Received April 27, 2019, accepted May 26, 2019, date of publication May 31, 2019, date of current version July 2, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2920175

FS-MAC: An Adaptive MAC Protocol With Fault-Tolerant Synchronous Switching for FANETs

MIN ZHANG, CHAO DONG, (Member, IEEE), AND YANG HUANG, (Member, IEEE)
Key Laboratory of Dynamic Cognitive System of Electromagnetic Spectrum Space, Ministry of Industry and Information Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Corresponding author: Chao Dong (dch@nuaa.edu.cn)

This work was supported by National Natural Science Foundation of China under Grant 61472445, Grant 61631020, and Grant 61827801.

ABSTRACT Medium access control (MAC) is significant for guaranteeing the quality of service of Flying Ad-hoc NETworks (FANETs). The adaptive MAC protocol is recognized as a promising solution, which is able to improve the flexibility and robustness of FANETs. In this paper, we propose a fault-tolerant synchronous-MAC (FS-MAC) protocol that can switch between CSMA/CA and TDMA protocols for the FANETs. In FS-MAC, we propose a distributed Q-learning-based MAC switching scheme which contains a MAC pre-selection operation and a practical byzantine fault tolerance (PBFT)-based consensus decision procedure to produce a MAC switching decision. By the MAC pre-selection operation, each UAV can evaluate its own performance accurately and determine which MAC protocol is more appropriate. Then, all UAVs in FANETs can implement fault-tolerant synchronous switching with the help of the PBFT-based consensus decision procedure. The simulations are conducted to evaluate the various performance of the FS-MAC. It is shown that FS-MAC can significantly outperform the baseline protocols in terms of the average throughput, delay, and packet retransmission ratio performance.

INDEX TERMS Adaptive MAC protocol, consensus algorithm, Q-learning, FANETs.

I. INTRODUCTION

Recently, Unmanned Aerial Vehicle (UAV) technology has made rapid progress and been applied in many fields including aerial photograph [1], 5G communication [2], agriculture and forest fire monitoring [3], search and rescue [4], and so forth. The communication and perception ability of a single UAV provides the basis of cooperative operation of multiple UAVs, which makes it possible for small UAVs to work cooperatively to replace the role of large aircraft [5], [6]. Due to the ability of supporting efficient, real-time and cooperative communications among multiple UAVs, Flying Ad-hoc NETworks (FANETs) have attracted attention from industry and academia. Similarly to Mobile Ad-hoc NETworks (MANETs) and Vehicular Ad-hoc NETworks (VANETs) [5], [7], the Medium Access Control (MAC) is critical for FANETs because it determines when the UAVs can transmit data packets on shared wireless channel.

MAC protocols used in FANETs can be categorized into two types: contention-based protocols, *e.g.*, Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) [8]–[10] and contention-free protocols, *e.g.*, Time Division Multiple Access (TDMA) [11]–[13]. However, different types of MAC protocols prefer different scenarios [13], [14]. For example, when many UAVs have data to be transmitted, TDMA is more appropriate, while CSMA/CA works better in scenarios of low contention [14]. For FANETs, there exist a large number of different application scenarios and different data traffic patterns [15], [16]. Therefore, an adaptive MAC protocol which can switch among different MAC protocols to satisfy various application requirements is appealing [13].

Similarly to other ad hoc networks, unified MAC protocol is important for guaranteeing the performance of FANETs. This requires the adaptive MAC protocol [16] to fault-tolerantly switches the MAC protocols of UAVs in a synchronous manner. However, the design is challenging. Firstly, the wireless environment of FANETs is sophisticated and dynamic, such that it is difficult for an UAV to determine which MAC protocol is more appropriate at this moment [13], [17], [18]. Moreover, FANETs are distributed networks such that it is non-trivial to implement a synchronous MAC protocol switching.

The associate editor coordinating the review of this manuscript and approving it for publication was Min Jia.

In recent years, there are a lot of works to study the adaptive MAC protocol. Z-MAC [14] can implement synchronous MAC protocol switching, but the switch sequence is fixed. AMAC [19] can continuously calculate the current throughput and compare with the previous stage, so the node can send a request packet to the neighbor node to apply for the MAC protocol switching. However, the synchronous switching cannot be obtained because there is no global consensus in AMAC. In [13], Wang *et al*. proposed an adaptive MAC protocol switching framework called CT-MAC which allows multiple MAC protocols to switch mutually based on several key information in FANETs. However, they only implement CT-MAC according to the Global Positioning System (GPS) information which is a centralized scheme essentially. In conclusion, although a lot have been done towards the adaptive MAC protocol, fault-tolerant synchronous MAC protocol switching in FANETs still needs to be studied further.

In this paper, we propose an FS-MAC protocol for FANETs which is essentially an adaptive MAC protocol with fault-tolerant synchronous switching. FS-MAC utilizes the distributed Q-Learning based MAC switching scheme which contains the MAC pre-selection operation and Practical Byzantine Fault Tolerance (PBFT) based consensus decision to produce MAC switching decision. Firstly, for a single UAV, FS-MAC utilizes a MAC pre-selection operation to evaluate the current state, so as to determine which MAC protocol is appropriate. Then, by the PBFT-based consensus decision procedure, the synchronous switching consensus among multiple UAVs can be reached with the faulty UAVs in consideration. The main contributions of this paper are listed as follows.

- We propose an FS-MAC protocol, which is essentially an adaptive MAC protocol with fault-tolerant synchronous switching for FANETs. In FS-MAC, we utilize the distributed Q-Learning based MAC switching scheme to produce MAC switching decision.
- We propose a MAC pre-selection operation for each UAV to determine which MAC protocol, *i.e.*, CSMA/CA or TDMA is more appropriate for the single UAV.
- We propose a PBFT-based consensus decision procedure to help the synchronous switching consensus among multiple UAVs to be reached with the faulty UAVs in consideration. The consensus selected MAC protocol can make the overall networks always using the appropriate MAC protocol.
- We evaluate the various performance of FS-MAC through simulations. The results show that FS-MAC can improve the average throughput, delay, and packet retransmission of DAMP [19] at most 35.43%, 29.71%, and 67.84%, respectively.

The rest of this paper is organized as follows: Section II introduces the related work about the adaptive MAC protocol for FANETs. In Section III, the proposed FS-MAC is introduced in detail, including system architecture, accurate pre-selection and fault-tolerant consensus decision. Section IV presents the simulation and performance evaluation. Section V concludes the paper.

II. RELATED WORKS

Over the past years, there has been increasing interest on the adaptive MAC protocols for MANETs. Meanwhile, with the development of UAVs, recently, several adaptive MAC protocols were proposed to improve the performance of FANETs.

Rhee *et al.* [14] put forward Z-MAC which can be viewed as the first hybrid MAC protocol that can switch between CSMA/CA and TDMA. In the network setup phase, Z-MAC selects appropriate MAC protocol based on the network topology. During the running phase, the protocol switching only happens when there is a significant change of network topology, *e.g.*, physical relocation of most nodes occurs. Hence, it is hard for Z-MAC to operate effectively in FANETs whose network topology change frequently.

Shrestha *et al.* [20] presented MCCA which is a hybrid TDMA-CSMA/CA protocol based on Markov Decision Process (MDP) in centralized wireless networks. Therefore, considering the distributed characters of FANETs, MCCA is not applicable for multi-hop FANETs.

RL-MAC [21] is an adaptive MAC protocol based on reinforcement learning, which is used to adjust the length of contention slot using Q-Learning algorithm so that the energy consumption can be effectively reduced. However, because CSMA/CA is a contention-based MAC protocol, only adjusting the length of contention slot is not enough when the traffic load is high.

Huang *et al.* [19] proposed AMAC in which the UAV can switch its MAC scheme between TDMA and CSMA/CA based on the acquired throughput. However, only relying on throughput to specify the switching criterion cannot adapt to the dynamics of FANETs.

