Cooperative Offloading Based on Online Auction for Mobile Edge Computing

Xiao Zheng[1], Syed Bilal Hussain Shah[2], Liqaa Nawaf[3], Omer F. Rana[4], Yuanyuan Zhu[5], and Jianyuan Gan[6]

[1] School of Computer Science and Technology, Shandong University of Technology, Zibo,China. xiao_zheng0910@163.com
[2] School of Computing and Mathematics Manchester Metropolitan University, United Kingdom princebilal250@gmail.com
[3] Computer Science School of Technologies, Cardiff Metropolitan University, UK LLLNawaf@cardiffmet.ac.uk
[4] School of Computer Science and Informatics, Cardiff University ,Cardiff, United Kingdom
ranaof@cardiff.ac.uk
[5] School of Software Dalian University of Technology, Dalian 116620, China zhuyan s@mail.dlut.edu.cn
[6] School of Software Dalian University of Technology, Dalian 116620, China jianyuan_g@mail.dlut.edu.cn

Abstract. In the field of edge computing, collaborative computing offloading, in which edge users offload tasks to adjacent mobile devices with rich resources in an opportunistic manner, provides a promising example to meet the requirements of low latency. However, most of the previous work has been based on the assumption that these mobile devices are willing to serve edge users, with no incentive strategy. In this paper, an online auction-based strategy is proposed, in which both users and mobile devices can interact dynamically with the system. The auction strategy proposed in this paper is based on an online approach to optimize the long-term utility of the system, such as start time, length and size, resource requirements, and evaluation valuation, without knowing the future. Experiments verify that the proposed online auction strategy achieves the expected attributes such as individual rationality, authenticity and computational ease of handling. In addition, the index of theoretical competitive ratio also indicates that the proposed online mechanism achieves near-offline optimal long-term utility performance.

Keywords: Online auction strategy · Collaborative computing offloading · Long-term utility.

1 Introduction

With the continuous development of advanced wireless communication technology in recent years, the number of mobile devices has also exploded. First, these applications are typically resource-intensive, latency-sensitive, and computationally intensive. Second, the computing power required by mobile devices is still severely limited by portability operations[1]. This presents a serious test for the future of mobile devices [2].

Offloading computing tasks is a fundamental solution to the problem of resource constraints[3]. Although cloud computing has made great achievements in the past many years, when users finally offload tasks to the public cloud, there is still the problem of long delay, especially in the environment of severe network congestion. In recent years, MEC has been designed as a promising computing paradigm for mobile services with ultra-low latency[4-5]. Rather than offloading tasks to a remote cloud, mobile users address ultra-low latency issues by cooperating with end users or by performing computationally intensive tasks at the network edge of nearby facilitie[6]. It is precisely because it is known that the MEC system performance can be effectively improved through the untapped resources of a large number of mobile devices, this paper studies the MEC framework of user cooperation[7]. Specifically, some mobile devices can share their untapped resources to assist other edge users in offloading computing tasks[8].

No more than the collaboration between edge users and MEC in this case requires an incentive mechanism that can stimulate edge users and resource-rich mobile devices to collaborate with each other in the system. In the absence of proper incentives, edge users will be reluctant to offload tasks to other users, and mobile devices will be reluctant to serve other users. Furthermore, designing online incentives requires addressing special challenges compared to static policies. Edge users and mobile devices will join and leave their systems over time, which means that this paper must refer to their systems as changing dynamically.

The biggest problem in the system of this paper is the dynamic pricing of resources, which is the crux of the online incentive strategy. The system in this paper allows its edge users to specify task lengths, which can cover multiple time periods. That is, where a mobile device decides whether to service a request, it must consider current profits and the impact of online policies on its ability to service future high-margin requests. To this end, the design of its dynamic price must refer to the arrival of dynamic tasks, thereby achieving the goal of optimizing long-term system performance in an online setting. This paper adopts the term social utility, which is defined as a representation of the total benefit of edge end users and mobile devices. Restricted by resource capacity, mobile devices can only serve a certain number of users. In the case where the price has been defined too low, mobile devices will serve more current low-value tasks, under the premise that their computing resources will be quickly exhausted, causing loss of social utilities, because their mobile devices may Opportunities to serve future high-value requests will be lost. When prices are defined so high that most requests are rejected, long-term societal utilities deteriorate. Therefore, setting the right price so that the long-term social utility online is close to optimal to offline optimal is really challenging.

Another existing challenge is to ensure the real reliability of the online strategy proposed in this paper. In addition to request-private conditions such as resource requirements and task evaluation involved in offline policies, this paper also needs to address new obstacles, namely ensuring the authenticity of start times and task durations [9]. False edge users purposefully use false reports of their private information to control market decision-making action to obtain high profits, which will deteriorate the long-term profit system[10-11]. This paper adopts the technical term social utility, which is defined as the total utility of edge end users and mobile devices. The key is to incentivize edge users to claim their actual information through the right price. Only when users report false conditions will they get less utility than if they report true information. Its early classic work - the vicery-clarke-groves (VCG) strategy was used to develop a mechanism to prove its auction[12-13]. However, the existing VCG algorithm is not suitable for the online situation, because its payment determination requires the optimal distribution results. In the case of uncertain future task requests, this paper cannot obtain those optimal solutions[14].

To address all the above problems and challenges, this paper develops an incentive mechanism for online auctions with the following properties: (1) The arrival and departure of computing tasks and mobile devices are dynamic at any time, and each task will be set up with a bundled resource package in the future; the auction decision-making behavior is matched between dynamic tasks and mobile devices. (2) The auction strategy designed in this paper is conducted in an online manner and does not make any assumptions about the arrival of future request information. Despite the premise that future information is not available, task assignment decision-making must be done instantaneously. The main contributions of this paper are summarized as follows.

