

# Distributed Fair Scheduling for Information Exchange in Multi-Agent Systems

Majid Raeis,<sup>1</sup> S. Jamaloddin Golestani<sup>2</sup>

<sup>1</sup> University of Toronto, Toronto, Canada

<sup>2</sup> Sharif University of Technology, Tehran, Iran  
m.raeis@utoronto.ca, golestani@sharif.edu

## Abstract

Information exchange is a crucial component of many real-world multi-agent systems. However, the communication between the agents involves two major challenges: the limited bandwidth, and the shared communication medium between the agents, which restricts the number of agents that can simultaneously exchange information. While both of these issues need to be addressed in practice, the impact of the latter problem on the performance of the multi-agent systems has often been neglected. This becomes even more important when the agents' information or observations have different importance, in which case the agents require different priorities for accessing the medium and sharing their information. Representing the agents' priorities by fairness weights and normalizing each agent's share by the assigned fairness weight, the goal can be expressed as equalizing the agents' normalized shares of the communication medium. To achieve this goal, we adopt a queueing theoretic approach and propose a distributed fair scheduling algorithm for providing weighted fairness in single-hop networks. Our proposed algorithm guarantees an upper-bound on the normalized share disparity among any pair of agents. This can particularly improve the short-term fairness, which is important in real-time applications. Moreover, our scheduling algorithm adjusts itself dynamically to achieve a high throughput at the same time. The simulation results validate our claims and comparisons with the existing methods show our algorithm's superiority in providing short-term fairness, while achieving a high throughput.

## Introduction

Many real-world multi-agent systems rely on the information exchange between the agents. However, the communication of the agents involves major practical limitations such as the shared communication medium and the limited bandwidth. Whether the information exchange is required for the coordination of agents, or solely the transfer of information from one point to another, the agents cannot transmit their messages simultaneously over the same communication medium. This becomes even more challenging when there is no coordinator to arbitrate access to the medium and therefore, a distributed algorithm is required for scheduling the agents' transmissions. Fairness is a key

concern in distributed scheduling algorithms, where some agents might hog the communication medium for a long period of time and therefore, lead to starvation of the other agents. Moreover, the agents' messages might have different importance or different latency requirements. Information exchange in multi-agent reinforcement learning (MARL) is one such example, where the importance of each agent's partially observed information can be different from one another (Kim et al. 2018). Weighted fair queueing, which has long been used as an effective basis for resource allocation and scheduling, is a natural candidate for dealing with these cases. The main goal of the weighted fair queueing algorithms is to provide service in proportion to some specified service shares (also known as fairness weights) to the competing users of a shared resource. The Generalized Processor Sharing (GPS) scheme (Parekh and Gallager 1993) serves as the reference for most of the existing fair queueing algorithms in the literature. The GPS algorithm uses a fluid flow model in which all the users can receive service simultaneously. However, because of the fluid flow assumption, the GPS scheme is not applicable in many real-world applications such as information exchange. Consequently, alternative fair queueing algorithms such as (Golestani 1994; Goyal, Vin, and Cheng 1997) have been proposed to mimic the behaviour of the GPS algorithm in real-world packet-based applications. However, the majority of these algorithms are centralized in implementation and cannot be used directly to provide fair queueing in a distributed manner.

In the information exchange context, the communication medium is the shared resource that needs to be scheduled among the agents. In wireless networks, the medium access control (MAC) protocol is responsible for scheduling the users. IEEE 802.11 DCF<sup>1</sup> (Wi-Fi) is the dominant distributed MAC protocol in wireless networks. Although there has been some effort in designing distributed fair scheduling algorithms based on 802.11 DCF, almost all these proposed methods take heuristic approaches to emulate a particular centralized fair queueing algorithm without providing any guarantees or theoretical analysis. Moreover, they suffer from short-term unfairness, which is particularly im-

<sup>1</sup>We use the terms 802.11 DCF and Wi-Fi interchangeably for referring to the Distributed Coordination Function (DCF) protocol in IEEE 802.11

portant in real-time applications. In contrast, we propose a distributed fair scheduling algorithm, which provides deterministic guarantees on the normalized share disparity among any pair of agents. Specifically, we propose a distributed scheduling algorithm for information exchange in single-hop wireless networks, which guarantees a bounded disparity between the normalized services received by any two agents. To the best of our knowledge, this is the first time that a distributed scheduling algorithm is capable of providing a bounded service disparity among the users. This particularly improves the short-term fairness of our proposed algorithm compared to the existing ones. Furthermore, our algorithm dynamically adjusts itself to balance the trade-off between the fairness and the network throughput.

### A Brief Background on 802.11 Wi-Fi

Wi-Fi is the dominant protocol for wireless communication scheduling, which is based on CSMA (Carrier-Sense Multiple Access). The main idea of CSMA is that each agent is required to listen to the medium before its transmission. The agent is allowed to transmit its message only if the medium is idle. However, to avoid collisions after busy periods, it uses the Binary Exponential Backoff (BEB) mechanism, which requires the agents to choose random backoffs in the interval  $[0, CW]$ . Whenever the medium becomes idle, each agent discretizes the time into time slots of predefined length (denoted by  $\sigma$ ) and decrements its backoff counter by one after each idle time slot. The backoff counter will be frozen during the busy periods. When the counter reaches zero, the agent is allowed to attempt a transmission. The parameter  $CW$  is called the contention window, which is doubled by the agent if its transmission is unsuccessful.  $CW$  is reset to its default value once the agent transmits its message successfully. Moreover, 802.11 uses interframe spaces (IFS), which are specified waiting periods between transmission of frames, to prioritize different traffic types (such as control and data).

