Joint Offloading and Resource Allocation Optimization for Time-Sensitive Mobile-Edge Computing Network

Jun-Jie Yu, Wen-Tao Li, Mingxiong Zhao

***Abstract*—In this paper, we investigate offloading scheme and resource allocation strategy for orthogonal frequency division multiple access (OFDMA) based mobile-edge computing (MEC) network to minimize the total system energy consumption. Partial data offloading is studied where mobile date can be computed at both local devices and the edge cloud with the consideration of time-sensitive tasks for users. Since the NP-hardness of the consideration optimization problem, we propose an iterative algorithm to design the resource allocation strategy and decide the proportion of data to offload in a sequence. Simulation results show that the proposed algorithm achieves the excellent performance.**

***Index Terms*—Mobile-edge computing (MEC), resource allocation, orthogonal frequency-division multiple access (OFDMA), time-sensitive.**

I. Introduction

Accompanying with the development of technology, a considerable number of novel applications, such as virtual reality(VR), online interactive games, smart city services, autonomous vehicle and industrial internet of things(IIoT), put forward stringent requirements of the communication performance, reliability, data rate capacity and other aspects of metrics [1]. For example, the VR applications and devices generally demand enormous computation workload while the expected delay should be lower than 1ms z[2]. Mobile Edge Computing(MEC), as a key technology of 5G network defined by the European telecommunications standardization institute (ETSI), makes it possible for these applications by providing IT infrastructures in close proximity to terminal user equipment(UE) within the radio access network (RAN) [3]. Comparing with offloading to traditional mobile cloud computing(MCC), applications on UE could save the dominated propagation delay and energy consumption by forwarding task requests to MEC servers for computation, thus avoiding core network(CN) congestion and long propagation delay.

MEC Offloading is the process by which tasks on UE can choose to be executed locally or sent to MEC server for computation, which is particularly helpful for energy constrained devices. A number of research have be addressed to energy optimization by offloading tasks to edge server. You *et al.* [4] formulate the offloading decision and resource allocation as a optimization problem and minimize system cost in terms of energy and time by a sub-optimal centralized solution on MEC server. In [5], [6], MEC system integrated with energy harvesting (EH) equipments have be discussed. Mao *et al.* [5] develope a Lyapunov optimization-based LODCO algorithm to jointly decide offloading option and CPU frequencies for a MEC system integrated with energy harvesting (EH) technologies. Energy optimization problems of Single user scenarios have be considered In [7], [8]. Dinh *et al.* [7] optimize single mobile device's energy consumption by joint considering task allocation decision and mobile device's CPU frequency. Wang *et al.* [9] propose an framework for offloading decision and resource allocation problem which is solved by graph coloring method.

On the other hand, problems of energy collaborative optimization in multi-user system have attracted extensive attention. Zhang *et al.* jointly consider energy consumption, delay and monetary cost in the formulated problem from terminal perspective and find the Nash equilibrium by a game theoretic algorithm In [10]. A solution of tradeoff between energy and latency combined iterative search algorithm with interior penalty function has be proposed by jointly optimizing communication and computation resource allocation in [11]. Li *et al.* [12] investigate the subcarrier and power allocation problems in an OFDMA based MEC system and propose a lower-bound algorithm to find optimal delay ensuring user devices' fairness. Recently, Cui *et al.* formulate a multiobjective constrained optimization problem considering tradeoff between energy consumption and delay cost and propose a modified NSGA-II algorithm which can always frequently find Pareto optimal set in [13].

In this paper, we propose a joint offloading and subcarrier allocation strategy. This strategy jointly optimizes the offload strategy, subcarrier and computing resource allocation scheme to reduce the total energy consumption of the entire MEC system. However, due to the NP-hard features and non-convexity of the joint problem. Since the optimal solution is difficult to obtain, we propose a bi-level iterative algorithm to find the optimal Offloading and resource allocation strategy. More detailed instructions, we propose an efficient iterative algorithm. The computing resource allocation strategy, the subcarrier allocation strategy, and the offload ratio are found in an alternating manner, and then the optimal solution is found through iteration. Numerical results show that the proposed algorithm is superior to traditional algorithms.

