IoT-U: Cellular Internet-of-Things Networks over Unlicensed Spectrum

Hongliang Zhang, *Student Member, IEEE*, Boya Di, *Student Member, IEEE*,
Kaigui Bian, *Member, IEEE*, and Lingyang Song, *Fellow, IEEE*

Abstract—In this paper, we consider an uplink cellular Internet-of-Things (IoT) network, where a cellular user (CU) can serve as the mobile data aggregator for a cluster of IoT devices. To be specific, the IoT devices can either transmit the sensory data to the base station (BS) directly by cellular communications, or first aggregate the data to a CU through Machine-to-Machine (M2M) communications before the CU uploads the aggregated data to the BS. To support massive connections, the IoT devices can leverage the unlicensed spectrum for M2M communications, referred to as IoT Unlicensed (IoT-U). Aiming to maximize the number of scheduled IoT devices and meanwhile associate each IoT devices with the right CU or BS with the minimum transmit power, we first introduce a single-stage formulation that captures these objectives simultaneously. To tackle the NP-hard problem efficiently, we decouple the problem into two subproblems, which are solved by successive linear programming and convex optimization techniques, respectively. Simulation results show that the proposed IoT-U scheme can support more IoT devices than that only using the licensed spectrum.

Index Terms—Internet-of-Things Unlicensed, Carrier aggregation, Machine-to-Machine Communication, Non-convex optimization

I. INTRODUCTION

We are rapidly connecting machines and other physical objects to Internet-enabled networks at unprecedented rates, accelerating towards the Internet-of-things (IoT). It is envisioned that billions of IoT devices will be connected by 2020 with transformative economic potentials for operators and stakeholders, reaching trillions of dollars [1]. One typical feature of the IoT network is the massive connectivity, which refers that a large number of IoT devices have accessed to the existing wireless networks. This feature has unlocked a various of opportunities to have applications across a variety of vertical sectors, such as city sensing [2], traffic control [3], smart home [4], heath care [5], and precision agriculture [6].

To support such a large number of IoT devices, the 3rd Generation Partnership Project (3GPP) has led the research on enabling IoT communications over the cellular systems. Its recent solutions, e.g., Narrow Band IoT (NB-IoT), provide Quality-of-Service (QoS) guaranteed connections and introduce power saving modes to improve battery life [7].

However, the licensed spectrum may not be sufficient for the mass deployments of IoT devices. To support more IoT devices, one promising solution is to extend the cellular IoT communications to the unlicensed spectrum using the LTE-Unlicensed (LTE-U) technology [8], [9], especially in the hotspot areas, which we refer to as IoT Unlicensed (IoT-U). Specifically, the characteristics of the IoT-U systems lie in that the unlicensed carriers are integrated with the licensed ones for data transmission of IoT devices by the existing carrier aggregation (CA) technology [10]. Besides, unlike traditional Machine-to-Machine (M2M) communications [11], the cellular users (CUs) in the IoT-U network can work as mobile data aggregators to further enhance the IoT connectivity, which can aggregate the data from IoT devices.

In this paper, we consider an uplink IoT-U network[1], where a CU can serve as the mobile data aggregator for a cluster of IoT devices. Specifically, the CU can collect sensory data from IoT devices through M2M communications, and aggregate the data to the base station (BS) via cellular communications. To further support more IoT devices, M2M communications can work as the underlay of cellular communications and share the unlicensed spectrum with the Wi-Fi systems. Different from the long-range communication techniques in unlicensed bands for IoT networks such as LoRa [13], which builds the network upon the IEEE 802.15.4 infrastructure with a mesh topology [14], IoT-U system requires the assist and control from the central BS.

However, the spectrum utilization brings new challenges to the scheduling of IoT devices in the IoT-U network. First, a suitable coexistence mechanism of the IoT-U and Wi-Fi systems is required due to the opportunistic feature of unlicensed channel access [15]. Second, the interference management among CUs, IoT devices and Wi-Fi users (WUs) becomes more complicated. To tackle the first challenge, we utilize a duty cycle [16] based protocol to share the unlicensed spectrum fairly. To cope with the second one, we optimize the association, scheduling, and power allocation to maximize the weighted scheduled number of IoT devices with the lowest power consumption. This problem is a mixed-integer non-linear programming (MINLP) problem, which is generally NP-hard. To solve this problem efficiently, we decouple it into two subproblems, i.e., IoT devices association and scheduling

The authors are with National Engineering Laboratory for Big Data Analysis and Applications, School of Electronics Engineering and Computer Science, Peking University, Beijing, China (email: hongliang.zhang@pku.edu.cn; diboya@pku.edu.cn; bkg@pku.edu.cn; lingyang.song@pku.edu.cn).

[1]A typical application for the IoT-U networks is smart home [12]. In the application, the sensory data generated by the IoT devices at home can be aggregated to a CU of the resident by M2M communications over both licensed and unlicensed spectrum. Then, the aggregated data is forwarded to the server by cellular communications for further data processing.

subproblem, and power allocation subproblem. For the first subproblem, we convert the non-linear constraints into linear ones and solved the transformed problem by the branch-and-bound algorithm [17]. For the second subproblem, we approximate the non-convex functions into a series of convex ones by successive convex approximation (SCA) [18] and solve them by existing convex techniques [19].

In literature, various techniques have been discussed for the spectrum sharing in cellular networks, such as cognitive radio [20], [21], Wi-Fi offloading [22], [23], LTE-U [24]–[26], and LTE-Licensed Assisted Access (LTE-LAA) [27], [28]. In [20], [21], a buffer-aided cognitive M2M communication network was considered, where the M2M communications share the spectrum with the cellular network in a cognitive manner. In [22], [23], Wi-Fi access points (APs) was deployed for cellular networks to offload the mobile data traffic to the unlicensed spectrum. Recently, LTE-U and LTE-LAA have been proposed for the coexistence of cellular networks and Wi-Fi networks by CA. LTE-U works on the duty cycle based protocol and can only be used in China, USA, South Korea and India [24]. In [25], an almost blank subframe scheme was proposed to mitigate the interference in the LTE-U network. The authors in [26] formulated a joint user association and power allocation problem for licensed and unlicensed spectrum to maximize the LTE-U system sum-rate subjected to the QoS constraints. For worldwide deployment, LTE-LAA uses the Listen-Before-Talk (LBT) and is the modified version of LTE-U [27]. In [28], the authors proposed a novel Markov chain-based analytic model to cope with the variation of the LTE-LAA frame structure. Different from the existing schemes that focused on either M2M or cellular links, our proposed IoT-U scheme considers the network assisted data aggregation on both the licensed and unlicensed spectrum, thereby utilizing the spectrum more efficiently.

Therefore, the major contributions of this paper are summarized as follows.

1) We consider an uplink IoT-U network to support more IoT devices, where the CUs can aggregate sensory data from IoT devices over both licensed and unlicensed spectrum, and utilize a feasible duty cycle based protocol to facilitate the fair unlicensed spectrum between the IoT-U and Wi-Fi systems.
2) We maximize the weighted scheduled number of IoT devices in each cycle by jointly optimizing the IoT device association, scheduling, and power allocation. To solve the problem efficiently, we decouple the problem into IoT devices association and scheduling as well as power allocation subproblems, and propose an iterative algorithm to solve these subproblems.
3) Simulation results show that our proposed IoT-U scheme can achieve a better performance than the IoT scheme without unlicensed band, decoupled unlicensed scheme, and the matching based unlicensed scheme.

The rest of the paper is organized as follows. In Section II, we first introduce the system model for the uplink IoT-U network. The spectrum sharing mechanism for IoT-U and Wi-Fi systems is introduced in Section III together with the
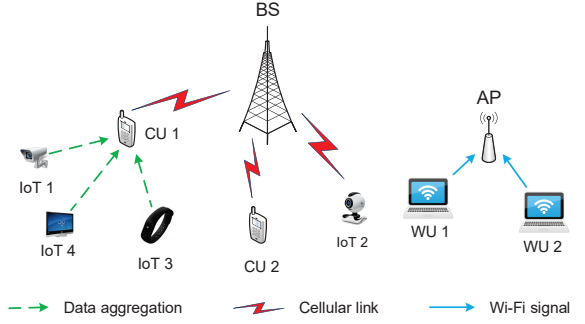

Fig. 1.   System model for the uplink IoT-U network.

interference analysis. In Section IV, we introduce the problem formulation and decouple the problem into two subproblems: IoT devices association and scheduling problem, and power allocation subproblem. Section V presents an iterative algorithm to solve these two subproblems. In Section VI, the system performance is discussed. Numerical results in Section VII evaluate the performance of our proposed IoT-U scheme. Finally, conclusions are drawn in Section VIII.

II. System Model

As shown in Fig. 1, we consider an uplink IoT-U network consisting of one BS, N CUs, denoted by $\mathcal{N} = \{1, \ldots, N\}$, and M IoT devices to collect sensory data, denoted by $\mathcal{M} = \{1, \ldots, M\}$. Among these CUs, there are Q idle CUs which have no data to transmit and can serve as aggregators to aggregate the data generated from IoT devices, denoted by $\mathcal{Q} = \{1, \ldots, Q\}$. To upload the sensory data efficiently, there exist two communication modes as described below:

- Cellular mode: The IoT devices can transmit their sensory data to the BS directly by cellular communications.
- Aggregation mode: An idle CU can aggregate the sensory data through M2M communications[2] before the CU uploads the aggregated data to the BS.

In Fig. 1, for example, the collected data of IoT device 2 is sent to BS directly, while IoT devices 1, 3, and 4 form a cluster and their sensory data is sent to the BS by data aggregation to CU 1. The network owns K licensed subchannels to support orthogonal frequency division multiple access (OFDMA) transmissions, denoted by $\mathcal{K} = \{1, \ldots, K\}$. Different cellular links need to utilize orthogonal channels to avoid mutual interference.

For the Wi-Fi system, we assume that there exist F active WUs marked by $\mathcal{F} = \{1, \ldots, F\}$. Besides, we assume that there are L unlicensed channels to support these WUs, e.g., there are 23 channels for IEEE 802.11n in the 5GHz band. In each cluster, the M2M links among IoT devices and the CUs can utilize both licensed and unlicensed spectrum via CA. Since the bandwidth of an unlicensed channel is much wider than a licensed subchannel, to utilize the unlicensed channel more efficiently, we divide the unlicensed channel

---

[2]The relaying network has been investigated in 3GPP to extend the coverage of IoT devices [29]. The relaying CU and IoT devices belong to the same group, for example, they subscribe to the same mobile operator, and thus, the CU is willing to relay the data traffic for IoT devices.
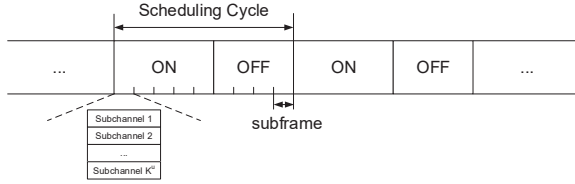
Fig. 2. Duty cycle based spectrum sharing mechanism for IoT devices in the unlicensed spectrum.

into K^u subchannels with the same bandwidth as the licensed one, denoted by $\mathcal{K}^u = \{1, \ldots, K^u\}$ [8]. To make full use of the licensed band, we assume that the M2M links can share the licensed band with the cellular links. Besides, we also assume that the CUs can upload the data to the BS and aggregate the data from IoT devices simultaneously on different subchannels.

We define $\boldsymbol{P}_{(M \times (N+1)+N) \times T \times (K+K^u)} = \left(\boldsymbol{P}^C_{N \times T \times (K+K^u)}, \boldsymbol{P}^I_{M \times (N+1) \times T \times (K+K^u)} \right)$ as the transmit power matrix, where $\boldsymbol{P}^C_{N \times T \times (K+K^u)} = [p_n^{k,t}]$ implies the transmit power of the link between CU n and the BS over subchannel k in subframe t, and $\boldsymbol{P}^I_{M \times (N+1) \times T \times (K+K^u)} = [p_{m,n}^{k,t}]$ implies the transmit power of the link between IoT device m and CU n over subchannel k in subframe t. Here, $n = 0$ if the IoT device m transmits to the BS directly. We also denote the total transmit powers of IoT device m and CU n by P^{iot} and P^c, respectively. The transmit power of WUs on the whole unlicensed channel is fixed as P^w. Therefore, we have

$$\begin{aligned} \sum_{k \in \mathcal{K}} p_{m,0}^{k,t} + \sum_{k \in \mathcal{K} \cup \mathcal{K}^u} \sum_{n \in \mathcal{N}} p_{m,n}^{k,t} &\leq P^{iot}, \\ \sum_{k \in \mathcal{K}} p_n^{k,t} &\leq P^c. \end{aligned} \quad (1)$$

The free space propagation path-loss model with Rayleigh fading is adopted to model the channel gain between two nodes in the network, i.e., for the link from nodes i to j over subchannel k, the received power can be expressed as

$$p_{i,j}^k = p_i^k \cdot |h_{i,j}^k|^2 = p_i^k \cdot G \cdot d_{i,j}^{-\alpha} \cdot |h_0^k|^2, \quad (2)$$

where p_i^k represents the transmit power of user i over subchannel k, $d_{i,j}$ is the distance between nodes i and j, α is the pathloss exponent, G is the constant power gains factor introduced by amplifier and antenna, and $h_0^k \sim \mathcal{CN}(0, 1)$ is a complex Gaussian variable representing Rayleigh fading. In addition, we assume that the thermal noise at each user satisfies independent Gaussian distribution with zero mean and the same variance σ^2.