### Related Work

Efficient and fair communication between agents of a system is a challenging task. Particularly, it is a key component in applications that require agents' coordination, such as cooperative multi-agent reinforcement learning. Most of the existing studies on multi-agent communication only focus on the bandwidth limitation of the medium and ignore the shared medium contention between the agents (Goldman and Zilberstein 2003; Mao et al. 2020; Foerster et al. 2016; Zhang and Lesser 2013). For instance, a message pruning mechanism is proposed by (Mao et al. 2020) to improve the bandwidth efficiency of the recent deep RL-based communication methods for multi-agent systems (Foerster et al. 2016; Jiang and Lu 2018; Peng, Zhang, and Luo 2018; Singh, Jain, and Sukhbaatar 2019). In another work, (Zhang, Zhang, and Lin 2019) introduce Variance Based Control (VBC) technique for efficient communication in MARL, which limits the variance of the exchanged messages between the agents in the training phase. (Goldman and Zilberstein 2003) have developed a theoretical model for decentralized control of a cooperative multi-agent system with communication. This

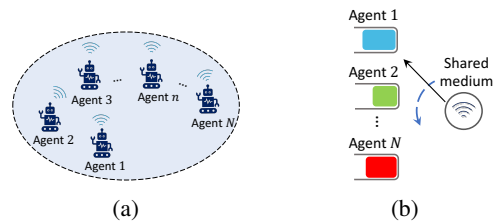


Figure 1: Modeling the communication medium as a shared server in a queuing model that needs to be scheduled between the agents.

model can be used for studying the trade-off between the value and cost (including bandwidth cost) of the exchanged information, as well as its effect on the joint utility of the agents. While almost all these works focus on the bandwidth efficiency of the communication and ignore the shared medium contention between the agents, (Kim et al. 2018) study the problem of communication scheduling in MARL. The authors propose SchedNet for scheduling the communication of the agents, which uses a weight-based scheduling algorithm to prioritize the agents with more important information.

On the other hand, the scheduling problem has been well studied in the area of wireless communication. Since Wi-Fi is the dominant distributed MAC protocol in wireless networks, most of the existing distributed scheduling algorithms have been proposed based on this protocol. In particular, these algorithms use one of the parameters of the 802.11, such as  $CW$  size (Qiao and Shin 2002), IFS (Shreedhar and Varghese 1996; Lee, Liao, and Chen 2007) or backoff intervals (Vaidya et al. 2005), to emulate centralized fair queueing service disciplines in a distributed manner. A common problem with these algorithms is the short-term unfairness, which is mainly caused by the random backoffs that are used in the collision resolution mechanism of 802.11 (Choi, Yoo, and Kim 2008). In order to improve the short-term fairness, (Dugar, Vaidya, and Bahl 2001) propose a protocol that uses pulse transmissions for giving priority to the agents that have experienced collision in their last attempt. Although this protocol improves the short-term fairness, the order in which packets are served could still drastically change from the reference centralized fair queueing algorithm. Moreover, almost all these methods use heuristic methods to mimic the behaviour of the centralized fair queueing algorithms and do not provide any guarantees or theoretical backings on the fairness or the network throughput.

### System Model and Problem Formulation

Consider a network consisting of  $N$  agents that need to communicate with each other to achieve a common goal. The set of all agents is represented by  $\mathcal{N}$  and each agent  $n \in \mathcal{N}$  has a fairness weight of  $\phi_n$ , which corresponds to its share of the communication medium. Fairness weights are chosen independently and are based on the importance of each agent's information. We assume that the agents can sense the medium and therefore, avoid transmitting their

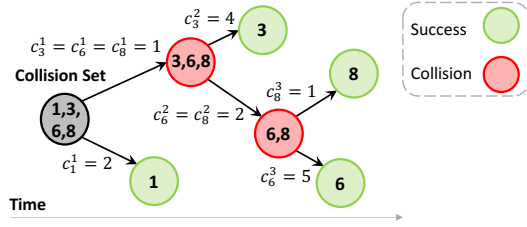


Figure 2: An example of using splitting algorithm (with  $m = 2$ ) for contention resolution.

messages when the medium is busy. Without loss of generality and solely to simplify the discussion, we assume that all the agents can instantaneously detect if the communication medium becomes busy or idle. Although in practice the agents are only able to detect the medium's status with some delay, the concept of interframe spaces in the 802.11 DCF protocol can be easily merged into our algorithm to address this issue for real-world applications.

In order to study this problem, we model the communication medium as a shared resource (server) that needs to be scheduled among the agents. Specifically, each agent has a transmission queue for storing the information that needs to be sent (Fig. 1). At a given time, the agents with non-empty queues are called *backlogged agents*, while the ones with no information to share are called *absent agents*. Let  $\mathcal{B}(t)$  denote the set of backlogged agents at time  $t$ , and  $\mathcal{A}(t)$  represent the absent agents. In the same way,  $\mathcal{B}(t_1, t_2)$  and  $\mathcal{A}(t_1, t_2)$  respectively represent the set of agents that are backlogged and absent during the whole interval  $(t_1, t_2)$ . Each time an agent successfully transmits a packet, it receives a service equal to the transmitted packet size. We define  $W_k(t)$  as the total service provided to agent  $k$  during  $(0, t)$ . Based on our definition, if a packet of agent  $k$  is under transmission at  $t$ ,  $W_k(t)$  includes any part of it transmitted by time  $t$ . The *normalized service* received by agent  $k$ ,  $w_k(t)$ , is defined as the aggregate service provided to this agent during  $(0, t)$  normalized to its fairness weight, i.e.,  $w_k(t) = W_k(t)/\phi_k, k \in \mathcal{N}$ . Furthermore, we define bivariate process  $w_k(t_1, t_2)$  as  $w_k(t_1, t_2) = w_k(t_2) - w_k(t_1), k \in \mathcal{N}, t_2 \geq t_1 \geq 0$ .

Now, our goal can be formulated as designing a distributed scheduling algorithm that guarantees a bounded normalized service disparity among any pair of backlogged agents, i.e.,

$$|w_k(t_1, t_2) - w_j(t_1, t_2)| \leq \varepsilon, \quad k, j \in \mathcal{B}(t_1, t_2), \quad (1)$$

while achieving a high throughput at the same time. Since our proposed algorithm implements the centralized *Self-Clocked Fair Queueing* algorithm (Golestani 1994) in a distributed manner to provide the above guarantee, we refer to it as Distributed SCFQ (DSCFQ) in the rest of the paper.

## Distributed SCFQ (DSCFQ)

An important source of short-term unfairness in the existing scheduling schemes is that the agents with unsuccessful last attempts have no priority over other agents. In order to

address this issue, we split the backlogged agents into two classes: class I that is defined as the set of agents that have to retransmit their messages because of their unsuccessful last attempts, and class II agents, which are the agents with new information to share. Moreover, each agent maintains a *collision counter*, *scaling factor* and a *compensation factor*. Collision counter keeps track of the number of collisions that the agent has experienced since its last successful transmission. Scaling factor is a parameter for suitable scaling of the backoff intervals, in order to reach an acceptable throughput. Finally, compensation factor is a parameter that is required for providing short-term fairness, which compensates the fairness deviations that are caused by the discrete backoff intervals.

