A supplementary document for the paper: Multi-channel Opportunistic Access for Heterogeneous Networks Based on Deep Reinforcement Learning

Xiaowen Ye[1], Yiding Yu [2], and Liqun Fu[1]
[1]School of Informatics, and the Key laboratory of Underwater Acoustic Communication and Marine Information Technology Ministry of Education, Xiamen University, Xiamen, China.
[2]Department of Information Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong Special Administrative Region, China.
E-mail: xiaowen@stu.xmu.edu.cn, 13058117978@163.com, liqun@xmu.edu.cn.

I. Introduction

This is a supplementary document to the paper: Multi-channel Opportunistic Access for Heterogeneous Networks Based on Deep Reinforcement Learning. In this document, we derive the optimal network throughputs when the Multi-channel Deep-reinforcement Learning Multiple Access (MC-DLMA) node coexists with the nodes using other protocols in various multi-channel heterogeneous wireless network (HetNet) scenarios. Then we use these optimal network throughputs as upper bound benchmarks for our paper.

We consider a multi-channel HetNet consisting of multiple types of networks (i.e., some radio networks and an emerging network) and an access point (AP). These radio networks use different MAC protocols, such as TDMA, q-ALOHA, FW-ALOHA, and MC-DLMA we designed using the deep reinforcement learning (DRL) technique. With the exception of the emerging network using the MC-DLMA protocol, each radio network has a dedicated channel for transmitting data packets to AP. We assume that different channels have different capacities: the capacity in Mbps of the TDMA channel is set to μ_1, the capacity in Mbps of the q-ALOHA channe is set to μ_2, and the capacity in Mbps of the FW-ALOHA channel is set to μ_3. Furthermore, we refer to the emerging network as the MC-DLMA network, and the nodes in the MC-DLMA network is referred to as the MC-DLMA nodes. The goal of the MC-DLMA nodes is to coexist in harmony with other nodes in multi-channel HetNets without knowing the MAC protocols of other nodes while efficiently employing the spectrum resources.

In order to get the optimal network throughputs in various scenarios, we replace the MC-DLMA nodes with the *model-aware* nodes. Unlike the MC-DLMA nodes which are *model-free*, the model-aware nodes know the MAC mechanisms of coexisting nodes in the network, and have different optimal spectrum access strategies in different coexistence scenarios. We will discuss the optimal spectrum access strategies achieved by the model-aware nodes in various scenarios in the following. *It is worth noting that the optimal spectrum access strategies in various scenarios that we deduced in this report are consistent with the simulation results (i.e., the spectrum access strategies of the DRL nodes in various scenarios) in our paper.*

A. Single-node Model-aware

1) Coexistence with TDMA networks: We consider the coexistence of one model-aware node with multiple TDMA nodes. We assume that all TDMA nodes use a total of X time slots out of Y time slots in a frame. In order to maximize the total network throughput, the optimal access policy of the model-aware node is *to take full advantage of the available time slots that are not used by TDMA nodes*. As a result, the optimal network throughput is μ_1.

2) Coexistence with q-ALOHA networks: We consider the coexistence of one model-aware node with N q-ALOHA nodes. The transmission probability of q-ALOHA node i ($i = 1, 2, \cdots, N$) in each time slot is denoted by q_i. To derive the optimal solution achieved by the model-aware node, we assume that the transmission probability of the model-aware node in each time slot is b. Then the total network throughput can be calculated as follows:

$$f(b) = \mu_2 b \prod_{i=1}^{N}(1-q_i) + \mu_2(1-b)\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big), \tag{1}$$

Thus

$$\frac{\mathrm{d}f(b)}{\mathrm{d}b} = \mu_2 \prod_{i=1}^{N}(1-q_i) - \mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big), \tag{2}$$

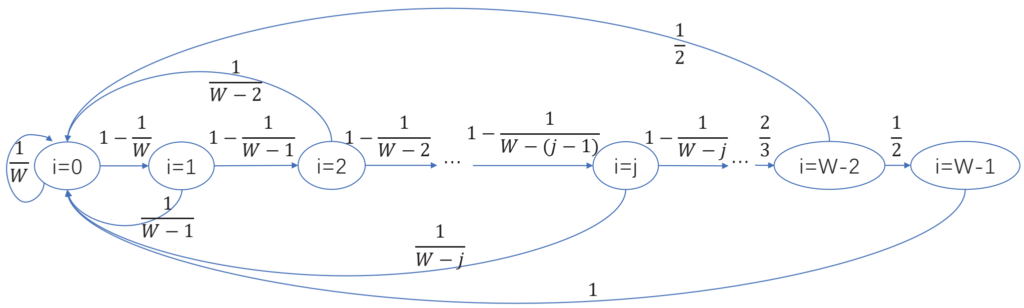

Fig. 1. Markov chain of the model-aware node when coexisting with the FW-ALOHA node.

$$\frac{\mathrm{d}^2 f(b)}{\mathrm{d}b^2} = 0, \tag{3}$$

indicating that $f(b)$ is convex in b. When $\mathrm{d}f(b)/\mathrm{d}b < 0$, if $b = 0$, $f(b)$ can get the maximum value; when $\mathrm{d}f(b)/\mathrm{d}b \geq 0$, if $b = 1$, $f(b)$ can get the maximum value. As a result, the optimal policy of the model-aware node is as follows:

The model-aware node does not transmit in each time slot if $\mathrm{d}f(b)/\mathrm{d}b < 0$*; it transmits in each time slot if* $\mathrm{d}f(b)/\mathrm{d}b \geq 0$.

Specifically, we have

$$b^* = \begin{cases} 0, & \text{if } \mathrm{d}f(b)/\mathrm{d}b < 0, \\ 1, & \text{otherwise.} \end{cases} \tag{4}$$

For convenience of illustration, let $z = \mathrm{d}f(b)/\mathrm{d}b = \mu_2 \prod_{i=1}^{N}(1-q_i) - \mu_2 \sum_{i=1}^{N} \big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)$. Then the optimal network throughputs are

$$\begin{cases} \mu_2 \sum\limits_{i=1}^{N} \big(q_i \prod\limits_{j=1_{j\neq i}}^{N} (1-q_j)\big), & \text{if } z < 0, \\ \mu_2 \prod\limits_{i=1}^{N} (1-q_i), & \text{otherwise.} \end{cases} \tag{5}$$

3) Coexistence with FW-ALOHA networks: We consider the coexistence of one model-aware node with one FW-ALOHA node with the contention window size W. As shown in Fig. 1, we construct a Markov chain to analyze the optimal network throughput. When the model-aware node stays in the t-th time slot, we define state i ($i = 0, 1, 2, \cdots, W-1$) of the Markov chain to be the number of continuous idle time slots of the FW-ALOHA node that has been observed by the model-aware node. The state transition probabilities are given in Fig. 1. We can get the state stationary probability p_i of state i by applying the balance equations as follows:

$$\begin{cases} p_0 = \frac{1}{W}p_0 + \frac{1}{W-1}p_1 + \cdots + \frac{1}{2}p_{W-2} + P_{W-1}, \\ p_1 = (1-\frac{1}{W})p_0, \\ p_2 = (1-\frac{1}{W-1})p_1, \\ \vdots \\ p_{W-2} = \frac{2}{3}p_{W-3}, \\ p_{W-1} = \frac{1}{2}p_{W-2}, \\ p_0 + p_1 + p_2 + \cdots + p_{W-1} = 1. \end{cases} \tag{6}$$

Thus, we can calculate the stationary probability p_i of state i ($i = 0, 1, 2, \cdots, W-1$) as

$$p_i = \frac{2(W-i)}{W(W+1)}. \tag{7}$$

When the model-aware node stays in state i and takes the action $a_i \in \{0, 1\}$ (i.e., if it transmits, $a_i = 1$; otherwise, $a_i = 0$), the FW-ALOHA node may transmit with the probability of $1/(W-i)$ or not transmit with the probability of $\big(1-1/(W-i)\big)$. Then, the total network throughput can be calculated as follows:

$$\begin{aligned} F(a_i) &= \sum_{i=0}^{W-1} \mu_3 \big(p_i a_i (1-\frac{1}{W-i})\big) + \mu_3 \sum_{i=0}^{W-1} \big(p_i(1-a_i)\frac{1}{W-i}\big) \\ &= \sum_{i=0}^{W-1} \mu_3 \big(\frac{2a_i(W-i-1)}{W(W+1)} + \frac{2(1-a_i)}{W(W+1)}\big) \\ &= \frac{2\mu_3}{W(W+1)} \sum_{i=0}^{W-1} \big((W-2)a_i - ia_i + 1\big). \end{aligned} \tag{8}$$

Thus, the goal of the model-aware node is to

$$\begin{aligned} &\text{maximize } F(a_i) \\ &\text{subject to } 0 \le \sum_{i=0}^{W-1} a_i \le W, \\ &a_i = 0 \text{ or } 1. \end{aligned} \tag{9}$$

