IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 11, 2019, accepted June 22, 2019, date of publication July 2, 2019, date of current version July 16, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2926161

Distributed Multichannel Access in High-Frequency Diversity Networks: A Multi-Agent Learning Approach With Correlated Equilibrium

WEN LI, (Student Member, IEEE), YUHUA XU, YUNPENG CHENG, YANG YANG, XUEQIANG CHEN, MENG WANG, (Member, IEEE), AND DIANXIONG LIU, (Student Member, IEEE)

College of Communications Engineering, Army Engineering University of PLA, Nanjing 210000, China

Corresponding author: Yuhua Xu (yuhuaenator@gmail.com)

This work was supported in part by the Natural Science Foundation for Distinguished Young Scholars of Jiangsu Province under Grant BK20160034, and in part by the National Natural Science Foundation of China under Grant 61671473 and Grant 61631020.

ABSTRACT This paper investigates the problem of multi-user multichannel access in distributed high-frequency (HF) diversity communication networks using a game-theoretic learning algorithm which is based on correlation equilibrium (CE). We formulate the channel access problem in the HF networks as a non-cooperative game. In the access game, each user equipment (UE) optimizes its access strategy without the information about other UEs, which makes the channel access problem challenging. It is proved that there is at least one CE point that makes all UEs' access strategy efficient and fair. We propose a distributed learning algorithm based on CE to achieve multi-user access with low cost and fairly in the distributed HF networks. We use coordination signals to help each UE learn the access strategy by themselves. When each UE receives and recognizes the right coordination signals, UEs will learn to transmit data on right channels without further collisions after the learning phase. The simulation results show that the proposed learning algorithm can not only completely avoid interference and get optimal throughput but also guarantee fairness among all UEs.

INDEX TERMS HF communication networks, multichannel access, game theory, correlation equilibrium (CE).

I. INTRODUCTION

High-frequency (HF) communication ($3-30$MHz) which realizes short distance transmission by ground-wave propagation and long distance transmission by sky-wave propagation with ionospheric reflection plays a significant role in military and emergency. Since HF channels suffer from multipath effect, deep fading and the jamming of environment or other users, the HF channels are always poor [1]–[3]. Thus, it is important for users to optimize their access strategy. Existing work about HF channel access mainly considered intelligent Auto Link Establishment (ALE) [4]–[6] and HF spectrum prediction [7]–[9]. However, they all considered the point-to-point communication. Since the ionospheric layer is unstable which is affected by many factors, like sunspot, season and region, the point-to-point communication is really unreliable. The HF communication networks which connect the HF infrastructures through the wired network, e.g., the IP networks, were proposed to enhance communication performance [10]–[12]. At the same time, diversity techniques are widely deployed in HF communications, among which channel diversity is a promising way by transmitting the same signal in several different channels [12]–[14]. However, all user equipments (UEs) have to access multichannel which increases the collision probability and the available HF channels are scant. Thus, channel diversity brings a new problem that how to optimize user's access strategy to improve network communication performance. In this article, the multichannel access problem with multiple UEs is investigated in distributed HF diversity networks.

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

The main challenge of the multichannel access problem is the limited HF channels. The available HF channels are scant for following reasons: 1) The high-frequency band (3-30 MHz) is narrow and shared by the whole world; 2) Lots of HF channels suffer from strong fading, complex multi-path or malicious jamming. The collisions happen when the UEs choose the same channel, which leads to severe through-put drop. However, the technique of channel diversity asks all UEs access several different channels which leads to more collisions. The optimal network state is that all channels are used and there is no collision. Since the available channels are scant, some users obtain high throughput with accessing more channels while others get low throughput in the optimal network state. However, this is unfair. HF communications are usually used in military and emergency and each user's communication needs have to be satisfied. Thus, the fairness also needs to be considered in HF networks. Therefore, we need an effective and fair algorithm to avoid collision among users and take advantage of available channels in HF distributed networks.

In this paper, the problem of multi-user multichannel access in distributed HF networks is discussed using a non-cooperative game. We analyze the problem of multi-user multichannel access using a game model for twofold. First, the users in the HF networks selects and executes the channel access strategy independently and distributively. Secondly, their decisions are interactive and less interference brings higher system throughput. For the non-cooperative game in distributed networks, a learning algorithm based on the Correlated Equilibrium (CE) which achieves effective channel access and guarantees fairness among all UEs is proposed.

As lots of researches have studied the synchronization for time-slot [15], [16], the learning algorithm can work using Time Division Multiple Address (TDMA). Motivated by [17], [18], we assume there is a coordination device to send the coordination signals at each time-slot which will instruct users to select the access strategy. All UEs will learn an efficient access strategy based on each corresponding coordination signal and Acknowledge (ACK) information. Each UE maintains a multichannel access strategy table (MAST) which maps the coordination signals with channel access strategy. For each transmission round, users draw channel access decisions according to the MAST after observing the coordination signals. If the transmission successes in one channel, the mapping in the strategy table remains unchanged, otherwise it will be updated. We conclude that the proposed algorithm gets the CE of the access game and guarantees fairness among all UEs according to theoretical analysis and simulation. The main contributions of this paper are summarized as follow:

- For the multi-user situation, since there is no information exchange in the network, we formulate the problem of multichannel access in distributed HF networks as a non-cooperative game. The formulated game model is proved that there is at least one CE point.
- We propose a learning algorithm based on CE to solve the multichannel access problem. Since there is no central control and information exchange, the proposed algorithm works in distributed HF networks. The strategy updating procedure is only relying on the received coordination signals and the ACK at each time slot. All UEs do not need any history information to update their access strategy.
- We analyze the convergence time and fairness of the proposed algorithm. The simulation results show that all UEs learn to play an efficient access strategy for each coordination signal. The upper bound of convergence time is calculated which is acceptable for actual networks. In the meantime, with the number of coordination signals increasing, the access strategy is increasingly fair.

In the rest of the paper, we present the system model and problem formulation in section III. The related work is reviewed in section II. The game model of the multichannel access problem in distributed HF communication network (HFCN) is analyzed in section IV, and the learning algorithm is proposed. Moreover, the convergence time and jain fairness index (JFI) of the algorithm are analyzed. In section V, the simulation results are presented, and the rationality of the results is investigated. Some discussions are made in VI. Finally, the conclusion is shown in section VII.

II. RELATED WORK

Recently, the reinforcement learning which works well under dynamic environment has been used to solve the problem of channel access [19]–[21]. However, above work mainly considered to enhance communication performance in single user situation. The problem of multi-user channel access was rarely studied in HF network. Applying cognitive radio to HF communications has been proposed in [22], [23]. They only discussed the challenges and opportunities, but the channel access method based on cognitive radio was not proposed. With the development of communication technology and the increased users, it is valuable to study the problem of multi-user multichannel access nowadays. The problem of multi-user multichannel access was extensively investigated in other networks [17], [18], [24]–[39], such as cellular network and IoT.

To cope with the interactions among multiple users, lots of different channel access approaches in other wireless networks using game theory have been proposed [24]–[30]. The concept of local interaction game was first presented in [25], in which each user not only considered its own playoffs but also its neighbors' playoffs. It did achieve best Nash Equilibrium (NE) and maximize the network throughput. Both cooperative and non-cooperative game were considered to solve the channel selection problem in [26], which also found the optimal Nash Equilibrium (NE). In [27], the multichannel multi-access problem in energy harvesting wireless networks was also formulated as a non-cooperative game and an online-learning algorithm was proposed to find the NE point. Reference [28] has explored the channel diversity in HF communications using a match-potential game. Based on game theory, [29] has considered variable-width channel allocation and [30] have investigated dynamic channel allocation for frequency-selective interference channels. All above work proved the existence of NE point and designed various learning algorithm to acquire the optimal pure-strategy NE which maximized the system utility. However, in distributed networks, the methods they have proposed could not ensure fairness among users with limited information of other users [17], [18].

Besides the Nash Equilibrium (NE), another more general solution concept for game theory was known as Correlated Equilibrium (CE) [33]. The CE has several nice advantages: it is easy to find in polynomial time by linear programming and every Nash Equilibrium is a Correlated Equilibrium. The CE is a suitable way to address the problem in distributed environment. The pioneer idea of using CE are found in a number of literatures [17], [18], [34]–[39]. The CE is proposed to cope with the problem that pure-strategy NE was not fair and mixed-strategy NE was not efficient [17], [18]. In [34], [35], they have investigated the simple two-user channel access game and analyzed how to get the efficient CE. The simulation results have shown that the CE improved the payoffs of both users than the mixed NEs. The users need to exchange information including the other users' access action and corresponding utility with each other. The users need to know global information and it is always unavailable in wireless networks. Thus, the multi-user access algorithms based on CE for different networks was proposed, e.g., multichannel allocation in OFDMA-based network [17], [18], energy-aware ad-hoc network [36], opportunistic spectrum access for cognitive radio networks [37], spectrum resource sharing in LTE/3Gpp network [38] and heterogeneous vehicular network [39]. The learning algorithm based on CE instructed the users to execute access action. The regret-matching algorithm used in [36]–[39] is based on the condition that all user can know other users' action. Since the point-to-point transmission is unreliable in HF networks and the transmission rate in HF networks is low, acquiring the information of other users' actions will bring lots of overhead and the information may be incorrect. Thus, it is difficult to directly apply the aforementioned approaches to our problem.

