

# To Wait or To Pay: A Game Theoretic Mechanism for Low-Cost M2M and Mission-Critical M2M

Mei-Ju Shih, *Student Member, IEEE*, Kevin Dowhon Huang, Chia-Yi Yeh, and Hung-Yu Wei, *Member, IEEE*

**Abstract**—When it comes to Machine-to-Machine (M2M) communications in advanced cellular networks, the resource allocation scheme should be re-examined to satisfy both low-cost M2M and mission-critical M2M. Because most M2M applications are uplink-dominated, we propose a mixed waiting-time auction and price-based dedicated uplink resource allocation framework for the low-cost and mission-critical M2M. The prioritized framework guarantees resources for traditional Human-to-Human (H2H) communications while meeting the needs of low-cost and mission-critical M2M devices on the basis of either time bids or direct price. In addition, the scheme ensures the existence and uniqueness of the Bayesian Nash equilibrium and the interregional and waiting-time-based truth-telling properties. This indirect mechanism holds with Bayesian-Nash incentive compatibility, interim efficiency, interim individual rationality and weakly budget balance. The results show that low-cost M2M devices with lower energy awareness are more willing to participate in the waiting-time auction, while mission-critical M2M with higher energy awareness turn to directly pay for guaranteed access. The delay in connected mode and the optimal price vary according to the M2M/H2H traffic loads and resource pool partitions. This work contributes insights that with proper mechanism design, low-cost M2M and mission-critical M2M can be served together, while the operator is financially compensated.

**Index Terms**—Machine-Type Communications (MTC), Machine-to-Machine (M2M) communications, Human-to-Human (H2H) communication, resource allocation, cellular network, game theory, mechanism design

## I. INTRODUCTION

M2M communications a.k.a. Machine-Type Communications (MTC), featuring as a wide range of autonomous devices to communicate wirelessly without human intervention, has become one of the most compelling requirements in the future communications systems. 5G is envisioned to satisfy the versatile massive M2M traffic characteristics, from best effort applications like water/gas metering systems and environmental monitoring, to ultra reliable ones such as healthcare, public safety and mission critical industry [1]. Thus, how to serve *low-cost M2M* and *mission-critical M2M* becomes significant in 5G. Low-cost M2M is also known

as massive MTC (mMTC) with low power consumption and extensive coverage [2], while mission-critical M2M as ultra-reliable MTC (uMTC) and low-latency M2M with latency and reliability as two key elements [3]. Before, the cellular network was designed for traditional H2H communications (e.g., voice calls), with monotonous and lasting traffic such as voice call and web browsing. The traffic types of H2H communications are asymmetrically downlink, whereas those of M2M communications are predominantly uplink-oriented. Due to the emerging diversified applications, a new M2M-oriented Quality-of-Service (QoS) categorization with eight classes is proposed to cover both H2H and varied M2M services [4]. It is an important but challenging issue on providing low-cost M2M and mission-critical M2M communications to cater for customized services.

Energy is another key aspect in M2M scenarios. Some M2M devices such as low-cost M2M requires energy efficiency and low power consumption for battery lifetime extension, whereas others such as mission-critical M2M requires reliable and delay sensitive traffic delivery at the cost of comparably higher power consumption. Thus, the degree to which M2M devices value their energy/traffic differs from devices to devices, and even from time to time. For example, several factors influence how M2M devices attach importance to energy, such as their residual energy levels, the power consumption, the difficulties to acquire energy and the electricity price. The literature focuses on energy efficiency and battery lifetime extension without considering the diversified M2M energy awareness. Energy efficiency maximization problem under statistical QoS guarantees is addressed for uplink resource allocation in M2M/H2H co-existence LTE networks [5]. Power efficient MAC such as coordinated/uncoordinated CDMA/FDMA/TDMA for cellular-based M2M are compared [6]. In this article, we capture *the energy awareness level* as a crucial factor for a variety of M2M devices. Low-cost M2M and mission-critical M2M are viewed to have different energy awareness levels.

One of the ways that M2M devices connect to the core network is by being equipped with cellular capability, called *Cellular M2M* [7]. Energy efficiency, scalability and cost effectiveness are essential for M2M MAC protocol issues [8]. The 3<sup>rd</sup> Generation Partnership Project (3GPP) has devoted efforts to M2M communications through infrastructure-based access networks and mainly focuses on scalability and cost effectiveness [2] [9] [10]. Recently, 3GPP begins to study the feasibility of using direct links for use cases such as vehicular systems [11] and IoT wearable devices [12]. Since the majority of M2M traffic is uplink-based, the procedure to perform

Manuscript received Feb. 10, 2016; revised Jun. 4, 2016; accepted Aug. 2, 2016. Date of current version Aug. 8, 2016. Corresponding author: Hung-Yu Wei.

Mei-Ju Shih is with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan (e-mail: d00942022@ntu.edu.tw).

Kevin Dowhon Huang and Chia-Yi Yeh were with the Department of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan (e-mail: solakh2@gmail.com; redwind191@gmail.com).

Hung-Yu Wei is with the Graduate Institute of Communication Engineering and Department of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan (e-mail: hywei@cc.ee.ntu.edu.tw).

uplink transmission in cellular networks is investigated [13]. Generally, the devices stay in idle mode to conserve energy if there is no queued uplink traffic. Once the uplink traffic arrives, the devices execute the network entry procedure in the control plane to acquire dedicated uplink resources. Therefore, the devices enter connected mode and uplink data to the evolved Node B (eNB) without collision in the user plane. The well-known network entry procedure in LTE-A is the RACH (Random Access Channel) procedure with four-message exchange steps: *Preamble Transmission*, *Random Access Response*, *Connection Request*, and *Connection Resolution*. When the same preamble sequence is transmitted on the same time-frequency resource due to massive M2M attempts, collision occurs in the *RRC Connection Request* step. The state of the art focuses on RACH overload congestion [13] such as RACH preamble detection and allocation [14], access class barring (ACB) [15], enhanced access barring (EAB) [16] and combination of them [17] [18]. Whether the RACH procedure is appropriate for M2M communications is comprehensively discussed [7]. The literature explores the suitability and modification of the network entry procedure to satisfy massive M2M, while leaves the uplink resource allocation for co-existence of low-cost M2M and mission-critical M2M in the user plane to be unaddressed.

The investigation of cellular networks with M2M/H2H co-existence has attracted attention recently, especially the preamble pool partition [19] [20]. The preferred partition: two disjoint subsets for H2H and M2M separately, or one small subset for both H2H and M2M and the remaining subset for H2H only, depends on the random access load [19]. Another game theoretic analysis is conducted for three disjoint small pools: one for H2H only, another for M2M only, and the other for hybrid M2M/H2H usage [20]. Queueing modeling and coalitional game analysis for M2M RACH coexisting with H2H is proposed [21]. So far, research on M2M/H2H co-existence has focused on the network entry procedure. Few researchers have investigated the impact of M2M/H2H co-existence on the user plane for dedicated uplink data transmission.

In this work, we aim to enable low-cost M2M and mission-critical M2M by designing a dedicated uplink resource allocation framework, considering devices' diversified energy-awareness levels and the procedure from idle mode, network entry to connected mode. In our preliminary work [22], we explored the M2M devices' energy-awareness properties in a proposed waiting-line auction resource allocation framework, which properly captures the LTE uplink procedure from the network entry to uplink data transmission. A waiting-line auction allows players to bid the scarce resources with time instead of money [23]. For example, this mechanism is adopted for the dynamic spectrum allocation in cognitive radio networks [24]. In this article, we further design the price-based resource pool to support mission-critical M2M, while keep low-cost M2M with lower priority to share the waiting-time-based resource pool with H2H. M2M devices can either bid the uncertain resources with time or pay for direct access. Thus, a prioritized resource allocation framework forms as shown in Fig. 1, which is modeled as Bayesian Stackelberg game with

waiting-time auction. We further investigate the properties of this framework through mechanism design analysis. This work contributes insights on the relationship between the resource allocation and devices' energy awareness. The energy-awareness property is modeled as a factor, *energy opportunity cost*, which means the value per time from the aspect of energy. A device with higher energy opportunity cost means that it values the spent time more, i.e., mission-critical and reliable transmission requirement. We exploit the way that the energy awareness plays a significant role in designing a mixed low-cost M2M and mission-critical M2M network. The price and waiting time vary according to M2M/H2H traffic loads and resource pool partitions.

The major contributions of this paper are as follow.

- A *prioritized resource allocation framework*, viewed as an *indirect mechanism design mixed with time and money*, is designed to support both *low-cost M2M* and *mission-critical M2M*.
- The framework presents the device heterogeneity by *energy awareness levels*, i.e., low-cost M2M with low energy awareness, and mission-critical M2M with high energy awareness.
- The framework is modeled as *Bayesian Stackelberg game with waiting-time auction*, where the eNB is the leader and auctioneer, and M2M devices are followers and bidders.
- Except for the eNB's announcement, the framework does not introduce extra signaling exchange compared to the LTE standard.
- The framework achieves *unique Bayesian Nash equilibrium*, and *interregional and waiting-time-based truth-telling*. The indirect mechanism holds with *Bayesian Nash incentive compatibility*, *interim efficiency*, *interim individual rationality and weakly budget balance*.
- The design of mission-critical M2M resource pool compensates the operator financially, with the price determined by the *eNB's expected revenue maximization*.
- This work implies that the operator can design *dynamic prices* according to M2M/H2H traffic loads and resource pool partitions, considering *the tradeoff between the expected revenue and the expected delay in connected mode*.

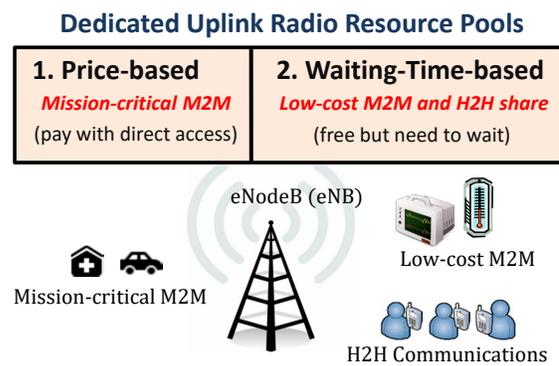


Fig. 1. An uplink cellular network with low-cost M2M, mission-critical M2M and H2H

## II. SYSTEM MODEL

### A. Problem Statement

We consider a cellular network with one eNB to serve two types of wireless apparatuses: H2H devices and M2M devices. H2H devices mainly intend for traditional human-to-human activities such as real-time voice calls, while M2M devices perform a variety of applications, namely from low-cost M2M to mission-critical M2M. However, the current uplink system is originally designed for H2H communications. When M2M devices begin to pop up, they share the dedicated uplink radio resources with H2H devices, which may not satisfy their diversified requirements, e.g., best effort, stringent latency and reliability. In this work, we propose a time-and-money hybrid prioritized resource allocation framework to not only guarantee the H2H priority but also fulfill the prerequisite of mixed low-cost M2M and mission-critical M2M.

### B. Proposed Prioritized Resource Allocation Framework

Here we design an innovative waiting-time-based and price-based resource allocation framework for distributing the deficient dedicated uplink radio resources to numerous rational M2M devices, while guaranteeing H2H communications as the traditional approach. According to the traditional uplink procedure, M2M devices wake up from idle mode to perform network entry procedure, and then enter connected mode for dedicated uplink resources. In the proposed scheme, M2M devices, upon completing the network procedure, can either directly acquire the resources in the price-based resource pool by paying, or line up in connected mode to compete for the free yet uncertain remaining resources in the waiting-time-based resource pool. The eNB designs the direct access price  $p_0$  per unit of resource based on the expected number of sold resources  $R_P$  in order to maximize its expected revenue.

First, the eNB prepares two independent resource pools: one with  $R_T$  units of resources for the waiting-time-based scheme, and another with  $\bar{R}_P$  units of resources for the price-based scheme. The total number of dedicated uplink resources is bounded by  $R_T + \bar{R}_P$ , where  $R_T, \bar{R}_P \in \mathbb{N}$ . The H2H devices use the waiting-time-based resource pool with the highest priority without extra payment, which is the same as the prevailing cellular system. Secondly, the eNB announces the price  $p_0$ , the size of waiting-time-based resource pool  $R_T$ , the anticipated H2H traffic load  $\lambda$ , and the reward time  $T$ , which is the time to allocate dedicated uplink resources for M2M communications. We assume that all devices experience the network entry procedure with an expected time  $k$  [25], send uplink scheduling requests, and enter connected mode for dedicated uplink resource allocation. Based on the scheduling requests from H2H devices, the eNB allocates  $R_H$  units of resources to the H2H communications timely, and then at the reward time  $T$  distributes the remaining  $R_W = R_T - R_H$  units of resources to the M2M communications, where  $R_W, R_H \in \mathbb{N}$ . Since the H2H traffic is mostly for voice calls,  $R_H$  is modeled as a random variable following  $Poisson(\lambda)$  according to the arrival of voice calls. The complete procedure from the network entry to the dedicated uplink resource allocation for

M2M devices is depicted in Fig. 2. The detailed notations can be referred to Table I.