Without loss of generality, let $\sum_{i=0}^{W-1} a_i = j$, then $j \in \{0, 1, 2, \cdots, W\}$. Because $a_i = 0$ or 1, thus

$$\sum_{i=0}^{W-1} i a_i \ge \big(0\times 1 + 1 \times 1 + \cdots + (j-1)\times 1\big) = \frac{j(j-1)}{2}. \tag{10}$$

Combining equations (8), (9) and (10), the objective becomes

$$\begin{aligned} &\text{maximize } F(j) = \frac{\mu_3\big(-j^2 + (2W-3)j + 2W\big)}{W(W+1)} \\ &\text{subject to } j \in \{0, 1, 2, \cdots, W\}. \end{aligned} \tag{11}$$

Further, we turn equation (11) into a convex optimization problem, i.e.,

$$\begin{aligned} &\text{minimize } -F(j) = -\frac{\mu_3\big(-j^2 + (2W-3)j + 2W\big)}{W(W+1)} \\ &\text{subject to } j \in \{0, 1, 2, \cdots, W\}, \end{aligned} \tag{12}$$

and thus,

$$-\frac{\mathrm{d}F(j)}{\mathrm{d}j} = \frac{\mu_3\big(2j-(2W-3)\big)}{W(W+1)}, \tag{13}$$

$$-\frac{\mathrm{d}^2F(j)}{\mathrm{d}j^2} = \frac{2\mu_3}{W(W+1)} > 0, \tag{14}$$

indicating that $-F(j)$ is convex in $j \in \{0, 1, 2, \cdots, W\}$. When $-\mathrm{d}F(j)/\mathrm{d}j = 0$, i.e., $j = (W-3/2)$, $-F(j)$ has the minimum value. Conversely, $F(j)$ has the maximum value in $j = (W-3/2)$. Specially, j should be an integer, thus $j = W-2$ or $j = W-1$. Especially, we let j be equal to $W-2$ in the following discussion. Therefore, for this coexistence scenario, the optimal access policy of the model-aware node can be summarized as follows:

After observing a transmission of the FW-ALOHA node, the model-aware node transmits in all subsequent time slots except after observing consecutive $W-2$ idle time slots of the FW-ALOHA node, in which case the model-aware node refrains from transmission in the next two time slots (i.e., if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will not transmit.).

Furthermore, by substituting $j = W-2$ into equation (11), we can get the optimal network throughput for the coexistence of one model-aware node with one FW-ALOHA node, i.e.,

$$\frac{\mu_3\big(W^2 - W + 2\big)}{W(W+1)}. \tag{15}$$

4) Coexistence with TDMA networks and q-ALOHA networks: We consider the coexistence of one model-aware node with multiple TDMA nodes and N q-ALOHA nodes. Let p denote the ratio of the number of time slots used by multiple TDMA nodes in a frame to the total number of time slots in a frame, and let q_i denote the transmission probability of q-ALOHA node i ($i = 1, 2, \cdots, N$).

At time slot t, if the TDMA channel is used by TDMA nodes, in order to maximize the total network throughput, the model-aware node will try to share the q-ALOHA channel with N q-ALOHA nodes as described in Section I-A2. We assume that the transmission probability of the model-aware node in each time slot is b when it coexists with q-ALOHA nodes, then the network throughput at time slot t can be calculated as

$$\mu_1 + \mu_2 b\Big(\prod_{i=1}^{N}(1-q_i)\Big) + \mu_2(1-b)\Big(\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big). \tag{16}$$

At time slot t, if the TDMA channel is not used by TDMA nodes, in order to maximize the total network throughput, the model-aware node either utilizes the TDMA channel or the q-ALOHA channel for transmission. Then the network throughput at time slot t can be calculated as

$$\begin{cases} \underbrace{\mu_1 + \mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)}_{\text{Case ①}}, & \text{if the model-aware node transmits on the TDMA channel,} \\ \underbrace{\mu_2\Big((1-b)\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big) + b\prod_{i=1}^{N}(1-q_i)\Big)}_{\text{Case ②}}, & \text{if the model-aware node transmits on the } q\text{-ALOHA channel.} \end{cases} \tag{17}$$

Therefore, the average network throughput $f(b)$ is

$$f(b) = \begin{cases} p\Big(\mu_1 + b\mu_2 \prod_{i=1}^{N}(1-q_i) + (1-b)\mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big) + (1-p)\Big(\mu_1 + \mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big), \\ p\Big(\mu_1 + b\mu_2 \prod_{i=1}^{N}(1-q_i) + (1-b)\mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big) + (1-p)\mu_2\Big((1-b)\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big) + b\prod_{i=1}^{N}(1-q_i)\Big). \end{cases}$$

Furthermore, we can get

$$f(b) = \begin{cases} pb\mu_2\Big(\prod_{i=1}^{N}(1-q_i) - \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big) + \mu_1 + \mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big), & \text{Case ①}, \\ b\mu_2\Big(\prod_{i=1}^{N}(1-q_i) - \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big) + p\mu_1 + \mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big), & \text{Case ②}. \end{cases} \tag{18}$$

and thus

$$\frac{\mathrm{d}f(b)}{\mathrm{d}b} = \begin{cases} p\mu_2\Big(\prod_{i=1}^{N}(1-q_i) - \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big), & \text{Case ①}, \\ \mu_2\Big(\prod_{i=1}^{N}(1-q_i) - \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big), & \text{Case ②}, \end{cases} \tag{19}$$

$$\frac{\mathrm{d}^2 f(b)}{\mathrm{d}b^2} = \begin{cases} 0, & \text{Case ①}, \\ 0, & \text{Case ②}, \end{cases} \tag{20}$$

indicating that $f(b)$ is convex in b. When $\mathrm{d}F(b)/\mathrm{d}b < 0$, if $b = 0$, $F(b)$ has the maximum value, i.e.,

$$\begin{cases} \mu_1 + \mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big), & \text{Case ①}, \\ p\mu_1 + \mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big), & \text{Case ②}. \end{cases} \tag{21}$$

Now, we compare Case ① with Case ② when $\mathrm{d}F(b)/\mathrm{d}b < 0$ and $b = 0$ as follows:

$$\text{Case ①} - \text{Case ②} = (1-p)\mu_1 > 0. \tag{22}$$

Thus, when $\mathrm{d}F(b)/\mathrm{d}b < 0$, i.e., $z = \prod_{i=1}^{N}(1-q_i) - \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big) < 0$, the optimal access policy of the model-aware node is as follows:

In the time slots when TDMA nodes do not transmit, the model-aware node transmits on the TDMA channel. In the time slots when TDMA nodes transmit, the model-aware node does not transmit.

Accordingly, the optimal network throughput when $z<0$ is

$$\mu_1 + \mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big). \tag{23}$$

When $\mathrm{d}F(b)/\mathrm{d}b \geq 0$, if $b = 1$, $F(b)$ has the maximum value, i.e.,

$$\begin{cases} \mu_1 + \mu_2 p\big(\prod_{i=1}^{N}(1-q_i)\big) + \mu_2(1-p)\Big(\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big), & \text{Case ①}, \\ p\mu_1 + \mu_2\big(\prod_{i=1}^{N}(1-q_i)\big), & \text{Case ②}. \end{cases} \tag{24}$$

Furthermore, we compare Case ① with Case ② when $\mathrm{d}F(b)/\mathrm{d}b \geq 0$ and $b = 1$ as follows:

$$\text{Case ①} - \text{Case ②} = (1-p)\Big(\mu_1 - \mu_2\Big(\prod_{i=1}^{N}(1-q_i) - \big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big)\Big) = (1-p)(\mu_1 - \mu_2 z). \tag{25}$$

Therefore, the optimal network throughput when $z\geq 0$ is

$$\begin{cases} \mu_1 + \mu_2 p\big(\prod_{i=1}^{N}(1-q_i)\big) + \mu_2(1-p)\Big(\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big), & 0\leq z<\mu_1/\mu_2, \\ p\mu_1 + \mu_2\big(\prod_{i=1}^{N}(1-q_i)\big), & z\geq \mu_1/\mu_2. \end{cases} \tag{26}$$

Accordingly, the optimal access policy of the model-aware node when $z\geq 0$ is as follows:

If $0\leq z<\mu_1/\mu_2$, in the time slots when TDMA nodes transmit, the model-aware node transmits on the q-ALOHA channel; in the time slots when TDMA nodes do not transmit, the model-aware node transmits on the TDMA channel. If $z\geq\mu_1/\mu_2$, the model-aware node transmits on the q-ALOHA channel in each time slot.

Finally, the optimal access policy of the model-aware node in this coexistence scenario can be summarized as follows:

- $z<0$. In the time slots when TDMA nodes do not transmit, the model-aware node transmits on the TDMA channel. In the time slots when TDMA nodes transmit, the model-aware node does not transmit.
- $0\leq z<\mu_1/\mu_2$. In the time slots when TDMA nodes transmit, the model-aware node transmits on the q-ALOHA channel; in the time slots when TDMA nodes do not transmit, the model-aware node transmits on the TDMA channel.
- $z\geq\mu_1/\mu_2$. The model-aware node transmits on the q-ALOHA channel in each time slot.

