

引用格式:吴迪,钱鹏智,陈勇. 恶意干扰下基于斯坦伯格博弈的无人机动态频谱分配[J]. 信息对抗技术, 2022, 1(1):73-80. [WU Di, QIAN Pengzhi, CHEN Yong. Dynamic spectrum allocation algorithm of UAV based on Stackelberg game in the condition of malicious interference [J]. Information Countermeasure Technology, 2022, 1(1):73-80. (in Chinese)]

恶意干扰下基于斯坦伯格博弈的无人机动态频谱分配

吴迪¹, 钱鹏智¹, 陈勇^{1,2*}

(1. 国防科技大学第63研究所, 江苏南京 210007; 2. 陆军工程大学通信工程学院, 江苏南京 210007)

摘要 针对恶意干扰场景下无人机群动态频谱分配问题, 构建了基于斯坦伯格博弈的动态频谱分配模型, 干扰机为斯坦伯格博弈的领导者, 无人机群为斯坦伯格博弈的跟随者, 设计了不同博弈参与者的效益函数, 并证明了该博弈存在稳定的斯坦伯格均衡解。在此基础上设计了一种分层动态频谱分配算法, 针对领导者采用Q学习选择干扰信道的场景下, 跟随者采用随机学习自动机来确定信道分配策略。仿真结果表明, 所提算法能够得到无人机用户的最优信道分配策略, 有效提升无人机用户的总吞吐量性能, 实现效益最大化。

关键词 动态频谱分配; 无人机; 恶意干扰; 斯坦伯格博弈

中图分类号 TN 929 文献标志码 A 文章编号 2097-163X(2022)01-0073-8

DOI 10.12399/j.issn.2097-163x.2022.01.007

Dynamic spectrum allocation algorithm of UAV based on Stackelberg game in the condition of malicious interference

WU Di¹, QIAN Pengzhi¹, CHEN Yong^{1,2*}

(1. The Sixty-third Research Institute, National University of Defense Technology, Nanjing 210007, China;

2. College of Communication Engineering, Army Engineering University of PLA, Nanjing 210007, China)

Abstract Aiming at the problem of dynamic spectrum allocation of unmanned aerial vehicle (UAV) group in malicious interference scenario, a dynamic spectrum allocation model based on Stackelberg game was constructed. The jammer was the leader of Stackelberg game and the UAV group was the follower of Stackelberg game. The benefit functions of different game participants were designed. It was proved that the game had a stable Stackelberg equilibrium solution. On the basis of this, a hierarchical dynamic spectrum allocation algorithm was designed. When the leader used Q-learning to select the interference channel, the follower used random learning automata (SLA) to determine the channel allocation strategy. Simulation results showed that the proposed algorithm could obtain the optimal channel allocation strategy of UAV users, effectively improved the throughput performance of UAV users and maximized benefits.

Keywords dynamic spectrum allocation; UAV; malicious interference; Stackelberg game

收稿日期: 2022-03-12

修回日期: 2022-03-28

通信作者: 陈勇, E-mail: chy63s@126.com

作者简介: 吴迪(1999—), 男, 硕士研究生, 研究方向为认知无线电、频谱管理; 钱鹏智(1994—), 男, 工程师, 研究方向为认知无线电、频谱管理; 陈勇(1975—), 男, 研究员, 全军学科拔尖人才, 江苏省中青年科技领军人才, 研究方向为无线网络、频谱管理

0 引言

随着高新技术的快速发展,无人机作为一种新型作战平台,以其机动性强、隐蔽性高、制造成本低等特点,逐渐应用到侦察监视、电子对抗、火力打击、通信指挥等任务中^[1]。无人机以大规模机群形式执行任务时往往会出现两方面的问题:一方面,大量无人机同时接入任务区域频段,导致该频段愈发拥挤,一个信道内多用户之间也可能存在干扰,频谱资源短缺问题日渐突出;另一方面,由于无线网络的开放特性,无人机通信面临欺骗、窃听和干扰等威胁,极大地影响了信息传输的安全性、可靠性。因此,研究恶意干扰条件下无人机群动态频谱分配技术对于无人机有效避免外部恶意干扰,实现频谱资源的高效利用具有重要意义。

在面临挑战上,恶意干扰条件下的动态频谱分配技术除了要解决传统动态频谱分配技术中频谱资源紧张、频谱感知受限、频谱环境变化快等问题,还要考虑恶意干扰对信息传输质量的影响;在场景模型上,未来干扰环境中的恶意干扰不仅局限于扫频干扰,而且是一种能够学习用户策略的新型干扰;在理论上,除了传统的动态频谱分配理论,博弈论还为解决干扰条件下的动态频谱分配问题提供了一种新的思路。

