基于NOMA的海洋物联网安全计算卸载

姜微，袁宵，王倩，钱丽萍
（浙江工业大学网络空间安全研究院，浙江 杭州 310014）

摘　要：针对海洋物联网（M-IoT, marine Internet of things）中存在多种恶意窃听设备（ED, eavesdropping device），为了确保无人水面艇（USV, unmanned surface vehicle）向高空平台（HAP, high altitude platform）的安全计算卸载，利用非正交多址接入（NOMA, non-orthogonal multiple access）辅助传输机制，将一组空闲的 USV 充当干扰 ED 窃听的干扰 USV（JU, jamming USV），与传输 USV（TU, transmitting USV）形成 NOMA 集群。考虑能耗约束和安全传输等要求，以最小化系统最大任务处理时延为目标，对 TU 的卸载比率、传输功率、计算资源分配和 NOMA 集群选择进行联合优化。为了解决混合整数非凸优化问题，提出了深度确定性策略梯度（DDPG, deep deterministic policy gradient）和交叉熵结合的算法。仿真结果显示，所提算法有效地降低了最大系统任务处理时延，并且保证了系统的安全性。

关键词：海洋物联网；移动边缘计算；非正交多址接入；物理层安全
中图分类号：TN929.5
文献标志码：A
doi：10.11959/j.issn.2096−3750.2024.00414

NOMA-based secure computation offloading in marine Internet of things

JIANG Wei, YUAN Xiao, WANG Qian, QIAN Liping
Institute of Cyberspace Security, Zhejiang University of Technology, Hangzhou 310014, China

Abstract: Regarding the presence of multiple malicious eavesdropping device (ED) in the marine Internet of things (M-IoT), to ensure the secure computation offloading of unmanned surface vehicle (USV) to a high altitude platform (HAP), a non-orthogonal multiple access (NOMA) assisted transmission policy was employed, where a set of idle USV acted as jamming USV (JU) that interfered with the eavesdropping of the ED and formed NOMA clusters with the transmitting USV (TU). Considering the requirements such as energy constraints and security transmission, the offloading ratio, transmitting power, computation resource allocation and NOMA cluster selection were jointly optimized with the objective of minimizing the maximum task processing latency. To solve the proposed mixed-integer non-convex optimization problem, an algorithm combining deep deterministic policy gradient (DDPG) and cross-entropy was proposed. Simulation results show that the proposed algorithm can effectively reduce maximum task processing latency and ensure the security of the system.

Key words: M-IoT, mobile edge computing, NOMA, physical layer security

0　引言

随着物联网技术的快速发展，海洋物联网（M-IoT, marine Internet of things）作为物联网技术在海洋领域的应用具有重要意义[1-3]。M-IoT 在海洋环境监测、海上勘探和航运安全方面发挥着至关重

收稿日期：2024−08−29；**修回日期：**2024−09−15
通信作者：钱丽萍，lpqian@zjut.edu.cn
基金项目：国家自然科学基金项目（No. 62122069，No. 62071431，No. 62302450）；浙江省自然科学基金项目（No. LQ24F020037）
Foundation Items: The National Natural Science Foundation of China (No. 62122069, No. 62071431, No. 62302450), The Zhejiang Provincial Natural Science Foundation of China (No. LQ24F020037)

要的作用。近年来，无人水面艇（USV, unmanned surface vehicle）以低成本、自主性和灵活部署的优势在M-IoT中扮演着重要角色[4-5]。随着M-IoT的发展，USV生成大量计算密集型任务，需要大量的通信和计算资源。但是，USV有限的传输功率和与陆地基站较长的传输距离抑制了无线传输的性能，导致传输延迟和成本的增加。

为了解决上述问题，研究者们在6G网络中提出了移动边缘计算（MEC, mobile edge computing）。MEC允许终端设备将计算任务卸载到部署在靠近终端设备的网络边缘服务器上，例如在高空平台（HAP, high altitude platform）上部署的MEC服务器，通过缩短传输距离和扩展终端设备的计算能力来降低传输延迟和成本消耗[6-8]。由于USV和MEC服务器之间的通信会产生额外的开销，因此卸载策略尤为重要。无线资源和计算资源的合理配置直接影响卸载策略的性能。

