Dynamic Computation Offloading and Resource Allocation for Multi-access Edge Computing Networks

Meng Wang[1](✉), Shuo Shi[1,2], and Xi Wang[1]

[1] School of Electronic and Information Engineering, Harbin Institute of Technology, Harbin 150001, Heilongjiang, China
wangmeng_hit@163.com, crcss@hit.edu.cn

[2] Peng Cheng Laboratory, Network Communication Research Centre, Shenzhen 518052, Guangdong, China

Abstract. This paper studies the computation offloading and resource allocation of Multi-access Edge Computing Networks (MEC) networks under stochastic task arrivals and imperfect channle state conditions (CSI). We aim to minimize the energy consumptions of all mobile users (MUs) while satisfying the stability of all MUs' queues by optimizing the transmit power, computational frequency and offloading ratio. The problem is modeled as a a long-term average stochastic optimization problem which cannot be solved trivially. To address this, we firstly apply the Lyapunov framework to transform the long-term average optimization problem into a series of per-slot optimization subproblems, and then an iterative optimization method is proposed to decompose the perslot optimization subproblem into three toy problems. The simulation results demonstrate the convergence and performance of the proposed algorithm.

Keywords: multi-access edge computing · computation offloading · resource allocation · imperfect CSI

1 Introduction

With the continuous demand for emerging services such as virtual reality, augmented reality, ultra-high-definition video live broadcast, and ultra-low-latency online games, mobile terminals need to run more and more tasks that are both latency-sensitive and consume a lot of computing resources [1]. However, the growth of the battery capacity of MUs is far from meeting the requirements for handling those tasks [2]. To address this, the European Telecommunications Standards Institute (ETSI) proposed a new paradigm namely Multi-access Edge

This work is supported by the National Natural Science Foundation of China under Grant 62171158 and the Research Fund Program of Guangdong Key Laboratory of Aerospace Communication and Networking Technology under Grant 2018B030322004.

A. Li et al. (Eds.): 6GN 2022, LNICST 504, pp. 14–25, 2023.
https://doi.org/10.1007/978-3-031-36011-4_2

Computing (MEC) in 2016. By deploying computing servers at the edge of the network which are close to MUs, MEC allows MUs to offload computational sensitive tasks to the MEC server, thereby greatly reducing the computing burden on the user side [3]. Meanwhile, compared with the high latency of mobile cloud computing (MCC) to complete tasks, MEC allows MUs to offload tasks to MEC servers through wireless interfaces (4G, 5G, WIFI, etc.), thereby reducing the transmission delay between users and servers. However, it should be noted that although MEC can greatly improve the MUs' quality of experience, it faces many technical challenges in the actual implementation process. The computation offloading and resource allocation strategy of MEC is the core problem among many theoretical problems to be solved [4]. Specifically, the objective of computation offloading is to design a strategy which aims to solve the problem of "Where to execute". Moreover, the objective of resource allocation is to solve the problem of "How to Execute" [1–4]. The connotation of the successful completion of the task has two aspects. One is that the task can be completed within the specified delay, the other is the MUs' queue can be kept stable during the long-term task execution process. The performance of MEC can be greatly improved by jointly optimizing the computation offloading and resource allocation strategy [5]

At present, many scholars have made contributions in this field [6] is a pioneering work in this direction. Sardellitti et al. considered the resource allocation problem in multi-cell MEC for the first time, respectively in single-user and multi-user scenarios. The objective is to minimize the total energy consumption, and the optimal strategy is obtained by the convex optimization method under the condition of satisfying the task processing delay. However, in [6], all tasks are processed by the edge server by default, and all MUs are considered to process only one fixed task without considering the dynamic task arrivals [7] further considers the queue stability problem caused by the dynamic arrival of tasks. Jeongho et al. introduced Lyapunov optimization into the joint offloading and resource allocation of MEC for the first time, and proposed the DREAM algorithm, which reduced the energy consumption by 35% compared with the benchmark algorithm. However, the mobile cloud computing scenario considered in [7] does not consider the fading effect brought by the wireless uplink. In [8], under the premise of considering the wireless uplink, considering the singleserver multi-user scenario, an online joint optimization strategy is designed by optimizing the offload ratio, computational frequency, transmit power and spectrum allocation strategy, and the average queue length and energy consumption are considerd. Simulation results demonstrate the relationship between average queue length and energy consumption. In [9], the terminal is further extended from the smart phone to the IoT device. First, the problem is modeled as a longterm average optimization problem, and the online joint computing and resource allocation algorithm is designed using Lyapunov optimization. The simulation results prove the convergence and effectiveness of the algorithm.