III. SPECTRUM SHARING MECHANISM AND INTERFERENCE ANALYSIS

In this section, we first introduce the spectrum sharing mechanism for IoT-U and Wi-Fi systems. Then, the interference issues over licensed and unlicensed spectrum are discussed, respectively.

A. Unlicensed Spectrum Sharing Mechanism

In the unlicensed spectrum, the IoT devices are allowed to share channels with WUs. To share the unlicensed spectrum fairly, a duty cycle method is adopted to support the coexistence of the IoT-U and Wi-Fi systems. Specifically, as shown in Fig. 2, the unlicensed band is slotted to subframes whose time length is the same as that in the cellular system for the sake of synchronization. To schedule the IoT devices efficiently, we define the concept of scheduling cycle containing T subframes, denoted by $\mathcal{T} = \{1, \ldots, T\}$, in which each IoT device will be scheduled at most once.

At the beginning of a scheduling cycle, the BS will select a channel with the least interference first according to the channel measurement by the IoT devices. The IoT devices will scan the unlicensed spectrum and identify a channel with the least interference from the set of available unlicensed channels for the uplink transmission. Specifically, the IoT devices will perform energy detection, and measure the interference level. If the interference is sensed less than the predefined threshold for a sensing duration, the channel will be regarded as idle for this IoT device. Then, these IoT devices will inform the BS whether they collide with WUs according to the measured result. If in the operating channel, the number of interfered users is larger than a given threshold, and there is another available channel with the less interference, the transmission will be switched to the new channel.

The scheduling cycle consists of two transmission modes, i.e., ON and OFF modes. In the ON mode, the IoT devices are enabled to transmit on unlicensed spectrum. The time period for the ON mode contains T_{ON} subframes, denoted by $\mathcal{T}_{ON} = \{1, \ldots, T_{ON}\}$, and each of them is divided into K^u subchannels to support the concurrent transmissions for different IoT devices. However, in the OFF mode, the IoT devices need to be silent in the unlicensed channel, which is handed over to the Wi-Fi system. When the reserved period for the Wi-Fi system expires, the IoT devices are activated in the unlicensed subchannels. In this way, the IoT-U system coexists with the Wi-Fi system in a time division multiplexing (TDM) manner. The duty cycle percentage T_{ON}/T captures the effect of IoT-U transmission on the Wi-Fi system, and should be carefully designed to provide fair unlicensed resource sharing.

To evaluate the signaling overhead for the spectrum sharing mechanism, we assume that α messages are required to inform the BS of the channel state sensed by an IoT device, β messages are required for an IoT device/CU to report its location and subchannel estimation results over a subchannel, and x messages are necessary for the BS to notify a user about the allocated subchannels. At the beginning of a scheduling cycle, each IoT device needs to report the sensing result over each channel, and thus, at most αML messages are required. And before each subframe, each IoT device/CU needs to report their locations and the subchannel estimation results for subchannel allocation, which requires $\beta(M + N)(T - T_{ON})K + \beta T_{ON}(M(K + K^u) + NK)$ messages for a cycle. Then, the BS will perform resource allocation process with extra information, and notify each user by sending at most $x(M + N)T$ messages. Note that

in one duty cycle, each IoT device only performs one energy detection over one channel. Thus, the signaling cost is under a tolerable level. In addition, the signaling cost of resource allocation is positively proportional to the number of IoT devices and CUs as well as the number of available unlicensed subchannels. When the number of IoT devices is numerous, we can also reduce the signaling overhead by constraining the number of available unlicensed subchannels for each IoT device.

B. Interference Analysis

In this part, we will discuss the interference over licensed and unlicensed spectrum.

1) Transmission Model over Licensed Band: In licensed spectrum, cellular and M2M links are able to occupy multiple subchannels for transmission, and a subchannel can be allocated to one cellular link. To guarantee reliable transmission of the control signaling, a M2M link is required to hold at least one licensed subchannel [10]. Besides, M2M links can work as an underlay of the cellular links. In other words, M2M links can utilize the licensed subchannels concurrently occupied by some cellular links.

Under these assumptions, the cellular links can only receive the co-channel interference from the underlaid M2M links, while the interference received by M2M links might be from cellular links and other co-channel M2M links. Define the set of IoT devices which upload sensory data by cellular communications as $\mathcal{S}^C$ and the set of IoT devices which transmit sensory data by M2M communications as $\mathcal{S}^M$, respectively. In addition, we define the IoT device association and subchannel allocation matrix for licensed band $\boldsymbol{A}_{(M \times (Q+1)+N) \times T \times K} = \left(\boldsymbol{\Phi}_{N \times T \times K}, \boldsymbol{\Theta}_{M \times (Q+1) \times T \times K} \right)$, where $\boldsymbol{\Phi}_{N \times T \times K} = [\phi_{n,t}^k]$, and $\boldsymbol{\Theta}_{M \times (Q+1) \times T \times K} = [\theta_{m,n}^{k,t}]$ stand for the subchannel allocation matrices for CUs and IoT devices, respectively. The values of $\phi_n^{k,t}$ and $\theta_{m,n}^{k,t}$ are defined as

$$\phi_n^{k,t} = \begin{cases} 1, & \text{when licensed subchannel } k \text{ is allocated to} \\ & \text{CU } n \text{ in subframe } t, \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

and

$$\theta_{m,n}^{k,t} = \begin{cases} 1, & \text{when licensed subchannel } k \text{ is allocated to} \\ & \text{link } m-n \text{ in subframe } t, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Using the aforementioned notations, the SINR at the receiver of BS from CU n over licensed subchannel k in subframe t can be given by

$$\gamma_n^{k,t} = \frac{p_n^{k,t} \phi_n^{k,t} |h_n^{k,t}|^2}{\sigma^2 + \sum\limits_{m \in \mathcal{S}^M} \sum\limits_{\substack{n' \in \mathcal{Q}, \\ n' \neq n}} p_{m,n'}^{k,t} \theta_{m,n'}^{k,t} |h_{m,0}^{k,t}|^2}, \tag{5}$$

where $h_n^{k,t}$ and $h_{m,0}^{k,t}$ represent the channel gains from CU n and IoT device m to the BS over subchannel k in subframe t, respectively. Likewise, the SINR at the receiver of the BS from

IoT device $m \in \mathcal{S}^C$ over licensed subchannel k in subframe t can be expressed as

$$\gamma_{m,0}^{k,t} = \frac{p_{m,0}^{k,t} \theta_{m,0}^{k,t} |h_{m,0}^k|^2}{\sigma^2 + \sum\limits_{m' \in \mathcal{S}^M} \sum\limits_{n \in \mathcal{Q}} p_{m',n}^{k,t} \theta_{m',n}^{k,t} |h_{m',0}^{k,t}|^2}. \tag{6}$$

In addition, the SINR at the associated CU $n \in \mathcal{Q}$ from IoT $m \in \mathcal{S}^M$ over licensed subchannel k in subframe t can be given by

$$\gamma_{m,n}^{k,t} = \frac{p_{m,n}^{k,t} \theta_{m,n}^{k,t} |h_{m,n}^{k,t}|^2}{\sigma^2 + I_{m,n}^{k,t}}, \tag{7}$$

where

$$I_{m,n}^{k,t} = \sum_{\substack{n' \in \mathcal{N}, \\ n' \neq n}} p_{n'}^{k,t} \phi_{n'}^{k,t} |h_{n',n}^{k,t}|^2 + \sum_{\substack{m' \in \mathcal{M}, \, n' \in \mathcal{Q} \cup \{0\}, \\ m' \neq m \quad n' \neq n}} P_{m',n}^{k,t} \theta_{m',n'}^{k,t} |h_{m',n}^{k,t}|^2. \tag{8}$$

Here, $h_{m,n}^{k,t}$ and $h_{n',n}^{k,t}$ represent the channel gains from IoT device m and CU n' to CU n over subchannel k in subframe t, respectively. Therefore, the achievable rates of the CU n - BS, IoT device m - BS or associated CU n links over subchannel k in subframe t can be obtained by

$$R_n^{k,t} = \log_2(1 + \gamma_n^{k,t}), \quad R_{m,n}^{k,t} = \log_2(1 + \gamma_{m,n}^{k,t}), \tag{9}$$

respectively.

2) Transmission Model over Unlicensed Band: In the unlicensed spectrum, only IoT devices which set up M2M communications with CUs are allowed to transmit. Therefore, the interference for a M2M link only comes from other M2M links utilizing the same unlicensed subchannel. Define the subchannel allocation matrix for unlicensed band $\boldsymbol{B}_{M \times Q \times T \times K^u} = [\beta_{m,n}^{k,t}]$, where

$$\beta_{m,n}^{k,t} = \begin{cases} 1, & \text{when unlicensed subchannel } k \text{ is allocated to} \\ & \text{link } m-n \text{ in subframe } t, \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$

The SINR at the associated CU from IoT device m over unlicensed subchannel k in subframe t can be written as

$$\gamma_{m,n}^{k,t,u} = \frac{p_{m,n}^{k,t} \beta_{m,n}^{k,t} |h_{m,n}^{k,t,u}|^2}{\sigma^2 + \sum\limits_{\substack{m' \in \mathcal{M}, \, n' \in \mathcal{Q}, \\ m' \neq m \quad n' \neq n}} p_{m',n'}^{k,t} \beta_{m',n'}^{k,t} |h_{m',n}^{k,t,u}|^2 + I_{m,n}^{k,t}}, \tag{11}$$

where $h_{m,n}^{k,t,u}$ is channel gain from IoT device m to CU n over unlicensed subchannel k and I_m^k is the total interference from Wi-Fi system to the associated CU of IoT device m. Here, the interference can be calculated as

$$I_{m,n}^{k,t} = \sum_{f \in \mathcal{F}} \frac{P^w}{K^u} |h_{f,n}^{k,t}|^2, \tag{12}$$

where $h_{f,n}^{k,t}$ is the channel gain from the active WU f to CU n over subchannel k in subframe t. Thus, the data rate of IoT device m over unlicensed subchannel k in subframe t is given by

$$R_{m,n}^{k,t,u} = \log_2(1 + \gamma_{m,n}^{k,t,u}). \tag{13}$$

IV. Weighted Scheduled Number Maximization of IoT Devices

In this section, we aim to maximize the weighted scheduled number of IoT devices in each cycle by jointly optimizing the IoT device association, scheduling, and power allocation.

A. Constraints for Clusters

To guarantee the performance of the IoT-U system, several constraints in terms of the data rate, subchannel allocation, and scheduling are illustrated below.

In this paper, we assume that the CUs can aggregate data from the IoT devices and upload the data to the BS simultaneously. To avoid the mutual interference, a cellular link and the M2M links associated to the same CU need to utilize different licensed subchannels, i.e.,

$$\phi_n^{k,t} + \sum_{m \in \mathcal{M}} \theta_{m,n}^{k,t} \leq 1, \forall n \in \mathcal{Q}, k \in \mathcal{K}, t \in \mathcal{T}. \quad (14)$$

Besides, following the OFDMA manner, the IoT devices associated with the same CU cannot utilize the same unlicensed subchannel, and thus, we have

$$\sum_{m \in \mathcal{M}} \beta_{m,n}^{k,t} \leq 1, \forall n \in \mathcal{Q}, k \in \mathcal{K}^u, t \in \mathcal{T}. \quad (15)$$

Similarly, we assume that a licensed subchannel can be allocated to at most one cellular link. Therefore, the subchannel allocation matrix for the licensed spectrum satisfies

$$\sum_{n \in \mathcal{N}} \phi_n^{k,t} + \sum_{m \in \mathcal{M}} \theta_{m,0}^{k,t} \leq 1, \forall k \in \mathcal{K}, t \in \mathcal{T}. \quad (16)$$

Define a association matrix $\boldsymbol{F}_{M \times (Q+1) \times T} = [f_{m,n}^t]$ to indicate whether IoT device m associates with CU n in subframe t, where

$$f_{m,n}^t = \begin{cases} 1, & \text{when IoT device } m \text{ associates with CU } n \\ & \text{in subframe } t, \\ 0, & \text{otherwise}, \end{cases} \quad (17)$$

Thus, we have

$$1 - f_{m,n}^t = \prod_{k \in \mathcal{K}} (1 - \theta_{m,n}^{k,t}), \forall m \in \mathcal{M}, \forall n \in \mathcal{Q} \cup \{0\}, \forall t \in \mathcal{T}. \quad (18)$$

Likewise, define a scheduling matrix $\boldsymbol{C}_{M \times T} = [c_{m,t}]$ to indicate whether IoT device m is scheduled in subframe t, where

$$c_m^t = \begin{cases} 1, & \text{when IoT device } m \text{ is scheduled in subframe } t, \\ 0, & \text{otherwise}, \end{cases} \quad (19)$$

Note that a IoT device is assigned at least one licensed subchannel if it is scheduled. Therefore, we have

$$1 - c_{m,t} = \prod_{n \in \mathcal{Q} \cup \{0\}} (1 - f_{m,n}^t), \forall m \in \mathcal{M}, \forall t \in \mathcal{T}. \quad (20)$$

According to the requirement of CA, each M2M link needs to occupy at least one licensed subchannel for control signaling, and thus, we have

$$\sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{Q}} \theta_{m,n}^{k,t} \geq c_{m,t}, \forall m \in \mathcal{S}^M, t \in \mathcal{T}_{ON}. \quad (21)$$

