

# 无人机自组网 SPMA 协议智能退避技术:DDQN 驱动的多维决策

王海军<sup>1,2</sup>,王 洁<sup>2</sup>,张 姣<sup>2</sup>,赵海涛<sup>2\*</sup>,任保全<sup>1</sup>,尹 浩<sup>1</sup>

(1. 军事科学院 系统工程研究院,北京 100082; 2. 国防科技大学 电子科学学院,湖南 长沙 410073)

**摘要:**现有基于统计优先级的多址接入(statistic priority-based multiple access, SPMA)协议退避机制依赖静态函数模型且优化参数维度单一,导致无法适应无人机自组网的动态传输和多优先级需求。为此,将 SPMA 协议中节点选择退避时间的动态决策过程建模为马尔可夫决策过程,创新提出了基于双重深度  $Q$  网络(double deep  $Q$ -network, DDQN)算法的 SPMA 协议智能退避策略。该策略综合考虑业务优先级、阈值和信道负载等因素,使用 DDQN 算法在有限、离散的动作空间中选择退避时间。仿真结果表明,相比传统二进制指数退避策略和基于对数函数的退避策略,所提策略对低优先级业务的传输时延最大可降低 33.3%、首次退避成功率可提升 18%,有效提高传输成功率并能适应网络规模的变化。

**关键词:**无人机自组网;基于统计优先级的多址接入协议;退避机制;多优先级业务;双重深度  $Q$  网络  
**中图分类号:**TN929.5 **文献标志码:**A **文章编号:**1001-2486(2026)03-096-11

## Intelligent backoff technology of SPMA protocol in UAV ad hoc networks: multi-dimensional decision driven by DDQN

WANG Haijun<sup>1,2</sup>, WANG Jie<sup>2</sup>, ZHANG Jiao<sup>2</sup>, ZHAO Haitao<sup>2\*</sup>, REN Baoquan<sup>1</sup>, YIN Hao<sup>1</sup>

(1. System Engineering Research Institute, Academy of Military Sciences, Beijing 100082, China;

2. College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China)

**Abstract:** The existing backoff mechanism of the SPMA (statistic priority-based multiple access) protocol relies on static function models and has single dimension of optimization parameters, making it unable to adapt to dynamic transmission and multi-priority requirements in UAV ad hoc networks. To address this issue, the dynamic decision-making process of node selection for backoff time in the SPMA protocol was modeled as a Markov decision process, and an intelligent backoff strategy based on the DDQN (double deep  $Q$ -network) was innovatively proposed. This strategy comprehensively considers factors such as service priority, thresholds, and channel load, and adopts the DDQN algorithm to select backoff time within a finite and discrete action space. Simulation results show that, compared to traditional binary exponential backoff strategies and logarithmic function-based backoff strategies, the proposed strategy can reduce the transmission delay for low-priority services by up to 33.3%, increase the initial backoff success rate by 18%, and effectively improve the transmission success rate and adapt to the variation of network scale well.

**Keywords:** unmanned aerial vehicles ad hoc network; statistic priority-based multiple access protocol; backoff mechanism; multi-priority service; double deep  $Q$ -network

近年来,无人机自组网凭借其快速部署、成本低廉和协同任务能力,在灾后救援、战术侦察、智慧农业等民用和军事场景展现出巨大应用潜力<sup>[1]</sup>。然而在上述场景中,无人机自组网因动态拓扑、时敏业务与资源受限等特性对媒体接入控制(media access control, MAC)协议提出了较高

要求,即既要保障高优先级业务的时效性,又要适应信道状态快速波动,这给传统竞争类 MAC 协议带来严峻挑战<sup>[2]</sup>。

一般来说,无人机自组网承载着多种类型的业务,如传感业务、控制业务、载荷业务、指控业务等,不同类型业务的优先级和服务质量(quality of

收稿日期:2025-07-29

基金项目:国家资助博士后研究人员计划和中国博士后科学基金资助项目(BX20240493);国家自然科学基金资助项目(61931020, 62201584, 62171449, 62371462)

第一作者:王海军(1993—),男,安徽淮北市人,副教授,博士,硕士生导师, E-mail: haijunwang14@nudt.edu.cn

\*通信作者:赵海涛(1981—),男,山东昌乐人,教授,博士,博士生导师, E-mail: haitaozhao@nudt.edu.cn

引用格式:王海军,王洁,张姣,等. 无人机自组网 SPMA 协议智能退避技术:DDQN 驱动的多维决策[J]. 国防科技大学学报, 2026, 48(3): 96-106.

Citation: WANG H J, WANG J, ZHANG J, et al. Intelligent backoff technology of SPMA protocol in UAV ad hoc networks: multi-dimensional decision driven by DDQN[J]. Journal of National University of Defense Technology, 2026, 48(3): 96-106.

service, QoS)需求具有差异性<sup>[3-4]</sup>。基于统计优先级的多址接入(statistic priority-based multiple access, SPMA)协议因其独特的业务优先级区分机制和信道负载统计特性,成为适配无人网络的新一代 MAC 协议<sup>[5]</sup>。该协议通过设置优先级阈值动态调控信道接入,使高优先级业务始终获得优先传输保障,相比传统随机竞争类 MAC 协议显著降低了信道竞争开销<sup>[6]</sup>。研究表明,SPMA 协议性能高度依赖于退避时间、统计窗口及优先级阈值三个参数的协同设计。其中,退避时间是提高 SPMA 协议效率的关键,它关系到节点接入信道前所需要的等待时间,直接影响着数据包时延和网络吞吐量性能:退避时间过短会导致接入冲突概率增大,引发频繁重传,产生不必要的信道资源消耗;退避时间过长会导致信道资源闲置,同时增加数据时延和降低吞吐量。因此,通过合理的退避策略来选择最优的退避时间是提高数据端到端时延性能的关键。同时,无人机自组网流量的随机性会导致信道负载与优先级阈值的大小关系不断变化,也就是信道状态在满足与不满足发送条件之间不断波动。因此退避时间决定着数据包能否恰好在网络状态健康时进行发送,对传输成功率也有较大影响。综上,高效的退避策略对提高 SPMA 协议的时延性能和传输成功率具有重要意义。

在 MAC 协议中,退避机制与发送机制互相作用决定着数据包的接入。这意味着,退避机制的设计只有契合 MAC 协议特点,才能最优化网络性能。现有成熟退避机制,如二进制指数退避(binary exponential backoff, BEB)及其改进策略,在传统通信网络中有着良好性能表现<sup>[7]</sup>。但随着研究的深入,发现 SPMA 协议中的节点基于传统退避规则很难获得恰当的退避时间。这是因为 SPMA 协议与传统基于竞争的 MAC 协议判断退避开始和结束的条件以及是否区分优先级均有所不同<sup>[5]</sup>。此外,无人机自组网存在拓扑动态变化、业务异构性强、节点负载不均、资源受限等特性,导致数据传输的过程较为复杂,对退避策略在适应瞬变环境、满足差异化业务需求、快速收敛、低开销等方面提出了特殊要求。因此,如何设计无人机自组网场景下 SPMA 协议的高效退避机制成为挑战。为此,文献[8]提出了一种时空数学模型来评估退避方案对 SPMA 协议性能的影响,首先对退避过程进行离散时间马尔可夫建模,然后将此模型集成到随机几何框架中,并分析了多址接入概率、突发成功概率和优先级分组成功概