As can be seen from the previous review, the current research in this direction mainly considers using the Lyapunov optimization framework to reduce the energy consumption of all users while ensuring queue stability. However, although the Lyapunov optimization framework does not rely on the priori distribution of the channel state, it relies on the accurate CSI at the current moment. Due to the sophisticated communication scenario, MUs cannot guarantee to obtain a perfect CIS at each moment [10]. In [11], the CSI at each moment is defined as the channel estimation value plus an estimation error, where the estimation error is a random variable. Therefore, the channel state at each moment is also a random variable. There is no work to prove whether the Lyapunov optimization framework can still obtain the trade-off between queue stability and energy consumption when the channel state is unknown at the current moment. Based on the above two points, the main contributions of this paper can be summarized as follows.

- This paper considers the joint computing offloading and resource allocation of a single cell MEC system under the premise of imperfect CSI between MUs and base stations, the problem is modeled as a long-term average stochastic optimization problem. We prove that the Lyapunov optimization framework can still maintain the stability of the queue and obtain a trade-off between the average energy comsumptions and the queue stability by applying Jenson's inequality.
- Then, the Lyapunov optimization framework is applied to transform the long-term average optimization problem into a per-slot optimization problem, and the corresponding per-slot optimization problem is solved by an iterative optimization method. Specifically, by dividing the original problem into sub-problems of calculation frequency optimization, offloading ratio optimization, and optimization of transmission power, each subproblem is solved separately.
- The simulation results show the convergence and performance superiority of the algorithm. It confirms the tradeoff relationship between average queue length and total energy consumptions in Lyapunov optimization under imperfect CSI. Moreover, we compare the performance of the proposed method with three benchmark algorithms.

The rest of the paper is organized as follows. Section 2 presents the system model and formulate the problem as long-term average optimization problem. Section 3 converts the problem into a per-slot optimization problem by applying the Lyapunov optimization framework, and propose an iterative optimization method to solve the per-slot optimization problem. Section 4 gives the numerical results of the proposed method. Finally, this paper is conclued in Section 5.

2 System Model and Problem Formulation

2.1 Network Model

Figure 1 gives an illustration of the considered network model. Specifically, we consider a single-cell network consists of one MEC server located in the BS and N MUs. We assume the tasks in this paper can be offloaded to the BS in the partial offloading manner, where tasks can be executed simultaneously both in MEC server and local processor under an offloading ratio [12]. We consider a discrete time slot system $\mathcal{T} = \{1, 2, 3, \cdots, T\}$ and ΔT denotes the time interval of a single time slot. In each time slot $t \in \mathcal{T}$ tasks arrive dynamically with $\mathbb{E}\{A_i(t)\} = \lambda, \forall \mathrm{i} \in [1, \mathrm{N}]$, where $A_i(t) = (S_i(t), C_i(t), L_i(t))$.$S_i(t)$, $C_i(t)$ and $L_i(t)$ denote the size (bits), CPU cycles per bit and the maximum tolerant latency, respectively. Each MU has a local execution queue and a remote offloading queue to buffer the incoming computational tasks.

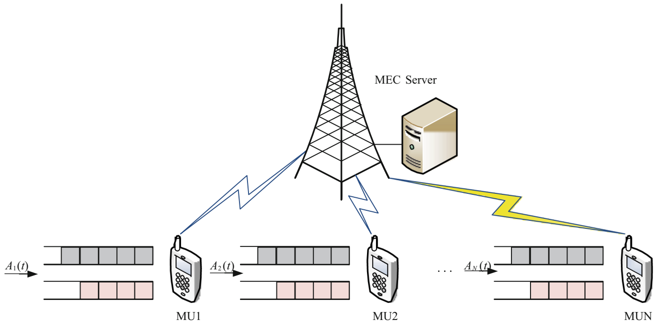

Fig. 1. An illustration of a multi-user MEC system under dynamic task arrivals.