Since an IoT device can be scheduled at most once in a cycle, we also have a scheduling constraint:

$$\sum_{t \in \mathcal{T}} c_{m,t} \leq 1, \forall m \in \mathcal{M}. \quad (22)$$

In addition, a scheduled IoT device can associate with at most one CU or BS, therefore, the association matrix needs to satisfy

$$\sum_{n \in \mathcal{Q} \cup \{0\}} f_{m,n}^t \leq c_{m,t}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T}. \quad (23)$$

Moreover, for each IoT device, the achievable data rate needs to exceed the generated data in order to upload the sensory data in a subframe. Denote the generated data of IoT device m by d_m, and therefore, for IoT devices in the cellular mode, we have

$$\sum_{k \in \mathcal{K}} R_{m,0}^{k,t} \geq d_m c_{m,t}, \forall m \in \mathcal{S}^C, t \in \mathcal{T}, \quad (24)$$

and for IoT devices in the aggregation mode, the following equations should be satisfied, which correspond to the subframes in the ON and OFF modes, respectively.

$$\sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{Q}} R_{m,n}^{k,t} + \sum_{k \in \mathcal{K}^u} \sum_{n \in \mathcal{Q}} R_{m,n}^{k,t,u} \geq d_m c_{m,t},$$
$$\forall m \in \mathcal{S}^M, t \in \mathcal{T}_{ON}, \quad (25)$$

$$\sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{Q}} R_{m,n}^{k,t} \geq d_m c_{m,t}, \forall m \in \mathcal{S}^M, t \in \mathcal{T} \backslash \mathcal{T}_{ON}. \quad (26)$$

In addition, to transmit the aggregated data to the BS successfully, the achievable data rate of the CU should be larger than the aggregated data in both the ON and OFF modes, i.e.,

$$\sum_{k \in \mathcal{K}} R_n^{k,t} \geq \sum_{m \in \mathcal{S}^M} \sum_{k \in \mathcal{K} \cup \mathcal{K}^u} R_{m,n}^{k,t}, \forall n \in \mathcal{Q}, t \in \mathcal{T}_{ON}, \quad (27)$$

$$\sum_{k \in \mathcal{K}} R_n^{k,t} \geq \sum_{m \in \mathcal{S}^M} \sum_{k \in \mathcal{K}} R_{m,n}^{k,t}, \forall n \in \mathcal{Q}, t \in \mathcal{T} \backslash \mathcal{T}_{ON}. \quad (28)$$

Define the minimum rate requirement for an active CU as R_{min}. To guarantee the QoS of active CUs, we have the constraint below:

$$\sum_{k \in \mathcal{K}} R_n^{k,t} \geq R_{min}, \forall n \in \mathcal{N} \backslash \mathcal{Q}, t \in \mathcal{T}. \quad (29)$$

B. Problem Formulation

To support the massive number of IoT devices, we aim to schedule more IoT devices in one cycle. However, those IoT devices with better channel conditions will be scheduled more frequently because they need fewer subchannels, and thus, more IoT devices can be scheduled. In this paper, we consider the user fairness in terms of scheduled opportunity, and introduce a weight factor w_m for IoT device m to adjust their scheduling priority. Based on the proportional fairness scheduler [30], the value of w_m is defined to be inversely proportional to the average scheduled times of IoT device m in the previous cycles. Therefore, the system performance can be evaluated by the weighted number of scheduled IoT devices

in a cycle. Then, the scheduling problem can be formulated as

$$\text{P1}: \max_{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{F},\boldsymbol{P}} \sum_{m\in\mathcal{M}} w_m \sum_{t\in\mathcal{T}} c_{m,t}, \tag{30a}$$

$$s.t. \sum_{k\in\mathcal{K}} p_{m,0}^{k,t} + \sum_{k\in\mathcal{K}\cup\mathcal{K}^u}\sum_{n'\in\mathcal{Q}} p_{m,n'}^{k,t} \le P^{iot}, \sum_{k\in\mathcal{K}} p_n^{k,t} \le P^c,$$
$$\forall m \in \mathcal{M}, n\in\mathcal{N}, t\in\mathcal{T}_{ON}, \tag{30b}$$

$$\sum_{k\in\mathcal{K}}\sum_{n'\in\mathcal{Q}\cup\{0\}} p_{m,n'}^{k,t} \le P^{iot}, \sum_{k\in\mathcal{K}} p_n^{k,t} \le P^c,$$
$$\forall m \in \mathcal{M}, n\in\mathcal{N}, t\in\mathcal{T}\backslash\mathcal{T}_{ON}, \tag{30c}$$

$$\phi_n^{k,t} + \sum_{m\in\mathcal{M}} \theta_{m,n}^{k,t} \le 1, \forall n\in\mathcal{Q}, k\in\mathcal{K}, t\in\mathcal{T}, \tag{30d}$$

$$\sum_{m\in\mathcal{M}} \beta_{m,n}^{k,t} \le 1, \forall n\in\mathcal{Q}, k\in\mathcal{K}^u, t\in\mathcal{T}, \tag{30e}$$

$$\sum_{n\in\mathcal{N}} \phi_n^{k,t} + \sum_{m\in\mathcal{M}} \theta_{m,0}^{k,t} \le 1, \forall k\in\mathcal{K}, t\in\mathcal{T}, \tag{30f}$$

$$\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{Q}} \theta_{m,n}^{k,t} \ge c_{m,t}, \forall m\in\mathcal{S}^M, t\in\mathcal{T}, \tag{30g}$$

$$\sum_{t\in\mathcal{T}} c_{m,t} \le 1, \forall m\in\mathcal{M}, \tag{30h}$$

$$\sum_{n\in\mathcal{Q}\cup\{0\}} f_{m,n}^t \le c_{m,t}, \forall m\in\mathcal{M}, k\in\mathcal{K}\cup\mathcal{K}^u, \forall t\in\mathcal{T}, \tag{30i}$$

$$1 - c_{m,t} = \prod_{n\in\mathcal{Q}\cup\{0\}} (1-f_{m,n}^t), \forall m\in\mathcal{M}, t\in\mathcal{T}, \tag{30j}$$

$$1 - f_{m,n}^t = \prod_{k\in\mathcal{K}} (1-\theta_{m,n}^{k,t}),$$
$$\forall m \in \mathcal{M}, n\in\mathcal{Q}\cup\{0\}, t\in\mathcal{T}, \tag{30k}$$

$$\sum_{k\in\mathcal{K}} R_{m,0}^{k,t} \ge d_m c_{m,t}, \forall m\in\mathcal{S}^C, t\in\mathcal{T}, \tag{30l}$$

$$\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{Q}} R_{m,n}^{k,t} + \sum_{k\in\mathcal{K}^u}\sum_{n\in\mathcal{Q}} R_{m,n}^{k,t,u} \ge d_m c_{m,t},$$
$$\forall m \in \mathcal{S}^M, t\in\mathcal{T}_{ON}, \tag{30m}$$

$$\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{Q}} R_{m,n}^{k,t} \ge d_m c_{m,t}, \forall m\in\mathcal{S}^M, t\in\mathcal{T}\backslash\mathcal{T}_{ON}, \tag{30n}$$

$$\sum_{k\in\mathcal{K}} R_n^{k,t} \ge \sum_{m\in\mathcal{S}^M}\sum_{k\in\mathcal{K}\cup\mathcal{K}^u} R_{m,n}^{k,t}, \forall n\in\mathcal{Q}, t\in\mathcal{T}_{ON}, \tag{30o}$$

$$\sum_{k\in\mathcal{K}} R_n^{k,t} \ge \sum_{m\in\mathcal{S}^M}\sum_{k\in\mathcal{K}} R_{m,n}^{k,t}, \forall n\in\mathcal{Q}, t\in\mathcal{T}\backslash\mathcal{T}_{ON}, \tag{30p}$$

$$\sum_{k\in\mathcal{K}} R_n^{k,t} \ge R_{min}, \forall n\in\mathcal{N}\backslash\mathcal{Q}, t\in\mathcal{T}, \tag{30q}$$

$$\phi_n^{k,t}, \theta_{m,n}^{k,t}, \beta_{m,n}^{k,t}, f_{m,n}^t, c_{m,t} \in \{0,1\}, p_n^{k,t} \ge 0, p_{m,n}^{k,t} \ge 0. \tag{30r}$$

Constraints (30b) and (30c) are the power allocation constraints for the ON and OFF modes, respectively. Constraints (30d)- (30f) are the subchannel constraints corresponding to (14)-(16). According to the property of CA, each M2M link needs to occupy at least one licensed subchannel for control

signals, and thus, constraint (30g) needs to be satisfied. Constraints (30h)-(30k) are scheduling and association constraints corresponding to (18), (20), (22), and (23), respectively. Constraints (30l)-(30q) are rate constraints corresponding to (24)-(29).

In general, the solution to problem (P1) is not unique, since different IoT device association strategies may result in the same number of scheduled IoT devices. Denote the optimal solution set of problem (P1) as Ω, where

$$\Omega = \{(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{F},\boldsymbol{P})|\forall(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{F},\boldsymbol{P}) \text{ is a solution}$$
$$\text{to problem (P1)}\}. \tag{31}$$

To reduce power consumption of the IoT system, it is desirable to find one solution in Ω that consumes the minimum amount of power for the IoT devices and CUs. This can be achieved by solving the following power control problem:

$$\text{P2}: \min_{(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{F},\boldsymbol{P})\in\Omega} \sum_{m\in\mathcal{M}}\sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{K}\cup\mathcal{K}^u}\sum_{n\in\mathcal{Q}\cup\{0\}} p_{m,n}^{k,t} +$$
$$\sum_{n\in\mathcal{N}}\sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{K}} p_n^{k,t}. \tag{32}$$

Through solving problems (P1) and (P2) sequentially, we can achieve the maximum system utility with the minimum transmission power. Problem (P1) is a mixed-integer non-convex problem due to the products of resource allocation variables in the constraints. In this paper, we adopt the single-stage formulation technique [31] to combine problems (P1) and (P2) into a single-stage optimization problem, which is easier to tackle with.

The utility of the single-stage optimization problem can expressed as

$$U = \sum_{m\in\mathcal{M}}\sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{K}\cup\mathcal{K}^u}\sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon p_{m,n}^{k,t} + \sum_{n\in\mathcal{N}}\sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{K}} \epsilon p_n^{k,t}$$
$$+(1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1-c_{m,t}), \tag{33}$$

where λ_m's are positive integers obtained by scaling up w_m by the same constant ρ, i.e.,

$$\rho \triangleq \frac{\lambda_m}{w_m}, \forall m\in\mathcal{M}. \tag{34}$$

Thus, the single-stage problem be formulated as

$$\text{P3}: \min_{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{F},\boldsymbol{P}} U,$$
$$s.t. \quad (30b)\text{-}(30r), \tag{35}$$

The following theorem shows that solving problem (P3) is equivalent to solving problems (P1) and (P2) sequentially, as long as ϵ is a constant satisfying

$$0 < \epsilon < \frac{1}{MP^{iot} + NP^c + 1}. \tag{36}$$

Theorem 1. *Given ϵ as in (36), the solution to problem (P3) is equivalent to solving problems (P1) and (P2) sequentially.*

Proof: See Appendix A. ∎

C. Problem Decomposition

Since problem (P3) is an MINLP problem, it is NP-hard. Besides, the IoT devices association, scheduling, and power allocation variables are nested. To solve problem (P3) efficiently, we decompose it into two subproblems: IoT device association and scheduling subproblem, and power allocation subproblem.

1) IoT Device Association and Scheduling Subproblem: Given the power allocation $\boldsymbol{P}$, problem (P3) can be written by

$$\text{SP1}: \quad \min_{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{F}} (1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1-c_{m,t}),$$
$$s.t. \quad (30d)\text{-}(30q), \qquad\qquad (37)$$
$$\phi_n^{k,t}, \theta_{m,n}^{k,t}, \beta_{m,n}^{k,t}, f_{m,n}^t, c_{m,t} \in \{0,1\}.$$

Motivated by the decomposition technique in [32], define δ_m as the Lagrangian multiplier corresponding to constraint (30h). Thus, the Lagrangian dual problem can be written by

$$\text{SP2}: \max_{\delta_m\geq 0} \quad \sum_{t\in\mathcal{T}} \min_{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{F}} \left((1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m(1-c_{m,t}) - \right.$$
$$\left. \sum_{m\in\mathcal{M}} \delta_m(c_{m,t}-1/T) \right),$$
$$s.t. \quad (30d)\text{-}(30g), (30i)\text{-}(30q),$$
$$\phi_n^{k,t}, \theta_{m,n}^{k,t}, \beta_{m,n}^{k,t}, f_{m,n}^t, c_{m,t} \in \{0,1\}. \qquad (38)$$

To solve problem (SP2), we need to solve $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{F}$ given δ_m first. Since the constraints in (SP2) are independent in each subframe t, the problem (SP2) can be decomposed into T concurrent subproblems given δ_m. In subframe $t \in \mathcal{T}_{ON}$, the problem can be given by

$$\text{SP3}: \quad \min_{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{F}} \left((1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m(1-c_{m,t}) - \right.$$
$$\left. \sum_{m\in\mathcal{M}} \delta_m(c_{m,t}-1/T) \right), \qquad (39)$$
$$s.t. \quad (30d)\text{-}(30g), (30i)\text{-}(30m), (30o), (30q),$$
$$\phi_n^{k,t}, \theta_{m,n}^{k,t}, \beta_{m,n}^{k,t}, c_{m,t} \in \{0,1\}.$$

In the subframe with the OFF mode, constraints (30m) and (30o) will be replaced by constraints (30n) and (30p). Since they are in the same function form, the algorithm to solve problem (SP3) can also be used in the subframe with the OFF mode. For brevity, we will only discuss problem (SP3) in the following.