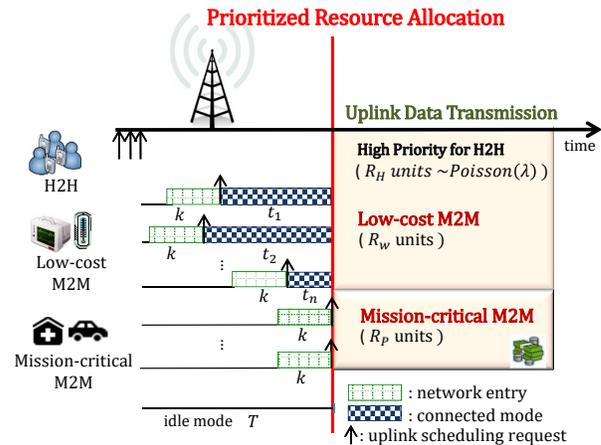


Fig. 2. A prioritized resource allocation framework: (1) pay with direct access (2) waiting-time auction

For M2M devices to acquire the dedicated uplink resources, the proposed framework is composed of (1) pay for guaranteed resource approach and (2) intuitive first-come first-served waiting-line approach. We assume that each M2M device requests one unit of resource because the M2M traffic also features as small data transmission. If the M2M device chooses the first approach, it pays  $p_0$  to the eNB for a unit of resource in the price-based resource pool. On the other hand, if the M2M device selects the second approach, the eNB distributes the remaining resources in the waiting-time-based resource pool to the M2M device in a first-come first-served manner. The earlier the M2M device lines up in connected mode, with higher probability the device wins the resource. That is, each M2M device decides to pay for pledged resources, line up in connected mode to compete for unsure resources, or give up to stay in idle mode. Except that the eNB broadcasts the announcement, the proposed framework does not induce extra signaling exchange compared to the LTE standard. M2M devices do not send extra signaling to reveal their choices. The eNB notices the devices' arrival time in connected mode upon receiving the network entry complete message. Regarding of paying with direct access, the devices can indicate its willingness to pay by reusing the network entry complete message, which already exists in the prevailing system.

The first-come first-served principle provides fairness among M2M devices at the sacrifice of energy consumption due to waiting time in connected mode. We model the M2M devices' energy-awareness property as *energy opportunity cost*  $w_i$  for device  $i$ . The value of  $w_i$  reflects device  $i$ 's value per time (i.e., opportunity cost) from the aspect of energy. *Energy opportunity cost*  $w_i$  is defined as an increasing value function of transmission power  $P_i$ , i.e.,  $w_i = \alpha_i(P_i)$ . The value function  $\alpha_i(\cdot)$  reflects how importance that device  $i$  attaches to energy.  $\alpha_i(\cdot)$  and  $P_i$  vary from devices to devices. Low-cost M2M devices tend to have lower energy opportunity costs owing to the requirements of low power consumption and energy efficiency. On the other hand, mission-critical

TABLE I  
MATHEMATICAL NOTATIONS AND DEFINITION

Notation	Definition
$N$	Total M2M device number
$R_T$	Total units of dedicated uplink resources in waiting-time-based resource pool
$R_H$	Dedicated uplink resources allocated for H2H communications, $R_H \sim \text{Poisson}(\lambda)$
$R_W$	Dedicated uplink resources allocated for M2M communications for free, $R_W = R_T - R_H$
$\overline{R_P}$	Upper bound of dedicated uplink resources in the price-based resource pool
$R_P$	Expected number of sold dedicated uplink resources, $R_P \leq \overline{R_P}$
$T$	Reward time, $T > 0$
$k$	Expected network entry time, $k > 0$
$w_i$	Energy opportunity cost of M2M device $i$ , $w_i \sim G_W(w_i)$
$t_i$	Waiting time in connected mode of M2M device $i$
$t_{(j)}$	The $j^{\text{th}}$ -largest waiting time among all M2M devices
$p_0$	The direct access price per unit of resource
$v(w_i)$	Utility function of obtaining a unit of resource
$a_i = \frac{v(w_i)}{w_i}$	Time-valued utility function of obtaining a unit of resource, $a_i \sim G_A(a_i)$
$a_{(j)}$	The $j^{\text{th}}$ -largest time-valued utility among all M2M devices
$G_{A(j)}(a)$	Cumulative distribution function for the order statistic of rank $j$ among $N - 1$ independent drawings from the CDF $G_A(a)$
$\pi_i^P$	M2M device $i$ 's payoff in time units when choosing to pay
$\pi_i^W$	The M2M winner's payoff in time units when choosing to wait
$\pi_i^L$	The M2M loser's payoff in time units when choosing to wait
$\pi_i^d$	M2M device $i$ 's payoff in time units when choosing to be idle
$P_i^W = F(a_i)$	The resource accessibility probability of M2M device $i$ , with PDF $f(a_i)$
$\sigma(a)$	Waiting-time equilibrium strategy function
$a^*$	The condition that motivates an M2M device to participate in the waiting-time auction

M2M devices are inclined to have higher energy opportunity costs with comparably higher power consumption to deliver delay sensitive mission-critical traffic. Without loss of the generality, we assume that  $w_i$  follows a general distribution with CDF (Cumulative Distribution Function)  $G_W(w_i)$  and PDF (Probability Density Function)  $g_W(w_i)$ . The exact value  $w_i$  is private information only known to device  $i$  itself.

### III. PROBLEM FORMULATION

#### A. Stackelberg Game with Incomplete Information

The proposed mechanism is formulated as Bayesian Stackelberg game  $\mathcal{G} = (\mathcal{N}, \Theta, \mathcal{S}, \mathcal{U})$ , summarized in Table II, where  $\mathcal{N} = \{0, 1, \dots, N\}$  is the player set,  $\Theta = \{\theta_i\}_{i \in \mathcal{N}}$  is the type space,  $\mathcal{S} = \{s_i\}_{i \in \mathcal{N}}$  is the strategy space, and  $\mathcal{U} = \{u_i\}_{i \in \mathcal{N}}$  is the payoff space. The eNB, i.e.,  $i = 0$ , as the Stackelberg leader, specifies the price  $p_0$  to maximize its expected revenue for

introducing the reliable transmission. The M2M devices, i.e.,  $i \in \mathcal{N} \setminus \{0\}$ , as the Stackelberg followers, choose whether to ask for resources and the way of obtaining the resources, i.e., either pay or wait, based on the rule specified by the eNB. H2H communications are always guaranteed with highest priority without the extra payment. The Stackelberg game with the waiting-time auction is a form of Bayesian game because the players do not know others' types, i.e., followers' energy opportunity costs and the leader's exact remaining resource number. The detailed game formulation is described in the following subsections.

#### B. Knowledge of Information

In a game with incomplete information, the players' private information is their types, which are known to the players themselves. The eNB's type is the exact amount of remaining waiting-time-based resources  $R_w$  for M2M communications, and the M2M devices' types are their energy opportunity costs  $w_i$ . M2M devices consider that the number of remaining waiting-time-based resources is a random variable  $R_W = R_T - R_H$ , where  $R_W$  follows a probability mass function (PMF)  $P_{R_W}(R_w)$ . All players have a common belief that energy opportunity cost  $w_i$  follows a general distribution with CDF  $G_W(w_i)$  and PDF  $g_W(w_i)$ , and the number of resources for H2H  $R_H$  follows PMF  $\text{Poisson}(\lambda)$ . The utility function of obtaining one unit of resource,  $v(w_i)$ , is assumed to be a continuously differentiable and positively valued non-decreasing concave function of energy opportunity cost. Except for the players' types, it is assumed that each player has the complete information of the system parameters such as  $N$ ,  $k$ ,  $T$ ,  $R_T$ ,  $v(w_i)$  and  $\lambda$ . The notations are summarized in Table I.

#### C. The eNB's Payoff

The eNB prepares two independent resource pools for M2M communications. The total units of waiting-time-based resources are predetermined to be  $R_T$ , and the maximum size of price-based resources are  $\overline{R_P}$ . Since the eNB has to satisfy the undetermined amount of mission-critical M2M devices, the price  $p_0$  is decided in advance to maximize the eNB's expected revenue. After the price is announced, each M2M device decides to accept the price for direct access or participate in the waiting-time auction. Upon the eNB receiving the uplink scheduling requests from the H2H devices, it allocates resources to them as the existing system. At the reward time  $T$ , the remaining waiting-time-based resources are allocated to M2M devices in a first-come first-served manner. The strategy set of the eNB is  $s_0 = \{p_0 : p_0 \in \mathbb{R}^+\}$ . The payoff set of the eNB is  $u_0 = \{p_0 R_P : p_0 \in s_0, R_P \in \mathbb{N}, R_P \leq \overline{R_P}\}$ , where  $R_P$  is the expected number of allocated resources for direct access. The expected revenue maximization problem can be formulated as

$$\begin{aligned} \underset{s_0}{\text{maximize}} \quad & \mathbb{E}(U_0) = \underset{p_0 \in \mathbb{R}^+, R_P(p_0)}{\text{maximize}} \quad p_0 R_P, \\ \text{subject to} \quad & 0 \leq R_P \leq \overline{R_P}, R_P \in \mathbb{N}, p_0 \in \mathbb{R}^+. \end{aligned} \quad (1)$$

TABLE II  
STACKELBERG GAME WITH WAITING-TIME AUCTION

Player set $\mathcal{N} = \{0, \dots, N\}$	Type space $\Theta = \prod_{i=0}^N \theta_i$	Strategy space $\mathcal{S} = \prod_{i=0}^N s_i$	Utility space $\mathcal{U} = \prod_{i=0}^N u_i$
One eNB $i = 0$	$\theta_0 = \{R_w : 0 \leq R_w \leq R_T\}$	$\mathcal{S}_0(\theta_0, \theta_{-0}) = \{p_0 : p_0 \in \mathbb{R}^+\}$	$u_0 = \{p_0 R_P : p_0 \in \mathcal{S}_0, R_P \in \mathbb{N}\}$
M2M devices $i \in \mathcal{N} \setminus \{0\}$	$\theta_i = \{w_i : w_i \in \mathbb{R}^+\}$	$\mathcal{S}_i(\theta_i, \theta_{-i}) = \{p_0, t_i : t_i \geq 0\}$	$u_i \in \{U_i^p, \mathbb{E}[U_i^t], U_i^d\}$ and $U_i^t \in \{U_i^w, U_i^l\}$

#### D. M2M Devices' Payoff

M2M devices decide whether to pay announced price  $p_0$  for guaranteed resources, to participate in the waiting-time auction with time bid  $t_i > 0$ , or to stay idle with waiting time  $t_i = 0$ . The strategy set is  $s_{i \in \mathcal{N} \setminus \{0\}} = \{p_0, t_i : t_i \geq 0\}$ .

1) *Price-based resource allocation*: If one determines to pay  $p_0$ , it obtains the pledged resource and has more time in idle mode for power saving, as shown in Fig. 2. The payoff of choosing to pay is  $U_{i \in \mathcal{N} \setminus \{0\}}^p(p_0) = v(w_i) - p_0 - w_i k + w_i(T - k)$ , where  $w_i k$  is the cost of executing the network entry procedure and  $w_i(T - k)$  is the utility of saving power in idle mode. The price  $p_0$  has an impact on the devices' willingness to directly pay.

2) *Waiting-time-based resource allocation*: The waiting-time-based resource allocation is also viewed as a non-monetary auction with sealed time bids. The bidders bid for the uncertain amount of remaining resources, i.e.,  $R_w = R_T - R_H$ , with their waiting time  $t_i$  in connected mode. The bidder  $i$  may win the resource with probability  $P_i^w$ , defined as resource accessibility probability. However, the bidder may also lose with probability  $P_i^l$  due to limited resources and a large amount of competitors ( $N \gg R_T$ ). Thus, the payoff of participating in the waiting-time-based resource allocation is  $U_i^t \in \{U_i^w, U_i^l\}$ . The winner's payoff is  $U_{i \in \mathcal{N} \setminus \{0\}}^w(t_i, t_{-i}) = v(w_i) - w_i t_i - w_i k + w_i(T - k - t_i)$ , where  $t_{-i}$  represents the time bids from other devices,  $w_i k$  is the cost of executing the network entry procedure,  $w_i t_i$  is the cost of waiting in connected mode, and  $w_i(T - k - t_i)$  is the utility of saving power in idle mode. The loser's payoff is  $U_{i \in \mathcal{N} \setminus \{0\}}^l(t_i, t_{-i}) = -w_i t_i - w_i k + w_i(T - k - t_i)$ . The expected payoff is

$$\begin{aligned} \mathbb{E}[U_{i \in \mathcal{N} \setminus \{0\}}^t(t_i, t_{-i})] &= U_i^w(t_i, t_{-i})P_i^w + U_i^l(t_i, t_{-i})P_i^l \\ &= v(w_i)P_i^w - 2w_i(k + t_i) + w_i T. \end{aligned} \quad (2)$$

M2M device  $i$  chooses an optimal waiting time with the aim of maximizing its expected payoff in the waiting-time auction.

$$\begin{aligned} &\underset{t_i}{\text{maximize}} \quad \mathbb{E}[U_{i \in \mathcal{N} \setminus \{0\}}^t(t_i, t_{-i})] \\ &\text{subject to} \quad R_T = R_H + R_w, R_H \sim \text{Poisson}(\lambda), \\ & \quad t_i > 0, k > 0, T > 0, N \gg R_T \end{aligned} \quad (3)$$

The device's payoff is affected by the resource accessibility probability and energy cost. The resource accessibility probability relies on the number of available resources  $R_w$  and the competitor number  $N$ . The longer the device waits, the higher its resource accessibility probability, albeit with the sacrifice of higher energy cost. The cost depends on both the waiting duration and the energy opportunity cost.