In [17], [18], a novel multi-agent learning algorithm is proposed, but they only solved the problem of single channel selection. Different from previous work, channel diversity is considered in this paper, which allows each user to access several different HF channels. Besides, previous works thought all channel work the same. However, HF channel is time-varying and different channel has different communication performance. Therefore, we propose a learning algorithm based on CE to solve the multichannel access problem in HF networks. Two important information are used in the proposed algorithm: the coordination signals and transmission ACK. Each coordination signal only contains an integer number. Thus, it is easy to be transmitted and has no requirement for transmission rate. Each user only needs to detect the collision in their transmission channels or monitoring channels, and the transmission ACK can be obtained without too much overhead.

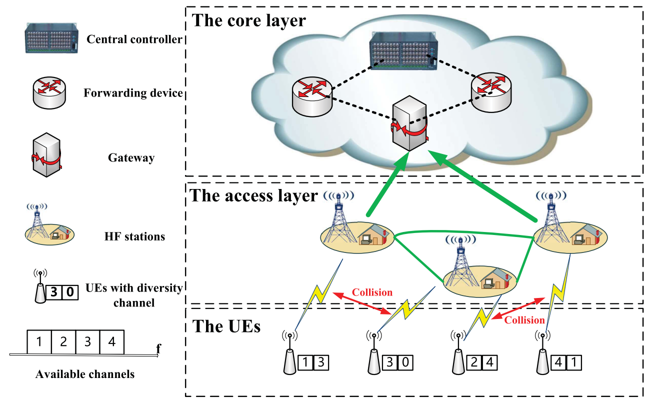


FIGURE 1. A distributed HF diversity network.

III. SYSTEM MODEL AND PROBLEM FORMULATION

A. SYSTEM MODEL

We consider a HF communication network (HFCN) consisting of a certain number of stations. All stations are connected by wired networks. Each station equipped with wideband transceiver demodulates signals over HF band (3-30MHz). The HFCN is divided into core layer (wired (IP) network) and access layer [12], which is shown in Fig. 1. The core layer is composed of central controller, forwarding device and gateway, and the access layer contains the HF stations. The core wired network receives data packets from transmitter through uplink and helps forward data packets to receivers through downlink. The user equipment (UE) supplied with several antennas transmits the same signal in several channels by channel diversity. Since the downlink is deployed by the central controller and the uplink is decided by a distributed way, we mainly discuss the multi-user access problem in the uplink. In each time round, the data transmission is executed and UEs contend for channels to access.

Assume that there are M UEs, and the set of UEs is denoted by $\mathcal{M}$, i.e., $\mathcal{M} = \{1, 2, \ldots, M\}$. Each UE is distributed in different place and has no information interaction with others. Channel diversity deployed in the networks allows each UE to access N channels at most, on which same signals are transmitted. The set of all available channels is denoted by $\mathcal{F} = \{1, 2, \ldots, F\}$. The set of available channels which has better communication performance can be obtained by perceiving or predicting by various methods [7]–[9]. When $F > MN$, it is feasible for channel-user allocation and there is no collision in the network. The proposed algorithm in this paper is applicable for both $F > MN$ and $F <= MN$ cases. We mainly consider the situation with $F <= MN$.

Since the state of HF channels is time-varying and unstable, the communication fails on certain time slots. As shown in Fig. 2, channel 1 is unavailable in time slots 2 and 5

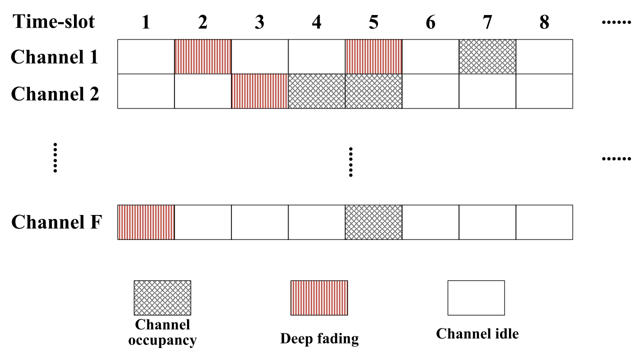


FIGURE 2. The HF channel status.

with deep fading, 7 with channel occupancy. In the wireless communication, when the signal to noise ratio (SNR) in the receiver is larger than the SNR threshold, the data transmission is successful. Since channel fading and the noise exist in wireless communication, the SNR in receiver becomes a random variable. The probability of received SNR greater than the threshold means the probability of successful data transmission. Thus, we can actually use a communication probability to describe channel quality. We define the communication probability of channels to describe the communication quality of channels. It means that the transmission succeeds with a statistical probability. The communication probability is a statistical variable and obtained by long-term measurement. The communication probability of all UEs is denoted by $P = (p_f^m)_{m\in\mathcal{M},f\in\mathcal{F}}$, where $p_f^m \in [0, 1]$ indicates the communication probability in channel f for UE m. The throughput of each UE is jointly determined by communication probability of access channels and the collisions caused by other UEs. The main problem is the selection of access channel, which determines the collision between UEs.

B. PROBLEM FORMULATION

When two and more UEs access the same channel simultaneously, a collision occurs and their data transmissions fail. Each UE finds out whether their transmission is successful according to the feedback ACK. For UE m, the set of its available access strategies access strategy is $\mathcal{A}_m$, i.e.,

$$\mathcal{A}_m = \{(f_1, f_2, \cdots, f_n, \cdots, f_N)|f_n \in \mathcal{F} \cup \{0\}\}. \tag{1}$$

Then, the set of strategy profiles for network is denoted by $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \cdots \times \mathcal{A}_M$. The element of $\mathcal{A}$ is $a = (a_m)_{m\in\mathcal{M}}$, where a_m is the access strategy of UE m. Each UE m chooses an access strategy $a_m = (a_m(1), a_m(2), \cdots, a_m(N)) \in \mathcal{A}_m$, where $a_m(n)$ denotes the n-th channel UE m access. When $a_m(n) = 0$, the n-th antenna does not access channel and keep silent. Specially, the UE m does not access any channel with $a_m = (0, 0, \cdots, 0)$. Then when UE m chooses a strategy a_m, the individual average packets transmitted successfully is denoted by

$$s_m(a_1, a_2, \cdots, a_M) = \sum_{n=1}^{N} p^m_{a_m(n)} \prod_{j=1,j\neq m}^{M} f(a_m(n), a_j), \tag{2}$$

where M and N denote the number of UEs and the number of channels each UE can access at most, respectively, and $p^m_{a_m(n)}$ is the communication probability in channel $a_m(n)$ for UE m, and $f\left(a_m(n), a_j\right)$ is the Kronecker delta function defined as

$$f(a_m(n), a_j) = \begin{cases} 1, & a_m(n) \notin a_j \\ 0, & a_m(n) \in a_j \ or \ a_m(n) = 0. \end{cases} \tag{3}$$

where $a_m(n) \notin a_j$ means UE j does not access the channel $a_m(n)$. UE m with strategy $a_m = (0, 0, \cdots, 0)$ has to wait for next time slot and zero packet is transmitted successfully, i.e., $s_m = 0$. If the communication probability of UE m for all available channels is 1 and no collision happens in all selected channel $a_m(n)$, the maximum value is $s_m = N$. According to (2), the average packets transmitted successfully for whole network is given by

$$T_0 = \sum_{m=1}^{M} s_m = \sum_{m=1}^{M}\sum_{n=1}^{N} p^m_{a_m(n)} \prod_{j=1,j\neq m}^{M} f(a_m(n), a_j). \tag{4}$$

Then, the global goal is to find the optimal channel access profile to maximize the average packets transmitted successfully for whole network, i.e.,

$$P: \quad \max T_0(a_1, a_2, \cdots, a_M). \tag{5}$$

The number of access strategy profiles is calculated as $(F + 1)^{MN}$ and it is really huge with large-scale network. At the same time, we have to solve the problem in a distributed way, because the UE is lack of information about others. Thus, the problem of channel access is challenging due to huge strategy space and distributed decision making.

IV. GAME MODEL AND MULTICHANNEL ACCESS ALGORITHM

A. MULTICHANNEL ACCESS GAME MODEL

Since the access strategy is self-determined by the UEs, the channel access problem is formulated as a non-cooperative game. The multichannel access game is denoted as $\mathcal{G} = (\mathcal{M}, \{\mathcal{A}_k\}_{k\in\mathcal{M}}, \{u_k\}_{k\in\mathcal{M}})$, where $\mathcal{M}$ is the set of UEs, $\mathcal{A}_k$ is the action space of UE k and u_k is the utility function of UE k. The available action of all UEs is same, which is defined in (1). The set of strategy profiles is denoted by $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \cdots \times \mathcal{A}_M$ and the element of $\mathcal{A}$ is $a = (a_n)_{n\in\mathcal{M}}$, where a_n is the access strategy of UE n. The utility function of UE k is denoted by $u_k(a_k, a_{-k}) : a \to \mathbb{R}$, where a_k is action of UE k and a_{-k} is the action profile of all the players excluding player k. The utility of UE k is defined as

$$u_k(a_k, a_{-k}) = s_k = \sum_{n=1}^{N} p^k_{a_k(n)} \prod_{j=1,j\neq k}^{M} f(a_k(n), a_j), \tag{6}$$

where s_k is the average number of successful packets for user k defined in (2). Then, the definition of Nash Equilibrium (NE) and Correlated Equilibrium (CE) is given.

