

文章编号: 2095-4980(2024)01-0080-08

基于边缘计算技术的智能无人系统

李飞翔¹, 李 宁¹, 刘明哲¹, 贺思源¹, 孙 甜²

(1. 中国电子科技集团公司 第十五研究所, 北京 100089; 2. 北京清华同衡规划设计研究院, 北京 100089)

摘 要: 智能无人系统是信息化战争中夺取信息优势、实施精确打击、完成特殊作战任务的重要手段之一, 是未来军事力量的倍增器。如何应用新兴技术提升智能无人系统在未来作战场景的有效性是当前重要的研究方向之一。介绍了智能无人系统的相关背景, 引入边缘计算技术以及其应用在智能无人系统中的优势, 对边缘计算技术在智能无人系统中的问题进行建模及算法求解, 最后通过仿真实验验证了算法的有效性。

关键词: 智能无人系统; 边缘计算; 深度强化学习

中图分类号: TN914.42

文献标志码: A

doi: 10.11805/TKYDA2021402

Intelligent unmanned system based on edge computing technology

LI Feixiang¹, LI Ning¹, LIU Mingzhe¹, HE Siyuan¹, SUN Tian²

(1. The 15th Research Institute of China Electronics Technology Group Corporation, Beijing 100089, China;

2. Beijing Tsinghua Tongheng Urban Planning & Design Institute, Beijing 100089, China)

Abstract: Intelligent unmanned system is one of the important means to seize the information advantage, implement the precision attack and complete special combat tasks in information war. It is also a multiplier of future military forces. How to apply emerging technologies to improve the effectiveness of intelligent unmanned systems in future combat scenarios is one of the important research directions. The relevant background of intelligent unmanned system is introduced. The edge computing technology and its advantages in intelligent unmanned system are analyzed, and then the problems during the application of edge computing technology into intelligent unmanned system are modeled and solved. The effectiveness of the algorithm is verified by simulation experiments.

Keywords: intelligent unmanned system; edge computing; deep reinforcement learning

军事领域的智能无人化发展, 是“加快军事智能化发展”的重要内容, 也是军事智能“由人向物”迁移的关键领域, 其推力源泉, 主要来自科技突破引发的军事价值认知整体转变。一般意义上的军事智能, 内涵极为广博, 是知识结构、智力水平、经验灵感与谋略艺术在军事实践中的综合运用^[1], 形成主体在“人”, 衍生于军事技术与武器装备, 实际效用随军事实践产生。从科技变革角度看, 随着智能技术、网络信息、生物交叉、微纳材料等前沿技术的高速发展, 各产业方向不断加强与传感、网络、智能的交叉融合, 量子通信、蓝领机器人、智能语音、深度学习、移动协作等前沿方向多点突破^[2], 人类智能的物化迁移领域不断加大, 速度不断加快, 程度不断加深, 水平不断加强, 为新一轮军事革命的智能无人化发展创造了极为丰富的理论基础、技术条件、产业环境、人力储备与社会价值观念认同。传统军事智能, 正在经历“无人化”的革命性迁移, 催生“人与机器”战场分工的重新配置和新的作战能力与作战机制。

智能无人作战系统, 目前尚无统一定义^[3]。从功能特征把握, 主要包括模拟军事人员简单战场行为、形成作战知识表示机制、拓展军事人员作战认知理解能力与生理机能等方面, 使传统的完全由人操控的武器装备, 具有自主/半自主的运动控制、任务规划、指挥决策、任务执行等方面的智能特征, 即“使一部机器的反应方式像人类行动时所依据的智能”。如物理域的无人机、无人舰、无人车等物理平台和智能弹药, 信息与认知域的无人值守战场感知系统、智能情报分析与决策支持系统等。

在未来作战场景中, 随着智能无人系统^[4]设备迅速增长, 无线网络^[5]的传输容量面临着前所未有的挑战。为了满足该网络架构对计算服务的需求, 边缘计算技术^[6]被认为是将计算能力从网络中心转移到边缘的有效技术。

与传统云计算相比，边缘计算^[7]既能够有效缩短终端设备与数据中心之间的数据传输距离，又能避免大量数据交换造成的网络拥塞，提高网络回程能力，但边缘计算并不拥有足够的计算能力和存储能力。因此，在多用户场景中，如何在通信和计算资源的约束下智能有效地卸载多个任务，满足智能无人系统设备的服务质量是一个重大的挑战。

