$$P_0^{\text{cac}} = \sigma \sum_{l \in L} I_l d^l \quad (9)$$

$I_l = \{1, 0\}$ is the indicator variable indicating whether the l -th type of content is cached. One content copy can be reused many times, thus the cache power-consumption only needs to be calculated once for multiple UEs requesting the same content. Due to the limited storage capacity of SBSs, good caching decision should prioritize caching popular content as much as possible.

2.3 Problem Formulation

In general, the total power consumption of a SBS includes three parts: static power consumption P_k^{sta} , computational power consumption P_k^{cmp} for service 0 and downlink transmission power consumption for service 1.

$$P_k = P_k^{sta} + P_k^{cmp} + P_k^d \quad (10)$$

Similarly, according to the cache decision, the power consumption of MBS can be divided into the power consumption for transmission and cache respectively, namely P_0^{tra} and P_0^{cac} .

In our scenario, SBSs are powered by both power grid and green energy in the same time, while the MBS is powered by power grid only, the green energy is collected by solar panels. Therefore, refer to the green energy acquisition model in [27], the on-grid power consumption of BS can be derived as

$$H_k = \begin{cases} P_0^{tra} + P_0^{cac}, & k = 0 \\ \max(P_k - G_k, 0), & k \in \mathcal{K} \end{cases} \quad (11)$$

H_0 is the on-grid power consumption of MBS and G_k is the green energy generation rate of SBS k . The total on-grid power consumption of the system can be obtained as follows:

$$H_{total} = \sum_{k=0}^K H_k \quad (12)$$

The decision matrix for computing offloading of service 0 and cache placement of service 1 can be denoted as $A = \{\alpha_{n,k}, \forall n \in N, k \in \mathcal{K}\}$. Once A is determined, the association scheme of heterogeneous UEs is also determined. Similarly, the bandwidth including uplink and downlink allocated to UEs can be written as $B^u = \{B_{n,k}^u, \forall n \in N_0, k \in \mathcal{K}\}$ and $B^d = \{B_{n,k}^d, \forall n \in \Theta_k, k \in \mathcal{K}\}$, the allocated computing resources can also be expressed as $F = \{f_{n,k}, \forall n \in N_0, k \in \mathcal{K}\}$. In heterogeneous services situation, the contention between two services for common resources should be considered. Based on that, we formed the following optimization problem to optimize user association so that UEs can choose green energy-rich SBSs as much as possible to provide services. In order to improve the utilization of green energy, the on-grid power consumption of the whole system is minimized under the constraints of limited resources and time delay.

$$P = \min_{A, F, B^u, B^d} H_{total} \quad (13)$$

$$s.t. \text{ C1: } \alpha_{n,k} = \{0, 1\}, \forall n \in N, k \in \mathcal{K}$$

$$\text{C2: } \sum_{k \in \mathcal{K}} \alpha_{n,k} \leq 1, n \in N$$

$$\text{C3: } t_n^c \leq T_n^{c, \max}, \forall c \in \{0, 1\}, n \in N$$

$$\text{C4: } \sum_{n \in N_0} \alpha_{n,k} f_{n,k} \leq F_k^{\max}, \forall k \in \mathcal{K}$$

$$\text{C5: } \sum_{n \in N_0} \alpha_{n,k} B_{n,k}^u \leq B_k^{u, \max}, \forall k \in \mathcal{K}$$

$$\text{C6: } \sum_{n \in \Theta_k} B_{n,k}^d \leq B_k^{d, \max}, \forall k \in \mathcal{K}$$

$$\text{C7: } \sum_{n \in \Theta_k} \rho_k B_{n,k}^d \leq P_k^{\max}, \forall k \in \mathcal{K}$$

$$\text{C8: } \sum_{c=0}^1 \sum_{n \in N_c} \alpha_{n,k} d_n^c \leq S_k^{\max}, \forall k \in \mathcal{K}$$

The constraint C1 restricts the offloading decision and caching decision of two heterogeneous services and C2 indicates that each UEs with service 0/1 can only be served by one SBS. Constraint C3 ensures that the service delay is tolerable. Constraint C4-C6 guarantee that the resources (bandwidth resource and computing resource) allocated to UEs are non-negative and do not exceed the total amount of resources of the SBSs. C7-C8 ensure that the maximum transmitting power limit and storage capability allowance of SBSs are not exceeded. Note that both constraint C6 and C7 are essentially restrictions on the downlink bandwidth, so we can merge the two constraints as follows

$$\text{C9: } \sum_{n \in \Theta_k} B_{n,k}^d \leq \min \left(\frac{P_k^{\max}}{\rho_k}, B_k^{d,\max} \right), \forall k \in \mathcal{K} \quad (14)$$