The rest of this paper is organized as follows. In section II, we described the system model and formulated the original problem. In section III, we introduce our proposed alogrithms. Simulation results are showed in section IV. Finally, we conclude the article in section V

II. System model and problem formulation

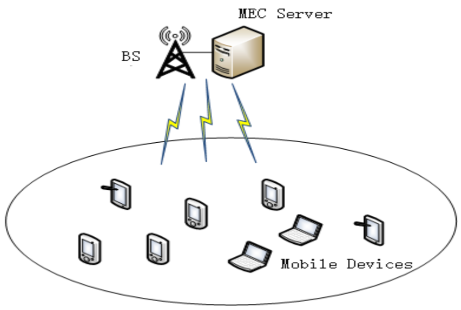

Fig. 1. The MEC system model of multiuser offloading with OFDMA.

We consider an OFDMA-based MEC system with K users and one base-station (BS) integrated with an MEC server to execute the offloaded data of users, and all nodes are equipped with a single antenna. Denote $\mathcal{K} \triangleq \{1, 2, \cdots, K\}$ as the set of users, and let $\mathcal{N} \triangleq \{1, 2, \cdots, N\}$ be the index for multiple orthogonal subcarriers, each of which has bandwidth B and can be assigned to only one user. In this system, we assume that user k has a task described by a tuple of four parameters $\{R_k, c_k, \lambda_k, t_k\}$, where R_k indicates the amount of input data to be processed, c_k represents the number of CPU cycles for computing 1-bit of input data, $\lambda_k \in [0, 1]$ is the proportion of R_k offloading to MEC, while the rest $(1 - \lambda_k)R_k$ bits are processed by its local CPU, and t_k is the maximum tolerable latency. In this paper, it is assumed that the maximum tolerable latency for user k, t_k is no longer than the channel coherence time, such that the wireless channels remain constant during a time slot with length T, i.e., $t_k \le T, \forall k$, but can vary from time to time. The local CPU frequency of user k is characterized by f_k, and $f_{k,m}$ is the computational speed of the edge cloud assigned to user k, where both of them are measured by the number of CPU cycles per second. Herein, a practical constraint that the total computing resources allocated to all the associated users must not excess the server's computing capacity is given by $\sum_{k \in \mathcal{K}} f_{k,m} \le F$.

In the following, the time latency and the energy consumption of user k for our considered system are given in details.

A. Time Latency

1) Local Computing at Users: Consider the local computing for executing the residual $(1 - \lambda_k)R_k$ input bits at user k, the time consumption for local computing at user k is

$$t_{k,l} = \frac{c_k\left(1 - \lambda_k\right)R_k}{f_k}. \tag{1}$$

2) Computation Offloading: According to the OFDMA mechanism, the inter-interference is ignored in virtue of the exclusive subcarrier allocation. Therefore, the aggregated transmission rate to offload $\lambda_k R_k$ input bits from user k to MEC server is expressed as

$$r_k = B \sum_{n \in \mathcal{N}} x_{k,n} \log_2\left(1 + \frac{p_k g_{k,n}}{\sigma_n^2}\right), \tag{2}$$

where $g_{k,n}$ and σ_n^2 are the channel gain between user k and BS, and the variance of the additive white Gaussian noise at BS on subcarrier n, respectively, where we set $\sigma_n^2 = \sigma^2, \forall n$. Denote p_k as the transmission power, which can be allocated by the user and constrained by the maximum transmission power $p_k^{\max}$. Apparently, any power optimization solution has a good influence on the system performance. For the sake of simplicity, p_k remains at a random level in this paper [14]. Meanwhile, denote $x_{k,n}$ as the channel allocation indicator, specifically $x_{k,n} = 1$ means that subcarrier n is assigned to user k, otherwise $x_{k,n} = 0$.