本文讨论在一个新的网络场景中，考虑面向智能化的多用户边缘服务问题。具体而言，考虑边缘计算中的通信资源、计算资源和缓存资源的框架，采用深度强化学习方法帮助管理任务卸载。实验结果表明，该框架能有效提高边缘计算的回报率，快速收敛到满意解。

1 系统模型

1.1 网络架构

如图 1 所示，智能无人系统的网络架构由 3 部分组成：宏蜂窝基站、小蜂窝基站和终端设备。宏蜂窝基站和小蜂窝基站分别与各自的服务器连接，称作宏蜂窝基站边缘云和小蜂窝基站边缘云。宏蜂窝基站边缘云可以提供比小蜂窝基站边缘云更强大的计算能力，小蜂窝基站边缘云则更接近终端设备，可以提供更快的计算服务。终端设备与边缘云之间的边缘计算服务包括 3 个步骤：a) 选择终端设备中的任务卸载到小蜂窝基站边缘云或宏蜂窝基站边缘云；b) 边缘云选择卸载任务的执行决策，如果任务尚未缓存在边缘云中，则计算此任务；否则，将直接返回缓存结果；c) 边缘云将结果返回终端设备。

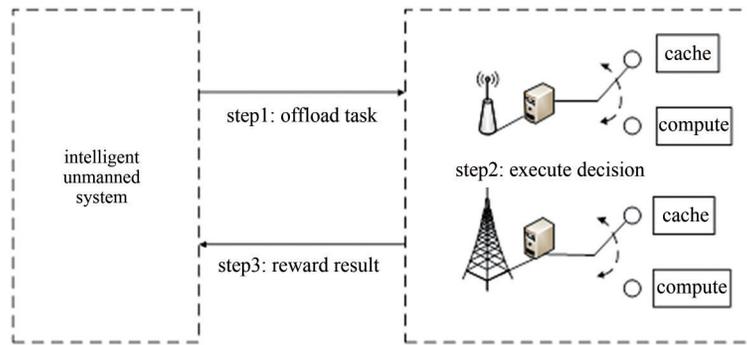


Fig.1 Edge computing service process in network architecture
图 1 网络架构下的边缘计算服务过程

1.2 通信模型

通信模型架构基于智能无人系统试验鉴定环境的异构网络，在此网络架构下，每个小蜂窝基站边缘云都能够通过无线链路与宏蜂窝边缘云通信。每个终端设备中的任务可被卸载到不同边缘云进行计算，且每个终端设备同时只能与一种边缘云通信。假设终端设备 m 中有计算任务 $J_m = \{A_m, A'_m, B_m, T_m^{\max}\}$ ，其中， A_m 和 A'_m 分别表示任务 J_m 计算前后的数据量大小； B_m 表示任务计算所需的 CPU 核数； T_m^{\max} 表示任务的最大时延容忍值。

1) 任务卸载到宏蜂窝基站边缘云。假设 w_{m0} 为作战终端设备 m 到宏蜂窝基站边缘云的传输功率， σ^2 为高斯白噪声常量， h_{m0} 为终端设备 m 到宏蜂窝基站边缘云的信道增益， $k_{m0} \in \{0, 1\}$ 为终端设备 m 和宏蜂窝基站边缘云的影响系数，则终端设备 m 到宏蜂窝基站边缘云的传输速率如式(1)所示：

$$r_{m0} = w_{m0} \log_2 \left(1 + \frac{p_{m0} |h_{m0}|^2}{\sigma^2 + \sum_{m=1}^M k_{m0} p_{0l} |h_{m0}|^2} \right) \quad (1)$$

在以上模型中，假设 α_{m0} 表示作战终端设备的接入成本， β_{m0} 表示宏基站边缘云的频谱使用成本，则通信收益可以建模为 $R_{m0}^{\text{comm}} = \alpha_{m0} r_{m0} - \beta_{m0} w_{m0}$ 。

2) 任务卸载到小蜂窝基站边缘云。假设 w_{mn} 表示终端设备 m 到小蜂窝基站边缘云 n 的传输功率， h_{mn} 表示作战终端设备 m 到小蜂窝基站边缘云 n 的信道增益， $k_{mn} \in \{0, 1\}$ 表示终端设备 m 和小蜂窝基站边缘云 n 的影响系数，则终端设备 m 到小蜂窝基站边缘云 n 的传输速率如式(2)所示：

$$r_{mn} = w_{mn} \log_2 \left(1 + \frac{P_{mn} |h_{mn}|^2}{\sigma^2 + \sum_{m=1}^M k_{mn} P_n |h_{mn}|^2} \right) \quad (2)$$