Furthermore, we summarize the optimal network throughputs for the coexistence of one model-aware node with multiple TDMA nodes and N q-ALOHA nodes in the following TABLE I.

TABLE I. The optimal network throughputs for the coexistence of one model-aware node with multiple TDMA nodes and N q-ALOHA nodes.

$z<0$	$\mu_1 + \mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)$
$0\leq z<\mu_1/\mu_2$	$\mu_1 + \mu_2 p\big(\prod_{i=1}^{N}(1-q_i)\big) + \mu_2(1-p)\Big(\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big)$
$z\geq\mu_1/\mu_2$	$p\mu_1 + \mu_2\big(\prod_{i=1}^{N}(1-q_i)\big)$

5) Coexistence with TDMA networks and FW-ALOHA networks: We consider the coexistence of one model-aware node with multiple TDMA nodes and one FW-ALOHA node. Let p denote the ratio of the number of time slots used by TDMA nodes to the total number of time slots in a frame, and let W denote the contention window size of the FW-ALOHA node.

At time slot t, if TDMA nodes transmit on the TDMA channel, in order to maximize the spectrum utilization, the model-aware node will coexist with the FW-ALOHA node as described in Section I-A3 (i.e., if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will not transmit.). Then the network throughput at time slot t can be calculated as follows:

$$\mu_1 + \underbrace{\mu_3 \frac{W^2-W+2}{W(W+1)}}_{\text{As given in (15)}}, \tag{27}$$

At time slot t, if TDMA nodes do not transmit on the TDMA channel, the model-aware node either utilizes the TDMA channel or the FW-ALOHA channel for transmission. Then the network throughput at time slot t can be calculated as

$$\begin{cases} \underbrace{\mu_1 + \mu_3 \frac{2}{W+1}}_{\text{Case ①}}, & \text{if the model-aware node transmits on the TDMA channel,} \\ \underbrace{\mu_3 \frac{W^2-W+2}{W(W+1)}}_{\text{Case ②}}, & \text{if the model-aware node transmits on the FW-ALOHA channel,} \end{cases} \tag{28}$$

where $2\mu_3/(W+1)$ is the average transmission probability of the FW-ALOHA node in each time slot [1], which can be obtained by substituting $j=0$ into equation (11). $\mu_3(W^2-W+2)/(W(W+1))$ is the average network throughput when the model-aware node coexists with the FW-ALOHA node, which is given in equation (15).

Therefore, when the model-aware node coexists with multiple TDMA nodes and one FW-ALOHA node, the optimal network throughput is the sum of the throughput of the TDMA channel and the throughput of the FW-ALOHA channel, i.e.,

$$\begin{cases} p\big(\mu_1 + \mu_3 \frac{W^2-W+2}{W(W+1)}\big) + (1-p)\big(\mu_1 + \mu_3 \frac{2}{W+1}\big), & \text{Case ①}, \\ p\big(\mu_1 + \mu_3 \frac{W^2-W+2}{W(W+1)}\big) + (1-p)\mu_3 \frac{W^2-W+2}{W(W+1)}, & \text{Case ②}, \end{cases} \tag{29}$$

Furthermore, we can get

$$\begin{cases} \mu_1 + \mu_3 p \frac{W^2-3W+2}{W(W+1)} + \frac{2\mu_3}{W+1}, & \text{Case ①}, \\ p\mu_1 + \mu_3 \frac{W^2-W+2}{W(W+1)}, & \text{Case ②}, \end{cases} \tag{30}$$

Now, we compare Case ① and Case ② as follows:

$$\text{Case ①} - \text{Case ②} = (1-p)\big(\mu_1 - \mu_3 \frac{W^2-3W+2}{W(W+1)}\big). \tag{31}$$

Therefore, the optimal network throughput is

$$\begin{cases} \mu_1 + \mu_3 p \frac{W^2-3W+2}{W(W+1)} + \frac{2\mu_3}{W+1}, & \frac{\mu_1}{\mu_3} \geq \frac{W^2-3W+2}{W(W+1)}, \\ p\mu_1 + \mu_3 \frac{W^2-W+2}{W(W+1)}, & \frac{\mu_1}{\mu_3} < \frac{W^2-3W+2}{W(W+1)}. \end{cases} \tag{32}$$

In summary, for this coexistence scenario, the optimal access policy of the model-aware node is as follows:

- $\mu_1/\mu_3 \geq (W^2-3W+2)/(W(W+1))$. In the time slots when TDMA nodes do not transmit, the model-aware node transmits on the TDMA channel; in the time slots when TDMA nodes transmit, the model-aware node coexists with the FW-ALOHA node as in Section I-A3, i.e., if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will not transmit.
- $\mu_1/\mu_3 < (W^2-3W+2)/(W(W+1))$. In each time slot, the model-aware node coexists with the FW-ALOHA node as in Section I-A3, i.e., if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will not transmit.

6) Coexistence with q-ALOHA networks and FW-ALOHA networks: We first consider the coexistence of one model-aware node with one q-ALOHA node and one FW-ALOHA node. Let q denote the transmission probability of the q-ALOHA node, and let W denote the contention window size of the FW-ALOHA node. As in Section I-A3, we also use the Markov chain shown in Fig. 1 to analyze the optimal network throughput in this coexistence scenario. State i ($i = 0, 1, 2, \cdots, W-1$) of the Markov chain is the same as defined in Section I-A3, and the corresponding state stationary probability p_i of state i is given in equation (7) (i.e., $p_i = 2(W-i)/(W(W+1))$).

When the model-aware node stays in state i, it may transmit on the FW-ALOHA channel, or transmit on the q-ALOHA channel, or not transmit on any channel. Specifically, let a_i denote the transmission decision of the model-aware node on the FW-ALOHA channel (i.e., if the model-aware node transmits on the FW-ALOHA channel, $a_i = 1$; otherwise, $a_i = 0$), and let b_i denote the transmission decision of the model-aware node on the q-ALOHA channel (i.e., if the model-aware node transmits on the q-ALOHA channel, $b_i = 1$; otherwise, $b_i = 0$). When the model-aware node stays in state i, the FW-ALOHA node may transmit with probability $1/(W-i)$ or not transmit with probability $(1 - 1/(W-i))$. For the q-ALOHA node, it may transmit with probability q or not transmit with probability $1-q$ in each time slot. Then, the total network throughput can be calculated as follows:

$$F(a_i, b_i) = \underbrace{\sum_{i=0}^{W-1} \big(p_i \mu_3 (1-a_i) \frac{1}{W-i}\big)}_{\text{The throughput of the FW-ALOHA node}} + \underbrace{\sum_{i=0}^{W-1} \big(p_i \mu_2 (1-b_i) q\big)}_{\text{The throughput of } q\text{-ALOHA nodes}} + \underbrace{\sum_{i=0}^{W-1} p_i \big(\mu_3 a_i (1 - \frac{1}{W-i}) + \mu_2 b_i (1-q)\big)}_{\text{The throughput of the model-aware node}}, \tag{33}$$

where $a_i + b_i \leq 1$, i.e., the model-aware node can transmit data packets on at most one channel in each time slot. By substituting equation (7) into equation (33), we can get

$$F(a_i, b_i) = \frac{2}{W(W+1)} \sum_{i=0}^{W-1} \big(\mu_3 (W-2) a_i - \mu_3 i a_i + \mu_2 (1-2q) W b_i - \mu_2 (1-2q) i b_i + \mu_3 + \mu_2 q W - \mu_2 q i\big). \tag{34}$$

Thus, the goal of the model-aware node is to maximize $F(a_i, b_i)$, i.e.,

$$\begin{gathered} \text{maximize } F(a_i, b_i) \\ \text{subject to } a_i, b_i \in \{0,\ 1\} \text{ and } a_i + b_i \leq 1. \end{gathered} \tag{35}$$

Without loss of generality, let $\sum_{i=0}^{W-1} a_i = n$, then $n \in \{0,\ 1,\ 2,\ \cdots,\ W\}$. Because $a_i = 0$ or 1, thus

$$\sum_{i=0}^{W-1} i a_i \geq \big(0 \times 1 + 1 \times 1 + \cdots + (n-1) \times 1\big) = \frac{n(n-1)}{2}. \tag{36}$$

Similarly, let $\sum_{i=0}^{W-1} b_i = m$, then $m \in \{0,\ 1,\ 2,\ \cdots,\ W\}$. Because $b_i = 0$ or 1, thus

$$\sum_{i=0}^{W-1} i b_i \geq \big(0 \times 1 + 1 \times 1 + \cdots + (m-1) \times 1\big) = \frac{m(m-1)}{2}. \tag{37}$$