2.2 Communication Model

In the beginning of each time slot, MUs offload the task to the MEC server through wireless uplink. In this paper, we assume the imperfect CSI between MUs and the BS, which is given by

$$h_i(t) = \hat{h}_i(t) + \phi, \forall i \in [1, N] \tag{1}$$

where $\hat{h}_i(t)$ is the channel estimate value of i-th MU, which considers the large scale fading and small scale fading. ϕ is the channel estimation error denoted by $\phi \sim \mathcal{CN}\left(0, \sigma_z^2\right)$, where σ_z^2 is the variance. According to [14], the uplink transmission rate can be given by

$$R_i(t) = \frac{B}{N} \log_2\left(1 + \frac{|h_i(t)|^2 p_i(t)}{|\phi|^2 p_i(t) + \frac{N_0 B}{N}}\right), \forall \text{i} \in [1, \text{N}] \tag{2}$$

where B denotes the bandwidth of MUs, $p_i(t)$ is the transmit power, N_0 is the white noise spectrum density. In this paper, we assume the bandwidth is equally allocated to N MUs, the spectrum resource allocation issue will be discussed in our future work. The probability density function (PDF) of $|\phi|^2$ can be formulated as

$$p\left(|\phi|^2 = r\right) = \frac{1}{2\sigma_z^2} \exp\left(-\frac{r}{2\sigma_z^2}\right), 0 < r < \infty. \tag{3}$$

The uplink transmission time between i-th MU and the BS can thus be modelled as:

$$T_i^U(t) = \frac{\beta_i(t) S_i(t)}{R_i(t)}, \forall \text{i} \in [1, \text{N}]. \tag{4}$$

The transmit energy consumption in each time slot can thus be given as:

$$E_i^U(t) = p_i(t) T_i^U(t), \forall \text{i} \in [1, \text{N}]. \tag{5}$$

According [4], we do not consider the downlink transmission time in this paper, the reason is that compared with the uplink offloading process, the size of the result after processing can be neglected, and therefore the downlink transmission time is negligible.

2.3 Computational and Queueing Model

According to [5], the computational energy consumption of i-th MU can be given as

$$E_i^L(t) = \epsilon(1 - \beta_i(t))C(t)f_i(t)^2, \tag{6}$$

where ϵ is the compuataion coefficient the CPU, $f_i(t)$ is the computation frequency of i-th MU. In this paper, we focus on the energy consumption in the MUs' side, and thus the energy consumption in the MEC server is not taken into consideration. Moreover, we assume the computation capacity in the MEC server is sufficient so that the exeuction delay in MEC server is negligiable. Therefore, the local execution delay can be given as

$$T_i^L(t) = \frac{(1 - \beta_i(t))C_i(t)S_i(t)}{f_i(t)} \tag{7}$$

According to [15], the local execution queue can be modeled as

$$Q_i^l(t+1) = \max\left[Q_i^l(t) + I_i^l(t) - O_i^l(t), 0\right], i \in [1, N], \tag{8}$$

where $I_i^l(t) = (1 - \beta_i(t)) \sum_{k=1}^{A_i(t)} S_k$ represents the input data in t-th time slot.

$O_i^l(t) = \frac{f_i(t)\Delta T}{C_i}$ represents the output data in t-th time slot. Similarly, the remote offloading queue can be modeled as

$$Q_i^r(t+1) = \max\left[Q_i^r(t) + I_i^r(t) - O_i^r(t), 0\right], i \in [1, N], \tag{9}$$

where $I_i^r(t) = \sum_{k=1}^{A_i(t)} S_i\beta_i(t)$, $O_i^r(t) = R_i(t)\Delta T$.

2.4 Problem Formulation

According to the communication and computational model, the total energy consumption in a single time slot can be given as

$$E_i^T(t) = E_i^U(t) + E_i^L(t) \tag{10}$$

The total execution delay in time slot t can be given as

$$T_{total}(t) = \max(T_U(t), T_l(t)) \tag{11}$$

Therefore, the long-term average energy-efficient computation offloading problem can be formulated as

$$\mathbf{P1} \min_{\boldsymbol{f}(t),\boldsymbol{\beta}(t),\boldsymbol{p}(t),} \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{N} \left\{E_i^T(t)\right\} \tag{12a}$$

$$s.t. 0 \le f_i(t) \le f_{\max}, \forall \text{t} \in [1, \text{T}], \forall \text{i} \in [1, \text{N}] \tag{12b}$$