In this paper, an online incentive strategy is developed in a collaborative MEC environment for multi-type resource users. This paper deals with the generality of collaborative task execution: (1) Tasks are heterogeneous and require different amounts of different resources; (2) The number of tasks a mobile device can perform is limited by its resource capacity; (3) The performance of resource supply and demand can affect its unit resource price. Therefore, two mechanisms are designed in this paper. One is an offline mechanism based on VCG, which is optimal as a benchmark. The other is a true online mechanism that only refers to the current request status to make decisions. Unlike the previously studied online variant of the VCG mechanism, which is premised on future availability, this paper does not make any assumptions about the arrival of future request information[15]. This paper provides theoretical proofs that both strategies are designed to satisfy desired properties such as individual rationality, authenticity, and computational tractability. In addition, this paper presents the derivation results for the competition ratio of the online mechanism, and its performance is shown to be close to the offline optimal. It is verified by a large number of real trajectory experiments that the designed algorithm shows better efficiency.

2 System Model

A. Mobile Edge Computing

Consider here a MEC system involving M edge users, indicated by $\mathcal{M} = \{1, 2, ..., M\}$, microbase stations for mobile devices N, indicated by $\mathcal{N} = \{1, 2, ..., N\}$

devices serving the users. Mobile devices can be thought of as smartphones, mobile microclouds, ipads and Internet of Things devices. It is assumed that the system is run in timeslot mode, and each timeslot is represented as $t \in \mathcal{T}$, $\mathcal{T} = \{1, 2, ...T\}$. User i's j-th task is denoted as $\mathcal{T}_{ij} = \{t_{ij}, l_{ij}\}$, where t_{ij} stands for the start time of the task, l_{ij} stands for the length of the task, that is, the number of time slots used to complete the task. For example, $t_{ij} = 2$ and $l_{ij} = 4$ indicate that the task t_{ij} starts at slot 1 and takes 2 slots to complete. Therefore, the index number required to complete the time slot is expressed as $t'_{ij} = t_{ij} + l_{ij} - 1$. Z-Resources such as CPU, RAM, and bandwidth are assumed. Let $a^z_{ij}(t)$ be the amount of z-resources required for slottime t, whose variable $a^z_{ij}(t)$ varieties with time, and its varieties will be different due to the heterogeneity of computing tasks. $A_{ij} = \left\{a^1_{ij}, a^2_{ij}, ..., a^Z_{ij}\right\}$ is defined as a computing resource as a specified bundle, where $a^z_{ij} = \left\{a^z_{ij}(t) : \forall t \in [t_{ij}, t'_{ij}]\right\}$.

To give a mode for mobility, set t_nTN and s_n to the interval time and service duration of mobile devices $n \in N$ respectively. The computing resources of each mobile device are limited. Defining C^z_n indicates the maximum capacity of type z-type resources on mobile device N. Because the microbase station can access the all network state, it is a system controller that controls the decision making of task scheduling.

B. Auction Theory

In this paper, the interaction between edge users and mobile devices is modeled as an auction strategy, where edge users are regarded as bidders and mobile devices as sellers. The microbase-station is a trusted third-party auction manager who manages both parties and makes online decisions. Users on the edge ask nearby mobile devices to assist with tasks and provide some immediate reward when the task is completed. The stages of the auction process are as follows:

Set b^j_i to the bid of task $\mathcal{T}_{ij}$. The bidding model of the task $\mathcal{T}_{ij}$ should denoted as $\sigma^j_i = \left\{t_{ij}, l_{ij}, A_{ij}, b^j_i\right\} \in \Sigma_i$, where Σ_i is the bidding group of edge user i.

There are $\mathcal{M}, \mathcal{N}$ and $\Sigma = \{\Sigma_1, \Sigma_2, ..., \Sigma_M\}$, the auction manager can control a winning bid set $\mathcal{W}$ and a task assignment scheme, i.e., to search a mapping: $\left\{\sigma^j_i : \sigma^j_i \in \mathcal{W}\right\} \to \{n : n \in \mathcal{N}\}$ and the payment of each winning bidder $\sigma^j_i \in \mathcal{W}$. Note here that each bid σ^j_i is private info for edge user i.

In a fake auction, the bidder will present the difference between his request and his actual request. For the purpose of distinguishing, the submitted bids are indicated by $\sigma^j_i = \left\{t_{ij}, l_{ij}, A_{ij}, b^j_i\right\}$, and the actual request info is indicated by $\bar{\sigma}^j_i = \left\{\bar{t}_{ij}, \bar{l}_{ij}, \bar{A}_{ij}, q^j_i\right\}$.

C. Offline Revenue Maximization Problem The entire information about bidding and mobile devices is available in an offline environment. There is a tradeoff between the utility and the cost of completing a task, which in turn creates some utilities for the bidder. The bid assignment variable $y_n(\sigma^j_i)$ is given here, and $y_n(\sigma^j_i) = 1$ when the bid σ^j_i is assigned to the mobile device N. And the overall bid allocation strategy is $\mathcal{Y} = (y_n(\sigma^j_i) : \forall n \in \mathcal{N}, \forall \sigma^j_i \in \Sigma)$.

Bidding allocation strategy $\mathcal{Y}$ is defined, $\Lambda = (\lambda_{ij})$ is a payment rule, and λ_{ij} indicates the payment of task $\mathcal{T}_{ij}$. In order to explore this tradeoff, this paper adopts welfare benefit maximization index, which is mainly characterized via system completion utility and mobile device service cost.