Zheng *et al.* [22] proposed PPMAC which is a novel Position-Prediction based directional MAC protocol for FANETs. The access control scheme of PPMAC is similar to CSMA/CA which is contention-based. Hence, under high contention scenario in FANETs, the performance will degrade.

In [15], Wang *et al.* proposed an adaptive MAC framework for UAV ad hoc networks. Under this framework the UAVs can choose and switch to the most appropriate MAC protocol based on some kind of information. To verify this framework, the authors designed an adaptive MAC protocol called CT-MAC which allows UAVs to switch between CSMA/CA and TDMA based on their own GPS location information. However, the switching criterion of CT-MAC is only based on the GPS location information without considering dynamic network and channel state. Besides, because that GPS information is in fact a global one, CT-MAC cannot implement synchronous MAC switching in distributed FANETs where there is no global control information.

In conclusion, current adaptive MAC protocols for MANETs and FANETs cannot meet the dynamic and distributed characters of FANETs well. How to select

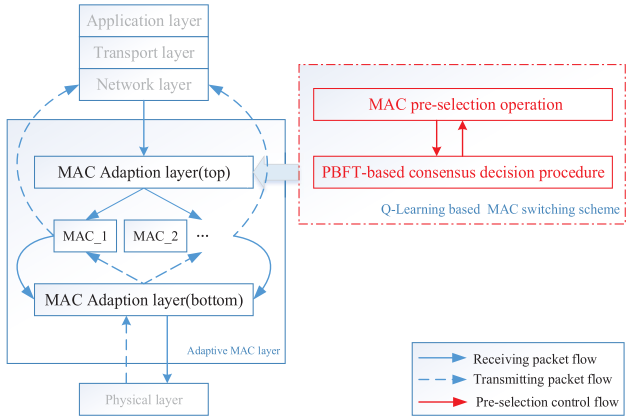


FIGURE 1. The framework of FS-MAC.

appropriate MAC protocol and implement synchronous switching to the UAVs needs to be studied further.

III. FS-MAC FOR FANETs

In this section, we propose the FS-MAC. We first overview the overall architecture of the FS-MAC. Then two main functions of FS-MAC, including the pre-selection operation and the PBFT-based consensus decision procedure, are proposed.

A. OVERVIEW OF FS-MAC

As shown in Fig. 1, following the adaptive MAC framework proposed in [15], FS-MAC utilizes a distributed Q-Learning based MAC switching scheme which contains a MAC pre-selection operation and a PBFT-based consensus decision procedure to produce MAC switching decision, which is the inputt of the bottom MAC adaption layer. Specially, when the top MAC adaption layer receives a packet from the network layer, the top MAC adaption layer selects an appropriate MAC protocol according to the output of distributed Q-Learning based MAC switching scheme. Then the top MAC adaption layer passes the packets to the selected MAC protocol for encapsulation. In the distributed Q-Learning based MAC switching scheme, the MAC pre-selection operation is employed to select an appropriate MAC protocol for a UAV according to the UAV's current state, which consists of Successful Transferred Amount of Data (STAD), delay and Packet Retransmission Ratio (PRR). Meanwhile, the PBFT-based consensus decision procedure, which would be discussed in Section III-C, is used for fault-tolerant and synchronous MAC protocol switching among the UAVs in FANETs.

Fig. 2 depicts the procedure of MAC protocol selection scheme. First, at the beginning of a slot, each UAV runs Q-Learning algorithm to determine which MAC protocol between TDMA and CSMA/CA is more appropriate. Once the UAV finds that the MAC pre-selection operation yields a decision of switching the MAC protocol, to avoid performance degradation caused by protocol asynchronization among all UAVs (*i.e.*, operating different MAC protocol

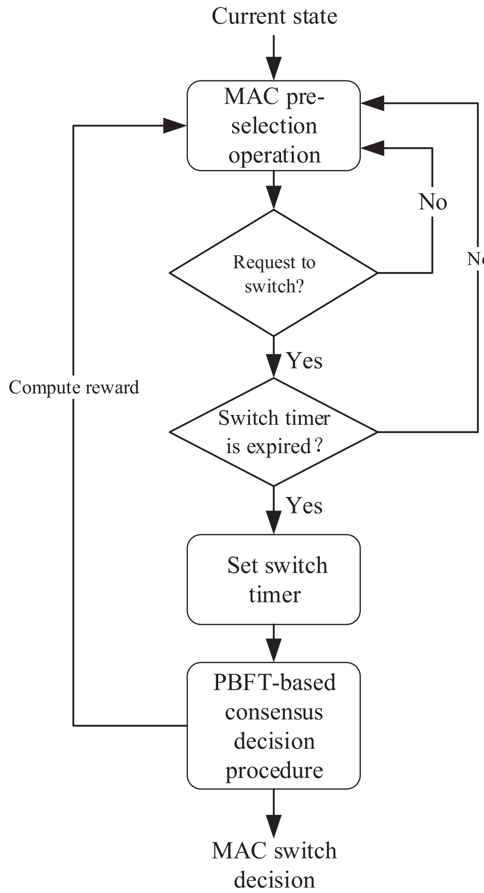


FIGURE 2. The procedure of the distributed Q-Learning based switching in one time slot.

simultaneously), this UAV needs to negotiate with other UAVs to reach a consensus. Then, this UAV will check its switching timer which is used to avoid frequent negotiation request of MAC protocol switching. If the switching timer is expired, this UAV will reset the switching timer and ask the other UAVs to run the PBFT-based consensus decision procedure. Meanwhile, the MAC switching scheme can obtain a reward value after the PBFT-based consensus decision procedure is finished.

We propose the distributed Q-Learning based switching scheme, which consists of the two main functions of the FS-MAC, *i.e.*, the MAC pre-selection operation and the PBFT-based consensus decision procedure. Afterwards, we elaborate on the PBFT-based consensus decision procedure. Table 1 lists the major notations used in this paper.

B. DISTRIBUTED Q-LEARNING BASED SWITCHING

In this section, we propose the distributed Q-Learning based switching scheme, which includes the MAC pre-selection operation.

Suppose there are N UAVs in FS-MAC, in order to preselect a MAC protocol at each UAV, *i.e.*, UAV$_i$ where $i \in \{1, 2, \ldots, N\}$, we design the MAC pre-selection operation,

TABLE 1. A list of major notations.

Notation	Definition
S	State space
$s_i(t)$	State of UAV$_i$ at slot t
A	Action space
$a_i(t)$	Action of UAV$_i$ at slot t
R	Reward function
$r_i(t)$	Reward value of UAV$_i$ at slot t
π	Policy
π^*	Optimal policy
$\bar{b}$	The average throughput of network
$b_i(t)$	The STAD of UAV$_i$ at slot t
$B_i(t)$	The STAD level of UAV$_i$ at slot t
$\bar{d}$	The average packet delay of network
$d_i(t)$	The packet delay of UAV$_i$ at slot t
$D_i(t)$	The packet delay level of UAV$_i$ at slot t
$\bar{l}$	The average PRR of network
$l_i(t)$	PRR of UAV$_i$ at slot t
$L_i(t)$	PRR level of UAV$_i$ at slot t
$m_i(t)$	Current MAC protocol of UAV$_i$ at slot t
α	The update coefficient of Q-function
τ	The size of randomness
γ	Discount factor
ψ, β, θ and ω	The coefficients in reward function
V	Value of state
N	The number of all the UAVs
f	The number of faulty UAVs
n_p	The number of successful transmitted date packets
l	The length of the transmitted packet
η	Level number of STAD
μ	Level number of delay
φ	Level number of PRR
w	Consensus message
$U(w)$	Digest of consensus message
u	Digest
v	View number

which can pre-select an appropriate MAC protocol according to its current state. Then through the PBFT-based consensus decision procedure, pre-selections at various UAVs can reach a consensus and UAV$_i$ can obtain a reward to update its Q-table.