The offloading time $t_{k,\text{off}}$ of user k mainly consists of two parts[1]: the uplink transmission time $t_{k,u}$ from user k to MEC server and the corresponding execution time at MEC server $t_{k,m}$. Therefore, the offloading time $t_{k,\text{off}}$ is given by

$$t_{k,\text{off}} = t_{k,u} + t_{k,m} = \frac{\lambda_k R_k}{r_k} + \frac{\lambda_k R_k c_k}{f_{k,m}}. \tag{3}$$

Due to the parallel computing at users and MEC server, the total time latency for user k depends on the the larger one between $t_{k,l}$ and $t_{k,\text{off}}$, and can be expressed as

$$T = \max\{t_{k,l}, t_{k,\text{off}}\}. \tag{4}$$

B. Energy Consumption

According to the strategy of computation offloading at user k, the total energy consumption comprises two parts[1]: the energy for local computing and for offloading, given in details as follow.

1) Local Computing Mode: Given the processor's computing speed f_k, the power consumption of the processor is modeled as $\kappa_k f_k^3$ (joule per second), where κ_k represents the computation energy efficiency coefficient related to the processor's chip of user k [16]–[18]. Taking consideration of (1), the energy consumption at this mode is given by

$$E_{k,l} = \kappa_k f_k^3 t_{k,l} = \kappa_k c_k \left(1 - \lambda_k\right) R_k f_k^2. \tag{5}$$

2) Computation offloading mode: In this mode, the energy consumption includes the cost of uplink transmitting and remote computing for offloaded $\lambda_k R_k$ input bits, which can be obtained as

$$\begin{aligned} E_{k,\text{off}} &= E_{k,u} + E_{k,m} \\ &= p_k \frac{\lambda_k R_k}{r_k} + \kappa_m \lambda_k c_k R_k f_{k,m}^2, \end{aligned} \tag{6}$$

where κ_m is the computation energy efficiency coefficient related to the processor's chip of MEC server.

Therefore, the total energy consumption for user k related with its computation offloading strategy in our system is

$$E_k = E_{k,l} + E_{k,u} + E_{k,m}. \tag{7}$$

[1] In practice, the MEC-integrated BS will provide sufficient transmit power, while the amount of output data from MEC server to user k is usually much less than that of the input data, the time consumed and the transmission energy for delivering the computed results are negligible [15].

In this paper, we minimize the overall energy consumption of the considered system, which is related to resource allocation on subcarriers, offloading communication and computation. Mathematically, the energy consumption minimization problem can be written as

$$\begin{aligned}
\mathbf{P1}: \ \min_{\boldsymbol{\lambda},\boldsymbol{f},\boldsymbol{X}} \ & \sum_{k\in\mathcal{K}} E_k && \text{(8a)}\\
\text{s.t.} \ & 0 \le \lambda_k \le 1, \forall k, && \text{(8b)}\\
& \max\{t_{k,l}, t_{k,\text{off}}\} \le T, \forall k, && \text{(8c)}\\
& 0 \le f_{k,m} \le f_{k,m}^{max}, \forall k, && \text{(8d)}\\
& \sum_{k\in\mathcal{K}} f_{k,m} \le F, && \text{(8e)}\\
& \sum_{k\in\mathcal{K}} x_{k,n} \le 1, \forall n, && \text{(8f)}\\
& x_{k,n} \in \{0,1\}, \forall k,n, && \text{(8g)}
\end{aligned}$$

where $\boldsymbol{\lambda} \triangleq \{\lambda_k\}$, $\boldsymbol{f} \triangleq \{f_{k,m}\}$ and $\boldsymbol{X} \triangleq \{x_{k,n}\}$. The constraints above can be explained as follows: constraint (8c) states that the task of user k must be completely executed within a time slot; constraint (8d) and (8e) show that MEC server must allocate a positive computing resource to user associated with it, and the sum of which cannot exceed the total computational capability of MEC server; constraint (8f) and (8g) enforce that each subcarrier can only be used by one user to avoid the multi-user interference.

III. Offloading and Resource Allocation Strategy

On the observation of $\mathbf{P1}$, $\mathbf{P1}$ is a mixed integer programming and thus is NP-hard and non-convex, finding the optimal solution is generally prohibitively due to the complexity. However, the duality gap becomes zero in multi-carrier systems as the number of subcarriers goes to infinity according to the time-sharing condition [19], [20]. Thus, the optimal solution for a non-convex resource allocation problem in multi-carrier system can be achieved in the dual domain. In this section, the joint offloading and resource allocation strategy will be designed in accordance with the iterative approach as follows.