$$0 \le p_i(t) \le p_{\max}, \forall \text{t} \in [1, \text{T}], \forall \text{i} \in [1, \text{N}] \tag{12c}$$

$$0 \leq \beta_{i,k}(t) \leq 1, \forall t \in [1,T], \forall i \in [1,N], \forall k \in [1,A(t)] \tag{12d}$$

$$\frac{\left(1-\beta_{i,k}(t)\right)C_i(t)S_i(t)}{f_i(t)} \leq L_i(t), \forall \mathrm{t} \in [1,\mathrm{T}], \forall \mathrm{i} \in [1,\mathrm{N}], \forall \mathrm{k} \in [1,\mathrm{A(t)}] \tag{12e}$$

$$\Pr\left(\frac{\beta_{i,k}(t)S_i(t)}{R_i(t)} > L_i(t)\right) < \eta, \forall t \in [1,T], \forall i \in [1,N], \forall k \in [1,A(t)] \tag{12f}$$

$$\lim_{T\to\infty}\frac{E\left\{\left|Q_i^l(t)\right|\right\}}{T} = 0, \lim_{T\to\infty}\frac{E\left\{\left|Q_i^r(t)\right|\right\}}{T} = 0 \tag{12g}$$

The objective is to minimize the average total energy consumption of all MUs while satsifying the queue stability constraints. Specifically, (12b) and (12c) gives the constraints on computational frequency and transmit power of MUs, respectively. (12d) gives the constraint on local execution time. (12e) gives the constraint on remote offloading time, since $R_{\mathrm{i}}(t)$ is a random variable, (12e) is given in the probability form. (12f) gives the constarint on queue stability [15]. **P**1 is a stochastic optimization problem and cannot be solved directly through conventional optimization technique [8].

3 Proposed Iterative Optimization Method

To solve **P**1, we first apply Lyapunov optimization framework. Specifically, according to [15], we first define the Lyapunov function as follows

$$L(Q(t)) = \frac{1}{2}\sum\nolimits_{i=1}^{N}\left(Q_i^l(t)^2 + Q_i^r(t)^2\right) \tag{13}$$

Then we can define the Lyapunov Drift which is given by

$$L(\boldsymbol{Q}(t+1)) - L(\boldsymbol{Q}(t)) = \frac{1}{2}\sum_{i=1}^{N}\left(\Delta Q_i^l(t) + \Delta Q_i^r(t)\right) \tag{14a}$$

$$\Delta Q_i^l(t) = \max\left[Q_i^l(t) + I_i^l(t) - O_i^l(t), 0\right]^2 - Q_i^l(t)^2 \tag{14b}$$

$$\Delta Q_i^r(t) = \max\left[Q_i^r(t) + I_i^r(t) - O_i^r(t), 0\right]^2 - Q_i^r(t)^2 \tag{14c}$$

The conditional Lyapunov Drift is then expressed as

$$\Delta\boldsymbol{Q}(t) = \{L(\boldsymbol{Q}(t+1)) - L(\boldsymbol{Q}(t))|\boldsymbol{Q}(t)\} \tag{15}$$

By applying

$$\max[x-y+z,0]^2 \leq x^2+y^2+z^2+2x(z-y) \tag{16a}$$

$$\mathbb{E}\{A_i(t)|\mathrm{Q(t)}\} = \{A_i(t)\} = \lambda_i \tag{16b}$$

We can given an upper bound of ΔQ(t), which is given by

$$\begin{aligned}\Delta(\boldsymbol{Q}(t)) \leq K + \sum_{i=1}^{N} \lambda_i\big(Q_i^l(t) + Q_i^r(t)\big) \\ -\left\{\sum_{i=1}^{N} Q_i^l(t)O_i^l(t) + Q_i^r(t)O_i^r(t)|\mathrm{Q(t)}\right\}\end{aligned} \tag{17}$$

where K is a constant. To obtain a tradeoff between energy consumption and queue length, we define the Lyapunov Drift-Plus-Penalty function as follows.