The MAC switching design problem can be modeled as a Markov Decision Problem (MDP), which can be characterized with a four-tuple $\langle S, A, P, R\rangle$ [23]–[25], where S represents state space, A contains all the possible actions at each state, P represents a probability transition function and R is the reward function $S \times A \rightarrow R$. In addition, the objective is to maximize the expectation of total discounted rewards of each UAV [26], which can be written as

$$\max \left\{ E_\pi \left\{ \sum_{k=0}^{\infty} \left[\gamma^k r_k(s_k, \pi) \right] \right\} \right\}, \tag{1}$$

where the policy π means the mapping from state space S to action space A, *i.e.*, $a_i \in A_i(s_i)$. The parameter γ, which denotes the discount factor and ranges from 0 to 1, reflects that the future rewards are worth more or less than immediate rewards to update policy. Moreover, the optimal policy π_i^* of UAV$_i$ satisfies Bellman equation [23]. That is, the objective function can be written as

$$V(s_i) = \max_{a_i \in A_i(s_i)} \left[r(s_i, a_i) \right] + \gamma \sum_{s_i} \left[p_i(s_i' \,|s_i, a_i) V(s_i') \right], \tag{2}$$

where s_i' is the next state after s_i for UAV$_i$. However, $p_i(s_i' \,|s_i, a_i)$, which represents the transition probability from s_i to s_i', is an unknown statistic. Therefore, the well known Q-Learning algorithm is exploited to solve problem of Equation (2). The Q-value updating function can be written as [24]

$$\begin{aligned} Q_{t+1}(s, a) = & \, Q_t(s, a) \\ & + \alpha (r_t + \gamma \max_{a' \in \mathbf{A}(s')} Q(s', a') - Q_t(s, a)), \end{aligned} \tag{3}$$

remarkably, the max $Q(s', a')$ in the above equation is nothing else than the MAC pre-selection operation. In (3), α is the update rate of Q-table, ranging from 0 to 1. If α vanishes, it means that the Q-values remain the same over time slots, thereby nothing is learned. A high value α such as 0.9 can make UAV$_i$ learn efficiently, while it makes the learning, *i.e.*, the update in Equation (3), focus too much on future rewards and ignore previous knowledge.

The detailed definition of S, A, and R are as follows.

1) STATE SPACE

S is the state space of UAV. For UAV$_i$, at slot t, state $s_i(t)$ is defined as a three tuple $\langle B_i(t), D_i(t), L_i(t)\rangle$, where $B_i(t)$, $D_i(t)$ and $L_i(t)$ are the STAD, delay and PRR levels for UAV$_i$ at the beginning of slot t, respectively. In order to guarantee the computational feasibility, we divide these performance indexes into different levels as follows.

$$\begin{gathered} \left[0, \frac{B^{\mathrm{m}}{}_i}{\eta}\right), \left[\frac{B^{\mathrm{m}}{}_i}{\eta}, \frac{2B^{\mathrm{m}}{}_i}{\eta}\right), \ldots, \left[\frac{(\eta-1)B^{\mathrm{m}}{}_i}{\eta}, B^{\mathrm{m}}{}_i\right] \\ \left[0, \frac{D^{\mathrm{m}}{}_i}{\mu}\right), \left[\frac{D^{\mathrm{m}}{}_i}{\mu}, \frac{2D^{\mathrm{m}}{}_i}{\mu}\right), \ldots, \left[\frac{(\mu-1)D^{\mathrm{m}}{}_i}{\mu}, D^{\mathrm{m}}{}_i\right] \\ \left[0, \frac{L^{\mathrm{m}}{}_i}{\varphi}\right), \left[\frac{L^{\mathrm{m}}{}_i}{\varphi}, \frac{2L^{\mathrm{m}}{}_i}{\varphi}\right), \ldots, \left[\frac{(\mu-1)L^{\mathrm{m}}{}_i}{\varphi}, L^{\mathrm{m}}{}_i\right], \end{gathered} \tag{4}$$

where η, μ and φ are the number of levels for STAD, delay, and PRR, respectively. $B^{\mathrm{m}}{}_i$, $D^{\mathrm{m}}{}_i$ and $L^{\mathrm{m}}{}_i$ are the max values of STAD, delay and PRR, respectively. Hence, the number of state in S is $\eta\mu\varphi$.

TABLE 2. Action and its corresponding reward of FS-MAC.

Action	Consensus selection	Reward type
CSMA/CA	CSMA/CA	Type 1
	TDMA	Type 2
TDMA	CSMA/CA	Type 2
	TDMA	Type 1

2) ACTION SPACE

A is the action space and $a_i(t)$ represents actions for UAV$_i$ in slot t. As shown in Table 2, in FS-MAC, the action $a_i(t) \in A$ of UAV$_i$ at each state is defined as $A = \langle$CSMA/CA, TDMA$\rangle$. However, an action of CSMA/CA means that this single UAV pre-selects CSMA/CA MAC protocol but does not use it immediately, while all the UAVs must use the protocol determined by the PBFT-based consensus decision procedure.

3) REWARD

R is the reward function which considers the difference between the of STAD, delay and PRR performance of UAV$_i$ in the state before pre-selection and those after consensus selection. That is, after UAV$_i$ takes an action and requests a switching consensus, it will not obtain a reward value immediately until the consensus operation is finished. The proposed consensus decision will be discussed in the Section III-C. As shown in Table 2, the pre-selection can be different from the consensus selection, hence the rewards are divided into two types, Type 1 and Type 2, and FS-MAC reduces the reward of Type 2 through multiplying the type 2 reward by ψ. Therefore, for UAV$_i$ at slot t, the reward function can be given by

$$R_i(t) = \begin{cases} R'_i(t), & \text{Type 1} \\ \psi R'_i(t), & \text{Type 2,} \end{cases} \tag{5}$$

where

$$R'_i(t) = \beta \Delta p_i(t) - \theta \Delta d_i(t) - \omega \Delta l_i(t) m_i(t). \tag{6}$$

In the above Equation (6),

$$\begin{cases} \Delta b_i(t) &= b_i(t) - b_i(t-1) \\ \Delta d_i(t) &= d_i(t) - d_i(t-1) \\ \Delta l_i(t) &= l_i(t) - l_i(t-1), \end{cases} \tag{7}$$

where $b_i(t)$, $d_i(t)$ and $l_i(t)$ represent the STAD, delay, PRR of UAV$_i$ during the slot t, respectively. $m_i(t)$ is the current MAC protocol of UAV$_i$ at the beginning of slot t. ψ, θ, β and ω represent the coefficients of rewards and each one ranges from 0 to 1. In addition, β, θ and ω satisfy $\beta + \theta + \omega = 1$. In contrast to $\beta \Delta p_i(t)$, $\theta \Delta d_i(t)$ and $\omega \Delta l_i(t)$ are multiplied by -1 because that lower $\theta \Delta d_i(t)$ and $\omega \Delta l_i(t)$ values can yield a higher reward. Moreover, the throughput and delay are the common key indicators of current state for UAV$_i$, but the PRR performance in TDMA is always better than CSMA/CA. Therefore, we use ω and $m_i(t)$ to adjust $\Delta l_i(t)$.

Now we introduce how to select the action for a single UAV, *i.e.*, UAV$_i$. In order to obtain the tradeoff between exploration and exploitation, pre-selection operation uses Gibs (or Boltzmann) distribution to nonlinearize $Q[s(t), a(t)]$ of UAV$_i$ [26], and then uses Roulette Wheel Selection [26] approach to select action. It chooses action a in the tth time slot with:

$$p_i(a) = \frac{e^{Q_i[s(t),a(t)]}/\tau}{\sum_{b=1}^{n} e^{Q_i[s(t),b(t)]}/\tau}, \tag{8}$$

where τ represents the size of randomness and the larger the τ, the greater the randomness. When τ is close to 0, action selection is almost equivalent to greedy selection [26] [27]. Considering the special running environment of FANETs and handle the tradeoff issue between exploration and exploitation of the MAC pre-selection operation, we adopt this Boltzmann approach to pre-select actions under a certain state.