Define $\mathcal{R}$ as all sets of possible $\boldsymbol{\lambda}$ that satisfy constraint (8b), $\mathcal{F}$ as all sets of possible $\boldsymbol{f}$ that satisfy constraint (8d) and $\mathcal{X}$ as all sets of possible $\boldsymbol{X}$ that satisfy constraints (8f) and (8g).

However, it is still intractable to deal with (11) due to the coupled variants based on our observation. To decouple these variants, we can iteratively update $\boldsymbol{f}$ and $\boldsymbol{X}$ with fixed $\boldsymbol{\lambda}$ and given dual variables $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)$ at one iteratively. the next iteration to obtain the optimal $\boldsymbol{\lambda}$. The process is repeated until $\boldsymbol{\lambda}$, $\boldsymbol{f}$ and $\boldsymbol{X}$ converge, which is known as the block coordinate descent (BCD) method [21] and given detailedly as follows.

A. Resource allocation strategy

In this subsection, the resource allocation strategy will be designed to assign the computational capabilities of MEC server, and allocate the subcarriers for users.

First, for given $\boldsymbol{\lambda}$, the $\mathbf{P1}$ can be written in the form of $\mathbf{P2}$ as below.

$$\begin{aligned}
\mathbf{P2}: \ \min_{\boldsymbol{f},\boldsymbol{X}} \ & \sum_{k\in\mathcal{K}} E_k && \text{(9a)}\\
\text{s.t.} \ & \max\{t_{k,l}, t_{k,\text{off}}\} \le T, \forall k, && \text{(9b)}\\
& \sum_{k\in\mathcal{K}} f_{k,m} \le F, && \text{(9c)}\\
& && \text{(9d)}
\end{aligned}$$

The Lagrangian function for $\mathbf{P1}$ is given by

$$\begin{aligned}
\mathcal{L}(\boldsymbol{f}, \boldsymbol{X}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) = & \sum_{k\in\mathcal{K}} E_k + \sum_{k\in\mathcal{K}} \alpha_k \left[\frac{(1-\lambda_k) R_k c_k}{f_k} - T\right] \\
& + \sum_{k\in\mathcal{K}} \beta_k \left(\frac{\lambda_k R_k}{r_k} + \frac{\lambda_k R_k c_k}{f_{k,m}} - T\right) + \gamma\left(\sum_{k\in\mathcal{K}} f_{k,m} - F\right),
\end{aligned} \tag{10}$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots \alpha_k]^T$, $\boldsymbol{\beta} = [\beta_1, \beta_2, \cdots \beta_k]^T$ and γ are the non-negative Lagrange multipliers. The dual function is then defined as

$$g(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) = \min_{\boldsymbol{f}\in\mathcal{F}, \boldsymbol{X}\in\mathcal{T}} \mathcal{L}(\boldsymbol{f}, \boldsymbol{X}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma). \tag{11}$$

Furthermore, the dual problem is given by

$$\begin{aligned}
\max \ & g(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) \\
\text{s.t.} \ & \boldsymbol{\alpha} \succeq \mathbf{0}, \ \boldsymbol{\beta} \succeq \mathbf{0}, \ \gamma \ge 0.
\end{aligned} \tag{12}$$

we may solve P2 in order to obtain the optimal computing resource and subcarrier allocation strategies with fixed offloading strategy.

1) Computational capabilities assignment: First, for given $\boldsymbol{\lambda}$, Therefore the following condition is both necessary and sufficient for the computing resource allocation's optimality:

$$\frac{\partial \mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{X}, \boldsymbol{f}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)}{\partial f_{k,m}^*} = 2 f_{k,m}^* \kappa_m \lambda_k R_k c_k - \frac{\beta_k c_k R_k \lambda_k}{f_{k,m}^{2*}} + \gamma = 0 \tag{13}$$

Unfortunately, It is difficult to find a closed-form experssin for the optimal $f_{k,m}^*$. However, since $\mathcal{L}$ is a convex function of $f_{k,m}$, and $\frac{\partial \mathcal{L}}{\partial f_{k,m}}$ monotonically increases as $f_{k,m}$ increases, we can adopt the bisection search method to solve $f_{k,m}^*$ over $0 \le f_{k,m} \le f_{k,m}^{max}$. The detailed solution process of $f_{k,m}$ is in Algorithm 1. The whole procedure to solve $f_{k,m}$ is summarized in Algorithm 1 in the following.