率等性能指标。文献[9]针对 SPMA 协议区分业务优先级的特点,提出了一种基于优先级的退避窗口计算方式。在此基础上,文献[10]首先分析了该退避算法的局限性,然后进一步考虑信道负载值和优先级阈值因素,同时对比分析了指数模型和对数模型的退避时间计算关系式在 SPMA 系统中的准确性,最后提出了一种基于信道占用及优先级的自适应退避算法。文献[11]提出了一种细粒度信道接入机制,通过细化同一优先级不同数据的退避,可以使整个网络保持更高的信道利用率。然而,现有关于 SPMA 协议退避策略的研究大多将退避窗口与影响参数的关系定义为一个特定初等数学模型,这在动态变化的无人机自组网中显然并不是最优的。

近年来,快速发展的机器学习为解决通信场景里的各类问题提供了新思路,尤其是机器学习中的强化学习被广泛用于解决分布式路由、资源分配、信道接入等动态决策问题<sup>[12-15]</sup>。无人机自组网节点通过随机退避竞争信道的过程正是一个需要对网络状态进行自适应的动态决策过程,十分契合强化学习解决问题的初衷。文献[16-17]考虑了信道占用率的变化趋势,结合深度强化学习构建了 SPMA 协议的退避模型,提出了全联通网络下基于深度强化学习的智能退避策略。然而,总体来说,目前将机器学习应用到 SPMA 协议退避策略的研究较少,或者现有基于机器学习的退避策略没有针对无人机自组网的复杂性和动态组网需求做专门的优化<sup>[16-17]</sup>。基于此,本文引入深度强化学习技术并结合 SPMA 协议特点和无人机自组网的需求,提出了一种基于深度强化学习的智能退避策略。该策略能够保证每个无人机节点获取到不同优先级数据的准确退避时间,提高了网络运行效率。主要创新点包括:

1) 将 SPMA 协议中节点选择退避时间的动态决策过程抽象为马尔可夫决策过程(Markov decision process, MDP),建立“环境—状态—动作—奖励”动态映射关系,克服了传统 BEB 策略依赖静态函数模型的局限性。其中,MDP 是一种描述动态决策的数学模型,其核心假设是未来状态和奖励只取决于当前状态和选择的动作,与之前的历史状态和动作无关。

2) 提出了基于双重深度  $Q$  网络(double deep  $Q$ -network, DDQN)算法的智能退避策略,该策略综合考虑业务优先级、阈值和信道负载等因素,实现在有限、离散的动作空间中选择退避时间。

3) 仿真结果表明,该策略能够有效减少退避

次数和时延,同时提高传输成功率并能适应网络规模的变化,有望为无人机自组网任务关键型通信提供低时延、高可靠的新一代 MAC 层解决方案。

## 1 无人机自组网系统模型

考虑如图 1 所示的单跳全联通无人机自组网,包含  $N$  个无人机节点,所有无人机节点地位平等,任意节点间的数据传输均单跳可达,无须中继节点转发。设从频率维度划分出  $N_f$  ( $N_f < N$ ) 个正交信道,每个信道具有相同的带宽与接入条件,所有节点共享这些有限信道资源。节点数据业务的产生服从参数为  $\lambda$  (数据包/s) 的泊松分布,进一步将这些数据分为  $p$  个优先级,其中 0 表示最高优先级,  $p-1$  表示最低优先级。每种优先级数据的产生同样服从泊松过程,速率分别为  $\lambda_0, \lambda_1, \dots, \lambda_{p-1}$  (数据包/s),满足  $\lambda = \lambda_0 + \lambda_1 + \dots + \lambda_{p-1}$ 。网络中所有节点具有相同的传输速率,所有数据包的大小和通信链路的衰落、干扰情况也相同。

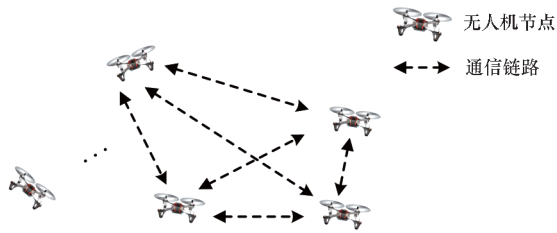


图 1 无人机自组网拓扑结构

Fig. 1 Topology of UAV ad hoc network

在 SPMA 协议中,节点退避的目的是通过延迟发送数据,使得信道负载  $C$  逐渐降低到优先级阈值  $T_b^i$ ,而  $T_b^i$  的计算原则是保证信道负载维持在满载门限  $C_0$ 。满载门限体现的是在满足 99% 传输成功率要求下的最大网络容量,而信道负载取决于已发送的数据,受到退避机制的影响。具体来说,信道负载大于满载门限会导致数据的冲突碰撞概率增大,反之会增加数据的时延。可见,在基于 SPMA 协议的通信系统中,信道负载是否维持在满载门限是影响网络性能的关键。因此,可以用两者的大小关系来衡量数据传输的可靠性和时延。需要说明的是,高优先级业务具有时延敏感性,而低优先级业务对时延具有一定的容忍程度。综上,可将目标优化问题定义为:

$$\min \left| \sum_{l=1}^{N_f} C_l - C_0 \right|$$

$$\text{s. t. } \begin{cases} \text{C1: } T_b^0 < T_b^1 < \dots < T_b^{p-1} \\ \text{C2: } 0 \leq T_b^i \leq T_{b,\max}, 0 \leq i < p \end{cases} \quad (1)$$

其中,  $C_l$  表示信道  $l$  上的负载统计,  $C_0$  表示满载门限,  $T_b^i$  表示优先级  $i$  分组的退避时间,约束条件 C1 表示各优先级分组退避时间的关系,约束条件 C2 表示各优先级分组的退避时间应该在一个合理区间,  $T_{b,\max}$  表示各优先级业务的最大有效期。有效期是指数据包从生成时刻起,可以被网络尝试传输的最大时间。为了更真实地反映不同业务对时延的容忍度,并在网络拥塞时优先保证高优先级业务的及时性,需要为不同优先级业务设置不同的有效期,这正是体现业务 QoS 差异的关键设计之一。一般来说,业务优先级越高,有效期越短。

## 2 退避策略

现有 SPMA 协议的相关研究大多直接沿用了 BEB 算法。但特殊的发送机制和基于优先级区分的特点,使得 SPMA 协议在本质上区别于传统 MAC 协议,这意味着 BEB 策略应用在 SPMA 通信系统中会存在局限性。为进一步提高 SPMA 协议数据传输性能和高效利用网络资源,部分研究提出了针对性的退避机制,但无论是退避时间的计算模型还是考虑的影响参数都不够成熟。下面将对该问题进行详细说明。

### 2.1 二进制指数退避及其改进策略

BEB 策略能有效降低数据包碰撞的概率,是网络中应用最广泛的退避策略之一。Bianchi 在饱和输入条件下,首先利用马尔可夫链对 BEB 策略进行建模并提出了著名的 Bianchi 模型<sup>[18]</sup>,然后定量分析了节点传输概率和网络吞吐量。在此基础上,后续有大量研究对该模型的假设条件或存在的问题做出了深入探讨<sup>[19-20]</sup>。至此,BEB 算法的理论分析研究已较为成熟。

BEB 算法规定节点需要维持一个竞争窗口 (contention window, CW),并在尝试接入信道时确定 CW 的取值。节点通常将 CW 的大小初始化为  $W_{\min}$ ,然后在通信过程中根据传输结果对 CW 的大小进行调整。具体调整策略为:如果上一次传输发生了冲突,说明多个节点同时选择了相同的退避时间,为了减小这种可能性,节点将 CW 翻倍直至最大值  $W_{\max}$ ;一旦数据传输成功,节点就将 CW 重新设置为最小值  $W_{\min}$ 。CW 大小和退避次数  $j$  的关系为:

$$W_j = \begin{cases} 2^j W_{\min} & 0 \leq j < m \\ W_{\max} & j \geq m \end{cases} \quad (2)$$