针对无人机群动态频谱分配问题,学者们开展了大量的研究工作。文献[2-4]为缓解无人机间共信道干扰,制定了干扰缓解对策,并证明其为精确势能博弈,但是没有考虑无人机受外部恶意干扰的情况。文献[5]提出了一种基于马尔可夫决策过程的抗干扰方案,并给出了一种对抗干扰攻击的信道选择策略,但是该研究中没有关注用户之间对有限频谱资源的相互竞争以及用户信道选择与外部干扰攻击之间的时序行为。文献[6]采用马尔可夫博弈对抗干扰问题进行建模和分析,提出了一种协同多智能体抗干扰算法,以获得最优的抗干扰策略,但仅考虑了扫频干扰,不具备智能决策能力。文献[7]研究了干扰对抗场景下频谱接入优化问题,提出基于分层博弈的自主信道选择机制,设计一种基于Q学习的分层强化学习算法来求解分层博弈的均衡解,但没有考虑存在共信道干扰的情况。文献[8]将抗干扰描述为一个分层博弈问题,提出了分层学习算法,考虑了用户之间的恶意干扰和共信道干扰,

然而对于效益函数的设计和用户共信道干扰时存在的干扰门限问题还有待改进。鉴于当前无人机群对抗中面临的挑战和研究的不足,本文提出了一种基于斯坦伯格博弈的分层动态频谱分配算法,以期有效降低或者避免干扰机对无人机群的影响。与现有研究相比,本文的特点如下:

(1) 将干扰机作为博弈的领导者,无人机用户作为博弈的跟随者,构建了基于斯坦伯格博弈的动态频谱分配模型,设计以吞吐量为目标的效益函数,证明斯坦伯格均衡(Stackelberg equilibrium, SE)解的存在性。

(2) 在无人机群层级中引入局部互利模型,证明了在干扰机干扰策略给定时,下层子博弈同样是一个精确势能博弈(exact potential game, EPG),即至少存在一个纳什均衡(Nash equilibrium, NE)解。同时,证明了本文所构建的博弈将个体优化与全局优化联系起来,单个无人机用户效益的提升引起博弈势能函数的提升,即无人机群吞吐量的提升。

(3) 对于所构建的博弈模型求解,领导者采用Q学习信道选择算法,跟随者采用随机学习自动机(stochastic learning automata, SLA)信道选择算法,并分析本文算法的收敛性。该算法使干扰机和无人机用户根据历史经验和当下环境信息调整最佳信道分配策略,最终达到稳定均衡解。

1 系统模型与问题构建

1.1 无人机网络模型

图1所示为恶意干扰条件下无人机网络的动态频谱分配模型,该模型中,干扰机能够根据用户的策略和动态环境自适应地调整干扰策略,使其干扰效果最大化。假设在网络覆盖范围内有 N 个无人机用户对(发射端—接收端),其集合为 $N = \{1, 2, \dots, N\}$,为表述方便在本文中使用无人机用户 n 代表无人机用户对 n 的收发端; M 个可用信道,其集合为 $M = \{1, 2, \dots, M\}$,且 $M < N$ 。在网络边缘处有一干扰机J工作在时分多址模式,即每一时隙选择1个信道以固定功率进行干扰。假设无人机1个时隙只能接入1个信道,每个信道同一时隙可以分配给多个无人机。

1.2 优化问题构建

恶意干扰条件下信道模型示意图如图2所示。图2中, $h_{n,n}$ 表示无人机用户 n ($n = 1, \dots,$

N)发射端到无人机用户 n 接收端的信道增益; $h_{n,i}$ 表示无人机用户 n 发射端到无人机用户 i ($i=1, \dots, N$)接收端的信道增益; $h_{i,n}$ 表示无人机用户 i 发射端到无人机用户 n 接收端的信道增益; $h_{i,i}$ 表示无人机用户 i 发射端到无人机用户 i 接收端的信道增益; $h_{j,n}$ 表示干扰机 J 到无人机用户 n 接收端的信道增益; $h_{j,i}$ 表示干扰机 J 到无人机用户 i 接收端的信道增益。信道增益表示如下:

$$h_{x,y} = \beta_0 (d_{x,y})^{-\alpha} \quad (1)$$

其中, $d_{x,y}$ 为设备之间的距离; α 为路径损耗指数; β_0 为参考距离 $d_0=1$ m 处的信道功率增益。

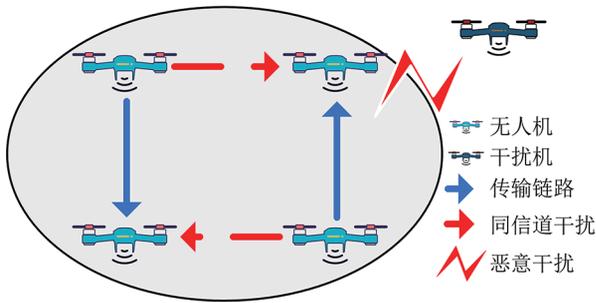


图1 恶意干扰条件下无人机网络模型示意图

Fig. 1 Schematic diagram of network model of UAV in the condition of malicious interference