2) Subcarrier allocation strategy: When the optimal computing resource $f^*$is calculated, and with fixed offloading ratio $\boldsymbol{\lambda}$, we may solve P1 in order to obtain the optimal subcarrier allocation strategy.

Then, because (10) can be rewritten as

$$\begin{aligned}
& \mathcal{L}(\boldsymbol{X}, \boldsymbol{f}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) = \\
& \sum_{k\in\mathcal{K}} \left(p_k \frac{\lambda_k R_k}{r_k} + \beta_k \frac{\lambda_k R_k}{r_k}\right) + \\
& \sum_{k\in\mathcal{K}} \Bigg[\kappa_k c_k (1-\lambda_k) R_k f_k^2 + \kappa_m \lambda_k R_k f_{k,m}^2 + \gamma f_{k,m} \\
& + \alpha_k \left(\frac{(1-\lambda_k) R_k c_k}{f_k} - T\right) + \beta_k \left(\frac{\lambda_k R_k c_k}{f_{k,m}} - T\right)\Bigg] - \gamma F
\end{aligned} \tag{14}$$

Algorithm 1 Proposed Iterative Algorithm

Input: The Offloading Ratio $\boldsymbol{\lambda}$.

Initialize: ρ.

1: **for** $k \in \mathcal{K}$ **do**
2: **Initialize:** $f_{k,m}^{UB} = f_{k,m}^{max}$ **and** $f_{k,m}^{LB} = 0$.
3: **repeat**
4: **Set** $f_{k,m}$ = $\frac{1}{2}(f_{k,m}^{UB} + f_{k,m}^{LB})$.
5: **Compute** $\frac{\partial L}{\partial f_{k,m}}$ **according to** (13).
6: **if** $\frac{\partial \mathcal{L}}{\partial f_{k,m}} > 0$ **then**
7: **Set** $f_{k,m}^{UB} = f_{k,m}$.
8: **else**
9: **Set** $f_{k,m}^{LB} = f_{k,m}$.
10: **until** $\left|\frac{\partial \mathcal{L}}{\partial f_{k,m}}\right| < \rho$, **where** ρ **is a very small constant for controlling accuracy.**
11: **Obtain the optimal** $f_{k,m}^*$

So, the (10) can be transformed into

$$\begin{aligned} &\mathcal{L}(\boldsymbol{X}, \boldsymbol{f}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) = \\ &= \mathcal{L}^{sub}(\boldsymbol{X}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) + \sum_{k\in\mathcal{K}} \Bigg[\kappa_k c_k (1-\lambda_k) R_K f_k^2 + \\ &\kappa_m \lambda_k R_k f_{k,m}^2 + \gamma f_{k,m} + \alpha_k \left(\frac{(1-\lambda_k) R_k c_k}{f_k} - T \right) + \\ &\beta_k \left(\frac{\lambda_k R_k c_k}{f_{k,m}} - T \right) \Bigg] - \gamma F \end{aligned} \tag{15}$$

where

$$\mathcal{L}^{sub}(\boldsymbol{X}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) = \sum_{k\in\mathcal{K}} \left(p_k \frac{\lambda_k R_k}{r_k} + \beta_k \frac{\lambda_k R_k}{r_k} \right) \tag{16}$$

And once the optimal computing resource $\boldsymbol{f}^*$ is calculated, and with fixed offloading ratio $\boldsymbol{\lambda}$. the energy consumed for MEC calculations and local calculations has been fixed. So we only need to consider the energy consumed by user uploads.