Definition 1 (Nash Equilibrium (NE) [25]): An action profile $A^* = (a_1^*, \ldots, a_N^*)$ is a pure strategy NE of game $\mathcal{G} =$

$(\mathcal{N}, \{A_n\}_{n\in\mathcal{N}}, \{u_n\}_{n\in\mathcal{N}})$ *if and only if no player n can improve its utility by deviating unilaterally, i.e.,*

$$u_n(a_n^*, a_{-n}^*) \geq u_n(a_n, a_{-n}^*), \quad \forall n \in \mathcal{N},\ \forall a_n \in A_n,\ a_n \neq a_n^*. \tag{7}$$

The multichannel access algorithms can be proposed to get the various NE points, but it is not desirable for two reasons. On the one hand, the pure-strategy NE brings higher payoff, but they may be unfair to the user who always uses less aggressive action [17]. On the other hand, the mixed-strategy NE is fair, but not efficient: the expected payoff may be small [18]. In [33], the concept of Correlated Equilibrium (CE) is proposed to cope with the above problems.

Definition 2 (Correlated Equilibrium (CE) [36]): A probability distribution ψ on $\mathcal{A}$ is a correlated equilibrium of $\mathcal{G} = (\mathcal{M}, \{\mathcal{A}_m\}_{m\in\mathcal{M}}, \{u_m\}_{m\in\mathcal{M}})$, if $\forall m \in \mathcal{M}$, $\forall j \in \mathcal{A}_m$ and $\forall k \in \mathcal{A}_m$, we have

$$\sum_{a_{-m}} \psi(j, a_{-m})[u_m(k, a_{-m}) - u_m(j, a_{-m})] \leq 0. \tag{8}$$

The specific theoretical analysis and proof can be found in [36]. In other word, a CE is a probability distribution over strategy profiles $\mathcal{A}$. Each player chooses the recommending action according to coordination device. If all players do not intend to deviate from the recommended action, the strategy profile is a CE [18].

Theorem 1: There always exists a CE in the multichannel access game $\mathcal{G} = (\mathcal{M}, \{\mathcal{A}_k\}_{k\in\mathcal{M}}, \{u_k\}_{k\in\mathcal{M}})$.

Proof: In [36], [40], the authors figure out that every finite game has at least one CE. Since the set $\mathcal{M}$ and $\mathcal{A}_k$ are finite, the set of Correlation Equilibrium is nonempty in multichannel access game, which confirms our theorem. The more detailed analysis about the existence of CE can be found in [40]. □

At present, there are three main methods for obtaining the CE: linear programming [34], [35], regret-matching [36]–[39], and multi-agent learning [17], [18]. The linear programming works with all information available for optimization. It is not possible for distributed networks. The regret-matching proposed in [40] is a distributed learning algorithm based on the history information including other users' actions and updates its policy using the regret matrix. However, the history information may not be observable or correct. In [18], a multi-agent learning-based algorithm which need not rely on the history is proposed. The coordination device in the network does not need to know anything about users and it just send a randomly chosen integer at each time slot. Motivated by [18], we propose a multichannel access algorithm for HF networks in next section.

In HFCN, since each UE has no knowledge of others, the CE of the multichannel access game is difficult to reach without coordination. Inspired by the learning-based algorithm in [18], which achieved an efficient and fair CE, we propose an algorithm to solve the channel access problem in HFCN. The HF core network just sends a random integer called coordination signal from a set $\{1, 2, \ldots, K\}$ to UEs at begin of each time slot and each UE observes the random integer from time to time.[1] The HF core network need not know anything about the UEs, and just provide coordination signals. For each UE, each coordination signal is mapped to an access strategy. UEs receive the coordination signals before each time slot and decide which access strategy to choose according to the multichannel access strategy table (MAST) or keep silent. If the transmission on a channel fails, then they may change their access strategy. Otherwise, they keep access this channel at next time slot. By above way can we achieve an efficient and fair channel access strategy profile. Next, we will explain the algorithm in detail.

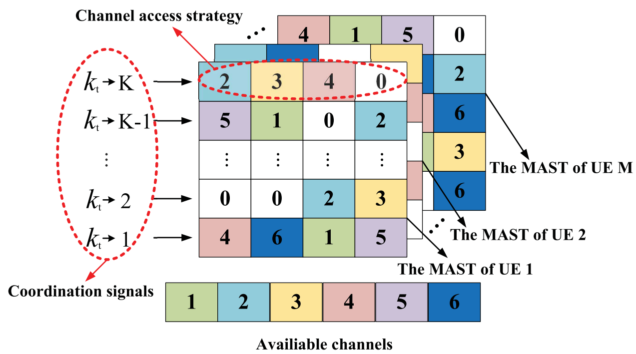


FIGURE 3. **The multichannel access strategy table (MAST) of UEs.**

B. MULTICHANNEL ACCESS ALGORITHM (MAA) DESCRIPTION

In section III, we define the set of UEs and available channels, i.e., $\mathcal{M} = \{1, 2, \cdots, M\}$ and $\mathcal{F} = \{1, 2, \cdots, F\}$, respectively. The channel diversity allows each UE to access N channels at most. The space of available coordination signals is denoted by $\mathcal{K} = \{1, 2, \cdots, K\}$. In HFCN, the HF core network chooses a random coordination signal and transmits it at begin of each time-slot. Each UE possesses a multichannel access strategy table (MAST). The MAST of UE m is defined as $f_m : \mathcal{K} \rightarrow \mathcal{A}_m$, where $\mathcal{A}_m$ is the set of access strategy. The $\mathcal{A}_m$ contains all available access strategy of UE m. In the MAST of each UEs, it contains the access strategies corresponding to each coordination signal. For example, in Fig. 3, k_t is the coordination signal in t-th transmission and each row represents the channel access strategy. When the coordination signal is K, the access strategy UE 1 will access is $(2, 3, 4, 0)$ and UE M will access $(4, 1, 5, 0)$. Each coordination signal corresponds to an access strategy for every UE. The MAST instructs each UE m to select access strategy $a_m = (a_m(1), a_m(2), \cdots, a_m(N))$ in every time slot according

[1] At the algorithm description, we assume that each UE can observe the coordination signals correctly. However, all UEs are hard to observe the coordination signals correctly at each time slot, because the communication environment is complex and dynamic. We give the simulations in section V that UEs may observe the incorrect coordination signals. We find that though UEs may accept the incorrect coordination signals, the proposed algorithm is still effective.

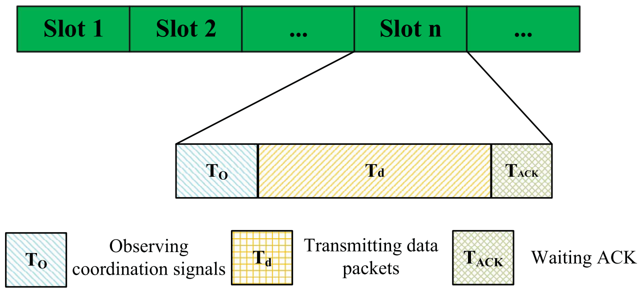


FIGURE 4. The structure of time slot.

to the received coordination signal k_t, i.e., $f_m(k_t) = a_m$, where $a_m(N)$ means the N-th antenna of UE m sends signal in channel $a_m(N)$. If i_{th} element of $f_m(k_t)$ is 0, the i-th antenna of UE keeps silent in this time-slot. Otherwise, the antenna sends signal immediately. At the beginning of learning phase, each UE m initializes its MAST randomly. For each coordination signal in $\mathcal{K}$, UE m chooses N different channels from $\mathcal{F} \bigcup \{0\}$ randomly. By randomized manner can the UEs access each channel fairly. In the meantime, the collisions are unavoidable with stochastic MAST at the beginning. As shown in Fig. 4, the time slot can be divided into three phases: observing coordination signals T_O, transmitting data packets T_d and waiting ACK T_{ACK}. Each UE adapts its strategy as follows:

1) **Observing coordination signals**: At time slot t, UEs wait for the coordination signal k_t transmitted by the HF core network. According to signal k_t and MAST, each UE obtains its access strategy $f_m(k_t)$.
2) **Transmitting data packets**: The i-th antenna of UE m accesses the channel $f_m^i(k_t)$. When $f_m^i(k_t) > 0$, The i-th antenna begins to transmit over channel $f_m^i(k_t)$. Otherwise, if $f_m^i(k_t) = 0$, the i-th antenna keeps silent and UE m chooses a channel $c \in \mathcal{F}$ with probability q_c to monitor, where q_c is normalized probability of channel communication probability, i.e,
$$q_c = \frac{\exp(\beta \times p_c^m)}{\sum_{f=1}^{F} \exp(\beta \times p_f^m)}, \tag{9}$$
where $\beta \geq 0$ is the adjustment factor. Specially, if $\beta = 0$, UE m chooses a monitoring channel randomly. UE m will monitor the channel c and observe whether the channel c is idle.
3) **Waiting ACK**: In the time slot t, the UEs wait for the ACK. There are two cases,
   - The i-th antenna transmitting in channel $f_m^i(k_t)$: if the transmission succeeds, UE will keep this channel in his channel strategy unchanged. if the transmission fails, it will set $f_m^i(k_t) = 0$ with switching probability p_d, which means to avoid collisions in future time slots.
   - The i-th antenna keeping silent in channel $f_m^i(k_t)$: If there is any UEs transmitting in channel c, UE m keeps $f_m^i(k_t) = 0$. If channel c is idle, then UE m sets $f_m^i(k_t) = c$.