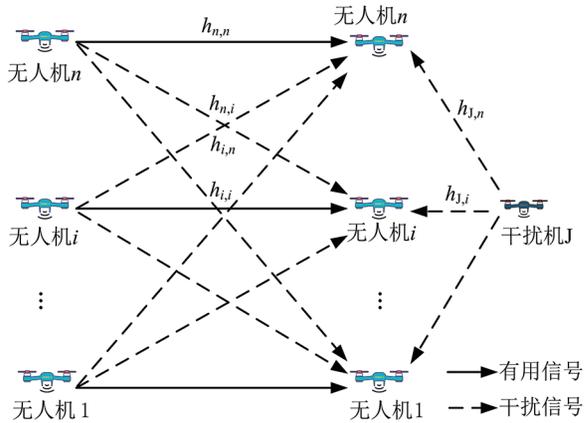


图2 恶意干扰条件下信道模型示意图

Fig. 2 Schematic diagram of channel model in the condition of malicious interference

无人机用户 n 的传输速率为:

$$R_n = B \log_2 \left(1 + \frac{P_n h_{n,n}}{I_{U,n} + I_{J,n} + N_0} \right) \quad (2)$$

其中, B 为信道带宽; N_0 为背景噪声功率; P_n 为无人机用户 n 的发射功率; $I_{U,n}$ 为无人机用户之间的干扰; $I_{J,n}$ 为干扰机对无人机用户干扰, 2 种干扰量表示如下:

$$I_{U,n} = \sum_{i \in \Gamma_n} P_i h_{i,n} \delta(a_i, a_n) \quad (3)$$

$$I_{J,n} = P_j h_{j,n} \delta(c_j, a_n) \quad (4)$$

式中, a_n 为无人机用户 n 选择的信道; c_j 为干扰机选择的信道; Γ_n 为无人机用户 n 的邻居集, 可以定义为:

$$\Gamma_n = \{i \in N, d_{i,n} < d_{th}\} \quad (5)$$

式中, d_{th} 为无人机通信阈值, 即只有在通信阈值内的无人机才会对其造成共信道干扰。 $\delta(x, y)$ 为二元指示函数, 用来判断二者在当前时刻是否选取相同信道, 其定义为:

$$\delta(x, y) = \begin{cases} 1, & x = y \\ 0, & x \neq y \end{cases} \quad (6)$$

综上所述可以发现, 无人机用户的效益除了和自己的决策有关, 还受干扰机决策的影响。因此, 这里将该问题优化为一个竞争博弈问题求解。对于无人机用户来说需要优化信道分配策略 $a = \{a_1, a_2, \dots, a_N\}$ 以最小化恶意干扰影响, 获得最大传输速率; 对于干扰机来说, 需要优化干扰信道分配策略 c_j 以最大化其干扰破坏效果, 符合斯坦伯格博弈模型。因此, 可以通过找到一个斯坦伯格最优解, 来解决此问题。

2 信道选择博弈

2.1 博弈模型

本文将干扰机作为领导者、无人机作为跟随者, 构建斯坦伯格博弈模型为 $G = \{J, N, \{C_j, A_n\}, \{U_j, U_n\}\}$, 其中 J 和 $N = \{1, 2, \dots, N\}$ 是博弈的参与者, 分别表示为干扰机(领导者)和无人机用户(跟随者); C_j 和 A_n 分别为干扰机和无人机用户的策略分配集; U_j 和 U_n 分别为干扰机和无人机用户的效益函数。

从无人机用户的角度出发, 每个用户独立自主地选择信道, 目的是最大化其传输速率, 其通信效益可以表示为:

$$W_n = R_n \quad (7)$$

因为无人机用户的最佳收益受到其共信道无人机用户的影响, 所以每个无人机都应考虑自身和其邻居的效益, 以寻求更大的效益。受文献[9]启发, 在这里引入局部互利模型来表示每个无人机用户的博弈效益:

$$U_n = W_n + \sum_{i \in \Gamma_n} W_i \quad (8)$$

从干扰机的角度出发, 其目的是对无人机用户实现最大程度的干扰, 使其总吞吐量最小, 效益函数可以表示为:

$$U_J = - \sum_{n \in N} R_n \quad (9)$$

综上所述,斯坦伯格博弈中领导者(上层)和跟随者(下层)的优化问题分别表示为:

$$\begin{cases} C_1: \max_{c_J} U_J(c_J, \partial) \\ C_2: \max_{a_n} U_n(c_J, a_n, \partial_{-n}) \end{cases} \quad (10)$$