1. Computation Completion Utility: Set bid Σ and bid allocation strategy $\mathcal{Y}$, and the system utility that can be completed by computing the task will be expressed as:

$$U(\mathcal{Y}) = \sum_{\sigma_i^j \in \Sigma} \sum_{n \in \mathcal{N}} y_n(\sigma_i^j) \cdot b_i^j \tag{1}$$

2. Mobile Device Service Costs: The service cost of mobile devices mainly comes from its battery energy consumption. This paper applies a linear energy consumption model based on resource consumption. It is understood that in the case of not using dynamic voltage frequency scaling, its energy consumption and CPU, RAM usage approximately show a linear relationship. Set $r_n^z(t)$ to represent the z-type resource usage on mobile device N at time t, and its corresponding execution cost can be expressed as

$$E_n^z(r_n^z(t)) = \begin{cases} g_n^z r_n^z(t) & 0 \leq r_n^z(t) \leq C_n^z \\ +\infty & otherwise \end{cases} \tag{2}$$

which g_n^z indicates the energy consumption required to use unit z-type resource in each slottime on mobile device n.

The total resource consumption in $\mathcal{T}$ is summarized by $\boldsymbol{r} = (r_n^z(t)) : \forall n \in \mathcal{N}, \forall z \in \mathcal{Z}, \forall t \in \mathcal{T}$. Therefore, its operating cost is:

$$\Omega_E(\boldsymbol{r}) = \sum_{n \in \mathcal{N}} \sum_{t \in \mathcal{T}} \sum_{z \in \mathcal{Z}} E_n^z(r_n^z(t)) \tag{3}$$

3. Utility Maximization Problem:

Set Σ_{-i}^{-j} to the entire set of claimed bid profiles for all tasks except bid σ_i^j. That $(\sigma_i^j, \Sigma_{-i}^{-j})$ stand for the entire bidding situation. The user i's untility function is: $\mu_{ij}(\sigma_i^j, \Sigma_{-i}^{-j}) = b_i^j - \lambda_{ij}(\sigma_i^j, \Sigma_{-i}^{-j})$, when exist $x_n(\sigma_i^j) = 1$. The total utility of a mobile device is to receive the total payment minus the cost of service. Social utility maximization problem (SUM) is the difference between the utility completed after task aggregation and the service cost. In short, the problem of maximizing social benefits in the system model in this paper can be transformed into the mixed integer programming problem as follows:

$$\max_{\mathcal{Y}, \boldsymbol{r}} SUM(\mathcal{Y}, \boldsymbol{r}) = U(\mathcal{Y}) - \Omega_E(\boldsymbol{r}) \tag{4}$$

s.t:

$$\sum_{\sigma_i^j \in \Sigma : t_{ij} \le t \le t'_{ij}} y_n(\sigma_i^j) a_{ij}(t) \le r_n^z(t) \qquad \forall n, \forall t, \forall z \tag{4a}$$

$$\sum_{n \in \Psi_{ij}} y_n(\sigma_i^j) \le 1 \qquad \forall \sigma_i^j \tag{4b}$$

$$y_n(\sigma_i^j) \in \{0,1\} \qquad \forall \sigma_i^j \quad \forall n \in \Psi_{ij} \tag{4c}$$

which $\Psi_{ij} = \{n \in N : t_n \le t_{ij}, l_{ij} \le s_n\}$.

Note: Condition (4a) indicates the resource capacity limit, and condition (4b) indicates the resource allocation limit. Although in offline environment, because mixed integer programming is related to combinatorial optimization of discrete decision variables, it is concluded that the optimal task scheduling scheme is not so easy to solve. By reducing the multidimensional knapsack problem, we can verify that the mixed integer programming, namely formula (4), is strongly NP-hard. Unluckily, the article must be more challenging in an online scenario where immediate decision-making behavior is made without knowing future information. Because of the resource problem that tasks may occupy multiple decision slots, the current decision and future decision will have a strong coupling relationship under the resource capacity limitation.

3 Offline Auction Strategy Formed

A. Ideal Attribute

The strategy of dedicated hopes to achieve the following important properties.

Defininational. (Single Rational Mechanism) : For each bid σ_i^j and other defined bids Σ_{-i}^{-j} , the auction mechanism is said to be single rational if the utility of the bidding strategy is non-negative in the whole strategy. i.e.,

$$\mu_{ij}(\sigma_i^j, \Sigma_{-i}^{-j}) \ge 0 \qquad \forall \sigma_i^j \in \Sigma. \tag{5}$$

Definiation2. (Authenticity): The auction strategy is credible if other bids Σ_{-i}^{-j} are given and bidders claim that their true bid σ_i^j is their dominant strategy. i.e.,

$$\mu_{ij}(\bar{\sigma}_i^j, \Sigma_{-i}^{-j}) \ge \mu_{ij}(\sigma_i^j, \Sigma_{-i}^{-j}) \qquad \forall \sigma_i^j \ne \bar{\sigma}_i^j, \quad \sigma_i^j \in \Sigma. \tag{6}$$

Definition 3. (Computationally tractable properties) : Computationally tractable if the allocation of auction strategies states that $\mathcal{Y}$ and the payment principle Λ can be computed in polynomial time.

B. VCG-enabled Offline Auction Strategy

The goal of this paper is to develop a VCG enabled offline optimal auction strategy in which the auctioneer has the entire future information situation. The optimal allocation profile is the optimal solution of the precisely maximized mixed integer programming, namely, equation (4).

Strategy 1 (VCG-OOA)

(1) The allocation strategy $\boldsymbol{Y}_n^O \triangleq (y_n^O(\sigma_i^j) \quad \forall \sigma_i^j \in \Sigma \quad \forall n \in \mathcal{N})$ is derived by optimal solution to the mixed integer programming problem with a union of global bid Σ.