So the minimization in (11) can be tuned into

$$\min_{\boldsymbol{X}\in\mathcal{T}} \mathcal{L}^{sub}(\boldsymbol{X}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) = \max_{\boldsymbol{X}\in\mathcal{T}} \mathcal{L}^{\frac{1}{sub}}(\boldsymbol{X}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) \tag{17}$$

where

$$\mathcal{L}^{\frac{1}{sub}}(\boldsymbol{X}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) = \sum_{k\in\mathcal{K}} \left[\frac{\sum_{n\in\mathcal{N}} x_{k,n} B \log_2 \left(1 + \frac{p_k g_{k,n}}{\sigma_n^2}\right)}{(p_k + \beta_k)\lambda_k R_k} \right] \tag{18}$$

Suppose that subcarrier n is assigned to user k, we have

$$\mathcal{L}^{\frac{1}{sub}} = \sum_{n\in\mathcal{N}} \mathcal{L}_n \tag{19}$$

where

$$\mathcal{L}_n = B \log_2 \left(1 + \frac{p_k g_{k,n}}{\sigma_n^2}\right) \left[\frac{1}{(p_k + \beta_k)\lambda_k R_k} \right] \tag{20}$$

Thus, the subproblem is given by

$$\max_{\boldsymbol{X_n}\in\boldsymbol{\mathcal{T}}} \mathcal{L}_n(\boldsymbol{X_n}, \boldsymbol{\beta}) \tag{21}$$

where $\boldsymbol{X}_n = \{x_{k,n}\}_{k=1}^K$, and this problem can be solved independently. By maximizing each $\mathcal{L}_n$, the optimal $\boldsymbol{X}$ can be obtained as

$$x_{k,n}^* = \begin{cases} 1, \text{if} \quad k = k^* = \mathrm{argmin}_k \, \mathcal{L}_n \\ 0, \text{ otherwise} \end{cases} \tag{22}$$

3) Lagrange Multipliers Update: Then, we have determined the optimal $\boldsymbol{f}^*, \boldsymbol{X}^*$ for given $\boldsymbol{\alpha}, \boldsymbol{\zeta}$ and $\gamma, \boldsymbol{\beta}$, we can discuss how to update Lagrange Multipliers in the following. The dual problem can be expressed as $\max \mathrm{g}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)$,where $\boldsymbol{\alpha} \succeq 0$ and $\gamma \geq 0, \boldsymbol{\beta} \succeq 0$. We can easily prove that the dual problem is a convex one. Thus, a subgradient of $g(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)$is given by

$$\begin{aligned} \Delta\alpha_k &= \frac{(1-\lambda_k) R_k c_k}{f_k} - T \\ \Delta\gamma &= \sum_{k\in\mathcal{K}} f_{k,m} - F \\ \Delta\beta_k &= \frac{\lambda_k R_k}{r_k} + \frac{\lambda_k R_k c_k}{f_{k,m}} - T \end{aligned} \tag{23}$$

Thus,the sunbgradient projection method for(23) is as

$$\begin{aligned} \alpha_k(z+1) &= [\alpha_k(z) - \zeta_k \Delta\alpha_k]^+ \\ \gamma(z+1) &= [\gamma(z) - \rho\Delta\gamma]^+ \\ \beta_k(z+1) &= [\beta_k(z) - \eta_k \Delta\beta_k]^+ \end{aligned} \tag{24}$$

B. Offloading strategy

Given the computional capabilities assigment $\boldsymbol{f}^*(\boldsymbol{\lambda})$ and subcarrier allocation strategies $\boldsymbol{X}^*(\boldsymbol{\lambda})$ with fixed $\boldsymbol{\lambda}$ as obtained from the alogrithm in Section III, the problem $\mathbf{P1}$ can be written as $\mathbf{P3}$

$$\mathbf{P3}: \min_{\boldsymbol{\lambda}} \sum_{k\in\mathcal{K}} E_k \tag{25a}$$

$$\text{s.t. } 0 \leq \lambda_k \leq 1, \forall k, \tag{25b}$$

$$\max\{t_{k,l}, t_{k,\text{off}}\} \leq T, \forall k, \tag{25c}$$

So we can get the subproblem for user $k \in \mathcal{K}$,

$$\mathbf{P3-A}: \min_{\lambda_k} E_k \tag{26a}$$

$$\text{s.t. } 0 \leq \lambda_k \leq 1, \tag{26b}$$

$$\max\{t_{k,l}, t_{k,\text{off}}\} \leq T, \tag{26c}$$

Then we hava the value range of λ_k from (26c) is expressed as

$$1 - \frac{T f_k}{c_k R_k} \leq \lambda_k \leq \frac{T r_k f_{k,m}}{R_k f_{k,m} + r_k R_k c_k} \tag{27}$$

and for $\mathbf{P3-A}$, we have

$$\frac{\partial E_k}{\partial \lambda_k} = \frac{r_k R_k c_k \left(f_{k,m}^2 \kappa_m - f_k^2 \kappa_k\right) + p_k R_k}{r_k} \tag{28}$$