其中,  $m$  表示 CW 翻倍的次数。确定了 CW

大小之后,节点便在 $[0, W_j]$ 区间选择一个随机值作为退避时间。节点在单位时隙内如果检测到信道空闲就将退避时间减1,直到退避倒计时为0,才能接入信道。

现有研究充分表明<sup>[21-23]</sup>,CW 的取值范围和调整策略在很大程度上决定着网络性能的优劣。但在实际应用中,这两部分内容的确定却成为难题。一方面,应用场景的多样性和业务的复杂性会导致 CW 范围的确定十分困难;另一方面,BEB 机制中 CW 的调整策略会产生“竞争窗口振荡”的现象,这使得成功接入信道的节点下次竞争到信道的概率会更大,进而导致一段时间内信道一直被同一节点使用,最终出现节点接入不公平的问题<sup>[24]</sup>。针对此问题,一些研究提出了乘性增加线性减少(multiple increase linear decrease, MILD)和指数增加指数减少(exponential increase exponential decrease, EIED)等算法<sup>[25]</sup>。此类改进算法虽然在一定程度上有助于节点更公平地使用信道,但它们对不同负载的适应性欠佳。

## 2.2 现有 SPMA 协议退避策略

BEB 及其改进退避策略与 SPMA 协议的发送机制并不契合。为此,文献[8]提出一种首先根据业务优先级和负载统计周期计算退避窗口,然后在退避窗口内随机选择退避时间的退避策略,如式(3)所示。

$$T_b = \text{random} \left[ \frac{1}{3} \times T_w \times (i+1) \right] \quad (3)$$

其中, $T_b$ 表示计算得到的退避时间, $T_w$ 表示负载统计周期, $i$ 表示优先级, $\text{random}(x)$ 表示在 $[0, x]$ 区间内随机取值。虽然该退避机制无须根据退避结果对退避窗口做动态调整,并且不再继续沿用 BEB 算法的框架,但在计算退避窗口时仅仅考虑了业务优先级,忽略了 SPMA 协议中数据的发送也直接受到信道负载统计值与优先级阈值这两个参数的影响。

文献[26]提出了一种考虑更多影响因素的退避窗口的计算方式,如式(4)所示。

$$T_b = \text{random} \left[ T_w \times (i+1) \times \frac{C - T_h^i}{H} \right] \quad (4)$$

其中, $H$ 表示常量系数。虽然该机制将优先级阈值和信道负载统计纳入了退避窗口的计算,但依然简单地将退避窗口与影响参数的函数关系抽象为线性,这并不能适应无人机自组网中复杂的数据传输过程。

文献[9]不仅考虑了业务优先级、信道负载统计和优先级阈值因素,还基于指数函数和对数

函数的凹凸性比较,提出了一种对数模型的退避策略,退避窗口的计算如式(5)所示。

$$T_b = M \cdot \ln [ N \times (i+1) \times (C - T_h^i) \times C ] \quad (5)$$

其中, $M$ 和 $N$ 均是常量系数。但该机制的问题是没有引入随机性,是一种确定性的退避策略,并且对数模型也仅仅是基于函数凹凸性得出,其科学性有待进一步提高。除此之外, $M$ 和 $N$ 取值与具体的网络结构、传输速率、优先级业务量比例以及负载统计周期等有关,确定过程较为复杂。

综上,现有 SPMA 协议退避策略的相关研究虽然在影响因素和函数模型两方面不断深入,但是它们的研究方法本身具有一定的局限性,即 SPMA 协议中数据传输过程的复杂性较高,确定的初等函数关系不足以映射退避时间的准确变化。

## 3 基于 DDQN 的退避算法

由于现有退避策略存在局限性,本文将针对 SPMA 协议提出一种基于深度强化学习的智能退避策略。在 SPMA 协议中,节点退避时间的选择是一个动态决策的过程,即某一时刻无人机节点在需要退避的时候采取确定退避时间的动作,在退避结束后,判断此时信道负载与优先级阈值的大小关系,产生退避成功或失败的决策结果,并根据通信环境的奖励反馈及时地增大或减小下一次退避时间,以追求长期的接入效益,最终提高网络整体性能。

在 SPMA 协议的退避过程中,虽然离散的状态空间和动作空间表明,基于动作状态价值的  $Q$ -Learning 算法(一种无模型的强化学习算法,通过与环境实时交互学习最优策略,采用状态-动作值函数即  $Q$  函数来评估行动价值)可用于此问题的求解,但是数据不同的有效期和传输调度的复杂性会导致  $Q$  表占据过大内存空间。这极大地增加了状态动作对的复杂程度和查找时间,将导致无人机退避时间的选择不能很快收敛到最佳状态。此外,深度  $Q$  网络(deep  $Q$ -network, DQN)算法(一种结合深度神经网络和  $Q$ -Learning 的强化学习算法,用于解决高维状态空间下的 MDP 问题)虽然可以扩展  $Q$ -Learning 的状态和动作空间,但是存在严重的高估问题。相比之下,DDQN 算法具有更好的性能<sup>[27]</sup>。这是因为该算法引入了两个  $Q$  网络,其中主网络(也称当前网络)用于选择动作,目标网络用于估计该动作的价值,并且目标网络的参数更新相对缓慢。这种延迟更新的策略有助于减少高估现象的发生和智能体(Agent)

不必要的探索,提高算法的稳定性和收敛性。

在 SPMA 协议中,高效的退避策略需要将信道负载降低到优先级阈值的同时尽可能地减小退避时间。因此,SPMA 协议退避机制的优化需要综合考虑业务优先级、优先级阈值和信道负载等因素,才能进一步提高网络性能。

### 3.1 退避算法设计

深度强化学习的主要组成部分是交互环境、状态空间、动作空间、动作选择策略和奖励函数,下面将基于上文已构建的优化问题对每部分的设计进行详细说明,并给出基于 DDQN 的智能退避策略的模型图。

1) 交互环境:基于 DDQN 的 SPMA 协议智能退避策略将无人机节点视作智能体,整个无人机自组网是 Agent 学习的环境。数据流以泊松分布到达 MAC 层不同优先级队列,然后节点根据 SPMA 协议对其进行发送或退避。

2) 状态空间:基于区分业务优先级和信道负载统计的 SPMA 协议发送机制,该策略定义状态由两部分构成——一部分是此时需要发送的数据包的优先级  $i_t$ ,另一部分是当前信道负载统计  $C_t$ 。本文设定所有无人机节点均共享单信道并工作于全双工模式,信道负载定义为统计周期内节点发送的数据包与接收到的数据包之和。特别地,无人机自组网流量虽呈现随机波动特性,但观测表明,信道负载变化在短时间尺度上存在显著的自相关性和局部平稳性。因此,为了更好地选择退避时间,基于 DDQN 的 SPMA 协议智能退避策略希望能够学习信道负载的变化趋势,所以状态的设计引入了过去一定时刻的信道负载信息。综上所述,将某一时刻的状态定义为:  $s_t = \{i_t, C_t, C_{t-1}, \dots, C_{t-m}\}$ ,其中  $m$  表示考虑的信道负载统计值的历史数目,取正整数。