3) *Stay idle instead of asking for resources*: If device  $i$  decides not to acquire the resource at the reward time, it spends all the time  $T$  in idle mode for power saving. It does not

enter the network entry procedure, which means it neither pays nor waits in connected mode. The payoff of not requesting resources is  $U_{i \in \mathcal{N} \setminus \{0\}}^d(t_i = 0) = w_i T$ .

4) *Payoff maximization problem*: The payoff space of M2M devices can be formulated as  $\mathcal{U} = \prod_{i=1}^N u_i$ , where  $u_i \in \{U_i^p, \mathbb{E}[U_i^t], U_i^d\}$  and  $U_i^t \in \{U_i^w, U_i^l\}$ . The goal of each device is to select an optimal strategy  $s_i$  to maximize its payoff. Several factors should be considered for decision making such as the announced price, the remaining free resources in the waiting-time auction, and their energy opportunity costs. The payoff maximization problem is formulated as

$$\begin{aligned} &\underset{s_i}{\text{maximize}} \quad U_{i \in \mathcal{N} \setminus \{0\}}(s_i) = \underset{p_0 \in \mathbb{R}^+, t_i \geq 0}{\text{maximize}} \left( U_i^p, \max \mathbb{E}[U_i^t(t_i, t_{-i})], U_i^d \right) \\ &\text{subject to} \quad R_T = R_H + R_w, R_H \sim \text{Poisson}(\lambda), \\ & \quad R_P \leq \overline{R_P}, R_T, R_H, R_w, R_P \in \mathbb{N}, \\ & \quad t_i > 0, k > 0, T > 0, p_0 > 0, N \gg R_T. \end{aligned} \quad (4)$$

#### IV. BACKWARD INDUCTION OF WAITING-TIME-BASED BAYESIAN STACKELBERG GAME

In this section, we adopt the backward induction to analyze the proposed framework. Backward induction is a common and appropriate approach to deal with the Stackelberg game by first analyzing the followers' game under the assumption that the leader has announced its strategy. The followers' game is mainly composed of the waiting-time auction and the conditions to participate in. With the understanding or the followers' optimal strategies, the leader can thereby optimize its payoff and determine the price.

##### A. Followers' Game

To begin with the backward induction, we assume the eNB has announced the direct access price  $p_0^*$  and the resource allocation rule. The M2M devices aim to find their optimal strategy function  $s_i^*$  to maximize their payoff in Eq. (4).

1) *Resource pools*: The eNB provides two independent resource pools, i.e., price-based and waiting-time-based. If the M2M device accepts the direct access price  $p_0^*$ , it is guaranteed to have one unit of resource from the price-based resource pool. So, it does not care about the size of price-based resource pool. Regarding to the waiting-time-based resource pool, the size is known as  $R_T$  units. But, after the resource allocation for H2H, the exact remaining amount is unknown. Nevertheless, the distribution of H2H resource demand is assumed to be a

common knowledge, i.e.,  $R_H \sim \text{Poisson}(\lambda)$ . So, the PMF of the remaining resources,  $R_W = R_T - R_H$  is derived as

$$P_{R_W}(R_W) = \begin{cases} \sum_{l=R_T}^{\infty} \frac{\lambda^l e^{-\lambda}}{l!} & R_W = 0 \\ \frac{\lambda^{(R_T - R_W)} e^{-\lambda}}{(R_T - R_W)!} & 1 \leq R_W \leq R_T \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

All players have a common belief on the distribution of resources in the auction  $P_{R_W}(R_W)$ .

2) *Time-valued payoff*: The M2M devices have an option to gain the resources by waiting in connected mode and being served according to their waiting time. That is, the more time they spend in connected mode, the more likely that they win the resources. But, the waiting time they are willing to spend rely on how time-valuable the resources are. Thus, we transform the payoff to be in time units. The payoff profile in time units is defined as  $\pi_i \in \{\pi_i^p, \pi_i^t, \pi_i^d\}$ ,  $\pi_i^t \in \{\pi_i^w, \pi_i^l\}$ ,  $i \in \mathcal{N} \setminus \{0\}$ . In the waiting-time auction, the winner's payoff is  $\pi_i^w$  and the loser's payoff is  $\pi_i^l$ . The time-valued payoff  $\pi_i$  is derived by dividing the monetary payoff  $U_i$  by  $w_i$ , i.e.,  $\pi_i = \frac{U_i}{w_i}$ . A new attribute  $a_i = \frac{v(w_i)}{w_i} = h(w_i)$  is defined as device  $i$ 's utility on a resource in time units (time-valued utility). The time-valued payoff functions when M2M devices pay ( $\pi_i^p$ ), wait ( $\pi_i^t \in \{\pi_i^w, \pi_i^l\}$ ), and be idle ( $\pi_i^d$ ) are derived as follows.

$$\begin{aligned} \pi_i^p &= \frac{U_i^p(p_0, s_{-i})}{w_i} = a_i - \frac{p_0}{w_i} - k + (T - k) \\ \pi_i^w &= \frac{U_i^w(t_i > 0, s_{-i})}{w_i} = a_i - k - t_i + (T - k - t_i) \\ \pi_i^l &= \frac{U_i^l(t_i > 0, s_{-i})}{w_i} = -k - t_i + (T - k - t_i) \\ \pi_i^d &= \frac{U_i^d(t_i = 0, s_{-i})}{w_i} = T \end{aligned} \quad (6)$$

The CDF of the attribute  $a_i$  is denoted as  $G_A(a_i)$ , and its PDF as  $g_A(a_i)$ . Under the assumption that  $w_i$  for all M2M devices  $i$  follows the same distribution,  $a_i$  of each M2M device has the same distribution  $G_A(a_i)$  as well. In addition,  $a_i = h(w_i)$  is a monotonically decreasing function because  $v(w_i)$  is a continuously differentiable and positive-valued non-decreasing concave function and  $w_i$  is always positive. The PDF of the attribute  $a_i$  can be derived in Eq. (7) by univariate transformation. Since the energy opportunity cost  $w_i$  is private, the time-valued utility function  $a_i$  is also the private information. All players have a common belief on  $g_A(a_i)$ .

$$g_A(a_i) = g_W(h^{-1}(a_i)) \left| \frac{dh^{-1}(a_i)}{da_i} \right| \quad (7)$$

If an M2M device  $i$  participates in the waiting-time auction, the best response is an optimal waiting time in connected mode, i.e.,  $t_i$ , aiming to maximize the expected time-valued payoff  $\mathbb{E}[\pi_i^t(t_i, t_{-i})] = P_i^w \pi_i^w + P_i^l \pi_i^l = a_i P_i^w - 2(k + t_i) + T$ . The waiting-time strategy function is determined based on the device's own private information  $a_i$ . With higher time-valued utility on a resource, the device is expected to line

up earlier. Thus, the waiting-time strategy function is defined as  $t_i = \sigma(a_i, a_{-i})$ , which must be positive-valued, strictly increasing, and differentiable.

3) *Resource accessibility probability  $P_i^w$* : Resource accessibility probability means the probability that an M2M device wins the resource if it lines up. This probability plays a key role in determining the expected payoff in the waiting-time auction, as indicated in Eq. (2). There is a tradeoff between the waiting time and the resource accessibility probability. If the device lines up earlier, the cost of time increases but the resource accessibility probability also rises. In other words, the order of waiting time is essential to determine the resource accessibility probability because the waiting-time-based resources are allocated in a first-come first-served manner. At the reward time, the eNB sorts all received time bids (i.e.,  $t_i \in \mathcal{N} \setminus \{0\}$ ), in descending order:  $t_{(1)} \geq t_{(2)} \geq t_{(3)} \geq \dots \geq t_{(N)}$ . Note that the devices choosing to pay or be idle are viewed to submit time bids  $t_i = 0$ . The devices with  $t_i$  greater than  $t_{(R_W+1)}$  win resources. Recall that the waiting-time strategy function  $t_i = \sigma(a_i, a_{-i})$  is strictly increasing. Given  $R_W$  resources for low-cost M2M, the bidder wins the resource if its attribute  $a_i$  is greater than or equal to the  $R_W^{\text{th}}$ -largest attribute, i.e.,  $a_{(R_W)}$ , among other  $N - 1$  competitors. Specifically, the conditional resource accessibility probability given  $R_W$  units of waiting-time-based M2M resources is  $G_{A_{(R_W)}}(a_i) = P(a_i \geq a_{(R_W)})$ , which is the  $R_W^{\text{th}}$  order statistic of the time-valued utility on a resource in Eq. (8).

$$G_{A_{(R_W)}}(a_i) = \sum_{m=0}^{R_W-1} \binom{N-1}{m} [1 - G_A(a_i)]^m [G_A(a_i)]^{N-1-m} \quad (8)$$

Due to the symmetry property, the resource accessibility probability should be equivalent to all devices; i.e.,  $P_i^w = P^w$ . With the common belief on the distribution of the remaining resources in Eq. (5) and the conditional resource accessibility probability in Eq. (8), the resource accessibility probability  $P^w$  is

$$\begin{aligned} P^w &= \sum_{R_W=1}^{R_T} P(\text{win} | R_W \text{ resources}) \times P(R_W = R_W) \\ &= \sum_{R_W=1}^{R_T} G_{A_{(R_W)}}(a_i) \times P_{R_W}(R_W) = F(a_i). \end{aligned} \quad (9)$$

For simplicity,  $P^w$  is rewritten as  $F(a_i)$  with PDF  $f(a_i)$ , which implies that different time-valued utility on a resource results in distinctive values of resource accessibility probability. The time-valued expected payoff to join the waiting-time auction is  $\mathbb{E}[\pi_i^t(t_i, t_{-i})] = a_i F(a_i) - 2(k + t_i) + T$ .

4) *Waiting-time strategy function  $t_i$* : The Bayesian Nash equilibrium is characterized by (i) conditions under which M2M devices decide to attend the direct access, waiting-time auction, or none, and (ii) an equilibrium strategy function  $s_i(a_i, a_{-i})$ . In this part, we focus on discussing the boundary conditions and deriving the waiting-time equilibrium strategy function  $t_i^* = \sigma(a_i, a_{-i})$ .