Then, we will update the Lagrangian multiplier δ_m according to the following rule until it converges. The update rule can be expressed as

$$\delta_m^{(l+1)} = \left[\delta_m^{(l)} - \eta_\delta \left(1 - \sum_{t\in\mathcal{T}} c_{m,t} \right) \right]^+, \qquad (40)$$

where $[x]^+ = \max\{x,0\}$, l is the iteration indicator, and η_δ is the step size.

2) Power Allocation Subproblem: Given the IoT device association and subchannel allocation $\boldsymbol{A},\boldsymbol{B}$ and scheduling $\boldsymbol{C}$, problem (P3) can be converted into

$$\text{SP4}: \quad \min_{\boldsymbol{P}} \sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{N}\cup\{0\}} \epsilon p_{m,n}^{k,t} +$$
$$\sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon p_n^{k,t}, \qquad (41)$$
$$s.t. \quad (30b), (30c), (30l)\text{-}(30q),$$
$$p_n^{k,t} \geq 0, p_{m,n}^{k,t} \geq 0.$$

Since the power allocation in different subframes is independent, we decouple the problem (SP4) into T subproblems, which can be solved in parallel. For simplicity, we only solve the subprolem in the subframe with the ON mode $t \in \mathcal{T}_{ON}$. The subproblem can be written by

$$\text{SP5}: \quad \min_{\boldsymbol{P}} \sum_{m\in\mathcal{M}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{N}\cup\{0\}} \epsilon p_{m,n}^{k,t} + \sum_{n\in\mathcal{N}} \sum_{k\in\mathcal{K}} \epsilon p_n^{k,t}$$
$$s.t. \quad (30b), (30l), (30m), (30o), (30q),$$
$$p_n^{k,t} \geq 0, p_{m,n}^{k,t} \geq 0.$$
$$(42)$$

The same approach can also be applied to the power allocation in the licensed band.

V. Iterative Association, Scheduling, and Power Allocation Algorithm

In this section, we will design a low-complexity iterative association, scheduling, and power allocation (IASPA) algorithm to obtain the suboptimal solutions to problem (P3)[3]. Specifically, we solve (SP3) given the power allocation $\boldsymbol{P}$, and then solve (SP5) given the IoT device association $\boldsymbol{F}$ and scheduling $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$. Finally, we present the overall algorithm. Without loss of generality and for notational brevity, we omit the time index t in the following discussions.

A. IoT Device Association and Scheduling Algorithm

Note that all the variables in problem (SP3) are integers, and thus, it is an NP-hard problem. To solve this problem effectively, we transform the non-linear constraints into the linear ones. Thus, the problem (P3) is converted to a series of linear integer programming problems, which can be solved by the branch-and-bound algorithm [17].

Consider the polynomial constraint (30j). The following theorem [33] shows that the non-linear constraint is equivalent to two linear inequalities.

Theorem 2. *Equation (30j) holds if and only if*

$$- \sum_{n\in\mathcal{Q}\cup\{0\}} (1-f_{m,n}) + (Q+1)(1-c_m) \leq 0, \qquad (43)$$
$$\sum_{n\in\mathcal{Q}\cup\{0\}} (1-f_{m,n}) - (1-c_m) \leq Q. \qquad (44)$$

Proof: If any $f_{m,n} = 1$, then $c_m = 1$. In this case, (43) becomes $c_m \geq \sum_{n\in\mathcal{Q}\cup\{0\}} \frac{f_{m,n}}{Q+1} > 0$ which implies $c_m = 1$, and (44) is redundant. If all $f_{m,n} = 0$, then $c_m = 0$. In

---

[3]The IASPA algorithm is performed in a centralized manner. The BS obtains the information of the IoT devices and CUs such as locations and channel conditions, and then performs the IASPA algorithm to obtain the association, scheduling, and power allocation results. After that, the BS will inform the IoT devices and CUs of the results over the control channel.

such a situation, (43) is redundant, while (44) becomes $c_m \leq \sum_{n \in \mathcal{Q} \cup \{0\}} f_{m,n} = 0$ which means $c_m = 0$. ∎

Similarly, equation (30k) holds if and only if

$$-\sum_{k \in \mathcal{K}} (1 - \theta_{m,n}^k) + K(1 - f_{m,n}) \leq 0, \tag{45}$$

$$\sum_{k \in \mathcal{K}} (1 - \theta_{m,n}^k) - (1 - f_{m,n}) \leq K - 1. \tag{46}$$

In addition, consider that rate constraints (30l), (30m), (30o) and (30q) are non-convex due to the interference term. To tackle these constraints effectively, we can utilize the a first order Taylor expansion to approximate the data rate into a linear one.

Note that the data rate R_n^k can be rewritten by

$$
\begin{aligned}
R_n^k &= \log_2 \left(\sigma^2 + p_n^k \phi_n^k |h_n^k|^2 + \sum_{m \in \mathcal{S}^M} \sum_{\substack{n' \in \mathcal{Q}, \\ n' \neq n}} p_{m,n'}^k \theta_{m,n'}^k |h_{m,0}^k|^2\right) \\
&\quad - \log_2 \left(\sigma^2 + \sum_{m \in \mathcal{S}^M} \sum_{\substack{n' \in \mathcal{Q}, \\ n' \neq n}} p_{m,n'}^k \theta_{m,n'}^k |h_{m,0}^k|^2\right), \\
&\triangleq f_n^k(\phi_n^k, \boldsymbol{\theta}_{m,-n}^k) - g_n^k(\boldsymbol{\theta}_{m,-n}^k),
\end{aligned}
\tag{47}
$$

where $\boldsymbol{\theta}_{m,-n}^k \triangleq (\theta_{m,n'}^k)_{m \in \mathcal{M}, n' \in \mathcal{Q} - \{n\}}$ represents all $\theta_{m,n'}^k$'s except $\theta_{m,n}^k$. Denote $\tilde{\phi}_n^k$ and $\tilde{\theta}_{m,n}^k$ by the existing solutions for $\boldsymbol{A}$. According to the results in [34], we have the following approximations:

$$
\begin{aligned}
f_n^k(\phi_n^k, \boldsymbol{\theta}_{m,-n}^k) &= f_n^k(\phi_n^k, \tilde{\boldsymbol{\theta}}_{m,-n}^k) \\
&\quad + \sum_{m \in \mathcal{S}^M} \sum_{\substack{n' \in \mathcal{Q}, \\ n' \neq n}} a_{m,n'}^k (\theta_{m,n'}^k - \tilde{\theta}_{m,n'}^k),
\end{aligned}
\tag{48}
$$

$$
\begin{aligned}
g_n^k(\boldsymbol{\theta}_{m,-n}^k) &= g_n^k(\tilde{\boldsymbol{\theta}}_{m,-n}^k) \\
&\quad + \sum_{m \in \mathcal{S}^M} \sum_{\substack{n' \in \mathcal{Q}, \\ n' \neq n}} b_{m,n'}^k (\theta_{m,n'}^k - \tilde{\theta}_{m,n'}^k)
\end{aligned}
\tag{49}
$$

where $a_{m,n'}^k$ and $b_{m,n'}^k$ are the first order gradient over $\theta_{m,n'}^k$ for functions f_n^k and g_n^k, respectively. Note that the value of $\theta_{m,n'}^k$ can be 0 or 1. Thus, the gradient can be replaced by the difference of function values with $\theta_{m,n'}^k = 0$ and $\theta_{m,n'}^k = 1$. That is,

$$
\begin{aligned}
a_{m,n'}^k &= f_n^k(\tilde{\phi}_n^k, \tilde{\boldsymbol{\theta}}_{m,-\{n,n'\}}^k, \theta_{m,n'}^k = 1) \\
&\quad - f_n^k(\tilde{\phi}_n^k, \boldsymbol{\theta}_{m,-\{n,n'\}}^k, \theta_{m,n'}^k = 0),
\end{aligned}
\tag{50}
$$

$$
\begin{aligned}
b_{m,n'}^k &= g_n^k(\boldsymbol{\theta}_{m,-\{n,n'\}}^k, \theta_{m,n'}^k = 1) \\
&\quad - g_n^k(\boldsymbol{\theta}_{m,-\{n,n'\}}^k, \theta_{m,n'}^k = 0).
\end{aligned}
\tag{51}
$$

Similarly, we can also approximate $f_n^k(\phi_n^k, \tilde{\boldsymbol{\theta}}_{m,-n}^k)$ by

$$
\begin{aligned}
f_n^k(\phi_n^k, \tilde{\boldsymbol{\theta}}_{m,-n}^k) &= \left(f_n^k(1, \tilde{\boldsymbol{\theta}}_{m,-n}^k) - f_n^k(0, \tilde{\boldsymbol{\theta}}_{m,-n}^k)\right) \phi_n^k \\
&\quad + f_n^k(0, \tilde{\boldsymbol{\theta}}_{m,-n}^k) \\
&\triangleq a_n^k \phi_n^k + f_n^k(0, \tilde{\boldsymbol{\theta}}_{m,-n}^k).
\end{aligned}
\tag{52}
$$

As such, the data rate R_n^k can be approximated by

$$
\begin{aligned}
R_n^k &= a_n^k \phi_n^k + f_n^k(0, \tilde{\boldsymbol{\theta}}_{m,-n}^k) - g_n^k(\tilde{\boldsymbol{\theta}}_{m,-n}^k) \\
&\quad + \sum_{m \in \mathcal{S}^M} \sum_{\substack{n' \in \mathcal{Q}, \\ n' \neq n}} (a_{m,n'}^k - b_{m,n'}^k)(\theta_{m,n'}^k - \tilde{\theta}_{m,n'}^k)
\end{aligned}
\tag{53}
$$

Follow the same approach, we can also approximate $R_{m,n}^k$ and $R_{m,n}^{k,u}$ as a linear function of $\theta_{m,n}^k$, ϕ_n^k and $\beta_{m,n}^k$. Therefore, constraints (30l), (30m), (30o), and (30q) can be converted into linear constraints, and thus, problem (SP3) is transformed into a linear integer programming, which can be solved by the existing branch-and-bound algorithm.

B. Power Allocation Algorithm

Due to the existence of the interference term in constraints (30l), (30m), (30o), and (30q), the problem (SP5) is a non-convex optimization problem. However, we can utilize the SCA technique [18] to approximate the non-convex functions into a series of convex ones. As such, the non-convex problem can be transformed into a convex one and solved by the Lagrange dual technique [19].

1) Successive Convex Approximation: According to the results in [18], the data rate R_n^k can be approximated by

$$\tilde{R}_n^k = d_n^k \log(\gamma_n^k) + e_n^k, \tag{54}$$

where

$$d_n^k = \frac{\tilde{\gamma}_n^k}{1 + \tilde{\gamma}_n^k}, \quad e_n^k = \frac{1}{\log 2} \log(1 + \tilde{\gamma}_n^k) - d_n^k \log(\tilde{\gamma}_n^k). \tag{55}$$

Here, $\tilde{\gamma}_n^k$ is calculated based on the current power allocation result. Likewise, $R_{m,n}^k$ and $R_{m,n}^{k,u}$ can be also expressed by

$$
\begin{aligned}
\tilde{R}_{m,n}^k &= d_{m,n}^k \log(\gamma_{m,n}^k) + e_{m,n}^k, \\
\tilde{R}_{m,n}^{k,u} &= d_{m,n}^{k,u} \log(\gamma_{m,n}^{k,u}) + e_{m,n}^{k,u}.
\end{aligned}
\tag{56}
$$