3) 动作空间:Agent 的动作  $a_t$  表示数据需要退避时,节点在相应退避窗口内选取的退避时间。具体来说,SPMA 协议中不同优先级业务具有不同有效期,因此它们的退避窗口是不同的。各优先级业务的退避窗口由不同数目的单位时隙构成,表明深度强化学习模型里的动作空间是离散且有限的。设优先级的集合为  $\mathbf{P} = \{0, 1, \dots, p-1\}$ ,该优先级可选的动作数量为  $N_a^i$ ,对于优先级  $i \in \mathbf{P}$  的数据来说,该优先级业务的动作子空间可以表示为:

$$A_i = \left\{ \sum_{l=0}^i N_a^{l-1} + 1, \sum_{l=0}^i N_a^{l-1} + 2, \dots, \sum_{l=0}^i N_a^{l-1} + N_a^i \right\} \cdot \tau \quad (6)$$

其中,  $\tau$  是单位时隙长度,并设  $N_a^{-1} = 0$ 。因此,整个网络的动作空间表示为  $\mathbf{A} = \{A_0, A_1, \dots, A_{p-1}\}$ 。在 SPMA 协议中,数据包退避是为了将信道负载降低到优先级阈值,并且在退避倒计时结束后,节点需要再次判断此时网络状态是否满足数据发送条件。因此,退避窗口不再像 BEB 及其改进策略那样根据退避结果以一定规则变化。

4) 动作选择策略:Agent 根据贪婪策略选择动作,即 Agent 以  $1 - \varepsilon$  的概率选择估计网络计算出来的最优动作  $a^*$ ,以  $\varepsilon$  的概率在所有动作中随机选择。为了保证 Agent 能够探索到更多的动作和状态,  $\varepsilon$  的初始值为 1,然后随着训练迭代次数的增加而减小,直到训练到一定的迭代次数时,  $\varepsilon$  的值将保持不变。具体地,  $\varepsilon$  的取值如下:

$$\varepsilon(t) = \begin{cases} \varepsilon_{\text{start}} - (\varepsilon_{\text{start}} - \varepsilon_{\text{end}}) \times \frac{t}{T_{\text{decay}}} & t < T_{\text{decay}} \\ \varepsilon_{\text{end}} & t \geq T_{\text{decay}} \end{cases} \quad (7)$$

其中,  $\varepsilon_{\text{start}} = 1$ ,  $\varepsilon_{\text{end}} = 0.1$ ,  $T_{\text{decay}} = 5\ 000$  episodes,  $t$  为当前训练 episode 的编号。

5) 奖励函数:表示 Agent 在状态  $s$  执行动作  $a$  后获得的奖励,应体现退避机制需要优化的网络目标。具体来说,SPMA 协议对数据进行退避使得信道负载降低到优先级阈值的同时还需要避免信道闲置情况的发生。因此,奖励函数应体现退避机制对数据接入结果和时延的影响。

退避完成之后,网络状态发生变化,智能体从网络环境中观测退避结果,如果当前信道负载小于优先级阈值则可以接入信道,反之需要再次进行退避。记退避结果因子为  $\varphi \in \{1, -1\}$ ,其中 1 表示退避成功, -1 表示退避失败。通过合理的动作选择,可在高负载时增大退避时间以抑制冲突,在低负载时缩短退避时间以提高响应速度,从而实现信道负载与退避时间的动态匹配。因此,可以用退避结束时刻的信道负载衡量退避时间的长短。同时,基于优先级阈值  $T_h^i$  大于信道负载  $C$  的发送条件,选择用两者差值来衡量时延对奖励的影响。差值越小则表明退避时延越有利于数据的发送,应获得更大的奖励;差值越大表明信道闲置情况越明显,则获得的奖励应越小。特别地,为了尽力满足高优先级业务的低时延、高可靠性需求,将数据优先级也考虑到了奖励函数的影响因素中。综上,单步即时奖励函数设计为:

$$r_t = \varphi \cdot e^{\frac{c - r_h^i}{i+1}} \quad (8)$$

采用标准 DDQN 的目标  $Q$  值<sup>[27]</sup>,表示为:

$$y_t = r_t + \gamma Q(s_{t+1}, \arg \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}; \theta); \theta^-) \quad (9)$$

其中: $\gamma=0.9$  为折扣因子; $\theta$  是当前网络参数, $\theta^-$  为目标网络参数,该双网络结构可有效缓解传统 DQN 的过高估计问题。因此,累积奖励可表示为:

$$J(\theta) = \mathbb{E} \left( \sum_{k=0}^{\infty} \gamma^k r_{t+k} \right) \quad (10)$$

决策目标就是要最大化长期累积奖励。根据以上分析,基于 DDQN 的智能退避策略模型示意图如图 2 所示。

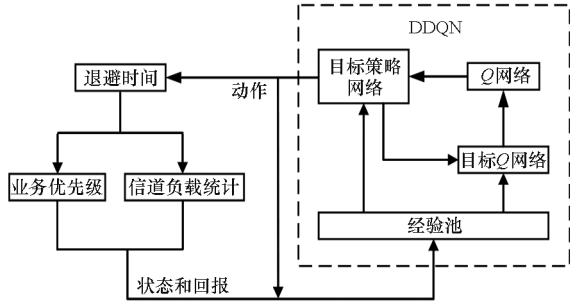


图2 基于 DDQN 的智能退避策略模型

Fig.2 Model of intelligent backoff policy based on DDQN

### 3.2 退避算法实现

基于 DDQN 的 SPMA 协议智能退避算法系统框图如图 3 所示。上层业务以泊松分布到达 SPMA 协议模块的发送队列,节点以信道负载与优先级阈值的大小关系判断队首数据包是否发送,如果发送条件得到满足则通过天线发送该数据包,反之进行退避。在网络训练阶段,SPMA 协议在需要退避时,首先将收集到的当前自组网状态信息传递给 DDQN 模块,然后 DDQN 模块输出一个退避时间,SPMA 协议模块在完成退避之后返回更新的状态和奖励。网络训练完成之后,SPMA 协议模块将直接根据 DDQN 模块输出的退避时间进行退避,不再需要反馈后续状态和奖励值。

DDQN 的输入是智能体的状态空间,因此 DDQN 输入神经元的数量为状态空间的大小。DDQN 的输出为每个可能动作的  $Q$  值,因此将 DDQN 输出神经元的数量定义为最低优先级业务有效期包含的单位时隙的数目。智能体以  $1 - \epsilon$  的概率选择  $Q$  值最大的动作,即退避时间的选择结果。在输入和输出之间有两个隐藏层,层与层之间是全连接的。

在训练阶段,基于 DDQN 的 SPMA 协议智能退避算法首先初始化无人机自组网系统模型和 DDQN 网络参数,并且在每个训练的 episode 更新数据优先级和信道负载统计信息;然后,在每个  $t$

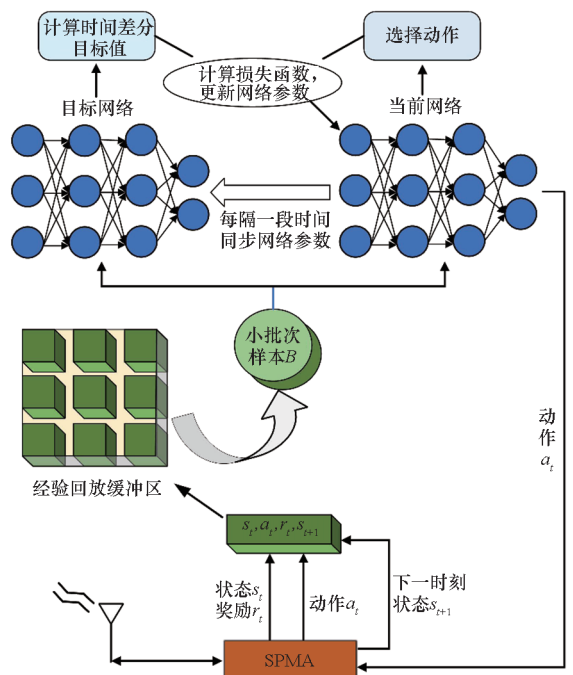


图3 基于 DDQN 的 SPMA 协议智能退避算法系统框图  
Fig.3 System block diagram of intelligent backoff algorithm of SPMA protocol based on DDQN