Free resources are the incentive for the devices with high resource accessibility probability to join the waiting-time

auction. But for the devices with less resources accessibility probability, pay with direct access is another attracting alternative. Regarding to the devices with least time-valued utility, fixed costs of network entry prevents them from asking for resources. As a result, we denote 3 indifferent boundary points:  $a_{p,d}^*$ ,  $a_{t,p}^*$ , and  $a_{t,d}^*$ . A device with  $a_{p,d}^*$  satisfying Eq. (10) is neutral between paying and being idle. A device  $i$  with  $a_i > a_{p,d}^*$  prefers paying to idling.

$$a_{p,d}^* - \frac{P_0^*}{w_{p,d}^*} + (T - 2k) = T \quad (10)$$

A device with time-valued utility  $a_{t,d}^*$  would be indifferent between being awarded with zero time spent in the auction and remaining idle. The second boundary condition indicates Eq. (11) holds. A device  $i$  with  $a_i > a_{t,d}^*$  favors attending the waiting-time auction than being idle.

$$a_{t,d}^* F(a_{t,d}^*) + (T - 2k) = T \quad (11)$$

In the third boundary condition, a device with  $a_{t,p}^*$  is fair-minded whether to bid with zero time in the auction or to directly pay for the resource. In such a case, Eq. (12) holds. A device  $i$  with  $a_i > a_{t,p}^*$  chooses to join the waiting-time auction rather than pay with direct access.

$$a_{t,p}^* F(a_{t,p}^*) + (T - 2k) = a_{t,p}^* - \frac{P_0^*}{w_{t,p}^*} + (T - 2k) \quad (12)$$

In summary, the three boundary points  $a_{p,d}^*$ ,  $a_{t,d}^*$ , and  $a_{t,p}^*$  satisfy Eq. (13).

$$\begin{cases} BC1 : a_{p,d}^* = \frac{2k}{1 - \frac{P_0^*}{v(w_{p,d}^*)}} \\ BC2 : a_{t,d}^* F(a_{t,d}^*) = 2k \\ BC3 : a_{t,p}^* = F^{-1}\left(1 - \frac{P_0^*}{v(w_{t,p}^*)}\right) \end{cases} \quad (13)$$

Next, we aim to derive the waiting-time equilibrium strategy function based on the boundary conditions.  $a^*$  is denoted as a boundary attribute. If an M2M device has  $a_i > a^*$ , it always regards the waiting-time auction as the most appealing approach. Thus, Eq. (14) holds.

$$a^* = \max\{a_{t,d}^*, a_{t,p}^*\} \quad (14)$$

According to Eq. (9) in [23], the waiting-time equilibrium strategy function should satisfy a first order differential equation, which can be expressed as Eq. (15).

$$\sigma'(a_i, a_{-i}) = \sigma'(a_i) = \frac{1}{2} a_i f(a_i), \quad a_i > a^* \quad (15)$$

A device  $i$  with private attribute  $a_i$  slightly greater than the cutoff value  $a^*$  would be willing to participate in the auction, but would only wait for an extremely short time. This observation gives the initial condition for Eq. (15) as follows:

$$\lim_{a_i \rightarrow a^{*+}} \sigma(a_i) = 0 \quad (16)$$

Thus, the specific waiting-time equilibrium strategy function can be derived as

$$t_i^* = \sigma(a_i) = \int_{a^*}^{a_i} \sigma'(y) dy = \frac{1}{2} \int_{a^*}^{a_i} y f(y) dy. \quad (17)$$

5) *Equilibrium strategy function  $s_i^*$* : According to Eq. (14),  $a_{t,d}^*$  and  $a_{t,p}^*$  affect the condition to attend the waiting-time auction. The devices' actions also depend on their private attributes  $a_i$ . Therefore, their private attributes and three boundary conditions play a key role to determine the equilibrium strategy function for M2M devices' payoff maximization problem in Eq. (4). The boundary conditions are determined based on the network entry cost  $k$ , direct access price  $p_0^*$ , utility of acquiring resources  $v(w_i)$ , and resource accessibility probability  $F(a_i)$ . So, we begin with discussing the relation between  $a_{t,d}^*$ ,  $a_{t,p}^*$  and  $a_{p,d}^*$  so as to have insights into the behaviors of all the M2M devices in the equilibrium.

**Lemma 1.**  $a_{t,d}^* = a_{t,p}^* = a_{p,d}^*$  holds if and only if  $a_{t,d}^* = a_{t,p}^*$ ,  $a_{t,p}^* = a_{p,d}^*$  or  $a_{t,d}^* = a_{p,d}^*$ .

**Proof.** Considering the case  $a_{t,d}^* = a_{t,p}^*$  holds, we assume  $\tilde{a} = a_{t,d}^* = a_{t,p}^*$ . It means that  $\tilde{a}$  satisfies Eq. (11) and Eq. (12), i.e.,  $\pi^d(\tilde{a}) = \mathbb{E}[\pi^t(\tilde{a}, t_i = 0)]$  and  $\pi^P(\tilde{a}) = \mathbb{E}[\pi^t(\tilde{a}, t_i = 0)]$ . Thus,  $\tilde{a}$  also satisfies  $\pi^P(\tilde{a}) = \pi^d(\tilde{a})$  in Eq. (10), representing that  $\tilde{a}$  can be also viewed as  $a_{p,d}^*$ . As a result, we prove  $a_{t,d}^* = a_{t,p}^* = a_{p,d}^*$  if  $a_{t,d}^* = a_{t,p}^*$  holds. For the other two cases when  $a_{t,p}^* = a_{p,d}^*$  and  $a_{t,d}^* = a_{p,d}^*$  hold, the results can be showed by following the same rationale.  $\square$

**Proposition 1.** If  $a_{t,d}^* = a_{t,p}^* = a_{p,d}^*$  holds, the devices either attend the waiting-time auction or stay idle, depending on their attributes.

**Proof.** Assume  $\tilde{a} = a_{t,d}^* = a_{t,p}^* = a_{p,d}^*$ . The cutoff value  $a^*$  to wait in the auction should be  $a^* = \tilde{a}$  according to Eq. (14). An M2M device with  $a_i > \tilde{a}$  has the payoff order:  $\mathbb{E}[U_i^t(t_i^*, t_{-i}^*)] > U_i^P > U_i^d$ , so it participates in the waiting-time auction. An M2M device with  $a_i < \tilde{a}$  always stays in idle mode because its payoff order is  $U_i^d > U_i^P > \mathbb{E}[U_i^t(t_i^*, t_{-i}^*)]$ .  $\square$

**Lemma 2.**  $a_{p,d}^* > a_{t,d}^* > a_{t,p}^*$  if and only if  $a_{t,d}^* > a_{t,p}^*$  holds.

**Proof.** First, we argue that if  $a_{t,d}^* > a_{t,p}^*$  holds, then  $\frac{2k}{1 - \frac{P_0^*}{v(w_{t,p}^*)}} > a_{t,d}^* > a_{t,p}^*$  can be derived. The reason is that  $a_{t,d}^* > a_{t,p}^*$  implies  $F(a_{t,d}^*) > F(a_{t,p}^*)$  because  $F(\cdot)$  is a strictly increasing function. So, by multiplying  $a_{t,d}^*$ ,  $2k = a_{t,d}^* F(a_{t,d}^*) > a_{t,d}^* F(a_{t,p}^*)$  is derived. Because  $F(a_{t,p}^*) = 1 - \frac{P_0^*}{v(w_{t,p}^*)}$  shown in Eq. (13), the inequality is rewritten as  $2k > a_{t,d}^* (1 - \frac{P_0^*}{v(w_{t,p}^*)})$ . Thus, with  $\frac{2k}{1 - \frac{P_0^*}{v(w_{t,p}^*)}} > a_{t,d}^*$  and  $a_{t,d}^* > a_{t,p}^*$ , we conclude  $\frac{2k}{1 - \frac{P_0^*}{v(w_{t,p}^*)}} > a_{t,d}^* > a_{t,p}^*$ .

Next, we consider two cases:  $a_{p,d}^* > a_{t,p}^*$  and  $a_{p,d}^* < a_{t,p}^*$ . In the first case, i.e.,  $a_{p,d}^* > a_{t,p}^*$ ,  $\frac{2k}{1 - \frac{P_0^*}{v(w_{p,d}^*)}} \geq \frac{2k}{1 - \frac{P_0^*}{v(w_{t,p}^*)}}$  must hold because  $h(\cdot)$  is a monotonically decreasing function and  $v(\cdot)$  is a non-decreasing continuously differentiable positive-valued function. So, considering both  $a_{t,d}^* > a_{t,p}^*$  and  $a_{p,d}^* > a_{t,p}^*$ , we derive  $a_{p,d}^* = \frac{2k}{1 - \frac{P_0^*}{v(w_{p,d}^*)}} > \frac{2k}{1 - \frac{P_0^*}{v(w_{t,p}^*)}} > a_{t,d}^* > a_{t,p}^*$  according to Eq. (13). Next, we argue the second case, i.e.,  $a_{p,d}^* < a_{t,p}^*$ , does not exist. The second case results in  $a_{t,d}^* > a_{t,p}^* > a_{p,d}^*$ , which is not valid. The reason is that a device  $i$  with its

attributes  $a_{t,d}^* > a_i > a_{t,p}^*$  regards  $\pi_i^d > \mathbb{E}[\pi_i^t] > \pi_i^p$ , which contradicts with the condition that  $\pi_i^p > \pi_i^d$  when  $a_i > a_{p,d}^*$  holds. In short, the condition  $a_{t,d}^* > a_{t,p}^*$  only results in  $a_{p,d}^* > a_{t,d}^* > a_{t,p}^*$ .  $\square$

**Proposition 2.** *If  $a_{p,d}^* > a_{t,d}^* > a_{t,p}^*$  exists, M2M device  $i$ 's best response is either to attend the waiting-time auction or to stay in idle mode, depending on its attribute  $a_i$ .*

**Proof.** Consider a device  $i$  with its attribute  $a_i$ . If  $a_i > a_{p,d}^* > a_{t,d}^* > a_{t,p}^*$ , the device regards  $\mathbb{E}[\pi_i^t(a_i)] > \pi_i^p(a_i) > \pi_i^d(a_i)$ . Thus, participating in the auction is its best response. If its attribute falls in the region  $a_{p,d}^* > a_i > a_{t,d}^* > a_{t,p}^*$ , attending the auction is the best response because  $\mathbb{E}[\pi_i^t(a_i)] > \pi_i^d(a_i) > \pi_i^p(a_i)$ . If its attribute is in the region  $a_{p,d}^* > a_{t,d}^* > a_i > a_{t,p}^*$ , keeping in idle mode is its best response because  $\pi_i^d(a_i) > \mathbb{E}[\pi_i^t(a_i)] > \pi_i^p(a_i)$ . Finally, if its attribute falls in the region  $a_{p,d}^* > a_{t,d}^* > a_{t,p}^* > a_i$ , staying in idle mode is its best response because  $\pi_i^d(a_i) > \pi_i^p(a_i) > \mathbb{E}[\pi_i^t(a_i)]$ . To sum up, the device's equilibrium strategy function depends on the value of its attribute  $a_i$ . If  $a_{p,d}^* > a_{t,d}^* > a_{t,p}^*$  holds, the device  $i$ 's best response is to join the waiting-time auction or to remain in idle mode, depending on its attribute.  $\square$

**Lemma 3.**  *$a_{t,p}^* > a_{t,d}^* > a_{p,d}^*$  if and only if  $a_{t,p}^* > a_{t,d}^*$  holds.*

**Proof.** First, we argue that if  $a_{t,p}^* > a_{t,d}^*$  exists, then  $a_{t,p}^* > a_{t,d}^* > \frac{2k}{1 - \frac{p_0^*}{v(w_{t,p}^*)}}$  holds. The reason is that  $a_{t,p}^* > a_{t,d}^*$  implies  $F(a_{t,p}^*) > F(a_{t,d}^*)$  because  $F(\cdot)$  is a strictly increasing function. Next, we can derive  $a_{t,d}^* F(a_{t,p}^*) > a_{t,d}^* F(a_{t,d}^*) = 2k$  by multiplying  $a_{t,d}^*$ . Because of  $F(a_{t,p}^*) = 1 - \frac{p_0^*}{v(w_{t,p}^*)}$  shown in Eq. (13), the inequality is rewritten as  $a_{t,d}^* (1 - \frac{p_0^*}{v(w_{t,p}^*)}) > 2k$ . So, we conclude  $a_{t,p}^* > a_{t,d}^* > \frac{2k}{1 - \frac{p_0^*}{v(w_{t,p}^*)}}$ .