(2) The payment strategy $\boldsymbol{\Lambda}_n^O \triangleq (\lambda_n^O(\sigma_i^j) \quad \forall \sigma_i^j \in \Sigma \quad \forall n \in \mathcal{N})$, which $\lambda_n^O(\sigma_i^j)$ is descripted as :

$$\lambda_n^O(\sigma_i^j) = SUM(\mathcal{Y}^o(\Sigma), \boldsymbol{r}^o(\Sigma)) - b_i^j - SUM(\mathcal{Y}^o(\Sigma - \left\{\sigma_i^j\right\}), \boldsymbol{r}^o(\Sigma - \left\{\sigma_i^j\right\})) \quad (7)$$

$\Sigma - \left\{\sigma_i^j\right\}$ indicates all bid sequences except bid b_i^j, and $\mathcal{Y}^o(\Sigma - \left\{\sigma_i^j\right\})$ indicates the optimal solution obtained when $\Sigma - \left\{\sigma_i^j\right\}$ treats as the input. The premise here is that the whole group of bid Σ is given. In this paper, some standard algorithms are used to find the optimal solution of mixed integer programming problem (4) to obtain the optimal value of the objective function $SUM(\mathcal{Y}^o(\Sigma), \boldsymbol{r}^o(\Sigma))$. The optimal solution $\mathcal{Y}^o(\Sigma)$ is used to denote the allocation decision. To get the payment for accepting the bid σ_i^j, here delete the bid σ_i^j from Σ to obtain $\Sigma - \left\{\sigma_i^j\right\}$. In this paper, $\Sigma - \left\{\sigma_i^j\right\}$ is used as input to find the optimal solution of mixed integer programming, and the optimal value of the objective function $SUM(\mathcal{Y}^o(\Sigma - \left\{\sigma_i^j\right\}), \boldsymbol{r}^o(\Sigma - \left\{\sigma_i^j\right\}))$ is obtained. The payment σ_i^j from obtaining the bid is calculated by formula (7). When the payment is 0, it is a rejected bid. Since VCG-OAA strategy adopts VCG payment strategy, it is independently reasonable and authentic.

4 Online Auction Strategy Formation

A. Ideal Attribute

This paper first gives a theorem -Myerson's theorem [30], which shows that the behavior of payment strategy with monotone allocation strategy and controlled by critical value is true and effective. Here the preference attribute $\succeq$ is defined at bid union Σ.

Definination1: Given two unequal bids σ_i^j and $\hat{\sigma}_i^j$ with same t_{ij}, $\hat{\sigma}_i^j \succeq \sigma_i^j$ iff $l_{ij} \geq \hat{l}_{ij}$, $\hat{b}_i^j \geq b_i^j$ and $A_{ij} \geq \hat{A}_{ij}$, which $A_{ij} \geq \hat{A}_{ij}$ indicates that $a_{ij}^z(t) \geq \hat{a}_{ij}^z(t)$ $\forall z \in \mathcal{Z}, t \in \mathcal{T}$.

Definination2. Monotonicity: Given two independent bids σ_i^j and $\hat{\sigma}_i^j$ where $\sigma_i^j \succeq \hat{\sigma}_i^j$, if $\hat{\sigma}_i^j$ wins, then σ_i^j also wins, indicating that the allocation strategy is monotonous.

Definination3. Criticality: Given a monotonic allocation strategy, the critical value of the bid σ_i^j is denoted by c_i^j. In the case of a winning bid σ_i^j, the value must be less than or equal to the bid price generated by c_i^j.

B. OAP-SUM Strategy

According to the applicable rules of Myerson's theorem, this paper mainly describes the scheme implementation and allocation rules of SUM online auction policy (OAP-SWM), as shown in Algorithm 1. This paper first develops an event

Algorithm 1 OAP-SUM Strategy

1: **Input:** Current Event;
2: $\tilde{t} \leftarrow Now \quad timeslot$;
3: $N(\tilde{t}) \leftarrow \{n|$the collection of mobile devices participating in the auction at$\tilde{t}\}$;
4: $\Psi(\tilde{t}) \leftarrow \{\sigma_i^j|$bid has been authorized but work has not yet been processed within$\tilde{t}\}$
5: **if** Event=='Mobile device n reaches' **then**
6: $\quad N(\tilde{t}) \leftarrow N(\tilde{t}) \cup n, \quad (t_n \leq t \leq t'_n)$;
7: **end if**
8: **if** Event=='Bid $\sigma_i^j reaches$' **then**
9: $\quad$ Computing the union Ψ_{ij} for bid σ_i^j based on $N(\tilde{t})$;
10: $\quad \mathcal{Y}(\sigma_i^j), \lambda_{ij} \leftarrow$ OAP-SUM-A $(\tilde{t}, N(\tilde{t}), \sigma_i^j, \Psi_{ij})$
11: $\quad \Psi(\tilde{t}) \leftarrow \Psi(\tilde{t}) \cup \sigma_i^j$
12: $\quad \mathcal{Y} \leftarrow \mathcal{Y} \cup \mathcal{Y}(\sigma_i^j), \Lambda \leftarrow \Lambda \cup \lambda_{ij}$
13: **end if**
return: $\mathcal{Y}$ and Λ

processing application, which involves the call processing of events such as bid arrival, bid acceptance, task completion, mobile device arrival and mobile device departure (line 2-4). OAP-SUM computes two collections, $N(\tilde{t})$ indicates the collection of mobile devices available at auction time $\tilde{t}$. $\Psi(\tilde{t})$is the collection of accepted bids for unfinished tasks at $\tilde{t}$ (line 5-6). The union $N(\tilde{t})$ is updated as soon as the new mobile device reaches. In the case of submitting a new bid, OAP-SUM first computes the union Ψ_{ij} of bid σ_i^j based on $N(\tilde{t})$ (line 7-9). OAP-SUM-A is controlled by function based on allocation decision and payment decision at lines 10. The OAP-SUM set is then updated at lines 11-12.