时刻,智能体利用贪婪策略选择退避动作并执行于无人机自组网环境中,在执行完动作后,计算奖励函数得到当前动作的奖励并转移当前状态为下一状态;接着,将当前步骤所得四元组  $(s_t, a_t, r_t, s_{t+1})$  存储于经验回放池  $D$ ,完成一次数据样本的采集,其中,经验回放是一种强化学习中的技术,它通过存储和重用过去的经验来提高训练的稳定性和效率;最后,从经验回放池中随机取出小批次样本进行训练,计算损失函数  $\mathcal{L}(\theta) = \mathbb{E}_{(s, a, r, s') \sim D} \{ [y_t - Q(s_t, a_t; \theta)]^2 \}$ ,根据梯度下降更新当前网络参数  $\theta$ ,并在一定的迭代回合后同步给目标网络。上述过程意味着,经过这一轮的更新,新的策略模型可能学习到适合当前无人机自组网环境的退避时间选择策略,进而采取获得更大奖励的动作。基于 DDQN 的 SPMA 协议智能退避算法伪代码如算法 1 所示。

### 3.3 参数规模和复杂度分析

所提算法的主网络和目标网络均采用全连接神经网络,其参数规模主要取决于神经网络架构的层数和每层的神经元数量。其中,输入层的神经元数量等于状态空间的维度  $(m + 2)$ ,输出层的神经元数量等于最低优先级业务动作子空间的维度  $N_a^{p-1}$ ,均已由协议约束。设有  $h$  个隐藏层,每个隐藏层的神经元数量为  $N_h$ ,因此参数规模采用式(11)计算。

## 算法 1 基于 DDQN 的智能退避算法

Alg. 1 Intelligent backoff algorithm based on DDQN

输入: 最大迭代次数  $T$ , 状态特征维度  $n$ , 动作集合  $A$ , 学习步长  $\alpha$ , 折扣因子  $\gamma$ , 探索率  $\varepsilon$ , 批量样本数  $B$ , 目标  $Q$  网络更新步长  $k$ , 历史值数目  $m$

输出: 当前  $Q$  网络参数

1. 初始化无人机自组网环境, 包括组网结构和数据业务特性等参数;
2. 使用正态分布随机初始化当前网络参数  $\theta$  和目标网络参数  $\theta^-$ , 更新所有状态行为对应的  $Q$  值, 初始化经验回放列表  $D$ ;
3. **for** each episode do:
4. 更新当前时刻无人机发送数据包的优先级和当前信道负载统计  $C$ ;
5. 初始化智能体观察到的状态空间  $s_t = \{i_t, C_t, C_{t-1}, \dots, C_{t-m}\}$ ;
6. **for** each step  $t$  do:
7. 根据状态  $s_t$  的特征向量, 利用  $\varepsilon$  贪婪策略选择动作  $a_t$ ;
8. 智能体执行动作  $a_t$ , 即无人机节点执行确定时间的退避;
9. 更新数据优先级、传输完成状态、统计负载并记录  $m$  个历史值;
10. 智能体与环境交互得到新的状态  $s_{t+1}$  和奖励  $r_t$ ;
11. 将  $(s_t, a_t, r_t, s_{t+1})$  放入经验回放列表中;
12. 如果经验回放池的样本数大于  $|D|$ , 则用新的转移数据替换最旧的样本数据;
13. 从经验回放列表  $D$  中随机抽取  $B$  个样本, 计算当前的目标  $Q$  值;
14. 根据目标网络, 计算目标  $Q$  值  $y_t$ :
 
$$y_t = \begin{cases} r_t, & s_{t+1} \text{ 是终止状态} \\ r_t + \gamma Q(s_{t+1}, \arg \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}; \theta); \theta^-), & s_{t+1} \text{ 不是终止状态} \end{cases}$$
15. 最小化损失函数, 通过批量随机梯度下降更新主网络参数  $\theta$ ;
16. 每隔固定步长  $k$  更新目标网络参数  $\theta^- \leftarrow \theta$ ;
17. **end for**
18. **end for**

$$N_{\text{para}} \approx (m+2+1) \times N_h + (N_h+1) \times N_h \times (h-1) + (N_h+1) \times N_a^{p-1} \quad (11)$$

需要说明的是, 基于 DDQN 的智能退避算法的训练阶段在地面站或边缘服务器上完成, 仅将收敛后的轻量推理模型下载至机载设备, 属于“离线训练 + 在线部署”模式, 降低了对机载终端算力的要求, 可实现轻量化部署。

由于无人机端仅执行推理功能, 只需要考虑

前向传播的计算量, 计算量约等于  $(m+2) \times N_h + N_h \times N_h + N_h \times N_a^{p-1}$  次浮点运算。对于典型无人机自组网参数和双隐藏层神经网络结构, 以 NVIDIA Jetson Nano 为例, 单次推理时间在微秒级, 可满足 SPMA 协议的决策需求。

## 4 仿真验证与结果分析

在不同业务量场景下, 将所提的基于 DDQN 的智能退避策略与二进制指数退避策略、基于对数函数的退避策略进行性能对比。SPMA 协议采用 OPNET Modeler 14.5 进行仿真, 机器学习算法采用 Python 进行仿真。为了模拟无人机自组网的高动态性、体现信道负载变化, 在 OPNET 中设置了节点移动模型、动态拓扑演化、突发流量注入、动态链路质量模型等。主要仿真参数如表 1 所示, 其中, 默认节点个数为 15。另外需要特别说明的是, 基于对无人机自组网典型应用和业务 QoS 需求的调研<sup>[3-4]</sup>, 设置了 3 类典型优先级业务, 优先级由高到低分别对应: ①指控业务(优先级 0), 即操作员向无人机发送的指挥和控制指令, 如任务更新指令、突发威胁规避指令等, 是极端时延和可靠性敏感型业务, 业务有效期设置为 10 ms; ②控制和载荷业务(优先级 1), 即维持无人机安全飞行和编队控制所要交互的信息(如实时位置、航路点信息), 以及任务相关载荷采集的关键数据(如侦察无人机需要回传的目标图片、视频等), 是时延和可靠性敏感型业务, 业务有效期设置为 20 ms; ③传感业务(优先级 2), 如常规

表 1 主要仿真参数

Tab. 1 Main simulation parameters

参数	取值
传输速率	2 Mbit/s
数据包长度	1 000 bit
单位时隙长度 $\tau$	20 $\mu$ s
业务生成速率 $\lambda$	[3 400 个/s, 4 100 个/s]
业务优先级数 $p$	3
负载统计周期 $T_w$	10 ms
各优先级分组业务量比例	1 : 2 : 3
各优先级分组业务有效期	10 ms/20 ms/100 ms
节点个数	[5, 55]
折扣因子 $\gamma$	0.9
经验回放池大小 $ D $	20 000
采样样本数 $B$	256

状态报告、周期性环境监测等,是时延和可靠性不敏感型或尽力而为型业务,对时延有较大容忍度,业务有效期设置为 100 ms。在不同的场景中,如军事侦察、民用物流、工业巡检、农业监测等,不同优先级业务所占的比例也不同,这里以民用场景为例,设置三种业务的比例为 1:2:3,中等和低优先级业务所占比例较高。