与上一个模型类似, 假设 α_{mn} 表示终端设备的接入成本, β_{mn} 表示小蜂窝基站边缘云 n 的频谱使用成本, 则通信收益可以建模为 $R_{mn}^{\text{comm}} = \alpha_{mn} r_{mn} - \beta_{mn} w_{mn}$ 。

1.3 计算模型

1) 宏蜂窝基站边缘云中的计算模型。宏蜂窝基站边缘云的计算时延建模为 $T_{m0}^c = \frac{B_m}{C_{m0}}$, 其中 C_{m0} 为宏基站边缘云的计算能力, 则终端设备的任务 m 在宏蜂窝基站边缘云中的计算速率建模为:

$$q_{m0} = \frac{A_m}{B_m / C_{m0}} \quad (3)$$

宏蜂窝基站边缘云中的计算能耗表示为 $e_{m0} = \omega_{m0} T_{m0}^c$, 其中 ω_{m0} 表示宏蜂窝基站边缘云中单位时间内的能耗速率。

假设 ϕ_{m0} 表示终端设备 m 支付的计算成本, 宏蜂窝基站边缘云的计算成本表示为 φ_{m0} , 则计算收益可以建模为 $R_{m0}^{\text{comp}} = \phi_{m0} q_{m0} - \varphi_{m0} e_{m0}$ 。

2) 小蜂窝基站边缘云 n 中的计算模型。小蜂窝基站边缘云 n 的计算时延建模为 $T_{mn}^c = \frac{B_m}{C_{mn}}$, 其中 C_{mn} 表示小蜂窝基站边缘云 n 的计算能力, 则终端设备的任务 m 在小蜂窝基站边缘云 n 中的计算速率建模为:

$$q_{mn} = \frac{A_m}{B_m / C_{mn}} \quad (4)$$

小蜂窝基站边缘云 n 中的计算能耗表示为 $e_{mn} = \omega_{mn} T_{mn}^c$, 其中 ω_{mn} 表示小蜂窝基站边缘云 n 中单位时间内的能耗速率。

假设 ϕ_{mn} 表示终端设备 m 支付的计算成本, 小蜂窝基站边缘云 n 的计算成本表示为 φ_{mn} , 则计算收益可以建模为 $R_{mn}^{\text{comp}} = \phi_{mn} q_{mn} - \varphi_{mn} e_{mn}$ 。

1.4 缓存模型

假设边缘云中缓存 D 个内容文件, 缓存策略由二进制数值确定, 1 表示该内容缓存在边缘云中, 0 表示该内容没有缓存在边缘云中。在此模型中, 内容流行度定义为 $G = \{g_1, g_2, \dots, g_D\}$, 其中 D 表示内容文件的最大类型个数。每个作战终端设备以概率 g_d 请求内容文件 d 。其中, G 服从 Zipf 分布, 公式建模如下:

$$g_d = \frac{1/d^\varepsilon}{\sum_{d=1}^D 1/d^\varepsilon} \quad (5)$$

式中 ε 为内容流行度参数, 范围在 $[0.5, 1.5]$ 之间。通过缓存内容 A'_m 节省回传带宽的成本可建模为:

$$l_{A'_m} = \frac{g_{A'_m} A'_m}{T_{A'_m}} \quad (6)$$

式中 $T_{A'_m}$ 为下载请求缓存内容文件的持续时间。回传成本记作 γ_{m0} , ψ_{m0} 表示内容文件缓存的存储成本, 则宏蜂窝基站边缘云的缓存收益可建模为: $R_{m0}^{\text{cache}} = \psi_{m0} l_{A'_m} - \gamma_{m0} A'_m$ 。

与宏蜂窝基站边缘云类似, 小蜂窝边缘云 n 回传成本记作 γ_{mn} , ψ_{mn} 表示内容文件缓存的存储成本, 则宏蜂窝基站边缘云的缓存收益可以建模为: $R_{mn}^{\text{cache}} = \psi_{mn} l_{A'_m} - \gamma_{mn} A'_m$ 。

