DRAIM: A Novel Delay-constraint and Reverse Auction-based Incentive Mechanism for WiFi Offloading

Huan Zhou, *Member, IEEE*, Xin Chen, Shibo He, *Member, IEEE*, Jiming Chen, *Fellow, IEEE*, and Jie Wu, *Fellow, IEEE*

***Abstract*—Offloading cellular traffic through WiFi Access Points (APs) has been a promising way to relieve the overload of cellular networks. However, data offloading process consumes a lot of resources (e.g., energy, bandwidth, etc.). Given that the owners of APs are rational and selfish, they will not participate in the data offloading process without receiving the proper reward. Hence, there is an urgent need to develop an effective incentive mechanism to stimulate APs to take part in the data offloading process. This paper proposes a novel Delay-constraint and Reverse Auction-based Incentive Mechanism, named DRAIM. In DRAIM, we model the reverse auction-based incentive problem as a nonlinear integer problem from the business perspective, aiming to maximize the revenue of the Mobile Network Operator (MNO), and jointly consider the delay constraint of different applications in the optimization problem. Then, two low-complexity methods: Greedy Winner Selection Method (GWSM), and Dynamic Programming Winner Selection Method (DPWSM) are proposed to solve the optimization problem. Furthermore, an innovative standard Vickrey-Clarke-Groves scheme-based payment rule is proposed to guarantee the individual rationality and truthfulness properties of DPWSM. At last, extensive simulation results show that the proposed DPWSM is superior to the proposed GWSM and the Random Winner Selection Method in terms of the MNO's utility and traffic load under different scenarios.**

***Index Terms*—Cellular Network; Data Offloading; WiFi Access Point; Reverse Auction; Incentive Mechanism.**

I. Introduction

With the rapid popularity of mobile devices (e.g., i-pad, laptops, smart-phones) in recent years, mobile Internet services are experiencing an explosive growth [1], [2], [3]. Cellular networks are the most popular way to provide mobile Internet services today, especially with the coming of 5G networks [4], [5]. However, the explosive growth of mobile services and user demands will very likely make the cellular network overload and congest in the near future. Especially during peak time or in urban area, mobile users may face extreme performance hits in terms of low network bandwidth, missed voice calls, unreliable coverage, and so on. According to Cisco's report, it is stated that global mobile traffic was 1.2 ZB per year in 2016, but by 2021, it will reach 3.3 ZB per year [6]. Therefore, it is very urgent for the Mobile Network Operator (MNO) to offer quick and promising methods to ease the traffic burden of cellular networks.

As mobile network traffic continues to grow rapidly, mobile data offloading has become a key industry area, which uses complementary network communication technologies to offload traffic originally planned for delivery through cellular networks [7], [8], [9]. Cellular traffic can be offloaded through other complementary networks, such as Small Base Stations (SBSs), Opportunistic Mobile Networks, WiFi Access Points (APs), or Heterogeneous Networks [10], [11], [12], [13], [14]. Data offloading through SBSs, or SBSs offloading, uses low power Small Base Stations (SBSs) to offload cellular traffic. Data offloading through Opportunistic Mobile Networks, or opportunistic offloading, utilizes Opportunistic Mobile Networks to offload cellular traffic. Data offloading through WiFi Access Points (APs), or WiFi offloading, switches cellular traffic to WiFi APs to reduce traffic burden of the cellular networks when mobile users enter into a WiFi-covered area. Meanwhile, data offloading through Heterogeneous Networks is the combination of the above three data offloading methods.

Researcher find that WiFi traffic from both mobile devices and WiFi-only devices altogether accounts for more than 60% of mobile traffic in 2017 [15]. Due to the widely deployment of WiFi APs, offloading overloaded cellular traffic to WiFi APs has become an efficient and promising method. Recent studies [16], [17], [18] have demonstrated that WiFi offloading can effectively ease the traffic burden of cellular networks. However, WiFi APs may be reluctant to participate in the data offloading process without receiving the appropriate economic incentives (e.g., payment or reward) [19]. The main reason is that providing data offloading services for the MNO will incur additional resource consumption inevitably, e.g., energy consumption, bandwidth consumption, and so on. In addition, if WiFi APs assist the MNO in data offloading, their own service experience, e.g., bandwidth, transmission rate, QoS, and so on, may be affected. Hence, there is an urgent need to develop effective incentive mechanism to stimulate APs to take part in the data offloading process.

To solve the above problem, this paper proposes DRAIM, a novel Delay-constrained and Reverse Auction-based Incentive

Manuscript received August 27, 2019; accepted November 11, 2019. This work was supported in part by the National Natural Science Foundation of China under Grants No. 61602272 and 61872221. (Corresponding Authors: H. Zhou and S. He.)

H. Zhou and X. Chen are with the College of Computer and Information Technology, China Three Gorges University, Yichang 443002, China. E-mail: zhouhuan117@gmail.com, sexychenxin@gmail.com.

S. He and J. Chen are with the College of Control Science and Technology, Zhejiang University, Hangzhou 310027, China. E-mail: s18he@zju.edu.cn, cjm@zju.edu.cn.

J. Wu is with the Department of Computer and Information Sciences, Temple University, Philadelphia PA19122, USA. E-mail: jiewu@temple.edu.

Mechanism to stimulate WiFi APs to take part in the data offloading process. In DRAIM, we model the reverse auction-based incentive problem as a nonlinear integer problem from the business perspective, aiming to maximize the revenue of the MNO, and jointly consider the delay constraint of different applications in the optimization problem. In specific, we consider a simple scenario which consists of a base station of the MNO, multiple mobile users, and several WiFi APs deployed by some third party companies. The MNO acts as a buyer (i.e., an auctioneer) and the Wi-Fi APs acts as a seller (i.e., bidders). The Wi-Fi APs lease their bandwidth resources as commodities and submit bids to the MNO. After evaluating the bids from all the Wi-Fi APs, the MNO purchases the required bandwidth resources to meet the traffic demands and maximize its revenue. Finally, the MNO determines: (1) the allocation (i.e., which bidder is the winner), and (2) the price (i.e., how much to pay for each winner).

The contributions of this paper can summarize as follows:

1) The WiFi offloading problem is formulated as a reverse auction-based incentive problem from the business perspective, aiming to maximize the revenue of the MNO, and a novel Delay-constrained and Reverse Auction-based Incentive Mechanism, named DRAIM is proposed to stimulate WiFi APs to participate in the data offloading process.
2) The optimization problem is modeled as a nonlinear integer programming problem by considering the delay constraint of different applications, and two low-complexity algorithms: Greedy Winner Selection Method (GWSM), and Dynamic Programming Winner Selection Method (DPWSM) are proposed to solve the optimization problem.
3) A standard Vickrey-Clarke-Groves (VCG) scheme-based payment rule is proposed, which can guarantee the individual rationality and truthfulness properties of D-PWSM.
4) Extensive simulation results verify that the proposed DPWSM is superior to the proposed GWSM and the Random Winner Selection Method in terms of the M-NO's utility and traffic load under different scenarios.

The remainder of this paper is organized as follows. After reviewing the related work in Section II, Section III introduces the system model related to this paper. Section IV introduces the optimization problem of this paper. Section V introduces the proposed methods, and a standard VCG scheme-based payment rule. Section VI evaluates the performance of our proposed methods. At last, Section VII concludes the paper.

II. Related Work

Some studies have been developed to exploit WiFi offloading. In [20], Lee et al. introduced a quantitative study about the performance of data offloading through WiFi APs, and showed that WiFi APs can offload about 65% of the total mobile traffic and save 55% of battery power. The authors in [21] proposed a collaborative WiFi offloading architecture in Metropolitan Advanced Delivery Network, so as to increase the energy efficiency of smartphones. The authors in [22] presented an analytical model for multi-path WiFi offloading to derive the aggregate offloading time via an alternative path for multi-path offloading. Some studies also investigate the optimum deployment of WiFi APs in WiFi offloading. The authors in [23] investigated nodes' mobility patterns to deploy Wi-Fi APs, so as to maximize the continuous WiFi coverage for mobile nodes. The authors in [24] deployed Wi-Fi APs according to the density of nodes' data requests by using some real user mobility traces. However, these studies assume that WiFi APs are cooperative.

Many economic theory-based incentive mechanisms have been developed to motivate WiFi APs to provide data offloading services. The authors in [25] proposed a market-based data offloading method, in which a multi-leader multi-follower game was proposed to study the amount of traffic WiFi APs offload for the MNO and the pricing strategy of the MNO. A one-to-many bargaining game was proposed in [26] to model and analyze the amount of the MNO's offloading traffic and the WiFi APs' payment. A three-stage Stackelberg game was proposed in [19] to investigate the data offloading with price-taking and price-setting APs. The authors in [27] investigated the data offloading problem through third-party WiFi APs from the business perspective. In specific, they model the problem as a utility maximization problem and design different data offloading methods for the MNO under three scenarios. In [28], Lee et al. proposed a two-stage sequential game between the MNO and WiFi APs, and investigated how much economic revenues can be generated by WiFi offloading.