3. Proposed Algorithm

The formulated problem is complicated due to the coupling of user association and resource allocation. In this part, we designed a DRL-based scheme and solve the above problem in two stages for simplicity. First, we determine the offload decision and cache decision, and then we fine-tune the allocation of resources. Considering the latency constraint of heterogeneous services, it is necessary to allocate at least a certain amount of resources to UEs to meet the QOS assurance. For UEs with service 1, the delay is only related to the downlink bandwidth. Therefore, according to Eq. (2) the downlink bandwidth allocated to UE n for content delivery should satisfy $B_{n,k}^d \geq B_{n,k}^{d,\min}$, the minimum downlink bandwidth required for content transmission can be given as

$$B_{n,k}^{d,\min} = \frac{d_n^1}{\left(T_n^{1,\max} - (1-a_{n,k}) \frac{d_n^1}{r^b} \right) \log_2 \left(1 + \frac{h_{n,k'} \rho_{k'}}{N_0} \right)} \quad (15)$$

which contains two cases: the required content is cached or not be cached. For UEs with service 0, however, the delay is composed of transmission time and computation time which are determined respectively by the given uplink bandwidth and computing resources, thus the allocation of both resources need to be considered simultaneously. Here, we divided the uplink bandwidth of SBSs in proportion to the task size of associated UEs in the first stage like

$$B_{n,k}^u = \frac{d_n^0 B_k^{d,\max}}{\sum_{n' \in N_0} \alpha_{n',k} d_{n'}^0} \quad (16)$$

and then carried out more elaborate resource allocation in the second stage under the condition that part of constraints have been met. Therefore, given the uplink bandwidth allocation, the delay budget for computation can be obtained, so the computing resources should satisfy $f_{n,k} \geq f_{n,k}^{\min}$ at least, $f_{n,k}^{\min}$ can be given as follows

$$f_{n,k}^{\min} = \frac{\omega_n^0}{T_n^{0,\max} - \frac{d_n^0}{r_{n,k}^u}} \quad (17)$$

In the first stage, we make a rough allocation of resources according to Eq. (15) ~ Eq. (17), and then use DRL model to determine user association, the detailed definition of the three elements of DRL can be given as follow, i.e., states, actions and reward functions.

1) States: The state consists of two parts: total on-grid power consumption H_{total} and the proportion of SBSs exceeding the resource budget (τ_B, τ_F, τ_S) , which are the ratios of SBSs that do not meet the constraint C9, C4 and C8 respectively. The reason we don't monitor the uplink bandwidth usage of the SBSs is that we fixed the allocation in the first stage as shown in Eq. (18).

$$s = (H_{total}, \tau_B, \tau_F, \tau_S) \quad (18)$$

2) Actions: For heterogeneous services, if SBS k is selected by UE n to execute computing task or cache the desired file, then take action a_k which corresponds to $\alpha_{n,k} = 1$; Otherwise, if UE n chooses to compute task locally or get the desired content directly from the core network without caching then take action a_0 , which corresponds to $\alpha_{n,k} = 0, \forall k \in \mathcal{K}$. The action set can be expressed as

$$\rho = \{a_0, a_1, \dots, a_K\} \quad (19)$$

3) Reward Function: Since the optimization goal is to reduce system on-grid power consumption, by making reward function negatively correlated with the power consumption, we realized the transformation of the original problem. Moreover, for the action collection which exceeds SBSs' resource budget, σ is set to a larger value to give a certain penalty, which ensure that the model can continuously reduce the on-grid power consumption while satisfy the resource limit as much as possible in the optimization process. Here, we severely penalize the overallocation of downlink bandwidth and storage by setting a large value of σ while setting a small value of δ imposes a smaller penalty for overallocation of computing resources, because in the second stage, we will reallocate computing resources and uplink bandwidth.

$$r = \frac{1}{H_{total} + \sigma(\tau_B + \tau_S) + \delta\tau_F} \quad (20)$$

Table 1. Deep reinforcement learning based association algorithm

Initialize: replay memory \mathcal{D} to capacity c_1 , action-value function Q with random weights θ , target action-value function \hat{Q} with random weights $\theta = \theta$, UEs' state s .