Combining equations (35), (36), and (37), we can get

$$F(a_i, b_i) \leq \frac{2}{W(W+1)} \Big(\underbrace{\mu_3 \big(-\frac{n^2}{2} + (W - \frac{3}{2}) n\big)}_{①} + \underbrace{\mu_2 (1-2q) \big(-\frac{m^2}{2} + (W + \frac{1}{2}) m\big)}_{②} + \mu_3 W + \mu_2 q \frac{W(W+1)}{2}\Big). \tag{38}$$

First, from equation (38), we can find that if $q > 1/2$, in order to maximize the total network throughput (i.e., $F(a_i, b_i)$), the model-aware node will never transmit on the q-ALOHA channel (i.e., $m = 0$). Thus, our goal becomes

$$\begin{gathered} \text{maximize } \frac{2}{W(W+1)} \Big(\mu_3 \big(-\frac{n^2}{2} + (W - \frac{3}{2}) n\big) + \mu_3 W + \mu_2 q \frac{W(W+1)}{2}\Big) \\ \text{subject to } n \in \{0,\ 1,\ \cdots,\ W\}. \end{gathered} \tag{39}$$

Obviously, when $n = W - 2/3$, the objective function in equation (39) will be maximized. Specially, n should be an integer, thus $n = W-2$ or $n = W-1$. Especially, we let n be equal to $W-2$ in the following discussion. Therefore, for this coexistence scenario, the optimal access policy of the model-aware node can be summarized as follows: if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will not transmit. Specifically, when the model-aware node stays in state i, its transmission decisions a_i and b_i are

$$\begin{cases} a_i = 1, & \text{for } i = 0,\ 1,\ \cdots,\ W-3, \\ b_i = 0, & \text{for any } i. \end{cases} \tag{40}$$

By substituting equation (40) into equation (34), we can get the optimal network throughput for the case where $q{>}1/2$ in this coexistence scenario as follows:

$$\mu_3\frac{W^2-W+2}{W(W+1)}+\mu_2 q. \tag{41}$$

Now, we consider the optimal access strategy for the case where $q{\leq}1/2$. It is obvious that part ① in equation (38) will be maximized when $n=W-2$[1]. In other words, the model-aware node transmits on the FW-ALOHA channel when it observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, thus part ① in equation (38) will be maximized. Similarly, part ② in equation (38) will be maximized when $m=W$ (i.e., the model-aware node transmits on the q-ALOHA channel in each time slot.). However, the model-aware node can only access one channel in each time slot. In other words, in the same state i, a_i and b_i cannot be equal to 1 at the same time. Thus, for the case where $n=m=W-2$, we compare part ① with part ② as follows:

$$\begin{aligned}\text{Part ①} - \text{part ②} &= \mu_3\Big(-\mu_2\frac{(W-2)^2}{2}+(W-\frac{3}{2})(W-2)\Big)-(1-2q)\Big(-\frac{(W-2)^2}{2}+(W+\frac{1}{2})(W-2)\Big)\\ &=\frac{1}{2}(W-2)\big(\mu_3(W-1)-\mu_2(1-2q)(W+3)\big),\end{aligned} \tag{42}$$

where $W{\geq}2$. We can find that when $q{>}(W\mu_2-W\mu_3+3\mu_2+\mu_3)/2\mu_2(W+3)$, the value of part ① is greater than the value of part ②; when $q{\leq}(W\mu_2-W\mu_3+3\mu_2+\mu_3)/2\mu_2(W+3)$ the value of part ① is not greater than the value of part ②. Therefore, we can get the optimal access strategy of the model-aware node for the case where $q{\leq}1/2$ as below:

* If $(W\mu_2-W\mu_3+3\mu_2+\mu_3)/2\mu_2(W+3){<}q{\leq}1/2$, the optimal access strategy is as follows: if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will transmit on the q-ALOHA channel. Specifically, when the model-aware node stays in state i, its transmission decisions a_i and b_i are

$$\begin{cases} a_i=1, & \text{for } i=0,\,1,\,\cdots,\,W-3,\\ b_i=1, & \text{for } i=W-2,\,W-1.\end{cases} \tag{43}$$

By substituting equation (43) into equation (34), we can get the corresponding optimal network throughput for this case as follows:

$$\frac{\mu_3(W^2-W+2)+\mu_2(qW^2+qW-12q+6)}{W(W+1)}. \tag{44}$$

* If $q{\leq}(W\mu_2-W\mu_3+3\mu_2+\mu_3)/2\mu_2(W+3)$, the optimal access strategy is as follows: the model-aware node transmits on the q-ALOHA channel in each time slot, and the FW-ALOHA channel is only used by the FW-ALOHA node. Specifically, when the model-aware node stays in state i, its transmission decisions a_i and b_i are

$$\begin{cases} a_i=0, & \text{for any } i,\\ b_i=1, & \text{for any } i.\end{cases} \tag{45}$$

By substituting equation (45) into equation (34), we can get the corresponding optimal network throughput for this case as follows:

$$\mu_2(1-q)+\frac{2\mu_3}{1+W}. \tag{46}$$

Overall, the optimal spectrum access strategies of the model-aware node can be summarized as follows:

* $q{>}1/2$*: if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than* $W-2$*, it will transmit on the FW-ALOHA channel; otherwise, it will not transmit. The q-ALOHA channel is only used by the q-ALOHA node.*
* $(W\mu_2-W\mu_3+3\mu_2+\mu_3)/2\mu_2(W+3){<}q{\leq}1/2$*: if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than* $W-2$*, it will transmit on the FW-ALOHA channel; otherwise, it will transmit on the q-ALOHA channel.*
* $0{<}q{\leq}(W\mu_2-W\mu_3+3\mu_2+\mu_3)/2\mu_2(W+3)$*: the model-aware node transmits on the q-ALOHA channel in each time slot, and the FW-ALOHA channel is only used by the FW-ALOHA node.*

With the above optimal access strategies, the optimal network throughputs for the coexistence of one model-aware node with one q-ALOHA node and one FW-ALOHA node can be summarized in the following TABLE II.

Extensively, we further consider the case of N q-ALOHA nodes in the q-ALOHA network[2]. Let q_j denote the transmission probability of q-ALOHA node j ($j=1,2,\cdots,N$), and let $z=\prod_{j=1}^{N}(1-q_j)-\sum_{j=1}^{N}\big(q_j\prod_{k=1_{k\neq j}}^{N}(1-q_k)\big)$. Similar to the

[1]Part ① in equation (38) will be also maximized when $n=W-1$, i.e., the total network throughput in the case of $n=W-1$ is the same as that in the case of $n=W-2$. For convenience of illustration, we let n be equal to $W-2$ in the following discussion.

[2]It is difficult to get the optimal analytical throughput for general multiple FW-ALOHA nodes, so this paper only shows the optimal analytical throughput when the model-aware node coexists with one FW-ALOHA node.

TABLE II. The optimal network throughputs for the coexistence of one model-aware node with one q-ALOHA node and one FW-ALOHA node.

$q>1/2$	$\mu_2 q+\mu_3\frac{W^2-W+2}{W(W+1)}$
$\frac{W\mu_2-W\mu_3+3\mu_2+\mu_3}{2\mu_2(W+3)}<q\leq 1/2$	$\frac{\mu_3(W^2-W+2)+\mu_2(qW^2+qW-12q+6)}{W(W+1)}$
$0<q\leq\frac{W\mu_2-W\mu_3+3\mu_2+\mu_3}{2\mu_2(W+3)}$	$\mu_2(1-q)+\frac{2\mu_3}{W+1}$

case where there is only a single q-ALOHA node in the network, when the model-aware node coexists with N q-ALOHA nodes and one FW-ALOHA node, its optimal access strategies are summarized as follows:

* $z<0$: *if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will not transmit. The q-ALOHA channel is only used by q-ALOHA nodes. Thus, the optimal network throughput is*

$$\underbrace{\mu_2\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)}_{\text{The throughput of the } q\text{-ALOHA channel}}+\underbrace{\mu_3\frac{W^2-W+2}{W(W+1)}}_{\text{The throughput of the FW-ALOHA channel}}. \tag{47}$$

* $0\leq z<\mu_3(W-1)/\mu_2(W+3)$: *if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will transmit on the q-ALOHA channel. Thus, the optimal network throughput is*

$$\underbrace{\mu_2\prod_{i=1}^{N}(1-q_i)\frac{6}{W(W+1)}+\mu_2\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\frac{(W+3)(W-2)}{W(W+1)}}_{\text{The throughput of the } q\text{-ALOHA channel}}+\underbrace{\mu_3\frac{W^2-W+2}{W(W+1)}}_{\text{The throughput of the FW-ALOHA channel}}. \tag{48}$$

* $z\geq\mu_3(W-1)/\mu_2(W+3)$: *the model-aware node transmits on the q-ALOHA channel in each time slot, and the FW-ALOHA channel is only used by the FW-ALOHA node. Thus, the optimal network throughput is*

$$\underbrace{\mu_2\prod_{i=1}^{N}(1-q_i)}_{\text{The throughput of the } q\text{-ALOHA channel}}+\underbrace{\frac{2\mu_3}{W+1}}_{\text{The throughput of the FW-ALOHA channel}}. \tag{49}$$

With the above optimal access strategies, the optimal network throughputs for the coexistence of one model-aware node with N q-ALOHA nodes and one FW-ALOHA node can be summarized in TABLE III.