However, $\log(\tilde{\gamma}_n^k)$, $\log(\tilde{\gamma}_{m,n}^k)$, and $\log(\tilde{\gamma}_{m,n}^{k,u})$ are still non-convex. Therefore, we can introduce the transformations of $p_{m,n}^k = \exp(\hat{p}_{m,n}^k)$ and $p_n^k = \exp(\hat{p}_n^k)$. As such, problem (SP5) can be converted into the following problem:

$$
\text{SP6}: \min_{\hat{\boldsymbol{P}}} \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K} \cup \mathcal{K}^u} \sum_{n \in \mathcal{Q} \cup \{0\}} \epsilon \exp(\hat{p}_{m,n}^k) +
$$

$$
\sum_{n \in \mathcal{N}} \sum_{k \in \mathcal{K}} \epsilon \exp(\hat{p}_n^k) \tag{57a}
$$

$$
s.t. \sum_{k \in \mathcal{K}} \exp(\hat{p}_{m,0}^k) + \sum_{k \in \mathcal{K} \cup \mathcal{K}^u} \sum_{n \in \mathcal{Q}} \exp(\hat{p}_{m,n}^k) \leq P^{iot},
$$

$$
\sum_{k \in \mathcal{K}} \exp(\hat{p}_n^k) \leq P^c, \forall m \in \mathcal{M}, n \in \mathcal{N}, \tag{57b}
$$

$$
\sum_{k \in \mathcal{K}} (d_{m,0}^k \log(\gamma_{m,0}^k) + e_{m,0}^k) \geq d_m c_{m,t}, \forall m \in \mathcal{S}^C, \tag{57c}
$$

$$
\sum_{n \in \mathcal{Q}} \left(\sum_{k \in \mathcal{K}} (d_{m,n}^k \log(\gamma_{m,n}^k) + e_{m,n}^k) + \right.
$$

$$
\left. \sum_{k \in \mathcal{K}^u} (d_{m,n}^{k,u} \log(\gamma_{m,n}^{k,u}) + e_{m,n}^{k,u}) \right) \geq d_m c_{m,t}, \forall m \in \mathcal{S}^M, \tag{57d}
$$

$$
\sum_{k \in \mathcal{K}} (d_n^k \log(\gamma_n^k) + e_n^k) \geq \sum_{m \in \mathcal{S}^M} \sum_{k \in \mathcal{K}} (d_{m,n}^k \log(\gamma_{m,n}^k) + e_{m,n}^k)
$$

$$
+ \sum_{m \in \mathcal{S}^M} \sum_{k \in \mathcal{K}^u} (d_{m,n}^{k,u} \log(\gamma_{m,n}^{k,u}) + e_{m,n}^{k,u}), \forall n \in \mathcal{Q}, \tag{57e}
$$

$$\sum_{k\in\mathcal{K}}(d_n^k\log(\gamma_n^k)+e_n^k)\geq R_{min}, \forall n\in\mathcal{N}\backslash\mathcal{Q}, \qquad (57f)$$

where

$$\log(\gamma_n^k)=\hat{p}_n^k+\log(|h_n^k|^2\phi_n^k)-$$
$$\log(\sigma^2+\sum_{m\in\mathcal{S}^M}\sum_{\substack{n'\in\mathcal{Q},\\n'\neq n}}\exp(\hat{p}_{m,n'}^k)\theta_{m,n'}^{k,t}|h_{m,0}^k|^2), \quad (58)$$

$$\log(\gamma_{m,0}^k)=\hat{p}_{m,0}^k+\log(|h_{m,0}^k|^2\phi_{m,0}^k)-$$
$$\log(\sigma^2+\sum_{m'\in\mathcal{S}^M}\sum_{n\in\mathcal{Q}}\exp(\hat{p}_{m',n}^k)\theta_{m',n}^k|h_{m',0}^k|^2), \quad (59)$$

$$\log(\gamma_{m,n}^k)=\hat{p}_{m,n}^k+\log(|h_{m,n}^k|^2\phi_{m,n}^k)-$$
$$\log\left(\sigma^2+\sum_{\substack{n'\in\mathcal{N},\\n'\neq n}}\exp(\hat{p}_{n'}^k)\phi_{n'}^k|h_{n',n}^k|^2+\right.$$
$$\left.\sum_{\substack{m'\in\mathcal{M},\,n'\in\mathcal{Q}\cup\{0\}\\m'\neq m\ \ n'\neq n}}\exp(\hat{p}_{m',n'}^k)\theta_{m',n'}^k|h_{m',n}^k|^2\right), \quad (60)$$

$$\log(\gamma_{m,n}^{k,u})=\hat{p}_{m,n}^k+\log(|h_{m,n}^{k,u}|^2\beta_{m,n}^{k,u})-$$
$$\log\left(\sigma^2+\sum_{\substack{m'\in\mathcal{M},\,n'\in\mathcal{Q},\\m'\neq m\ \ n'\neq n}}\exp(\hat{p}_{m',n'}^k)\beta_{m',n'}^k|h_{m',n}^{k,u}|^2\right.$$
$$\left.+I_{m,n}^{k,u}\right). \quad (61)$$

Since the log-sum-exp function is convex [19], problem (SP6) is a standard convex minimization problem, which can be solved by the Lagrange dual technique. Once the optimal solution is obtained, we may transform back with $p_m^k=\exp(\hat{p}_m^k)$ and $p_n^k=\exp(\hat{p}_n^k)$. To find the optimal power allocation, we need to iteratively update the approximation parameters according to the allocated power.

2) Lagrange Dual Technique: Define $\boldsymbol{\mu}$, $\boldsymbol{\nu}$, $\boldsymbol{\kappa}$, $\boldsymbol{\xi}$ and $\boldsymbol{\chi}$ are Lagrange multiplier vectors corresponding to constraints (57b), (57c), (57d), (57e) and (57f), respectively. Therefore,

the Lagrangian can be written as

$$L(\hat{\boldsymbol{P}},\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\kappa},\boldsymbol{\xi},\boldsymbol{\chi})=\sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}\cup\mathcal{K}^u}\sum_{n\in\mathcal{Q}\cup\{0\}}\epsilon\exp(\hat{p}_{m,n}^k)$$
$$+\sum_{n\in\mathcal{N}}\sum_{k\in\mathcal{K}}\epsilon\exp(\hat{p}_n^k)+\sum_{m\in\mathcal{M}}\mu_m\left(P^{iot}-\sum_{k\in\mathcal{K}}\exp(\hat{p}_{m,0}^k)\right.$$
$$\left.-\sum_{k\in\mathcal{K}\cup\mathcal{K}^u}\sum_{n\in\mathcal{Q}}\exp(\hat{p}_{m,n}^k)\right)+\sum_{n\in\mathcal{N}}\mu_n\left(P^c-\sum_{k\in\mathcal{K}}\exp(\hat{p}_n^k)\right)$$
$$+\sum_{m\in\mathcal{S}^C}\nu_m\left(\sum_{k\in\mathcal{K}}(d_{m,0}^k\log(\gamma_{m,0}^k)+e_{m,0}^k)-d_mc_{m,t}\right)$$
$$+\sum_{m\in\mathcal{S}^M}\kappa_m\left(\sum_{n\in\mathcal{Q}}(\sum_{k\in\mathcal{K}}(d_{m,n}^k\log(\gamma_{m,n}^k)+e_{m,n}^k)\right.$$
$$+\sum_{k\in\mathcal{K}^u}(d_{m,n}^{k,u}\log(\gamma_{m,n}^{k,u})+e_{m,n}^{k,u}))-d_mc_{m,t}\bigg)$$
$$+\sum_{n\in\mathcal{Q}}\xi_n\left(\sum_{k\in\mathcal{K}}(d_n^k\log(\gamma_n^k)+e_n^k)\right.$$
$$-\sum_{m\in\mathcal{S}^M}\sum_{k\in\mathcal{K}}(d_{m,n}^k\log(\gamma_{m,n}^k)+e_{m,n}^k)$$
$$\left.-\sum_{m\in\mathcal{S}^M}\sum_{k\in\mathcal{K}^u}(d_{m,n}^{k,u}\log(\gamma_{m,n}^{k,u})+e_{m,n}^{k,u})\right)$$
$$+\sum_{n\in\mathcal{N}\backslash\mathcal{Q}}\chi_n\left(\sum_{k\in\mathcal{K}}(d_n^k\log(\gamma_n^k)+e_n^k)-R_{min}\right). \quad (62)$$

and the dual problem can be given by

$$\max_{\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\kappa},\boldsymbol{\xi},\boldsymbol{\chi}\succeq\mathbf{0}}\ \min_{\hat{\boldsymbol{P}}}L(\hat{\boldsymbol{P}},\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\kappa},\boldsymbol{\xi},\boldsymbol{\chi}). \quad (63)$$

We can solve the dual problem iteratively by decomposing it into two subproblems. The master subproblem is the maximization of $\boldsymbol{\mu}$, $\boldsymbol{\nu}$, $\boldsymbol{\kappa}$, $\boldsymbol{\xi}$, and $\boldsymbol{\chi}$. The slave subproblem is to maximize $\hat{\boldsymbol{P}}$ given $\boldsymbol{\mu}$, $\boldsymbol{\nu}$, $\boldsymbol{\kappa}$, $\boldsymbol{\xi}$, and $\boldsymbol{\chi}$. In the following, we will discuss them in detail.

Slave Subproblem: According to the Karush-Kuhn-Tucker (KKT) conditions, we can obtain the optimal power allocation by equaling the first derivative of the Lagrangian function $L(\hat{\boldsymbol{P}},\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\kappa},\boldsymbol{\xi},\boldsymbol{\chi})$ over $\hat{\boldsymbol{P}}$ to 0. Therefore, we have

$$\exp(\hat{p}_n^k)=d_n^k\xi_n\bigg/\bigg(\mu_n-\epsilon+\sum_{m\in\mathcal{S}^M}\sum_{\substack{n'\in\mathcal{Q},\\n'\neq n}}(\kappa_md_{m,n'}^k-$$
$$\xi_{n'}d_{m,n'}^k)\frac{\gamma_{m,n'}^k\phi_n^k\theta_{m,n'}^k|h_{n,n'}^k|^2}{\exp(\hat{p}_{m,n'}^k)|h_{m,n'}^k|^2}\bigg),\forall n\in\mathcal{Q},k\in\mathcal{K}, \quad (64)$$

$$\exp(\hat{p}_n^k)=d_n^k\chi_n\bigg/\bigg(\mu_n-\epsilon+\sum_{m\in\mathcal{S}^M}\sum_{n'\in\mathcal{Q},}\kappa_md_{m,n'}^k$$
$$\frac{\gamma_{m,n'}^k\phi_n^k\theta_{m,n'}^k|h_{n,n'}^k|^2}{\exp(\hat{p}_{m,n'}^k)|h_{m,n'}^k|^2}\bigg),\forall n\in\mathcal{N}\backslash\mathcal{Q},k\in\mathcal{K}, \quad (65)$$

$$\exp(\hat{p}_{m,0}^k)=d_{m,0}^k\nu_m\bigg/\bigg(\mu_m-\epsilon+\sum_{\substack{m'\in\mathcal{S}^M,\\m'\neq m}}\sum_{n\in\mathcal{Q}}(\kappa_{m'}d_{m',n}^k$$
$$-\xi_nd_{m',n}^k)\frac{\gamma_{m',n}^k\theta_{m,0}^k|h_{m,n}^k|^2\theta_{m',n}^k}{\exp(\hat{p}_{m',n}^k)|h_{m',n}^k|^2}\bigg),\forall m\in\mathcal{S}^C,k\in\mathcal{K}, \quad (66)$$

$$
\begin{aligned}
\exp(\hat{p}_{m,n}^{k}) = {}& \left(\kappa_m d_{m,n}^{k} - \xi_n d_{m,n}^{k}\right)/\bigg(\mu_m - \epsilon + \\
& \sum_{\substack{m' \in \mathcal{S}^C, \\ m' \neq m}} \nu_{m'} d_{m',0}^{k} \frac{\gamma_{n'',0}^{k} \theta_{m,n}^{k} \theta_{m',0}^{k} |h_{m,0}^{k}|^2}{\exp(\hat{p}_{m',0}^{k})|h_{m',0}^{k}|^2} + \\
& \sum_{\substack{n' \in \mathcal{Q}, \\ n' \neq n}} \xi_{n'} d_{n'}^{k} \frac{\gamma_{n'}^{k} \theta_{m,n}^{k} \phi_{n'}^{k} |h_{m,0}^{k}|^2}{\exp(\hat{p}_{n'}^{k})|h_{n'}^{k}|^2} + \sum_{\substack{n' \in \mathcal{N}\backslash\mathcal{Q}, \\ n' \neq n}} \chi_{n'} d_{n'}^{k} \frac{\gamma_{n'}^{k} \theta_{m,n}^{k} \phi_{n'}^{k} |h_{m,0}^{k}|^2}{\exp(\hat{p}_{n'}^{k})|h_{n'}^{k}|^2} \\
& + \sum_{\substack{m' \in \mathcal{S}^M \\ m' \neq m}} \sum_{\substack{n' \in \mathcal{Q} \\ n' \neq n}} (\kappa_{m'} d_{m',n'}^{k} - \xi_{n'} d_{m',n'}^{k}) \frac{\gamma_{m',n'}^{k} \theta_{m,n}^{k} \theta_{m',n'}^{k} |h_{m,n'}^{k}|^2}{\exp(\hat{p}_{m',n'}^{k})|h_{m',n'}^{k}|^2} \bigg), \\
& \forall m \in \mathcal{S}^M, n \in \mathcal{Q}, k \in \mathcal{K},
\end{aligned}
\tag{67}
$$

$$
\begin{aligned}
\exp(\hat{p}_{m,n}^{k}) = {}& \left(\kappa_m d_{m,n}^{k,u} - \xi_n d_{m,n}^{k,u}\right)/\bigg(\mu_m - \epsilon + \\
& \sum_{\substack{m' \in \mathcal{S}^M \\ m' \neq m}} \sum_{\substack{n' \in \mathcal{N}, \\ n' \neq n}} (\kappa_{m'} d_{m',n'}^{k,u} - \xi_{n'} d_{m',n'}^{k,u}) \frac{\gamma_{m',n'}^{k,u} \beta_{m,n}^{k} \beta_{m',n'}^{k} |h_{m,n'}^{k,u}|^2}{\exp(\hat{p}_{m',n'}^{k})|hm',n'^{k,u}|^2}\bigg), \\
& \forall m \in \mathcal{S}^M, n \in \mathcal{N}, k \in \mathcal{K}^u.
\end{aligned}
\tag{68}
$$

As such, we can obtain the optimal power allocation results for all the scheduled IoT devices.