Algorithm 1 Distributed Q-Learning Based Switching

Require: Variables: $b_i(t)$, STAD value of UAV$_i$ at slot t; $d_i(t)$, delay value of UAV$_i$ at slot t; $l_i(t)$, PRR value of UAV$_i$ at slot t, $m_i(t)$, current MAC protocol of UAV$_i$ at slot t; γ discounted factor; α, learning rate, N_s, number of states, N_a, number of actions;
Ensure: *Pre_Selection*
for each $i \in \{1, 2, \ldots, N\}$ **do**
  initialize $s_i(t)$ and $a_i(t)$ to 0, and $Q_i(s_i(t), a_i(t))$ to 0-matrix with N_s rows and N_a columns;
end for
for each $t \in \{1, 2, \ldots, TotalTime\}$ **do**
  for each $i \in \{1, 2, \ldots, N\}$ **do**
    Select actions using Equation (8);
    if $a_i(t) \neq m_i(t)$ **then**
      Request to switch;
      $Pre\_Selection_i(t) \leftarrow [i, t, a_i(t)]$;
      Algorithm 2;
    end if
    Compute rewards using Equation (5);
    Update $Q_i(s_i(t), a_i(t))$ using Equation (3);
    $(s_i(t-1) \leftarrow s_i(t)$;
  end for
end for

The distributed Q-Learning based switching scheme is summarized in Algorithm 1. The inputs of the algorithm are the state of the UAV$_i$ including STAD, packet delay, PRR, current MAC type, *etc*. After the initialization of variables, each UAV repeats as follows: firstly, select action using Equation (8); secondly, if the action of UAV$_i$ is different from the current MAC protocol, *i.e.*, $a_i(t) \neq m_i(t)$, this UAV will request consensus, start Algorithm 2 and output $Pre\_Selection_i(t)$; thirdly, compute the reward using Equation (5); fourthly, update $Q_i(s, a)$ using Equation (3); at last, set $s_i(t)$ as the last iteration value $s_i(t-1)$. *Pre_Selection* is both the output of Algorithm 1 and the input of Algorithm 2, which is designed as a 3-tuple including UAV identifier, time slots and its corresponding action at the slot t. From the formulation and design, the appropriate pre-selection of each single UAV can be obtained.

C. PBFT-BASED CONSENSUS DECISION PROCEDURE

In this section, we propose the Practical Byzantine Fault Tolerance (PBFT) [28] based consensus decision procedure, which is actually a fast and fault-tolerant voting procedure.

There may exist some faulty UAVs in FANETs, some of them are controlled by adversary to interfere the operation of FANETs and the others are faulty due to their own factors, *e.g.*, breakdown. For the former, we can use cryptographic techniques to prevent spoofing and detect corrupt packets from these faulty UAVs [28], [29]. For the latter, we can

Algorithm 2 PBFT-Based Consensus Decision

Require: Variables: *Pre_Selection*; f, number faulty UAVs; N, total number of UAVs;
Ensure: *Consensus_Result*
initialize v, n, w, PU, $timer$, $Consensus\_Reslut$, n and $U(w)$ to 0;
Select PU using Equation (9);
$PU \leftarrow \langle \text{REQUEST}, switch, timestamp, SU \rangle_{\sigma_{SU}}$;
$w_{i_{PU}} \leftarrow a_{i_{PU}}$;
$w_{i_{BU}} \leftarrow a_{i_{BU}}$;
PU broadcast $\left\langle \langle \text{PRE-PREPARE}, v, n, u \rangle_{\sigma_{i_{PU}}}, w \right\rangle$;
Start timer;
if $timer > wait\_time$ **then**
  $Consensus\_Result \leftarrow 0$;
else
  while $timer \leqslant wait\_time$ **do**
    BU broadcast$\left\langle \langle \text{PREPARE}, v, n, u, i \rangle_{\sigma_{i_{BU}}}, w \right\rangle$;
    if $\langle \text{PREPARE} \rangle \sigma_{i_{BU}} \geqslant 2f$ **then**
      RU broadcast $\langle \text{COMMIT}, v, n, U(w), i \rangle \sigma_{i_{RU}}$;
    end if
    if $w_{i_{BU}} = w$ **then**
      $committed(w, v, n, i) \leftarrow 1$;
    end if
    if $\langle \text{COMMITTED} \rangle \sigma_{i_{non\_faulty}} \geqslant 2f + 1$ **then**
      $\text{RU}_{committed}$ broadcasts $\text{REPLY}(w, v, n, i)$;
    end if
    if $\text{REPLY}(w, v, n, i) \geqslant 1$ **then**
      $Consensus\_Resut \leftarrow 1$;
    else
      $Consensus\_Resut \leftarrow 0$;
    end if
  end while
end if

detect these faulty UAVs easily because that they cannot send or receive packets. Therefore, in this paper, we assume that the faulty UAVs cannot take part in the procedure of consensus decision. Meanwhile, we define f as the maximum number of UAVs that can be simultaneously faulty, and $f = (N-2)/3$ following [28].

In the PBFT-based consensus decision procedure, following [28], the replica information move through a succession of consensus decision called views. In one view, the UAV that requests a voting is defined Start UAV (SU), and the other UAVs are Replica UAVs (RUs). Among RUs, one is the Primary UAV (PU), and the others are Backup UAVs (BUs). Besides, We define the number of PU in one view is

$$k = v \bmod (N-1), \tag{9}$$

where v is the view number. The PU can receive the request packet from SU, and delivery the packet to the BUs through broadcast.

Noting that only when the result of pre-selection is different from the current used MAC protocol and the switching timer is expired, an UAV will become SU and start the voting procedure. In FS-MAC, we assume that the successful voting means to switch synchronously and the voting failure means to continue to use current MAC protocol. In order to avoid the interruption from the faulty PU, the BUs can detect whether PU is faulty through timeout mechanism [28]. If the PU is failed, the BU that first detect the PU is faulty will starts a view change to enter new consensus decision procedure, thus the PU will be re-selected according to Equation (9).

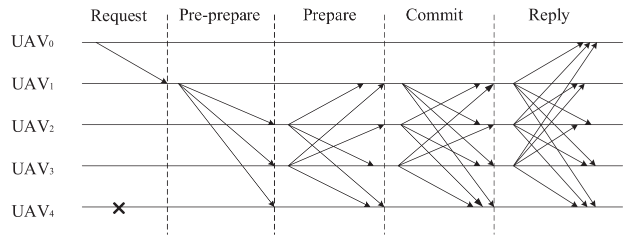


FIGURE 3. The operation of the PBFT-based consensus decision procedure.

Fig. 3 shows the operation of the PBFT-based consensus decision procedure. In the figure, UAV_0 is the SU, and $\text{UAV}_1 \sim \text{UAV}_4$ are the RUs. Among them, UAV_1 is the PU and $\text{UAV}_2 \sim \text{UAV}_4$ are the BUs. Meanwhile, we assume UAV_4 is faulty and the voting procedure is as follows.

- Request: UAV_0 sends a request packet to UAV_1 to initiate a vote procedure for MAC protocol switching. Note that the request packet includes the pre-ordered switching time to ensure the synchronous switching if the vote is successful finally.
- Pre-prepare: UAV_1 broadcasts the voting request packet to $\text{UAV}_2 \sim \text{UAV}_4$ after receiving it from UAV_0.
- Prepare: $\text{UAV}_2 \sim \text{UAV}_4$ handle the voting request and broadcast the voting prepare packets to the other RUs [28], *e.g.*, $\text{UAV}_2 \rightarrow \text{UAV}_1$, UAV_3 and UAV_4. However, in the figure, UAV_4 does not broadcast because it is faulty.
- Commit: if one of the RUs, *i.e.*, $\text{UAV}_1 \sim \text{UAV}_4$, receives $2f$ voting prepare packets from the BUs that agree to operate the switching request. This UAV will enter the commit phase as a prepared RU, and start to broadcast the voting commit packets including its own voting result.
- Reply: if one of RUs obtains $2f + 1$ voting commit packets (possibly including its own) to agree to switch from the other RUs, then this RU will broadcast voting reply packet to all the other UAVs.