TABLE III. The optimal network throughputs for the coexistence of one model-aware node with N q-ALOHA nodes and one FW-ALOHA node.

$z<0$	$\mu_2\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)+\mu_3\frac{W^2-W+2}{W(W+1)}$
$0\leq z<\frac{\mu_3(W-1)}{\mu_2(W+3)}$	$\mu_2\prod_{i=1}^{N}(1-q_i)\frac{6}{W(W+1)}+\mu_2\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\frac{(W+3)(W-2)}{W(W+1)}+\mu_3\frac{W^2-W+2}{W(W+1)}$
$z\geq\frac{\mu_3(W-1)}{\mu_2(W+3)}$	$\mu_2\prod_{i=1}^{N}(1-q_i)+\frac{2\mu_3}{W+1}$

7) Coexistence with TDMA networks, q-ALOHA networks, and FW-ALOHA networks: We consider the coexistence of one model-aware node with multiple TDMA nodes, N q-ALOHA nodes, and one FW-ALOHA node. Let p denote the ratio of the number of time slots used by TDMA nodes to the total number of time slots in a frame, q_i denote the transmission probability of q-ALOHA node i ($i=1,\cdots,N$), and W denote the contention window size of the FW-ALOHA node. For convenience of illustration, let $z=\prod_{i=1}^{N}(1-q_i)-\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)$ as defined in Section I-A4.

It is difficult to get the optimal analytical throughput for three types of channels with different capacities in this coexistence scenario, so this paper only shows the optimal analytical throughput when the model-aware node coexists with multiple TDMA nodes, N q-ALOHA nodes, and one FW-ALOHA node in the scenario of channels with the same capacity (i.e., $\mu_1=\mu_2=\mu_3=\mu$). Based on the results discussed in Section I-A6, the optimal access policy of the model-aware node in this coexistence scenario can be summarized as follows: in the time slots when TDMA nodes do not transmit, the model-aware node transmit on the TDMA channel; in the time slots when TDMA nodes transmit, the model-aware node coexists with q-ALOHA nodes and/or the FW-ALOHA node as analyzed in Section I-A6. Specifically, we discuss the optimal access strategy in this coexistence scenario in detail as below. Note that, the following classification conditions, i.e., $z<0$, $0\leq z<(W-1)/(W+3)$, and $(W-1)/(W+3)\leq z$, are obtained from Section I-A6 when $\mu_1=\mu_2=\mu_3=\mu$.

* $z<0$: As analyzed in Section I-A6, when one model-aware node coexists with N q-ALOHA nodes and one FW-ALOHA node, if $z<0$, the model-aware node will never occupy the q-ALOHA channel. Therefore, when one model-aware node coexists with multiple TDMA nodes, N q-ALOHA nodes, and one FW-ALOHA node, if $z<0$, the optimal access policy of the model-aware node is the same as in Section I-A5: in the time slots when TDMA nodes do not transmit, the model-aware node transmits on the TDMA channel; in the time slots when TDMA nodes transmit, the model-aware node will coexist with the FW-ALOHA node (i.e., if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will not transmit.). As a result, the q-ALOHA channel is only used by q-ALOHA nodes. Then the optimal network throughput can be calculated as follows:

$$\mu p\frac{(W^2-3W+2)}{W(W+1)}+\mu\frac{W+3}{W+2}+\mu\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big). \tag{50}$$

The sum of the first term in equation (50) and the second term in equation (50) is the optimal network throughput under the coexistence of the model-aware node with TDMA nodes and one FW-ALOHA node, as given in equation (32) when $\mu_1=\mu_3=\mu$. The last term in equation (50) is the optimal throughput of N q-ALOHA nodes on the q-ALOHA channel.

* $0\leq z<(W-1)/(W+3)$: According to Section I-A6, we can get the optimal access policy of the model-aware node as follows: in the time slots when TDMA nodes do not transmit, the model-aware node transmits on the TDMA channel. As a result, the q-ALOHA channel is always occupied by q-ALOHA nodes, and the FW-ALOHA channel is always occupied by the FW-ALOHA node. In the time slots when TDMA nodes transmit, if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will transmit on the q-ALOHA channel. Thus, the TDMA channel is used in every time slot, and the average throughput of the TDMA channel is 1. For the q-ALOHA channel, the average throughput of it can be calculated as follows:

$$\underbrace{\mu(1-p)\Big(\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big)}_{\text{The case where the TDMA channel is used by the model-aware node.}}\quad\underbrace{+\mu p\Big(\underbrace{\prod_{i=1}^{N}(1-q_i)\frac{6}{W(W+1)}+\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\frac{(W+3)(W-2)}{W(W+1)}}_{\text{As given in the first two items of equation (48) when }\mu_2=1}\Big)}_{\text{The case where the TDMA channel is used by the TDMA node.}}. \tag{51}$$

For the FW-ALOHA channel, the average throughput of it can be calculated as follows:

$$\underbrace{\mu(1-p)\Big(\frac{2}{W+1}\Big)}_{\text{The case where the TDMA channel is used by the model-aware node.}}\quad+\quad\underbrace{\mu p\frac{W^2-W+2}{W(W+1)}}_{\text{The case where the TDMA channel is used by the TDMA node.}}\quad, \tag{52}$$

where $2/(W+1)$ is the average transmission probability of the FW-ALOHA node in eahc time slot, which can be obtained by substituting $j=0$ into equation (11) when $\mu_3=\mu$. $(W^2-W+2)/(W(W+1))$ is the average network throughput when the model-aware node coexists with the FW-ALOHA node, which is given as equation (15) when $\mu_3=\mu$.
Then the average throughput of the overall network is the sum of the average throughput of the TDMA network, the average throughput of the q-ALOHA network, and the average throughput of the FW-ALOHA network, i.e.,

$$\begin{aligned}&\mu+\mu(1-p)\Big(\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big)+\mu p\Big(\prod_{i=1}^{N}(1-q_i)\frac{6}{W(W+1)}+\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\frac{(W+3)(W-2)}{W(W+1)}\Big)+\\&\mu(1-p)\Big(\frac{2}{W+1}\Big)+p\frac{W^2-W+2}{W(W+1)}\\=&\mu\frac{W+3}{W+1}+\mu\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)+\mu p\frac{W^2-3W+2+6\prod_{i=1}^{N}(1-q_i)-6\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)}{W(W+1)}.\end{aligned} \tag{53}$$

* $z\geq(W-1)/(W+3)$: According to Section I-A6, the optimal access policy of the model-aware node is as follows: in the time slots when TDMA nodes do not transmit, the model-aware node transmits on the TDMA channel; in the time slots

when TDMA nodes transmit, the model-aware node will transmit on the q-ALOHA channel. As a result, the FW-ALOHA channel is only used by the FW-ALOHA node. Then the optimal network throughput can be calculated as follows:

$$\begin{aligned} &\mu+\mu p\big(\prod_{i=1}^{N}(1-q_i)\big)+\mu(1-p)\Big(\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big)+\mu\frac{2}{W+1}\\ =&\mu p\big(\prod_{i=1}^{N}(1-q_i)\big)+\mu(1-p)\Big(\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big)+\mu\frac{W+3}{W+1}, \end{aligned} \tag{54}$$

The sum of the first two terms in equation (54) is the optimal network throughput under the coexistence of the model-aware node with multiple TDMA nodes and N q-ALOHA nodes, which is given as equation (24) when $\mu_1=\mu_2=\mu$. The last item in equation (54) is the average transmission probability of the FW-ALOHA node in each time slot, which can be obtained by substituting $j=0$ into equation (11) when $\mu_3=\mu$.

Overall, we summarize the optimal network throughputs for the coexistence of one model-aware node with multiple TDMA nodes, N q-ALOHA nodes, and one FW-ALOHA node in TABLE IV.

TABLE IV. The optimal network throughputs for the coexistence of one model-aware node with multiple TDMA nodes, q-ALOHA nodes, and one FW-ALOHA node.