Now, let us introduce our distributed scheduling algorithm for providing weighted fairness in a multi-agent system. Let  $p_k^i$  denote the  $i$ th message<sup>2</sup> (packet) of agent  $k$ . Moreover, the arrival time, transmission start time and transmission finish time of packet  $p_k^i$  are denoted by  $a_k^i, s_k^i$  and  $d_k^i$ , respectively. Now, each agent takes the following actions to transmit the backlogged messages in its queue.

### Class II agents:

1. At the time message  $p_k^i$  reaches the front of agent  $k$ 's transmitter queue, it is tagged with backoff tag  $B_k^i$ , calculated as follows:

$$B_k^i = \left\lceil \alpha \left( \frac{L_k^i}{\phi_k} - \epsilon_k^i \right) \right\rceil, \quad (2)$$

where  $L_k^i$  and  $\alpha$  are the size of packet  $p_k^i$  and scaling factor, respectively, and  $\epsilon_k^i$  represents the compensation factor, which is defined as

$$\epsilon_k^1 = 0, \quad \epsilon_k^i = \epsilon_k^{i-1} + \left( \frac{B_k^{i-1}}{\alpha} - \frac{L_k^{i-1}}{\phi_k} \right). \quad (3)$$

2. To send  $p_k^i$ , agent  $k$  picks a backoff interval equal to  $B_k^i$  slots. Class II agents can start decrementing their backoff counters only after sensing the medium idle for one time slot. The backoff counter is decremented by one after each idle slot and is frozen during the busy periods.
3. Whenever an agent's backoff reaches zero, it starts the transmission process for the intended destination.
4. If a transmitting class II agent detects a collision, it stops the transmission, increments its collision counter by 1 and starts following the procedure for class I agents.

### Class I agents:

1. **Splitting Algorithm for Collision Resolution:** In contrast to class II agents, each class I agent can access the medium as soon as it becomes idle, without waiting for an extra idle time slot. In this way, the agents involved in the collision get prior access to the medium. However, to avoid a guaranteed collision with other class I agents, we employ a mechanism based on the splitting algorithm (Bertsekas and Gallager 1992) for resolving the collision.

<sup>2</sup>Throughout the paper, we use the terms message and packet interchangeably.

In particular, agent  $k$  chooses a uniformly distributed integer,  $C_k^{q_k}$ , in the interval  $[(q_k - 1)m + 1, q_k m]$ , where  $q_k$  is agent  $k$ 's collision counter and  $m$  denotes the number of branches into which agents involved in the latest collision are divided. As soon as the medium becomes idle, agent  $k$  transmits a pulse for a duration of  $C_k^{q_k}$  slots and then starts sensing the medium. If it concludes that the following slot is busy, it defers its transmission and repeats this step, without changing its collision counter. Otherwise, if the medium is idle after its pulse transmission, the agent starts its transmission. Hence, the agent with the largest pulse transmission period among class I agents will be the winner of the contention.

2. If a collision happens during the collision resolution period, agents involved in the collision increase their collision counters and go to the previous step.
3. Each time an agent transmits successfully, it resets its collision counter to 0 and follows the procedure for the class II agents.

### Fairness Analysis of DSCFQ

Now let us introduce some definitions, which facilitate our fairness analysis of the proposed scheduling algorithm.

**Definition 1.** We define *Collision Resolution Period* as the whole period of the splitting algorithm, during which all class I agents receive service, and will refer to it as CRP in the rest of the paper. Furthermore, we define the set of agents that were involved in the latest collision as *Collision Set*.

**Definition 2.** A generalized time slot is defined as the interval between two consecutive times at which class II agents are allowed to transmit a packet.

In other words, the generalized time slot equals an idle slot ( $\sigma$ ) if the medium has been idle at least for one time slot. Otherwise, the generalized time slot will be equal to the time interval during which the medium is entirely busy until it gets idle for one time slot. These cases are shown in Fig. 3.

**Definition 3.** We define the network's virtual time,  $v(t)$ , as the number of idle slots during interval  $(0, t)$ , normalized by the scaling factor. Furthermore, for any interval  $(t_1, t_2)$ ,  $v(t_1, t_2)$  is defined as  $v(t_2) - v(t_1)$ .

It is obvious from Definition 3 that the system's virtual time is a cumulative function and as a result, it is a nondecreasing function of time. Whenever an agent transmits at the beginning of a slot, all the other agents will be able to detect that transmission. As a result, all the agents have similar observations about the status of a given slot and therefore, all the agents are capable of tracking the network's virtual time.

In order to compare the normalized services received by different agents, we define virtual time and service deviation for each agent, as follows:

**Definition 4.** The virtual time of agent  $k$ ,  $v_k(t)$ , is defined

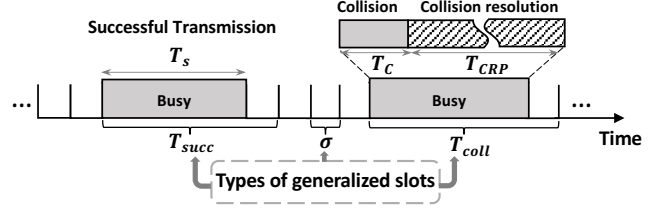


Figure 3: Illustration of the possible states of a generalized time slot.

as

$$v_k(0) = 0, \quad \forall k \in \mathcal{N},$$

$$v_k(t_1, t_2) = \begin{cases} w_k(t_1, t_2), & \forall k \in \mathcal{B}(t_1, t_2), \\ v(t_1, t_2), & \forall k \in \mathcal{A}(t_1, t_2). \end{cases}$$

**Definition 5.** We define normalized service deviation of agent  $k$ , which represents the difference between the network's and the agent's virtual times, as follows

$$\delta_k(t) = v(t) - v_k(t), \quad \forall k \in \mathcal{N}.$$

**Theorem 1.** The maximum normalized service disparity between any pair of agents  $k$  and  $j$ ,  $k, j \in \mathcal{B}(t_1, t_2)$ , in a shared medium with our proposed scheduling algorithm, is bounded as

$$|w_k(t_1, t_2) - w_j(t_1, t_2)| \leq \frac{L_k^{max}}{\phi_k} + \frac{L_j^{max}}{\phi_j} + \frac{2}{\alpha}, \quad (4)$$

where  $L_k^{max}$  and  $L_j^{max}$  are the maximum packet sizes that can be transmitted by agents  $k$  and  $j$ , respectively.

We prove this theorem through the following set of lemmas.