Finally, if one of the UAVs, including SU, receives any of voting reply packets, then this UAV considers that the consensus has been reached. Meanwhile, this UAV will switch the MAC protocol at the pre-ordered switching time. Since the PBFT-based consensus decision procedure does not consider broadcast failure [13], [28], when one UAV receives voting reply packet, all the non-faulty UAVs can receive the voting reply packets. Therefore the synchronous switching of

TABLE 3. Simulation parameters.

Parameter	Value
Number of UAVs	5 to 12
Number of state	18
Number of action	2
Slot time	20 (μs)
Physical date rate	2 (Mbit/s)
Bit error rate	10^{-6}
ACK length for CSMA/CA	56 (μs)
SIFS length for CSMA/CA	10 (μs)
DIFS length for CSMA/CA	50 (μs)
Average arrival time of packets	5 slots

FANETs can be reached. Besides, after the synchronously switching, MAC switching scheme can obtain a reward value according to Equation (5).

The PBFT-based consensus decision algorithm is summarized in Algorithm 2, in which its inputs include *Pre_Selection*, f and N, and the output is a Boolean variable *Consensus_Result* whose true value represents voting success and conversely false value represents voting failure. Besides, $U(w)$ is the digest of packet w [30]. Meanwhile, *timestamp* is used for the RUs to ensure exactly-once voting for the procedure and that no more SU appears before the consensus timer expires.

We have proved the convergence analysis in the Appendix. Briefly, the application environment of distributed Q-Learning based switching scheme is a closed-cycle control system, in which the state and action space are both finite. Moreover, we designed a consensus timer mechanism to avoid the endless consensus request. As for the convergence of the proposed algorithm, we have provided the convergence proof in the Appendix. According to the proof, we can know that the proposed algorithm is convergent. Thus, the stability can be guaranteed indeed.

IV. PERFORMANCE EVALUATION

In this section, we first introduce the simulation setting. Then, we compare the performance of FS-MAC with other three protocols.

A. SIMULATION SETTING

We setup a FANETs which consists of $5 \sim 12$ UAVs for performance evaluation. We consider that the UAVs are evenly located in spherical space, and all the UAVs can communicate with others. Following IEEE 802.1609, the channel is divided into Control CHannel (CCH) and Service CHannel (SCH). Consensus procedure uses CCH, and SCH is used for packet transmission. We assume that the clock among UAVs is synchronous. Besides, following [13], [15], [31]–[33], the arrival of the packet is designed as obeying Poisson distribution and its average packet arrival time λ is five slots.

The following three metrics are utilized to evaluate the performance of FS-MAC protocol, which are as follows.

(a) *Average throughput*: in this paper, the average throughput in the total time is defined as

$$\bar{b} \triangleq \frac{\sum_{n}^{N_p} b_n}{T} \tag{10}$$

where l_n represents the nth packet length, N_p is the number of total packets and T presents the total running time.

(b) *Average packet delay*: the average packet delay is defined as

$$\overline{d} \triangleq \frac{\sum_{n}^{N_p} d_n}{n_p} \tag{11}$$

where d_n is the packet delay of nth data packet, including MAC queueing delay, backoff delay, transmission delay and propagation delay. Note that only the packets that have successful transmitted are analyzed statistically for the average delay, and the packets that failed to transmit are not analyzed.

(c) *Average packet retransmission ratio*: the average Packet Retransmission Ratio (PRR) is defined as

$$\overline{l} \triangleq \frac{\sum_{n}^{N_p} l_n}{n_p} \tag{12}$$

where l_n denotes the packet retransmission of nth data packet, including the retransmission caused by the over backoff time and packed loss.

Different MAC protocols are investigated and run on the same simulation environment with our proposed protocols, and we compare FS-MAC with other three protocol schemes as follows.

- Distributed Adaptive MAC Protocol (DAMP) [19], which allows the UAVs to select MAC protocol based on some performance, *e.g.*, delay. It can be seen that, because there is no synchronous switching scheme, multiple MAC protocols maybe operate simultaneously.
- FS-MAC without Pre-selection (FSWP), in which the PBFT-based consensus decision procedure is still in use but the MAC pre-selection operation is replaced by another mechanism as follows, if the delay value of transmitting UAVs is below than a threshold value [19], then this UAV will request to switch to another MAC protocol.
- FS-MAC without Consensus selection (FSWC), the MAC pre-selection operation is still in use but the PBFT-based consensus decision procedure is replaced by another mechanism as follows. Firstly, each UAV will transmit the packets using the protocol determined by themselves. Secondly, if the protocol using by the transmitting UAVs is different from the receiving UAVs, transmitting UAVs will try to switch to another protocol repeatedly until the packet transmitted successfully [19]. That is to say, because there is no protocols consensus selection, multiple MAC protocols maybe operate simultaneously.

B. PERFORMANCE COMPARISONS

The performance comparisons of the four protocol schemes are as follows.

1) AVERAGE THROUGHPUT

First, we set the time to 200s and increase the number of UAVs from 5 to 12. Fig. 4 shows that the average throughput

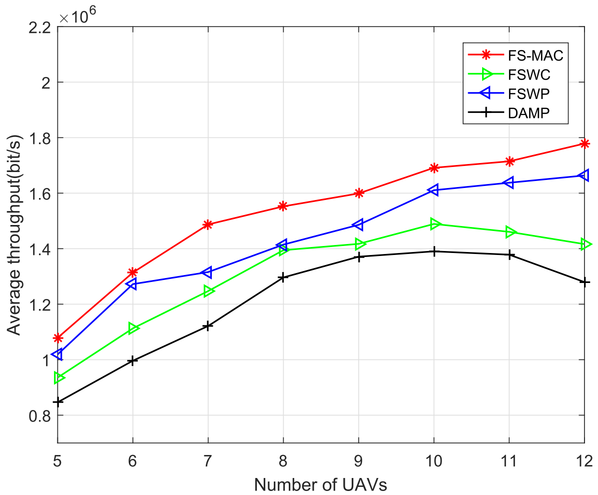


FIGURE 4. **Average throughput comparison under different number of UAVs.**

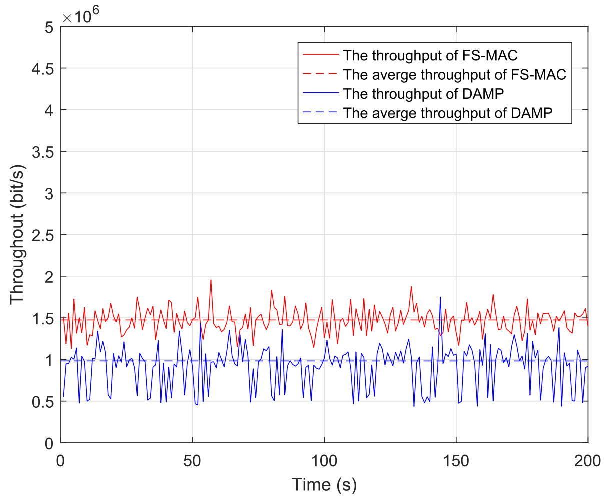


FIGURE 5. **Throughput comparison under different time slots (taking 6 UAVs as an example).**

performance of FS-MAC and FSWP is improved as the increase of the data communication amount and the TDMA percentage. The average packet arrival time is five slots, with the increasing number of UAVs, especially more than 10, FS-MAC outperforms the other three protocols because it can switch to appropriate MAC protocol. when the number of UAVs $N > 10$ in this simulation, the average throughput performance of FSWC and DAMP is reduced by the increasing number of UAVs, because the packets delay caused by the asynchronous switching is increased when $N > 10$. Moreover, the average throughput performance of FSWP and FSWC is inferior to FS-MAC due to the lack of pre-selection and consensus selection, respectively.