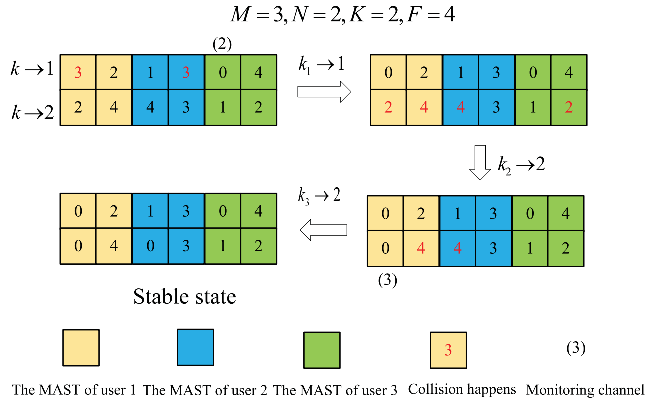


FIGURE 5. A schematic diagram of the proposed algorithm.

In the Fig. 5, we give an easy example to show how our proposed algorithm works. The different color represents different user's MAST and the different row maps to the different coordination signals. At the first transmission, the coordination signal is 1. User 1 and user 2 chooses the same channel 3, then their transmissions in channel 3 fail. According to the updating process, the user 1 gives up channel 3 and set access channel as 0. Since the user 3 has a channel 0, it chooses channel 2 to monitor. After the transmission, it finds that the channel 2 is occupied and then keeps the channel 0 unchanged. We can find that the access strategies of all users keep unchanged, when the coordination signal is 1. All users change their MASTs following above updating process. After 3-th transmission, the network convergence to the stable state.

The flow chart of the MAA algorithm is shown in Fig. 6. The normalized probability in (9) is adopted in the phase of transmitting data packets, where UE chooses the monitoring channel c with probability p_c. It is different from the algorithm in [18] which chooses the monitoring channel with the same probability. In [18], all channels were considered to be same. However, it is not true in HF networks, as HF communication is sensitive to the communication frequency. Different HF channels will provide different communication performance. Since the communication probability of same channel is different for different UEs, a channel may be bad for one user but good for others. Selecting monitoring channel with same probability can not make full use of the above characteristic. Thus, we introduce normalized probability to improve network throughput. The UE will select a better HF channel with a higher probability.

C. ALGORITHM ANALYSIS

Theorem 1 have proved that there existed at least one CE point in the multichannel access game. The CE point means that no UE wants to change its access strategy. In this section, we first prove that the network converges to an absorbing state, and

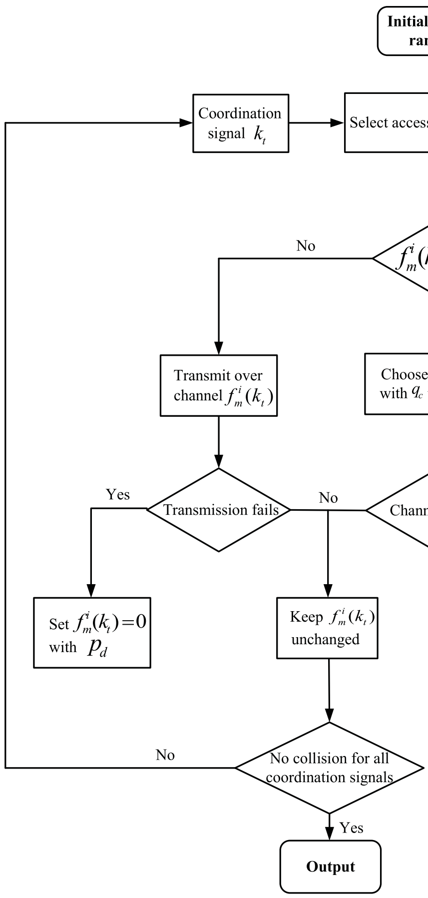


FIGURE 6. **Flow chart of the proposed algorithm.**

then the expected convergence time to the state and fairness among UEs are analyzed.

Firstly, we prove the following theorem:

Theorem 2: The proposed algorithm converges to an absorbing state, and each UE keeps its access strategy unchanged for all coordination signals in the absorbing state.

Proof: The network contains M UEs which can access N channels, available channels $F \geq 1$ and coordination signals $K \geq 1$. For each coordination signal $k \in \mathcal{K}$, we define the network state as $s_k(t) = (n_1^k(t), n_2^k(t), \cdots, n_F^k(t))$, where $n_F^k(t)$ is the number of antennas that will transmit packet in t-th time slot on the channel F. If all elements in $s_k(t)$ are equal to 1, i.e., $s_k(t) = (1, 1, 1, \cdots, 1)$, all channels are occupied and there is no collision in network. According to the proposed algorithm, since no collision happens, the antennas which occupy one channel will not change its access channel. Simultaneously, others will find that the monitoring channel is busy and keep their access strategy. Thus, the state $s_k(t) = (1, 1, 1, \cdots, 1)$ is an absorbing state and each UE will not change its access strategy. When all UEs update their access strategy using the MAA algorithm, we find that all states are interoperable. Thus, the network will definitely converge to the absorbing state $s_k(t) = (1, 1, 1, \cdots, 1)$ [41]. As the coordination signal is selected randomly, the network will go into the absorbing state for each coordination signal. Therefore, each UE keeps its access strategy unchanged for all coordination signals after several steps, which confirms our theorem. □

The convergence time is the number of steps to the absorbing state for all coordination signals. The fairness among UEs indicates whether each UE has equal chance to access all channels, when the network converges to the absorbing state.

1) CONVERGENCE TIME

We analyze the upper bound of convergence time when the multichannel access game converges to a pure-strategy NE. The HFCN including M UEs, K coordination signals and each UE can access N channels. The analysis is divided into three different scenarios. According to analysis, the following theorems are given.

Theorem 3: For M UEs which can access N channels and $F = 1$, $K = 1$, $0 < P_d < 1$, the expected number of steps to converge to a pure-strategy NE of multichannel access game is $O(\frac{1}{P_d(1-P_d)} \log M)$.

Proof: Since there is only one available channel in HFCN, only one UE can transmit in the stable state. We define the network state which is how many UEs are transmitting data on this channel. Since the state is only related to the state of previous time slot, a Markov chain is used to describe its execution. The state is $X_t \in \{0, 1, 2, \cdots, M\}$ at time slot t. For each coordination signal, the transition probability is expressed as:

$$\begin{cases} P(X_{t+1} = N | X_t = 0) = 1 \\ P(X_{t+1} = 1 | X_t = 1) = 1 \\ P(X_{t+1} = j | X_t = i) = C_i^j p_d^{i-j} (1 - p_d)^j \ i > 1, j \leq i \\ 0 \ others. \end{cases} \tag{10}$$

The state transition diagram is shown in Fig. 7. The state $X_t = 1$ is the absorbing state. We want to know the number of steps that the chain first reaches the state $X_t = 1$, which is called as the hitting time. In fact, the expected hitting time can be found by following linear equations.

$$u_i(C) = 1 + \sum_{j \in X_t} p_{ij} u_j(C), \tag{11}$$

where $u_i(C)$ denotes the expected hitting time from state i to the absorbing state C, i.e., $X_t = 1$, and p_{ij} is the transition probability from i to j. However, it is difficult to solve them analytically. In [18], the authors calculate the upper bound of the hitting time by a modified Markov chain. According to its calculation, the expected number of steps before reaching the absorbing state $X_t = 1$ is $O(\frac{1}{P_d(1-P_d)} \log M)$. □

In general, each UE will choose N different channels. It is noted that more collisions will occur, if one UE choose same channel to transmit. Thus, the expected number of steps is less than MN UEs who accesses single channel. According to [18], the following theorems are given:

Theorem 4: For M UEs which can access N channels and $F \geq 1$, $K = 1$, $0 < P_d < 1$, the expected number of steps to

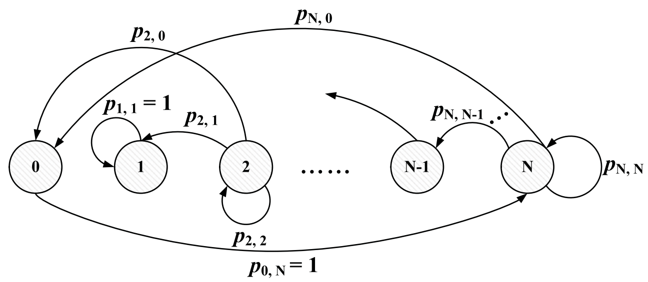


FIGURE 7. The state transition diagram of the proposed Markov chain.

converge to a pure-strategy NE of multichannel access game is $O(F\frac{1}{1-P_d}(\frac{1}{P_d}\log MN+F))$.