**Lemma 1.** Whenever the transmission of an agent's message finishes, the normalized service deviation of the corresponding agent satisfies

$$-\frac{1}{\alpha} < \delta_k(d_k^i) \leq 0 \quad \forall i, \forall k \in \mathcal{N}. \quad (5)$$

*Proof.* Let us define  $b_k^i$  as  $\max\{a_k^i, d_k^{i-1}\}$ . If packet  $p_k^i$  arrives before that packet  $p_k^{i-1}$  has finished its service,  $b_k^i = d_k^{i-1}$  and hence

$$\delta_k(b_k^i) = \delta_k(d_k^{i-1}), \quad a_k^i < d_k^{i-1}. \quad (6)$$

Otherwise,  $b_k^i = a_k^i > d_k^{i-1}$  and therefore we have

$$\delta_k(b_k^i) = \delta_k(a_k^i) = \delta_k(d_k^{i-1}) + \delta_k(d_k^{i-1}, a_k^i). \quad (7)$$

Since agent  $k$  is not backlogged during the interval  $(d_k^{i-1}, a_k^i)$ , we get  $\delta_k(d_k^{i-1}, a_k^i) = v(d_k^{i-1}, a_k^i) - v_k(d_k^{i-1}, a_k^i) = 0$ . We conclude from Eqs. (6) and (7) that

$$\delta_k(b_k^i) = \delta_k(d_k^{i-1}). \quad (8)$$

On the other hand, according to the definition of  $b_k^i$ , all the previous messages of agent  $k$  have been sent by time  $b_k^i$ . So,  $p_k^i$  is the only message that will be transmitted by agent  $k$

during  $(b_k^i, d_k^i)$ . Furthermore, transmitting  $p_k^i$  requires sensing the medium idle for  $B_k^i$  generalized slots. Consequently, during the interval  $(b_k^i, d_k^i)$ , normalized service of agent  $k$  and the network's virtual time will be increased by  $L_k^i/\phi_k$  and  $B_k^i/\alpha$ , respectively, i.e.,

$$v_k(b_k^i, d_k^i) = w_k(b_k^i, d_k^i) = \frac{L_k^i}{\phi_k}, \quad (9)$$

$$v(b_k^i, d_k^i) = \frac{B_k^i}{\alpha} = \frac{[\alpha(\frac{L_k^i}{\phi_k} - \epsilon_k^i)]}{\alpha}. \quad (10)$$

Using Eqs. (9) and (10), we get

$$\delta_k(b_k^i, d_k^i) = \frac{[\alpha(\frac{L_k^i}{\phi_k} - \epsilon_k^i)]}{\alpha} - \frac{L_k^i}{\phi_k}. \quad (11)$$

Let us return to the definition of  $\epsilon_k^i$ . Comparing Eqs. (3) and (11) we get  $\epsilon_k^{i+1} - \epsilon_k^i = \delta_k(b_k^i, d_k^i)$ . Moreover, we have:

$$\begin{aligned} \epsilon_k^i &= \epsilon_k^1 + \sum_{n=1}^{i-1} (\epsilon_k^{n+1} - \epsilon_k^n) = \sum_{n=1}^{i-1} \delta_k(b_k^n, d_k^n) \\ &= \delta_k(b_k^1, d_k^1) + \sum_{n=2}^{i-1} \delta_k(b_k^n, d_k^n). \end{aligned}$$

Using Eq. (8) and considering the fact that  $\delta_k(b_k^1) = 0$ , since  $b_k^1$  is the first time that agent  $k$  gets backlogged and consequently  $v(b_k^1) = v_k(b_k^1)$ , we get the following equation:

$$\epsilon_k^i = \delta_k(d_k^1) + \sum_{n=2}^{i-1} \delta_k(d_k^{n-1}, d_k^n) = \delta_k(d_k^{i-1}). \quad (12)$$

Therefore, *CompensationFactor* for the  $i$ th message of agent  $k$  is actually equal to the normalized service deviation after the transmission of message  $p_k^{i-1}$ . Using Eq. (8), we have  $\delta_k(d_k^i) = \delta_k(d_k^{i-1}) + \delta_k(b_k^i, d_k^i)$ . Now, defining  $\gamma = \alpha(\frac{L_k^i}{\phi_k} - \delta_k(d_k^{i-1}))$ , and substituting  $\epsilon_k^i = \delta_k(d_k^{i-1})$  into Eq. (11), we get

$$\delta_k(d_k^i) = \delta_k(d_k^{i-1}) + \delta_k(b_k^i, d_k^i) = \frac{[\gamma] - \gamma}{\alpha}. \quad (13)$$

Since  $-1 < [\gamma] - \gamma \leq 0$ , the proof is complete.  $\square$

**Lemma 2.** *The normalized service deviation of agent  $k$ ,  $k \in \mathcal{N}$ , in a shared medium with our proposed scheduling algorithm, gets bounded as*

$$-\frac{1}{\alpha} \leq \delta_k(t) \leq \frac{L_k^{max}}{\phi_k}, \quad \forall k \in \mathcal{N}. \quad (14)$$

*Proof.* We consider two possible cases. In the first case, agent  $k$  is absent at time  $t$ , i.e.  $k \in \mathcal{A}(t)$ . If  $p_k^{i-1}$  represents the last packet that was transmitted by this agent before  $t$ , we can easily conclude that this agent's queue has been empty in the interval  $(d_k^{i-1}, t)$ , which results in  $v_k(d_k^{i-1}, t) = v(d_k^{i-1}, t)$ , and therefore  $\delta_k(d_k^{i-1}, t) = 0$ . Now,  $\delta_k(t)$  can be calculated as follows:

$$\delta_k(t) = \delta(d_k^{i-1}) + \delta(d_k^{i-1}, t) = \delta(d_k^{i-1}).$$

Finally, using Lemma 1 we get

$$-\frac{1}{\alpha} < \delta_k(t) \leq 0, \quad k \in \mathcal{A}(t). \quad (15)$$

Hence, while an agent is absent, its normalized service deviation equals  $\delta(d_k^{i-1})$ , where  $p_k^{i-1}$  is the last packet sent by it before becoming absent.