In Fig. 4, the average throughput of FSWC performs better than that of DAMP, since UAVs can evaluate their states accurately and take the appropriate MAC protocol for their pre-selection. Similarly, the performance of FSWP is better than that of DAMP, since the synchronous switching of FSWP can reduce the trial times and packets delay. Besides, the average throughput performance of FSWP is better than that of FSWC, due to the operation of multiple MAC protocols at the same time in FSWC. In other words, when the MAC protocol used by the transmitting UAV is different from the receiving UAV, the trial operations [19] of the transmitting UAV will occupy a lot of time slots and may fail to transmit packets. Therefore, the average throughput performance of FSWP is better than that of FSWC. In general, FS-MAC can improve the average throughput of DAMP at most 35.43%, since the DAMP protocol fails to accurately evaluate the current state and switch by consensus.

Taking 6 UAVs as an example, the throughput performance of FS-MAC and DAMP for varying time slots is shown in Fig. 5. The throughput performance is generally stable as the time slots increases and the proposed FS-MAC has better performance than that of DAMP due to the synchronously switching among the protocols. The reason is that, when the throughput value reduces, the MAC pre-selection operation can request a switching consensus in time, and the PBFT-based consensus decision procedure can give a consensus decision rapidly. In other words, in FS-MAC, the MAC pre-selection operation can make a single UAVs always use the appropriate MAC protocol to transmit packets. In addition, the MAC consensus selection is synchronous in the FANETs and the communication failure caused by different MAC protocols is non-existing.

2) AVERAGE DELAY

We set the time to 200s and increase the number of UAVs from 5 to 12. Fig. 6 displays the average delay performance comparison of varying numbers of UAVs. It can be observed that the average delay value of the four protocols separately increases with the growing number of UAVs, because the data communication amount and the TDMA percentage increases. Similar to the average throughput performance, the proposed scheme achieves better performance than others, since distributed Q-Learning based MAC switching scheme can select an appropriate MAC protocol for FANETs.

As shown in Fig. 6, FSWC and FSWP outperform DAMP in average delay, due to the implementation of pre-selection and the consensus selection for appropriate MAC protocol, respectively. Besides, the average delay performance of FSWP is better than that of FSWC. In one hand, FSWP takes delay threshold [19] as the switching criterion. That is, a transmitting UAV in FSWP will request to switch when the packets delay is under a certain threshold, hence the average delay performance of FSWP can always keep excellent. Similar to throughput performance in Fig. 4, the delay performance of FS-MAC when N>10 is also better than the others. In the other, the synchronous switching of FSWP can reduce trial times and packets delay. In general, FS-MAC can reduce at most 29.71% of the average packet delay of DAMP, since the DAMP protocol fails to accurately evaluate the current state and switch by consensus.

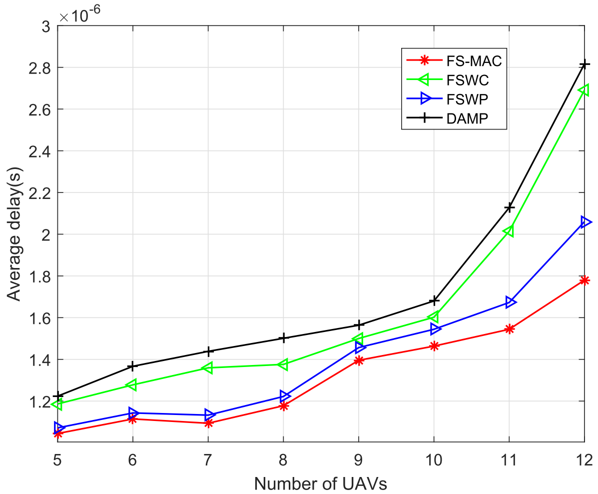


FIGURE 6. Average delay comparison under different number of UAVs.

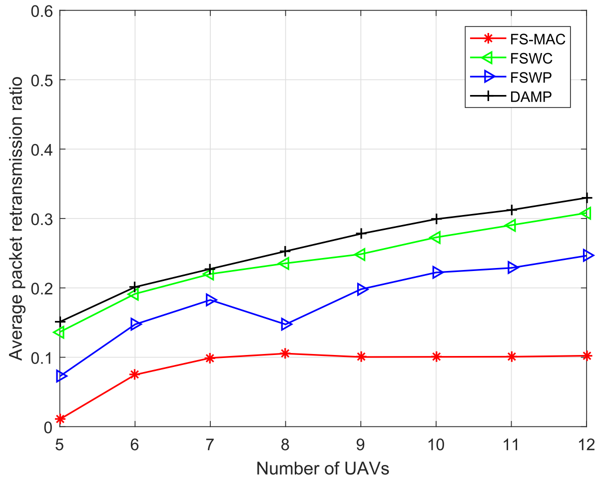


FIGURE 8. Average PRR comparison under different number of UAVs.

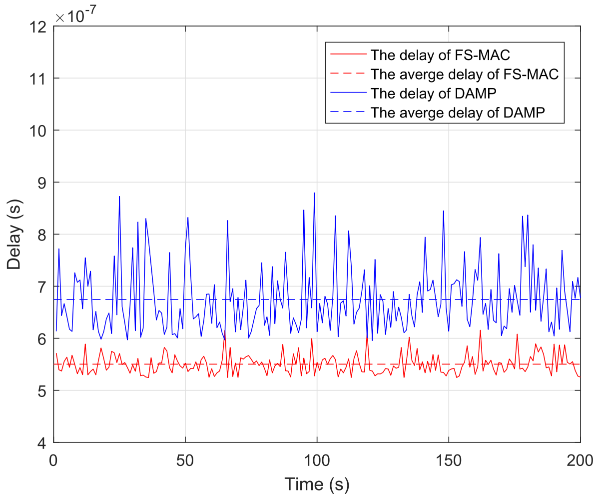


FIGURE 7. Delay comparison under different time slots (taking 6 UAVs as an example).

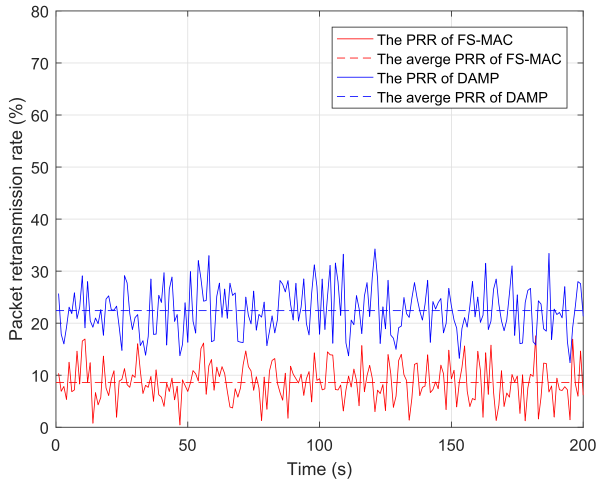


FIGURE 9. PRR comparison under different time slots (taking 6 UAVs as an example).

Taking 6 UAVs as an example, Fig. 7 illustrates the delay performance of FS-MAC and DAMP in varying time slots. The delay performance is generally stable as the time slots increase and the proposed scheme has better delay performance than DAMP due to the adaptive switching among protocols. From the above discussion, we know that DAMP does not have the ability to synchronously switch between the MAC protocols and the trial operation will cause a lot of delay, therefore the delay performance of DAMP at different time slots is more unsteady than FS-MAC.

3) AVERAGE PACKET RETRANSMISSION RATIO

We set the time to $200s$ and increase the number of UAVs from 5 to 12. Fig. 8 shows the Packet Retransmission Ratio (PRR) comparison curves of the four protocols for varying numbers of UAVs. Since FS-MAC has made a balanced overall selection in both CSMA/CA and TDMA modes and the average PRR of TDMA mode is particularly low, the overall average PRR of FS-MAC have been significantly reduced. Since FS-MAC can always select the appropriate MAC protocol and synchronously switch to it rapidly, the average PRR performance of both FSWP and FSWC are better than that of DAMP. Besides, the average PRR performance of FSWP also outperforms FSWC, since the trial operation in FSWC brings much retransmission of packets. Like Fig. 4, the PRR performance of FS-MAC when N>10 is also better than the others. In general, FS-MAC outperforms DAMP in average PRR at most 67.84%, since the DAMP protocol fails to accurately evaluate the current state and switch according to the consensus.