$$\begin{aligned}\Delta(\boldsymbol{Q}(t)) + V\big\{E_i^T(t)|\boldsymbol{Q}(\mathrm{t})\big\} \leq K + V\{\hat{e}((t))|\boldsymbol{Q}(\mathrm{t})\} \\ + \sum_{i=1}^{N} \lambda_i\big(Q_i^l(t) + Q_i^r(t)\big) - \left\{\sum_{i=1}^{N} Q_i^l(t)O_i^l(t) + Q_i^r(t)O_i^r(t)|\boldsymbol{Q}(\mathrm{t})\right\}\end{aligned} \tag{18}$$

According to the max-weight algorithm proposed in [15], we can convert the long-term stochastic optimization problem into per-slot optimization problem by minimizing the following term in each time slot.

$$\begin{aligned}&\omega(\mathbb{X}(t)) = V\big(E_i^T(t)\big) \\ &- \sum_{i=1}^{N} Q_i^l(t)\Big(\hat{b}_i^l(\mathbb{X}(t)) - I_i^l(t)\Big) - \sum_{i=1}^{N} Q_i^r(t)\Big(\hat{b}_i^{r'}(\mathbb{X}(t)) - I_i^r(t)\Big)\end{aligned} \tag{19}$$

where $\mathbb{X}(t) = [\boldsymbol{f}(t), \boldsymbol{\beta}(t), \boldsymbol{p}(t)]$, $\hat{b}_i^l(\mathbb{X}(t)) = f_i(t)\Delta T$ and $\hat{b}_i^{r'}(\mathbb{X}(t)) = \Delta T Q_i^r(t)R_i(t)$.

Thus, $\boldsymbol{P}$ can be transformed into

$$\mathbf{P2} \min_{\boldsymbol{f}(t), \boldsymbol{\beta}(t), \boldsymbol{p}(t),} w(\mathbb{X}(t)) \tag{20a}$$

$$s.t. 0 \leq f_i(t) \leq f_{\max}, \forall \mathrm{t} \in [1, \mathrm{T}], \forall \mathrm{i} \in [1, \mathrm{N}] \tag{20b}$$

$$0 \leq p_i(t) \leq p_{\max}, \forall \mathrm{t} \in [1, \mathrm{T}], \forall \mathrm{i} \in [1, \mathrm{N}] \tag{20c}$$

$$0 \leq \beta_{i,k}(t) \leq 1, \forall \mathrm{t} \in [1, \mathrm{T}], \forall \mathrm{i} \in [1, \mathrm{N}], \forall \mathrm{k} \in [1, \mathrm{A(t)}] \tag{20d}$$

$$\frac{\big(1 - \beta_{i,k}(t)\big)C_i(t)S_i(t)}{f_i(t)} \leq L_i(t), \forall \mathrm{t} \in [1, \mathrm{T}], \forall \mathrm{i} \in [1, \mathrm{N}], \forall \mathrm{k} \in [1, \mathrm{A(t)}] \tag{20e}$$

$$\Pr\left(\frac{\beta_{i,k}(t)S_i(t)}{R_i(t)} > L_i(t)\right) < \eta, \forall \mathrm{t} \in [1, \mathrm{T}], \forall \mathrm{i} \in [1, \mathrm{N}], \forall \mathrm{k} \in [1, \mathrm{A(t)}] \tag{20f}$$

However, **P**2 is still a non-convex optimization problem due to (20a) and (20f). To address this, we propose an iterated optimzation method to optimize $\mathbb{X}(t)$ separately. Specifically, the computation frequency optimization subproblem can be formulated as

$$\mathbf{P3}\text{: } \min_{f(t)} \Delta T V \sum_{i=1}^{N} \Big(\varepsilon f_i(t)^3\Big) - \sum_{i=1}^{N} Q_i^l(t)\left(\frac{f_i(t)\Delta T}{C_i(t)}\right) \tag{21a}$$

$$\text{s.t.}0 \leq f_i(t) \leq f_{\max}, \forall \text{i} \in [1, \text{N}] \tag{21b}$$

$$\frac{\left(1-\beta_{i,k}(t)\right)C_i(t)S_i(t)}{f_i(t)} + \frac{Q_i^l(t-1)}{f_{\max}} \leq L_i(t), \forall i \in [1, N], \forall k \in [1, A(t)] \tag{21c}$$

Note that $f(t)$ can be decoupled in **P**3, we can define the component-wise ob jective function as follows

$$F_1(f_i(t)) = \varepsilon \Delta T V f_i^3(t) - Q_i^l(t)\Delta T f_i(t)/C_i(t) \tag{22}$$