上层子博弈和下层子博弈可以分别定义为:

$$\begin{cases} G_t = \{J, C_J, U_J(c_J, \partial)\} \\ G_d = \{N, A_n, U_n(c_J, a_n, \partial_{-n})\} \end{cases} \quad (11)$$

式中, $\partial = (a_1, a_2, \dots, a_N)$; $\partial_{-n} = (a_1, a_2, \dots, a_{n-1}, a_{n+1}, \dots, a_N)$; C_J, A_n 与可用信道集 M 相同。

2.2 斯坦伯格博弈求解

为了证明下层子博弈存在稳定的 NE 解,这里引入精确势能博弈 EPG 的概念,即通过证明在给定干扰机干扰策略的情况下,下层子博弈同样是一个精确势能博弈。

定义 1 博弈是一个精确势能博弈^[10],当且仅当其存在一个势能函数满足:

$$U_n(a'_n, \partial_{-n}) - U_n(a_n, \partial_{-n}) = \Phi(a'_n, \partial_{-n}) - \Phi(a_n, \partial_{-n}) \quad (12)$$

式中, a'_n 为无人机用户 n 改变其信道分配策略,并且根据文献[8]有引理 1 和定理 1。

引理 1 任意一个精确势能博弈都至少具有一个纯策略 NE 解。

定理 1 给定干扰机的信道分配策略 c_J 后,下层子博弈为一个精确势能博弈,并且存在稳定 NE 解。

证明 根据文献[8],构建势能函数为:

$$\Phi(c_J, a_n, \partial_{-n}) = \sum_{n \in N} W_n(c_J, a_n, \partial_{-n}) \quad (13)$$

当无人机用户 n 单方面改变其信道分配策略:

$$\begin{aligned} & \Phi(c_J, a'_n, \partial_{-n}) - \Phi(c_J, a_n, \partial_{-n}) = \\ & [W_n(c_J, a'_n, \partial_{-n}) - W_n(c_J, a_n, \partial_{-n})] + \\ & \left[\sum_{i \in \Gamma_n} W_i(c_J, a_i, \partial'_{-i}) - \sum_{i \in \Gamma_n} W_i(c_J, a_i, \partial_{-i}) \right] + \\ & \left[\sum_{i \in \{N \setminus (n \cup \Gamma_n)\}} W_i(c_J, a_i, \partial'_{-i}) - \sum_{i \in \{N \setminus (n \cup \Gamma_n)\}} W_i(c_J, a_i, \partial_{-i}) \right] \end{aligned} \quad (14)$$

又因为

$$\begin{aligned} & \sum_{i \in \{N \setminus (n \cup \Gamma_n)\}} W_i(c_J, a_i, \partial'_{-i}) - \\ & \sum_{i \in \{N \setminus (n \cup \Gamma_n)\}} W_i(c_J, a_i, \partial_{-i}) = 0 \end{aligned} \quad (15)$$

所以

$$\begin{aligned} & \Phi(c_J, a'_n, \partial_{-n}) - \Phi(c_J, a_n, \partial_{-n}) = \\ & [W_n(c_J, a'_n, \partial_{-n}) - W_n(c_J, a_n, \partial_{-n})] + \\ & \left[\sum_{i \in \Gamma_n} W_i(c_J, a_i, \partial'_{-i}) - \sum_{i \in \Gamma_n} W_i(c_J, a_i, \partial_{-i}) \right] = \\ & [W_n(c_J, a'_n, \partial_{-n}) + \sum_{i \in \Gamma_n} W_i(c_J, a_i, \partial'_{-i})] - \\ & [W_n(c_J, a_n, \partial_{-n}) + \sum_{i \in \Gamma_n} W_i(c_J, a_i, \partial_{-i})] = \\ & U_n(c_J, a'_n, \partial_{-n}) - U_n(c_J, a_n, \partial_{-n}) \end{aligned} \quad (16)$$

因此证得下层子博弈同样是一个精确势能博弈,并且至少存在一个纯策略 NE 点,即没有无人机可以通过单方面改变动作来提升其效益,称此时的信道分配策略 ∂^* 为无人机用户 n 的最佳信道分配策略,即:

$$U_n(c_J, \partial^*) \geq U_n(c_J, \partial) \quad (17)$$

因为在精确势能博弈中,任意用户单方面改变策略所引起的个体效益函数变化与博弈势能函数的变化是相同的,即将个体优化与全局优化联系起来,根据式(13)所构建的势能函数表达式,其物理意义恰好为无人机用户的总吞吐量,所以当无人机用户 n 改变策略提高效益后,博弈的势能函数增加,无人机用户的总吞吐量得到提升。