Recently, some auction-based incentive mechanisms have been proposed to solve resource allocation, D2D communications and many other applications [29], [30], [31], [32], [33], [34]. Particularly, a few studies have proposed auction-based incentive mechanisms to solve the incentive problems in data offloading. The authors in [35] introduced a reverse auction-based incentive mechanism, named Win-Coupon, to stimulate mobile nodes with high delay tolerance and large offloading potential to offload their traffic to other intermittently connected networks. In [36], Paris et al. designed a combinatorial reverse auction mechanism to select the cheapest WiFi APs and offload the maximum amount of traffic from the MNO. In [37], Dong et al. proposed a novel reverse auction-based incentive framework, named iDEAL, to enable the MNO to buy third-party resources on demand, with significant savings. In [38], Lu et al. proposed a new auction model, named EasyBid, to provide guarantees for truthfulness even when a system with imprecise valuations is considered, and designed a dynamic programming based algorithm to maximize the MNO's utility with partial truthfulness and imprecision loss. In [39], Song et al. proposed a reverse auction-based incentive mechanism with a cost constraint in content distribution via D2D communications. In [40], Hou et al. proposed a social relationship-based auction to offload cellular traffic through WiFi APs efficiently, in which the social relationship of mobile nodes is considered. In [41], Zhao et al. proposed an optimal and truthful reverse auction-based incentive framework, which can minimize the cost of the MNO and meet the traffic demand in each time period.

Different from the above existing studies, in our study, we model the reverse auction-based incentive problem as a nonlinear integer problem from the business perspective, aiming to maximize the revenue of the MNO. Furthermore, we also jointly consider the delay constraint of different applications in the optimization problem. In specific, mobile users may run different applications, and thus the maximum delay that can be tolerated for different applications should be considered.

III. System Model of DRAIM

This section introduces the system model related to DRAIM.

A. Data Offloading Model

Fig. 1 shows a mobile data offloading scenario in a single cell, which is formed by a Base Station (BS) of the MNO, several WiFi Access Points (APs) and a set of Mobile Users (MUs). Here, the BS is deployed by the Mobile Network Operator (MNO) and WiFi APs are deployed by different third-party companies. In addition, the MUs and the APs are within the service coverage of the BS, and each AP only covers a part of the MUs due to its smaller transmission power. Each MU can download its preferred content from the BS. Since the BS has a limited backhaul and radio access capacity, the MNO may select some MUs to be served by the APs to improve the overall network performance, especially when the network congestion occurs.

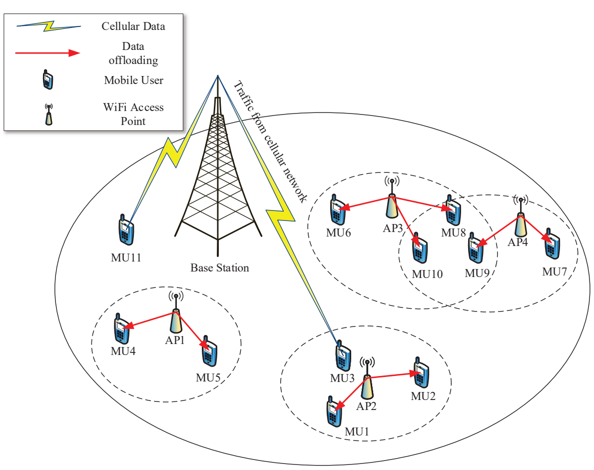


Fig. 1. Network scenario of WiFi offloading.

Since providing data offloading services will inevitably incur additional resource consumptions, the APs deployed by third-party companies usually do not actively participate in the data offloading process. To stimulate APs to participate in the data offloading process, the MNO needs to compensate for the additional resource consumptions of the APs. Moreover, when the MNO makes an offloading decision, the maximum delay that the MUs can tolerate should be considered.

APs lease their available spectrum resource blocks to the MNO, in exchange for remuneration. In detail, each AP reports its available spectrum resource blocks periodically and asks for a price from the MNO, whereas the MNO chooses some APs based on the collected information and pays them the corresponding remuneration.

B. The MNO's Revenue Model

In the network, the MNO obtains a revenue by providing content services to the MUs. We assume that the total mobile data requested by the MUs in the system is q, where the amount of traffic that can be offloaded by APs is f. Therefore, the total traffic transmitted by the BS is $(q-f)$. Since the cost invested by the MNO in the construction and maintenance of BSs in the early stage is fixed, we can set the cost of unit mobile data as e and the price as d. Then, we can obtain the revenue that the MNO receives from the unit mobile data as $(d-e)$. If the payment is not considered for the APs participating in the data offloading process, the MNO's revenue function is expressed as follows:

$$U(q,f) = (d-e)(q-f) + df, \tag{1}$$

where the first term is the revenue generated by the traffic that is transmitted by the BS, and the second term is the revenue generated by the traffic that is offloaded by APs.

We divide the data offloading process into multiple time slots, where each time slot is at least larger than the maximum tolerated delay of MUs. The MNO launches an auction and collects bids from APs in each time slot. Based on the requests of the surrounding MUs, the MNO can evaluate each AP's offloading potential according to its available spectrum resource blocks. Then, the MNO selects the most valuable APs to provide data offloading services.

C. Transmission Model

Using $\mathcal{N} = \{1, 2, 3, ..., N\}$ to denote the set of MUs. Each MU $j \in \mathcal{N}$ has the following attributes:

- The traffic demand s_j: Each MU in the network has its own traffic demand, which is based on different types of applications that need to be transferred in the specified delay.
- The maximum delay threshold δ_j: It depends on the service types of each MU. We consider that MUs running different types of applications tolerate different maximum delays. For example, applications (such as file) are more tolerant of latency than video applications (such as Webcast).
- The channel transmission rate R_{ij}: For the channel model between each AP and MU, both the path loss and small-scale fading are considered. The channel transmission rate between AP $i \in \mathcal{K}$ and MU $j \in \mathcal{N}$ is expressed as follows:

$$R_{ij} = B_{ij} \log_2 \left\{ 1 + \frac{P_t |h_j \left(d_{ij}\right)^{-\alpha}|^2}{N_0} \right\}, \tag{2}$$

where B_{ij} is the bandwidth of the leased transmission channel between AP i and MU j, P_i is the transmission power of the AP i, N_0 denotes the channel noise power, d_{ij} denotes the distance between AP i and MU j, α is the path-loss exponent, and h_j is the small-scale channel fading which is Rayleigh distributed.

D. AP's Bidding Model

Using $\mathcal{K} = \{1, 2, 3, ..., K\}$ to denote the set of APs. Each AP $i \in \mathcal{K}$ has the following attributes:

- The available spectrum resource blocks $B_i^{\max}$: It represents the available spectrum resource blocks of AP i leased to the MNO for remuneration.
- True value of the unit leased spectrum resource block v_i: It represents the true consumption per unit time of AP i's unit spectrum resource block by serving the data offloading service, and v_i is the private information belonging to the AP i.
- The bid of the unit leased spectrum resource block ϕ_i: Similar to v_i, we use ϕ_i to represent the compensation that AP i requests from the MNO, compensating for the resource consumption generated by the unit spectrum resource block per unit time. Therefore, based on bandwidth of the leased spectrum B_{ij}, and the occupied time duration $t_{ij} = \frac{s_j}{R_{ij}}$, the communication cost of MU j with AP i, defined as $\mathcal{E}_{ij}$, can be calculated as:

$$\mathcal{E}_{ij} = \phi_i B_{ij} t_{ij}. \tag{3}$$

During the auction process, each AP $i \in \mathcal{K}$ submits its bid vector $[\phi_i, B_i^{\max}]$ to the MNO, denoting its bid value ϕ_i and available spectrum resource blocks $B_i^{\max}$. Due to the nature of selfishness and rationality, AP i may raise a bid ϕ_i higher than its true value v_i to get more compensation. To make the auction process fair and reasonable, and the trade between the APs and the MNO mutually beneficial, appropriate incentive mechanism should be designed to stimulate the APs to participate in the data offloading process.

E. Reverse Auction Model

This paper uses the reverse auction to stimulate APs to take part in the data offloading process, where the MNO acts as the auctioneer and the APs act as the bidders. The APs lease their bandwidth resources as commodities and submit the bid vector to the MNO. After evaluating the bids for all APs, the MNO purchases the required bandwidth resources to meet the traffic demands and maximize its revenue. Specifically, the auction procedure includes three steps:

- Each AP submits its bid vector $[\phi_i, B_i^{\max}]$ to the MNO.
- Each MU reports to the MNO its available WiFi connection and the maximum delay that can be tolerated as well as the requested data size (i.e., s_j and δ_j, $\forall j \in \mathcal{N}$). Based on the reported information, the MNO can generate an AP-MU association set $\mathcal{F} = \{F_i\}_{i\in\mathcal{K}}$ which reflects the set of MUs within each AP's coverage, where F_i denotes the set of MUs covered by AP i. For instance, MU 6, MU 8, MU 9, MU 10 are covered by AP 3 as shown in Fig. 1. Therefore, the set of MUs covered by AP 3 is $\mathcal{F}_3 = \{6, 8, 9, 10\}$. Similarly, the set of MUs covered by AP 4 is $\mathcal{F}_4 = \{7, 8, 9, 10\}$.
- The MNO determines the subset of APs which will provide data offloading services, and calculates the payment granted to each AP.