while $\xi \leq \xi^{\max}$

for $n = 1 : N$

if $\text{rand}(0,1) < \varepsilon$

 Select an action a randomly from ρ for the n -th UE

else

 Select an action $a = \underset{a \in \rho}{\operatorname{argmax}} Q(s, a; \theta)$ for the n -th UE

end if

 Allocate bandwidth and computing resources in Eq. (15) ~ Eq. (17) according to UEs' action

 Obtain reward r (Eq. (20)), get new state s' and store the new sample (s, a, r, s') in \mathcal{D}

 update state $s = s'$

 Sample random minibatch of samples (s_j, a_j, r_j, s_{j+1}) from \mathcal{D}

 Set $y_j = r_j + \gamma \max_{a'} \hat{Q}(s_{j+1}, a'; \theta^-)$

Perform a gradient descent step on $\left(y_j - Q(s_j, a_j; \theta)\right)^2$ with respect to the network parameters θ

Every E steps reset $\hat{Q} = Q$

end for

$\xi = \xi + 1$

end while

The decision matrix A can be determined from the action collection of all UEs

It should be noted if $\frac{\omega_n^0}{F_n^l} \leq T_n^{0, \max}$ UEs with service 0 can only choose to offload its task.

The detailed implementation of the DRL-based algorithm is illustrated in Tab. 1, we initialize the model parameters at first. Then, at each training episode, UEs are required to select an action according to the ε -greedy policy [29]. Once the collection of all UEs' action is obtained, we allocate bandwidth and computing resources according to Eq. (15) ~ Eq. (17). After getting rewards r and new states s' , we store the new sample (s, a, r, s') and take a random sample from \mathbf{D} to update the parameters θ of Q network. At the same time, the target Q network $\bar{\theta} = \theta$ is updated every E steps. Repeat the above process until the maximum episode we set is reached. With the increase of training episodes the algorithm gradually converges, and finally we get the action collection which can reduce the on-grid power consumption of system.

In the second stage, we divide resources more finely to improve service performance for UEs with service 0. For service 1, its power consumption mainly comes from content transmission and pre-caching of popular content, moreover the downlink transmission power from SBS to UE is proportional to the allocated downlink bandwidth $B_{n,k}^d$ as shown in Eq. (7). In the first stage, we set the downlink bandwidth $B_{n,k}^d$ to the minimum value in Eq. (15) that can satisfy UE's delay constraint, and strictly punish the excessive allocation of resources through the reward function, so as to restrict the total downlink bandwidth allocated by SBSs within the budget. Therefore, once the first stage is over, which indicates that under the current caching decision and offloading decision, the allocation of downlink bandwidth can meet the delay requirement while minimize the downlink transmission power, so we still retain the downlink bandwidth allocation scheme of the first stage and omit constraints C6 and C7. In addition, once caching decision and offloading decision are determined, UEs with different service served by which SBS can also be determined. Therefore, the three part of power consumption of service 1 as shown in Eq. (7) ~ Eq. (9) is fixed in the second stage, according to Eq. (10) ~ Eq. (11), the expression of the on-grid

power consumption for BS can be changed to $H_k = \begin{cases} C_0, k = 0 \\ \max(P_k^{cmp} - C_k, 0), k \in \mathcal{K} \end{cases}$, where

$C_k, k \in \{0, \mathcal{K}\}$ is constant. In the same way, the reward function also penalizes the decision that exceeds the storage capacity of SBSs, so it can ensure that the final user association scheme meets constraint C8. In conclusion, the original problem can be degenerate to

$$P_1 = \min_{F, B^d} \sum_{k \in \mathcal{K}} P_k^{cmp} \quad (21)$$

$$s.t. \text{ C1: } t_n^0 \leq T_n^{0, \max}, n \in N$$

$$\text{C2: } \sum_{n \in N_0} \alpha_{n,k} f_{n,k} \leq F_k^{\max}, \forall k \in \mathcal{K}$$

$$\text{C3: } \sum_{n \in N_0} \alpha_{n,k} B_{n,k}^u \leq B_k^{u,\max}, \forall k \in \mathcal{K}$$

According to Eq. (3), the computing power consumption of SBSs depends only on the computing resources allocated to associated UEs. The problem \mathcal{P}_1 can be easily solved by Lagrange multiplier method.