Next, we consider two cases:  $a_{p,d}^* > a_{t,p}^*$  and  $a_{p,d}^* < a_{t,p}^*$ . In the first case, i.e.,  $a_{p,d}^* > a_{t,p}^*$ , we argue that  $a_{p,d}^* > a_{t,p}^* > a_{t,d}^*$  does not exist. A device  $i$  with its attributes  $a_{p,d}^* > a_{t,p}^* > a_i > a_{t,d}^*$  regards  $\pi_i^p > \mathbb{E}[\pi_i^t] > \pi_i^d$ , which contradicts with the condition that  $\pi_i^d > \pi_i^p$  because of  $a_{p,d}^* > a_i$ . When it comes to the second case, i.e.,  $a_{p,d}^* < a_{t,p}^*$ , we find  $\frac{2k}{1 - \frac{p_0^*}{v(w_{p,d}^*)}} < \frac{2k}{1 - \frac{p_0^*}{v(w_{t,p}^*)}}$

holds because  $h(\cdot)$  is a monotonically decreasing function and  $v(\cdot)$  is a non-decreasing continuously differentiable positive-valued function. Taking both  $a_{t,p}^* > a_{t,d}^*$  and  $a_{t,p}^* > a_{p,d}^*$  into consideration, we derive  $\frac{2k}{1 - \frac{p_0^*}{v(w_{t,p}^*)}} > a_{t,p}^* > a_{t,d}^* >$

$a_{p,d}^* = \frac{2k}{1 - \frac{p_0^*}{v(w_{p,d}^*)}}$ . In short, if  $a_{t,p}^* > a_{t,d}^*$  holds, then only  $a_{t,p}^* > a_{t,d}^* > a_{p,d}^*$  exists.  $\square$

**Proposition 3.** *If  $a_{t,p}^* > a_{t,d}^* > a_{p,d}^*$  exists, M2M device  $i$ 's best response can be to attend the waiting-time auction, to pay with direct access, or to stay idle, depending on its attribute  $a_i$ .*

**Proof.** Consider a device  $i$  with its attribute  $a_i$ . If  $a_i > a_{t,p}^* > a_{t,d}^* > a_{p,d}^*$ , the device regards  $\mathbb{E}[\pi_i^t(a_i)] > \pi_i^p(a_i) > \pi_i^d(a_i)$ . Thus, participating in the auction is its best response. If

its attribute falls in the region  $a_{t,p}^* > a_i > a_{t,d}^* > a_{p,d}^*$ , paying with direct access is its best response because of  $\pi_i^p(a_i) > \mathbb{E}[\pi_i^t(a_i)] > \pi_i^d(a_i)$ . If its attribute is in the region  $a_{t,p}^* > a_{t,d}^* > a_i > a_{p,d}^*$ , paying with direct access is its best response because of  $\pi_i^p(a_i) > \pi_i^d(a_i) > \mathbb{E}[\pi_i^t(a_i)]$ . Finally, if its attribute falls in the region  $a_{t,p}^* > a_{t,d}^* > a_{p,d}^* > a_i$ , staying in idle mode is its best response because of  $\pi_i^d(a_i) > \pi_i^p(a_i) > \mathbb{E}[\pi_i^t(a_i)]$ . To sum up, the device's equilibrium strategy function depends on the value of its attribute  $a_i$ . If  $a_{t,p}^* > a_{t,d}^* > a_{p,d}^*$  holds, the device  $i$ 's best response is to join the waiting-time auction, to pay with direct access, or to remain idle, counting on its attribute  $a_i$ .  $\square$

**Proposition 4.** *The M2M devices' Bayesian Nash equilibrium strategy function  $s_i^*$ ,  $\forall i \in \mathcal{N} \setminus \{0\}$  is summarized in Table III.*

**Proof.** After analyzing the relation among the three boundary points  $a_{t,d}^*$ ,  $a_{t,p}^*$  and  $a_{p,d}^*$ , the best responses of a device with its attribute  $a_i$  are derived in Proposition 1, 2 and 3. The device's best response is to maximize its payoff in Eq. (4) according to its attribute. Depending on different scenarios and the regions  $a_i$  falls in, the Bayesian Nash equilibrium strategies are summarized in Table III. The Bayesian Nash equilibrium strategy function  $s_i^*$  satisfies  $U_i(s_i^*, s_{-i}^*) \geq U_i(s_i', s_{-i}^*)$ ,  $s_i' \neq s_i^*$ , where  $U_i(s_i) = \left( U_i^p, \max \mathbb{E}[U_i^t(t_i, t_{-i})], U_i^d \right)$ .  $\square$

The boundary points, determined by  $k$ ,  $v(\cdot)$  and  $p_0^*$ , play a key role in the device's equilibrium strategy function. In other words, by properly designing the price for direct access, the eNB can induce the devices to operate in the third scenario so as to maximize its expected revenue.

### B. Leader's Game

The eNB's strategy is to determine the direct access price  $p_0$  per price-based resource, aiming to maximize its expected revenue in Eq. (1). The announced price  $p_0$  affects the boundary conditions in the followers' game indicated in Eq. (13) and the followers' equilibrium strategies.

From the eNB's viewpoints, it provides two disjoint resource pools. Waiting-time-based resources ( $R_T$  units) are free to guarantee H2H and the remaining ones to serve low-cost M2M. Price-based resources ( $\overline{R_P}$  units) aim to enhance mission-critical M2M with reliability, so devices pay and are directly allocated with resources without waiting. The expected resource number to be sold is  $R_P$  units,  $R_P \leq \overline{R_P} \leq N$ , in case that all M2M devices urge for ultra reliable communications. The eNB assumes that each M2M device requests the direct access with probability  $P_b$ . So, the expected revenue from the price-based resource pool is

$$\mathbb{E}(U_0) = p_0 R_P = p_0 \left[ \sum_{i=1}^n i P_b^i (1 - P_b)^{(N-i)} \right] = p_0 N P_b. \quad (18)$$

If the direct access price is high, most devices would turn to the free waiting-time-based resources. Based on the boundary conditions in Eq. (13), the direct access price  $p_0$  satisfies two inequalities: (1)  $p_0 < v(w_{t,p})[1 - P^w]$ , an increasing function, and (2)  $p_0 < v(w_{p,d}) - 2kw_{p,d}$ , an decreasing function, as shown in Fig. 3. The value of  $p_0$  motivates the devices with

TABLE III  
M2M DEVICE  $i$ 'S BAYESIAN NASH EQUILIBRIUM STRATEGY FUNCTION

Scenario	Equilibrium Strategy
i $a^* = a_{t,p}^* = a_{t,d}^* = a_{p,d}^*$	$s_i^* = t_i^* = \sigma(a_i, a_{-i})$ in Eq. (17) if $a_i > a^*$ wait $s_i^* = t_i^* = 0$ if $a_i < a^*$ idle
ii $a_{p,d}^* > a_{t,d}^* > a_{t,p}^*$ $a^* = a_{t,d}^*$	$s_i^* = t_i^* = \sigma(a_i, a_{-i})$ in Eq. (17) if $a_i > a^*$ wait $s_i^* = t_i^* = 0$ if $a_i < a^*$ idle
iii $a_{t,p}^* > a_{t,d}^* > a_{p,d}^*$ $a^* = a_{t,p}^*$	$s_i^* = t_i^* = \sigma(a_i, a_{-i})$ in Eq. (17) if $a_i > a^*$ wait $s_i^* = p_0^*$ if $a^* > a_i > a_{p,d}^*$ pay $s_i^* = t_i^* = 0$ if $a_i < a_{p,d}^*$ idle

$w_i$  satisfying  $w_{t,p} < w_i < w_{p,d}$  would like to pay. That is, the portion of devices choosing to pay is  $P_m = G_W(w_{p,d}) - G_W(w_{t,p})$ . The expected number of devices falling in the range is  $N(G_W(w_{p,d}) - G_W(w_{t,p}))$ . The expected total payment to the eNB is  $p_0 N(G_W(w_{p,d}) - G_W(w_{t,p}))$ . Because the M2M devices' expected total payment is equivalent to the eNB's expected revenue,  $p_0 N P_b = p_0 N P_m$  must hold. Thus, we find  $P_b = P_m = G_W(w_{p,d}) - G_W(w_{t,p})$ .

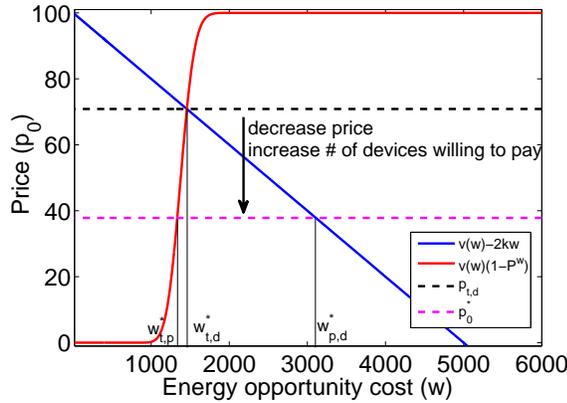


Fig. 3. Solution concept ( $w_i \sim N(1800, 2000^2)$ ,  $N = 175$ ,  $\overline{R_P} = 90$ ,  $R_T = 50$ ,  $v(w_i) = 100$ ,  $k = 0.01$ ,  $T = 0.1$ ,  $R_P^* = 72$ )

The expected revenue maximization problem is

$$\begin{aligned}
 & \underset{s_0}{\text{maximize}} \quad \mathbb{E}(U_0) \approx \underset{p_0, R_P, p_0}{\text{maximize}} \quad p_0 R_P \\
 & \text{subject to} \quad w_{p,d}^* = \frac{v(w_{p,d}^*) - p_0}{2k}, w_{t,p}^* = \frac{v(w_{t,p}^*)}{F^{-1}(1 - \frac{p_0}{v(w_{t,p}^*)})}, \\
 & \quad R_P \approx N \int_{w_{t,p}^*}^{w_{p,d}^*} g_W(w) dw, 0 \leq R_P \leq \overline{R_P}, R_P \in \mathbb{N}, \\
 & \quad p_0 \in \mathbb{R}^+.
 \end{aligned} \tag{19}$$

However, the functions defining the equality constraints are not affine. It turns out that the expected revenue maximization problem is a non-convex optimization problem, which is difficult to solve. Thanks to the boundary conditions and finite integer  $\overline{R_P}$ , we come out an iterative algorithm to find the optimal direct access price  $p_0^*$  and expected revenue, as shown in Algorithm 1. Given each  $R_P \in [1, \overline{R_P}]$ , we can find the corresponding price by setting the initial price to be  $p_{t,d}$ , which satisfies  $a_{t,d}^* = a_{t,p}^* = a_{p,d}^*$ . Any price higher than

$p_{t,d}$  results in the situation that all devices choose to wait or be idle. The price iteratively decreases from  $p_{t,d}$  based on Newton's method to achieve the condition that the number of paying devices equals to the fixed resource amount. Algorithm 1 includes two parts. The goal of the first part is to find the initial step size of Newton's method, by ensuring the portion of devices willing to pay falls within a certain range, e.g., between  $0.05 \frac{R_P}{N}$  and  $0.5 \frac{R_P}{N}$  in Algorithm 1. The initial step size is very important to make sure the corresponding price can be found. In the second part, the price continues to decrease with a variable step size  $p_u$  so that the portion of devices willing to pay is extremely close to  $\frac{R_P}{N}$ . Finally, the optimal expected revenue, resource number and price are selected. The order of Newton's method is 2. The complexity of this algorithm is linear with  $\overline{R_P}$ , and quadratic order of convergence to find the corresponding price given each  $R_P$ .

**Algorithm 1** Revenue Maximization Iterative Algorithm with Newton's Method

**Input:**  $U_0(p_0)$ : objective function;  $p_0$ : direct access price;  $p_{t,d}$ : initial price;  $p_u$ : step size;  $R_P$ ;  $N$ ;  
**Output:**  $U_0^*(p_0^*)$ : maximized revenue;  $p_0^*$ : optimal direct access price;  $R_P^*$ : optimal resource amount  
1: initialize  $U_0^* = 0$  and  $p_0 = p_{t,d}$ ;  
2: **for all**  $R_P \in [1, \overline{R_P}]$  **do**  
3:   initial  $p_u = 10^{-6} \times p_0$ ;  
4:   **repeat**  
5:      $p_1 = p_0 - p_u$ ;  
6:      $Q(p_1) = G_A(a_{t,p}(p_1)) - G_A(a_{p,d}(p_1))$ ;  
7:     **if**  $Q(p_1) < 0.05 \times \frac{R_P}{N}$  **then**  
8:        $p_u = 1.1 \times p_u$ ;  
9:     **else if**  $Q(p_1) > 0.5 \times \frac{R_P}{N}$  **then**  
10:        $p_u = 0.5 \times p_u$ ;  
11:     **end if**  
12:     **until**  $(0.05 \times \frac{R_P}{N} < Q(p_1) < 0.5 \times \frac{R_P}{N})$   
13:      $Q(p_0) = G_A(a_{t,p}(p_0)) - G_A(a_{p,d}(p_0))$ ;  
14:     **while**  $|N \times Q(p_0) - R_P| > 10^{-10}$  **do**  
15:        $p_1 = p_0 - p_u$ ;  
16:        $Q(p_1) = G_A(a_{t,p}(p_1)) - G_A(a_{p,d}(p_1))$ ;  
17:        $p_u = \frac{\frac{R_P}{N} - Q(p_1)}{5(Q(p_1) - Q(p_0))} \times p_u$ ;  
18:        $p_0 = p_1$ ;  
19:     **end while**  
20:     **if**  $U_0(p_0) > U_0^*$  **then**  
21:        $U_0^* = U_0$ ,  $p_0^* = p_0$ , and  $R_P^* = R_P$ ;  
22:     **end if**  
23: **end for**  
24: **return**  $U_0^*$ ,  $p_0^*$  and  $R_P^*$ ;

### C. Properties

We show three properties of the prioritized framework: existence, uniqueness, and interregional and waiting-time-based truth-telling. Also, the framework is an indirect mechanism [26], which possesses Bayesian Nash incentive compatibility, interim efficiency, interim individual rationality and weakly budget balance.

**Proposition 5.** *The Bayesian Nash equilibrium of the prioritized framework exists. Considering the eNB's expected revenue maximization, the equilibrium strategy function  $s_{i \neq 0}^* = \{p_0^*, t_i^* = \sigma(a_i)\}$  in Scenario iii of Table III is a Bayesian Nash equilibrium.*

**Proof.** We discuss the equilibrium strategy function in three cases since  $s_{i \neq 0}^*$  in Scenario iii of Table III is region-wise.