TABLE I
NOTATIONS AND SYMBOLS

Notation	Explanation
$\mathcal{N}$	The set of all MUs
$\mathcal{K}$	The set of all APs
q	The total mobile data requested by MUs in the system model
f	The amount of mobile data that can be offloaded
e	The MNO unit mobile data cost
d	The MNO unit mobile data price
$B_i^{\max}$	The available spectrum resource blocks of AP i
v_i	True value of the unit leased spectrum resource block of AP i
ϕ_i	The bid of the unit leased spectrum resource block of AP i
s_j	The traffic demand of MU j
δ_j	The maximum delay threshold of MU j
R_{ij}	The channel transmission rate between MU j's and AP i
$\mathcal{N}_i$	The set of MUs covered by AP i
$\mathcal{E}_{ij}$	The communication cost of MU j with AP i
t_{ij}	The communication duration of MU j with AP i under the maximum delay constraint
B_{ij}	The bandwidth of the leased transmission channel between AP i and MU j
$\mathcal{F}$	AP-MU association set
F_i	The set of MUs within AP i's coverage
x_i	Binary variable that indicates if AP i wins the auction
a_{ij}	Binary variable that indicates if MU j is assigned to AP i
u_i	AP i's contribution
$\mathcal{L}_w$	The set of MUs served by winning APs
$\mathcal{T}$	The optimal AP-MU association set
T_i	The optimal set of MUs served by AP i, which can maximize the mobile data that AP i offloads in each time slot.
b_i	The total asking price of AP i
$\mathcal{W}$	The set of winning APs
$\mathcal{M}$	Winner set contribution
$\mathcal{M}_i$	AP i's marginal contribution

IV. Problem Formulation

This section describes the objective function of the MNO, and models the delay-constrained and reverse auction-based incentive problem as a nonlinear integer programming problem, aiming to maximize the revenue of the MNO.

Assuming that the MNO knows all MUs' channel state information and transmission power. Then, two indicator functions $x_i \in \{0, 1\}$ and $a_{ij} \in \{0, 1\}$ are used to indicate whether AP i is selected to offload traffic, and whether AP i is selected to provide data offloading service for MU j, respectively. If AP i is selected to conduct the offloading task, $x_i = 1$; otherwise, $x_i = 0$. Similarly, if MU j connects to AP i, $a_{ij} = 1$; otherwise, $a_{ij} = 0$. Then, the MNO's utility can be given as:

$$H_{\mathcal{K}}(x_i, a_{ij}) = U\left(\sum_{j\in\mathcal{N}} s_j, \sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{N}} a_{ij}s_j\right) - \sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{N}} x_i a_{ij}\mathcal{E}_{ij}. \tag{4}$$

where $U(q, f)$ is the revenue function of the MNO as shown in Eq. (1), but it does not consider the payment given to the winning APs. Therefore, we need to use $U(q, f)$ minus the payment to the winner.

The MNO aims to maximize its revenue with fewer payments. Hence, the optimization problem can be formulated as follows:

$$
\begin{aligned}
\max \quad & H_{\mathcal{K}}\left(x_i, a_{ij}\right) && (5)\\
\text{s.t.} \quad & \sum_{j\in\mathcal{N}} a_{ij}B_{ij} \le B_i^{\max}, \quad \forall i\in\mathcal{K}, \forall j\in\mathcal{N}, && (6)\\
& t_{ij} \le \delta_j, \qquad \forall i\in\mathcal{K}, \forall j\in\mathcal{N}, && (7)\\
& a_{ij} \le x_i, \qquad \forall i\in\mathcal{K}, \forall j\in\mathcal{N}, && (8)\\
& x_i, a_{ij} \in \{0,1\}, \qquad \forall i\in\mathcal{K}, \forall j\in\mathcal{N}. && (9)
\end{aligned}
$$

The meaning of the constraints is as follows:

- Constraint (5) makes sure that the bandwidth of the spectrum leased by AP to MU matches the maximum available spectrum resource blocks.
- Constraint (6) makes sure that MU transmission delay does not exceed its maximum delay threshold.
- Constraint (7) expresses that only winners can be assigned to assist with MU data offloading.
- Constraint (8) guarantees the integer nature of binary variables.

It is very hard to solve the problem which has integer constraints $x_i \in \{0,1\}$, $a_{ij} \in \{0,1\}$ and division mathematical operation of the variables. Compared with the 0-1 knapsack problem, the proposed optimization problem is more complicated, so it certainly belongs to the NP-hard problem. Since no algorithm in polynomial time can solve this problem, we introduce heuristic algorithms to reduce the time complexity and obtain the approximate optimal solution.

V. Main Approach of DRAIM

This section introduces heuristic algorithms to solve the above optimization problem, and reduce the computational complexity. We first propose a selection method based on the greedy algorithm, and further propose a selection method based on the dynamic programming to solve the defects in the greedy selection method. In the end, we introduce the MNO's payment determination for the winning APs.

A. Greedy Winner Selection Method

A simple way to solve the above optimization problem is to choose the APs with the maximum increase of the MNO's utility. Therefore, in this part, we first introduce a Greedy Winner Selection Method, named GWSM.

Before introducing the proposed GWSM, We first give some related definitions.

Definition 1. *(AP's contribution) The contribution of AP $i \in \mathcal{K}$ is defined as the increment of the MNO's revenue after selecting AP i, which is given as:*

$$u_i = \sum_{j\in F_i} ds_j, \tag{10}$$

where F_i denotes the set of MUs within AP i's coverage.

Definition 2. *(AP's total bid price) The total bid price of AP $i \in \mathcal{K}$ is defined as the sum of the communication costs of the MU served by AP i , which can be calculated as:*

$$b_i = \sum_{j\in F_i} \mathcal{E}_{ij}, \tag{11}$$

where $\mathcal{E}_{ij}$ denotes the communication cost of MU j with AP i .

Definition 3. *(Winner set contribution) Using $\mathcal{W}$ to denote the set of winning APs, the contribution of the winner set is defined as the sum of the contribution of all APs in $\mathcal{W}$, which can be calculated as:*

$$\mathcal{M}(\mathcal{W}) = \sum_{i\in\mathcal{W}} u_i. \tag{12}$$

Definition 4. *(AP's Marginal Contribution) Using $\mathcal{W}$ to denote the set of winning APs, the marginal contribution of AP $i \notin \mathcal{W}$ is defined as the increase in the winner set contribution caused by the winning AP i, which can be calculated as:*

$$\mathcal{M}_i(\mathcal{W}) = \mathcal{M}(\mathcal{W}\cup\{i\}) - \mathcal{M}(\mathcal{W}). \tag{13}$$

Since the MNO aims to select a set of APs to maximize its utility, according to Eq. (4), the MNO's objective function can be changed to:

$$
\begin{aligned}
H(\mathcal{W}) &= \max\left(\sum_{j\in\mathcal{N}\setminus\mathcal{L}_w}(d-e)s_j + \sum_{i\in\mathcal{W}} u_i - \sum_{i\in\mathcal{W}} b_i\right)\\
&= \max\left(\sum_{j\in\mathcal{N}\setminus\mathcal{L}_w}(d-e)s_j + \sum_{i\in\mathcal{W}}\sum_{j\in F_i} ds_j - \sum_{i\in\mathcal{W}}\sum_{j\in F_i}\mathcal{E}_{ij}\right)
\end{aligned}
\tag{14}
$$

$$
\begin{aligned}
\text{s.t.} \quad & \sum_{j\in F_i} B_{ij} \le B_i^{\max}, \quad \forall i\in\mathcal{W}, \forall j\in F_i, && (15)\\
& t_{ij} \le \delta_j, \qquad \forall i\in\mathcal{W}, \forall j\in F_i, && (16)
\end{aligned}
$$

where $\mathcal{L}_w$ denotes the set of MUs served by the winning APs.

We introduce a greedy algorithm to select the AP with the largest marginal contribution value to effectively solve the optimization problem. The AP is sorted according to the marginal contribution value of the AP minus the total bid price, as shown in Algorithm 1. The winning APs can be sorted as:

$$\mathcal{M}_{\theta(1)} - b_{\theta(1)} \ge \cdots \ge \mathcal{M}_{\theta(n)} - b_{\theta(n)} \ge \cdots \ge \mathcal{M}_{\theta(N)} - b_{\theta(N)} \tag{17}$$

where $\theta(n)$ denotes the index of the AP at the position n in the ordering. In Eq. (17), $\mathcal{M}_{\theta(n)}$ is used instead of $\mathcal{M}_{\theta(n)}(\mathcal{W})$ to simplify the notation.

Moreover, we initialize the index of the AP that has the maximum value of the marginal contribution minus the total bid price (Lines 2 in Algorithm 1). Then, we repeatedly select the winning APs whose marginal contribution value is greater than its total bid price (Lines 3-7 in Algorithm 1).

Note that although the objective of GWSM is to maximize the utility of the MNO by constantly searching APs with the largest value of the marginal contribution minus total asking price, it does not take into account the constraints (15) (16). Furthermore, there is no consideration that the same served MU may exist in the overlapping areas of the APs. For instance, as shown in Fig. 1, MU 8 $\in F_3$ and MU 8 $\in F_4$, so MU 8 may be served by both AP 3 and AP 4, but each MU can only connect to one AP in each time slot.