接下来分析所构建的斯坦伯格博弈存在 SE 解。根据文献[11]的论述,每个有限策略博弈均存在一个混合策略均衡。因此,上层子博弈也存在平稳意义上的均衡。

基于上述分析,假设干扰机的混合策略为 ϵ_J 。

定义 2 策略组合 $(\epsilon_J^*, \partial^*)$ 称为该博弈的 SE 解,当且仅当 ϵ_J^* 是使干扰机效益最大的信道分配策略, ∂^* 是使无人机效益最大的信道分配策略,满足:

$$\begin{cases} U_J(\epsilon_J^*, \partial^*) \geq U_J(\epsilon_J, \partial^*) \\ U_n(\epsilon_J^*, a_n^*, \partial_{-n}^*) \geq U_n(\epsilon_J^*, a_n, \partial_{-n}^*) \end{cases} \quad (18)$$

定理 2 在所构造的斯坦伯格博弈中,存在一个由于干扰机平稳策略和无人机用户 NE 策略所构成的 SE 解。

证明 由上可知,当给定干扰机信道分配策略 ϵ_J ,斯坦伯格博弈 G 变成一个非合作博弈。基于定理 1,下层子博弈是一个精确势能博弈,且至少存在一个纯策略 NE 解 (ϵ_J, ∂^*) 。

因此,给定干扰策略 ϵ_J ,在斯坦伯格博弈 G

中总是存在 $E_{NE}(\partial^*(\epsilon_j))$, 于是干扰机的平稳策略表示为 $\epsilon_j^* = \arg \max_{\epsilon_j} U_j(\epsilon_j, E_{NE}(\partial^*(\epsilon_j)))$, 根据文献[11]的论述, 每个有限策略博弈都有一个混合策略均衡, 因此 $(\epsilon_j^*, E_{NE}(\partial^*(\epsilon_j^*)))$ 构成该问题平稳意义上的 SE。

2.3 基于斯坦伯格博弈的分层动态频谱分配算法

为了获得所构造斯坦伯格博弈的均衡解, 提出一种分层动态频谱分配算法, 领导者采用 Q 学习信道选择算法, 跟随者采用 SLA 信道选择算法。SLA^[12] 是一种概率选择算法, 通常用于信道分配。该算法不需要进行信息交互, 每个用户根据获得的回报进行动态策略更新, 即根据所选动作的效用改变所有动作的选择概率。例如, 当一个无人机用户 n 通过动作 a_n 提高其效益时, SLA 会增加该动作的选择概率, 同时降低该用户其他动作的选择概率。

本文算法中, 干扰机和无人机分别以时间尺度 t 和 τ 更新信道分配策略, 即干扰机 J 在时期 t 的信道分配概率为:

$$p_J(t) = \{p_{J,1}(t), p_{J,2}(t), \dots, p_{J,M}(t)\} \quad (19)$$

式中, $p_{J,i}(t)$ 表示干扰机在时期 t 选择信道 i 进行干扰的概率, 并且满足 $\sum_{i \in M} p_{J,i}(t) = 1$ 。无人机用户 n 在时隙 τ 的信道分配概率为:

$$p_n(\tau) = \{p_{n,1}(\tau), p_{n,2}(\tau), \dots, p_{n,M}(\tau)\} \quad (20)$$

式中, $p_{n,i}(\tau)$ 表示无人机用户 n 在时隙 τ 选择信道 i 的概率, 并且满足 $\sum_{i \in M} p_{n,i}(\tau) = 1$ 。

其算法流程如图 3 所示。

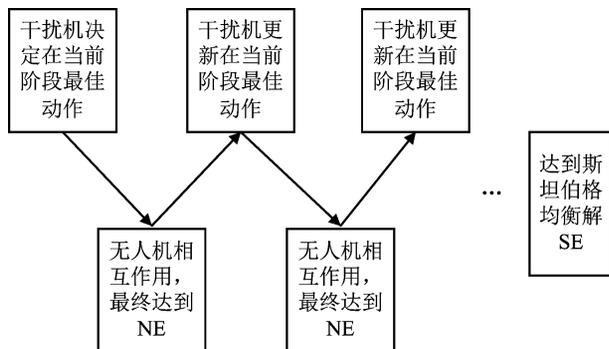


图 3 斯坦伯格博弈流程图

Fig. 3 The flow chart of Stackelberg game

干扰机选择在当前周期要攻击的信道, 然后选择无人机学习环境, 直到其子博弈收敛到 NE 解; 重复该过程, 直到达到最大迭代次数或相邻

时隙的信道选择策略相同。多次迭代后, 上层子博弈和下层子博弈均收敛到 NE 解, 即 SE 解。

分层动态频谱分配算法的具体流程如下:

(1) 初始化。

设定 $t=0$ 和 $\tau=0$, 同时初始化干扰机和无人机的信道分配概率 $p_J(0)$ 和 $p_n(0)$ 。

(2) 循环。

步骤 1 在 t 时期, 干扰机根据当前时刻的信道分配概率 $p_J(t)$ 进行信道初分配。

步骤 2 无人机更新决策动作。

① 在第 τ 个时隙, 无人机根据当前时刻的信道分配概率 $p_n(\tau)$ 为其分配信道 $a_n(\tau)$;