$z<0$	$\mu\frac{p(W^2-3W+2)+W(W+3)}{W(W+1)}+\mu\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)$
$0\le z<\frac{W-1}{W+3}$	$\mu\frac{W+3}{W+1}+\mu\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)+\mu p\frac{W^2-3W+2+6\prod_{i=1}^{N}(1-q_i)-6\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)}{W(W+1)}$
$z\ge\frac{W-1}{W+3}$	$\mu\big(\prod_{i=1}^{N}(1-q_i)\big)+\mu(1-p)\Big(\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big)+\mu\frac{W+3}{W+1}$

B. *Multi-node model-aware*

In this subsection, we derive the optimal network throughput achieved by multiple model-aware nodes. In order to maximize the total network throughput, we assume that multiple model-aware nodes are aware of each other and can fully cooperate to employ the under-utilized spectrum resource. The cooperation fashion among the model-aware nodes is centralized, i.e., one model-aware node in the network is designated as the gateway, which associates with other model-aware nodes and coordinates the model-aware network with other radio networks. In each time slot, the designated gateway decides whether the model-aware nodes should access the channel or not. If the designated gateway finds K $(0\le K\le N)$ channels that can be accessed, it will choose K model-aware nodes out of all the model-aware nodes in a round-robin manner to access K channels, respectively. As a result of the transmission, each selected model-aware node will receive a corresponding feedback/reward from the environment (i.e., AP), and then communicates with its gateway with this information. The coordination information between each designated gateway and other model-aware nodes belonging to its part can be transmitted through a control channel. For example, the control channel can be implemented as a quite short time slot after each time slot of information transmission [2]. If the designated gateway decides not to transmit, all model-aware nodes will keep silent. We give the optimal access strategies of the model-aware nodes in various coexistence scenarios as follows.

1) Coexistence with TDMA networks and q-ALOHA networks: We consider the coexistence of multiple model-aware nodes with multiple TDMA nodes and N q-ALOHA nodes. Let p denote the ratio of the number of time slots used by TDMA nodes to the total number of time slots in a frame, and q_i denote the transmission probability of q-ALOHA node i $(i=1,\cdots,N)$. For this coexistence scenario, we assume that multiple model-aware nodes will cooperate to employ the under-utilized spectrum resource as described above. In order to maximize the total network throughput, the model-aware nodes will transmit on the TDMA channel in a round-robin fashion in the time slots when TDMA nodes do not transmit, and they will do not transmit in the time slots when TDMA nodes transmit. Meanwhile, for those model-ware nodes that are not transmitting on the TDMA channel, let b denote their total probability of transmitting on the q-ALOHA channel. Then the optimal network throughput can be calculated as follows:

$$\begin{aligned} f(b)=&\underbrace{\mu_1 p+\mu_1(1-p)}_{\text{The throughput of the TDMA channel}}\underbrace{+\mu_2 b\prod_{i=1}^{N}(1-q_i)+\mu_2(1-b)\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)}_{\text{The throughput of the } q\text{-ALOHA channel}}\\ =&\mu_1+\mu_2 b\prod_{i=1}^{N}(1-q_i)+\mu_2(1-b)\sum_{i=1}^{N}\big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\big), \end{aligned} \tag{55}$$

thus,

$$\frac{\mathrm{d}f(b)}{\mathrm{d}b} = \mu_2 \prod_{i=1}^{N}(1-q_i) - \mu_2 \sum_{i=1}^{N} \big(q_i \prod_{j=1_{j\neq i}}^{N} (1-q_j)\big), \tag{56}$$

$$\frac{\mathrm{d}^2 f(b)}{\mathrm{d}b^2} = 0, \tag{57}$$

indicating that $f(b)$ is convex in b. When $\mathrm{d}f(b)/\mathrm{d}b < 0$, if $b=0$, $f(b)$ can get the maximum value; when $\mathrm{d}f(b)/\mathrm{d}b \geq 0$, if $b=1$, $f(b)$ can get the maximum value. For convenience of illustration, let $z = \prod_{i=1}^{N}(1-q_i) - \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)$. Thus, the optimal spectrum access strategy of the model-aware nodes is as follows:

The model-aware nodes transmit on the TDMA channel in a round-robin fashion in the time slots when TDMA nodes do not transmit, and they do not transmit in the time slots when TDMA nodes transmit. For those model-aware nodes that are not transmitting on the TDMA channel, if z<0, they transmit on the q-ALOHA channel in a round-robin manner in each time slot; otherwise, they do not transmit in each time slot.

Furthermore, TABLE V summarizes the optimal network throughputs for this coexistence scenario.

TABLE V. The optimal network throughputs for the coexistence of multiple model-aware nodes with multiple TDMA nodes and N q-ALOHA nodes.

z<0	$\mu_1 + \mu_2 \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)$
z≥0	$\mu_1 + \mu_2 \prod_{i=1}^{N}(1-q_i)$

2) Coexistence with TDMA networks and FW-ALOHA networks: We consider the coexistence of multiple model-aware nodes, multiple TDMA nodes, and one FW-ALOHA node. Let p denote the ratio of the number of time slots to the total number of time slots in a frame, and let W denote the contention window size of the FW-ALOHA node. In order to make full use of the underutilized spectrum, we assume that multiple model-aware nodes will cooperate to employ the under-utilized spectrum resource as described above. For this coexistence scenario, the optimal spectrum access policy of multiple model-aware nodes is as follows:

The model-aware nodes transmit on the TDMA channel in a round-robin fashion in the time slots when TDMA nodes do not transmit, and they do not transmit in the time slots when TDMA nodes transmit. For the model-aware nodes that are not transmitting on the TDMA channel, if they observe that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, they transmit on the FW-ALOHA channel in a round-robin manner; otherwise, they do not transmit.

Then the optimal network throughput can be calculated as follows:

$$\underbrace{\mu_1}_{\text{The throughput of the TDMA channel}} + \underbrace{\mu_3 \frac{W^2 - W + 2}{W(W+1)}}_{\text{The throughput of the FW-ALOHA channel}}, \tag{58}$$

where $\mu_3(W^2 - W + 2)/(W(W+1))$ is the average network throughput when the model-aware node coexists with the FW-ALOHA node, which is given as equation (15).

3) Coexistence with q-ALOHA networks and FW-ALOHA networks: We consider the coexistence of multiple model-aware nodes with N q-ALOHA nodes and one FW-ALOHA node. Let q_i denote the transmission probability of q-ALOHA node i ($i = 1, \cdots, N$), and let W denote the contention window size of the FW-ALOHA node. For convenience of illustration, let $z = \prod_{i=1}^{N}(1-q_i) - \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)$ as defined in Section I-A2. As analyzed in Section I-A2, when the model-aware node coexists with q-ALOHA nodes, if z<0, the optimal access strategy of the model-aware node is not to transmit on the q-ALOHA channel in each time slot; otherwise, the optimal access strategy of the model-aware node is to transmit in each time slot. In order to take full advantage of the underutilized spectrum, we assume that multiple model-aware nodes in this coexistence scenario will cooperate to employ the under-utilized spectrum resource as described above. Therefore, for this coexistence scenario, the optimal spectrum access policy of the model-aware nodes can be summarized as follows:

The model-aware nodes transmit on the q-ALOHA channel in a round-robin fashion in each time slot if z<0, and otherwise they do not transmit in each time slot. For those model-aware nodes that are not transmitting on the q-ALOHA channel, if they observe that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, they transmit on the FW-ALOHA channel in a round-robin fashion; otherwise, they do not transmit.

With the above optimal access strategy, the optimal network throughputs for this coexistence scenario are summarized in TABLE VI, where $\mu_3(W^2 - W + 2)/(W(W+1))$ is the average network throughput when the model-aware nodes coexist with the FW-ALOHA node, which is given in equation (15).

TABLE VI. The optimal network throughputs for the coexistence of multiple model-aware nodes with N q-ALOHA nodes and one FW-ALOHA node.

$z<0$	$\mu_2 \sum_{i=1}^{N} \left(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\right) + \mu_3 \frac{W^2-W+2}{W(W+1)}$
$z\geq 0$	$\mu_2 \prod_{i=1}^{N}(1-q_i) + \mu_3 \frac{W^2-W+2}{W(W+1)}$

4) Coexistence with TDMA networks, q-ALOHA networks, and FW-ALOHA networks: We consider the coexistence of L model-aware nodes with multiple TDMA nodes, N q-ALOHA nodes, and one FW-ALOHA node. Let p denote the ratio of the number of time slots used by TDMA nodes to the total number of time slots in a frame, q_i denote the transmission probability of q-ALOHA node i ($i = 0, 1, \cdots, N$), and W denote the contention window size of the FW-ALOHA node. For convenience of illustration, let $z = \prod_{i=1}^{N}(1-q_i) - \sum_{i=1}^{N}\left(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\right)$ as defined in Section I-A4. In this coexistence scenario, the optimal access policy is different when $L = 2$ and when $L>2$. Therefore, we will discuss these two coexistence scenarios, respectively.

(1) $L = 2$. *It is difficult to get the optimal analytical throughput for three types of channels with different capacities in this coexistence scenario, so this paper only shows the optimal analytical throughput when two model-aware nodes coexist with multiple TDMA nodes, N q-ALOHA nodes, and one FW-ALOHA node in the scenario of channels with the same capacity (i.e.,* $\mu_1 = \mu_2 = \mu_3 = \mu$*).* When two model-aware nodes coexist with TDMA nodes, q-ALOHA nodes, and the FW-ALOHA node, the optimal policy of the model-aware nodes depends on z and W. In particular, we consider three cases in the following.