本节主要从传输时延、首次退避成功率和传输成功率、策略可扩展性四个方面对比评估所提算法的性能。其中,传输时延是指从数据包生成时刻到其被目的节点成功接收并确认(回复确认字符 ACK)之间的时间间隔。通常来说,传输成功率是指目的节点成功接收到的数据量与源节点发送数据量之间的比值,但基于 SPMA 协议的特殊性,本文将传输成功率定义为:目的节点成功接收到的某一优先级且在有效期内的数据包数与源节点产生的同一优先级数据包数之比。后文若无特殊说明,传输成功率均使用此定义,下面将详细说明这一定义的合理性。

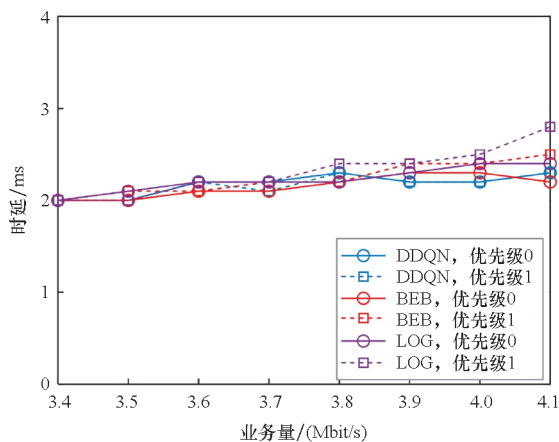
SPMA 协议通过采用一系列特殊机制,使得高优先级分组在被发送前拥有对信道资源更大的使用权,从而保证了其低时延性能。在之后的无线信道传播过程中,所有数据面临的冲突碰撞情况是相同的。在传输资源有限的情况下,网络总是优先考虑高优先级分组的传输,会导致部分低优先级分组受到排队等待和传输中断的影响更大,其时延会随之增高。当低优先级分组时延持续增高并且超过相应有效期时,传输将变得无意义。在以任务为导向的通信系统中,时延超过有效期的数据即使没有被发送,也将被视为“传输失败”。因为一个即使最终送达但严重超时的数据包(如迟到的紧急指令),对于实时应用已失去价值,所以计入失败是合理的。这比单纯使用物理层是否成功送达更能反映网络对上层应用的服务质量。因此,为了体现上述 SPMA 协议的本质特征,本文对传输成功率进行定义和说明。传输成功率越大,表明传输可靠性越强,在差异化传输过程中该优先级数据被优先考虑,获得了更优质的服务。后续仿真结果分析均基于此定义。

#### 4.1 传输时延

图 4 展示了三种退避策略在不同业务量下各优先级分组的传输时延。其中,DDQN 表示所提的基于 DDQN 的智能退避策略,BEB 指二进制指数退避策略,LOG 指基于对数函数的退避策略。

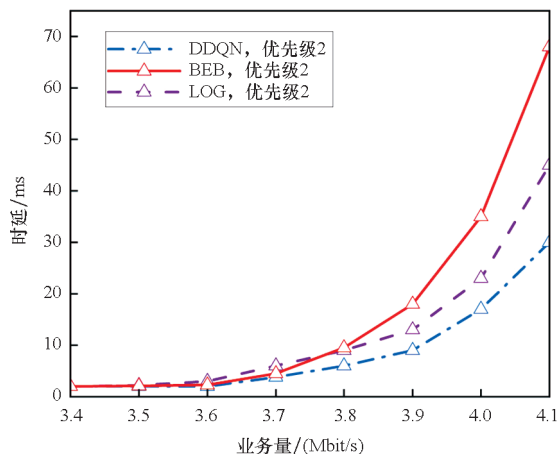
图 4(a) 仿真结果显示,在 3.4 ~ 4.1 Mbit/s 业务量范围内,高优先级 0、1 分组可以几乎无退

避地直接发送。总的来说,不同策略下的时延差别不大,不过随着业务量的增大,时延有增大的趋势并且不同策略下时延的差别也逐渐明显。图 4(b) 仿真结果显示,低优先级 2 分组在业务量大于 3.6 Mbit/s 时开始出现退避,时延随之增加并且数据在不同退避机制下有着不同的时延性能。综合两幅图的结果表明,在 3.4 ~ 4.1 Mbit/s 业务量变化范围内,所提的 DDQN 策略具有最优的时延性能,并且随着业务量增大,性能提升效果愈发明显。特别是对于低优先级 2 分组来说,在业务量为 4.0 Mbit/s 时,DDQN 策略相对于 LOG 策略带来了 26% 的时延性能提升;在业务量为 4.1 Mbit/s 时,DDQN 策略相对于 LOG 策略带来了 33.3% 的性能提升。这是因为随着业务量增加,SPMA 协议中数据调度机制的复杂性得以体现,此时恰当的退避时间在更大程度上影响着数据传输性能。



(a) 优先级 0、1 分组的传输时延

(a) Transmission delay of priority 0 and 1 packets



(b) 优先级 2 分组的传输时延

(b) Transmission delay of priority 2 packets

图 4 三种退避策略下不同优先级分组的时延

Fig. 4 Delay of different priority packets under three backoff strategies

在 SPMA 协议中,若退避策略性能较差,过长或过短的退避时间都会导致数据时延的增加,甚至影响传输成功率,并且这种负面影响会随着信道负载的增大而增大。因此,现有的以确定性初等函数计算退避窗口或以倍数增大退避窗口的方法无法选择到合适的退避时间,而基于 DDQN 的智能退避策略可以很好地适应 SPMA 协议的复杂传输过程和网络状态。

### 4.2 首次退避成功率

由于在 3.4 ~ 4.1 Mbit/s 业务量范围内,SPMA 协议主要通过对优先级 2 分组进行退避以维持网络健康状态,因此接下来以优先级 2 数据包的退避次数来衡量不同退避策略选择退避时间的准确性。

图 5 展示了优先级 2 分组在三种退避策略下的各退避次数的数据包占需要退避的总数据包的比例。仿真结果表明,DDQN 策略可以将退避次数控制在 5 次之内,并且退避一次就成功接入的比例为 83%;而在 BEB 和 LOG 策略下,部分数据包的退避次数会达到 6 次,首次退避成功的比例分别为 65% 和 72%,均低于 DDQN 策略。这是因为基于对数函数的退避策略仅考虑了当前时隙的信道占用率,BEB 策略的退避窗口增长机制不符合 SPMA 协议的退避特性。同时,DDQN 策略将退避次数控制在 3 次以内的概率高达 95%,而 BEB 和 LOG 策略的概率分别为 90% 和 92%。这意味着基于 DDQN 的智能退避策略相对于其他策略可以更准确地找到退避时间,无须多次退避尝试。值得注意的是,SPMA 协议引入的帧抢占机制会导致部分数据在退避时被高优先级数据的到来抢占,进而造成退避次数的增加,所以 DDQN 策略的一次退避成功率只有 83%。

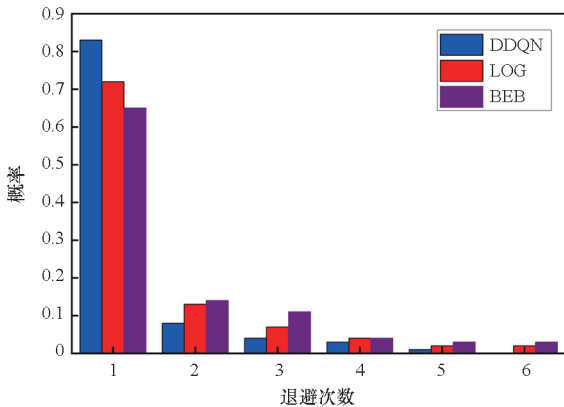
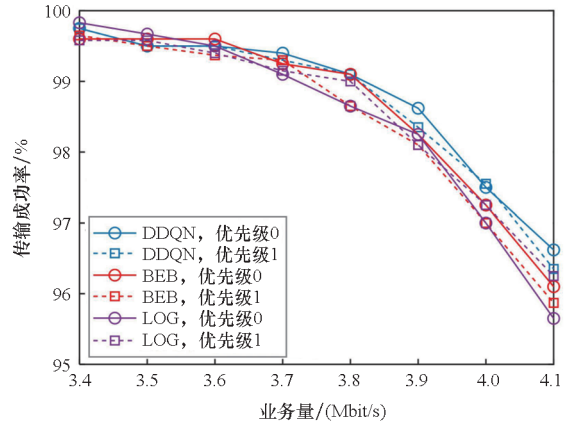