Taking 6 UAVs as an example, Fig. 9 shows the PRR performance of FS-MAC and DAMP in varying time slots. The PRR performance is generally stable as the time slots increase and the proposed scheme has better PRR steady performance than DAMP, since the PRR value is considered as a key factor in Equation (5) and can significantly influence the pre-selection of FS-MAC in the pre-selection operation.

In summary, from Fig. 4 to Fig. 9, FS-MAC shows an increasing gain over other three protocols as number of UAVs increases, and accurate evaluation and appropriate selection for MAC protocol can be achieved. Due to the adaptive switching function, the stable performance also outperform DAMP. Besides, it confirms the observations made in the 6 figures mentioned above and validates the Q-Learning based switching scheme and the PBFT-based consensus decision procedure can synchronously switch MAC protocols and gain the performance of FANETs.

V. CONCLUSION

Considering the variety of missions and application scenarios, the adaptive MAC protocol for FANETs is desired. In this paper, we propose FS-MAC which can synchronously switch appropriate MAC protocol between TDMA and CSMA/CA with the faulty UAVs. In FS-MAC, we propose the distributed Q-Learning based switching scheme for FANETs to procedure a switching decision. Firstly, a single UAV uses the MAC pre-selection operation to determine an appropriate MAC protocol. Secondly, using proposed the PBFT-based consensus decision procedure, the synchronous switching agreement among multiple UAVs can be reached. Extensive simulations reveal that, the performance of average throughput, delay and packet retransmission ratio are improved compared to the other three MAC protocols. In the future works, we will focus on the event-trigger method [34] applied in the FANETs, which can model real FANETs event-trigger scenario, reduce communication amount and increase the quality of service.

APPENDIX
CONVERGENCE PROOF OF DISTRIBUTED Q-LEARNING BASED SWITCHING SCHEME

This proof is based on the observation that the Q-Learning algorithm can be viewed as a stochastic process to which techniques of stochastic approximation are generally applicable [35]. Thus, we begin with the following theorems.

Theorem 1 [35, Th. II-1]: A random iterative process $\Delta_{n+1}(x) = (1-\alpha_n(x))\Delta_n(x) + \beta_n(x)F_n(x)$ converges to zero w.p.1 under the following assumptions:

1) The state space S is finite.
2)
$$\sum_n a_n(x) = \infty,$$
$$\sum_n {a_n}^2(x) < \infty,$$
$$\sum_n \beta_n(x) = \infty,$$
$$\sum_n {\beta_n}^2(x) < \infty,$$
and $E\{F_n(x)\,|P_n\} \leqslant E\{\alpha_n(x)\,|P_n\}$ *uniformly w.p.*1.
3)
$$\|E\{\,Fn(x)\,|Pn\,\}\,\|\,\mathrm{w} \leqslant \gamma\,\|\Delta_n\|\,\mathrm{w}, \quad where\ \gamma \in (0,1).$$
4)
$$\mathrm{Var}\{F_n(x)\,|P_n\} \leqslant C(1+\|\Delta_n\|\,\mathrm{w})^2, \quad where\ C\ is\ some\ constant.$$

Where $P_n = \{\Delta_n, \Delta_{n-1}, \ldots, F_{n-1}, \ldots, \alpha_{n-1}, \ldots, \beta_{n-1}, \ldots\}$ stands for the past at step n. $F_n(x)$, $\alpha_n(x)$, $\beta_n(x)$ are allowed to depend on the past insofar as the above conditions remain valid. The notation $\|\cdot\|$ w refers to some weighted maximum norm.

In applying the theorem, the Δ_n process will generally represent the difference between a stochastic process of interest and some optimal value (e.g., the optimal value function). The formulation of the theorem therefore requires knowledge to be available about the optimal solution to the learning problem before it can be applied to any algorithm whose convergence is to be verified. In the case of Q-Iearning the required knowledge is available through the theory of dynamic programming (DP) and Bellman's equation [25] in particular.

The convergence of the Q-Learning algorithm now follows easily by relating the algorithm to the converging stochastic process defined by Theorem 1.

Theorem 2 [35, Th. II-2]: The Q-Learning algorithm given by

$$Q_{t+1}(s_t, a_t) = (1-\alpha_t(s_t, a_t))Q_t(s_t, a_t) + \alpha_t(s_t, a_t)[c_{st}(a_t) + \gamma V_t(s_t+1)], \tag{13}$$

converges to the optimal $Q^*(s, a)$ vales if

1) S and A are finite.
2) $\sum_t \alpha t(s, a) = \infty$ and $\sum_t {\alpha_t}^2(s, a) < \infty$ *uniformly w.p.*1.
3) $\mathrm{Var}\{c_s(a)\}$ is bounded.
4) If $\gamma = 1$, all policies lead to a cost free terminal state w.p.1.

Proof: For each single UAV, by subtracting $Q^*(s, a)$ from both sides of the learning rule and by defining $\Delta t(s, a) = Q_t(s, a) - Q^*(s, a)$ together with

$$F_t(s, a) = c_s(u) + \gamma V_t(s') - Q^*(s, a).$$

Since the consensus selection operation would not affect the iteration of distributed Q-Learning algorithm in FS-MAC, and the state and action space are both finite, the distributed algorithm implemented on the single UAV can be seen to have the form of the process in Theorem 1 with $\beta_t(s, a) = \alpha_t(s, a)$.

To verify that $F_t(s, a)$ has the required properties we begin by showing that it is a contraction mapping with respect to some maximum norm. This is done by relating F_t to the DP value iteration operator for the same Markov chain [25]. More specifically,

$$\begin{aligned}
&\max_a |E\{F_t(i, a)\}| \\
&= \gamma \max_a \left| \sum_j p_{ij}(a)\left[V_t(j) - V^*(j)\right] \right| \\
&\leqslant \gamma \max_a \sum_j p_{ij}(a) \max_v \left| Q_t(j, v) - Q^*(j, v) \right| \\
&= \gamma \max_a \sum_j p_{ij}(a) V^{\Delta}(j) = T\left(V^{\Delta}\right)(i),
\end{aligned} \tag{14}$$

where we have used the notation

$$V^{\Delta}(j) = \max_{v} \left| Q_t(j, v) - Q^*(j, v) \right|$$

and T is the DP value iteration operator for the case where the costs associated with each state are zero [35]. If $\gamma < 0$ the contraction property of $E\{F_t(i, a)\}$ can be obtained by bounding $\sum_j p_{ij}(a) V^{\Delta}(j)$ by $\max_j V^{\Delta}(j)$ and then including the γ factor. When the future costs are not discounted, i.e., $\gamma = 1$ but the chain is absorbing and all policies lead to the terminal state w.p.1 there still exists a weighted maximum norm with respect to which T is a contraction mapping [35] thereby forcing the contraction of $E\{F_t(i, a)\}$. The variance of $F_t(s, a)$ given the past is within the bounds of Theorem 1 as it depends on $Q_t(s, a)$ at most linearly and the variance of $c_s(u)$ is bounded [23].

REFERENCES

[1] A. S. Laliberte and A. Rango, "Texture and scale in object-based analysis of subdecimeter resolution unmanned aerial vehicle (UAV) imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 761–770, Mar. 2009.

[2] A. Merwaday and I. Guvenc, "UAV assisted heterogeneous networks for public safety communications," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops*, Mar. 2015, pp. 329–334.

[3] E. Honkavaara, H. Saari, J. Kaivosoja, I. Pölönen, T. Hakala, P. Litkey, J. Mäkynen, and L. Pesonen, "Processing and assessment of spectrometric, stereoscopic imagery collected using a lightweight UAV spectral camera for precision agriculture," *Remote Sens.*, vol. 5, no. 10, pp. 5006–5039, 2013.