2 基于深度强化学习算法的边缘计算技术

2.1 深度强化学习算法

强化学习算法是机器学习领域的一个重要分支, 与通常的有监督学习和无监督学习不同, 该算法模拟智能体与环境之间的交互作用, 并从动态变化的环境中获取观察状态而采取下一步行动, 通过多次尝试适当的决策达到目标。强化学习包括实验状态、行为、回报函数和状态转换概率 4 个部分。 Q 学习算法是强化学习算法中应

用最广泛的算法之一，采用 Q 值表示强化学习的回报值，并通过不断的尝试和调整误差寻找最优行为。 Q 学习算法的决策过程基于马尔科夫决策过程，可以通过五元组 $\{s_i, a_i, P(s_i, a_i, s_{i+1}), R(s_i, a_i), Q(s_i, a_i)\}$ 表示。其中 s_i 表示状态， a_i 表示动作，概率 $P(s_i, a_i, s_{i+1})$ 用来帮助动作 a_i 将状态 s_i 转移到下一个状态 s_{i+1} ， $R(s_i, a_i)$ 表示处于状态 s_i 时动作 a_i 取得的立即回报， $Q(s_i, a_i)$ 表示状态 s_i 选择动作 a_i 的累积回报值。

假设状态 s_i 的累积回报函数记作 $V(s_i)$ ：

$$V(s_i) = R_i + \lambda R_{i+1} + \lambda^2 R_{i+2} + \dots \quad (7)$$

式中： R_i 为第 i 步的回报值； $\lambda(0 < \lambda < 1)$ 为折扣因子，用于评估历史状态回报值对未来回报的影响， λ 取值越大，表示训练的智能体与历史经验越相关；否则，与立即回报越相关。在 Q 学习算法中， Q 值为状态和动作的评估，表示为：

$$Q(s_i, a_i) = R_i + \lambda V(s_{i+1}) \quad (8)$$

基于以上公式，第 $i+1$ 步的 $Q_{i+1}(s_i, a_i)$ 的更新如式(9)所示，并通过学习率 η 控制学习速率。

$$Q_{i+1}(s_i, a_i) = (1 - \eta)Q_i(s_i, a_i) + \eta[R_i + \lambda \max_{a'} Q_i(s_{i+1}, a_{i+1})] \quad (9)$$

深度 Q 网络的思想是利用前馈人工神经网络逼近 Q 值函数 $Q(s, a; \theta)$ ，与一次计算一个状态动作对 (s, a) 的 Q 值不同，深度 Q 网络中的 Q 值可与紧密的网络结构同时计算。 Q 网络的输入层为状态 s ，输出层为 Q 值，对应于在状态 s 执行的动作 a 。参数 θ 代表神经网络的权值， Q 网络通过最小化损失函数更新 θ 进行训练：

$$L(\theta) = E \left[\left(y(s, a, s'; \hat{\theta}) - Q(s, a; \theta) \right)^2 \right] \quad (10)$$

式中目标函数 $y(s, a, s'; \hat{\theta}) = R + \lambda \max_{a'} Q(s', a'; \hat{\theta})$ ，根据参数 $\hat{\theta}$ 的变化而更新。

2.2 基于深度强化学习算法的边缘计算技术

在联合试验环境网络场景下，上行信道条件和计算能力都在动态变化，导致系统状态和行为空间很大，传统的方法很难找到最优解。深度强化学习算法无需提前建立良好模型，能够根据环境状态的变化自适应调整改进策略。本节引入深度 Q 学习算法，根据作战任务是否缓存在边缘云中，提出不同策略。下面详细介绍深度 Q 学习算法模型状态、行为和回报函数的设计。

1) 状态空间。不同时刻的动作 a_m^{comm} 和 a_m^{cache} 决定不同的系统状态。

深度 Q 学习算法

输入：状态 s ，动作 a ；

输出： $Q(s, a)$ ；

初始化：

根据参数 θ 初始化深度 Q 网络；

Begin

For $i < T$ do

If 概率 $P < \delta$ then

选择动作 a_i ；

Else

$a_i = \operatorname{argmax} Q(s, a; \theta)$ ；

End If

执行动作 a_i ，获得回报函数值 R_i ，下一代状态 s_{i+1} ；

计算目标 Q 值；

$y(s, a, s'; \hat{\theta}) = R + \lambda \max_{a'} Q(s', a'; \hat{\theta})$ ；

根据式(10)最小化损失值 $L(\theta)$ ，更新神经网络；

End For

End

2) 动作空间。在联合试验边缘云系统中，智能体根据有限的通信资源确定任务卸载到边缘云的位置，不管

卸载的任务是否已经缓存在服务器中。 t 时刻的动作 $a_m(t)$ 定义为:

$$a_m(t) = \{a_m^{\text{comm}}(t), a_m^{\text{cache}}(t)\} \quad (11)$$

其中 $a_m^{\text{comm}}(t)$ 和 $a_m^{\text{cache}}(t)$ 的具体描述为:首先,定义行向量 $a_m^{\text{comm}}(t) = [a_{m,1}^{\text{comm}}(t), a_{m,2}^{\text{comm}}(t), \dots, a_{m,N}^{\text{comm}}(t)]$, $a_{m,i}^{\text{comm}}(t)$, $i \in \{1, 2, \dots, N\}$ 表示终端设备 m 是否连接到边缘云。 $a_{m,i}^{\text{comm}}(t)$ 的取值为 $\{0, 1\}$, $a_{m,i}^{\text{comm}}(t)=0$ 表示在时刻 t 终端设备 m 将任务卸载到宏蜂窝基站边缘云;否则, $a_{m,i}^{\text{comm}}(t)=1$ 表示在时刻 t 终端设备 m 将任务卸载到小蜂窝基站边缘云。然后,定义行向量 $a_m^{\text{cache}}(t) = [a_{m,1}^{\text{cache}}(t), a_{m,2}^{\text{cache}}(t), \dots, a_{m,N}^{\text{cache}}(t)]$, 其中 $a_{m,j}^{\text{cache}}(t)$, $j \in \{1, 2, \dots, N\}$ 表示终端设备 m 是否缓存在边缘云中。 $a_{m,j}^{\text{cache}}(t)$ 的取值为 $\{0, 1\}$, $a_{m,i}^{\text{cache}}(t)=0$ 表示在时刻 t 终端设备 m 的内容文件没有缓存在边缘云;否则, $a_{m,i}^{\text{cache}}(t)=1$ 表示在时刻 t 终端设备 m 的内容文件缓存在边缘云。

3) 回报函数。边缘云系统的回报函数设计目标是最大化通信模型、计算模型和缓存模型的回报值。终端设备 m 的回报函数建模为:

$$R_m(t) = R_m^{\text{comm}}(t) + R_m^{\text{cache}}(t) \quad (12)$$

式中: $R_m^{\text{comm}}(t)$ 表示从边缘云中通信模型所获得的回报值,包括宏蜂窝基站边缘云和小蜂窝基站边缘云中通信模型回报值; $R_m^{\text{cache}}(t)$ 表示内容文件是否缓存在边缘云中所获得的回报值,包括宏蜂窝基站边缘云和小蜂窝基站边缘云中的计算模型回报值和缓存模型回报值。

$$R_m^{\text{comm}}(t) = [1 - a_m^{\text{comm}}(t)]R_{m0}^{\text{comm}}(t) + a_m^{\text{comm}}(t)R_{mn}^{\text{comm}}(t) = [1 - a_m^{\text{comm}}(t)](\alpha_{m0}r_{m0} - \beta_{m0}w_{m0}) + a_m^{\text{comm}}(t)(\alpha_{mn}r_{mn} - \beta_{mn}w_{mn}) \quad (13)$$

$$R_m^{\text{cache}}(t) = [1 - a_m^{\text{comm}}(t)]a_m^{\text{cache}}(t)R_{m0}^{\text{cache}} + [1 - a_m^{\text{comm}}(t)][1 - a_m^{\text{cache}}(t)]R_{m0}^{\text{comp}} + a_m^{\text{comm}}(t)a_m^{\text{cache}}(t)R_{mn}^{\text{cache}} + a_m^{\text{comm}}(t)[1 - a_m^{\text{cache}}(t)]R_{mn}^{\text{comp}} =$$

$$[1 - a_m^{\text{comm}}(t)]a_m^{\text{cache}}(t)(\psi_{m0}l_{A'_m} - \gamma_{m0}A'_m) + [1 - a_m^{\text{comm}}(t)][1 - a_m^{\text{cache}}(t)](\phi_{m0}q_{m0} - \varphi_{m0}e_{m0}) +$$

$$a_m^{\text{comm}}(t)a_m^{\text{cache}}(t)(\psi_{mn}l_{A'_m} - \gamma_{mn}A'_m) + a_m^{\text{comm}}(t)[1 - a_m^{\text{cache}}(t)](\psi_{mn}l_{A'_m} - \gamma_{mn}A'_m) \quad (14)$$