* *The first access policy* is analyzed as follows: one of two model-aware nodes coexists with TDMA nodes and q-ALOHA nodes as described in Section I-A4, i.e., in the time slots when TDMA nodes do not transmit, it transmits on the TDMA channel; in the time slots when TDMA nodes transmit, it transmits on the q-ALOHA channel. For the other model-aware node, it coexists with the FW-ALOHA node as described in Section I-A3. Then the optimal network throughput can be calculated as follows:

$$\underbrace{\mu + \mu + p\prod_{i=1}^{N}(1-q_i) + \mu(1-p)\sum_{i=1}^{N}\left(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\right)}_{\text{As given in equation (24) when } \mu_1=\mu_2=\mu} + \underbrace{\mu\frac{W^2-W+2}{W(W+1)}}_{\text{As given in equation (15) when } \mu_3=\mu}$$
$$= \mu p\prod_{i=1}^{N}(1-q_i) + \mu(1-p)\sum_{i=1}^{N}\left(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\right) + \mu\frac{2(W^2+1)}{W(W+1)}. \tag{59}$$

* *The second access policy* is analyzed as follows: one of two model-aware nodes coexists with TDMA nodes and the FW-ALOHA node as described in Section I-A5, i.e., in the time slots when TDMA nodes do not transmit, it transmits on the TDMA channel; in the time slots when TDMA nodes transmit, it coexists with the FW-ALOHA node. For the other model-aware node, it transmits on the q-ALOHA channel in each time slot because of $z\geq 0$, as described in Section I-A2. Then the optimal network throughput can be calculated as follows:

$$\underbrace{\mu\frac{W+3}{W+1} + \mu p\frac{W^2-3W+2}{W(W+1)}}_{\text{As given in equation (32) when } \mu_1=\mu_3=\mu} + \underbrace{\mu\prod_{i=1}^{N}(1-q_i)}_{\text{As given in equation (5) when } \mu_2=\mu}. \tag{60}$$

* *The third access policy* is analyzed as follows: one of two model-aware nodes coexists with TDMA nodes on the TDMA channel as analyzed in Section I-A1. For the other model-aware node, it coexists with q-ALOHA nodes and the FW-ALOHA node in different time slots as analyzed in Section I-A6, i.e.,
  (1) When $z<0$, if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will not transmit. The q-ALOHA channel is only used by q-ALOHA nodes.
  (2) When $0\leq z<(W-1)/(W+3)$, if the model-aware node observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will transmit on the q-ALOHA channel.
  (3) When $z\geq(W-1)/(W+3)$, the model-aware node transmits on the q-ALOHA channel in each time slot, and the FW-ALOHA channel is only used by the FW-ALOHA node.

Then the optimal network throughput can be calculated as follows:

$$\begin{cases} \underbrace{\mu}_{\text{TDMA channel}} + \underbrace{\mu\frac{W^2-W+2}{W(W+1)} + \mu\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)}_{\text{The sum throughput of the } q\text{-ALOHA channel and the FW-ALOHA channel, as given in (47) when } \mu_2=\mu_3=\mu}, & \text{if } z<0, \\ \underbrace{\mu}_{\text{TDMA channel}} + \underbrace{\mu\frac{W^2-W+2}{W(W+1)} + \prod_{i=1}^{N}(1-q_i)\frac{6\mu}{W(W+1)} + \mu\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\frac{W^2+W-6}{W(W+1)}}_{\text{The sum throughput of the } q\text{-ALOHA channel and the FW-ALOHA channel, as given in (48) when } \mu_2=\mu_3=\mu}, & \text{if } 0\leq z<\frac{W-1}{W+3}, \\ \underbrace{\mu}_{\text{TDMA channel}} + \underbrace{\mu\prod_{i=1}^{N}(1-q_i) + \mu\frac{2}{W+1}}_{\text{The sum throughput of the } q\text{-ALOHA channel and the FW-ALOHA channel, as given in (49) when } \mu_2=\mu_3=\mu}, & \text{if } z\geq\frac{W-1}{W+3} \end{cases} \tag{61}$$

Furthermore, equation (61) can be simplified to

$$\begin{cases} \mu\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big) + \mu\frac{2W^2+2}{W(W+1)}, & \text{if } z<0, \\ \mu\prod_{i=1}^{N}(1-q_i)\frac{6}{W(W+1)} + \mu\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\frac{(W+3)(W-2)}{W(W+1)} + \mu\frac{2W^2+2}{W(W+1)}, & \text{if } 0\leq z<\frac{W-1}{W+3}, \\ \mu\prod_{i=1}^{N}(1-q_i) + \mu\frac{W+3}{W+1}, & \text{if } z\geq\frac{W-1}{W+3}. \end{cases} \tag{62}$$

First, we compare equation (59) and equation (60) as follows:

$$\begin{aligned} (59)-(60) &= \mu(1-p)\Big(\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big) - \big(\prod_{i=1}^{N}(1-q_i)\big) + \frac{W^2-3W+2}{W(W+1)}\Big) \\ &= \mu(1-p)\big(\frac{(W-1)(W-2)}{W(W+1)} - z\big) \end{aligned} \tag{63}$$

Thus, when $z<\big((W^2-3W+2)/(W(W+1))\big)$, the network throughput achieved using the first access policy is greater than that using the second access policy; otherwise, the network throughput achieved using the second access policy is greater than that using the first access policy.

Second, we compare equation (59) and equation (62) as follows:

$$(59)-(62) = \begin{cases} \mu pz<0, & \text{if } z<0, \\ \mu\big(p-\frac{6}{W(W+1)}\big)z, & \text{if } 0\leq z<\frac{W-1}{W+3}, \\ \mu\frac{(W-1)(W-2)}{W(W+1)} - z(1-p), & \text{if } z\geq\frac{W-1}{W+3}. \end{cases} \tag{64}$$

Finally, we compare equation (60) and equation (62) as follows:

$$(60)-(62) = \begin{cases} \mu\big(z-\frac{(1-p)(W-1)(W-2)}{W(W+1)}\big)<0, & \text{if } z<0, \\ \mu\frac{W-2}{W(W+1)}\big((W+3)z-(1-p)(W-1)\big), & \text{if } 0\leq z<\frac{W-1}{W+3}, \\ \mu p\frac{(W-1)(W-2)}{W(W+1)}>0, & \text{if } z\geq\frac{W-1}{W+3}. \end{cases} \tag{65}$$

We can find from equation (64) and equation (65) that when $z<0$, the third policy is optimal, i.e., the optimal access policy of two model-aware nodes is as follows: one of two model-aware nodes coexists with TDMA nodes on the TDMA channel as analyzed in Section I-A1; for the other model-aware node, if it observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$, it will transmit on the FW-ALOHA channel; otherwise, it will not transmit.

We can find from equation (65) that when $z\geq\big((W-1)/(W+3)\big)$, the network throughput achieved using the second access policy is greater than the network throughput achieved using the third access policy. Furthermore, according to equation (63), when $z \geq \big((W-1)(W-2)/(W(W+1))\big)$, the network throughput achieved using the second access policy is greater than the network throughput achieved using the first access policy. Since

$$\frac{W-1}{W+3} > \frac{W^2-3W+2}{W(W+1)}, \tag{66}$$

we can get that when $z \geq \big((W-1)/(W+3)\big)$, the second access policy is optimal, i.e., one of two model-aware nodes coexists with TDMA nodes and the FW-ALOHA node as described in Section I-A5, and the other model-aware node transmits on the q-ALOHA channel in each time slot. Therefore, the corresponding optimal network throughput is given in equation (60).

Now, we discuss the case when $0 \leq z < \big((W-1)/(W+3)\big)$ in the following.

We can find from equation (63) that when $z < \big((W-1)(W-2)/(W(W+1))\big)$, the network throughput achieved using the first access policy is greater than that using the second access policy; otherwise, the network throughput achieved using the first access policy is not greater than that using the second access policy. Further, according to equation (64), when $p > \big(6/W(W+1)$, the network throughput achieved using the first access policy is greater than that using the third access policy; otherwise, the network throughput achieved using the first access policy is not greater than that using the third access policy. According to equation (65), we can get that when $z > \big((W-1)(1-p)/(W+3)\big)$, the network throughput achieved using the second access policy is greater than that using the third access policy; otherwise, the network throughput achieved using the second access policy is not greater than that using the third access policy. In summary, when $0 \leq z < \big((W-1)/(W+3)\big)$, we can get

$$\begin{cases} \text{The first policy} > \text{the second policy}, & \text{if } z < \frac{(W-1)(W-2)}{W(W+1)}; \\ \text{The first policy} > \text{the third policy}, & \text{if } p > \frac{6}{W(W+1)}; \\ \text{The second policy} > \text{the third policy}, & \text{if } z > \frac{(W-1)(1-p)}{W+3}. \end{cases} \tag{67}$$

Furthermore, we find that when $p \geq \big(6/(W(W+1))\big)$,

$$0 \leq \frac{(W-1)(1-p)}{W+3} < \frac{(W-2)(W-1)}{W(W+1)} < \frac{W-1}{W+3}. \tag{68}$$

Similarly, when $p < \big(6/(W(W+1))\big)$, we can get

$$0 \leq \frac{(W-2)(W-1)}{W(W+1)} < \frac{(W-1)(1-p)}{W+3} < \frac{W-1}{W+3}. \tag{69}$$

Thus, we can get that when $0 \leq z < \big((W-1)/(W+3)\big)$,

$$\text{if } p \geq \frac{6}{W(W+1)} \begin{cases} \text{The first policy is optimal}, & \text{if } 0 \leq z < \frac{(W-1)(W-2)}{W(W+1)}; \\ \text{The second policy is optimal}, & \text{if } \frac{(W-1)(W-2)}{W(W+1)} \leq z < \frac{W-1}{W+3}. \end{cases} \tag{70}$$

$$\text{if } p < \frac{6}{W(W+1)} \begin{cases} \text{The third policy is optimal}, & \text{if } 0 \leq z < \frac{(W-1)(1-p)}{W+3}; \\ \text{The second policy is optimal}, & \text{if } \frac{(W-1)(1-p)}{W+3} \leq z < \frac{W-1}{W+3}. \end{cases} \tag{71}$$

Overall, we summarize the optimal spectrum access strategies for different cases as below.