1) Allocation policy: This paper uses the primitive dual technique to develop allocation policy. Firstly, the relaxed integer constraint is adopted in equation (4c), and $y_n(\sigma_i^j) \in \{0,1\}$ is converted to $y_n(\sigma_i^j) \geq 0$. Based on the standard Fenchel duality principle, the cost function or the conjugate function form $\hat{E}_n^z(x)$ of $E_n^z(r_n^z(t))$ is first given.

$$\hat{E}_n^z(x) = \underset{r_n^z(t) \geq 0}{max} \{x r_n^z(t) - E_n^z(x)\} \tag{8}$$

Lemma: The conjugate function

$$\hat{E}_n^z(x) = \begin{cases} 0, & if \quad x \leq g_n^z \\ (x - g_n^z) C_n^z, & otherwise \end{cases} \tag{9}$$

Proof: After rearrangement, it was rewritten as:

$$\hat{E}_n^z(x) = \underset{0 \leq r_n^z(t) \leq C_n^z}{max} \{(x - g_n^z) r_n^z(t)\}. \tag{10}$$

Therefore when $x \leq g_n^z$, $\hat{E}_n^z(x) = 0$, which $r_n^z(t) = 0$. Else, $\hat{E}_n^z(x) = (x - g_n^z) C_n^z$, which $r_n^z(t) = C_n^z$.

The dual variables $\eta_n^z(t)$ and ν_{ij} are added to the equations (4a) and (4b) in the form of constraint conditions. The duality problem is developed as follows:

$$min \sum_{n\in\mathcal{N}}\sum_{z\in\mathcal{Z}}\sum_{t\in\mathcal{T}}\hat{E}_n^z(\eta_n^z(t)) + \sum_{\sigma_i^j\in\Sigma}\nu_{ij} \tag{11}$$

s.t

$$\nu_{ij} \geq b_i^j - \sum_{z\in\mathcal{Z}}\sum_{t_{ij}\leq t\leq t'_{ij}} \eta_n^z(t)a_{ij}^z(t) \qquad \forall n, \forall\sigma_i^j \tag{11a}$$

$$\eta_n^z(t) \geq 0 \qquad \forall n, \forall z, \forall t \tag{11b}$$

$$\nu_{ij} \geq 0 \qquad \forall\sigma_i^j \tag{11c}$$

Based on the online mode, with the constant occurrence of the original variable y and the dual variable ν_{ij}, $\eta_n^z(t)$, it is because the bid σ_i^j and the mobile device n both arrive dynamically with time T. In this paper, $r_n^z(t,\tilde{t})$, $\eta_n^z(t,\tilde{t})$ is used instead of $r_n^z, \eta_n^z(t)$ to obtain the occupied z-type resources at the future time t of the current decision time $\tilde{t}$. $\Psi(\tilde{t})$ Set $\Psi(\tilde{t})$ to the bid collection of tasks received but not yet completed at $\tilde{t}$. Its concentration on z-type resources at t and mobile device n can be expressed as: $r_n^z(t,\tilde{t}) = \sum a_{ij}^z(t) \quad \forall\sigma_i^j \in \Psi(\tilde{t})$.

Based on the principle of complementary relaxation primal duality in the Karush -Kuhn -Tucker (KKT) condition, the primal variable $y_n(\sigma_i^j) = 1$ iff the dual constraint, i.e., equation (11a) is valid in the optimal solution. In order to realize the feasibility of the double constraint of formula (11a), for each new bid σ_i^jcase, this paper defines:

$$\nu = [y]^+, \quad y = \max_{n\in\Psi_{ij}}(b_i^j - \sum_{z\in\mathcal{Z}}\sum_{t}\eta_n^z(t_{ij},t)a_{ij}^z(t)) \tag{12}$$

which $[y]^+$ indicates max $\{y,0\}$.

From this comes the allocation rule. In the case of $\nu_{ij} > 0$, bids σ_i^j will be accepted and $y_{n'}(\sigma_i^j) = 1$, otherwise, they will be rejected.

(2) Payment Principle:

The dual variable $\eta_n^z(t,\tilde{t})$ is taken as the optimal planned price for a z-type resource in time slot $t \geq \tilde{t}$ on mobile device n. Based on the off-line environment, the dual problem can be easily resolved directly to obtain these prices. However, it is difficult to get these prices when bids change dynamically over time. An auxiliary price function of $r_n^z(t) \in [0, C_n^z]$ is developed to realize online decision as soon as possible.

$$\Gamma_n^z = \frac{P_z - g_n^z}{2Z}\left(\frac{2Z(Q_z - g_n^z)}{P_z - g_n^z}\right)^{\frac{r_n^z(t)}{C_n^z}} + g_n^z \tag{13}$$

which $P_z = \min_{\sigma_{ij}} \frac{b_i^j}{\sum_{t\in[t_{ij},t'_{ij}]} a_{ij}(t)}$, $Q_z = \max_{\sigma_{ij}} \frac{b_i^j}{\sum_{t\in[t_{ij},t'_{ij}]} a_{ij}(t)}$ are respectively the lower and upper limits of the bidder's valuation of unit z-type resources, which are obtained from previous information. According to each bid σ_i^j, the bid will be rejected if its valuation is below the corresponding service cost, mainly because the existence of such a bid has the potential to benefit society at the lowest cost. So let's define $\underset{m}{max}\, g_n^z \leq P_z$. In the case of $r_n^z(t)$ starting from 0, the price is processed at the lowest price, resulting in all bids being accepted. Resource prices increase exponentially with the number of bids accepted. This paper adopts the method of raising the price to guarantee that the vendue will not offer a low price to bid. As a result, it can have enough energy to service higher value bids in the future for obtaining higher utility. As long as $r_n^z(t) = C_n^z$, the price will reach the maximum value Q_z. On this basis, all bids will be rejected. This ensures that z-type resource capacity constraints are enforceable on mobile devices. In addition, according to the given price function, it can be verified that the ratio of the change of the original objective function to the change of the dual objective function is bounded in the process of each iteration, which will be given in the later proof. Then the payment principle of the winning bidder's bid σ_i^j is realized according to the product of its resource demand and the sum of the relevant resource prices, namely,

$$\omega_{ij} = \sum_{z \in \mathcal{Z}} \sum_{t \in [t_{ij}, t'_{ij}]} \eta_n^z(t, t'_{ij}) a_{ij}^z(t). \tag{14}$$

In a word, the specific details of the allocation principle appear in OAP-SUM-A of Algorithm 2.