目前，大多数关于MEC卸载的研究[9-11]都集中在计算卸载决策和计算资源的联合优化。文献[9]研究了MEC网络中异构程序的卸载选择问题，并提出了一种联盟博弈机制，以更好地降低延迟。文献[10]联合优化了MEC网络中的计算卸载策略和无人机轨迹，以实现用户与无人机之间的实时通信，并最大限度地降低总能量和延迟。文献[11]研究了单无人机辅助下的USV在不同任务类型的卸载策略和资源分配方案中的最小化能耗，并扩展至多无人机的多接入情况。但是，上述工作都忽略了计算卸载过程中的安全问题。

由于无线通信的开放性，M-IoT中的USV在进行计算卸载的过程中容易受到窃听攻击，恶意实体可以对无线信号进行拦截和解密，造成信息泄露[12-13]。因此，USV的计算卸载安全性已成为备受关注的问题。过去通常使用加密方法来防止窃听设备（ED, eavesdropping device）对信息进行任意解码。然而，随着各种设备计算性能的飞速发展，仅靠密钥已经很难保障信息的安全。为了解决这一问题，许多研究探索了物理层安全（PLS, physical layer security）的潜力。PLS理论利用无线信道固有的物理特性，从信息论的角度提供了一种无法被任意恶意节点窃听的吞吐量的基本度量，从而增强了网络的保密能力[14-17]。

非正交多址接入（NOMA, non-orthogonal multiple access）作为第三代合作伙伴项目（3GPP, Third Generation Partnership Project）认可的5G新技术，对于提高保密能力、降低系统延迟和能耗有重要意义。NOMA允许多个用户在同一频带信道上同时传输数据，并使用连续干扰消除（SIC, successive interference cancellation）减轻信道间干扰[18-21]。在这种背景下，许多研究[22-24]探索了利用NOMA技术在PLS方面增强MEC卸载安全性的研究。文献[22]研究了块坐标下降和逐次凸近似算法，以提高基于NOMA的无人机−MEC网络计算的安全性，该方案最大限度地提高了网络的安全计算能力。文献[23]考虑了MEC在增强机器类型通信能力方面的作用，在存在ED的情况下，提出了一种基于NOMA的MEC系统，该系统进一步使用保密率来衡量卸载的安全性。文献[24]在考虑安全通信和能耗的情况下，提出了一种基于混合协作NOMA的安全边缘计算传输方法，这种方法可以根据不同用户的信道条件设置卸载决策，保证了用户间的公平。

上述研究表明，通过采用NOMA通信可以实现更高级别的MEC卸载数据加密和访问控制，从而提高系统的安全性。然而，现有的研究暂时没有在复杂M-IoT中利用NOMA技术在考虑信息安全和系统能耗的情况下降低任务处理时延的方案。因此，本文研究M-IoT中USV的安全计算卸载问题。具体而言，本文考虑在多个ED试图窃听信息的情况下，多个USV将数据信息卸载到HAP上部署的MEC服务器中。本文将空闲的USV作为干扰USV（TU, jamming USV）与传输USV（TU, transmitting USV）组成的NOMA集群，发射干扰信号，从而干扰ED对数据的接收，达到安全计算卸载的目的。本文的主要贡献如下。

1) 考虑存在多个ED的M-IoT中USV向HAP的安全计算卸载问题，其中空闲的USV充当JU与TU形成NOMA集群，用于增强ED接收到的干扰信号的强度，以保护卸载的信息免受窃听。本文建立了一个联合优化问题，通过优化TU的卸载比率、发射功率、计算资源分配和NOMA集群选择，最小化最大系统延迟。

2) 提出了一种深度确定性策略梯度（DDPG, deep deterministic policy gradient）和交叉熵结合的算法，对所提的联合优化问题进行求解。具体而言，将联合优化问题分解成两个子问题，即NOMA集群

选择子问题和计算卸载与资源分配子问题。NOMA集群选择子问题是根据DDPG优化的值提出的一种基于交叉熵的目标选择学习（CEtsL, cross-entropy-based target selection learning）算法。计算卸载与资源分配子问题在给定的NOMA集群选择的情况下使用DDPG联合优化TU的发射功率、TU的卸载比率和MEC服务器上的计算资源分配。两个子问题所使用的算法相互迭代得到优化目标。