4. Simulation Results

In this section, we compare the proposed scheme with other schemes to evaluate the performance of the proposed scheme from different perspectives. In the simulation, the channel pathloss model $L[\text{dB}] = 140.7 + 36.7 \log_{10} d[\text{km}]$ is considered with reference to [30]. We consider a single cell covered with $400\text{m} \times 400\text{m}$, the MBS is located in the center of the cell, SBSs and UEs are randomly distributed in this space. We assume that the proportion of UEs with heterogeneous service is same, i.e., $|N_0| = |N_1| = \frac{N}{2}$. For service 0, the size of the input data and the number of CPU cycles required to perform the computing task are uniformly distributed between $[100, 1000]\text{KB}$ and $[10^8, 10^9]$ cycles respectively, and the size of the requested content for service 1 also follow the uniform distribution within the range $[4, 8]\text{MB}$. As for the maximum allowable delay of UE's tasks, we set both heterogeneous services in the range of $[1, 2]\text{s}$. The uplink transmission power P_n^u and the computing capacity of local CPU F_n^l are set to 100mW and 0.5GHz , other network parameters are shown in Tab. 2. For DRL model parameters, we set ϵ -greedy policy probability and reward decay to 0.9, learning rate to 0.01 and the maximum episode $\xi^{\max} = 10000$, as for the penalty factors σ and δ in Eq. (20), we set them as 1000 and 100 respectively. The simulation results are averaged by multiple experiments.

Table 2. Simulation parameters setting

Parameters	Settings
The maximum CPU computing capability of SBS F_k^{\max}	[6,8]GHz
Total available bandwidth of SBS B_k^{\max}	10MHz
Total storage capacity of SBS S_k^{\max}	[1,2]GB
Maximum transmitting power of SBS P_k^{\max}	35dBm
Energy consumption per CPU cycle of SBS δ	1 W/GHz
Power consumption per Mb for pre-caching σ	0.05W/Mb
Transmit power-spectral density of SBS ρ_k	0.2uW/HZ
The static power consumption of SBS P_k^{sta}	33W
The average wireline transmission rate r^b	1Gbps
Allocated power to links with SBSs p^0	1W
Green energy generation rate of SBS G_k	[35,40]W

As mentioned above, the proposed scheme uses DRL-based algorithm to determine the offloading decision and caching decision, and allocate resources at the same time. As a comparison, we use other strategies to obtain the offloading decision and caching decision respectively. For the offloading decision we adopt SNR-based offloading strategy, if the local computation fails to meet the delay requirements, the nearby SBSs are selected for offloading according to the channel quality status, i.e. signal-noise ratio (SNR). As for caching decision, we adopt random caching strategy which means SBSs are randomly selected to cache the requested content without consideration of the content popularity, note that the storage capacity of the SBSs should not be exceeded. The proposed scheme and comparison schemes are elaborated as follows:

- **DRL-based offloading strategy and caching strategy (DODC):** DODC scheme uses the proposed DRL-based algorithm described in tab. 1 to determine the offloading decision and caching decision, and then allocate resources by solving P_1 as shown in Eq. (21).
- **SNR-based offloading strategy and DRL-based caching strategy (SODC):** SODC scheme uses the above-mentioned SNR-based offloading strategy to determine the offloading decision instead of the DRL-based offloading strategy in DODC scheme, while the caching strategy and the resource allocation is the same as the DODC scheme.
- **DRL-based offloading strategy and random caching strategy (DORC):** DORC scheme uses the above-mentioned random caching strategy to determine the caching decision, the offloading strategy and resource allocation is the same as the DODC scheme.
- **SNR-based offloading strategy and random caching strategy (SORC):** SORC scheme uses SNR-based offloading strategy and random caching strategy to determine the offloading decision and caching decision respectively.

Fig. 2 and **Fig. 3** show the influence of the number of UEs and SBSs to the system on-grid power consumption for four schemes. In **Fig. 2** the number of SBSs K is set to 4, it can be seen that with the increase of the number of UEs N , the on-grid power consumption of four schemes basically presents an increasing trend, but the proposed scheme DODC is obviously superior to the other three comparison schemes. In addition, the figure shows that when the number of UEs is small, the on-grid power consumption of DODC is 0, and then presents a slow growth trend. This is because the proposed scheme can distribute the traffic load according to the green energy status of SBSs, which make UE's tasks consume as much green energy as possible thus the on-grid power consumption can be reduced. As a result, at the beginning, SBSs' green energy is enough to support the completion of UE's service requests, so the on-grid power consumption is 0. Then, after the green energy is exhausted, the proposed scheme can also adjust the association strategy to make the on-grid power consumption as small as possible.