图 5 三种退避策略下优先级 2 分组退避次数的分布  
Fig. 5 Distribution of backoff times of priority 2 packets under three backoff strategies

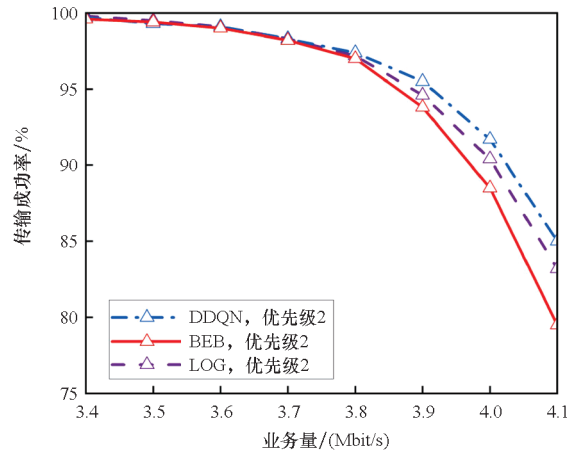
### 4.3 传输成功率

图 6 展示了三种退避策略下各优先级分组在 3.4 ~ 4.1 Mbit/s 业务量范围内的传输成功率,仿真结果表明 SPMA 协议可优先为高优先级分组提供高可靠的传输服务。



(a) 优先级 0、1 分组的传输成功率

(a) Transmission success rate of priority 0 and 1 packets



(b) 优先级 2 分组的传输成功率

(b) Transmission success rate of priority 2 packets

图 6 三种退避策略下不同优先级分组的传输成功率

Fig. 6 Transmission success rate of different priority packets under three backoff strategies

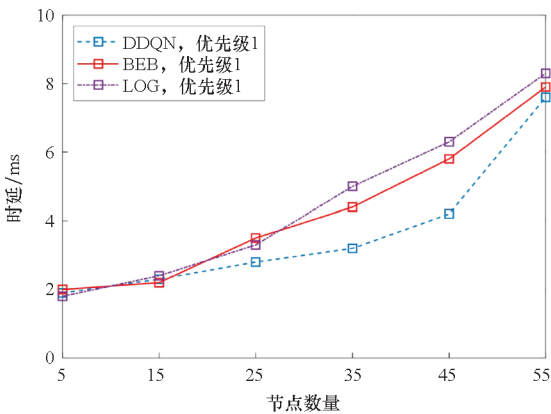
图 6(a) 仿真结果表明,在 3.4 ~ 4.1 Mbit/s 业务量范围内,优先级 0、1 分组的传输成功率随着业务量的增加而略微减小,三种退避策略下传输成功率的差别不大,并且全都可达到 95% 以上。图 6(b) 仿真结果表明,不同退避策略对传输成功率的影响主要体现在低优先级 2 分组上。具体来说,当业务量在 3.4 ~ 3.6 Mbit/s 范围时,此时网络处于轻负载,优先级 2 分组可以进行无退避发送,传输成功率可达到 99% 以上;当业务量在 3.6 ~ 4.1 Mbit/s 逐渐增大时,SPMA 协议对数据的发送情况发生改变,因此低优先级 2 分组

的传输成功率会有较大程度的减小,并且三种策略下的传输成功率的差别逐渐明显,大小关系是  $BEB < LOG < DDQN$ 。这是因为当业务量增大后,随机竞争类 MAC 协议主要通过退避来避免数据冲突。同时,SPMA 协议无信道预约过程,完全依赖对信道状态的估计来判断数据的发送,因此合理的退避时间在很大程度上影响着 SPMA 协议数据的传输成功率。具体来说,过长的退避时间会导致优先级队列中队尾数据包的长时间等待,过短的退避时间会导致数据经历多次退避,这两方面都最终会造成低优先级分组时延的增加。随着时间的推进,这种现象造成的负面影响越来越明显,部分低优先级数据的时延会超过其有效期,信息不再具有价值而被节点直接丢弃,从而导致传输失败。

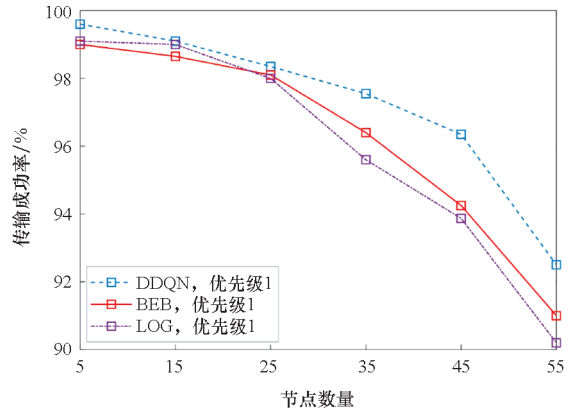
#### 4.4 策略可扩展性

为了验证所提策略在不同节点规模下的鲁棒性和可扩展性,本小节以中等优先级业务(优先级 1)为例,评估在给定的业务量(3.8 Mbit/s)下,时延和传输成功率随着节点数量(5、15、25、35、45、55)的变化趋势,仿真结果如图 7 所示。

图 7(a)和图 7(b)的结果表明,本文提出的 DDQN 退避策略在节点数量为 5 ~ 55 时均能有效工作。在小规模网络(5 ~ 25 个节点)中,不同策略下的时延差别不大,并且传输成功率都可以达到 98%;随着节点数量的增大,在中等规模网络(25 ~ 45 个节点)中,出现时延快速上升、传输成功率快速下降的趋势,但是所提 DDQN 退避策略相比传统的 BEB、LOG 策略展现出的优势较为明显,其中时延平均降低 31%,展现出了较好的可扩展性;当网络规模较大时(45 ~ 55 个节点),网络整体负载加重导致所有策略性能均明显下降,



(a) 优先级 1 分组传输时延随节点数量变化趋势  
(a) Transmission delay versus the number of nodes for priority 1 packets



(b) 优先级 1 分组传输成功率随节点数量变化趋势  
(b) Transmission success rate versus the number of nodes for priority 1 packets

图 7 三种退避策略下的网络性能随节点数量变化趋势  
Fig. 7 Network performance versus the number of nodes for priority 1 packets under three backoff strategies

所提 DDQN 策略的优势减小,这表明在超大规模网络中可能需要结合其他组网优化机制,如分层分簇网络架构等。

## 5 结论

无人机自组网的高动态性与业务差异化需求对 MAC 协议提出了双重挑战。针对现有 SPMA 协议的退避策略在动态网络适应性及多优先级区分能力上的不足,本文提出了一种基于 DDQN 的智能退避方法。首先,分析了无人机自组网系统模型和退避优化问题,将 SPMA 协议的退避时间选择过程抽象为 MDP,建立了无人机节点退避行为与网络状态的动态映射关系;然后,选择了合适的 DDQN 算法对问题进行求解,并对深度强化学习的交互环境、动作空间、状态空间和奖励函数进行了设计,其中不同于传统问题求解的是,在 SPMA 协议中,信道负载是否稳定在满载门限是影响网络性能的关键;最后,仿真实验表明,所提的智能退避算法在传输时延、传输成功率和首次退避成功率方面具有更优的性能,并且在节点规模增大后,具有较强的可扩展性,为高动态自组网提供了一种轻量化、自适应的退避决策新方法。未来工作将进一步探索跨层级联合优化机制,推动智能退避技术的工程化应用。

## 参考文献 (References)

- [1] CEVIZ O, SEN S, SADIOGLU P. A survey of security in UAVs and FANETs: issues, threats, analysis of attacks, and solutions [J]. IEEE Communications Surveys & Tutorials, 2025, 27(5): 3227–3265.