Master Subproblem: Once the optimal power allocation is achieved, the solution of the dual problem can be obtained by a subgradient method [35] as follows:

$$
\mu_n^{(l+1)} = \left[\mu_n^{(l)} - \eta_\mu \left(P^c - \sum_{k \in \mathcal{K}} \exp(\hat{p}_n^{k})\right)\right]^+, \forall n \in \mathcal{Q},
\tag{69}
$$

$$
\begin{aligned}
\mu_m^{(l+1)} = {}& \bigg[\mu_m^{(l)} - \eta_\mu \bigg(P^{iot} - \sum_{k \in \mathcal{K}} \exp(\hat{p}_{m,0}^{k}) - \\
& \sum_{k \in \mathcal{K} \cup \mathcal{K}^u} \sum_{n \in \mathcal{Q}} \exp(\hat{p}_{m,n}^{k})\bigg)\bigg]^+, \forall m \in \mathcal{M},
\end{aligned}
\tag{70}
$$

$$
\begin{aligned}
\nu_m^{(l+1)} = {}& \bigg[\nu_m^{(l)} - \eta_\nu \bigg(\sum_{k \in \mathcal{K}} (d_{m,0}^{k} \log(\gamma_{m,0}^{k}) + e_{m,0}^{k}) - \\
& d_m c_{m,t})\bigg]^+, \forall m \in \mathcal{S}^C,
\end{aligned}
\tag{71}
$$

$$
\begin{aligned}
\kappa_m^{(l+1)} = {}& \bigg[\kappa_m^{(l)} - \eta_\kappa \bigg(\sum_{n \in \mathcal{Q}} \sum_{k \in \mathcal{K}} (d_{m,n}^{k} \log(\gamma_{m,n}^{k}) + e_{m,n}^{k}) \\
& + \sum_{k \in \mathcal{K}^u} (d_{m,n}^{k,u} \log(\gamma_{m,n}^{k,u}) + e_{m,n}^{k,u})) - d_m c_{m,t}\bigg)\bigg]^+, \\
& \forall m \in \mathcal{S}^M,
\end{aligned}
\tag{72}
$$

$$
\begin{aligned}
\xi_n^{(l+1)} = {}& \bigg[\xi_n^{(l)} - \eta_\xi \bigg(\sum_{k \in \mathcal{K}} (d_n^{k} \log(\gamma_n^{k}) + e_n^{k}) - \\
& \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}} (d_{m,n}^{k} \log(\gamma_{m,n}^{k}) + e_{m,n}^{k}) - \\
& \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}^u} (d_{m,n}^{k,u} \log(\gamma_{m,n}^{k,u}) + e_{m,n}^{k,u})\bigg)\bigg]^+, \forall n \in \mathcal{Q},
\end{aligned}
\tag{73}
$$

$$
\chi_n^{(l+1)} = \left[\chi_n^{(l)} - \eta_\chi \left(\sum_{k \in \mathcal{K}} (d_n^{k} \log(\gamma_n^{k}) + e_n^{k}) - R_{min}\right)\right]^+,
$$

Algorithm 1 The Procedure of IASPA Algorithm

Input: The set of IoT devices $\mathcal{M}$; The set of CUs $\mathcal{N}$; The set of idle CUs $\mathcal{Q}$; The number of licensed subchannels K and unlicensed subchannels K^u; The number of subframes for one cycle T;

Output: IoT association and scheduling matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{F}$; Transmit power matrix $\boldsymbol{P}$;

1: Initialize transmit power $\boldsymbol{P}_{int}^{0}$ by setting $p_{m,n}^{k} = P^{iot}/(K + K^u)$ and $p_n^{k} = P^c/K$, and $l = 0$;
2: **repeat**
3:     $l = l + 1$;
4:     Solve subproblem (SP3) given the initial power allocation $\boldsymbol{P}_{int}^{l-1}$, and obtain the association and scheduling result $\boldsymbol{A}^l$, $\boldsymbol{B}^l$, $\boldsymbol{C}^l$, and $\boldsymbol{F}^l$;
5:     Solve subproblem (SP5) given the association and scheduling result $\boldsymbol{A}^l$, $\boldsymbol{B}^l$, $\boldsymbol{C}^l$, and $\boldsymbol{F}^l$, and then obtain the power allocation result $\boldsymbol{P}^l$;
6:     Use the solved power allocation result to replace the transmit power of the scheduled IoT devices in $\boldsymbol{P}_{int}^{l}$ and obtain an updated initial transmit power matrix $\boldsymbol{P}_{int}^{l+1}$ for the next iteration;
7: **until** $|O(\boldsymbol{A}^l, \boldsymbol{B}^l, \boldsymbol{C}^l, \boldsymbol{F}^l, \boldsymbol{P}^l) - O(\boldsymbol{A}^{l-1}, \boldsymbol{B}^{l-1}, \boldsymbol{C}^{l-1}, \boldsymbol{F}^{l-1}\boldsymbol{P}^{l-1})| \leq \pi$;

$$
\forall n \in \mathcal{N}\backslash\mathcal{Q},
\tag{74}
$$

where l is the iteration indicator, and η_μ, η_ν, η_κ, η_ξ and η_χ are sufficiently small step sizes to guarantee convergence. As such, the optimal solution for the dual problem when the multiplier vectors converge.

C. Overall Algorithm

Based on the results presented in the previous two subsections, we propose an overall iterative algorithm, i.e., IASPA algorithm for problem (P3) by applying the alternating optimization method. Initially, we set the power for each link are allocated to all the subchannels equally. Then the IoT device association and scheduling variables $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{F}$ are optimized by solving subproblem (SP3) while keeping the allocated power fixed. After obtaining the association and scheduling results, we will optimize the power allocation $\boldsymbol{P}$ on these scheduled IoT devices by solving subproblem (SP5). Those scheduled IoT will utilize the optimized transmit power as the initial power in the next iterations. Define $O(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{F}, \boldsymbol{P})$ as the value of the objective function in problem (P3). The two subproblems will be solved alternatively until the value difference of the objective functions between two successive iterations is less than a predefined threshold π, i.e., $|O(\boldsymbol{A}^{l+1}, \boldsymbol{B}^{l+1}, \boldsymbol{C}^{l+1}, \boldsymbol{F}^{(l+1)}, \boldsymbol{P}^{l+1}) - O(\boldsymbol{A}^l, \boldsymbol{B}^l, \boldsymbol{C}^l, \boldsymbol{F}^{(l)}, \boldsymbol{P}^l)| \leq \pi$, where l denotes the iteration. The details of IASPA algorithm are summarized in Algorithm 1.

VI. PERFORMANCE ANALYSIS

In this section, we analyze the effectiveness and efficiency of the proposed algorithm, and remark some key properties of

the IoT-U network. In the first part, the convergence and the complexity of the proposed IASPA algorithm is proved. Then, we discuss the IoT devices association and the offloading gain of unlicensed spectrum.

A. Convergence

First, in the IoT association and scheduling subproblem, we can obtain a better association and scheduling solution given $\boldsymbol{P}^l$. Therefore, we have

$$O(\boldsymbol{A}^{l+1}, \boldsymbol{B}^{l+1}, \boldsymbol{C}^{l+1}, \boldsymbol{F}^{l+1}, \boldsymbol{P}^l) \leq O(\boldsymbol{A}^l, \boldsymbol{B}^l, \boldsymbol{C}^l, \boldsymbol{F}^l, \boldsymbol{P}^l). \tag{75}$$

Second, given the association and scheduling $\boldsymbol{A}^{l+1}$, $\boldsymbol{B}^{l+1}$, $\boldsymbol{C}^{l+1}, \boldsymbol{F}^{l+1}$, we minimize the allocated power, and thus, the following inequality holds:

$$\begin{aligned} O(\boldsymbol{A}^{l+1}, \boldsymbol{B}^{l+1}, \boldsymbol{C}^{l+1}, \boldsymbol{F}^{l+1}, \boldsymbol{P}^{l+1}) \leq \\ O(\boldsymbol{A}^{l+1}, \boldsymbol{B}^{l+1}, \boldsymbol{C}^{l+1}, \boldsymbol{F}^{l+1} \boldsymbol{P}^l). \end{aligned} \tag{76}$$

Based the inequalities (75) and (76), we can obtain

$$O(\boldsymbol{A}^{l+1}, \boldsymbol{B}^{l+1}, \boldsymbol{C}^{l+1}, \boldsymbol{F}^{l+1}, \boldsymbol{P}^{l+1}) \leq O(\boldsymbol{A}^l, \boldsymbol{B}^l, \boldsymbol{C}^l, \boldsymbol{F}^l, \boldsymbol{P}^l), \tag{77}$$

which implies that the objective value of problem (P3) is non-decreasing after each iteration of the ISAPA Algorithm. Since the objective value of problem (P3) is upper bounded by a finite value, the proposed ISAPA algorithm is guaranteed to converge.

B. Complexity

In the IoT device association and scheduling subproblem, we approximate it as a sequence of integer linear programming problems. Based on the results in [36], the integer linear programming problem is a polynomial computational complexity related to the number of variables, i.e., $M(N+1)(K+1)T + MNK^uT$.

In power allocation subproblem, we approximate it as a series of convex problems and solve it by the primary-dual interior point method. According to the results in [19], the scale of the computational complexity of each convex problem is $O(\sqrt{Z}\log(\frac{1}{\tau}))$, where $Z = (M+1)(N+1)$ is the number of constraints and τ is the tolerance threshold for convergence.

C. Association

Proposition 1. *Define the data rate of a CU is R. The number of the IoT devices associating with this CU cannot exceed* $\min\{K-1, R/d\}$.

Proof: On the one hand, note that the associated IoT devices cannot share the same spectrum with the CU, and thus, these IoT devices can only utilize at most $K-1$ licensed subchannels. Besides, due to the CA requirement, each IoT device need to be allocated to at least one licensed subchannel. Therefore, the number of the associated IoT devices cannot be larger than $K-1$. On the other hand, each IoT device has d data traffic to transmit, and therefore, the CU can support at most R/d IoT devices. ∎

Remark 1. *The IoT devices with high data traffic will tend to associate with the BS while those with low data traffic will tend to associate with the CUs.*

Proof: When the data traffic is high, for example, the data rate for the cellular link can support the aggregated data from only one IoT device, the data rate in the cellular link will be the bottleneck. Therefore, the IoT device will associate with the BS directly since it will not bring the interference. However, when the data traffic is low, the cellular link can support the data aggregated from more than one IoT devices. Even the IoT device is located close to the BS, the IoT device will still associate with a neighboring CU because a subchannel can support more than one IoT devices in this case. ∎

D. Offloading Gain

Proposition 2. *The number of scheduled IoT devices in the IoT-U network is no less than that using the licensed scheme where only the licensed spectrum is utilized.*

Proof: Note that the number of scheduled IoT devices using the licensed scheme is equivalent to solve (P1) by setting $\beta_{m,n}^{k,t} = 0$. Therefore, the solution obtained by the licensed scheme is also a feasible solution to (P1). Since the proposed unlicensed scheme is to find the optimal solution of (P1), the value of the objective obtained by the proposed unlicensed scheme is no less than that obtained by the licensed scheme. ∎

Remark 2. *Define R_L as the expectation of the data rate for a cellular link over one subchannel using the licensed scheme and R_U as that using the scheme where there does not exist a D2D link over the licensed spectrum. The number of scheduled IoT devices in the IoT-U network can increase $R_U/R_L - 1$ at most compared to the licensed scheme.*

Proof: In the unlicensed scheme, the IoT devices associated with a CU will transmit with the most of power over the unlicensed band. Therefore, the less transmit power over the licensed band alleviates the interference to the cellular links, which brings the offloading gain. For this reason, the IoT-U network can schedule at most R_U/d where d is the expectation of the data traffic for an IoT device when the interference is neglected, while the licensed scheme schedules R_L/d IoT devices. Thus, the IoT-U network can schedule at most $R_U/R_L - 1$ more IoT devices than the licensed scheme. ∎

VII. SIMULATION RESULTS

In this section, we evaluate the performance of our proposed IoT-U scheme. For comparison, the following schemes are also performed:

- Traditional cellular (TC) networks: All IoT devices transmit their sensory data to the BS in cellular mode.
- Cluster based IoT networks over licensed spectrum (IoT-L): The IoT devices adopt the cluster framework where only the licensed spectrum can be utilized.
- Decoupled association, scheduling, and power allocation (DASPA) scheme: The IoT devices association and

TABLE I
PARAMETERS FOR SIMULATION

Parameters	Values
Number of WUs	3
WU's transmit power P^w	23 dBm
Transmission bandwidth	180 kHz
Interval for a time slot	100 ms
Number of time slots for a cycle T	6
Duty cycle percentage T_{ON}/T	0.5
Carrier frequency of licensed band	1.9 GHz
Carrier frequency of unlicensed band	5 GHz
Noise figure	5 dB
Error tolerance level π	0.1

scheduling are performed first, and then the transmission power on each subchannel is allocated to minimize the total energy consumption.

- Matching based association, scheduling, and power allocation (MASPA) scheme [37]: The IoT devices association is formulated as one-to-one matching, the scheduling is many-to-one matching, and the power allocation is determined by geometric programming.

In our simulation, we set the radius of the cell as 500 m. The number of CUs, idle CUs, and IoT devices are set as $N = 15$, $Q = 10$, and $M = 300$, respectively. These CUs and IoT devices are uniformly distributed in the cell. The maximum transmit powers of each CU and each IoT device are set as $P^c = 23$ dBm and $P^{iot} = 13$ dBm, respectively. For simplicity, we assume that the data requirement for different IoT devices are the same, i.e., $d_1 = \ldots = d_M = d$. To satisfy the requirements for different applications, the value of d varies from 200 kbps to 1 Mbps. The minimum data rate requirement for a CU is set as $R_{min} = 1$ Mbps. We set the numbers of licensed and unlicensed subchannels as $K = 9 \sim 15$ and $K^u = K$. The small-scale fading is modeled as the Rayleigh fading. We adopt the path loss model in [38] where the decay factor $\alpha = 3.76$ and power gains factor $G = -17.7$ dB. Other simulation parameters are set according to [8], as listed in Table I.