[4] A. Ryan and J. K. Hedrick, "A mode-switching path planner for UAV-assisted search and rescue," in *Proc. IEEE 44th Conf. Decis. Control Eur. Control Conf.*, Dec. 2005, pp. 1471–1476.

[5] I. Bekmezci, O. K. Sahingoz, and Ş. Temel, "Flying ad-hoc networks (FANETs): A survey," *Ad Hoc Netw.*, vol. 11, no. 3, pp. 1254–1270, 2013.

[6] J. Chen, Q. Wu, Y. Xu, Y. Zhang, and Y. Yang, "Distributed demand-aware channel-slot selection for multi-UAV networks: A game-theoretic learning approach," *IEEE Access*, vol. 6, pp. 14799–14811, 2018.

[7] V. Nguyen, O. T. T. Kim, C. Pham, T. Z. Oo, N. H. Tran, C. S. Hong, and E.-N. Huh, "A survey on adaptive multi-channel MAC protocols in VANETs using Markov models," *IEEE Access*, vol. 6, pp. 16493–16514, 2018.

[8] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[9] A. Woo and D. E. Culler, "A transmission control scheme for media access in sensor networks," in *Proc. 7th Annu. Int. Conf. Mobile Comput. Netw.*, 2001, pp. 221–235.

[10] J. Polastre, J. Hill, and D. Culler, "Versatile low power media access for wireless sensor networks," in *Proc. 2nd ACM Int. Conf. Embedded Netw. Sensor Syst.*, Nov. 2004, pp. 95–107.

[11] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 12, no. 3, pp. 493–506, Jun. 2004.

[12] Z. Zhixian, W. Yajun, Y. Yuan, and W. Yuzhou, "Implementation of UAVs communication network based on dynamic TDMA MAC protocol," in *Proc. 3rd Int. Conf. Adv. Comput. Theory Eng. (ICACTE)*, vol. 6, Aug. 2010, pp. V6-551–V6-554.

[13] W. Wang, C. Dong, H. Wang, and A. Jiang, "Design and implementation of adaptive MAC framework for UAV ad hoc networks," in *Proc. Int. Conf. Mobile Ad-Hoc Sensor Netw.*, Dec. 2016, pp. 195–201.

[14] I. Rhee, A. Warrier, M. Aia, J. Min, and M. L. Sichitiu, "Z-MAC: A hybrid MAC for wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 16, no. 3, pp. 511–524, Jun. 2008.

[15] W. Wang, C. Dong, S. Zhu, and H. Wang, "DFRA: Demodulation-free random access for UAV ad hoc networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[16] W. Hu, X. Li, and H. Yousefi'zadeh, "LA-MAC: A load adaptive MAC protocol for MANETs," in *Proc. IEEE Global Telecommun. Conf.*, Nov./Dec. 2009, pp. 1–6.

[17] Q. Wu, G. Ding, J. Wang, and Y.-D. Yao, "Spatial-temporal opportunity detection for spectrum-heterogeneous cognitive radio networks: Two-dimensional sensing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 516–526, Feb. 2013.

[18] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y.-D. Yao, "Opportunistic spectrum access in unknown dynamic environment: A game-theoretic stochastic learning solution," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1380–1391, Apr. 2012.

[19] K.-C. Huang, X. Jing, and D. Raychaudhuri, "MAC protocol adaptation in cognitive radio networks: An experimental study," in *Proc. 18th Int. Conf. Comput. Commun. Netw.*, Aug. 2009, pp. 1–6.

[20] B. Shrestha, E. Hossain, and K. W. Choi, "Distributed and centralized hybrid CSMA/CA-TDMA schemes for single-hop wireless networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 4050–4065, Jul. 2014.

[21] Z. Liu and I. Elhanany, "RL-MAC: A reinforcement learning based MAC protocol for wireless sensor networks," *Int. J. Sensor Netw.*, vol. 1, nos. 3–4, pp. 117–124, 2006.

[22] Z. Zheng, A. K. Sangaiah, and T. Wang, "Adaptive communication protocols in flying ad hoc network," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 136–142, Jan. 2018.

[23] W. B. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Hoboken, NJ, USA: Wiley, 2009.

[24] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, nos. 3–4, pp. 279–292, 1992.

[25] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ, USA: Wiley, 1994.

[26] R. S. Sutton and A. G. Barto, *Reinforcement Learning—An Introduction*, vol. 1, no. 1. Cambridge, MA, USA: MIT Press, 1998.

[27] Z.-X. Wang, M. Zhang, X. Gao, W. Wang, and X. Li, "A clustering WSN routing protocol based on node energy and multipath," *Cluster Comput.*, to be published. doi: 10.1007/s10586-017-1550-8.

[28] M. Castro and B. Liskov, "Practical Byzantine fault tolerance," in *Proc. OSDI*, vol. 99, 1999, pp. 173–186.

[29] S. Chen, D. W. C. Ho, L. Li, and M. Liu, "Fault-tolerant consensus of multi-agent system with distributed adaptive protocol," *IEEE Trans. Cybern.*, vol. 45, no. 10, pp. 2142–2155, Oct. 2015.

[30] R. Rivest, "The MD5 message-digest algorithm," *Internet RFC*, vol. 473, no. 10, p. 492, 1992.

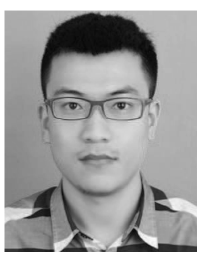
[31] L. Ma, X. Han, and C.-C. Shen, "Dynamic open spectrum sharing MAC protocol for wireless ad hoc networks," in *Proc. IEEE 1st Int. Symp. New Frontiers Dyn. Spectr. Access Netw.*, Nov. 2005, pp. 203–213.

[32] K. Gulati, B. L. Evans, J. G. Andrews, and K. R. Tinsley, "Statistics of co-channel interference in a field of Poisson and Poisson-Poisson clustered interferers," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6207–6222, Dec. 2010.

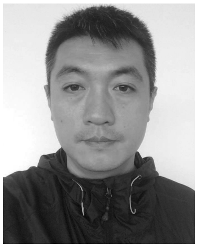
[33] H. Yao and I. Sasase, "Effect of delay-dependent priority discipline on the queueing delay of packet switch with batch poisson arrivals," in *Proc. IEEE Pacific Rim Conf. Commun., Comput., Signal Process.*, May 1995, pp. 177–180.

[34] S. Li, L. Li, Z. Li, X. Chen, T. Fernando, H. H.-C. Iu, G. He, Q. Wang, and X. Liu, "Event-trigger heterogeneous nonlinear filter for wide-area measurement systems in power grid," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2752–2764, May 2019.

[35] T. Jaakkola, M. I. Jordan, and S. P. Singh, "Convergence of stochastic iterative dynamic programming algorithms," in *Proc. Adv. Neural Inf. Process. Syst.*, 1994, pp. 703–710.



MIN ZHANG received the M.S. degree in signal and information processing from the School of Opto-Electronic Information Science and Technology, Yantai University, Yantai, China, in 2017. He is currently pursuing the Ph.D. degree in signal and information processing with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics. His research interests include wireless communication and UAVs swarm networking.



CHAO DONG received the Ph.D. degree in communication engineering from the PLA University of Science and Technology, China, in 2007. From 2008 to 2011, he held a postdoctoral position with the Department of Computer Science and Technology, Nanjing University, China. From 2011 to 2017, he was an Associate Professor with the Institute of Communications Engineering, PLA University of Science and Technology, Nanjing, China. He is currently a Full Professor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His current research interests include D2D communications, UAVs swarm networking, and anti-jamming network protocol. He is a member of ACM and IEICE.

YANG HUANG received the B.S. and M.S. degrees from Northeastern University, China, in 2011 and 2013, respectively, and the Ph.D. degree from Imperial College London, in 2017. He is currently an Associate Professor with the Department of Information and Communication Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His research interests include wireless communications, MIMO systems, convex optimization, machine learning, signal processing for communications, 5G networks, the Internet of Things, wireless power transfer, and wireless powered communications.