3 数值实验

在Anaconda3 Spyder上仿真和评估基于深度强化学习的边缘计算技术。

3.1 参数设置

网络仿真场景由宏蜂窝基站边缘云、小蜂窝基站边缘云和终端设备组成。宏蜂窝基站边缘云数设为1,小蜂窝基站边缘云数设为10,并提前设定在固定位置。终端用户在宏蜂窝基站覆盖范围内随机分布,仿真场景如图2所示。

终端设备至宏蜂窝基站的传输信道带宽设为10 MHz,终端设备至小蜂窝基站的传输信道带宽设为1 MHz;终端设备的传输功率在50~100 mW范围内随机产生;高斯噪声的方差设为 $\sigma^2 = -100$ dBm;干扰因子为-100 dBm;通道增益 $h = d^{-\chi}$,其中 χ 表示通过损耗因子,设为4。表1为与此实验有关的其他参数的更多详细信息。

该方案由1个输入层、2个隐藏层和1个输出层组成。同时,第一隐层设置为120个隐神经元,第二隐层设置为80个。训练间隔设置为10,训练批大小为128。此外,Adam优化器的学习率为0.005,内存大小为1 024。

3.2 仿真实验

对基于深度强化学习的边缘计算技术的实验效果进行验证并研究参数对仿真结果的影响。图3为不同设备数量下的回报函数值变化,设备数量分别设为 $\{5, 10, 20\}$ 。通过图3可以发现,设备数量为5时,收敛后的回报函数值最小;设备数量为20时,回报函数值最高,但收敛速度在350代左右才开始迅速增长,略晚于设备数量为10的函数曲线。

研究深度强化学习算法中3个重要参数(学习率、批量大小和训练间隔)对仿真结果的影响。学习率控制模型的学习过程,学习率取值过高,可能导致快速收敛;学习率过低,则容易带来过度拟合。批量大小影响模型的

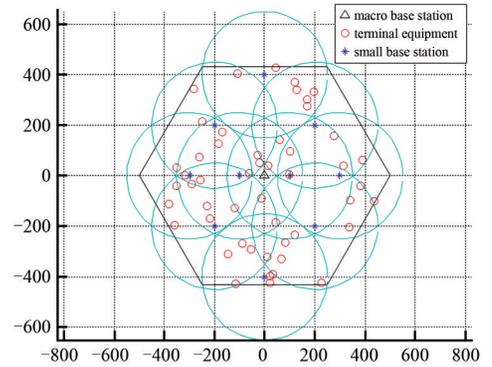


Fig.2 Experimental simulation scenario

图2 实验仿真场景

优化程度和速度，批量大小的值太小，算法不会在 epoches 中收敛；否则，很容易陷入局部收敛。同样，训练间隔也是深部神经网络的一个重要参数。

表 1 参数设置

Table 1 Parameter setting

parameter	value
access fee charged by the macro edge cloud/(unit/bps)	3
access fee charged by the small edge cloud n /(unit/bps)	1
the usage cost of edge cloud payment for macro edge cloud/(unit/Hz)	3×10^{-4}
the usage cost of edge cloud payment for small edge cloud n /(unit/Hz)	1×10^{-4}
compute fee charged by the macro edge cloud/(unit/J)	0.8
compute fee charged by the small edge cloud n /(unit/J)	0.4
the compute cost of edge cloud payment for macro edge cloud/(unit/J)	0.2
the compute cost of edge cloud payment for small edge cloud n /(unit/J)	0.1
storage fee charged by the macro edge cloud/(unit/byte)	20
storage fee charged by the small edge cloud n /(unit/byte)	10
the reward cost of edge cloud payment for macro edge cloud/(unit/bps)	0.2
the reward cost of edge cloud payment for small edge cloud n /(unit/bps)	0.1

图 4 为学习率对仿真结果的影响，学习率设置为 {0.005,0.01,0.05,0.1}。当学习率为 0.01 时，深度强化学习算法所需的回报函数值在整个迭代过程中表现得比其他算法好；学习率为 0.1 时，不能达到学习率为 0.01 的效果，但性能优于其他两种情况；学习率为 0.05 时的性能最差。