By taking the first derivative, we can find the stationary point $f_i^*(t) = \sqrt{\frac{Q_i^l(t)}{3\varepsilon V C_i(t)}}$ (21c) can be rearranged as

$$f_i(t) \geq \frac{\left(1-\beta_{i,k}(t)\right)C_i(t)S_i(t)}{L_i(t) - \frac{Q_i^l(t-1)}{f_{\max}}} \triangleq \Omega \tag{23}$$

The optimal $f^{\star}(t)$ is then determined by the relationship of $f_{\text{i}}(t)$ between Ω and $f_{\max}$. Specifically, when $\Omega \leq f_{\max}$, (21b) can be rewritten as $\Omega \leq f_{\text{i}}(t) \leq f_{\max}$. Then, when $f_{\text{i}}^*(t) \in [\Omega, f_{\max}]$, we have $f^{\star}(t) = f_{\text{i}}^*(t)$. When $f_{\text{i}}^*(t) \notin [\Omega, f_{\max}]$, $f^{\star}(t) \in \{\Omega, f_{\max}\}$, $F^{\star} = \min\{F_1(\Omega), F_1(f_{\max})\}$. Furthermore, when $\Omega > f_{\max}$, (21c) can be ignored. By letting $\Omega = 0$ in the previous results, we can obtain $f^{\star}(t)$ when $\Omega > f_{\max}$.

Now we consider the offloading ratio optimization subproblem, which is formualted as

$$\text{P4} : \min_{(t)} \sum_{n=1}^{N} Q_i^l(t)\left(\sum_{k=1}^{A_i(t)}\left(1-\beta_{i,k}(t)\right)S_i(t)\right) + \sum_{n=1}^{N} Q_i^r(t)\left(\sum_{k=1}^{A_i(t)} S_i(t)\beta_{i,k}(t)\right) \tag{24a}$$

$$s.t.0 \leq \beta_{i,k}(t) \leq 1, \forall k \in [1, A_i(t)], \forall i \in [1, N] \tag{24b}$$

$$\frac{\left(1-\beta_{i,k}(t)\right)C_i(t)S_i(t)}{f_i(t)} + \frac{Q_i^l(t-1)}{f_{\max}} \leq L_i(t), \forall k \in [1, A_i(t)], \forall i \in [1, N] \tag{24c}$$

$$\Pr\left(\frac{\beta_{i,k}(t)S_i(t)}{R_i(t)} + \frac{Q_i^r(t-1)}{R_{\max}(t)} \geq L_i(t)\right) \leq \eta, k \in [1, A_i(t)], \forall i \in [1, N] \tag{24d}$$

It is not difficult to observed that (24a) is a convex objective function, (24b) and (24c) are convex constraints. According to (3), (24d) can be reformulated as

$$\beta_{i,k}(t)S_i(t) + \frac{x_1 \log_2(1+c)B_g}{N} < \frac{\log_2(1+c)BL_i(t)}{N} \tag{25}$$

where $x_1 = \frac{Q_i^r(t-1)}{R_{\max}(t)}$, $c = \frac{|h_i(t)|^2}{\frac{N_0 B}{N p_i(t)} - 2\sigma_z^2 \ln \eta}$ are not related to $\beta_{i,k}(t)$. Therefore, (24d) is also a convex constratint, which shows **P**4 is a convex optimization problem, which can be solved by optimization toolbox, e.g., CVX.

Finally, we focus on the transmit power optimization problem, which is given by

$$\mathbf{P5}: \min_{p(t)} \Delta TV \sum_{i=1}^{N} (p_i(t)) - \sum_{i=1}^{N} Q_i^r(t)(\Delta T(R_i(t))) \tag{26a}$$

$$\text{s.t.} 0 \le p_i(t) \le p_{\max}, \forall i \in [1, N], \forall t \in [1, T] \tag{26b}$$

$$\Pr\left(\frac{\beta_{i,k}(t)S_i(t)}{R_i(t)} + \frac{Q_i^r(t-1)}{R_{\max}(t)} \ge L_i(t)\right) \le \eta, k \in [1, A(t)], \forall i \in [1, N] \tag{26c}$$