Algorithm 1 AP Selection in GWSM

Require: $\mathcal{K}$, $\mathcal{N}$, $B_i^{\max}$,ϕ_i,s_j,δ_j
Ensure: $\mathcal{W}$

```
W ← ∅;
i ← argmax_{k∈K} (M_k(W) − b_k);
while b_i < M_i(W) do
    W ← W ∪ {i};
    Update each winner's marginal contribution in W;
    i ← argmax_{k∈K\W} (M_k(W) − b_k);
end while
return W
```

We then take the constraints (15) (16) and the overlapped coverage into account and propose a dynamic programming based algorithm.

B. Dynamic Programming Winner Selection Method

This section proposes a Dynamic Programming Winner Selection Method to select winning APs and allocate MUs, named DPWSM. In specific, the solution is divided into two parts. The first part is to use a dynamic programming algorithm to obtain the optimal AP-MU association set in each AP's coverage area, which can maximize the mobile data that each AP offloads in each time slot. The second part is to use the greedy algorithm to select the winning AP during the auction process.

Firstly, we modify some of the definitions in GWSM. Then, we elaborate the AP selection rule in the proposed DPWSM.

Definition 5. *we modify Definition 1 and Definition 2 in GWSM as follows:*

$$u_i = \sum_{j \in T_i} ds_j, \quad b_i = \sum_{j \in T_i} \mathcal{E}_{ij}, \tag{18}$$

where T_i denotes the optimal set of MUs served by AP i, which can maximize the mobile data that AP i offloads in each time slot.

To obtain the optimal AP-MU association set T_i in each AP i's coverage area, we consider using a dynamic programming algorithm to solve. Similar to the 0-1 knapsack problem, the optimization objective of this sub-problem is to find a subset T_i in F_i, which can maximize the mobile data that AP i offloads in each time slot.

In each AP $i \in \mathcal{K}$, sorting MU $j \in F_i$ according to j in an ascending order, we can get MU $\varphi(x) \in F_i$ as:

$$\varphi(1) < \varphi(2) < \cdots < \varphi(Q). \tag{19}$$

where $\varphi(x)$ is the index of MU at the position x in the ordering and the last one in the ordering is Q.

Then, we can describe this sub-problem as when the available spectrum resource block of the AP i is $B_i^{\max}$, which MU $\varphi(x) \in F_i$ can make AP i offload the maximum amount of mobile traffic in each time slot. We can define the state of the sub-problem as:

Let $J^i[x, y]$ indicate the maximum mobile data that can be offloaded by AP i if MUs set $\{MU\ \varphi(1),\ ...\ ,\ MU\ \varphi(x)\}$ with the available spectrum resource block $y \in [0, B_i^{\max}]$ are in its coverage area. Then, the state transition equation can be expressed as:

$$J^i[x, y] = \max\left\{J^i[x-1, y], J^i\left[x-1, y - V^i_{\varphi(x)}\right] + s_{\varphi(x)}\right\}, \tag{20}$$

where $V^i_{\varphi(x)}$ is the spectrum resource block required by the MU $\varphi(x)$ under the maximum delay, which can be calculated as:

$$V^i_{\varphi(x)} = \frac{s_{\varphi(x)}}{\delta_{\varphi(x)} \log_2\left\{1 + \frac{P_t |h_{\varphi(x)} \left(d_{i\varphi(x)}\right)^{-\alpha}|^2}{N_0}\right\}}, \tag{21}$$

where $s_{\varphi(x)}$ is the traffic demand of MU $\varphi(x)$.

According to Eq. (20), if MU $\varphi(x)$ is not served by AP i, then $J^i[x, y] = J^i[x-1, y]$; if the MU $\varphi(x)$ is served by AP i, then $J^i[x, y] = J^i\left[x-1, y - V^i_{\varphi(x)}\right] + s_{\varphi(x)}$. Finally, the maximum mobile data that AP i can offload is $J^i[Q, B_i^{\max}]$, which can be obtained by traversing the state of all MU $\varphi(x) \in F_i$ according to the state equation. Since our objective is to obtain the optimal AP-MU association set T_i, we can get T_i based on the state transition of the final state $J^i[Q, B_i^{\max}]$. More details are shown in Algorithm 2.

Algorithm 2 MUs Allocation in DPWSM

Require: $\mathcal{K}$, $\mathcal{N}$, $B_i^{\max}$, ϕ_i, s_j, δ_j
Ensure: T_i

```
for each AP i ∈ K do
    T_i ← ∅;
    for each MU j ∈ F_i do
        Calculate V^i_j according to Eq. (21);
    end for
    After sorting MU j ∈ F_i according to j in an ascending
    order, let φ(x) denote the index of MU at the position
    x in the ordering and the last one in the ordering is Q.
    for x ← 1 to Q do
        for y ← 1 to B_i^max do
            J^i[x, y] ← J^i[x − 1, y];
            if y ≥ V^i_{φ(x)} then
                Calculate J^i[x, y] according to Eq. (20);
            end if
        end for
    end for
    x' = Q, y' = B_i^max;
    while x' > 0 and y' > 0 do
        if J^i[x', y'] = J^i[x' − 1, y' − V^i_{φ(x)'}] + s_{φ(x)'} then
            T_i ← T_i ∪ φ(x');
        end if
        x' = x' − 1;
    end while
end for
return T_i
```

According to the dynamic programming algorithm, we can obtain the optimal AP-MU association set $\mathcal{T} = \{T_i\}_{i \in \mathcal{K}}$ which represents the optimal set of MUs within each AP's coverage. Now, we can formulate the problem of maximizing the MNO's utility by combining GWSM as follows:

$$H(\mathcal{W}) = \max\left(\sum_{j\in\mathcal{N}\backslash\mathcal{L}_w}(d-e)s_j + \sum_{i\in\mathcal{W}}u_i - \sum_{i\in\mathcal{W}}b_i\right)$$
$$= \max\left(\sum_{j\in\mathcal{N}\backslash\mathcal{L}_w}(d-e)s_j + \sum_{i\in\mathcal{W}}\sum_{j\in T_i}ds_j - \sum_{i\in\mathcal{W}}\sum_{j\in T_i}\mathcal{E}_{ij}\right) \tag{22}$$

$$\text{s.t.} \quad \sum_{j\in T_i}B_{ij} \le B_i^{\max}, \quad \forall i\in\mathcal{W}, \forall j\in T_i, \tag{23}$$
$$t_{ij} \le \delta_j, \qquad \forall i\in\mathcal{W}, \forall j\in T_i. \tag{24}$$

To solve this problem, we first calculate the spectrum resource block demand V_j^i for each MU j under the maximum delay constraint in the coverage area of AP i. For each AP $i\in\mathcal{K}$, we use dynamic programming algorithm to obtain T_i under the constraints of (23) and (24), which are shown in Algorithm 2. Then, we select the winning APs according to the order of their marginal contributions minus total bid prices until the total bid prices of the selected APs are larger than or equal to their marginal contributions. In each while-loop of the selection of winning APs, $\mathcal{T}' = \{T_i\}_{i\in\mathcal{K}\backslash\mathcal{W}}$ needs to be updated since the same MU may exist in the optimal AP-MU association set of both the selected APs and unselected APs (Lines 9-11). For instance, it can be found that $T_3 = \{6, 8, 10\}$, $T_4 = \{7, 8\}$ as shown in Fig. 1. Once AP 3 is selected as the winning AP, $T_4 = \{7, 8\}$ should be updated as $T_4' = \{7, 9\}$ since MU 8 can only be served by one AP in each time slot. The detail of the proposed AP selection method is shown in Algorithm 3.

Algorithm 3 AP Selection in DPWSM

Require: $\mathcal{K}$, $\mathcal{N}$, $B_i^{\max}$, ϕ_i, s_j, δ_j
Ensure: $\mathcal{W}$

1: $\mathcal{W} \leftarrow \emptyset$, $\mathcal{L}_w \leftarrow \emptyset$;
2: Calculate T_i according to Algorithm 2;
3: $i \leftarrow \arg\max_{k\in\mathcal{K}\backslash\mathcal{W}} (\mathcal{M}_k(\mathcal{W}) - b_k)$;
4: **while** $b_i < \mathcal{M}_i(\mathcal{W})$ **do**
5:     $\mathcal{W} \leftarrow \mathcal{W} \cup \{i\}$,
6:     $\mathcal{L}_w \leftarrow \mathcal{L}_w \cup T_i$,
7:     $\mathcal{K} \leftarrow \mathcal{K}\backslash\{i\}$;
8:     Update each winner's marginal contribution in $\mathcal{W}$;
9:     **for** each AP $i\in\mathcal{K}$ **do**
10:         $F_i \leftarrow F_i\backslash\mathcal{L}_w$;
11:     **end for**
12:     Update T_i according to Algorithm 2;
13:     $i \leftarrow \arg\max_{k\in\mathcal{K}} (\mathcal{M}_k(\mathcal{W}) - b_k)$;
14: **end while**
15: **return** $\mathcal{W}$

C. Payment Determination

After selecting the winning APs, the MNO needs to determine the payment to compensate for their cost. As we mentioned in the AP's bidding model, each AP wants to get a higher reward which is not equal to the real value they provided. In this part, we propose a standard VCG scheme-based payment rule for DPWSM. The proposed payment rule can stimulate the involvement of APs in the data offloading process and guarantee the individual rationality and truthfulness properties.