Proof: In general, each UE will choose N different channels, when $F \geq 1$. We propose a new scenario where there are MN UEs which only choose one channel. In this paper, each UE will not choose a same channel except 0. However, all UEs in the new scenario are independent. The more same channels will be selected by different UEs in the new scenario. Thus, it is noted that more collisions will occur in the new scenario than our scenario. Thus, the expected number of steps for our scenario is less than the new scenario. Firstly, we analyze the convergence time of the new scenario with MN UEs. According to the analysis in [18], the expected number of steps to the absorbing state for our scenario is $O(F\frac{1}{1-P_d}(\frac{1}{P_d}\log MN+F))$. □

Theorem 5: For M UEs which can access N channels and $F \geq 1$, $K \geq 1$, $0 < P_d < 1$, the expected number of steps to converge to a pure-strategy NE of multichannel access game is less than $O(K^2F\frac{1}{1-P_d}(\frac{1}{P_d}\log MN+F))$.

Proof: The problem of calculating convergence time with $k > 1$ can be analyzed by a more general problem called *Coupon collector's problem* [18]. Similar to *Theorem* 3, we can find that the expected number of steps is $O(K^2F\frac{1}{1-P_d}(\frac{1}{P_d}\log MN+F))$. □

From [18], we conclude that any Nash equilibrium (NE) is a correlated equilibrium (CE). Moreover, the MAA algorithm can converge to efficient pure-strategy NE where no collision occurs. Therefore, the proposed algorithm MAA converges to an efficient CE of multichannel access game in expected polynomial time, which is proportional to K, F, M and N.

2) FAIRNESS

As UEs quit the certain channel with the same probability P_d, the access strategy in the stable state is related to the initial access strategy for each coordination signal. Simultaneously, the initial strategy is generated randomly. Thus, when the number of coordination signals is large enough, each UE has the same chance to access all channels. For a given correlation signal, if one antenna of the UE can access one channel, this antenna wins this time slot. For F available channels and M UEs with N antennas, an antenna wins a time slot with probability $p = C_{MN-1}^{F-1}/C_{MN}^{F} = \frac{F}{MN}$. Denote random variable $X_{m(i)}$ as the number of slots won by antenna $m(i)$ of UE m for all coordination signals. Thus, $X_{m(i)}$ follows a binomial distribution, i.e., $X_{m(i)} \sim B(K,p)$. Denote random variable X_m as the sum of time slots won by all its antennas, i.e., $X_m = \sum_{i=1}^{N} X_{m(i)}$. Since all antennas are independent and identically distributed (i.i.d.), X_m also follows a binomial distribution, i.e., $X_m \sim B(NK,p)$. We use the *Jain index* defined in [42] to calculate the fairness. For a random variable X_m, the *jain index* is calculated as:

$$J(X) = \frac{(E[X])^2}{E[X^2]}. \tag{12}$$

Where its first and second moments are

$$E[X] = K\frac{F}{M}, \tag{13}$$

$$E[X^2] = (K\frac{F}{M})^2 + K\frac{F}{MN}\frac{MN-F}{M}. \tag{14}$$

The *Jain index* $J(X)$ can be expressed as:

$$J(X) = \frac{NFK}{NFK+(MN-F)}. \tag{15}$$

The *Jain index* satisfies that $0 < J(X) \leq 1$. The larger $J(X)$ is, the better the fairness of access strategy is. The access strategy is absolutely fair if $J(X) = 1$, which means that all UEs win equal number of channel across all the time slot. As $MN - F$ is finite, it holds that $\lim_{K\to\infty}(MN-F)/K \to 0$. We can obtain the fact that

$$\lim_{K\to\infty} J(X) = \lim_{K\to\infty}\frac{NF}{NF+(MN-F)/K} \to 1. \tag{16}$$

If we choose a relatively large K, the access strategy of UEs become absolutely fair. Thus, the fairness is proportional to the coordination signals space.

V. SIMULATION RESULTS AND DISCUSSION

In this section, the performance of the proposed algorithm (MAA) is evaluated through simulations. The simulations work on the MATLAB. All simulation results are obtained by performing 300 independent trials and then taking the expectation. The UEs are randomly located in a $500km \times 500km$ square region and all UEs transmit in each time slot, if they can access the channels. The adjustment factor β in (9) is set to 1. The communication probability is randomly generated within (0, 1). We mainly discuss the performance of the MAA algorithm from following aspects. First, we are interested in how many steps does the HFCN need to converge to the stable state with different UEs M, coordination signals K and switching probability P_d. Then, the normalized throughput is obtained and compared with multichannel Slotted-ALOHA [43] with different number of UEs. Finally, we simulate how to choose K to achieve a reasonable bound on fairness.

A. CONVERGENCE PERFORMANCE

Different from the MAA algorithm (the proposed algorithm), when the access strategy of i-th antenna satisfies $f_m^i(k_t) = 0$, the comparison algorithms choose the monitoring channel in different ways. In the proposed algorithm, we choose a

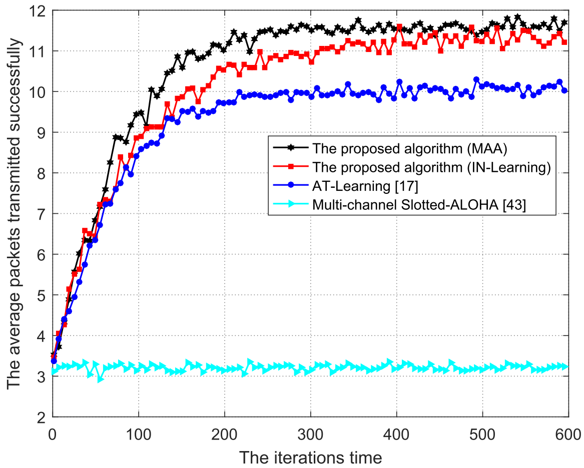


FIGURE 8. The average packets transmitted successfully through different algorithms ($M = 10, N = 3, F = 20, K = 20$).

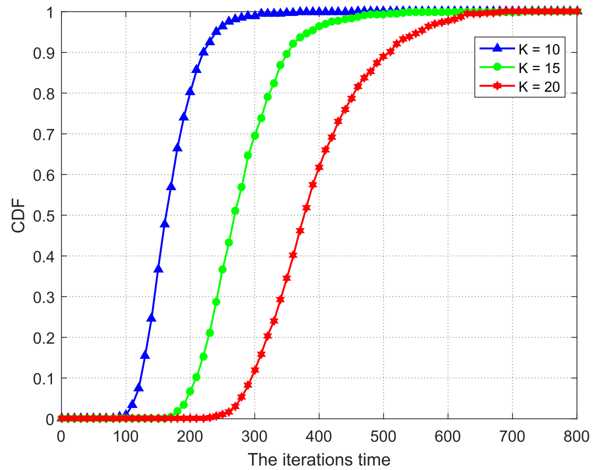


FIGURE 9. The convergence time of the MAA for different network scales ($M = 10, N = 3, F = 20$).

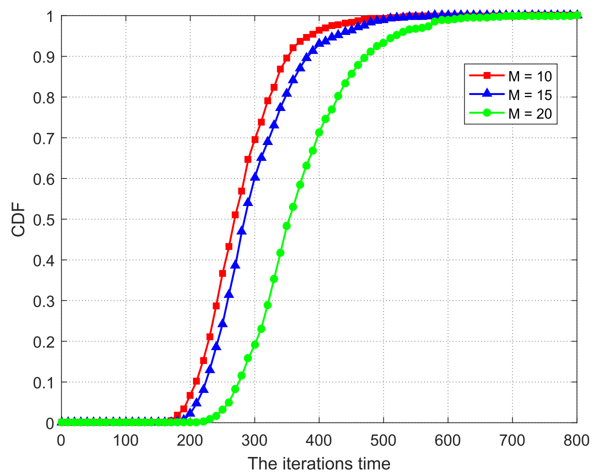


FIGURE 10. The convergence time of the MAA for different network scales ($N = 3, F = 20, K = 15$).

monitoring channel c from all available channels with normalized probability q_c which defined in (9). The comparison algorithms are as follow: (1)

1) Attachment-learning algorithm (AT-Learning) [17]: The UE chooses the monitoring channel from all available channels with the same probability.
2) Idle_normalize learning algorithm (IN-Learning): The UE chooses the monitoring channel h from the idle channels with the normalized probability q_h and the idle channels can be obtained through wideband spectrum sensing (WBSS). The q_h is calculated as follow

$$q_h = \frac{\exp(\beta \times p_h^m)}{\sum\limits_{h \in I} \exp(\beta \times p_h^m)}, \tag{17}$$

where I is the set of idle channels and p_h^m means the communication probability in channel h for UE m.
3) Multichannel slotted-ALOHA [43]: The UE randomly chooses access channels from all available channels without learning at each time slot.