1) Device  $i$  with  $a_i > a_{t,p}^*$ : This device may deviate from  $t_i^* = \sigma(a_i)$  to either  $t_i' \neq \sigma(a_i)$ , to  $p_0^*$ , or to remain idle. Since  $\sigma(\cdot)$  is a monotonically increasing function, there exists a unique reverse function  $\sigma^{-1}(\cdot)$ , which determines the fraudulent time-valued utility  $a_i' = \sigma^{-1}(t_i')$ . The time-valued expected payoff with deviation is

$$E[\pi_i^t(t_i', a_i)] = a_i F(\sigma^{-1}(t_i')) - \int_{a_{t,p}^*}^{\sigma^{-1}(t_i')} y f(y) dy + T - 2k$$

$$, \text{ where } t_i' = \int_{a^*}^{\sigma^{-1}(t_i')} y f(y) dy. \quad (20)$$

First, we compare the difference between the expected time-valued payoff of applying the waiting-time equilibrium strategy function  $\sigma(a_i)$  and deviation with  $t_i'$  in Eq. (21).

$$E[\pi_i^t(t_i^*, t_{-i}^*)] - E[\pi_i^t(t_i', t_{-i}^*)] = a_i \left( \int_{\sigma^{-1}(t_i')}^{a_i} \left(1 - \frac{y}{a_i}\right) f(y) dy \right) \quad (21)$$

Since the time-valued attribute  $a_i$  is always positive, we need to justify the result of Eq. (21) is always positive in two cases: (i)  $t_i' > \sigma(a_i)$  and (ii)  $t_i' < \sigma(a_i)$ , so that the Bayesian Nash equilibrium exists. In case (i),  $\sigma^{-1}(t_i') > a_i$  holds because  $\sigma(\cdot)$  is a strictly increasing function. Taking the integral range into consideration, we observed that  $\frac{y}{a_i} \geq 1$ . The integral part can be transformed into  $\int_{a_i}^{\sigma^{-1}(t_i')} \left(\frac{y}{a_i} - 1\right) f(y) dy$ , which is always positive. Similarly,  $a_i > \sigma^{-1}(t_i')$  and  $\frac{y}{a_i} \leq 1$  hold in the case (ii) so that the integral part is always positive. Thus, we have proved Eq. (21) always to be positive. That is, waiting  $t_i' \neq \sigma(a_i)$  cannot make the device  $i$  better.

Based on the condition of boundary point  $a_{t,p}^*$ , we find that deviation from waiting time  $t_i^* = \sigma(a_i)$  to paying  $p_0^*$  results in less time-valued utility, as shown in Eq. (22).

$$E[\pi_i^t(t_i^* = \sigma(a_i), t_{-i}^*)] > E[\pi_i^t(t_i' = \sigma(a_{t,p}^*), t_{-i}^*)]$$

$$= a_i F(a_{t,p}^*) - 2k + T = a_i \left(1 - \frac{p_0^*}{v(w_{t,p}^*)}\right) - 2k + T \quad (22)$$

$$\geq a_i \left(1 - \frac{p_0^*}{v(w_i)}\right) - 2k + T = \pi_i^p$$

Moreover, the time-valued expected payoff with  $\sigma(a_i)$  is better than that to stay idle shown in Eq. (23), because of the

increasing function  $F(\cdot)$ , i.e.,  $a_i > a_{t,p}^*$  and  $F(a_{t,p}^*) > F(a_{t,d}^*)$ .

$$E[\pi_i^t(t_i^* = \sigma(a_i), t_{-i}^*)] > E[\pi_i^t(t_i' = \sigma(a_{t,p}^*), t_{-i}^*)]$$

$$= a_i F(a_{t,p}^*) - 2k + T > a_i F(a_{t,p}^*) - a_{t,d}^* F(a_{t,d}^*) + T \quad (23)$$

$$> T = \pi_i^d$$

2) Device  $i$  with  $a_{t,p}^* > a_i > a_{p,d}^*$ : In such a case,  $w_{t,p}^* < w_i < w_{p,d}^*$  and  $v(w_{t,p}^*) \leq v(w_i) \leq v(w_{p,d}^*)$  since  $v(\cdot)$  is a non-decreasing concave function. The device may deviate from paying  $p_0^*$  to either wait  $t_i''$  or stay idle. According to the boundary condition  $a_{t,p}^*$  in Eq. (13), we show in Eq. (24) that the time-valued expected payoff with  $t_i''$  is not better because  $\frac{y}{a_i} \geq 1$  and the integral is not positive.

$$\pi_i^p = a_i \left(1 - \frac{p_0^*}{v(w_i)}\right) - 2k + T$$

$$\geq a_i \left(1 - \frac{p_0^*}{v(w_{t,p}^*)}\right) - 2k + T = a_i F(a_{t,p}^*) - 2k + T \quad (24)$$

$$\geq a_i F(a_{t,p}^*) + \int_{a_{t,p}^*}^{\sigma^{-1}(t_i'')} a_i \left(1 - \frac{y}{a_i}\right) f(y) dy - 2k + T$$

$$= E[\pi_i^t(t_i'', t_{-i}^*)]$$

Afterwards, based on the boundary condition  $a_{p,d}^*$ , the monetary payoff to pay is rewritten as

$$U_i^p(a_i) = v(w_i) - p_0^* - 2kw_i + Tw_i$$

$$= (w_{p,d}^* - w_i) \left(2k - \frac{v(w_{p,d}^*) - v(w_i)}{w_{p,d}^* - w_i}\right) + Tw_i. \quad (25)$$

We further show that  $\frac{v(w_{p,d}^*) - v(w_i)}{w_{p,d}^* - w_i}$  is less than  $2k$  in Eq. (26) because of the concave function  $v(\cdot)$  and boundary condition  $a_{p,d}^*$ .

$$\frac{v(w_{p,d}^*) - v(w_i)}{w_{p,d}^* - w_i} \leq \frac{v(w_{p,d}^*) - v(w_{t,p}^*)}{w_{p,d}^* - w_{t,p}^*} = \frac{2kw_{p,d}^* + p_0^* - v(w_{t,p}^*)}{w_{p,d}^* - w_{t,p}^*}$$

$$< \frac{2kw_{p,d}^* + v(w_{t,p}^*) - 2kw_{t,p}^* - v(w_{t,p}^*)}{w_{p,d}^* - w_{t,p}^*} = 2k \quad (26)$$

According to  $w_{p,d}^* - w_i > 0$  and Eq. (26), paying  $p_0^*$  brings more payoff than staying idle.

3) Device  $i$  with  $a_{t,p}^* > a_{p,d}^* > a_i$ : This device with  $w_i > w_{p,d}^* > w_{t,d}^*$  and  $v(w_i) \geq v(w_{p,d}^*) \geq v(w_{t,d}^*)$  may deviate from being idle to either pay  $p_0^*$  or wait  $t_i''$ . Because of Eq. (26) and the concave function  $v(\cdot)$ , the payoff of paying  $p_0^*$  in Eq. (25) is worse than that of staying idle. Considering Eq. (24), the payoff of waiting  $t_i''$  is further less than that of paying  $p_0^*$ . Thus, staying idle is the optimal action.

In summary, the Bayesian Nash equilibrium exists with  $s_0^* = p_0^*$  and  $s_{i \neq 0}^*$  following Scenario iii of Table III, and  $\sigma(\cdot)$  is the waiting-time equilibrium strategy function.  $\square$

**Proposition 6.** *The Bayesian Nash equilibrium of the prioritized framework is unique.*

**Proof.** The eNB's revenue function is strictly concave with respect to the direct access price, where the price  $p_0$  is finite within  $[0, p_{t,d}]$ . Thus, the optimal price  $p_0^*$  to maximize the

revenue is unique. According to Eq. (24), (25) and (26), only accepting the unique  $p_0^*$  is the best response for M2M device  $i$  with  $a_{i,p}^* > a_i > a_{p,d}^*$ . Similarly, based on the discussion in Proposition 5, staying idle is the best response for device  $i$  with  $a_i < a_{p,d}^*$ . In addition, contending for waiting-time-based resources instead of paying and being idle is the best response for device  $i$  with  $a_i > a_{i,p}^*$ . We further assume there exists another waiting-time equilibrium strategy function  $\zeta(a_i)$  better than  $\sigma(a_i)$  to reach Pareto improvement. Based on the relation between affordable waiting time and the time-valued utility on a resource,  $\zeta(a_i)$  is a positive-valued, strictly increasing and differentiable function. Since it should satisfy the boundary condition in Eq. (15) whether to participate in the waiting-time auction,  $\zeta(a_i)$  must have the similar structure as Eq. (17). However,  $\zeta(a_i) < \sigma(a_i)$  results in worse expected payoff as indicated in Eq. (21). By contradiction,  $\zeta(a_i)$  should not exist so that the unique waiting-time equilibrium strategy function is  $\sigma(a_i)$ . To sum up, the direct access price is unique and the Bayesian Nash equilibrium in Scenario iii of Table III is unique, which applies for all M2M devices.  $\square$

**Proposition 7.** *The interregional truth-telling property is guaranteed in the prioritized resource allocation framework.*

**Proof.** Assume device  $i$  with  $a_i$  applies the Bayesian Nash equilibrium strategy function in Scenario iii of Table III. If the device with  $a_i > a_{i,p}^*$  lies its time-valued utility to be  $\tilde{a}_i < a_{i,p}^*$ , it will choose to pay or to be idle, resulting in worse payoff shown in Eq. (22) and (23). The device with  $a_{i,p}^* > a_i > a_{p,d}^*$ , lying its time-valued utility to be  $\tilde{a}_i > a_{i,p}^*$  or  $\tilde{a}_i < a_{p,d}^*$ , will not decide to pay and thus have less payoff. Similarly, if the device with  $a_i < a_{p,d}^*$  lies its time-valued utility to be  $\tilde{a}_i > a_{p,d}^*$ , it then decides not to be idle and hence has less payoff. Therefore, a rational device should apply the interregional decisions according to the proof in Proposition 5. From its revealed action, i.e., wait, pay or stay idle, we can judge the region of its time-valued utility.  $\square$

**Proposition 8.** *The M2M device with  $a_i > a_{i,p}^*$  truly reveals its private information in the waiting-time auction.*

**Proof.** Assume device  $i$  with the true time-valued utility on the resource  $a_i$  applies the equilibrium strategy  $\sigma(a_i)$ . However, if it lies the time-valued utility to be  $\tilde{a}_i \neq a_i$ , the waiting time will be  $\tilde{t} \neq \sigma(a_i)$  by following the strictly increasing equilibrium strategy function. As shown in Eq. (21), the expected payoff will not be better off. Furthermore, paying or staying idle will not make it better off, according to Eq. (22) and (23). Therefore, a rational device should apply the waiting-time equilibrium strategy  $\sigma(a_i)$  and honestly reveal its private time-valued utility  $a_i$ .  $\square$

**Definition 1.** *The prioritized resource allocation framework is an indirect (revelation) mechanism  $\mathcal{M} = (\mathcal{S}_{i \neq 0}, \mathcal{L}(\cdot))$ , where  $\mathcal{S}_{i \neq 0}$  is a set of possible actions for M2M device  $i$ .  $\mathcal{L} : \prod_{i=1}^N \mathcal{S}_i(\theta_i) \rightarrow X$  is a function that maps each action profile to an outcome  $x \in X$ .*

The indirect mechanism provides a choice of actions, i.e., wait, pay and be idle, to each M2M device and specifies a social outcome for each action profile. The function  $\mathcal{L}(\cdot)$

allocates the waiting-time-based resources according to the order of waiting time, and distributes the price-based resources if the device pays. The outcome profile  $x = (x_1, \dots, x_N)$ , where  $x_i \in \{0, 1\}$  and  $\sum_{i=1}^N x_i \leq \overline{R}_P + R_W$ .