5 Experimental Analysis

A. Experimental Environment Setting

The task data in this paper is taken from Google database, which consists of task start time, execution time and resource requirement conditions. This article converts the request into a bid, as shown below. This paper assumes two types of resources, namely $Z = 2$. In order to obtain the bidding value of the task, the unit z-type resource valuation is taken from the random selection in P_z, Q_z, whose bidding value corresponds to its resource requirements within the quantization range of the unit valuation. The value range of P_z, Q_z varies with different experiments. $Q_z = 8$ and $P_z = 1$ are the default cases.

Under other default conditions, the edge computing system in this paper contains $N = 20$ mobile devices. In this paper, the cycle T of 300 timeslots is tried to run, and then the trajectory of the mobile device is randomly generated. It is supposed that mobile devices have the property of Poisson process arrival in average arrival interval N/T, and their service time interval s_n is selected uniformly and randomly in $[15, 35]$, and the normalized resource capacity of each mobile device comes from the uniform and random distribution within the range of $[0.25, 0.35]$. g_n^z is uniformly and randomly distributed between the range $[0.6 - 1]$.

B. Actual Acquisition of Competitive Ratio Analysis

This paper adopts online auction strategy to realize the comparison between the actual competition ratio and the corresponding theoretical one. The actual

Algorithm 2 OAP-SUM-A $(\tilde{t}, N(\tilde{t}), \sigma_i^j, \Psi_{ij})$

Initialize:$\mathcal{Y}(\sigma_i^j) = (y_n(\sigma_i^j) \quad \forall n \in \mathcal{N}), \lambda_{ij} = 0$;
for $\forall n \in \mathcal{N}(t)$ **do**
  for $\forall z \in \mathcal{Z}$ **do**
    for $\forall t = \tilde{t} : T$ **do**
      $r_n^z(t, \tilde{t}) = \sum a_{ij}^z(t) \quad \forall \sigma_i^j \in \Psi(\tilde{t})$
      $\eta_n^z(t, \tilde{t}) = \Gamma_n^z(r_n^z(t, \tilde{t}))$
    end for
  end for
end for
Obtaining n' via resolving the Eq. (12);
Computing the dual variable value ν_{ij};
Computing the union Ψ_{ij} for bid σ_i^j based on $N(\tilde{t})$;
$\nu_{ij} \leftarrow b_i^j - \sum_{z \in \mathcal{Z}} \sum_{t_{ij} \leq t \leq t'_{ij}} \eta_{n'}^z(t, \tilde{t}) a_{ij}^z(t)$
$\Psi(\tilde{t}) \leftarrow \Psi(\tilde{t}) \cup \sigma_i^j$
$\mathcal{Y} \leftarrow \mathcal{Y} \cup \mathcal{Y}(\sigma_i^j), \Lambda \leftarrow \Lambda \cup \lambda_{ij}$
if $\nu_{ij} > 0$ **then**
  $y_{n'}(\sigma_i^j) \leftarrow 1$ and $y_n(\sigma_i^j) \leftarrow 0 \quad \forall n \in \{\mathcal{N} - n'\}$
  $\lambda_{ij} \leftarrow \sum_{z \in \mathcal{Z}} \sum_{t_{ij} \leq t \leq t'_{ij}} \eta_{n'}^z(t, \tilde{t}) a_{ij}^z(t)$
else
  $\nu_{ij} \leftarrow 0$
  $y_n(\sigma_i^j) \leftarrow 0 \quad n \in \mathcal{N}$
end if
return: $\mathcal{Y}(\sigma_i^j)$ and λ_{ij}

competition ratio is based on the OAP-SUM Algorithm to realize the ratio between the maximum actual social utility and the optimal offline social utility. The value of the theoretical competition ratio is $ln(2Z\gamma)$, which $\gamma = \max_{n,z} \frac{Q_z - g_n^z}{P_z - g_n^z}$.

Fig.1(a) verifies the comparison between the actual and theoretical competition ratios of OAP-SUM Algorithm as the number of tasks grows. $Q_z = 8$, this paper learned that most of the actual competitive ratio is around 1.4, which is far less than the upper limit of the actual theoretical value, which denotes that the online strategy developed shows superior performance. However, the value of the actual competition ratio increased slightly with the grow in the number of tasks. The real reason is that with the grow of tasks, the future uncertainty will be more difficult to control, and the possibility of task allocation will be more and more, and the corresponding decision difficulty will be more and more uncontrollable. Furthermore, it is specifically understood that the theoretical ratio is not correlated with the corresponding number of tasks.

In Fig. 1(b), the functional forms of actual and theoretical competition ratios of OAP-SUM are studied when parameter γ ranges from 6 to 12. The number of tasks given here is 120 and the number of mobile devices is 20. Thus, with the increase of γ, the actual ratio will also increase. This makes sense because higher unit resource prices will lead to real improvements in competitiveness.