- [2] XIE L F, XU J, ZENG Y. Common throughput maximization for UAV-enabled interference channel with wireless powered communications[J]. *IEEE Transactions on Communications*, 2020, 68(5): 3197–3212.
- [3] WANG H J, ZHAO H T, ZHANG J, et al. Survey on unmanned aerial vehicle networks: a cyber physical system perspective[J]. *IEEE Communications Surveys & Tutorials*, 2020, 22(2): 1027–1070.
- [4] 尹浩, 魏急波, 赵海涛, 等. 面向有人/无人协同的智能通信与组网关键技术: 现状与趋势[J]. *通信学报*, 2024, 45(1): 1–17.  
YIN H, WEI J B, ZHAO H T, et al. Intelligent communication and networking key technologies for manned/unmanned cooperation: states-of-the-art and trends[J]. *Journal on Communications*, 2024, 45(1): 1–17. (in Chinese)
- [5] 王洁, 王海军, 张晓瀛, 等. 无人机自组网中基于优先级和发送概率的低时延 MAC 协议设计[J]. *信号处理*, 2024, 40(4): 661–670.  
WANG J, WANG H J, ZHANG X Y, et al. Design of low latency MAC protocol based on priority and transmission probability in UAV ad-hoc network[J]. *Journal of Signal Processing*, 2024, 40(4): 661–670. (in Chinese)
- [6] HE Y, ZHANG Y, SUN H G, et al. Stochastic geometry modeling and analysis of statistical priority-based multiple access protocol [C]//*Proceedings of the IEEE/CIC International Conference on Communications in China (ICCC)*, 2021: 1059–1064.
- [7] ZHENG Z C, JIANG S M, FENG R Y, et al. An adaptive backoff selection scheme based on Q-learning for CSMA/CA[J]. *Wireless Networks*, 2023, 29(4): 1899–1909.
- [8] ZHANG Y, ZHANG Z K, SUN H G, et al. On the backoff scheme for SPMA network: a spatio-temporal mathematical model[J]. *IEEE Communications Letters*, 2023, 27(9): 2541–2545.
- [9] 王希洋. 基于统计优先级的数据链 MAC 协议研究[D]. 西安: 西安电子科技大学, 2017.  
WANG X Y. Research on statistical priority-based MAC protocol for data link[D]. Xi'an: Xidian University, 2017. (in Chinese)
- [10] 郑文庆, 金虎, 郭建蓬, 等. 基于信道占用及优先级的 MAC 协议退避算法[J]. *计算机工程与应用*, 2019, 55(11): 80–84, 116.  
ZHENG W Q, JIN H, GUO J P, et al. Backoff algorithm for MAC protocol based on channel occupancy and priority[J]. *Computer Engineering and Applications*, 2019, 55(11): 80–84, 116. (in Chinese)
- [11] 包叶锦. SPMA 协议接入控制机制研究[D]. 哈尔滨: 哈尔滨工程大学, 2022.  
BAO Y J. Research on the access control mechanism of SPMA protocol[D]. Harbin: Harbin Engineering University, 2022. (in Chinese)
- [12] IACOVELLI G, COLUCCIA A, GRIECO L A. Multi-UAV IRS-assisted communications: multinode channel modeling and fair sum-rate optimization via deep reinforcement learning[J]. *IEEE Internet of Things Journal*, 2024, 11(3): 4470–4482.
- [13] LUONG N C, HOANG D T, GONG S M, et al. Applications of deep reinforcement learning in communications and networking: a survey[J]. *IEEE Communications Surveys & Tutorials*, 2019, 21(4): 3133–3174.
- [14] ZHU C X, DASTANI M, WANG S H. A survey of multi-agent deep reinforcement learning with communication[J]. *Autonomous Agents and Multi-Agent Systems*, 2024, 38(1): 4.
- [15] 胡帅. 基于深度强化学习的移动 Ad-Hoc 网络路由选择[D]. 北京: 北京邮电大学, 2023.  
HU S. Mobile Ad-Hoc network routing selection based on deep reinforcement learning[D]. Beijing: Beijing University of Posts and Telecommunications, 2023. (in Chinese)
- [16] 刘派. SPMA 协议智能优化设计技术研究[D]. 杭州: 浙江大学, 2021.  
LIU P. Research on intelligent optimization design technology of SPMA protocol [D]. Hangzhou: Zhejiang University, 2021. (in Chinese)
- [17] 章晓豪. SPMA 优先级阈值与退避策略优化方法研究[D]. 杭州: 浙江大学, 2022.  
ZHANG X H. Research on SPMA priority threshold and backoff strategy optimization [D]. Hangzhou: Zhejiang University, 2022. (in Chinese)
- [18] BIANCHI G. Performance analysis of the IEEE 802.11 distributed coordination function [J]. *IEEE Journal on Selected Areas in Communications*, 2000, 18(3): 535–547.
- [19] KWAK B J, SONG N O, MILLER L E. Performance analysis of exponential backoff [J]. *IEEE/ACM Transactions on Networking*, 2005, 13(2): 343–355.
- [20] ZHENG Y, LU K J, WU D P, et al. Performance analysis of IEEE 802.11 DCF in imperfect channels [J]. *IEEE Transactions on Vehicular Technology*, 2006, 55(5): 1648–1656.
- [21] SEPTA N. The performance analysis of 802.11p with cooperative communication and dynamic contention window[J]. *Wireless Personal Communications*, 2023, 131(1): 431–454.
- [22] TORGUNAKOV V A, LOGINOV V A, KHOROV E M, et al. Algorithm for adaptative selection of a contention window in IEEE 802.11bd networks[J]. *Journal of Communications Technology and Electronics*, 2022, 67: S241–S247.
- [23] NANDYALA C S, JIN S. An adjustable contention window management for dense IEEE 802.11 networks[J]. *IEICE Transactions on Communications*, 2022, E105.B(3): 270–274.
- [24] AL-HUBAISHI M, ALAHDAL T, ALSAQOUR R, et al. Enhanced binary exponential backoff algorithm for fair channel access in the IEEE 802.11 medium access control protocol[J]. *International Journal of Communication Systems*, 2014, 27(12): 4166–4184.
- [25] RAHAMAN M M, ASHRAFUZZAMAN K, CHOWDHURY M S, et al. Performance measurement of different backoff algorithms in IEEE 802.15.4 [C]//*Proceedings of the International Conference on Innovations in Science, Engineering and Technology (ICISSET)*, 2017: 1–4.
- [26] LI Y J, LI Z M. Research and backoff algorithm improvement of statistical priority-based multiple access protocol [C]//*Proceedings of the International Conference on Space-Air-Ground Computing (SAGC)*, 2021: 67–72.
- [27] VAN HASSELT H, GUEZ A, SILVER D. Deep reinforcement learning with double Q-learning [EB/OL]. (2015–12–08) [2025–05–15]. <https://arxiv.org/abs/1509.06461>.