Fig. 8 presents the convergence behavior of four different algorithms with $M = 10$ UEs, $N = 3$ antennas, $F = 20$ available channels and $K = 20$ coordination signals. We can see that the proposed algorithms using coordination signals converge to the stable state with stochastic initial access strategy and have a significant improvement than the multichannel Slotted-ALOHA. This is reasonable and consistent with the previous analysis. At the same time, the MAA algorithm has the best communication performance. The MAA algorithm converges to the stable state faster than the IN-Learning algorithm. The reason is that all UEs can obtain the idle channels by WBSS, for which there are more collisions for the same idle channels. Thus, choosing monitoring channel from all available channels is better when UEs' channel strategy is 0. Simultaneously, it is noted that the MAA and IN-Learning using the normalized probability perform better than the AT-Learning [17]. The normalized probability proposed in this paper can bring a higher throughput, because UEs are more inclined to choose better channels. It is noted that the convergence time to the stable state is less than 400.

We compare the convergence time of different network scales in Fig. 9 and Fig. 10. The cumulative distribution function (CDF) of the iterations time needed to converge to the stable state with different UEs and coordination signals are shown in Fig. 9 and Fig. 10, respectively. It is noted from the figures that the iterations to the stable state is proportional to the number of UEs and coordination signals. With the number of UEs increasing, the more collisions may happen and UEs have to adapt their MAST more frequently, which results in a longer convergence time. With the number of coordination signals increasing, the UEs have to wait for a certain coordination signal to be transmitted with a longer steps. This is consistent with our previous theoretical analysis. The convergence time is acceptable even for a large-scale network.

Fig. 11 presents the number of convergence iterations with different switching probability P_d. It is noted form the pic-

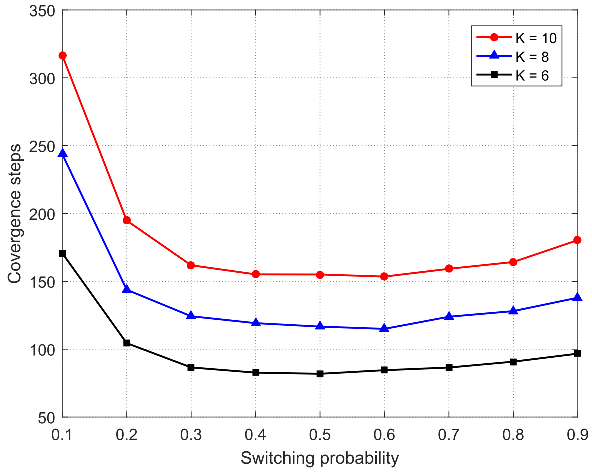


FIGURE 11. Number of iterations that network converges to the stable state with different coordination signals and switching probability ($M = 10, N = 3, F = 20$).

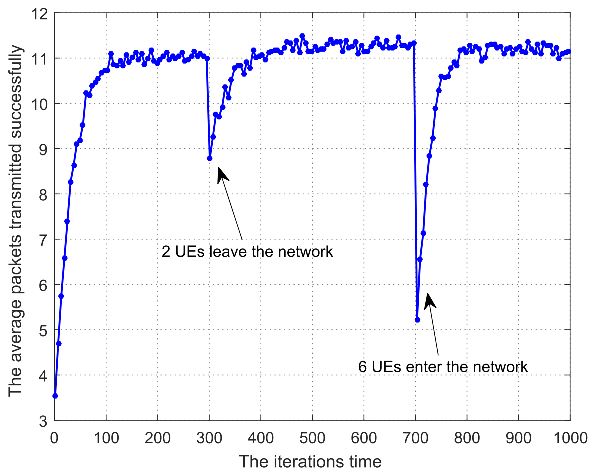


FIGURE 12. Convergence behavior of dynamic networks scales ($N = 3, F = 20, K = 10, P_d = 0.4$).

ture that the convergence time increases as the coordination signals K increase. The convergence time is smaller, with $P_d = 0.5$ than other probabilities, since it provide the balance between switching to 0 and keeping the access strategy. When the probability P_d is small, few UEs change their access strategy to 0 and the collisions keep happening in next time slot. If more UEs change their strategy to 0, however, the channel may become free and the more collisions may happen in this channel.

In the actual HFCN, the UEs may leave or join the HFCN, which make the scale of HFCN dynamic and time-varying. The convergence behavior of dynamic HFCN is shown in Fig. 12. It can be noted that the number of packets transmitted successfully decreases, when two UEs leave the network or six UEs join the network with stochastic access strategy. With two UEs leaving, the channels occupied by the two UEs become free, then other UEs will compete for these channels. When six UEs join the network, their strategy has conflict with other UEs and it makes the throughput drop

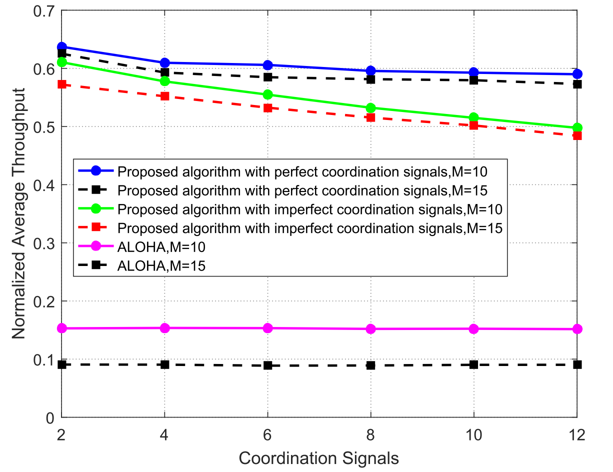


FIGURE 13. The comparison of the normalized throughput with varying coordination signals K ($F = 20, N = 3$).

rapidly. However, from the picture, we see that the network converges to the stable state quickly with few steps after the number of UEs changes. Therefore, the MAA can work well in the dynamic environment.

B. AVERAGE THROUGHPUT PERFORMANCE

In this subsection, we evaluate the throughput performance of the MAA comparing with multichannel Slotted-ALOHA and the proposed algorithm with imperfect coordination signals. The proposed algorithm with imperfect coordination signals means the UEs receive no coordination signal or misjudge into another signal. When users receive no coordination, they randomly choose access strategy from the MAST. When some users observe the incorrect coordination signals (including receiving no signal and misjudging), they will choose a wrong access strategy which may make collisions happen.. The normalized throughput is used to evaluate the performance of UEs, which is defined as

$$Th_n = \frac{N_{succ}}{N_{slot} \times F}, \tag{18}$$

where N_{slot} and F are the number of time slots UEs transmit and available channels, respectively. N_{succ} is the number of packets transmitted successfully during N_{slot}, which is defined in (4). When the UEs use the multichannel Slotted-ALOHA, they choose access strategy randomly. Fig. 13 presents the normalized throughput with varying coordination signals space after reaching the stable state. It can be noted that the throughput of the Slotted-ALOHA is really poor, with only ten percent of data packets transmitted successfully. Since the UEs use the random access strategy, the collisions happen frequently with the number of UEs increasing. However, the MAA can avoid collisions and achieve sixty percent normalized throughput with random communication probability. When UEs receive the incorrect coordination signals, the stable state will be broken and more collisions occur. In the simulation, we set the incorrect probability as 5%. The normalized throughput of the

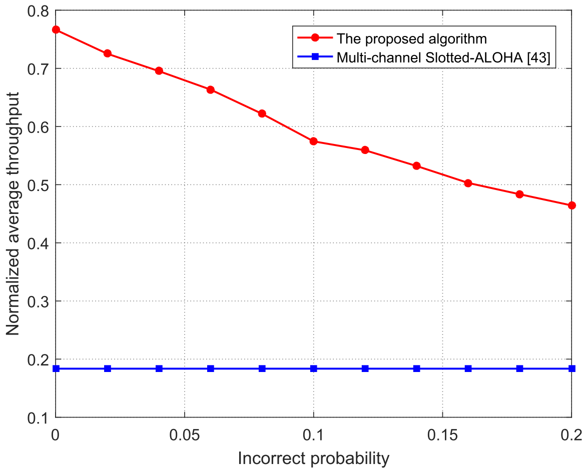


FIGURE 14. The comparison of the normalized throughput with varying error probability ($F = 20, N = 3, K = 8, M = 15$).