Fig. 3(a) illustrates the number of scheduled IoT devices v.s. the number of licensed subchannels K with $d = 500$ kbps. From this figure, we can observe that the number of scheduled IoT devices grows as the number of licensed subchannels and the proposed IoT-U scheme outperforms IoT-L scheme. This implies that utilizing the unlicensed spectrum to support M2M communications can effectively mitigate the interference to the cellular links, and thus improving the system performance. In addition, compared to the DASPA and MASPA schemes, the IASPA algorithm in IoT-U scheme can schedule more IoT devices.

Fig. 3(b) shows the energy consumption v.s. the number of licensed subchannels K with $d = 500$ kbps. It is observed that the energy consumption increases as the number of licensed subchannels because more IoT devices are scheduled. Combining these two figures, we can infer that the TC network is the most energy efficient, that is, costing the least energy for one scheduled IoT device. This implies that the cluster based IoT networks can schedule more IoT devices at the cost of more energy for the CU relay. Moreover, we can find out

that the energy consumption of an IoT device in the IoT-U scheme is lower than that in the DASPA and MASPA schemes because the IASPA algorithm in the IoT-U scheme can utilize the resource more efficiently.

Fig. 4 illustrates how the data traffic influence the system performance with respect to the number of scheduled IoT devices and the IoT device association strategies. In Fig. 4(a), the number of scheduled IoT devices is evaluated for both IoT-L and IoT-U schemes with different data traffic of a IoT device. When the data traffic is low, a CU can aggregate the data traffic from multiple IoT devices. However, due to the subchannel sharing, the interference may be significant when a relatively large number of M2M links coexist on the same licensed subchannel. IoT-U scheme offloads M2M communications into the unlicensed band and alleviates the interference. Therefore, more IoT devices are scheduled. As the data traffic grows, the number of M2M links in a cluster decreases and the benefit from utilizing unlicensed subchannels for M2M communications also decreases. In addition, when the data traffic is low, i.e., $d = 200$ kbps, we can find out that the growth rate for IoT-U scheme with $K < 6$ is higher than that with $K \geq 6$. When $K < 6$, the number of scheduled IoT devices is bounded by the number of subchannels and it is bounded by the data rate of cellular links when $K \geq 6$, which are in accordance with Proposition 1 in Section VI-C.

Fig. 4(b) further evaluates how different data traffic influences the association strategy of IoT devices. We can easily observe that the proportion of scheduled IoT devices increases when the data traffic of each IoT devices becomes larger. Note that the data rate for the cellular link becomes the bottleneck when the data traffic for a IoT devices is high. Therefore, more IoT devices are associated with the BS using the orthogonal subchannels to avoid interference to other cellular links. This is consistent with Remark 1 in Section VI-C. Besides, we can also observe that the proportion of scheduled IoT devices associating with the BS obtained by IoT-U scheme is less than that obtained by IoT-L scheme. Since M2M links may bring severe interference to the cellular links in IoT-L scheme, the IoT devices located far from the CUs need to associate with the BS. However, in IoT-U scheme, the interference from M2M links can be mitigated by allocating the transmit power to the unlicensed subchannels, and thus, a cluster can support more IoT devices than that in IoT-L scheme.

Fig. 5 shows the convergence of the proposed IASPA algorithm with 12 licensed subchannels for different values of d. It can be observed that 11 iterations are needed when $d = 200$ kbps while it requires 7 iterations when $d = 1$ Mbps, which implies that higher data rate for an IoT device requires more iterations. When the data rate for one IoT device is high, the remaining transmit power of CU may not support the data from another IoT device. Therefore, the algorithm with a higher data rate requires more iterations.

VIII. CONCLUSIONS

In this paper, we proposed the IoT-U scheme for the cluster based IoT network where the unlicensed spectrum can be utilized for the M2M communications in clusters. We first
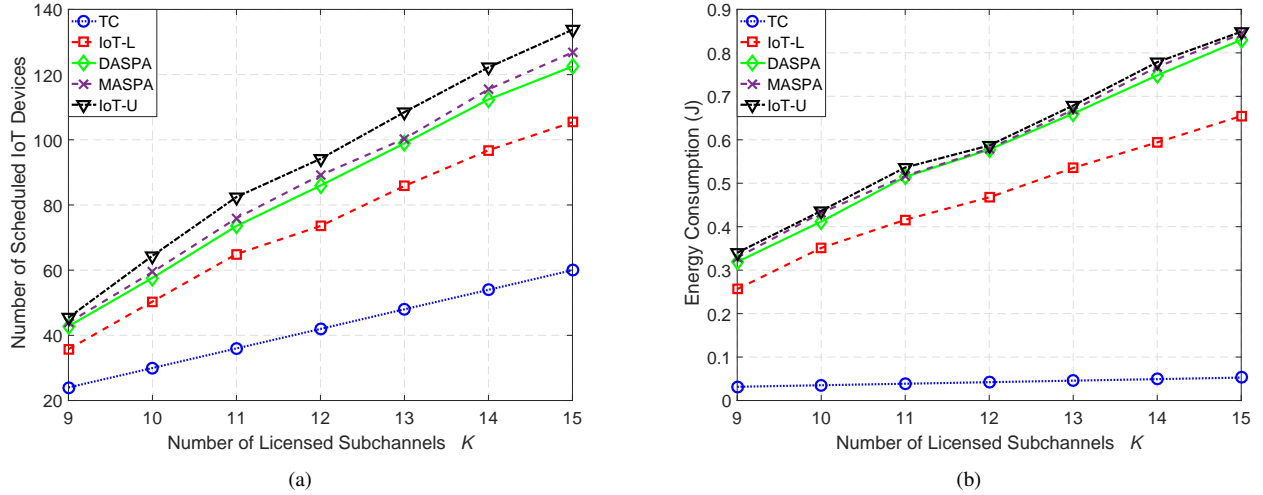
Fig. 3. (a) Number of scheduled IoT devices v.s. number of licensed subchannels; (b) Energy Consumption v.s. number of licensed subchannels with $d = 500$ kbps.
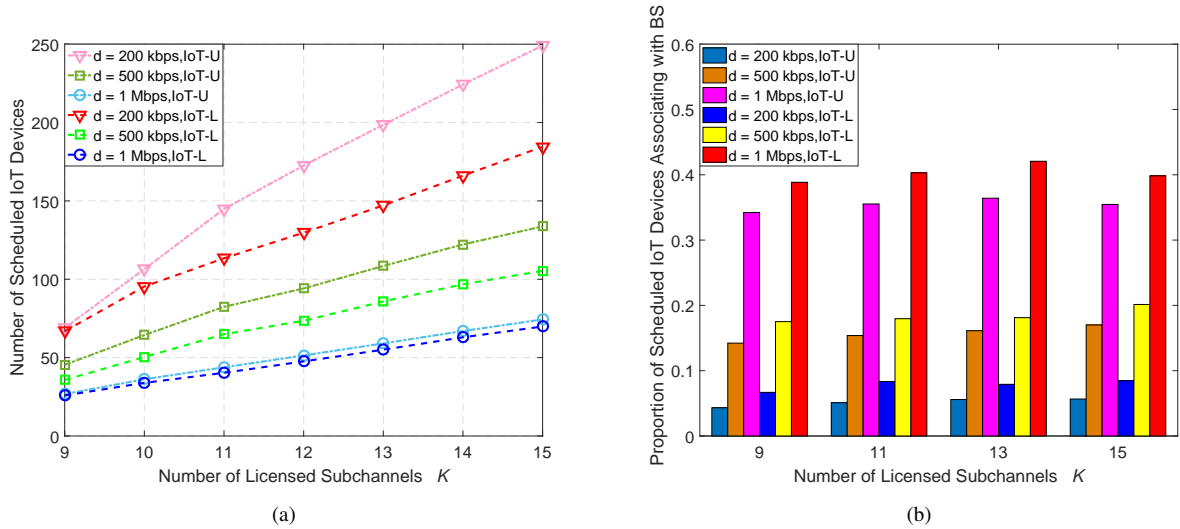

Fig. 4. (a) Number of scheduled IoT devices v.s. number of licensed subchannels; (b) Proportion of scheduled IoT devices associating with BS v.s. number of licensed subchannels.

designed a duty cycle based mechanism for fair unlicensed spectrum sharing, and then formulated an optimization problem to maximize the total weighted number of scheduled IoT devices with the minimum transmission energy consumption. Simulation results show that the proposed IoT-U scheme can support more IoT devices than the IoT-L one in one cycle. From analysis, we can obtain two important conclusions. First, since the IoT-U scheme offloads M2M communication into the unlicensed band and alleviates the co-channel interference, the CU can aggregate the data from more IoT devices in the unlicensed scheme. Second, the IoT devices are intended to associate with CUs when the data traffic is low. But when the data traffic grows, more IoT devices are inclined to associate with the BS to produce less interference.

APPENDIX A
PROOF OF THEOREM 1

Let $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$ be the optimal solution to problem (P3). We first prove that $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*)$ yields the maximum value of the objective in problem (P1), and then prove that $\boldsymbol{P}^*$ yields the minimum total transmission power in problem (P2).

To prove that $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*)$ yields the maximum value of the objective in problem (P1) is equivalent to prove that $\sum_{m \in \mathcal{M}} w_m \sum_{t \in \mathcal{T}} c_{m,t}^* = \sum_{m \in \mathcal{M}} w_m \sum_{t \in \mathcal{T}} \tilde{c}_{m,t}$, where $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{P}})$ is an optimal solution of problem (P1). Note that the constraints in problem (P3) is the same as that in problem (P1), and thus, the optimal solution of problem (P3) is also feasible to problem (P1). Suppose that $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}, \tilde{\boldsymbol{F}})$ yields a higher value of the objective than $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*,)$,
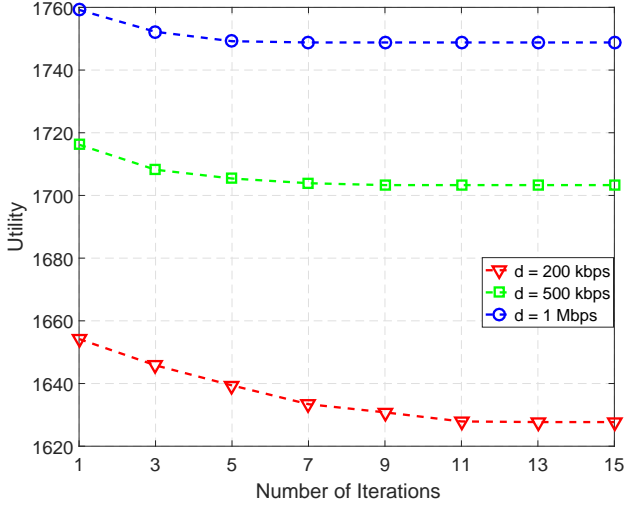
Fig. 5. Utility v.s. number of iterations with $K = 7$.

i.e.,

$$\sum_{m\in\mathcal{M}} w_m \sum_{t\in\mathcal{T}} \tilde{c}_{m,t} > \sum_{m\in\mathcal{M}} w_m \sum_{t\in\mathcal{T}} c_{m,t}^*. \tag{78}$$

By multiplying ρ to the both sides of (78), we can obtain

$$\sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} \tilde{c}_{m,t} > \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} c_{m,t}^*. \tag{79}$$

Since $c_{m,t}^*$, $\tilde{c}_{m,t}$, and λ_m are all positive integers, (79) can be tightened to

$$\sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} \tilde{c}_{m,t} - 1 \geq \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} c_{m,t}^*. \tag{80}$$

Note that $0 < \epsilon < \frac{1}{MP^{iot}+NP^c+1}$. Therefore, we can obtain the following inequality:

$$\sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon p_{m,n}^{k,t} + \sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon p_n^{k,t}$$
$$< \epsilon(MP^{iot} + NP^c) < 1 - \epsilon. \tag{81}$$

Thus, together with (80), the objective function values corresponding to $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$ and $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{P}})$ satisfy

$$\sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon \tilde{p}_{m,n}^{k,t} + \sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon \tilde{p}_n^{k,t}$$
$$+(1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1 - \tilde{c}_{m,t})$$
$$\leq \sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon \tilde{p}_{m,n}^{k,t} + \sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon \tilde{p}_n^{k,t}$$
$$+(1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1 - c_{m,t}^*) - (1-\epsilon)$$
$$< (1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1 - c_{m,t}^*). \tag{82}$$

We can observe that $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{P}})$ yields a smaller objective function value than $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$ does. Since $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}, \boldsymbol{F}^*, \tilde{\boldsymbol{P}})$ is also feasible to problem (P3), (82) violates the assumption that $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$ is the optimal solution of problem (P3). Therefore, we have

$$\sum_{m\in\mathcal{M}} w_m \sum_{t\in\mathcal{T}} \tilde{c}_{m,t} = \sum_{m\in\mathcal{M}} w_m \sum_{t\in\mathcal{T}} c_{m,t}^*. \tag{83}$$

Following the same approach, we can also prove that $\boldsymbol{P}^*$ yields the minimum total transmission power in problem (P2).