3) 提供的综合数值结果验证了所提算法的有效性。仿真结果表明，所提算法在收敛性和性能方面优于深度Q网络（DQN, deep Q-network）算法和近端策略优化（PPO, proximal policy optimization）算法。与完全本地计算、完全边缘计算和PPO算法方案相比，所提算法在不同网络参数下均表现出良好的性能优势。特别地，与完全本地计算相比，所提算法能够降低34.6%的任务处理延迟。

1　系统模型

M-IoT中USV的安全计算卸载模型如图1所示。海面上的USV分为有计算密集型任务的TU和空闲的JU，其中，JU可以向ED发送干扰信号来抑制ED的窃听。假设海面上有I个TU，所有TU的集合表示为$\mathcal{I}=\{1,2,\cdots,I\}$，由于TU的计算能力有限，因此，本文将计算任务适当地卸载到部署有MEC服务器的HAP中，从而减少任务处理延迟。为每个TU选择一些JU形成NOMA集群来保护TU的数据传输。JU的集合表示为$\mathcal{J}=\{1,2,\cdots,J\}$。假设海面存在$N$个ED窃听TU传输的数据，ED的集合表示为$\mathcal{N}=\{1,2,\cdots,N\}$。此外，NOMA组之间采用频分多址接入（FDMA, frequency division multiple access）技术保证NOMA集群间彼此正交互不干扰。

定义在三维笛卡尔坐标系中，HAP的位置为$o_{\mathrm{H}}=\{x_{\mathrm{H}},y_{\mathrm{H}},z_{\mathrm{H}}\}$，其中，$x_{\mathrm{H}}$、$y_{\mathrm{H}}$和$z_{\mathrm{H}}$分别表示HAP在$x$轴、$y$轴和$z$轴的坐标。第$i$个TU的位置表示为$o_i^{\mathrm{u}}=\{x_i^{\mathrm{u}},y_i^{\mathrm{u}},0\}$，第$j$个JU的位置表示为$o_j^{\mathrm{v}}=\{x_j^{\mathrm{v}},y_j^{\mathrm{v}},0\}$，第$n$个ED的位置为$o_n^{\mathrm{e}}=\{x_n^{\mathrm{e}},y_n^{\mathrm{e}},0\}$。此外，所有设备均使用单天线收发信息。

1.1　通信模型

由于海面空旷环境具有较大的视距，且在信号传输路径上不会有障碍物遮挡，信号损耗较低，信号传播条件有利于直接的视距（LoS, line-of-sight）链路通信。因此，HAP与USV之间采用LoS链路作为合适的信道模型[25]。令ζ_0表示参考距离为$d=1$ m的信道功率增益。因此，第i个TU到HAP的信道增益表示为$h_{i,\mathrm{H}}^{\mathrm{u}}=\zeta_0\cdot\|o_i^{\mathrm{u}}-o_{\mathrm{H}}\|^{-2}$，第$j$个JU到HAP的信道增益为$h_{j,\mathrm{H}}^{\mathrm{v}}=\zeta_0\cdot\|o_j^{\mathrm{v}}-o_{\mathrm{H}}\|^{-2}$。根据文献[26]可知，信号从USV传播到ED的信道为独立的瑞利衰落信道，因此从第i个TU和第j个JU到第n个ED的信道增益可以分别表示为$h_{i,n}^{\mathrm{u}}=\zeta_0\psi_0\cdot\|o_i^{\mathrm{u}}-o_n^{\mathrm{e}}\|^{-\epsilon}$和$h_{j,n}^{\mathrm{v}}=\zeta_0\psi_0\cdot\|o_j^{\mathrm{v}}-o_n^{\mathrm{e}}\|^{-\epsilon}$，其中，$\epsilon$表示路径损耗系数，$\psi_0$表示瑞利衰落值。

在上行链路中，HAP可以通过SIC对来自多个USV的叠加信号进行解码。本文考虑一个NOMA组中解码顺序为HAP优先解码JU，最后解码TU。假设第i个NOMA组的信道带宽为w_i，因此，从第