**Lemma 4.** *The followers' game  $\mathcal{G}^F$  induced by  $\mathcal{M}$  has a unique pure strategy Bayesian Nash equilibrium such that  $\mathcal{L}(s_1^*(\theta_1), \dots, s_N^*(\theta_N)) = \mathcal{F}(\theta_1, \dots, \theta_N)$ ,  $\forall (\theta_1, \dots, \theta_N) \in \Theta$ .*

**Proposition 9.** *According to Lemma 4, the proposed indirect mechanism  $\mathcal{M}$  implements the social choice function  $\mathcal{F}(\cdot)$  in Bayesian Nash equilibrium.*

**Proposition 10.** *The social choice function  $\mathcal{F}(\cdot)$  is Bayesian-Nash incentive compatible (BIC), i.e., truthfully implementable in Bayesian Nash equilibrium.*

**Proof.** The mechanism has Bayesian Nash equilibrium  $s^* = (s_1^*(\cdot), \dots, s_N^*(\cdot))$  where  $s_i^*(\theta_i) = s_i^*(a_i)|_{a_i=h(w_i)}$ . According to Proposition 7 and 8, true revelation of  $a_i$  for each device constitutes the Bayesian Nash equilibrium of the followers' game  $\mathcal{G}^F$ . Since  $h(\theta_i = w_i)$  is a strictly continuous decreasing function, truly revealing its type  $w_i$  for each device indeed results in the Bayesian Nash equilibrium. Thus, Bayesian-Nash incentive compatibility holds.  $\square$

**Proposition 11.** *The social choice function  $\mathcal{F}(\cdot)$  is interim efficient.*

**Proof.** Bayesian efficiency is the Pareto optimality with incomplete information. Suppose there are a waiting-time-based resource pool and a price-based resource pool. Distributing resources to the devices  $a_i < a_{p,d}^*$  will make their expected payoff worse. Allocating the devices  $a_{i,p}^* > a_i > a_{p,d}^*$  waiting-time-based resources make their expected payoff worse as well. In addition, providing the devices  $a_i > a_{i,p}^*$  with price-based resources reduces their expected payoff. Furthermore, no other waiting-time equilibrium strategy function exist for devices  $a_i > a_{i,p}^*$  to reach Pareto improvement. In other words, no other social choice functions achieves Pareto improvement. Since the social choice function is also incentive compatible, it is interim efficient provided that the devices already know their own types.  $\square$

**Proposition 12.** *Interim individual rationality (interim IR) holds for M2M devices.*

**Proof.** The time-valued expected payoff function in the Bayesian Nash equilibrium is continuously decreasing but always greater or equal to  $T$ . Suppose a device  $i$  is allowed to withdraw from the mechanism after it has learned its type but before it chooses the action. However, the time-valued payoff of withdrawing cannot be better than  $T$ , which is the lower value of its interim expected payoff  $\pi_i(w_i|\mathcal{F}(\cdot)) = E_{w_{-i}}[\pi_i(\mathcal{F}(w_i, w_{-i})|w_i)]$ .  $\square$

**Proposition 13.** *The social choice function  $\mathcal{F}(\cdot)$  is weakly budget balanced.*

**Proof.** In the Bayesian Nash equilibrium, at least one device pays  $p_0^*$  for the resource because the price is determined by the eNB's expected revenue maximization. It is the eNB instead of M2M devices that receives the payments. Thus,

considering the mechanism is designed for M2M devices, the total payments are greater than total receipts (i.e., zero).  $\square$

### V. NUMERICAL ANALYSIS

We implement the prioritized resource allocation framework on a MATLAB-based simulation platform. We not only verify the correctness of the mathematical model but also explore the system properties such as truth-telling and expected delay in connected mode caused by different H2H traffic loads, M2M device number, and energy awareness levels. The system properties due to the resource pool partition are further addressed. We assume  $R_T = 50$  units,  $\bar{R}_P = 250$  units,  $k = 0.01$  units,  $T = 0.1$  units,  $w_i \sim N(\mu, s^2)$  with  $\mu = 1800$  and  $s = 2000$ , and  $v(w_i) = 100$ , regardless of the energy opportunity costs.

First, we validate Algorithm 1 to find the optimal expected revenue and announced price, shown in Fig. 4. Given each  $R_P \in [1, \bar{R}_P]$ , the price is determined by Newton's method to satisfy the constraint that the expected sold resource number equals to the expected paying device number. As  $R_P$  increases, the corresponding price reduces to prompt more devices to pay. Afterwards, the optimal expected revenue and price are selected within the range of available resource supply. For instance,  $U_0^* = 4210$ ,  $p_0^* = 42.96$  and  $R_P^* = 98$  in Fig. 4.

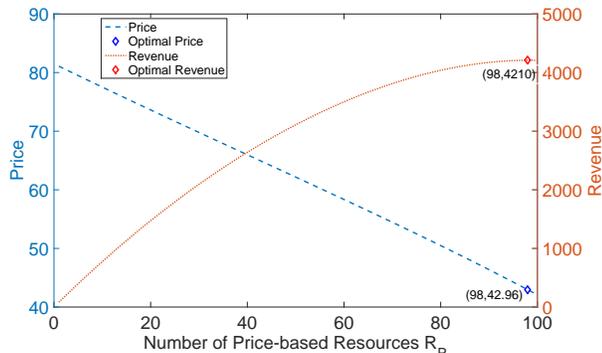


Fig. 4. Revenue vs. Price, with H2H traffic load  $\lambda = 20$

Next, we investigate the devices' optimal strategy according to the energy awareness levels (i.e., the energy opportunity costs), M2M device number and H2H traffic loads, as shown in Fig. 5 and 6. Generally, the low-cost M2M with less energy opportunity cost would prefer to wait because the resources are free. They care less about the energy wasted in connected mode. In contrast, the mission-critical M2M with more energy awareness directly pay for resources instead of wasting energy to line up in connected mode. The proposed mechanism serves both low-cost M2M and mission-critical M2M well by two types of resources: free but unreliable vs. costly but reliable. Trivially, the devices caring a lot about the consumed energy would not even ask for resources because they are too stingy to communicate, thereby staying idle. In other words, the mechanism also provides for admission control by keeping some devices in idle mode. When the device number increases in Fig. 5, the threshold of energy opportunity cost at which M2M devices are willing to wait ( $w_{t,p}^*$ ) decreases. This implies some devices surrender and change their actions from waiting to paying due to their stringent energy awareness. The

waiting time increases because more devices contend for free resources. Similarly, when the H2H load increases in Fig. 6, fewer free resources are available for low-cost M2M so that the waiting time increases. When the M2M device number rises or the H2H traffic load increases, the range of energy opportunity cost within which M2M devices will not participate in the auction expands. Therefore, the charges for direct resource access can be higher, since more M2M devices will ask for resources based on direct access price instead of time bids.

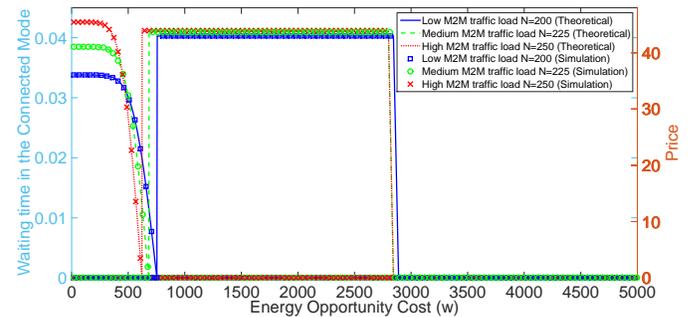


Fig. 5. Optimal Strategy in BNE, with H2H traffic load  $\lambda = 20$

In addition, when the H2H load increases, the time-valued payoff decreases because the winning devices have to spend more time waiting, as shown in Fig. 7. Overall, the time-valued payoff is a decreasing function of the energy opportunity cost. That is, a resource deserves less spent time due to devices' high energy awareness, so that the devices turn to pay directly. In Fig. 5, 6 and 7, we verify the mathematical results match the simulation results well.

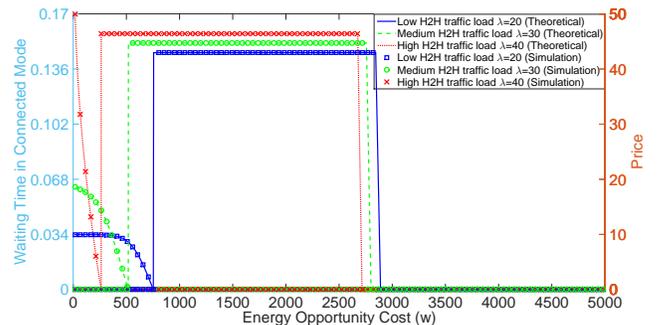


Fig. 6. Optimal Strategy in BNE, with  $N = 200$

Furthermore, we explore the expected waiting time (delay) in connected mode among all winners. According to Fig. 8, we observe that the expected delay in connected mode rises given the same H2H traffic load but increasing number of competitors. The reason is that the potential M2M winners with lower energy opportunity cost tends to increase their waiting time, as shown in Fig. 5. Hence, the expected delay of the fixed amount of winners tends to increase. In addition, when the H2H traffic load becomes heavy given the same number of competitors, the expected delay of the winners increases due to fewer available free resources and the longer waiting time, as shown in Fig. 6. That is, more competitors and heavy traffic H2H load not only make the optimal direct

access price higher, but also increase the expected delay in connected mode.

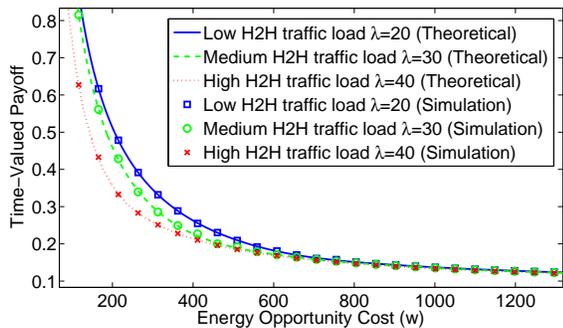


Fig. 7. Time-valued payoff in BNE, with  $N = 200$

In Fig. 9, we verify that the inter-regional and waiting-time-based truth-telling is guaranteed, as mentioned in Proposition 7 and 8. The blue line (early deviation) represents that each device pretends to have lower energy opportunity cost and higher time-valued payoff, so that it chooses to line up earlier, or line up instead of paying/being idle, or pay instead of being idle. The red line (late deviation) means the opposite case. Irrespective of the strategies, we find that the time-valued payoff function decreases with respect to the energy opportunity cost, same phenomena observed in Fig. 7. The payoff of deviation from the optimal strategy in BNE is worse whether the device arrives earlier or later. If it arrives earlier, the resource accessibility probability increases with the sacrifice of increasing waiting time cost. If it arrives later, the waiting time cost decreases with the sacrifice of the lower resource accessibility probability, and the device still devotes a fixed amount of entry cost. Therefore, truly revealing the private information by the corresponding decision is the optimal strategy in the prioritized framework.

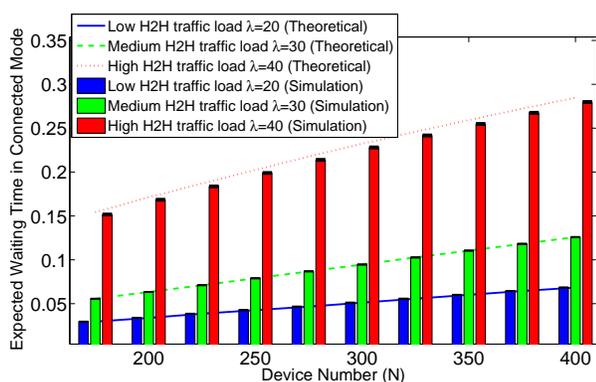


Fig. 8. Expected waiting time (delay) in connected mode

Finally, we further investigate the system properties owing to the resource pool partition. Assume the total resources  $R_{tot}$  is partitioned into  $R_T$  units for the waiting-time-based resource pool and the remaining  $R_{tot} - R_T = \bar{R}_P$  units for the price-based resource pool. We find that the optimal price varies under different resource pool partitions in Fig. 10. Given the fixed total resource size and M2M device number, as

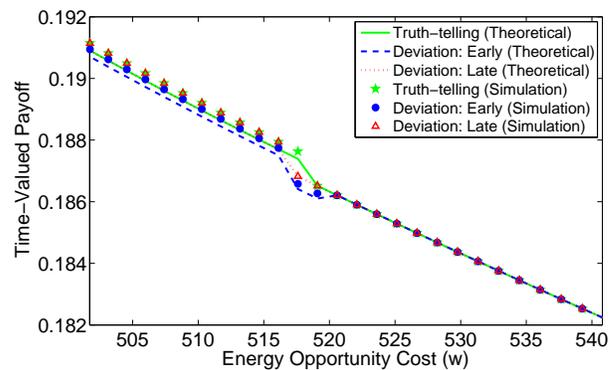


Fig. 9. Truth-telling verification ( $N = 200$ ,  $\lambda = 30$ )

$R_T$  increases, the price tends to decrease and the expected revenue decreases, as presented in Fig. 11. The reason is that more devices are willing to participate in the waiting-time auction when more free resources are provided, meaning the waiting-time auction is more attracting and less demand for the reliable resources. Thus, the eNB has to reduce the price. With the reduced price and less expected devices to pay for direct access, the eNB's expected revenue decreases. In the meanwhile, as the  $R_T$  increases, the expected delay in connected mode also declines, as shown in Fig. 12. That is, there is a tradeoff between the expected revenue and the expected delay in connected mode. If the expected revenue maximization is the eNB's only consideration, the optimal resource pool partition can be reducing  $R_T$  as much as possible to slightly higher than the H2H traffic loads, so that at least the H2H devices can be satisfied. In short, the optimization goals and several factors such as H2H traffic loads and M2M device number can have a great impact on the resource pool partition. Since the factors change from time to time, the resource pool partition can be achieved in a dynamic manner by context-aware estimation and prediction approaches, which is worthy of investigation as the future direction. In this way, the eNB can announce the updated price based on the appropriate resource pool partition every time the mechanism is held.