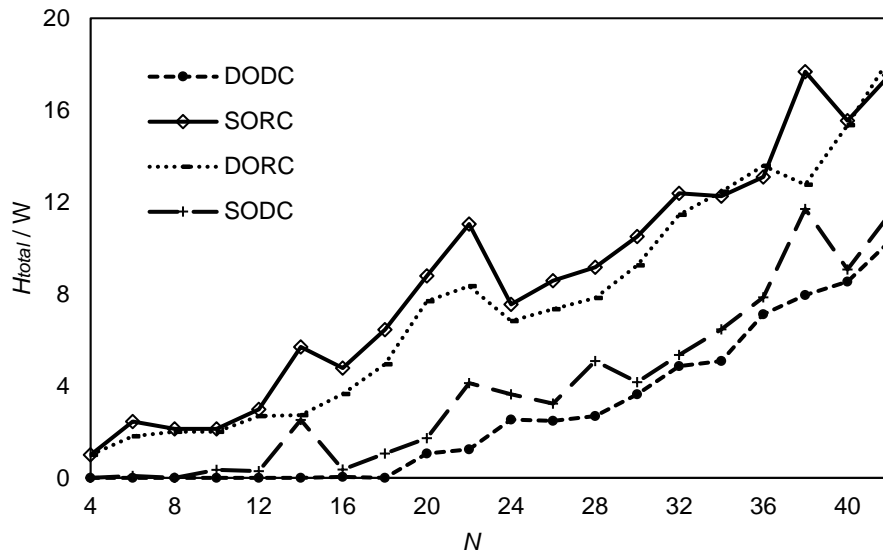


Fig. 2. System on-grid power computation of four schemes versus the number of UEs

Similarly, in **Fig. 3** we fixed the number of UEs N at 40 and gradually increased the number of SBSs K starting from 1. As expected, the system on-grid power consumption of DODC is gradually reduced to 0 with the increase of SBSs, this is because the number of UEs is fixed and more SBSs bring more green energy, the proposed DODC scheme can effectively distribute load to make full use of the green energy in the network to reduce the on-grid power consumption. Therefore, it can be seen that the on-grid power consumption of the system is actually negatively correlated with the number of SBSs, the effectiveness of DODC in energy saving can also be demonstrated by **Fig. 3**.

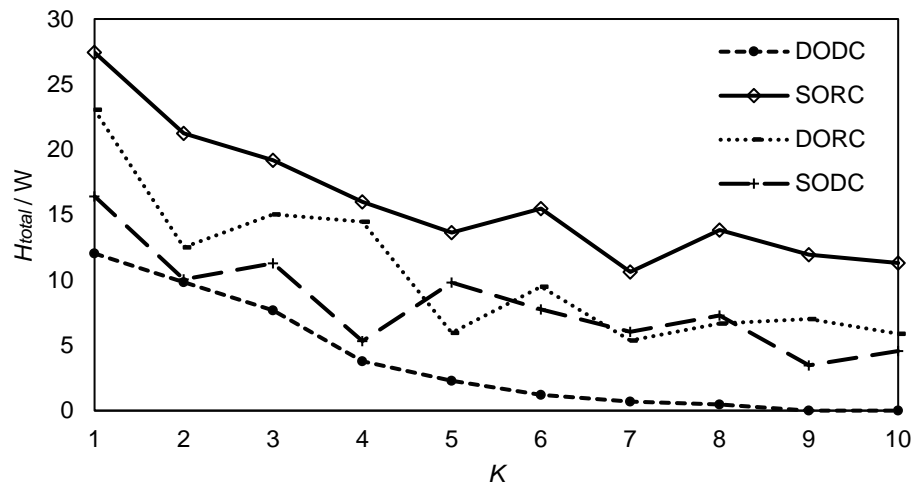


Fig. 3. System on-grid power computation of four schemes versus the number of SBSs

We define the load balancing index $\zeta = \frac{(\sum_{k \in K} H_k)^2}{K \sum_{k \in K} (H_k)^2}$ to measure the load distribution

of different SBSs. Fig. 4 illustrates the load balancing performance of different schemes with different number of UEs. According to the results, it can be seen that the load balancing performance of the proposed scheme DODC is better than the comparison scheme, which can effectively distribute the load to each SBS to reduce the total on-grid power consumption of the system.