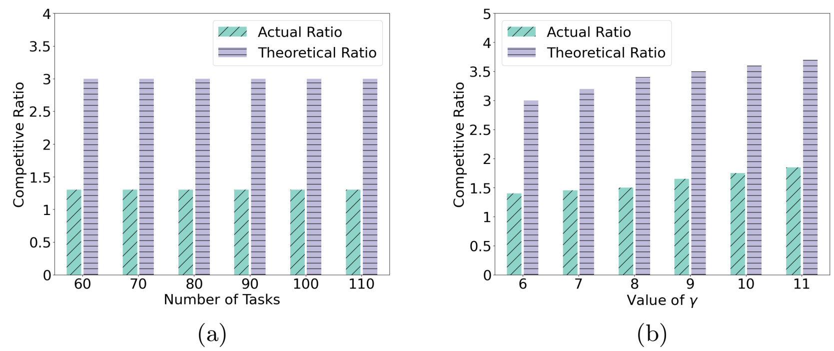


Fig. 1. Competitive Ratio of different number of tasks with γ.

The theoretical ratio is the same. This verification result is consistent with the analysis that the competition ratio is determined by the value of $ln(2Z\gamma)$.

C. Individual Rational Analysis

This paper studies the performance of OAP-SUM from the perspective of individual rationality, as shown in Fig.2. Here, 20 successful bids are randomly selected from the winning collection, and the submitted bids, actual payments, and actual execution costs are given. It can be seen from Fig.2 that the bid submitted is always greater than the actual payment price paid to mobile devices, i.e., the individual rationality is guaranteed by OAP-SUM.

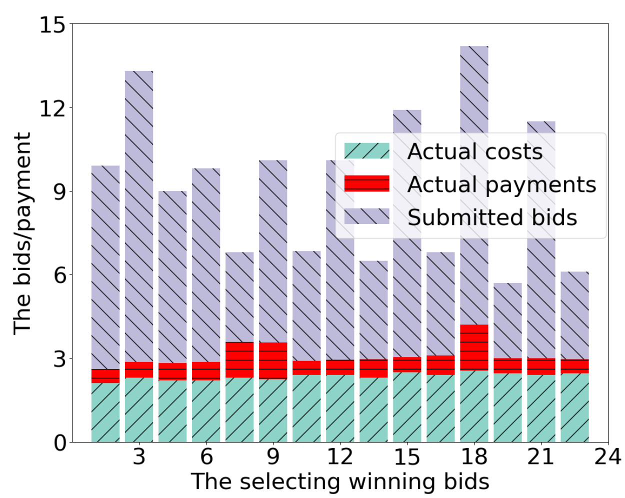


Fig. 2. Individual rational analysis

D. True Validity Analysis

Now we study the analysis of the true validity of OAP-SUM. Fig.3 shows the performance impact of unreal resource requirements and task execution time on user utility respectively. In this paper, a winning bid σ_i^j is randomly selected and its bid situation is adjusted at any time. Meanwhile, the OAP-SUM algorithm is run again with other bids unchanged. It is stated here that the user cannot declare that the execution time is shorter than the actual execution time and the actual resource demand is less. Therefore, this article only applies to the environment where the user claims that the execution time is longer and the resource demand is higher. The value on the x-axis refers to the ratio of claimed resource requirements to actual resource requirements. It follows from this that submitting more bids than actual resource requirements will reduce the user's utility, while the actual resulting true resource requirements will yield the highest utility.

Similarly, the added task execution time in Fig. 3(b) is a bid σ_i^j in the range from 10 to 20, and its real value is 10. As can be seen from the graph, submitting bids with longer execution times than actual times will reduce the user's utility, while actual verified true execution times will yield the highest utility.

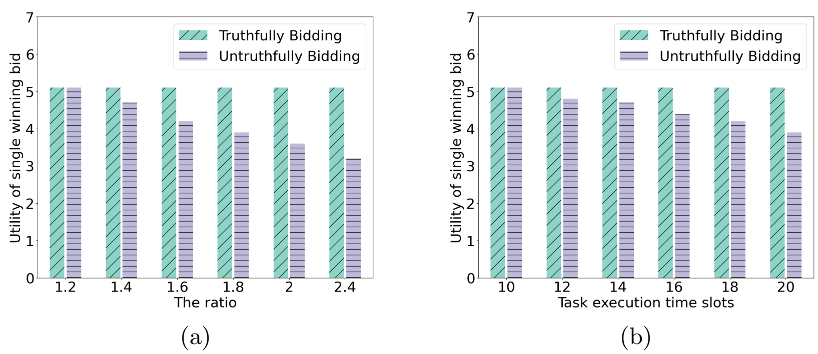
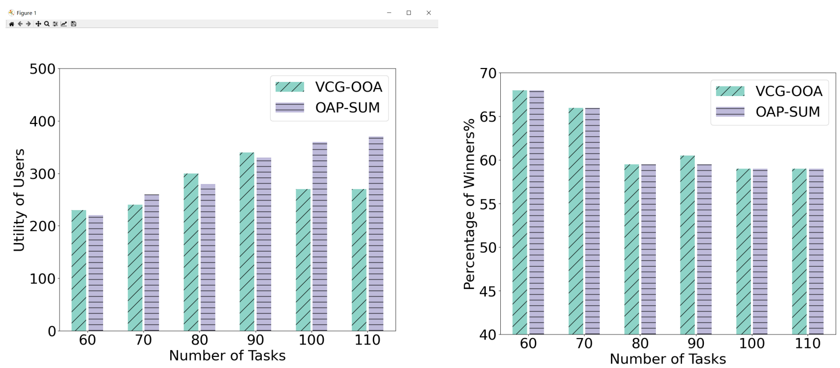


Fig. 3. Authenticity analysis

E. Comparison of the two Proposed Strategies

The online mechanism OAP-SUM is now compared with the offline mechanism VCG-OOA in terms of two performance metrics of user utility and winner percentage.