Note that (26a) has random variables, therefore we can only minmize its expectation. Furthermore, since $R_\mathrm{i}(t)$ is a convex function with respect to $|\phi|^2$, therefore we can apply Jenson's equality to convext (26a) into

$$\Delta TV \sum_{i=1}^{N} (p_i(t)) - \sum_{i=1}^{N} Q_i^r(t)\left(\Delta T\left(\frac{B}{N}\log_2\left(1 + \frac{|h_i(t)|^2 p_i(t)}{2\sigma_z^2 p_i(t) + \frac{N_0 B}{N}}\right)\right)\right) \tag{27}$$

It is trivial to prove that (27) is a convex function with respect to $p_i(t), \forall i \in [1, N]$ Moreover, (26c) can be converted as

$$\frac{\beta_{i,k}(t)S_i(t)}{L_i(t)B} - \frac{1}{N}\log_2\left(1 + \frac{p_i(t)\left|h_g^i(t)\right|^2}{N_0 B_g/N - p_i(t)2\sigma_z^2 \ln\eta}\right) \le 0 \tag{28}$$

By taking the second derivative of (28) with respect to $p_\mathrm{i}(t)$, we can prove (28) is a convex constraint. Therefore, **P**5 now can be solved by minimizing its upper bound. Until now, we can solve **P**2 in an iterative fashion.

4 Simulation Results

The main simulation parameters of this paper are given in Table 1. The simulation results are averaged over 10000 slots. Figure 2 shows the impact of V on the total average energy consumptions when $N = 5$, $\lambda = 3$. It is observed that the total energy consumptions decreases with the increase of V . When $V = 1e9$, we have the minimum energy consumption $3.4337e$ 6j and the algorithm converges. The reason is that when V becomes larger, the optimization goal foucus more on energy consumptions thereby resulting in a lower average energy consumptions which is consistent with the results in [15].

Figure 3 shows the impact of V on the total average queue length when $N = 5$, $\lambda = 3$. It is observed that the total average queue length increases with the increase of V . This results is consistent with the result of Fig. 2, which shows that by applying the Lyapunov optimization framework, we can obtain a $[O(1/V), O(V)]$ tradeoff. Therefore, we can dynamically adjust the value of V to acheive diffrent optimzation objectives.

Table 1. Parameters settings

Parameters	Values
Bandwidth B	$2*10^6$
Noise power density N_0	-174 dBm/Hz
Data input size $S(t)$	500–2000 KB
Maximum tolerant delay $L(t)$	0.01–0.1s
Maximum CPU frequency of MUs $f_{\max}$	1 GHz
Maximum transmit power $p_{\max}$	23 dBm
Computation factor $C(t)$	40
The variance of the channel estimation errorσ_Z^2	10^{-7}
Computation energy coefficient ϵ	10^{-26}
The successful transmission rate η	0.001
The threshold of the iterative method ξ	0.001

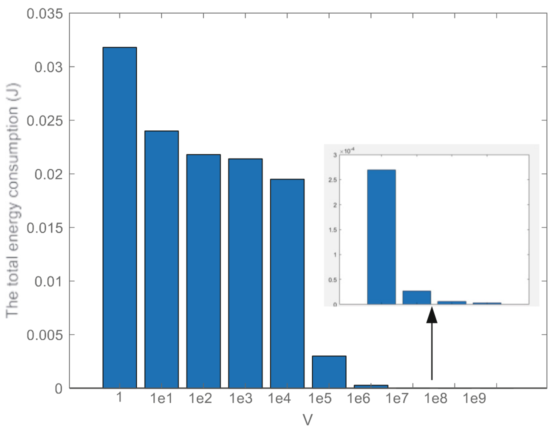

Fig. 2. Average total energy consumption versus V when $N = 5$, $\lambda = 3$.

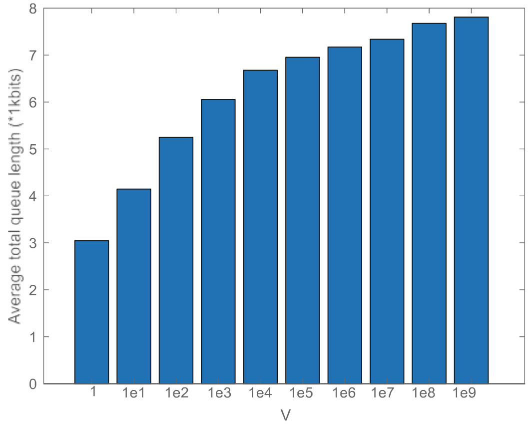

Fig. 3. Average total queue length versus V when $N = 5$, $\lambda = 3$.