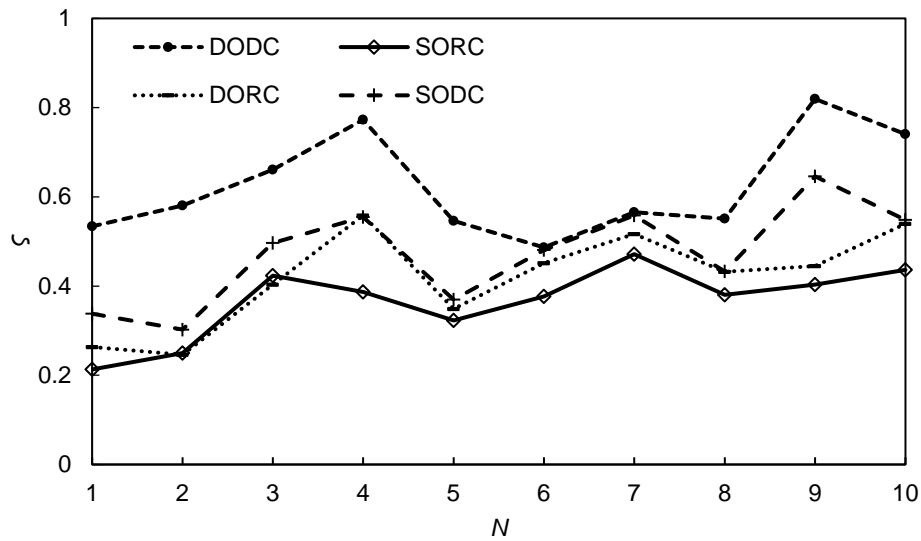


Fig. 4. Load balancing index of four schemes versus the number of UEs

5. Conclusions

In this article, we consider two typical heterogeneous service and hybrid energy supply pattern, and study the user association and resource allocation problem of a two-tier heterogeneous cellular network. Both computing offloading and content caching need to consume the power of BS, user association needs to perceive the status of green energy and make the green energy-rich SBSs bear more traffic load as far as possible to reduce traditional energy consumption. Therefore, the joint optimization problem of offloading decision, caching decision, computing resource and bandwidth allocation is formulated to minimize the on-grid power consumption of system with QOS guarantees. Since the above problem is NP-hard, we propose a DRL-based scheme to solve it in two stages. The first stage is to determine the optimal association assignment based on the offloading decision and caching decision, and then in second stage we make a more elaborate allocation of resources. The simulation results show that DODC scheme has good performance in reducing on-grid power consumption, and has good load distribution ability which can effectively balance the load.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (61871058).

References

- [1] Y. Lan, X. Wang, D. Wang, Y. Zhang and W. Wang, "Mobile-Edge Computation Offloading and Resource Allocation in Heterogeneous Wireless Networks," in *Proc. of 2019 IEEE Wireless Communications and Networking Conference (WCNC)*, Marrakesh, Morocco, pp. 1-6, 2019. [Article \(CrossRef Link\)](#)
- [2] Z. Tan, F. R. Yu, X. Li, H. Ji and V. C. M. Leung, "Virtual resource allocation for heterogeneous services in full duplex-enabled small cell networks with cache and MEC," in *Proc. of 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Atlanta, GA, pp. 163-168, 2017. [Article \(CrossRef Link\)](#)
- [3] L. Sun, Q. Yu, D. Peng, S. Subramani and X. Wang, "Fogmed: a fog-based framework for disease prognosis based medical sensor data streams," *Computers, Materials & Continua*, vol. 66, no.1, pp.603–619, 2021. [Article \(CrossRef Link\)](#)
- [4] Z. G. Qu, S. Y. Chen and X. J. Wang, "A Secure Controlled Quantum Image Steganography Algorithm," *Quantum Information Processing*, vol.19, no. 380, pp. 1-25, 2020. [Article \(CrossRef Link\)](#)
- [5] Z. G. Qu, S. Y. Wu, W. J. Liu and X. J. Wang, "Analysis and improvement of steganography protocol based on bell states in noise environment," *Computers, Materials & Continua*, vol. 59, no.2, pp.607–624, 2019. [Article \(CrossRef Link\)](#)
- [6] W. Jiang, G. Feng, S. Qin and Y. Liu, "Multi-Agent Reinforcement Learning Based Cooperative Content Caching for Mobile Edge Networks," *IEEE Access*, vol. 7, pp. 61856-61867, 2019. [Article \(CrossRef Link\)](#)
- [7] Y. He, N. Zhao and H. Yin, "Integrated Networking, Caching, and Computing for Connected Vehicles: A Deep Reinforcement Learning Approach," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 1, pp. 44-55, 2018. [Article \(CrossRef Link\)](#)
- [8] Y. Dai, D. Xu, S. Maharjan and Y. Zhang, "Joint Computation Offloading and User Association in Multi-Task Mobile Edge Computing," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 12313-12325, 2018.
- [9] L. Wang, P. Huang, K. Wang, G. Zhang, L. Zhang, N. Aslam and K. Yang, "RL-Based User Association and Resource Allocation for Multi-UAV enabled MEC," in *Proc. of 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, Tangier, Morocco, pp. 741-746, 2019. [Article \(CrossRef Link\)](#)
- [10] M. Merluzzi, P. D. Lorenzo and S. Barbarossa, "Dynamic Joint Resource Allocation and User Assignment in Multi-access Edge Computing," in *Proc. of ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, United Kingdom, pp. 4759-4763, 2019. [Article \(CrossRef Link\)](#)
- [11] S. Sardellitti, M. Merluzzi and S. Barbarossa, "Optimal Association of Mobile Users to Multi-Access Edge Computing Resources," in *Proc. of 2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, Kansas City, MO, pp. 1-6, 2018. [Article \(CrossRef Link\)](#)
- [12] T. X. Tran and D. Pompili, "Joint Task Offloading and Resource Allocation for Multi-Server Mobile-Edge Computing Networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 856-868, 2019. [Article \(CrossRef Link\)](#)
- [13] S. Seng, X. Li, H. Ji and H. Zhang, "Joint Access Selection and Heterogeneous Resources Allocation in UDNs with MEC Based on Non-Orthogonal Multiple Access," in *Proc. of 2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, Kansas City, MO, pp. 1-6, 2018. [Article \(CrossRef Link\)](#)
- [14] Y. Wei, Z. Wang, D. Guo and F. R. Yu, "Deep q-learning based computation offloading strategy for mobile edge computing," *Computers, Materials & Continua*, vol. 59, no. 1, pp. 89–104, 2019. [Article \(CrossRef Link\)](#)