* if $z<0$, the optimal access policy of two model-aware nodes is as follows: *one of two model-aware nodes coexists with TDMA nodes on the TDMA channel as analyzed in Section I-A1; for the other model-aware node, if it observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than* $W-2$*, it will transmit on the FW-ALOHA channel; otherwise, it will not transmit.*
* If $z \geq 0$ and $p \geq 6/\big(W(W+1)\big)$, then the optimal access strategy of two model-aware nodes is as follows: *when* $0 \leq z < (W-1)(W-2)/\big(W(W+1)\big)$*, one of two model-aware nodes coexists with TDMA nodes and q-ALOHA nodes as described in Section I-A4, and the other model-aware node coexists with the FW-ALOHA node as described in Section I-A3. When* $z \geq (W-1)(W-2)/\big(W(W+1)\big)$*, one of two model-aware nodes coexists with TDMA nodes and the FW-ALOHA node as described in Section I-A5, and the other model-aware node transmits on the q-ALOHA channel in each time slot as described in Section I-A2.*
* If $z \geq 0$ and $p < 6/\big(W(W+1)\big)$, the optimal access strategy of two model-aware nodes is as follows: *when* $0 \leq z < (W-1)(1-p)/\big(W(W+1)\big)$*, one of two model-aware nodes coexists with TDMA nodes on the TDMA channel as analyzed in Section I-A1. For the other model-aware node, if it observes that the number of consecutive idle time slots of the FW-ALOHA channel is less than* $W-2$*, it will transmit on the FW-ALOHA channel; otherwise, it will transmit on the q-ALOHA channel. When* $z \geq (W-1)(1-p)/\big(W(W+1)\big)$*, one of two model-aware nodes coexists with TDMA nodes and the FW-ALOHA node as described in Section I-A5, and the other model-aware node transmits on the q-ALOHA channel in each time slot as described in Section I-A2.*

With the above optimal access strategies, the corresponding optimal network throughputs for the coexistence of two model-aware nodes with multiple TDMA nodes, N q-ALOHA nodes, and one FW-ALOHA node are summarized in TABLE VII.

TABLE VII. The optimal network throughput for the coexistence of two model-aware nodes with multiple TDMA nodes, N q-ALOHA nodes, and one FW-ALOHA node.

$z<0$	$\mu\sum_{i=1}^{N}\left(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\right)+\mu\frac{2W^2+2}{W(W+1)}$
If $p\geq\frac{6}{W(W+1)}$	
$0\leq z<\frac{W^2-3W+2}{W(W+1)}$	$\mu p\prod_{i=1}^{N}(1-q_i)+\mu(1-p)\sum_{i=1}^{N}\left(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\right)+\mu\frac{2W^2+2}{W(W+1)}$
$z\geq\frac{W^2-3W+2}{W(W+1)}$	$\mu\frac{W+3}{W+1}+\mu p\frac{W^2-3W+2}{W(W+1)}+\mu\prod_{i=1}^{N}(1-q_i)$
If $p<\frac{6}{W(W+1)}$	
$0\leq z<\frac{(W-1)(1-p)}{W(W+1)}$	$\mu\frac{2W^2+2}{W(W+1)}+\mu\prod_{i=1}^{N}(1-q_i)\frac{6}{W(W+1)}+\mu\sum_{i=1}^{N}\left(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\right)\frac{(W+3)(W-2)}{W(W+1)}$
$z\geq\frac{(W-1)(1-p)}{W(W+1)}$	$\mu\frac{W+3}{W+1}+\mu p\frac{W^2-3W+2}{W(W+1)}+\mu\prod_{i=1}^{N}(1-q_i)$

(2) $L>2$. When more than two model-aware nodes coexist with multiple TDMA nodes, N q-ALOHA nodes, and one FW-ALOHA node, the optimal access strategy of the model-aware nodes depends on z. In particular, we consider two cases in the following.

Case a). $z<0$: As analyzed in Section I-A4, when one model-aware node coexists with TDMA nodes and q-ALOHA nodes, if $z<0$, in order to maximize the total network throughput, the model-aware node will never occupy the q-ALOHA channel. Therefore, when L model-aware nodes coexist with TDMA nodes, q-ALOHA nodes, and the FW-ALOHA node, if $z<0$, the optimal access policy of these model-aware nodes is as follows:

The model-aware nodes transmit on the TDMA channel in a round-robin fashion in the time slots when TDMA nodes do not transmit, and they do not transmit in the time slots when TDMA nodes transmit. For those model-aware nodes that are not transmitting on the TDMA channel, if they observe that the number of consecutive idle time slots of the FW-ALOHA channel is less than $W-2$*, they transmit on the FW-ALOHA channel in a round-robin manner; otherwise, they do not transmit. As a result, the* q*-ALOHA channel is only used by* q*-ALOHA nodes.*

Then the optimal network throughput is

$$\underbrace{\mu_1}_{\text{The throughput of the TDMA channel}} + \underbrace{\mu_2\sum_{i=1}^{N}\Big(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\Big)}_{\text{The throughput of the } q\text{-ALOHA channel}} + \underbrace{\mu_3\frac{W^2-W+2}{W(W+1)}}_{\text{The throughput of the FW-ALOHA channel}}. \tag{72}$$

Case b). $z\geq 0$: In order to maximize the network throughput, we assume that all model-aware nodes can cooperate with each other as described above. Then the optimal access policy of these model-aware nodes is as follows:

The model-aware nodes transmit on the TDMA channel in a round-robin fashion in the time slots when TDMA nodes do not transmit, and they do not transmit in the time slots when TDMA nodes transmit. The other model-aware nodes transmit on the q*-ALOHA channel in a round-robin fashion in each time slot. For those model-aware nodes that are neither transmitting on the TDMA channel nor on the* q*-ALOHA channel, if they observe that the number of consecutive idle time slots of the FW-ALOHA channel is less than* $W-2$*, they transmit on the FW-ALOHA channel in a round-robin manner; otherwise, they do not transmit.*

The corresponding optimal network throughput can be calculated as follows:

$$\underbrace{\mu_1}_{\text{The throughput of the TDMA channel}} + \underbrace{\mu_2\prod_{i=1}^{N}(1-q_i)}_{\text{The throughput of the } q\text{-ALOHA channel}} + \underbrace{\mu_3\frac{W^2-W+2}{W(W+1)}}_{\text{The throughput of the FW-ALOHA channel}}. \tag{73}$$

With the above optimal access strategies, the optimal total network throughputs for the coexistence of more than two model-aware nodes with multiple TDMA nodes, N q-ALOHA nodes, and one FW-ALOHA node are summarized in TABLE VIII.

TABLE VIII. The optimal network throughputs for the coexistence of more than two model-aware nodes with multiple TDMA nodes, N q-ALOHA nodes, and one FW-ALOHA node.

$z<0$	$\mu_1+\mu_2\sum_{i=1}^{N}\left(q_i\prod_{j=1_{j\neq i}}^{N}(1-q_j)\right)+\mu_3\frac{W^2-W+2}{W(W+1)}$
$z\geq 0$	$\mu_1+\mu_2\prod_{i=1}^{N}(1-q_i)+\mu_3\frac{W^2-W+2}{W(W+1)}$

REFERENCES

[1] G. Bianchi, L. Fratta, and M. Oliveri, "Performance evaluation and enhancement of the CSMA/CA MAC protocol for 802.11 wireless LANs," in *Proc. IEEE PIMRC*, Oct. 1996, pp. 392–396.

[2] X. Ye, Y. Yu, and L. Fu, "The optimal network throughputs when the model-aware node coexists with other nodes using different mac protocols," *arXiv preprint arXiv:2008.11621*, 2020.