② 每架无人机计算其效用 $U_n(\tau)$;

③ 无人机按照以下规则进行信道分配概率更新:

$$\begin{cases} p_{n,i}(\tau+1) = p_{n,i}(\tau) + \beta \frac{U_n(\tau)(1-p_{n,i}(\tau))}{U_{\max}}, \\ i = a_n(\tau) \\ p_{n,i}(\tau+1) = p_{n,i}(\tau) - \beta \frac{U_n(\tau)p_{n,i}(\tau)}{U_{\max}}, \\ i \neq a_n(\tau) \end{cases} \quad (21)$$

式中, β 为学习步长; U_{\max} 是无干扰情形下的无人机用户的效用;

④ 如果无人机状态稳定, 则进入步骤 3, 否则返回步骤 2。

步骤 3 干扰机决策更新过程。

① 干扰机计算当前时刻效用;

② 干扰机更新 Q 值, 即:

$$Q_0^{t+1}(i) = (1 - \alpha_t)Q_0^t(i) + \alpha_t U_J(t) \quad (22)$$

式中, α_t 表示学习速率, 满足:

$$\sum_{t=0}^{\infty} \alpha_t = \infty, \sum_{t=0}^{\infty} \alpha_t^2 < \infty \quad (23)$$

③ 干扰机根据玻尔兹曼分布规律更新其信道分配概率, 即:

$$p_{J,i}(t) = \frac{\exp[Q_0^t(i)/\tau_0]}{\sum_M \exp[Q_0^t(i)/\tau_0]} \quad (24)$$

式中, τ_0 为温度系数, 用于平衡决策空间的探索和利用, 具体来说 $\tau_0 \rightarrow 0$ 时, 干扰机倾向于选择 Q 值大的策略, 而 $\tau_0 \rightarrow \infty$ 时, 干扰机的信道分配策略是完全随机的。

(3) 结束。

当干扰机和无人机的效用保持不变或者达

到最大迭代次数,则循环结束。

根据文献[5],当学习步长 $\beta \rightarrow 0$ 时,下层子博弈收敛到一个纯 NE 点;根据文献[8],当上层 Q 学习满足式(21)时,上层子博弈也将收敛到最优解,即本文所提算法将收敛于斯坦伯格博弈的一个 SE 解。

3 仿真结果分析

仿真参数的设置如表 1 所示。

表 1 恶意干扰场景下系统仿真参数

Tab. 1 System simulation parameters under the condition of malicious interference scenario

参数	数值
无人机用户分布	400×400 m
无人机用户数 N /对	8
可用信道数 M /个	3
信道带宽 B /MHz	6
干扰机功率 P_j/W	5
无人机功率 P_n/W	1
无人机通信阈值 d_{th}/m	200
无人机间距离 $d_{n,n}/m$	50
背景噪声功率 N_0/dBm	-90
单位距离信道功率增益 β_0/dB	-30
路径损耗指数 α	2

设置 Q 学习算法的学习速率 $\alpha_t = \frac{1}{t}$, 初始温度 $\tau_0 = 10^7$, $\tau_i = 10^7 \exp(-i)$, 减小至终止温度 $\tau_0 = 1$ 后停止递减; SLA 算法学习步长 $\beta = 0.15$ 。

首先,给出 1 次仿真中无人机用户和干扰机的信道分配概率随时隙/期数的变化过程,仿真结果如图 4、图 5 所示。在初始时隙 $\tau = 0$, 系统以相同的概率 ($p_{2,1}(0) = p_{2,2}(0) = p_{2,3}(0) = \frac{1}{3}$) 为无人机随机分配信道,在经过大约 200 次迭代后,无人机用户分配信道 1 的概率收敛到 1,其他概率变为 0,说明系统在第 1 个时期结束时为无人机用户 2 分配信道 1 接入;对于干扰机来说,在初始时期 $t = 0$,以相同的概率随机分配信道进行干扰,经过大约 17 次迭代后达到收敛状态,最终选择了信道 3 进行攻击。无人机用户收敛迭代次数远大于干扰机,这是因为无人机用户作为跟随者