5 Conclusion

In this paper, we mainly studied the dynamic computation offloading and resource allocation problem of a single cell MEC networks. Considering the stability of the queue, the problem is modeled as a long-term average stochastic optimization problem. In order to solve this problem, we leverage the Lyapunov optimization framework to obtain the trade-off between queue stability and total energy consumption under the condition of imperfect channel state. Specifically, Lyapunov optimization is used to transform the long-term average problem into a single-time optimization problem, which is solved iteratively by decomposing the optimization problem into three sub-problems. In order to deal with the non-convex constraint in the transmit power optimization subproblem, we relax the non-convex objective function by applying Jenson's equality, and transforms the non-convex problem into a convex optimization problem. Finally, simulation results verify the convergence and effectiveness of the algorithm.

References

1. Abbas, N., Zhang, Y., Taherkordi, A., et al.: Mobile Edge Computing: a Survey. IEEE Internet Things J. **5**(1), 450–465 (2018)
2. Mach, P., Becvar, Z.: Mobile edge computing: a survey on architecture and computation offloading. IEEE Commun. Surv. Tutor. **19**(3), 1628–1656 (2017)
3. Mao, Y., You, C., Zhang, J., Huang, K., Letaief, K.B.: A survey on mobile edge computing: the communication perspective. IEEE Commun. Surv. Tutor **19**(4), 2322–2358 (2017)
4. Dinh, T.Q., Tang, J., La, Q.D., Quek, T.Q.S.: Offloading in mobile edge computing: task allocation and computational frequency scaling. IEEE Trans. Commun. **65**(8), 3571–3584 (2017)
5. Chen, X., Jiao, L., Li, W., Fu, X.: Efficient multi-user computation offloading for mobile-edge cloud computing. IEEE/ACM Trans. Networking **24**(5), 2795–2808 (2016)
6. Sardellitti, S., Scutari, G., Barbarossa, S.: Joint optimization of radio and computational resources for multicell mobile-edge computing. IEEE Trans. Sig. Inf. Process. Over Netw. **1**(2), 89–103 (2015)

7. Kwak, J., Kim, Y., Lee, J., et al.: DREAM: dynamic resource and task allocation for energy minimization in mobile cloud systems. IEEE J. Sel. Areas Commun. **33**(12), 2510–2523 (2015)
8. Mao, Y., Zhang, J., Song, S.H., et al.: Stochastic joint radio and computational resource management for multi-user mobile-edge computing systems. IEEE Trans. Wireless Commun. **16**(9), 5994–6009 (2017)
9. Chen, Y., Zhang, N., Zhang, Y., et al.: Dynamic computation offloading in edge computing for internet of things. IEEE Internet Things J. **6**(3), 4242–4251 (2019)
10. Yang, Z., Ding, Z., Fan, P., et al.: On the performance of non-orthogonal multiple access systems with partial channel information. IEEE Trans. Commun. **64**(2), 654–667 (2016)
11. Fang, F., Wang, K., Ding, Z., et al.: Energy-efficient resource allocation for NOMAMEC networks with imperfect CSI. IEEE Trans. Commun. **69**(5), 3436–3449 (2021)
12. Mao, Y., Zhang, J., Song, S.H., Letaief, K.B.: Stochastic joint radio and computational resource management for multi-user mobile-edge computing systems. IEEE Trans. Wireless Commun. **16**(9), 5994–6009 (2017)
13. Guo, K., Gao, R., Xia, W., Quek, T.Q.S.: Online learning based computation offloading in MEC systems with communication and computation dynamics. IEEE Trans. Commun. **69**(2), 1147–1162 (2021)
14. Xie, R., Tang, Q., Liang, C., Yu, F.R., Huang, T.: Dynamic computation offloading in IoT Fog systems with imperfect channel-state information: a POMDP approach. IEEE Internet Things J. **8**(1), 345–356 (2021)
15. Neely, M.: Stochastic network optimization with application to communication and queueing systems. Morgan & Claypool (2010)