Fig.4 verifies the performance comparisons of VCG-OOA and OAP-SUM in terms of user utility and winner percentage, respectively, as the number of tasks increases. It can be concluded that the user utility of OAP-SUM is larger or smaller than that of VCG-OOA in Fig. 4(a). The reason for this is that although there is an allocation optimum, VCG-OOA may not necessarily be the payment optimum. Fig. 4(b) verifies the difference in performance comparison



(a) (b)

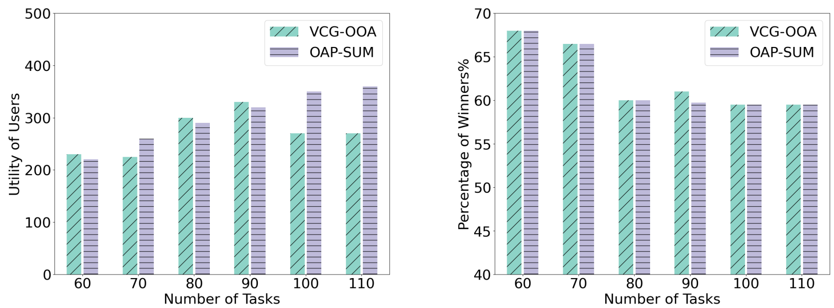
Fig. 4. Analysis and comparison of different number of tasks

of winner percentages. It is worth mentioning that the percentage of winners serves as a measure of distribution efficiency. Typically, the allocation strategy of OAP-SUM is close to optimal because its winner percentage is very close to the result of VCG-OOA. It can also be seen from the figure that as the number of tasks increases, the percentage of winners will decrease. The reason behind this is that due to the limited resource capacity of mobile devices, the increased task value is greater than the increased number of winners.

Fig.5 verifies how the winner's utility and the winner's percentage are affected when the ratio γ is increased from 6 to 12. Fig.5(a) shows that the utility of VCG-OOA is to maintain a kind of stability, while the utility of OAP-SUM decreases with the increase of γ value. The main reason is that the ratio σ is only a parameter of the marginal price function of the OAP-SUM strategy and will not affect the VCG-OOA strategy. OAP-SUM takes advantage of the increasing value of γ, which in turn charges the winner more to reduce user utility. Similarly, as shown in Fig.5(b), the percentage of winners shows a decreasing trend with the increase of γ value.

6 Conclusion

In this paper, cooperative computing offload performance in MEC is investigated. In this paper, task offloading scheduling is modeled as an NP-hard SUM problem, and an offline optimization strategy is first used as a reference benchmark. This paper further designs an online strategy that does not rely on future information, which not only schedules computing tasks and computing payments in polynomial time without involving future information, but also optimizes the long-term social utility problem in a near-optimal way . A large number of theoretical analyses show that the designed online auction achieves such properties as individual rationality, authenticity and computational tractability. Meanwhile,



(a) (b)

Fig. 5. Comparative analysis of different γ

performance evaluations on real-world trajectories also validate the effective performance of the online mechanism proposed in this paper.

References

1. Newman M . Mobile Trends 2020: mobile trends for the next 10 - a collaborative outlook[J]. 2013.
2. Slađana Jošilo, György Dán. Selfish Decentralized Computation Offloading for Mobile Cloud Computing in Dense Wireless Networks[J]. IEEE transactions on mobile computing, 2018.
3. Xu D , Yong L , Chen X , et al. A Survey of Opportunistic Offloading[J]. IEEE Communications Surveys Tutorials, 2018, 20(3): 2198–2236.
4. Zhang S , Zhang N , Zhou S , et al. Energy-Aware Traffic Offloading for Green Heterogeneous Networks[J]. IEEE Journal on Selected Areas in Communications, 2016, 34(5):1116-1129.
5. Ju R , Hui G , Xu C , et al. Serving at the Edge: A Scalable IoT Architecture Based on Transparent Computing[J]. IEEE Network, 2017, PP(5):12-21.
6. Liu G , Wang J , Tian Y , et al. Mobility-aware dynamic service placement for edge computing[J]. EAI Endorsed Transactions on Internet of Things, 2018, 5(19):163922.
7. Zhou Y , Zhang D , Xiong N , et al. Post-Cloud Computing Paradigms: A Survey and Comparison[J]. Tsinghua Science and Technology, 2017.
8. He J , Di Z , Zhou Y , et al. Towards a Truthful Online Auction for Cooperative Mobile Task Execution[C]// 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI). IEEE, 2018.
9. W Tang, K Zhang, J Ren,et.al. Flexible and Efficient Authenticated Key Agreement Scheme for BANs Based on Physiological Features[J]. Mobile Computing, IEEE Transactions on, 2018.
10. Wenjuan, Tang, Ju, et al. Enabling Trusted and Privacy-Preserving Healthcare Services in Social Media Health Networks[J]. IEEE Transactions on Multimedia, 2018.

11. Xu C , Ren J , Zhang D , et al. GANobfuscator: Mitigating Information Leakage Under GAN via Differential Privacy[J]. IEEE Transactions on Information Forensics and Security, 2019, PP(9):1-1.
12. Wang Y , Cai Z , Tong X , et al. Truthful incentive mechanism with location privacy-preserving for mobile crowdsourcing systems [J]. Computer Networks, 2018, 135:32-43.
13. Wang X , Chen X , Wu W . Towards Truthful Auction Mechanisms for Task Assignment in Mobile Device Clouds[C] IEEE INFOCOM 2017 - IEEE Conference on Computer Communications. IEEE, 2017.
14. Wu B , Chen X , Jiao L . Collaborative Computing Based on Truthful Online Auction Mechanism in Internet of Things[C] International Conference on Collaborative Computing: Networking, Applications and Worksharing. Springer, Cham, 2021.
15. Gershkov A , Moldovanu B . Efficient sequential assignment with incomplete information[J]. Games and Economic Behavior, 2010, 68(1):144-154.