在干扰机的策略基础上进行决策,因此其获得 NE 解所需的迭代次数更多。

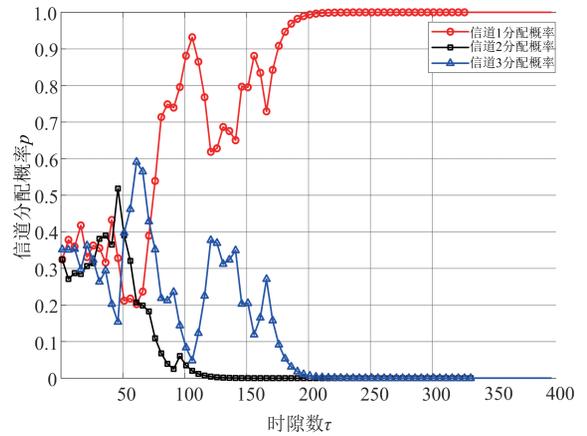


图 4 无人机用户 2 第 1 个时期内信道分配概率变化
Fig. 4 The change of channel allocation probability of UAV user 2 in the first period

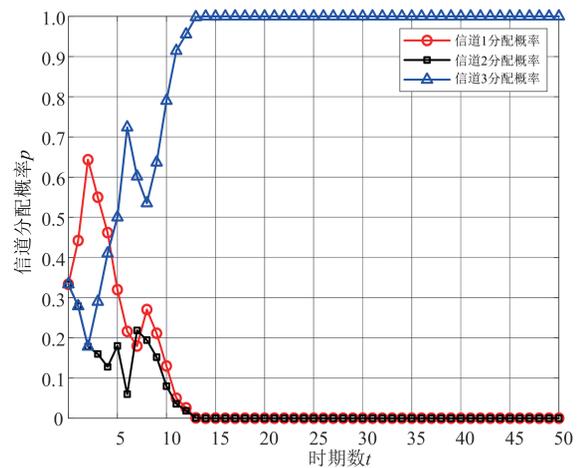


图 5 干扰机信道分配概率变化
Fig. 5 The change of jammer channel allocation probability

无人机用户总收益随无人机功率和干扰机功率的变化情况,如图 6、图 7 所示。图 6 中,无人机用户的总收益随着无人机用户功率的增加而增加,但是随着无人机用户功率到达一定程度后其收益增加会减缓,因为无人机用户可能存在共信道干扰,当功率较大时,共信道干扰增加,导致无人机的总收益提升较缓;图 7 中,随着干扰机功率的增加,无人机用户的总收益下降。显然,干扰机功率增加,无人机所受恶意干扰影响程度增加,并且可以发现,无人机功率越大,干扰对无人机用户总收益的影响越显著。

为了评估本文所提算法的性能,将其与随机分配算法和无干扰算法进行比较。在随机分配算法中,干扰机和无人机随机分配信道接入,然

后通过模拟 500 次实验取平均值得到。无干扰算法,即不存在干扰机,下层无人机用户采取 SLA 算法选择信道接入。仿真结果如图 8 所示,本文所提算法性能优于随机分配算法,差于无干扰情况性能。这是由于本文所提算法对无人机用户进行合理的信道分配,最大程度减少恶意干扰及共信道干扰的影响,而随机分配算法没有优化机制来进行资源的协调和分配,用户极大概率占用同一条信道进行传输,故其性能最差。还可以发现,随着信道数的增加,本文所提算法性能越接近于无干扰情况性能。但是当信道数达到 5 时,无人机用户的总收益并不会再随信道数的增大有较大提升,这主要与本文所设参数(如通信距离、无人机相对位置等)有关。