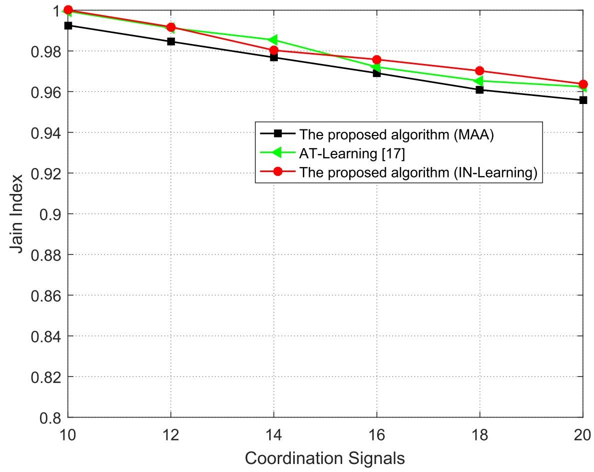


FIGURE 16. Fairness comparison of different algorithms ($N = 3, F = 30, K = 15$).

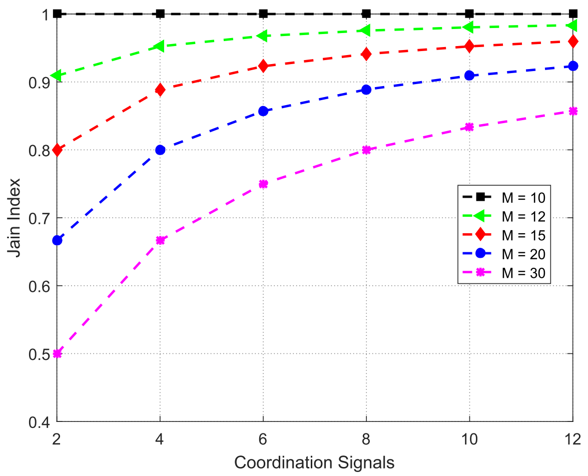


FIGURE 15. Jain fairness index for different settings of M and K ($F = 30, N = 3$).

proposed algorithm with imperfect coordination signals is smaller than the MAA. Nevertheless, it still achieve a high throughput. In Fig. 14, the normalized average throughput with varying incorrect probability is given. From the figure, we can see that the normalized average throughput drops quickly, when the incorrect probability increases. This is because more collisions happen with larger incorrect probability. However, it is noted that the network throughput using proposed algorithm is significantly higher than the Multichannel Slotted-ALOHA even if the incorrect probability is high. Thus, though all UEs may not correctly accept the coordination signals, the proposed algorithm is still effective.

C. FAIRNESS PERFORMANCE

In Sec. III, we have proved that the Jain Index $J(X)$ is close to 1 with a relatively large K, which proves the MAA algorithm is fair enough. Since the convergence time is really huge with a large K, we can not choose a large value of K in actual HFCN. Fig. 15 presents the *Jain Fairness Index* for different settings of M and F. It can be noted that when $MN = F$, the JFI approximates to 1, which means every UE can have the same chance to access all available channels. With the number of coordination signals increasing, the JFI increases and is close to 1, which meets our previous analysis. However, the JFI decreases, when the number of UEs increases. This is because there is not enough channels for each UE in every time slot. It is noted that K can not be too large since convergence time will be too long, nor can it be too small to ensure fairness. For different networks, we need balance convergence time and fairness. In our simulation network which contains 30 available channels, 20 UEs and 3 antennas, it is reasonable to choose 8 or 10 coordination signals that can reach the JFI about 0.9. We compare the fairness with Attachment-learning algorithm (AT-Learning) and Idle_normalize learning algorithm (IN-Learning) in Fig. 16. The HF network contains 30 available channels and the number of coordination signals is 15. It is noted that all three algorithms have high fairness and the Jain fairness index (JFI) is close to 1. we can observe that the proposed algorithm has similar JFI to other two algorithms and it also brings higher network throughput.

VI. DISCUSSION

The proposed algorithm in this paper are extensible and can be applied to other wireless networks, such as Cellular network, Internet of Things (IOT), and Unmanned aerial vehicle (UAV). The reason is as follows. 1) In this manuscript, we use a communication probability to describe channel quality and it is also suitable for other wireless networks. 2) The rapid development of Multi-input Multi-output (MIMO) technology and its high-capacity advantage have led to the use of multiple antennas in other networks. For example, MIMO technology will become a key technology in the 5G. Therefore, multichannel access issues also need to be considered in other networks. 3) The proposed algorithm does not need to rely on a complex coordination device. The device just sends an integer number at each time slot and all UEs learn to choose action for each signal value. Thus, the proposed algorithm is also very easy to implement in other networks. However, the proposed algorithm works under the condition that all users are in the same collision domains. If other wireless networks satisfy this condition, they can also use the proposed algorithm.

VII. CONCLUSION

In this paper, we have proposed a multi-user learning algorithm to deal with the problem of multichannel access in distributed HFCN. First, we formulated the channel access problem as a non-cooperative game and the Correlated Equilibrium (CE) of the access game was analyzed in terms of convergence time and fairness. Then, we used the coordination signals and the collision information to reduce the collision level. These coordination signals help UEs learn the access strategy by themselves. We verified the effectiveness and fairness of the proposed algorithm by simulation. We found that after the learning stage, all UEs know exactly when to transmit and on which channel according to their multichannel Access Strategy Table (MAST). The simulation results showed that the proposed learning algorithm can not only completely avoid interference and get optimal throughput, but also guarantee the fairness among all UEs. In this manuscript, we just give a numerical simulation based on MATLAB and a simplified channel model is used. For future work, the HF channel model will be considered and we intend to implement the proposed approach in practical system.

REFERENCES

[1] J. Wang, *Research and Development of HF Digital Communications* (in Chinese), Beijing, China: Science Press, 2013.

[2] R. Hanson, "Military applications of HF communications," *Mil. Technol.*, vol. 34, no. 10, pp. 70–77, 2010.

[3] E. Johnson, E. Koski, W. N. Furman, M. Jorgenson, and J. Nieto, *Third-generation and Wideband HF Radio Communications*. Norwood MA, USA: Artech House, 2013.

[4] W. N. Furman and E. Koski, "Next generation ALE concepts," in *Proc. Inst. Eng. Technol. 11th Int. Conf. Ionospheric Radio Syst. Techn.*, Edinburgh, U.K., 2009, pp. 152–156.

[5] J. Wang, G. Ding, and H. Wang, "HF communications: Past, present, and future," *China Commun.*, vol. 15, no. 9, pp. 1–9, 2018.

[6] Z. Qin, J. Wang, J. Chen, G. Ding, Y.-D. Yao, X. Ji, X. Chen, "Link quality analysis based channel selection in high-frequency asynchronous automatic link establishment: A matrix completion approach," *IEEE Syst. J.*, vol. 12, no. 2, pp. 1957–1968, Jun. 2018.

[7] L. Yu, J. Chen, Y. Zhang, H. Zhou, and J. Sun, "Deep spectrum prediction in high frequency communication based on temporal-spectral residual network," *China Commun.*, vol. 15, no. 9, pp. 25–34, 2018.

[8] L. Melián-Gutiérrez, N. Modi, C. Moy, F. Bader, I. Pérez-Álvarez, and S. Zazo, "Hybrid UCB-HMM: A machine learning strategy for cognitive radio in HF band," *IEEE Trans. Cogn. Commun. Netw.*, vol. 1, no. 3, pp. 347–358, Sep. 2015.

[9] X. Liu, Y. Xu, Y. Cheng, Y. Li, L. Zhao, and X. Zhang, "A heterogeneous information fusion deep reinforcement learning for intelligent frequency selection of HF communication," *China Commun.*, vol. 15, no. 9, pp. 73–84, 2018.

[10] A. Heritage, "HF 2000-The future of HF communications," *Mil. Technol.*, vol. 29, no. 3, pp. 81–82, 2005.

[11] P. Asenstorfer and J. B. Scholz, "'LongFish'—A HF radio network testbed," in *Proc. 7th Int. Conf. HF Radio Syst. Techn.*, Jul. 2000, pp. 198–200.

[12] K. Xu, B. Jiang, Z. Su, S. Wang, M. Guo, X. Li, and Z. Du, "High frequency communication network with diversity: System structure and key enabling techniques," *China Commun.*, vol. 15, no. 9, pp. 46–59, 2018.

[13] J. W. Nieto, "Spatial diversity and high data rate waveforms-is it worth doing?" in *Proc. 21st Century Mil. Commun., Archit. Technol. Inf. Superiority*, 2000, pp. 1144–1148.

[14] P. M. Ndao, Y. Erhel, D. Lemur, M. Oger, F. Marie, and J. Le Masson, "Test of HF (3–30 MHz) MIMO communication system based on polarisation diversity," *Electron. Lett.*, vol. 48, no. 1, pp. 50–51, 2012.

[15] H.-K. Song, "A frequency synchronization for TDMA system," *IEEE Commun. Lett.*, vol. 3, no. 4, pp. 113–115, Apr. 1999.

[16] F. Wang, C. Yu, X. Wu, and Y. Hu, "Dual time synchronisation method for wireless sensor networks," *Electron. Lett.*, vol. 51, no. 2, pp. 179–181, 2015.