So, we can get the λ_k^*

$$\lambda_k^* = \begin{cases} \left[1 - \dfrac{T_k f_k}{c_k R_k}\right]^+, & \dfrac{\partial E_k}{\partial \lambda_k} > 0 \\ \min\left\{\dfrac{T_k r_k f_{k,m}}{R_k f_{k,m} + r_k R_k c_k}, 1\right\}, & \dfrac{\partial E_k}{\partial \lambda_k} < 0 \end{cases} \tag{29}$$

where $[x]^+ = \max\{0, x\}$ and if $\frac{\partial E_k}{\partial \lambda_k}$ = 0, then the value of E_k is independent of λ_k.

TABLE I
SIMULATION PARAMETERS

MEC System Parameters	Values
The CPU frequency of MEC sever	10GHz
The CPU frequency of mobile users	0.6-0.7GHz
The tansmit power of users	0.6W
Input data size of users	$1*10^6 - 1.5*10^6$ bits
Background noise σ^2	10^{-13}W
Subcarrier bandwidth B	12.5KHz

C. Algorithm

The whole procedure to solve P1 is summarized in Algorithm 2 in the following.

Algorithm 2 Proposed Iterative Algorithm

initialize:

- **Set** $\boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma, \mathcal{Z}_{max}, \epsilon$**.**
- **Set z = 0**

repeat
  repeat
    Allocate computing resource $\boldsymbol{f}$ according to Algorithm 1.
    Determine subcarrier allocation $\boldsymbol{X}$ according to (22)
    Solve task offload ratio $\boldsymbol{\lambda}$ according to
    Update $\boldsymbol{\alpha}$, and $\gamma, \boldsymbol{\beta}$ from (24)
    z = z + 1
    if $\|\boldsymbol{\alpha}(z+1) - \boldsymbol{\alpha}(z)\| \leq \epsilon$ and $\|\gamma(z+1) - \gamma(z)\| \leq \epsilon$ and $\|\boldsymbol{\beta}(z+1) - \boldsymbol{\beta}(z)\| \leq \epsilon$ **then**
      break.
  until $z > \mathcal{Z}_{max}$
  Determine offloading ratio $\boldsymbol{\lambda}$ according to (29)
until The total energy consumption convergence.

IV. NUMERICAL RESULTS

In the section, we evaluate on the performance of the proposed algoritm with simulation results. In the simulations, we consider a single-server network, with Besides, other simulation parameters employed in the simulations, unless otherwise mentioned, are summarized in TABLE1.

V. CONCLUSIONS

REFERENCES

[1] "A Survey on Mobile Edge Computing: The Communication Perspective," *IEEE Commun. Surv. Tutorials*, vol. 19, no. 4, pp. 2322–2358, 2017, ISSN: 1553877X. DOI: 10.1109/COMST.2017.2745201.

[2] P. Porambage, J. Okwuibe, M. Liyanage, M. Ylianttila, and T. Taleb, "Survey on Multi-Access Edge Computing for Internet of Things Realization," *IEEE Commun. Surv. Tutorials*, vol. 20, no. 4, pp. 2961–2991, 2018.

[3] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile edge computing-A key technology towards 5G. ETSI White Paper," *ETSI White Pap.*, vol. 11, no. 11, pp. 1–16. 2015.

[4] C. You, K. Huang, H. Chae, and B. H. Kim, "Energy-Efficient Resource Allocation for Mobile-Edge Computation Offloading," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 3, pp. 1397–1411, 2017, ISSN: 15361276. DOI: 10.1109/TWC.2016.2633522.