- [15] E. Baccour, A. Erbad, A. Mohamed, K. Bilal and M. Guizani, "Proactive Video Chunks Caching and Processing for Latency and Cost Minimization in Edge Networks," in *Proc. of 2019 IEEE Wireless Communications and Networking Conference (WCNC)*, Marrakesh, Morocco, pp. 1-7, 2019. [Article \(CrossRef Link\)](#)
- [16] C. Li, L. Toni, J. Zou, H. Xiong and P. Frossard, "QoE-Driven Mobile Edge Caching Placement for Adaptive Video Streaming," *IEEE Transactions on Multimedia*, vol. 20, no. 4, pp. 965-984, 2018. [Article \(CrossRef Link\)](#)
- [17] K. Zhang, S. Leng, Y. He, S. Maharjan and Y. Zhang, "Cooperative Content Caching in 5G Networks with Mobile Edge Computing," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 80-87, 2018. [Article \(CrossRef Link\)](#)
- [18] T. Hou, G. Feng, S. Qin and W. Jiang, "Proactive Content Caching by Exploiting Transfer Learning for Mobile Edge Computing," in *Proc. of GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, Singapore, pp. 1-6, 2017. [Article \(CrossRef Link\)](#)
- [19] Y. Wei, F. R. Yu, M. Song and Z. Han, "Joint Optimization of Caching, Computing, and Radio Resources for Fog-Enabled IoT Using Natural Actor-Critic Deep Reinforcement Learning," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2061-2073, 2019.
- [20] B. Wang, Q. Kong, W. Liu and L. T. Yang, "On Efficient Utilization of Green Energy in Heterogeneous Cellular Networks," *IEEE Systems Journal*, vol. 11, no. 2, pp. 846-857, 2017. [Article \(CrossRef Link\)](#)
- [21] Y. Wei, F. R. Yu, M. Song and Z. Han, "User Scheduling and Resource Allocation in HetNets With Hybrid Energy Supply: An Actor-Critic Reinforcement Learning Approach," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 680-692, 2018. [Article \(CrossRef Link\)](#)
- [22] L. Li, Y. Wei, L. Zhang and X. Wang, "Efficient virtual resource allocation in mobile edge networks based on machine learning," *Journal of Cyber Security*, vol. 2, no. 3, pp. 141-150, 2020. [Article \(CrossRef Link\)](#)
- [23] J. Zhou, X. Zhang and W. Wang, "Joint Resource Allocation and User Association for Heterogeneous Services in Multi-Access Edge Computing Networks," *IEEE Access*, vol. 7, pp. 12272-12282, 2019. [Article \(CrossRef Link\)](#)
- [24] S. Bi and Y. J. Zhang, "Computation Rate Maximization for Wireless Powered Mobile-Edge Computing With Binary Computation Offloading," *IEEE Transactions on Wireless Communications*, vol.17, no.6, pp.4177-4190, 2018. [Article \(CrossRef Link\)](#)
- [25] Y. Zhou, F. R. Yu, J. Chen and Y. Kuo, "Resource Allocation for Information-Centric Virtualized Heterogeneous Networks With In-Network Caching and Mobile Edge Computing," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 11339-11351, 2017. [Article \(CrossRef Link\)](#)
- [26] Y. Jin, Y. Wen and C. Westphal, "Optimal Transcoding and Caching for Adaptive Streaming in Media Cloud: an Analytical Approach," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 12, pp. 1914-1925, 2015. [Article \(CrossRef Link\)](#)
- [27] Q. Fan and N. Ansari, "Green energy aware user association in heterogeneous networks," in *Proc. of 2016 IEEE Wireless Communications and Networking Conference*, Doha, pp. 1-6, 2016. [Article \(CrossRef Link\)](#)
- [28] M. M. Amiri and D. Gündüz, "Cache-aided data delivery over erasure broadcast channels," in *Proc. of 2017 IEEE International Conference on Communications (ICC)*, Paris, pp. 1-6, 2017. [Article \(CrossRef Link\)](#)
- [29] V. Mnih, K. Kavukcuoglu and D. Silver et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529-533, 2015. [Article \(CrossRef Link\)](#)
- [30] L. Wang, K. Wong, S. Jin, G. Zheng and R. W. Heath, "A New Look at Physical Layer Security, Caching, and Wireless Energy Harvesting for Heterogeneous Ultra-Dense Networks," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 49-55, 2018. [Article \(CrossRef Link\)](#)