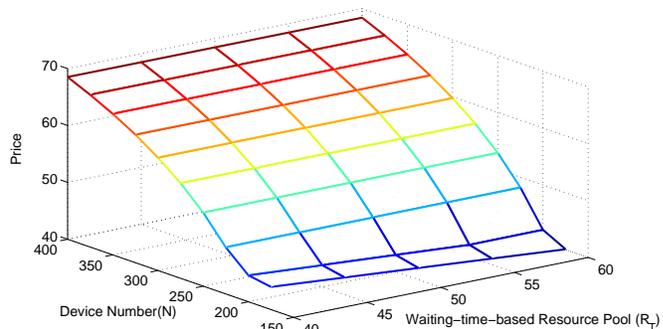


Fig. 10. Announced price under resource pool partitions ( $R_T + \bar{R}_P = 150$ ,  $\lambda = 20$ )

## VI. CONCLUSION

We design an uplink dedicated resource allocation framework with a mixture of waiting-time auction and direct access

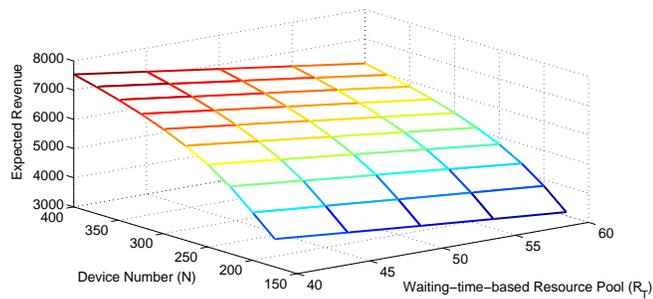


Fig. 11. Expected revenue under resource pool partitions ( $R_T + \overline{R_P} = 150$ ,  $\lambda = 20$ )

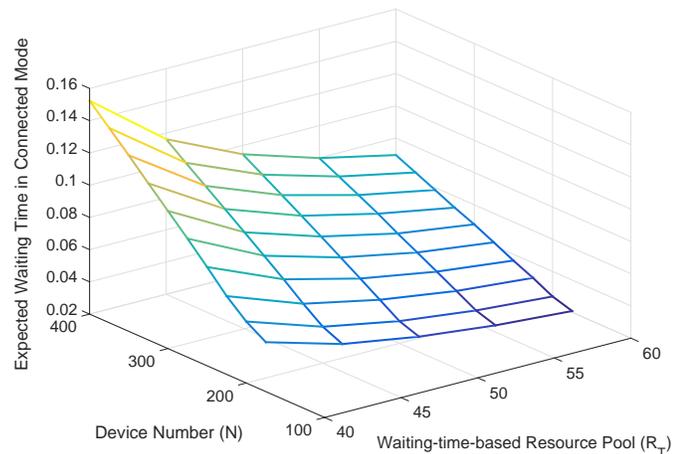


Fig. 12. Expected waiting time (delay) under resource pool partitions ( $R_T + \overline{R_P} = 150$ ,  $\lambda = 20$ )

payment for low-cost M2M and mission-critical M2M in a cellular network with H2H co-existence. H2H communications is always guaranteed with certain resources, whereas M2M communications has three alternatives: (1) being served with the free but unreliable resources in a first-come first-served manner, (2) paying for reliable direct access, and (3) staying idle. This framework not only serves low-cost M2M and mission-critical M2M together by time bidding and direct pay approaches, but also compensates the operator financially. Except for the eNB's announcement, no extra signaling exchange is required compared to the LTE standard. Furthermore, the existence and uniqueness of the Bayesian Nash equilibrium, and interregional and waiting-time-based truth-telling properties are showed. The proposed framework is an indirect mechanism with Bayesian-Nash incentive compatibility, interim efficiency, interim individual rationality and weakly budget balance. We further discuss the tradeoff between the expected revenue and the expected delay in connected mode. This work gives insights that dynamic pricing based on M2M/H2H traffic loads and resource pool partitions is one possible direction to serve mixed low-cost M2M and mission-critical M2M.

## REFERENCES

[1] METIS Official Website. Nov. 5, 2015. Retrieved from <https://www.metis2020.com/>, 2015.

[2] 3GPP TR 36.888 v12.0.0. Study on Provision of Low-cost Machine-Type Communications (MTC) User Equipments (UEs) based on LTE. Jun. 2013.

[3] R. Ratasuk, A. Prasad, Z. Li, A. Ghosh, and M. Uusitalo. Recent Advancements in M2M Communications in 4G Networks and Evolution Towards 5G. In *2015 18th ICIN*, pages 52–57, Feb.

[4] R. Liu, W. Wu, H. Zhu, and D. Yang. M2M-Oriented QoS Categorization in Cellular Network. In *7th WiCOM*, pages 1–5, Sept. 2011.

[5] A. Aijaz, M. Tshangini, M.R. Nakhai, X. Chu, and A-H. Aghvami. Energy-Efficient Uplink Resource Allocation in LTE Networks With M2M/H2H Co-Existence Under Statistical QoS Guarantees. *IEEE Trans. Commun.*, 62(7):2353–2365, Jul. 2014.

[6] H. S. Dhillon, H. C. Huang, H. Viswanathan, and R. A. Valenzuela. Power-Efficient System Design for Cellular-Based Machine-to-Machine Communications. *IEEE Trans. Wireless Commun.*, 12(11):5740–5753, Nov. 2013.

[7] A. Laya, L. Alonso, and J. Alonso-Zarate. Is the Random Access Channel of LTE and LTE-A Suitable for M2M Communications? A Survey of Alternatives. *IEEE Commun. Surveys Tuts.*, 16(1):4–16, First 2014.

[8] A. Rajandekar and B. Sikdar. A Survey of MAC Layer Issues and Protocols for Machine-to-Machine Communications. *IEEE Internet of Things Journal*, 2(2):175–186, Apr. 2015.

[9] 3GPP TR 37.868 v11.0.0. Study on RAN Improvements for Machine-type Communications. Sept. 2011.

[10] 3GPP RP-151621. New Work Item: NarrowBand IOT (NB-IOT). Sept. 2015.

[11] X. Cheng, L. Yang, and X. Shen. D2D for Intelligent Transportation Systems: A Feasibility Study. *IEEE Trans. Intell. Transp. Syst.*, 16(4):1784–1793, Aug. 2015.

[12] 3GPP RP-160677. SI: Further Enhancements LTE Device to Device, UE to Network Relays for Wearables. Mar. 2016.

[13] M.-Y. Cheng, G.-Y. Lin, H.-Y. Wei, and AC.-C. Hsu. Overload Control for Machine-Type-Communications in LTE-Advanced System. *IEEE Commun. Mag.*, 50(6):38–45, Jun. 2012.

[14] G.-Y. Lin and H.-Y. Wei. Auction-Based Random Access Load Control for Time-Dependent Machine-to-Machine Communications. *IEEE Internet of Things Journal*, PP(99):1–1, 2015.

[15] Z. Wang and V.W.S. Wong. Optimal Access Class Barring for Stationary Machine Type Communication Devices With Timing Advance Information. *IEEE Trans. Wireless Commun.*, 14(10):5374–5387, Oct. 2015.

[16] R.-G. Cheng, J. Chen, D.-W. Chen, and C.-H. Wei. Modeling and Analysis of an Extended Access Barring Algorithm for Machine-Type Communications in LTE-A Networks. *IEEE Trans. Wireless Commun.*, 14(6):2956–2968, Jun. 2015.

[17] C.-Y. Oh, D. Hwang, and T.-J. Lee. Joint Access Control and Resource Allocation for Concurrent and Massive Access of M2M Devices. *IEEE Trans. Wireless Commun.*, 14(8):4182–4192, Aug. 2015.

[18] G.-Y. Lin, S.-R. Chang, and H.-Y. Wei. Estimation and Adaptation for Bursty LTE Random Access. *IEEE Trans. Veh. Technol.*, 65(4):2560–2577, Apr. 2016.

[19] K.-D. Lee, S. Kim, and B. Yi. Throughput Comparison of Random Access Methods for M2M Service Over LTE Networks. In *2011 IEEE GC Wkshps*, pages 373–377, Dec. 2011.

[20] Y.-C. Pang, S.-L. Chao, G.-Y. Lin, and H.-Y. Wei. Network Access for M2M/H2H Hybrid Systems: A Game Theoretic Approach. *IEEE Commun. Lett.*, 18(5):845–848, Jun. 2014.

[21] D. Niyato, P. Wang, and D. I. Kim. Performance Modeling and Analysis of Heterogeneous Machine Type Communications. *IEEE Trans. Wireless Commun.*, 13(5):2836–2849, May 2014.

[22] M.-J. Shih, C.-Y. Yeh, K.D. Huang, and H.-Y. Wei. Energy-aware Waiting-line based Resource Allocation in Cellular Network with M2M/H2H Co-existence. In *IEEE ICC*, pages 2979–2984, Jun. 2015.

[23] Charles A. Holt Jr. and R. Sherman. Waiting-line Auctions. *The Journal of Political Economy*, pages 280–294, 1982.

[24] G. Wu, P. Ren, and C. Zhang. A Waiting-Time Auction Based Dynamic Spectrum Allocation Algorithm in Cognitive Radio Networks. In *IEEE GLOBECOM*, pages 1–5, Dec 2011.

[25] C.-H. Wei, G. Bianchi, and R.-G. Cheng. Modeling and Analysis of Random Access Channels With Bursty Arrivals in OFDMA Wireless Networks. *IEEE Trans. Wireless Commun.*, 14(4):1940–1953, Apr. 2015.

[26] Y. Narahari, D. Garg, R. Narayanam, and H. Prakash. *Game Theoretic Problems in Network Economics and Mechanism Design Solutions*. Springer-Verlag London, 1999.

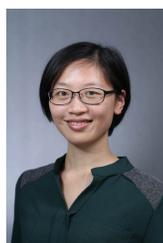


**Mei-Ju Shih** received the B.S. degree from the Department of Electrical Engineering, National Taiwan University (NTU), Taipei, Taiwan, in 2011. She is currently pursuing the Ph.D. degree in Graduate Institute of Communication Engineering at NTU. She was a graduate research intern in Intel Labs, Oregon, USA, during the summer of 2012 and 2013, respectively. She is a recipient of the Google Anita Borg Memorial Scholarship in 2014 and Pan Wen Yuan Foundation Scholarship in 2015. Her primary research interests include game theory, energy-

harvesting M2M communications, D2D communications, and protocol design for wireless communication networks.



**Kevin Dowhon Huang** received his B.S. degree from the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, in 2015. He joined Professor Hung-Yu Wei's lab since his junior year, researching on the topic of resource allocation problem in cellular networks. He is interested in the application of game theory in wireless networks. He got the admission to UCLA in 2016, where he will be pursuing his M.S. degree of Electrical Engineering, in the field of Signal and System.



**Chia-Yi Yeh** received the B.S. degree from the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, in 2014. She was a research assistant in Wireless Mobile Network Laboratory of National Taiwan University from 2014 to 2015. She is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA. Her research interests include protocol design and security issues in wireless communication systems.



**Hung-Yu Wei** received his B.S. degree in electrical engineering from National Taiwan University (NTU) in 1999. He received his M.S. and Ph.D. degrees in electrical engineering from Columbia University, in 2001 and 2005, respectively. He was a summer intern at Telcordia Applied Research in 2000 and 2001. He was with NEC Labs America from 2003 to 2005. He joined Department of Electrical Engineering at the National Taiwan University in July 2005. He is currently a Professor with the Department of Electrical Engineering and Graduate Institute of

Communication Engineering at National Taiwan University. His research interests include broadband wireless communications, vehicular networking, cross-layer design for wireless multimedia communications, Internet of Things, and game theoretic models for networking. He was a consulting member of the Acts and Regulation Committee of the National Communications Commission during 2008-2009. He actively participates in wireless communications standardization activities, and was a voting members in IEEE 802.16 working group.

He was the recipient of the Recruiting Outstanding Young Scholar Award from the Foundation for the Advancement of Outstanding Scholarship in 2006, K. T. Li Young Researcher Award from ACM Taipei Chapter and IICM in 2012, CIEE Excellent Young Engineer Award in 2014, and the NTU Excellent Teaching Award in 2008. He was awarded the Research Project for Excellent Young Scholars from Taiwan's Ministry of Science and Technology in 2014. He was also the recipient of the Wu Ta You Memorial Award from Ministry of Science and Technology in 2015. Currently, he is the chair of IEEE Vehicular Technology Society Taipei Chapter.