[5] Y. Mao, J. Zhang, and K. B. Letaief, "Dynamic Computation Offloading for Mobile-Edge Computing with Energy Harvesting Devices," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3590–3605, 2016, ISSN: 07338716.

[6] L. Ji and S. Guo, "Energy-efficient cooperative resource allocation in wireless powered mobile edge computing," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4744–4754, 2019, ISSN: 23274662. DOI: 10.1109/JIOT.2018.2880812.

[7] T. Q. Dinh, J. Tang, Q. D. La, and T. Q. Quek, "Offloading in Mobile Edge Computing: Task Allocation and Computational Frequency Scaling," *IEEE Wirel. Commun. Netw. Conf. WCNC*, vol. 65, no. 8, pp. 3571–3584, 2017, ISSN: 15253511. DOI: 10.1109/WCNC.2017.7925612.

[8] Z. Kuang, L. Li, J. Gao, L. Zhao, and A. Liu, "Partial Offloading Scheduling and Power Allocation for Mobile Edge Computing Systems," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6774–6785, 2019. DOI: 10.1109/jiot.2019.2911455.

[9] C. Wang, F. R. Yu, C. Liang, Q. Chen, and L. Tang, "Joint Computation Offloading and Interference Management in Wireless Cellular Networks with Mobile Edge Computing," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7432–7445, 2017, ISSN: 00189545. DOI: 10.1109/TVT.2017.2672701.

[10] J. Zhang, W. Xia, Y. Zhang, Q. Zou, B. Huang, F. Yan, and L. Shen, "Joint Offloading and Resource Allocation Optimization for Mobile Edge Computing," *2017 IEEE Glob. Commun. Conf. GLOBECOM 2017 - Proc.*, vol. 2018-Janua, pp. 1–6, 2018.

[11] J. Zhang, X. Hu, Z. Ning, E. C. Ngai, L. Zhou, J. Wei, J. Cheng, and B. Hu, *Energy-latency tradeoff for energy-aware offloading in mobile edge computing networks*, 2018. DOI: 10.1109/JIOT.2017.2786343.

[12] M. Li, S. Yang, Z. Zhang, J. Ren, and G. Yu, "Joint subcarrier and power allocation for OFDMA based mobile edge computing system," *IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC*, vol. 2017-Octob, pp. 1–6, 2018. DOI: 10.1109/PIMRC.2017.8292207.

[13] L. Cui, C. Xu, S. Yang, J. Z. Huang, J. Li, X. Wang, Z. Ming, and N. Lu, "Joint optimization of energy consumption and latency in mobile edge computing for internet of things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4791–4803, 2019, ISSN: 23274662. DOI: 10.1109/JIOT.2018.2869226.

[14] X. Yang, X. Yu, H. Huang, and H. Zhu, "Energy efficiency based joint computation offloading and resource allocation in multi-access MEC systems," *IEEE Access*, vol. 7, pp. 117 054–117 062, 2019. DOI: 10.1109/ACCESS.2019.2936435.

[15] X. Hu, K. K. Wong, and K. Yang, "Wireless powered cooperation-assisted mobile edge computing," *IEEE*

Transactions on Wireless Communications, vol. 17, no. 4, pp. 2375–2388, 2018.

[16] W. Zhang, Y. Wen, K. Guan, K. Dan, and D. O. Wu, "Energy-optimal mobile cloud computing under stochastic wireless channel," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4569–4581, 2013.

[17] Y. Wang, S. Min, X. Wang, W. Liang, and J. Li, "Mobile-edge computing: Partial computation offloading using dynamic voltage scaling," *IEEE Transactions on Communications*, vol. 64, no. 10, pp. 4268–4282, 2016.

[18] S. Bi, J. Ying, and 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.

[19] K. Seong, M. Mohseni, and J. M. Cioffi, "Optimal resource allocation for ofdma downlink systems," in *IEEE International Symposium on Information Theory*, 2006.

[20] Y. Wei and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, 2006.

[21] P. Richtărik and M. Takáč, "Iteration complexity of randomized block-coordinate descent methods for minimizing a composite function," *Mathematical Programming*, vol. 144, no. 1-2, pp. 1–38, 2014.