Siyuan Sun received bachelor's degree in communication engineering from Beijing University of Posts and Telecommunications in 2011 and received master's degree in electronic engineering from Dublin City University in Ireland in 2014. He is currently studying for a doctor's degree in electronic Information Engineering from Beijing University of Posts and telecommunications and works as a solution manager of the Industry Internet Operation Center of China Unicom Beijing Branch. He was involved in several enterprises 5G innovation projects in his work, and his research interests are green and energy-save 5G network, and 5G mobile edge network.



Ying Zheng received the B.Eng. degree in Electronic Information Science and Technology from Shandong University, China, in 2018 and the M.Eng. degree in Electronic and Communication Engineering from Beijing University of Posts and Telecommunications (BUPT), China, in 2021. Her current research interests are in intelligent resource allocation strategy, energy-efficient networking, deep reinforcement learning, multi-access edge computing and caching.



JunHua Zhou received master degree in Graduate School of the 2nd Academy of CASIC, Beijing, China, in 2009. He is a senior engineer of Beijing Institute of Electronic System Engineering and Simulation Center. His research interests include multidisciplinary virtual prototype modeling and simulation.



Jiuxing Weng graduated from the University of Hong Kong with a major in information strategy and enterprise transformation. In 2007, he served as the director of the information center and director of the information technology department of Shunyu optical technology Co., Ltd. He also holds many social positions: expert on the direction of the intelligent factory of the national key laboratory for intelligent manufacturing of complex products; director of the professional expert committee of Ningbo intelligent manufacturing association; expert committee of the Ningbo bureau of economics and information technology on intelligent.



Yifei Wei (M'12) received the B.Sc. and Ph.D. degrees in electronic engineering from Beijing University of Posts and Telecommunications (BUPT, China), in 2004 and 2009, respectively. He was a visiting scholar in the University of Houston (USA) from 2016 to 2017. He is currently a professor in school of electronic engineering at BUPT. His current research interests are in energy-efficient networking, software defined heterogeneous networks, machine learning, edge computing and caching.



Xiaojun Wang received the B.Eng. degree in Computer and Communications and the M.Eng. degree in Computer Applications from Beijing University of Posts and Telecommunications (BUPT), China, in 1984 and 1987 respectively. He received the Ph.D. degree in Electronic Engineering from Staffordshire University (then Staffordshire Polytechnic), England, U.K., in 1993. He joined the School of Electronic Engineering, Dublin City University, Ireland, as an Assistant Lecturer in 1992, where he is currently a Senior Lecturer. His current research interests include energy-efficient networking, network security, and hardware acceleration of cryptographic algorithms.