Now, let us consider the other case in which agent  $k$  is backlogged at  $t$ . Assuming that  $p_k^i$  is the first packet that finishes its service after  $t$ , we get  $b_k^i \leq t \leq d_k^i$ , where  $b_k^i = \max\{a_k^i, d_k^{i-1}\}$ . So, we can calculate  $\delta_k(t)$  as  $\delta_k(b_k^i) + \delta_k(b_k^i, t)$ . Since the transmission of packet  $p_k^i$  begins after sensing  $B_k^i$  idle slots, the network's virtual time will have an increase of  $B_k^i/\alpha$  by the time agent  $k$  starts transmitting  $p_k^i$ , i.e.,  $t = s_k^i$ . Furthermore, the medium will get busy during  $(s_k^i, d_k^i)$  and therefore, the network's virtual time will remain unchanged during this interval. Therefore, for  $t \in [b_k^i, d_k^i]$ ,  $v(b_k^i, t)$  satisfies the following:

$$0 \leq v(b_k^i, t) \leq v(b_k^i, s_k^i) = \frac{B_k^i}{\alpha} = \frac{[\alpha(\frac{L_k^i}{\phi_k} - \delta_k(d_k^{i-1}))]}{\alpha}, \quad (16)$$

where we have used  $\epsilon_k^i = \delta_k(d_k^{i-1})$  from Eq. (12). On the other hand, agent  $k$  is backlogged during  $(b_k^i, d_k^i)$  and therefore,  $v_k(b_k^i, t) = w_k(b_k^i, t)$ . Since  $s_k^i$  represents the time that transmission of packet  $p_k^i$  begins,  $w_k(b_k^i, t)$  will remain zero until  $s_k^i$ , and then reaches  $L_k^i/\phi_k$  at time  $d_k^i$  because of transmitting packet  $p_k^i$  during  $(s_k^i, d_k^i)$ . So we have

$$0 \leq v_k(b_k^i, t) = w_k(b_k^i, t) \leq w_k(b_k^i, d_k^i) = \frac{L_k^i}{\phi_k}. \quad (17)$$

Since  $\delta_k(b_k^i, t) = v(b_k^i, t) - v_k(b_k^i, t)$ , from Eqs. (16) and (17) we deduce that  $\delta_k(b_k^i, t)$  is zero for  $t = b_k^i$ , reaches its maximum at  $t = s_k^i$  and finally decreases to  $\delta_k(b_k^i, d_k^i)$  at  $d_k^i$ . So, this yields

$$\min\{0, \delta_k(b_k^i, d_k^i)\} \leq \delta_k(b_k^i, t) \leq \delta_k(b_k^i, s_k^i).$$

Adding  $\delta_k(d_k^{i-1})$  to each side of the above inequalities, we get

$$\min\{\delta_k(d_k^{i-1}), \delta_k(d_k^i)\} \leq \delta_k(t) \leq \delta_k(d_k^{i-1}) + \delta_k(b_k^i, s_k^i), \quad (18)$$

where we have used Eq. (8). Using Lemma 1, we have

$$-\frac{1}{\alpha} \leq \min\{\delta_k(d_k^{i-1}), \delta_k(d_k^i)\}. \quad (19)$$

Moreover, since  $v_k(b_k^i, s_k^i) = w_k(b_k^i, s_k^i) = 0$ , the upper-bound in Eq. (18) can be simplified as

$$\begin{aligned} \delta_k(d_k^{i-1}) + \delta_k(b_k^i, s_k^i) &= \delta_k(d_k^{i-1}) + v(b_k^i, s_k^i) \\ &= \delta_k(d_k^{i-1}) + \frac{[\alpha(\frac{L_k^i}{\phi_k} - \delta_k(d_k^{i-1}))]}{\alpha} = \frac{[\gamma] - \gamma}{\alpha} + \frac{L_k^i}{\phi_k} \\ &\leq \frac{L_k^{max}}{\phi_k}. \end{aligned}$$

Therefore, using (18) and (19) we have

$$-\frac{1}{\alpha} \leq \delta_k(t) \leq \frac{L_k^{max}}{\phi_k}, \quad k \in \mathcal{B}(t), \quad (20)$$

Finally, the lemma follows from (15) and (20).  $\square$

It can be shown from Lemma 2 that

$$|\delta_k(t_1, t_2)| \leq \frac{L_k^{max}}{\phi_k} + \frac{1}{\alpha}, \quad k \in \mathcal{N}. \quad (21)$$

Using Definitions 4 and 5, and inequality (21), we get the following corollary.

**Corollary 1.** *For any agent  $k \in \mathcal{B}(t_1, t_2)$ , we have*

$$|v(t_1, t_2) - w_k(t_1, t_2)| \leq \frac{L_k^{max}}{\phi_k} + \frac{1}{\alpha}. \quad (22)$$

Now, Theorem 1 can be easily obtained by using Corollary 1 for any pair of agents  $k$  and  $j$ ,  $k, j \in \mathcal{B}(t_1, t_2)$ .

It can be observed from Theorem 1 that the maximum normalized service disparity between a pair of backlogged agents is a function of the maximum packet sizes, fairness weights and the scaling factor. Furthermore, the effect of discretizing backoff times after scaling by  $\alpha$  is reflected in the term  $2/\alpha$  in Eq. (4). Hence, using small scaling factors results in large maximum service deviations and vice versa.

### Throughput Analysis of DSCFQ

In this section, we study the achievable network throughput by our proposed algorithm. We begin by introducing some assumptions and parameters that will be used in the analysis.

We concentrate on saturation throughput, which represents the network throughput in overloaded conditions that all the agents are always backlogged. Let us denote by  $g_k$  the probability that agent  $k \in$  class II attempts to transmit a packet in a given generalized slot. Furthermore, we define total attempt rate  $G$  as the expected number of transmission attempts in a generalized slot, i.e.,  $G = \sum_{n=1}^N g_k$ . Now, we analyze the network throughput for the case in which  $N \gg 1$ , and assume that each agent's average access probability is very small compared to the total transmission attempt rate, i.e.,  $g_k \ll G, \forall k \in \mathcal{N}$ . Since  $G$  represents the average number of transmission attempts in a given generalized slot, and also each collision wastes much more time compared to an idle slot, we expect the optimal total attempt rate, which maximizes the network saturation throughput, to be less than one. So, we conduct the analysis for the case in which  $G < 1$ , and consequently  $g_k \ll 1, \forall k \in \mathcal{N}$ .