图 5 为批量大小对仿真结果的影响，批量大小设置为 {32,64,128,256}。当批量大小等于 32 时，深度强化学习算法所要求的回报函数值在前期阶段的表现不好，但后期增长迅速，并最终取得了最好的性能；批量大小为 64 时，性能略低于批量大小为 32 的情况；批量大小为 128 和批量大小为 256 时，表现相对较差。

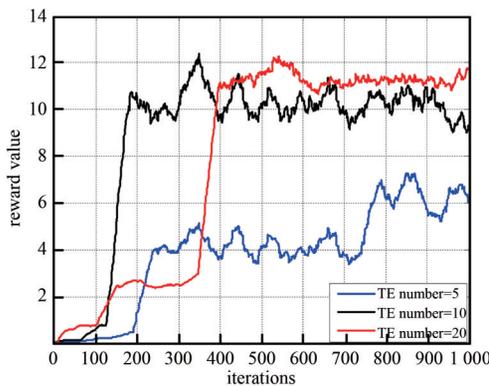


Fig.3 Change of reward function value under different number of equipment
图 3 不同设备数量的回报函数值变化

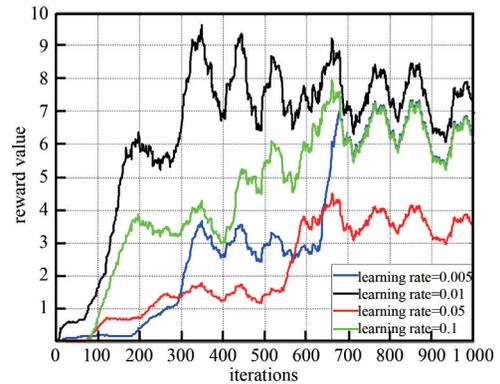


Fig.4 Change of reward function value under different learning rates
图 4 不同学习率的回报函数值变化

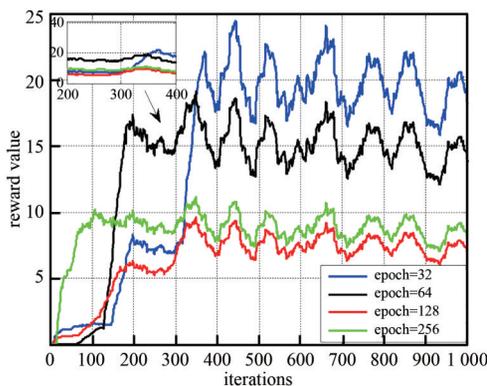


Fig.5 Variation of reward function value under different batch sizes
图 5 不同批量大小的回报函数值变化

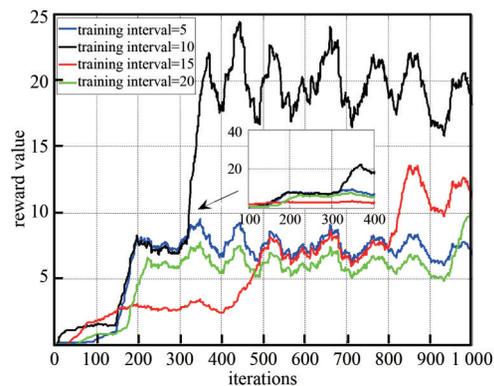


Fig.6 Variation of reward function value under different training intervals
图 6 不同训练间隔的回报函数值变化

图 6 为不同训练间隔随着迭代次数的回报函数值变化，训练间隔设置为 {5,10,15,20}。显然，当训练间隔为 10 时，深度强化学习算法所需的回报函数值在整个训练阶段表现良好；当训练间隔为 {5,15,20} 时，回报函数值表现相对较差，尤其是在训练间隔为 5 时，算法表现最差。

4 结论

针对智能无人系统中新技术的应用,本文提出了一种智能无人系统试验鉴定环境中的边缘计算技术。在该系统环境的网络架构中,所提出的算法能够根据边缘计算中通信资源、计算资源和缓存资源自适应地提供有效的边缘计算策略。仿真结果表明,该方案可以快速收敛到满意解,同时验证了该方案的不同参数对实验结果的影响。

参考文献:

- [1] 高富营. 军事智能论[M]. 北京:国防大学出版社, 2005. (GAO Fuying. Military intelligence theory[M]. Beijing: National Defense University Press, 2005.)
- [2] Massachusetts Institute of Technology. MIT technology review[EB/OL]. [2021-11-28]. <https://www.technologyreview.com/>.
- [3] 林聪榕. 智能化无人作战系统及其关键技术[J]. 国防技术基础, 2014(5):30-34. (LIN Congrong. Intelligent unmanned combat system and its key technologies[J]. Technology Foundation of National Defence, 2014(5):30-34.)
- [4] 赵欣怡,宗群,张睿隆,等. 类脑智能技术在无人系统上的应用[J]. 控制理论与应用, 2019,36(1):1-12. (ZHAO Xinyi,ZONG Qun,ZHANG Ruilong,et al. Application of brain-inspired intelligence technology in unmanned vehicles[J]. Control Theory & Applications, 2019,36(1):1-12.) doi:10.7641/CTA.2018.80418.
- [5] 范宇,赵书安. 基于ToA和误差抑制的大规模网络节点定位[J]. 太赫兹科学与电子信息学报, 2021,19(6):1020-1026. (FAN Yu,ZHAO Shu'an. Node localization based on time-of-arrival and error mitigation in large-scale network[J]. Journal of Terahertz Science and Electronic Information Technology, 2021,19(6):1020-1026.) doi:10.11805/TKYDA2020169.
- [6] 孙毅,宋乐,王桂芝. 2019年国外地面无人自主系统发展综述[J]. 飞航导弹, 2020(1):30-34. (SUN Yi,SONG Le,WANG Guizhi. Overview of the development of ground unmanned autonomous systems abroad in 2019[J]. Aerodynamic Missile Journal, 2020(1):30-34.) doi:10.16338/j.issn.1009-1319.20190868.
- [7] 张涛,李清,张长水,等. 智能无人自主系统的发展趋势[J]. 无人系统技术, 2018,1(1):11-22. (ZHANG Tao,LI Qing,ZHANG Changshui,et al. Current trends in the development of intelligent unmanned autonomous systems[J]. Unmanned Systems Technology, 2018,1(1):11-22.) doi:10.19942/j.issn.2096-5915.2018.01.002.

作者简介:

李飞翔(1991-),男,博士,工程师,主要研究方向为边缘计算、试验训练 .email:lifexiang126@126.com.

贺思源(1992-),男,硕士,工程师,主要研究方向为推荐算法、试验训练.

李 宁(1986-),男,硕士,高级工程师,主要研究方向为试验训练、边缘计算.

孙 甜(1993-),女,硕士,工程师,主要研究方向为边缘计算.

刘明哲(1987-),男,硕士,高级工程师,主要研究方向为试验训练.

(上接第79页)

- [10] 陈雷雨,周志杰,唐帅文,等. 融合多元信息的武器装备性能评估方法[J]. 系统工程与电子技术, 2020,42(7):1527-1533. (CHEN Leiyu,ZHOU Zhijie,TANG Shuaiwen,et al. Weapon equipment performance evaluation method with multi-information fusion[J]. Systems Engineering and Electronics, 2020,42(7):1527-1533.) doi:10.3969/j.issn.1001-506X.2020.07.14.
- [11] 王玉恒,杜太焦,刘峰,等. 高能激光武器系统效能评估方法[J]. 四川兵工学报, 2008,29(2):1-3,12. (WANG Yuheng,DU Taijiao,LIU Feng,et al. Effectiveness evaluation method of high energy laser weapon system[J]. Journal of Sichuan Ordnance, 2008,29(2):1-3,12.)
- [12] 杨菁,张小民,胡东霞,等. 高功率激光装置总体设计的综合评估方法[J]. 强激光与粒子束, 2005,17(5):660-664. (YANG Jing,ZHANG Xiaomin,HU Dongxia,et al. Comprehensive evaluation for system design of high-power laser facility[J]. High Power Laser and Particle Beams, 2005,17(5):660-664.) doi:CNKI:SUN:QJGY.0.2005-05-006.
- [13] 李君雅,锁斌,王琳,等. 考虑故障模式相关性的模糊TOPSIS风险评价方法[J]. 太赫兹科学与电子信息学报, 2021,19(4):705-711. (LI Junya,SUO Bin,WANG Lin,et al. Fuzzy TOPSIS risk evaluation method considering failure modes correlation[J]. Journal of Terahertz Science and Electronic Information Technology, 2021,19(4):705-711.) doi:10.11805/TKYDA2019339.

作者简介:

葛成良(1971-),男,博士,副研究员,主要研究方向为激光系统工程技术 .email:2692843634@qq.com.