In the standard VCG scheme, each winner will pay the "opportunity cost" caused to other participants. The "opportunity cost" of bidder i is denoted as total bids of all the other bidders that would win without the participation of bidder i, minus the sum of bids of all the other actual winning bidders.

According to Definition 5, we can know that u_i denotes the increment of the MNO's revenue after selecting AP i, and b_i denotes the sum of the communication costs of the MUs served by AP i. Mathematically, we define $H_{\mathcal{K}}^{-i}(x_i, a_{ij})$ as the optimal solution without considering the contribution of AP i, which can be formulated as:

$$H_{\mathcal{K}}^{-i}(x_i, a_{ij}) = H_{\mathcal{K}}(x_i, a_{ij}) - (u_i - b_i), \tag{25}$$

Furthermore, we use $H_{\mathcal{K}\backslash\{i\}}(x_i, a_{ij})$ to denote the new optimal solution without considering the participation of AP i. Then, the price paid to AP i is given as:

$$p_i = u_i - \left(H_{\mathcal{K}\backslash\{i\}}(x_i, a_{ij}) - H_{\mathcal{K}}^{-i}(x_i, a_{ij})\right). \tag{26}$$

Using $\gamma_i = \sum_{j\in T_i} v_i B_{ij} t_{ij}$ to denote the sum of the real value consumed by AP i in the data offloading process. Then, the utility of each AP $i\in\mathcal{W}$ is defined as:

$$\mu_i = p_i - \gamma_i. \tag{27}$$

The payment and utility of those AP $i\notin\mathcal{W}$ is defined as 0. The detail of the proposed payment rule is shown in Algorithm 4.

Algorithm 4 Payment Determination in DPWSM

Require: $\mathcal{W}$, T_i
Ensure: p_i

for each AP $i\in\mathcal{K}$ **do**
    $p_i \leftarrow 0$;
end for
for each AP $i\in\mathcal{W}$ **do**
    $H_{\mathcal{K}}^{-i}(x_i, a_{ij}) = H_{\mathcal{K}}(x_i, a_{ij}) - (u_i - b_i)$;
    $\mathcal{K} \leftarrow \mathcal{K}\backslash\{i\}$
    Update $\mathcal{W}$ according to Algorithm 3;
    Calculate p_i according to Eq. (26);
    $\mathcal{K} \leftarrow \mathcal{K} \cup \{i\}$;
end for
return $p_i, \forall i\in\mathcal{W}$

D. Proof of Properties

This part proves that our proposed DPWSM satisfies three crucial properties: individual rationality, truthfulness and computational efficiency. The individual rationality guarantees each winner can obtain a non-negative utility. The truthfulness prevents APs to obtain higher utility by bidding untruthfully.

Moreover, the computational efficiency of DPWSM can be completed in the polynomial time complexity.

Theorem 1. *(Individual Rationality). The payment rule defined in Eq. (26) satisfies the individual rationality property, i.e.,* $\forall i \in \mathcal{K}, p_i \geq \gamma_i$.

Proof: Based pm the payment rule Eq. (26), we can get:

$$\begin{aligned} p_i &= u_i - \left(H_{\mathcal{K}\setminus\{i\}}\left(x_i, a_{ij}\right) - H_{\mathcal{K}}^{-i}\left(x_i, a_{ij}\right)\right) \\ &= u_i - \left(H_{\mathcal{K}\setminus\{i\}}\left(x_i, a_{ij}\right) - H_{\mathcal{K}}\left(x_i, a_{ij}\right) - b_i + u_i\right) \\ &= H_{\mathcal{K}}\left(x_i, a_{ij}\right) - H_{\mathcal{K}\setminus\{i\}}\left(x_i, a_{ij}\right) + b_i. \end{aligned}$$

When each AP $i \in \mathcal{K}$ bids truthfully, i.e., $b_i = \gamma_i$, we can obtain:

$$\begin{aligned} \mu_i &= p_i - \gamma_i \\ &= H_{\mathcal{K}}\left(x_i, a_{ij}\right) - H_{\mathcal{K}\setminus\{i\}}\left(x_i, a_{ij}\right) \\ &\geq 0. \end{aligned}$$

Therefore, the individual rationality property is satisfied. ■

Theorem 2. *(Truthfulness). The payment rule defined in Eq. (26) satisfies the truthfulness property, i.e., it is a weakly dominant strategy for each AP to set the bid* $\phi_i = v_i$.

Proof: Assuming that a certain AP i declares the bid $\phi_i^{'}$ untruthfully, i.e., $\phi_i^{'} \neq v_i$. According to Eq. (27), the utility of AP i changes to:

$$\begin{aligned} &\mu_i^{'} = p_i^{'} - \gamma_i \\ &= u_i^{'} - \left(H_{\mathcal{K}\setminus\{i\}}\left(x_i, a_{ij}\right) - H_{\mathcal{K}}^{-i}\left(x_i^{'}, a_{ij}^{'}\right)\right) - \sum_{j\in T_i} v_i B_{ij} t_{ij}. \end{aligned}$$

Then, the difference of AP $i \in \mathcal{W}$'s utility after submitting the untruthful bid and the truthful bid can be expressed as:

$$\begin{aligned} &\Delta\mu_i = \mu_i^{\prime} - \mu_i \\ &= u_i^{'} - \left(H_{\mathcal{K}\setminus\{i\}}\left(x_i, a_{ij}\right) - H_{\mathcal{K}}^{-i}\left(x_i^{'}, a_{ij}^{'}\right)\right) - \sum_{j\in T_i} v_i B_{ij} t_{ij} \\ &\quad - \left[H_{\mathcal{K}}\left(x_i, a_{ij}\right) - H_{\mathcal{K}\setminus\{i\}}\left(x_i, a_{ij}\right)\right] \\ &= u_i^{'} + H_{\mathcal{K}}^{-i}\left(x_i^{'}, a_{ij}^{'}\right) - H_{\mathcal{K}}\left(x_i, a_{ij}\right) - \sum_{j\in T_i} \phi_i B_{ij} t_{ij} \\ &= \sum_{j\in\mathcal{N}\setminus\mathcal{L}_w^{\prime}} (d-e)s_j + \sum_{i\in\mathcal{W}^{\prime}} \sum_{j\in T_i^{\prime}} ds_j - \sum_{i\in\mathcal{W}^{\prime}\setminus\{i^{\prime}\}} \sum_{j\in T_i^{\prime}} \mathcal{E}_{ij} \\ &\quad - \left[\sum_{j\in\mathcal{N}\setminus\mathcal{L}_w} (d-e)s_j + \sum_{i\in\mathcal{W}} \sum_{j\in T_i} ds_j - \sum_{i\in\mathcal{W}\setminus\{i\}} \sum_{j\in T_i} \mathcal{E}_{ij}\right]. \end{aligned}$$

Since $\mathcal{W}$ is the optimal solution of the objective function Eq. (4), we can obatian:

$$\begin{aligned} &\sum_{j\in\mathcal{N}\setminus\mathcal{L}_w^{\prime}} (d-e)s_j + \sum_{i\in\mathcal{W}^{\prime}} \sum_{j\in T_i^{\prime}} ds_j - \sum_{i\in\mathcal{W}^{\prime}\setminus\{i^{\prime}\}} \sum_{j\in T_i^{\prime}} \mathcal{E}_{ij} \\ &\leq \sum_{j\in\mathcal{N}\setminus\mathcal{L}_w} (d-e)s_j + \sum_{i\in\mathcal{W}} \sum_{j\in T_i} ds_j - \sum_{i\in\mathcal{W}\setminus\{i\}} \sum_{j\in T_i} \mathcal{E}_{ij}. \end{aligned}$$

Therefore, $\Delta\mu_i \leq 0$, that is to say, APs cannot increase their utility by bidding untruthfully. ■

Theorem 3. *(Computational Efficiency). The proposed DPWSM is computationally efficient.*

Proof: To prove the computational efficiency of the proposed DPWSM, we just need to prove that DPWSM can be conducted in polynomial time. In DPWSM, the computational complexity primarily comes from two parts: the first part is to use Algorithm 2 to obtain the optimal AP-MU association set, the second part is to use the Algorithm 1 to select the winning APs.

Assuming that there are N MUs, K APs, W winning APs, and the maximum available spectrum resource blocks of the APs is B. In the first part, we can obtain that the computational complexity of obtaining the optimal AP-MU association set (Lines 2 in Algorithm 3) takes no more than $O(KNB)$ time. In the second part, the while-loop (Lines 4-14 in Algorithm 3) executes W times. In each loop, finding the AP with the largest marginal contribution minus total bid price (Lines 13 in Algorithm 3) takes no more than $O(K)$ time. Since we should update the APs' optimal AP-MU association set when a new AP is selected as a winner (Lines 12 in Algorithm 3), the computational complexity of Algorithm 3 is

$$O(WK^2NB).$$

Similar to Algorithm 3, the computational complexity of Algorithm 4 is $O(WK^2NB)$.