As discussed in the previous section, a generalized time slot may be an idle slot, may contain a successful transmission or it may contain collisions. So, a generalized time slot might have a length of  $T_{idle} = \sigma$ ,  $T_{succ}$  or  $T_{coll}$  as shown in Fig. 3. Let us denote by  $T_s$  and  $T_c$  the duration of a successful transmission and the length of the wasted time caused by the collision of class II agents, respectively. Hence,  $T_{coll}$  and  $T_{succ}$  can be calculated based on Fig. 3 as follows:

$$T_{coll} = T_c + T_{CRP} + \sigma, \quad T_{succ} = T_s + \sigma,$$

where  $T_{CRP}$  represents the duration of the contention resolution period (CRP). Now, we calculate the probabilities that a given generalized slot is idle ( $P_{idle}$ ), it contains a successful transmission ( $P_{succ}$ ) or it contains collisions ( $P_{coll}$ ). Since each agent  $k \in$  class II, transmits with probability  $g_k$  in a generalized slot, the total number of transmission attempts in a given slot has a Poisson Binomial distribution.

Furthermore, Le Cam's theorem (Le Cam 1960) suggests that the Poisson Binomial distribution with  $g_k \ll 1, \forall k \in \mathcal{N}$ , can be approximated by a Poisson distribution with mean  $G = \sum_{k=1}^N g_k$ . Therefore, considering the fact that a given slot remains idle (contains a successful transmission) only if zero (one) transmission attempt has been made in that slot, we can calculate  $P_{idle}$ ,  $P_{succ}$  and  $P_{coll}$  as follows

$$\begin{aligned} P_{idle} &\simeq e^{-G}, & P_{succ} &\simeq G e^{-G}, \\ P_{coll} &= 1 - P_{succ} - P_{idle} = 1 - G e^{-G} - e^{-G}. \end{aligned} \quad (23)$$

Now, let us define the normalized saturation throughput, which is represented by  $S$ , as the fraction of time the medium is used to successfully transmit payload bits. We can express  $S$  as follows:

$$S = \frac{(P_s + \bar{n}_c P_c)(\bar{L}/C)}{P_s T_{succ} + P_i \sigma + P_c T_{coll}}, \quad (24)$$

where  $\bar{L}$ ,  $C$  and  $\bar{n}_c$  denote the average message length, the transmission rate, and the average number of agents involved in a given collision of class II agents, respectively. Let  $n_t$  denote the number of agents attempting a transmission in a given generalized slot. Since a collision happens only if the number of transmitted packets in a given time slot is larger than one,  $\bar{n}_c$  can be calculated as follows

$$\bar{n}_c = E[n_t | \text{collision}] = E[n_t | n_t > 1] = G(1 - e^{-G}). \quad (25)$$

### Adaptive DSCFQ

As we discussed earlier, scaling factor is a parameter for backoff adjustment and it has a direct impact on the network throughput. Choosing a small scaling factor results in an increase of the collision probability, and reduces the idle intervals during which the medium is left unused. On the other hand, increasing the scaling factor causes less collisions, but large idle intervals. Therefore, it is important to pick the optimal scaling factor that maximizes the throughput. In the following, we propose a simple adaptive method for updating the scaling factor to reach the peak saturation throughput.

Let us first define the optimal total attempt rate,  $G^*$ , as the value for which the saturation throughput in Eq. (24) is maximized. Substituting  $G^*$  in Eq. (23), we get  $P_{idle}^*$ ,  $P_{succ}^*$  and  $P_{coll}^*$ , respectively. Now, we introduce an adaptive method to reach the maximum saturation throughput. As discussed earlier, each generalized slot has one of the three different states of successful transmission, collision or idle state. Assume that all the agents agree on a common initial scaling factor, which is denoted by  $\alpha_0$ . Then each agent senses the medium and updates  $\alpha$  based on each generalized slot's state as follows:

$$\alpha_{new} = \begin{cases} \alpha, & \text{success} \\ \alpha + \gamma, & \text{collision} \\ \alpha - \beta, & \text{idle} \end{cases} \quad (26)$$

where  $P_{idle}^* \beta = P_{coll}^* \gamma$ .

Now, let  $\alpha^*$  denote the optimal scaling factor that maximizes the saturation throughput. Since  $P_{idle}$  and  $P_{coll}$  are,

Parameter	Value	Definition
$C$	12 Mbps	data transmission rate
$C_{ctrl}$	6 Mbps	control transmission rate
$L_{ACK}$	112 bits	ACK frame length
$L_{RTS}$	160 bits	RTS frame length
$L_{CTS}$	112 bits	CTS frame length
$\bar{L}$	2016 bytes	average message length
$\sigma$	9 $\mu$ s	slot time
$SIFS$	10 $\mu$ s	SIFS time

Table 1: Simulation parameters.

respectively, monotonically decreasing and increasing functions of  $G$ , and  $G$  is inversely proportional to  $\alpha$ , we have

$$D_\alpha = \begin{cases} \gamma P_{coll} - \beta P_{idle} > 0, & \text{if } \alpha < \alpha^* \\ \gamma P_{coll} - \beta P_{idle} < 0, & \text{if } \alpha > \alpha^* \end{cases} \quad (27)$$

where  $D_\alpha$  is the expected drift of the scaling factor from one generalized slot to the next one. Hence, as long as  $\alpha < \alpha^*$ , there will be a positive drift which increases  $\alpha$  on average. On the other hand,  $\alpha > \alpha^*$  results in a negative drift and therefore, leads to the reduction of the scaling factor.

## Performance Evaluation

In this section, we present our evaluations of the proposed distributed scheduling algorithm. We first describe the experimental setup and the technical assumptions used in the experiments. We then explain the baselines and the metrics that have been used for the evaluations. Finally, we discuss the experiments and results.

**Experimental Setup:** Consider a network of 10 agents that are always backlogged and need to share their information with each other. Furthermore, the importance of the agents' information are different. Therefore, each agent chooses a fairness weight that reflects the importance of its messages. We consider a scenario in which three agents choose  $\phi = 10$ , three other agents choose  $\phi = 8$ , two agents pick  $\phi = 2$  and finally two agents pick  $\phi = 1$  as their fairness weights. Moreover, to study the performance of our algorithm in more realistic conditions, we implement our algorithm based on 802.11 DCF. Specifically, in our implementation we consider various practical requirements such as the control messages (RTS/CTS/ACK), inter frame spacing intervals, the imperfections caused by the packet headers and the propagation delay, etc. The parameters used in our implementation are summarized in Table 1.

**Baselines:** To compare our algorithm with the existing fair scheduling algorithms, we classify these baselines into two general types. The first type, which we call Type I, includes algorithms that emulate weighted fair queueing by taking backoff intervals (or IFS values) in proportion to the packet size of the head of the queue messages, i.e.