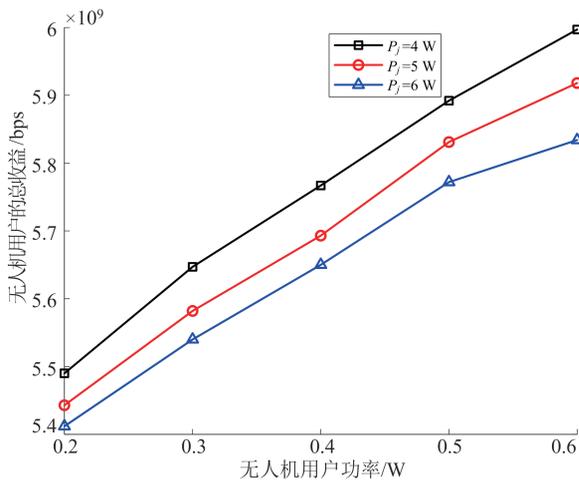


图 6 无人机用户总收益随无人机功率变化

Fig. 6 The total revenue of UAV users with the change of UAV power

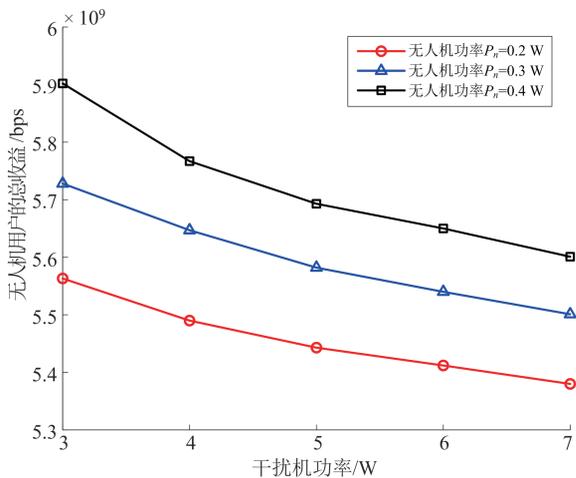


图 7 无人机用户总收益随干扰机功率变化

Fig. 7 The total revenue of UAV users with the change of the power of jammer

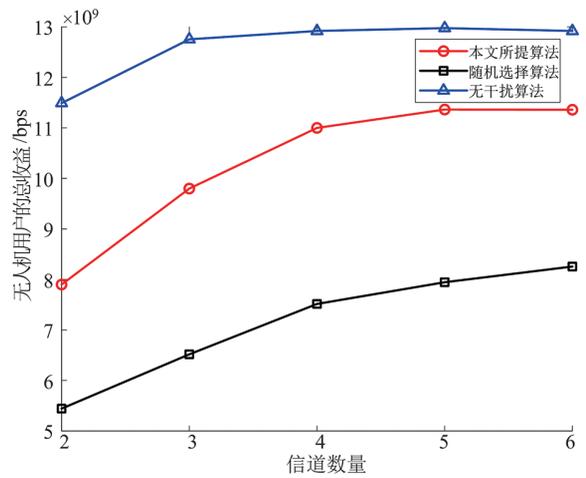


图 8 不同算法性能比较

Fig. 8 Performance comparison of different algorithms

4 结束语

无人机通信中如何有效避免外部恶意干扰,实现频谱资源的高效利用是目前面临的重要挑战。因此,本文主要研究了恶意干扰条件下的无人机群动态频谱分配问题,首先,针对无人机群受恶意干扰的场景,构建了以最大吞吐量效益函数的分层博弈模型。证明其存在 SE 解后提出了分层动态频谱分配算法,该分配算法中领导者采用 Q 学习信道选择算法,跟随者采用 SLA 信道选择算法。仿真结果表明,所提算法可以有效避免外部恶意干扰,实现无人机效益最大化,提升用户的总吞吐量性能。

参 考 文 献

- [1] GUPTA L, JAIN R, VASZKUN G. Survey of important issues in UAV communication networks [J]. IEEE Communications Surveys & Tutorials, 2015, 18 (2): 1123-1152.
- [2] WU Q, XU Y, WANG J, et al. Distributed channel selection in time-varying radio environment: interference mitigation game with uncoupled stochastic learning[J]. IEEE Transactions on Vehicular Technology, 2013, 62(9): 4524-4538.
- [3] ZHENG J, CAI Y, YANG W, et al. A fully distributed algorithm for dynamic channel adaptation in canonical communication networks[J]. IEEE Wireless Communications Letters, 2013, 2(5): 491-494.
- [4] ZHENG J, CAI Y, XU Y, et al. Distributed channel selection for interference mitigation in dynamic environment: a game-theoretic stochastic learning solution

- [J]. *IEEE Transactions on Vehicular Technology*, 2014, 63(9): 4757-4762.
- [5] CHEN C, SONG M, XIN C S, et al. A game-theoretical anti-jamming scheme for cognitive radio networks[J]. *IEEE Network*, 2013, 27(3): 22-27.
- [6] YAO F, JIA L. A collaborative multi-agent reinforcement learning anti-jamming algorithm in wireless networks[J]. *IEEE Wireless Communications Letters*, 2019, 8(4): 1024-1027.
- [7] 范超琼,赵成林,李斌. 无人机网络中基于分层博弈的干扰对抗频谱接入优化[J]. *通信学报*, 2020, 41(6): 26-33.
FAN Chaoqiong, ZHAO Chenglin, LI Bin. Spectrum access optimization of interference countermeasure based on hierarchical game in UAV network[J]. *Journal of Communication*, 2020, 41(6): 26-33. (in Chinese)
- [8] YAO F, JIA L, SUN Y, et al. A hierarchical learning approach to anti-jamming channel selection strategies [J]. *Wireless Networks*, 2019, 25(1): 201-213.
- [9] XU Y H, WANG J L, WU Q H, et al. Opportunistic spectrum access in cognitive radio networks: global optimization using local interaction games[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2012, 6(2):180-194.
- [10] MONDERER D, SHAPLEY L S. Potential games [J]. *Games and Economic Behavior*, 1996, 14(1): 124-143.
- [11] HAN Z. Game theory in wireless and communication networks[M]. Britain: Cambridge University Press, 2012.
- [12] XU Y H, XU Y T, ANPALAGAN A. Database-assisted spectrum access in dynamic networks: a distributed learning solution[J]. *IEEE Access*, 2015(3): 1071-1078.

责任编辑 钱 静