[17] L. Wang, K. Wu, M. Hamdi, and L. M. Ni, "Attachment-learning for multi-channel allocation in distributed OFDMA-based networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1712–1721, Apr. 2013.

[18] L. Cigler and B. Faltings, "Reaching correlated equilibria through multi-agent learning," in *Proc. 10th Int. Conf. Auton. Agents Multiagent Syst.*, 2011, pp. 509–516.

[19] Z. Qin, W. Liu, Y. Jiang, W. Lim, and J. Zhou, "Multiple armed bandits based opportunistic channel access with switching cost in block fading channel," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, 2015, pp. 1–5.

[20] Z. Qin, J. Wang, J. Chen, Y. Sun, Z. Du, and Y. Xu, "Opportunistic channel access with repetition time diversity and switching cost: A block multi-armed bandit approach," *Wireless Netw.*, vol. 24, no. 5, pp. 1683–1697, 2018.

[21] L. Melián-Gutiérrez, N. Modi, C. Moy, I. Pérez-Álvarez, F. Bader, and S. Zazo, "Upper confidence bound learning approach for real HF measurements," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, Jun. 2015, pp. 381–386.

[22] E. Koski and W. Furman, "Applying cognitive radio concepts to HF communications," in *Proc. 11th Int. Conf. Ionospheric Radio Syst. Techn. (IRST)*, Apr. 2009, pp. 1–6.

[23] T. Vanninen, T. Linden, M. Raustia, and H. Saarnisaari, "Cognitive HF—New perspectives to use the high frequency band," in *Proc. 9th Int. Conf. Cognit. Radio Oriented Wireless Netw. Commun. (CROWNCOM)*, Jun. 2014, pp. 108–113.

[24] K. Akkarajitsakul, E. Hossain, D. Niyato, and D. I. Kim, "Game theoretic approaches for multiple access in wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 3, pp. 372–395, 3rd Quart., 2011.

[25] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y.-D. Yao, "Opportunistic spectrum access in cognitive radio networks: Global optimization using local interaction games," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 2, pp. 180–194, Apr. 2012.

[26] X. Chen and J. Huang, "Database-assisted distributed spectrum sharing," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2349–2361, Nov. 2013.

[27] J. Zheng, H. Zhang, Y. Cai, R. Li, and A. Anpalagan, "Game-Theoretic multi-channel multi-access in energy harvesting wireless sensor networks," *IEEE Sensors J.*, vol. 16, no. 11, pp. 4587–4594, Jun. 2016.

[28] W. Li, L. Ruan, Y. Xu, Y. Zhang, Y. Xu, and X. Shao, "Exploring channel diversity in HF communication systems: A matching-potential game approach," *China Commun.*, vol. 15, no. 9, pp. 60–72, 2018.

[29] W. Yuan, P. Wang, W. Liu, and W. Cheng, "Variable-Width channel allocation for access points: A game-theoretic perspective," *IEEE Trans. Mobile Comput.*, vol. 12, no. 7, pp. 1428–1442, Jul. 2013.

[30] I. Bistritz and A. Leshem, "Game theoretic dynamic channel allocation for frequency-selective interference channels," *IEEE Trans. Inf. Theory*, vol. 65, no. 1, pp. 330–353, Jan. 2019.

[31] W. Cheng, W. Zhang, L. Liang, and H. Zhang, "Full-duplex for multi-channel cognitive radio ad hoc networks," *IEEE Netw.*, vol. 33, no. 2, pp. 118–124, Mar./Apr. 2019.

[32] W. Cheng, X. Zhang, and H. Zhang, "Optimal power allocation with statistical QoS provisioning for D2D and cellular communications over underlaying wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 151–162, Jan. 2016.

[33] R. J. Aumann, "Subjectivity and correlation in randomized strategies," *J. Math. Econ.*, vol. 1, no. 1, pp. 67–96, 1974.

[34] Z. Lin and M. V. D. Schaar, "On the Correlated Equilibrium Selection for Two-User Channel Access Games," *IEEE Signal Process. Lett.*, vol. 16, no. 3, pp. 156–159, Mar. 2009.

[35] J.-B. Seo and H. Jin, "Two-user NOMA uplink random access games," *IEEE Commun Lett.*, vol. 22, no. 11, pp. 2246–2249, Nov. 2018.

[36] D. Wu, Y. Cai, L. Zhou, Z. Zheng, and B. Zheng, "Cooperative strategies for energy-aware ad hoc networks: A correlated-equilibrium game-theoretical approach," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2303–2314, Jun. 2013.

[37] Z. Han, C. Pandana, and K. J. R. Liu, "Distributive opportunistic spectrum access for cognitive radio using correlated equilibrium and no-regret learning," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2007, pp. 11–15.

[38] J. W. Huang and V. Krishnamurthy, "Cognitive base stations in LTE/3GPP femtocells: A correlated equilibrium game-theoretic approach," *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3485–3493, Dec. 2011.

[39] Z. Xiao, X. Shen, F. Zeng, V. Havyarimana, D. Wang, W. Chen, and K. Li, "Spectrum resource sharing in heterogeneous vehicular networks: A noncooperative game-theoretic approach with correlated equilibrium," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9449–9458, Oct. 2018.

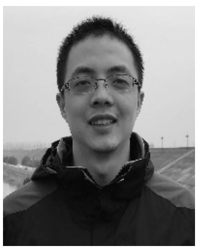
[40] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, Sep. 2000.

[41] J. G. Kemeny and J. L. Snell, *Markov Chains*. New York, NY, USA: Springer-Verlag, 1976.

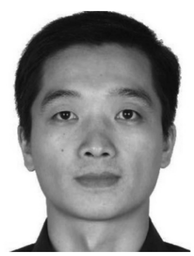
[42] R. Jain, D.-M. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," DEC Res., Hudson, MA, USA, Tech. Rep. TR-301, 1984, vol. 38.

[43] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1987.

WEN LI received the B.S. degree in electronic and information engineering from Tsinghua University, Beijing, China, in 2017. He is currently pursuing the M.S. degree with the College of Communication Engineering, Army Engineering University of PLA. His current research interests include game theory, dynamic optimization, and HF communication systems.

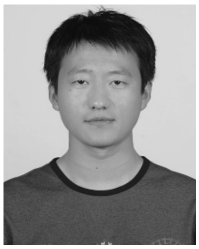


YUHUA XU received the B.S. degree in communications engineering and the Ph.D. degree in communications and information systems from the College of Communications Engineering, PLA University of Science and Technology, in 2006 and 2014, respectively. He is currently an Associate Professor with the College of Communications Engineering, Army Engineering University of PLA. He has published several papers in international conferences and reputed journals in his research area. His research interests include UAV communication networks, opportunistic spectrum access, learning theory, and distributed optimization techniques for wireless communications. He received the Certificate of Appreciation as an Exemplary Reviewer of the IEEE Communications Letters, in 2011 and 2012, respectively. He was selected to receive the IEEE Signal Processing Society 2015 Young Author Best Paper Award and the Funds for Distinguished Young Scholars of Jiangsu Province, in 2016.

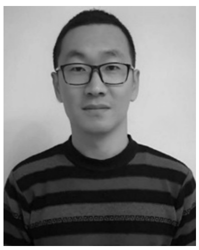


YUNPENG CHENG received the B.S., M.S., and Ph.D. degrees in communications and information systems from the Institute of Communications Engineering, Nanjing, China, in 1997, 2000, and 2003, respectively. He is currently a Professor with the Institute of Communications Engineering, Army Engineering University of PLA, China. His current research interests include HF communication and signal processing for wireless communications.

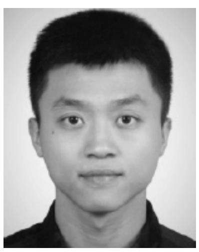
YANG YANG received the B.Eng. and Ph.D. degrees in communication and information engineering from the PLA University of Science and Technology (PLA UST), China, in 2005 and 2011, respectively. Since 2012, she has been a Teacher and a Researcher with the Institute of Communications Engineering, Army Engineering University of PLA. Her current research interests include UAV-assisted communication, multi-antenna transceiver optimization, dynamic spectrum sharing and power allocation in cognitive radio, and other wireless communication related topics.



XUEQIANG CHEN received the B.S. degree in electronics and communications engineering from the Qilu University of Technology, Jinan, China, in 2008, and the Ph.D. degree from the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China. He is currently a Lecturer with the College of Communication Engineering, Army Engineering University. His research interests include cognitive radio, opportunistic spectrum access, and game theory.



MENG WANG received the B.S. and M.S. degrees in instructional technology from the Nanjing Normal University, in 2009. His current research interests include wireless network security, game theory, and intelligent anti-jamming technology.



DIANXIONG LIU received the B.S. degree from the School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou, China, in 2014, and the M.S. degree from the College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2017. He is currently pursuing the Ph.D. degree in communications and information system with the College of Communications Engineering, Army Engineering University of PLA, Nanjing. His research interests include resource allocation in relay models, cognitive radio networks, UAV communication networks, and game theory.