To summarize, the computational complexity of the proposed DPWSM is $O(WK^2NB)$, and DPWSM is computationally efficient. ■

VI. Performance Evaluation

This section evaluates the performance of our proposed methods, and investigates the impact of some parameters on the performance of the proposed methods.

TABLE II
Simulation Configuration

Parameter	Configuration
The number of APs	[0, 30]
The number of MUs	[0, 100]
The available spectrum resource blocks of APs	[0, 80] MHz
The MNO unit mobile data cost	0.6
The MNO unit mobile data price	1.2
Transmission power of AP	2 W
The channel noise power	10^{-6} W
The traffic demand of each MU	20 MB
The MU's maximum delay	[0.1, 1]

A. Simulation Settings

In the simulation, we consider that several APs are randomly distributed within the coverage of the BS, with uniform transmission range over $[50, 100]m$. The bid of unit spectrum resource block is normally distributed over $[0.2, 0.5]$ monetary units (e.g., US dollars, or RMB)/($MHzsec$). Each MU's maximum delay is uniformly distributed over $[0.1, 1]$ second. We vary the number of MUs and APs in the ranges [0, 100] and [0, 30], and the maximum available spectrum resource blocks of APs B_{max} in the range [0, 80], respectively. The path-loss exponent is $\alpha = 2.5$. Both MUs and the APs are placed randomly in the area. We give the default simulation parameters in Table II.

We then compare the performance of our proposed DPWSM (S.1) and GWSM (S.2) with the Random Winner Selection Method (S.3), which selects a group of APs to provide data offloading services randomly. For fairness, the total number of selected APs in the Random Winner Selection Method is the same as that of our proposed methods, and the Random Winner Selection Method uses the dynamic programming algorithm to obtain the optimal AP-MU association set $\mathcal{T} = \{T_i\}_{i \in \mathcal{K}}$. According to the proposed optimization problem, we use the MNO's utility (or revenue) and the MNO's traffic load as the performance metrics. Here, the MNO's utility is defined as Eq. (4), and the MNO's traffic load denotes the overall traffic transmitted by the BS.

B. Performance Comparison

This part compares the performance of our proposed methods, DPWSM (S.1) and GWSM (S.2) with the Random Winner Selection Method (S.3) in terms of the MNO's utility (or revenue) and the MNO's traffic load under different scenarios.

1) Different numbers of APs: We first compare the performance of our proposed methods, DPWSM (S.1) and GWSM (S.2) with the Random Winner Selection Method (S.3) in terms of the MNO's utility and traffic load, when the numbers of APs is different. We set the number of MUs as $|N| = 100$, and the available spectrum resource blocks of APs as $20MHz$.

Fig. 2 shows the performance comparison of DPWSM (S.1), GWSM (S.2), and the Random Winner Selection Method (S.3) in terms of the MNO's utility with the increase of the number of APs. With the increase of the number of APs, more APs can help the MNO offload traffic, so the MNO's utility will increase continuously. It can be found that DPWSM performs much better than GWSM when the number of APs is different, especially when the number of APs is larger. This is because DPWSM uses a dynamic programming algorithm to obtain the optimal AP-MU association set in each AP's coverage area, so the MNO of DPWSM has greater possibility to select more valuable APs to offload traffic during the auction process, while GWSM does not consider the overlapping areas of the APs. It is obvious that the Random Winner Selection Method performs worst. The main reason is that the Random Winner Selection Method selects a set of APs to participate in the data offloading process randomly.

Fig. 3 shows the performance comparison of DPWSM (S.1), GWSM (S.2), and the Random Winner Selection Method (S.3) in terms of the MNO's traffic load with the increase of the number of APs. With the increase of the number of APs, more APs can help the MNO offload traffic, so the traffic load of the MNO will decrease continuously. It can be found that the traffic load of the MNO in DPWSM decreases significantly as the number of APs increases, compared with that of GWSM and the Random Winner Selection Method, which demonstrates that DPWSM performs best. The traffic load of the MNO in the Random Winner Selection Method is the largest, which demonstrates that the Random Winner Selection Method still performs worst.

2) Different numbers of MUs: This part compares the performance of our proposed methods, DPWSM (S.1) and GWSM (S.2) with the Random Winner Selection Method (S.3) in terms of the MNO's utility and traffic load, when the number of MUs is different. We set the number of APs as $|K| = 30$, and the available spectrum resource blocks of APs as $20MHz$.

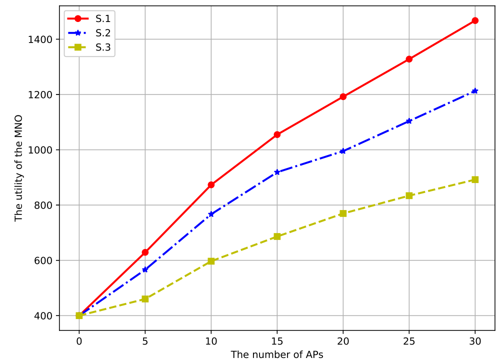


Fig. 2. The utility of the MNO with different numbers of APs.

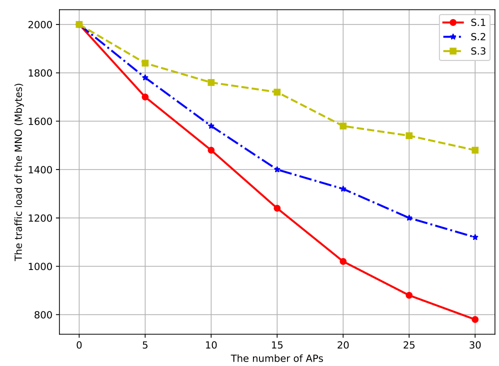


Fig. 3. The traffic load of the MNO with different numbers of APs.

Fig. 4 shows the performance comparison of DPWSM (S.1), GWSM (S.2), and the Random Winner Selection Method (S.3) in terms of the MNO's utility with the increase of the number of MUs. With the increase of the number of MUs, more traffic demands will be requested by MUs, which means that the traffic that can be offloaded by APs is also increasing. Therefore, the MNO's utility will increase continuously. Similarly, it can be found that with the increase of the number of MUs, DPWSM still performs much better than GWSM, especially when the number of APs is larger, and the Random Winner Selection Method performs worst. Furthermore, with the increase of the number of MUs, the utilities achieved by GWSM and the Random Winner Selection Method increase more slowly than DPWSM, which reflects that the proposed DPWSM has a better performance in high-density network scenarios. This is partly because the APs' spectrum resources are not being used efficiently, and there are more overlapped coverage areas among adjacent APs when the number of MUs increases.

Fig. 5 shows the performance comparison of DPWSM (S.1), GWSM (S.2), and the Random Winner Selection Method (S.3)

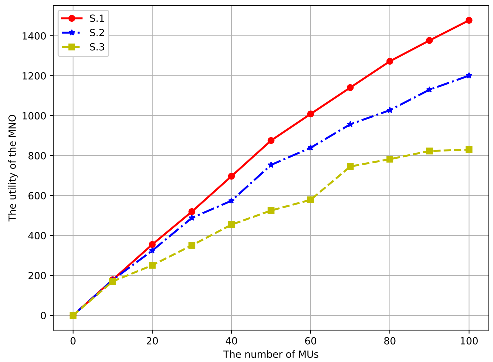


Fig. 4. The utility of the MNO with different numbers of MUs.

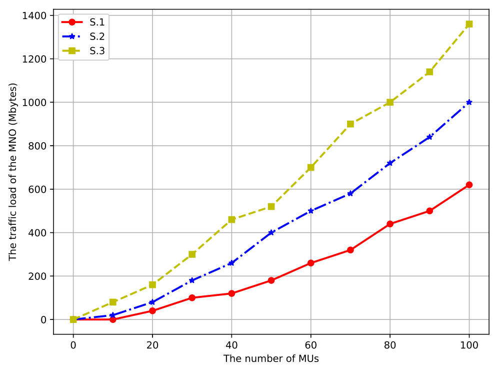


Fig. 5. The traffic load of the MNO with different numbers of MUs.

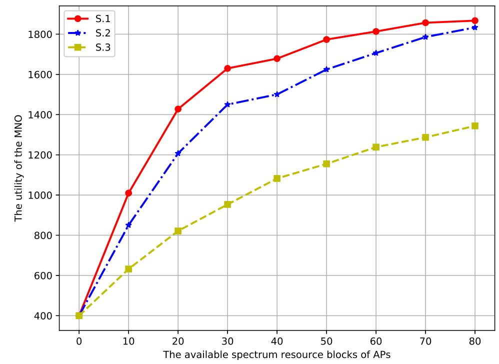


Fig. 6. The utility of the MNO with different available spectrum resource blocks of APs.