On the other hand, let $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$ be the optimal solution through solving problems (P1) and (P2) sequentially. We also prove that $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$ is the optimal solution to problem (P3). Define $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{P}})$ as the optimal solution of problem (P3). To prove that $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$ is the optimal solution to problem (P3) is equivalent to prove that $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$ can yield to the same value of U as $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{P}})$ does.

Note that the feasible set of problem (P2) is the optimal solution set of problem (P1). Therefore, $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$ yields the maximum value of the objective in problem (P1). Since the constraints in problem (P1) are the same as that in problem (P3), the optimal solution of problem (P1) is also feasible to problem (P3). Suppose that $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{P}})$ yields a lower value of the objective than $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$, that is,

$$\sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon \tilde{p}_{m,n}^{k,t} + \sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon \tilde{p}_n^{k,t}$$
$$+(1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1 - \tilde{c}_{m,t})$$
$$< \sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon (p_{m,n}^{k,t})^* + \sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon (p_n^{k,t})^*$$
$$+(1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1 - c_{m,t}^*). \tag{84}$$

Similarly, $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{P}})$ is also a feasible solution of problem (P1). If $\sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} \tilde{c}_{m,t} = \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} c_{m,t}^*$, then $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$ is also the optimal solution to problem (P3), because $\boldsymbol{P}^*$ yields to the minimum total transmission power. Therefore, we have

$$\sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} \tilde{c}_{m,t} < \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} c_{m,t}^*. \tag{85}$$

Since $c_{m,t}^*$, $\tilde{c}_{m,t}$, and λ_m are all positive integers, (85) can be tightened to

$$(1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1 - \tilde{c}_{m,t}) - (1-\epsilon) \geq$$
$$(1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1 - c_{m,t}^*). \tag{86}$$

Therefore, we have

$$(1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1 - \tilde{c}_{m,t}) - (1-\epsilon) +$$
$$\sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon (p_{m,n}^{k,t})^* + \sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon (p_n^{k,t})^*$$
$$\geq (1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1 - c_{m,t}^*) +$$
$$\sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon (p_{m,n}^{k,t})^* + \sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon (p_n^{k,t})^*$$
$$> \sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon \tilde{p}_{m,n}^{k,t} +$$
$$\sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon \tilde{p}_n^{k,t} + (1-\epsilon) \sum_{m\in\mathcal{M}} \lambda_m \sum_{t\in\mathcal{T}} (1 - \tilde{c}_{m,t}), \tag{87}$$

that is,

$$
\begin{aligned}
& \sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon(p_{m,n}^{k,t})^* + \\
& \sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon(p_n^{k,t})^* - (1-\epsilon) \\
> \ & \sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon\tilde{p}_{m,n}^{k,t} + \sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon\tilde{p}_n^{k,t} \\
\geq \ & 0.
\end{aligned}
\tag{88}
$$

Note that $0 < \epsilon < \frac{1}{MP^{iot}+NP^c+1}$. Therefore, we can obtain the following inequality:

$$
\begin{aligned}
& \sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}\cup\mathcal{K}^u} \sum_{n\in\mathcal{Q}\cup\{0\}} \epsilon(p_{m,n}^{k,t})^* + \sum_{n\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \epsilon(p_n^{k,t})^* \\
& < \epsilon(MP^{iot}+NP^c) < 1-\epsilon,
\end{aligned}
\tag{89}
$$

which violates (88). Therefore, $(\boldsymbol{A}^*, \boldsymbol{B}^*, \boldsymbol{C}^*, \boldsymbol{F}^*, \boldsymbol{P}^*)$ is also the optimal solution to problem (P3).

REFERENCES

[1] Mckinsey Global Institute, "The Internet of Things: Mapping the Value Beyond the Hype," Jun. 2015.

[2] M. Tao, K. Ota, and M. Dong, "Locating compromised data sources in IoT-enabled smart cities: A great-alternative-region-based approach," *IEEE Trans. Ind. Informat.*, vol. 14, no. 6, pp. 2579-2587, Jun. 2018.

[3] K. Ota, T. Kumrai, M. Dong, J. Kishigami, and M. Guo, "Smart infrastructure design for smart cities," *IT Professional*, vol. 19, no. 5, pp. 42-49, Sep. 2017.

[4] S. Feng, P. Setoodeh, and S. Haykin, "Smart Home: Cognitive Interactive People-Centric Internet of Things," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 34-39, Feb. 2017.

[5] U. Satija, B. Ramkumar, and M. S. Manikandan, "Real-Time Signal Quality-Aware ECG Telemetry System for IoT-Based Health Care Monitoring," *IEEE Internet Things J.*, vol. 4, no. 3, pp. 815-823, Jun. 2017.

[6] C. Brewster, I. Roussaki, N. Kalatzis, K. Doolin, and K. Ellis, "IoT in Agriculture: Designing a Europe-Wide Large-Scale Pilot," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 26-33, Sep. 2017.

[7] A. Hoglund, X. Lin, O. Liberg, A. Behravan, E. A. Yavuz, M. V. D. Zee, Y. Sui, T. Tirronen, A. Ratilainen, and D. Eriksson, "Overview of 3GPP Release 14 Enhanced NB-IoT," *IEEE Network*, vol. 31, no. 6, pp. 16-22, Nov. 2017.

[8] H. Zhang, Y. Liao, and L. Song, "D2D-U: Device-to-Device Communications in Unlicensed Bands for 5G System," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3507-3519, Jun. 2017.

[9] P. Wang, B. Di, H. Zhang, K. Bian, and L. Song, "Cellular V2X in Unlicensed Spectrum: Harmonious Coexistence with VANET in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5212-5224, Aug. 2018.

[10] R. Zhang, M. Wang, L. X. Cai, Z. Zeng, X. S. Shen, and L. Xie, "LTE-Unlicensed: The Future of Spectrum Aggregation for Cellular Networks," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 150-159, Jun. 2015.

[11] Z. Zhou, M. Dong, K. Ota, G. Wang, and L. T. Yang, "Energy-Efficient Resource Allocation for D2D Communications Underlaying Cloud-RAN-based LTE-A Networks," *IEEE J. Internet Things*, vol. 3, no. 3, pp. 428-438, Jun. 2016.

[12] F. Ramparany and Q. H. Cao, "A Semantic Approach to IoT Data Aggregation and Interpretation Applied to Home Automation," in *Proc. IOTA*, Pune, India, Jan. 2016.

[13] LoRa Alliance, "A Technical Overview of LoRa and LoRaWAN", Nov. 2015.

[14] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi, "Long-Range Communications in Unlicensed Bands: The Rising Stars in the IoT and Smart City Scenarios," *IEEE Wireless Commun.*, vol. 23, no. 5, pp. 60-67, Oct. 2016.

[15] T. Nihtilä, V. Tykhomyrov, O. Alanen, M. A. Uusitalo, A. Sorri, M. Moisio, S. Iraji, R. Ratasuk, and N. Mangalvedhe, "System Performance of LTE and IEEE 802.11 Coexisting on a Shared Frequency Band," in *Proc. IEEE WCNC*, Shanghai, China, Apr. 2013.

[16] Qualcomm Research, "LTE in Unlicensed Spectrum: Harmonious Coexistence with Wi-Fi," Jun. 2014.

[17] E. L. Lawler and D. E. Wood, "Branch-and-Bound Methods: A Survey," *Operations Research*, vol. 14, no. 4, pp. 699-719, Aug. 1966.

[18] J. Papandriopoulos and J. S. Evans, "SCALE: A Low-complexity Distributed Protocol for Spectrum Balancing in Multiuser DSL Networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711-3724, Aug. 2009.

[19] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.

[20] M. Darabi, B. Maham, W. Saad, A. Mehbodniya, and F. Adachi, "Joint Machine-Type Device Selection and Power Allocation for Buffer-Aided Cognitive M2M Communication," in *Proc. IEEE PIMRC*, Hong Kong, China, Sep. 2015.

[21] M. Darabi, A. Montazeri, B. Maham, W. Saad, and H. Zarrabi, "Packet Size Adjustment for Minimizing the Average Delay in Buffer-Aided Cognitive Machine-to-Machine Networks," *Elsevier J. Comput. Elect. Eng.*, vol. 68, pp. 298309, May 2018.

[22] X. Kang, Y.-K. Chia, S. Sun, and H. F. Chong, "Mobile Data Offloading Through A Third-Party Wi-Fi Access Point: An Operator's Perspective," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5340-5351, Oct. 2014.

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

[24] Nokia, "Nokia LTE for Unlicensed Spectrum," Jun. 2014.

[25] H. Zhang, X. Chu, W. Guo, and S. Wang, "Coexistence of Wi-Fi and Heterogeneous Small Cell Networks Sharing Unlicensed Spectrum," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 158-164, Mar. 2015.

[26] M. Ali, S. Qaisar, M. Naeem, and S. Mumtaz, "Joint User Association and Power Allocation for Licensed and Unlicensed Spectrum in 5G Networks," in *Proc. IEEE GLOBECOM*, Singapore, Singapore, Dec. 2017.

[27] B. Chen, J. Chen, Y. Gao, and J. Zhang, "Coexistence of LTE-LAA and Wi-Fi on 5GHz with Corresponding Deployment Scenarios: A Survey," *IEEE Commun. Surveys & Tutorials*, vol. 19, no. 1, pp. 7-32, First Quarter 2017.

[28] J. Yi, W. Sun, S. Park, and S. Choi, "Performance Analysis of LTE-LAA Network," *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1236-1239, Jun. 2018.

[29] 3GPP TR 36.746, "Study on Further Enhancements to LTE Device to Device (D2D), User Equipment (UE) to Network Relays for Internet of Things (IoT) and Wearables," Apr. 2017.

[30] B. Di, L. Song, and Y. Li, "Sub-Channel Assignment, Power Allocation, and User Scheduling for Non-Orthogonal Multiple Access Networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7686-7698, Nov. 2016.

[31] L. P. Qian, Y. J. Zhang, Y. Wu, and J. Chen, "Joint Base Station Association and Power Control via Benders' Decomposition," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1651-1665, Apr. 2013.

[32] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as Optimization Decomposition: A Mathematical Theory of Network Architectures," *Proc. IEEE*, vol. 95, no. 1, pp. 255-312, Jan. 2007.

[33] D. Li and X. Sun, *Nonlinear Integer Programming*. Boston, MA: Springer Verlag, 2006.

[34] A. Alvarado, G. Scutari, and J.-S. Pang, "A New Decomposition Method for Multiuser DC-Programming and Its Applications," *IEEE Trans. Signal Process.*, vol. 62, no. 11, pp. 2984-2998, Jun. 2014.

[35] S. Boyd, "Subgradient Methods", [Online]. Available: http://www.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf.

[36] A. Schrijver, *Theory of Linear and Integer Programming*, John Wiley & Sons, New Jersey, USA, 1998.

[37] T. LeAnh, N. H. Tran, W. Saad, L. B. Le, D. Niyato, T. M. Ho, and C. S. Hong, "Matching Theory for Distributed User Association and Resource Allocation in Cognitive Femtocell Networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 8413-8428, Sep. 2017.

[38] 3GPP TR 36.888, "Study on Provision of Low-cost Machine-Type Communications (MTC) User Equipments (UEs) based on LTE (Release 12)," Jun. 2013.

Hongliang Zhang (S'15) received the B. S. degree in Electronic Engineering from Peking University, Beijing, China, in 2014. He is currently pursuing his PhD's degree at School of Electrical Engineering and Computer Science in Peking University. His current research interest includes device-to-device communications, unmanned aerial vehicle networks, and hypergraph theory. He has also served as a TPC Member for GlobeCom 2016, ICC 2016, ICCC 2017, ICC 2018, and GlobeCom 2018.

Boya Di (S'17) is a PhD student at the School of Electrical Engineering and Computer Science, Peking University, China. She received the B.S. degree in electronic engineering from Peking University in 2014. Her current research interests include edge computing, vehicular networks, non-orthogonal multiple access, and 5G wireless networks. She has also served as a TPC member in GlobeCom 2016, ICCC 2017, ICC 2016 and ICC 2018.

Kaigui Bian (S'05, M'11) received his Ph.D. degree in Computer Engineering from Virginia Tech in 2011, and his B.S. degree in Computer Science from Peking University, Beijing, China in 2005. He was a Visiting Young Faculty in Microsoft Research Asia in 2013. He received the best paper awards of international conferences (IEEE ICC 2015, ICCSE 2017, BIGCOM 2018) and the best student paper award of IEEE DSC 2018. He was the recipient of IEEE Communication Society Asia-Pacific Board (APB) Outstanding Young Researcher Award in 2018. He serves as an Editor for IEEE Transactions on Vehicular Technology and IEEE Access, and the organizing committee member as well as technical program committee member of many international conferences. His research interests include wireless networking and mobile computing.

Lingyang Song (S'03-M'06-SM'12-F'19) received his Ph.D. from the the University of York, United Kingdom, in 2007, where he received the K. M. Stott Prize for excellent research. He worked as a research fellow at the University of Oslo, Norway, until rejoining Philips Research UK in March 2008. In May 2009, he joined the School of Electronics Engineering and Computer Science, Peking University, and is now a Boya Distinguished Professor. His main research interests include wireless communication and networks, signal processing, and machine learning. He was the recipient of the IEEE Leonard G. Abraham Prize in 2016 and the IEEE Asia Pacific (AP) Young Researcher Award in 2012. He has been an IEEE Distinguished Lecturer since 2015.