$B_k^i = \lfloor \alpha (\frac{L_k^i}{\phi_k}) \rfloor$ , while still using random mechanisms such as BEB for collision resolution. The DFS algorithm introduced by (Vaidya et al. 2005) belongs to this category. The other type (Type II) uses similar methods for providing fairness, with the difference of giving a higher priority to the

agents with unsuccessful last transmission attempts. The proposed algorithm by (Dugar, Vaidya, and Bahl 2001) is an example of type II algorithms. An advantage of the above classification is that the value and impact of each part of our algorithm on the performance of the system can be assessed separately.

**Evaluation Metrics:** For evaluating the fairness of our algorithm, we use the *fairness index* (Jain et al. 1996), which is defined as follows:

$$\text{fairness index} = \frac{(\sum_k T_k / \phi_k)^2}{N \cdot \sum_k (T_k / \phi_k)^2}, \quad (28)$$

where  $T_k$  denotes the throughput of flow  $k$ . Moreover, we compare the degree of fairness over different time-scales using Sliding Window Method (SWM), introduced by (Koksal, Kassab, and Balakrishnan 2000). In this method, we slide a window of size  $w$  across a packet trace of medium accesses. By calculating the fairness index for the transmissions that lie in this window after each sliding, we get a sequence of fairness indices. We report the average fairness index as the fairness metric for a given window size. The other metrics used in our evaluations are the throughput of the agents and the network throughput. The throughput of a particular agent in a given interval is defined as the ratio of the successfully transferred information to the length of the interval. The network throughput can be calculated by adding the throughput of the agents.

## Results and Discussions

Let us start our evaluations with the fairness analysis of the system. Fig. 4 shows the performance of our algorithm along with type I and II scheduling schemes for different time scales ( $w=30, 50, 100$  and  $1000$ ). As can be observed, our algorithm achieves the best performance among all the compared algorithms in both short-term ( $w=30, 50$ ) and long-term ( $w=100, 1000$ ) comparisons. While our algorithm achieves a relatively constant fairness index over a wide range of scaling factors ( $[0.0001, 0.02]$ ), the performance of type I and II algorithms deteriorate dramatically for the small scaling factors. This is particularly important since large scaling factors can result in more idle time slots and therefore, less efficient utilization of the medium. On the other hand, we can observe that the difference between the fairness indices of the compared algorithms starts to diminish as we increase the scaling factor. This can be explained by considering the main sources of unfairness in Type I and II algorithms. Specifically, the agents with unsuccessful last transmissions have no priority over the other agents in type I algorithms, which creates unfairness in the case of collisions. Since increasing the scaling factor results in less collisions, the effect of this factor becomes less dominant. On the other hand, discretization of the backoff intervals, i.e., using  $\lfloor \alpha (L_k^i / \phi_k) \rfloor$  instead of  $\alpha (L_k^i / \phi_k)$ , is another source of unfairness in both type I and II algorithms. In a similar way, the impact of this factor is more noticeable for small scaling factors, where a large range of packet lengths can be mapped into the same backoff intervals after scaling and discretization. The compensation factor  $\epsilon_k^i$  used in

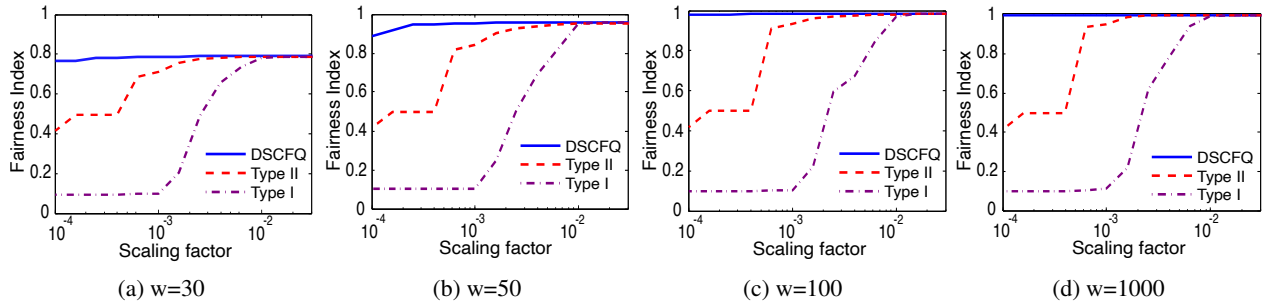


Figure 4: Average fairness index as a function of the scaling factor, for window sizes  $w=30, 50, 100, 1000$ .

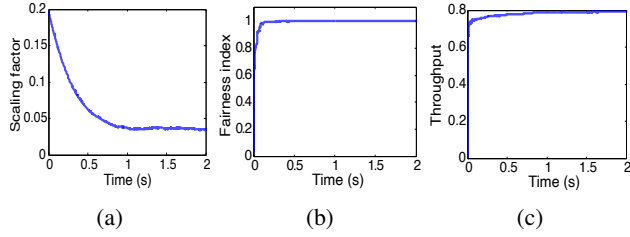


Figure 5: Convergence of the (a) scaling factor (b) fairness index and (c) normalized throughput to their optimal values.

the backoff calculation step in our algorithm (Eq. (2)) deals with this issue effectively. Since our algorithm eliminates the effect of the two sources of unfairness, it achieves the best performance among all the compared baselines in different time scales (Fig. 4). Therefore, the difference between the fairness indices of type I and II algorithms shows the amount of improvement in average fairness index, using an algorithm that gives priority to the agents with unsuccessful last transmission attempts. On the other hand, the difference between our algorithm and type II class demonstrates the increase in the fairness index by using compensation factor in the backoff calculation as in Eq. (2). Another observation we make is that the fairness index generally decreases for all the algorithms as we consider shorter time scales. This phenomenon can be also observed in the centralized scheduling algorithms, which is justified by the packet-based nature of the problem (in contrast to fluid flow models).