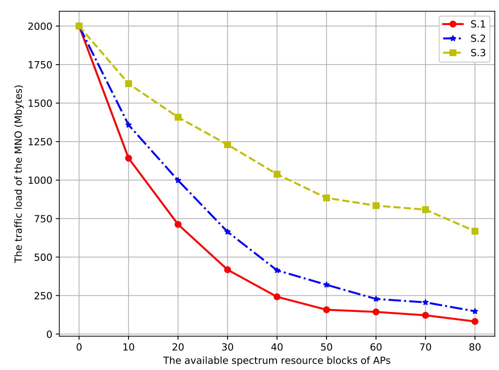


Fig. 7. The traffic load of the MNO with different available spectrum resource blocks of APs.

in terms of the MNO's traffic load with the increase of the number of MUs. With the increase of the number of MUs, more traffic demands are requested by MUs, so the traffic load of the MNO will increase continuously. It can be found that the traffic load of the MNO in DPWSM increases slowly as the number of MUs increases, compared with that of GWSM and the Random Winner Selection Method, which demonstrates that DPWSM performs best. The traffic load of the MNO in the Random Winner Selection Method increases very fast, which demonstrates that the Random Winner Selection Method still performs worst.

3) Different available spectrum resource blocks of APs: This part compares the performance of our proposed methods, DPWSM (S.1) and GWSM (S.2) with the Random Winner Selection Method (S.3) in terms of the MNO's utility and traffic load, when the available spectrum resource blocks of APs are different. We set the number of APs as $|K| = 30$, and the number of MUs as $|N| = 100$.

Fig. 6 shows the performance comparison of DPWSM (S.1), GWSM (S.2), and the Random Winner Selection Method (S.3) in terms of the MNO's utility with the increase of the available spectrum resource blocks of APs. With the increase of the available spectrum resource blocks of APs, more MUs can be served by APs, so the MNO's utility will increase continuously. It can be found that the utility of the MNO in DPWSM increases significantly as the available spectrum resource blocks of APs increase, compared with that of GWSM and the Random Winner Selection Method, which demonstrates that DPWSM performs best. When the available spectrum resource blocks of APs increase to a large value, i.e, 50 MHz, the MNO's utilities of our proposed methods increase very slowly, and are very close. This is because when the spectrum resources of the APs are sufficiently enough, each AP can meet the spectrum resource requirements of all MUs in its coverage area. Then, even if dynamic programming is not used to obtain the optimal AP-MU association set in each AP's coverage area, the performance difference of DPWSM and GWSM will not be obvious. Furthermore, the gaps between the Random Winner Selection Method and the other two curves indicate that the Random Winner Selection Method still performs worst.

Fig. 7 shows the performance comparison of DPWSM (S.1), GWSM (S.2), and the Random Winner Selection Method (S.3) in terms of the MNO's traffic load with the increase of the available spectrum resource blocks of APs. With the increase of the available spectrum resource blocks of APs, more MUs can be served by APs, so the MNO's traffic load will decrease continuously. Similarly, it can be found that the MNO's traffic load of DPWSM decreases significantly as the available spectrum resource blocks of APs increase, compared with that of GWSM and the Random Winner Selection Method, which demonstrates that DPWSM performs best. The traffic load of the MNO in the Random Winner Selection Method is the largest, which demonstrates that the Random Winner Selection Method still performs worst.

To summarize, we demonstrate that our proposed DPWSM performs best in terms of the MNO's utility and traffic load, and the proposed GWSM performs much better than the Random Winner Selection Method in terms of the MNO's utility and traffic load under different scenarios.

VII. Conclusion

In this paper, we proposed a delay-constrained and reverse auction-based incentive mechanism for offloading cellular traffic through WiFi APs. We formulate the data offloading problem as an optimization problem, and design an effective reverse auction-based incentive mechanism to stimulate WiFi APs to participate in the data offloading process, with consideration of the delay constraint of different applications. The optimization problem is modeled as a nonlinear integer programming problem, and two low-complexity methods: GWSM and DPWSM are proposed to solve the problem. Furthermore, an innovative standard Vickrey-Clarke-Groves (VCG) scheme-based payment rule is proposed to guarantee the individual rationality and truthfulness properties of DPWSM. Extensive simulation results illustrate that DPWSM outperforms the other two methods in terms of the MNO's utility and traffic load under different scenarios. In the future, we plan to consider some realistic mobility models, and analyze the impact of pricing strategy at different time periods on the utility of the MNO.

References

[1] G. Mao, Z. Zhang, and B. Anderson, "Cooperative content dissemination and offloading in heterogeneous mobile networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6573-6587, 2016.

[2] H. Zhou, J. Chen, J. Fan, Y. Du, and S. K. Das, "ConSub: Incentive-Based Content Subscribing in Selfish Opportunistic Mobile Networks," *IEEE Journal on Selected Areas in Communications*, vol.31, no. 9, pp. 669-679, 2013.

[3] M. Wu, S. He, Y. Zhang, J. Chen, Y. Sun, Y. Liu, J. Zhang, and H. Vincent Poor, "A Tensor-Based Framework for Studying Eigenvector Multicentrality in Multilayer Networks," *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, vol. 116, no. 31, pp. 15407-15413, 2019.

[4] B. Han, P. Hui, V. S. A. Kumar, M. Marathe, J. Shao, and A. Srinivasan, "Mobile data offloading through opportunistic communications and social participation," *IEEE Transactions on Mobile Computing*, vol. 11, no. 5, pp. 821-834, 2012.

[5] H. Zhou, H. Wang, X. Li, and V. C.M. Leung, "A Survey on Mobile Data Offloading Technologies, IEEE Access, vol. 6, no. 1, pp. 5101-5111, Jan. 2018.

[6] CISCO, Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016-2021, Tech. Rep., Feb. 2017.

[7] X. Chen, J. Wu, Y. Cai, H. Zhang, and T. Chen, "Energy-efficiency oriented traffic offloading in wireless networks: A brief survey and a learning approach for heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 4, pp. 627-640, 2015.

[8] P. Kortoci, L. Zheng, C. Joe-Wong, M. Francesco, and M. Chiang, "Fog-based Data Offloading in Urban IoT Scenarios," *in Proc. IEEE INFOCOM*, pp. 784-792, 2019.

[9] H. Zhou, H. Wang, X. Chen, X. Li, and S. Xu, "Data Offloading Techniques through Vehicular Ad Hoc Networks: A Survey," *IEEE Access*, vol.6, no. 1, pp. 65250-65259, 2018.

[10] F. Rebecchi, M. D. Amorim, V. Conan, A. Passarella, R. Bruno, and M. Conti, "Data offloading techniques in cellular networks: a survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 580-603, 2015.

[11] H. Zhou, V. C. M. Leung, C. Zhu, S. Xu, and J. Fan, "Predicting Temporal Social Contact Patterns for Data Forwarding in Opportunistic Mobile Networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10372-10383, 2017.

[12] M. Ruan, X. Chen, and H. Zhou, "Centrality Prediction Based on K-order Markov Chain in Mobile Social Networks," *Peer-to-Peer Networking and Applications*, vol. 12, no. 6, pp. 1662-1672, 2019.

[13] Y. Cao, T. Jiang, O. Kaiwartya, H. Sun, H. Zhou, and R. Wang, "Toward Preempted EV Charging Recommendation Through V2V Based Reservation System," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. PP, no. 99, pp. 1-1, 2020.

[14] H. Yu, M. Cheung, and J. Huang, "Cooperative Wi-Fi Deployment: A One-to-Many Bargaining Framework," *IEEE Trans. Mobile Comput.*, vol. 16, no. 6, pp. 1559-1572, June 2017.

[15] Juniper Research, Mobile Data Offload & Onload, Tech. Rep., December 2016.

[16] C. B. Sankaran, "Data offloading techniques in 3GPP Rel-10 networks: A tutorial," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 46-53, June 2012.

[17] M. H. Qutqut, F. M. Al-Turjman, and H. S. Hassanein, "MFW: Mobile femtocells utilizing WiFi: A data offloading framework for cellular networks using mobile femtocells," *in Proc. IEEE ICC*, pp. 6427-6431, Jun. 2013.

[18] J. Du, E. Gelenbe, C. Jiang, Z. Han, and Y. Ren, "Auction-based Data Transaction in Mobile Networks: Data Allocation Design and Performance Analysis," *IEEE Trans. Mobile Comput.*, vol. PP, no. 99, pp. 1-1, 2019.

[19] H. Shah-Mansouri, V. Wong, and J. Huang, "An Incentive Framework for Mobile Data Offloading Market Under Price Competition," *IEEE Trans. on Mobile Comput.*, vol. 16, no. 11, pp. 2983-2999, Nov. 2017.

[20] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong, "Mobile Data Offloading: How Much Can WiFi Deliver?" *IEEE/ACM Trans. on Networking*, vol. 21, no. 2, pp. 536-550, Apr. 2013.

[21] A. Y. Ding, B. Han, Y. Xiao, P. Hui, A. Srinivasan, M. Kojo, and S. Tarkoma, "Enabling Energy-Aware Collaborative Mobile Data Offloading for Smartphones," *in Proc. IEEE SECON*, pp. 487-495, Jun. 2013.

[22] S. I. Sou, and Y. T. Peng, "Performance Modeling for Multipath Mobile Data Offloading in Cellular/Wi-Fi Networks," *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3863-3875, Sept. 2017.

[23] Z. Zheng, P. Sinha, and S. Kumar, "Sparse WiFi deployment for vehicular internet access with bounded interconnection gap," *IEEE/ACM Trans. on Netw.*, vol. 20, no. 3, pp. 956-969, Jun. 2012.

[24] E. Bulut, and B. Szymanski, "Wifi access point deployment for efficient mobile data offloading" *ACM SIGMOBILE Mobile Computing andCommunications Review*, vol. 17, no. 1, pp. 71-78, 2013.

[25] L. Gao, G. Iosifidisy, J. Huang, and L. Tassiulasy, "Economics of Mobile Data Offloading," *in Proc. Infocom workshop on SDP*, Apr. 2013.

[26] L. Gao, G. Iosifidisy, J. Huang, L. Tassiulasy, and D. Li, "Bargaining-Based Mobile Data Offloading," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1114-1125, Jun. 2014.

[27] X. Kang, Y. Chia, S. Sun, and H. Chong, "Mobile Data Offloading through A Third-Party WiFi Access Point: An Operator's Perspective," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 10, pp. 5340-5351, 2014.

[28] J. Lee, Y. Yi, S. Chong, and Y. Jin, "Economics of wifi offloading: Trading delay for cellular capacity," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1540-1554, 2014.

[29] X. Zhou, and H. Zheng, "TRUST: a general framework for truthful double spectrum auctions," *in Proc. IEEE INFOCOM*, pp. 999-1007, Apr. 2009.

[30] D. Yang, X. Fang, and G. Xue, "Truthful auction for cooperative communications with revenue maximization," *in Proc. IEEE ICC*, pp. 4888-4892, Jun. 2012.

[31] Z. Chen, H. Huang, Y. Sun, and L. Huang, "True-MCSA: a framework for truthful double multi-channel spectrum auctions," *IEEE Trans. Wirel. Commun.*, vol. 12, no. 8, pp. 3838-3850, Aug. 2013.

[32] H. Huang, Y. Sun, X. Li, S. Chen, M. Xiao, and L. Huang, "Truthful auction mechanisms with performance guarantee in secondary spectrum markets," *IEEE Trans. Mobile Comput.*, vol. 14, no. 6, pp. 1315-1329, Jun. 2015.

[33] Q. Wang, B. Ye, T. Xu, S. Lu, and S. Guo, "Approximately truthful mechanisms for radio spectrum allocation," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2615-2626, Jun. 2015.

[34] Y. Wu, B. Wang, K.J.R. Liu, and T. C. Clancy, "A scalable collusion-resistant multi-winner cognitive spectrum auction game," *IEEE Trans. Wirel. Commun.*, vol. 57, no. 12, pp. 3805-3816, Dec. 2009.

[35] X. Zhuo, W. Gao, G. Cao, and Y. Dai, "Win-Coupon: an incentive framework for 3 G traffic offloading," *in Proc. IEEE ICNP*, pp. 206-215, Oct. 2011.

[36] S. Paris, F. Martignon, I. Filippini, and L. Chen, "An efficient auction-based mechanism for mobile data offloading," *IEEE Trans. Mobile Comput.*, vol.14, no. 8, pp. 1573-1586, Aug. 2015.

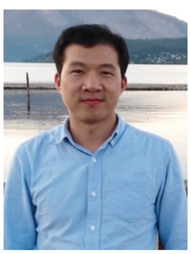
[37] W. Dong, S. Rallapalli, R. Jana, L. Qiu, K. Ramakrishnan, L. Razoumov, Y. Zhang, and T. Cho, "iDEAL: incentivized dynamic cellular offloading via auctions," *IEEE/ACM Trans. Networking*, vol.22, no. 4, pp. 1271-1284, Aug. 2014.

[38] Z. Lu, P. Sinha, and R. Srikant, "EasyBid: enabling cellular offloading via small players," *in Proc. IEEE INFOCOM*, pp. 691-699, May 2014.

[39] W. Song, and Y. Zhao, "A Randomized Reverse Auction for Cost-Constrained D2D Content Distribution," *in Proc. IEEE Globecom*, 2017.

[40] F. Hou, and Z. Xie, "Social-Aware Incentive Mechanism for AP Based Mobile Data Offloading," *IEEE Access*, vol. 6, no. 1, pp. 49408-49417, 2018.

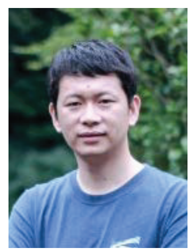
[41] F. Zhao, X. Xu, and H. Chen, "Reverse spectrum auction algorithm for cellular network offloading," *Ad Hoc Networks*, vol. 58, no. 2017, pp. 278-285, 2017.



Huan Zhou (M'14) received his Ph. D. degree from the Department of Control Science and Engineering at Zhejiang University. He was a visiting scholar at the Temple University from Nov. 2012 to May, 2013, and a CSC supported postdoc fellow at the University of British Columbia from Nov. 2016 to Nov. 2017. Currently, he is a full professor in the College of Computer and Information Technology, China Three Gorges University. He was a Lead Guest Editor of Pervasive and Mobile Computing, and Special Session Chair of the 3rd International Conference on Internet of Vehicles (IOV 2016), and TPC member of IEEE WCSP'13'14, CCNC'14'15, ICNC'14'15, ANT'15'16, IEEE Globecom'17'18, ICC'18'19, etc. He has published more than 50 research papers in some international journals and conferences, including IEEE JSAC, TPDS, TVT and so on. His research interests include Mobile Social Networks, VANETs, Opportunistic Mobile Networks, and mobile data offloading. He receives the Best Paper Award of I-SPAN 2014 and I-SPAN 2018, and is currently serving as an associate editor for IEEE ACCESS and EURASIP Journal on Wireless Communications and Networking.

Xin Chen received his B.S. Degree from China Three Gorges University. Currently, he is a graduate student at the College of Computer Information and Technology, China Three Gorges University. His main research interests are mobile data offloading and VANETs.

Shibo He (M'13-SM'19) received the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2012. He is currently a Professor with Zhejiang University. He was an Associate Research Scientist from March 2014 to May 2014, and a Postdoctoral Scholar from May 2012 to February 2014, with Arizona State University, Tempe, AZ, USA. From November 2010 to November 2011, he was a Visiting Scholar with the University of Waterloo, Waterloo, ON, Canada. His research interests include wireless sensor networks, crowdsensing, big data analysis, etc. Dr. He serves on the editorial board for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, Springer Peer-to-Peer Networking and Application and KSII Transactions on Internet and Information Systems, and is a Guest Editor for Elsevier Computer Communications and Hindawi nternational Journal of Distributed Sensor Networks. He was a Symposium Co-Chair for the IEEE ICC 2017, a Finance and Registration chair for ACM MobiHoc 2015, a TPC Co-Chair for the IEEE ScalCom 2014, a TPC Vice Co-Chair for ANT 2013ÍC2014, a Track Co-Chair for the Pervasive Algorithms, Protocols, and Networks of EUSPN 2013, a Web Co-Chair for the IEEE MASS 2013, and a Publicity Co-Chair of IEEE WiSARN 2010, and FCN 2014.



Jiming Chen is a Professor with College of control science and engineering, Zhejiang University. He was a visiting researcher at University of Waterloo from 2008 to 2010. Currently, He serves/served associate editors for ACM TECS, IEEE TPDS, IEEE Network, IEEE TCNS, IEEE TII, etc. He has been appointed as a distinguished lecturer of IEEE vehicular technology society since 2015. He also was the recipients of IEEE INFOCOME 2014 Best Demo Award, IEEE ICCC 2014 best paper award, IEEE PIMRC 2012 best paper award, and JSPS Visiting Fellowship 2011. He also received the IEEE Comsoc Asia-pacific Outstanding Young Researcher Award 2011. He is a Distinguished Lecturer of IEEE Vehicular Technology Society (2015-2018), and a Fellow of IEEE. His research interests include networked control, sensor networks, cyber security, IoT.

Jie Wu is the Director of the Center for Networked Computing and Laura H. Carnell professor at Temple University. He also serves as the Director of International Affairs at College of Science and Technology. He served as Chair of Department of Computer and Information Sciences from the summer of 2009 to the summer of 2016 and Associate Vice Provost for International Affairs from the fall of 2015 to the summer of 2017. Prior to joining Temple University, he was a program director at the National Science Foundation and was a distinguished professor at Florida Atlantic University. His current research interests include mobile computing and wireless networks, routing protocols, cloud and green computing, network trust and security, and social network applications. Dr. Wu regularly publishes in scholarly journals, conference proceedings, and books. He serves on several editorial boards, including IEEE Transactions on Services Computing and the Journal of Parallel and Distributed Computing. Dr. Wu was general co-chair for IEEE MASS 2006, IEEE IPDPS 2008, IEEE ICDCS 2013, ACM MobiHoc 2014, IEEE ICPP 2016, and IEEE CNS 2016, as well as program co-chair for IEEE INFOCOM 2011 and CCF CNCC 2013. He was an IEEE Computer Society Distinguished Visitor, ACM Distinguished Speaker, and chair for IEEE Technical Committee on Distributed Processing (TCDP). Dr. Wu is a CCF Distinguished Speaker and a Fellow of the IEEE. He is the recipient of the 2011 China Computer Federation (CCF) Overseas Outstanding Achievement Award.