Now, let us study the adaptive behaviour of our algorithm. As we discussed before, there is an intrinsic trade-off between the fairness and the network throughput, which can be controlled by the scaling factor. In previous section, we proposed an adaptive method for updating the scaling factor to achieve the maximum throughput attainable by our algorithm. Fig. 5a shows dynamic adaptation of the scaling factor for the case in which it is initially set to 0.2. We observe that as the scaling factor converges into its optimum value, the fairness index approaches one (Fig. 5b) and the network throughput reaches its maximum value (Fig. 5c). This can be validated from Fig. 6a, which shows the normalized network throughput obtained from the simulation and the theory (Eq. (24)). As can be seen, the analytically driven throughput from Eq. (24) closely approximates the

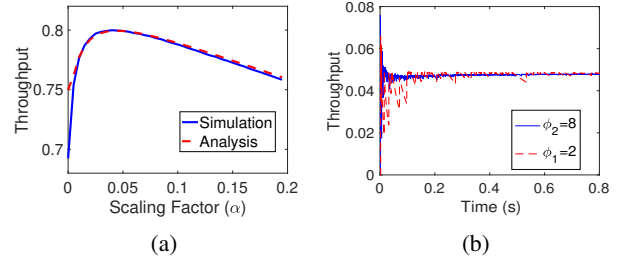


Figure 6: a) Comparison of the normalized network throughput obtained from the simulation and Eq. (24), as a function of the scaling factor. b) Convergence of the normalized throughput of agents 1 and 2 (with  $\phi = 2$  and  $\phi = 8$ ) into the same shares.

throughput obtained from the simulations. Moreover, we can observe that the maximum throughput ( $\simeq 0.8$ ) is achieved for a scaling factor around 0.04, which are the same values that our algorithm converges into in Figs. 5a and 5c. Finally, Fig. 6b shows the convergence of the normalized throughput of two agents with fairness weights of  $\phi_1 = 2$  and  $\phi_2 = 8$ . As can be observed, our algorithm equalizes the normalized throughput of these agents after its convergence.

## Conclusion

In this paper, we considered one of the important, but understudied challenges in the communication of multi-agent systems. Specifically, we studied the shared medium contention problem in the communication of the agents and proposed a new distributed scheduling algorithm for providing fairness in these systems. This becomes even more important when the agents' information or observations have different importance, in which case the agents require different priorities for accessing the medium and sharing their information. We showed that our proposed algorithm can provide a deterministic bound on the maximum service disparity among any pair of agents. This can particularly improve the short-term fairness, which is important in real-time applications. Moreover, we designed an adaptive mechanism that enables our scheduling algorithm to adjust itself and achieve a high throughput at the same time.



## References

- Bertsekas, D.; and Gallager, R. 1992. *Data Networks (2Nd Ed.)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.
- Choi, J.; Yoo, J.; and Kim, C.-K. 2008. A Distributed Fair Scheduling Scheme With a New Analysis Model in IEEE 802.11 Wireless LANs.  *Vehicular Technology, IEEE Transactions on 57(5)*: 3083–3093.
- Dugar, A.; Vaidya, N.; and Bahl, P. 2001. Priority and fair scheduling in a wireless LAN. In *MILCOM. Communications for Network-Centric Operations: Creating the Information Force. IEEE*, volume 2, 993–997 vol.2.
- Foerster, J.; Assael, I. A.; De Freitas, N.; and Whiteson, S. 2016. Learning to communicate with deep multi-agent reinforcement learning. In *Advances in neural information processing systems*, 2137–2145.
- Goldman, C. V.; and Zilberstein, S. 2003. Optimizing information exchange in cooperative multi-agent systems. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, 137–144.
- Golestani, S. 1994. A self-clocked fair queueing scheme for broadband applications. In *INFOCOM '94. Networking for Global Communications., 13th Proceedings IEEE*, 636–646 vol.2.
- Goyal, P.; Vin, H.; and Cheng, H. 1997. Start-time fair queueing: a scheduling algorithm for integrated services packet switching networks. *Networking, IEEE/ACM Transactions on 5(5)*: 690–704.
- Jain, R.; Babic, G.; Nagendra, B.; and Lam, C. 1996. Fairness, call establishment latency and other performance metrics. In *ATM Forum/96-1173, ATM Forum Document, August*.
- Jiang, J.; and Lu, Z. 2018. Learning attentional communication for multi-agent cooperation. In *Advances in neural information processing systems*, 7254–7264.
- Kim, D.; Moon, S.; Hostallero, D.; Kang, W. J.; Lee, T.; Son, K.; and Yi, Y. 2018. Learning to Schedule Communication in Multi-agent Reinforcement Learning. In *International Conference on Learning Representations*.
- Koksal, C. E.; Kassab, H.; and Balakrishnan, H. 2000. An Analysis of Short-Term Fairness in Wireless Media Access Protocols. In *ACM Sigmetrics 2000*. Santa Clara, CA.
- Le Cam, L. 1960. An approximation theorem for the Poisson binomial distribution. *Pacific J. Math.* 10(4): 1181–1197.
- Lee, J.-F.; Liao, W.; and Chen, M.-C. 2007. Interframe-Space (IFS)-Based Distributed Fair Queueing for Proportional Fairness in IEEE 802.11 WLANs.  *Vehicular Technology, IEEE Transactions on 56(3)*: 1366–1373.
- Mao, H.; Zhang, Z.; Xiao, Z.; Gong, Z.; and Ni, Y. 2020. Learning Agent Communication under Limited Bandwidth by Message Pruning. In *AAAI 2020*, 5142–5149.
- Parekh, A.; and Gallager, R. 1993. A generalized processor sharing approach to flow control in integrated services networks: the single-node case. *Networking, IEEE/ACM Transactions on 1(3)*: 344–357.
- Peng, Z.; Zhang, L.; and Luo, T. 2018. Learning to communicate via supervised attentional message processing. In *Proceedings of the 31st International Conference on Computer Animation and Social Agents*, 11–16.
- Qiao, D.; and Shin, K. 2002. Achieving efficient channel utilization and weighted fairness for data communications in IEEE 802.11 WLAN under the DCF. In *Quality of Service, 2002. Tenth IEEE International Workshop on*, 227–236.
- Shreedhar, M.; and Varghese, G. 1996. Efficient fair queuing using deficit round-robin. *Networking, IEEE/ACM Transactions on 4(3)*: 375–385.
- Singh, A.; Jain, T.; and Sukhbaatar, S. 2019. Individualized Controlled Continuous Communication Model for Multi-agent Cooperative and Competitive Tasks. In *International Conference on Learning Representations*.
- Vaidya, N.; Dugar, A.; Gupta, S.; and Bahl, P. 2005. Distributed fair scheduling in a wireless LAN. *Mobile Computing, IEEE Transactions on 4(6)*: 616–629.
- Zhang, C.; and Lesser, V. 2013. Coordinating multi-agent reinforcement learning with limited communication. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, 1101–1108.
- Zhang, S. Q.; Zhang, Q.; and Lin, J. 2019. Efficient communication in multi-agent reinforcement learning via variance based control. In *Advances in Neural Information Processing Systems*, 3235–3244.