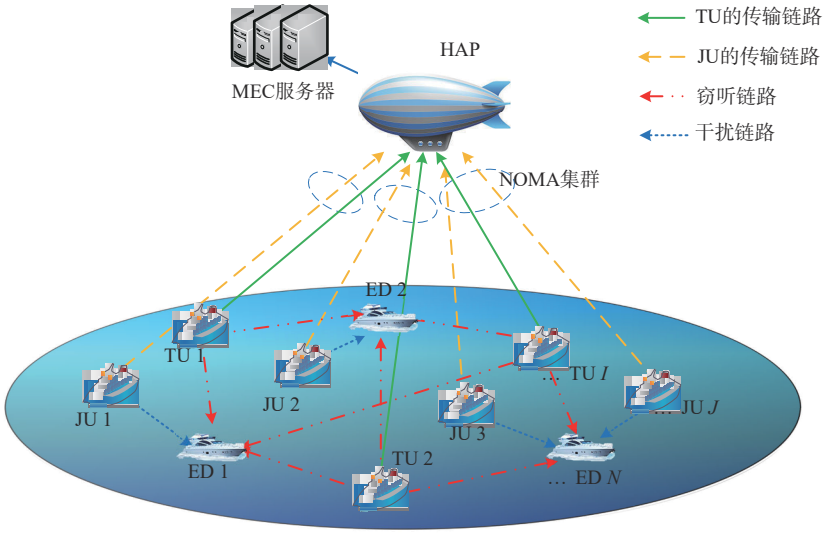

图1　M-IoT中USV的安全计算卸载模型

i个TU到HAP的信道吞吐量为

$$R_{i,\mathrm{H}} = w_i \,\mathrm{lb}\left(1 + \frac{p_i^{\mathrm{u}} h_{i,\mathrm{H}}^{\mathrm{u}}}{\sigma_0^2}\right), \forall i \in \mathcal{I} \tag{1}$$

其中，p_i^{u}是第i个TU的传输功率，σ_0^2是HAP处的噪声功率。

此外，定义$\boldsymbol{M} = \{\mu_{ij} \,|\, \mu_{ij} \in \{0,1\}, \forall i \in \mathcal{I}, \forall j \in \mathcal{J}\}$表示NOMA集群的选择，其中$\mu_{ij} = 1$表示第$i$个TU选择第$j$个JU组成NOMA组，否则$\mu_{ij} = 0$。由于每个JU只能被一个TU选择，因此可以得到如下约束

$$\sum_{i=1}^{I} \mu_{ij} \leqslant 1, \forall j \in \mathcal{J} \tag{2}$$

本文假设ED可以窃听到所有TU使用频段的信息，并且能够覆盖所有TU。此外，考虑ED是非串通窃听的。则第n个ED窃听第i个TU的卸载信息的吞吐量可以表示为

$$R_{i,n} = w_i \,\mathrm{lb}\left(1 + \frac{p_i^{\mathrm{u}} h_{i,n}^{\mathrm{u}}}{\sum_{j=1}^{J} \mu_{ij} p_j^{\mathrm{v}} h_{j,n}^{\mathrm{v}} + \sigma_n^2}\right), \quad \forall i \in \mathcal{I}, \forall n \in \mathcal{N} \tag{3}$$

其中，p_j^{v}是第j个JU的传输功率，σ_n^2是第n个ED的背景噪声功率。

本文考虑在最坏的窃听情况下的安全计算卸载场景，可以得出第i个TU到HAP的安全卸载吞吐量为

$$R_i^{\mathrm{sec}} = \left[R_{i,\mathrm{H}} - \max_{\forall n \in \mathcal{N}} R_{i,n}\right]^{+}, \forall i \in \mathcal{I} \tag{4}$$

1.2　计算卸载模型

本文考虑计算任务输入可以划分为任意大小的子集，并在本地或者HAP上并行执行。假设第i个TU需要处理的任务数据量为S_i（单位：bit），处理每比特数据所需的CPU周期数为c_i（单位：cycle/s）。引入向量$\boldsymbol{X} = \{x_i \,|\, 0 \leqslant x_i \leqslant 1, i \in \mathcal{I}\}$表示计算任务向HAP的卸载比率。因此，计算任务可以分为本地计算和远程执行两部分平行处理。

在本地计算中，第i个TU的本地计算能力设为f_i^{loc}，本地执行时间可以表示为

$$t_i^{\mathrm{loc}} = \frac{(1 - x_i) c_i S_i}{f_i^{\mathrm{loc}}}, \forall i \in \mathcal{I} \tag{5}$$

根据CPU功耗模型[27]，则本地计算能耗为

$$E_i^{\mathrm{loc}} = (1 - x_i) \eta_1 c_i S_i \left(f_i^{\mathrm{loc}}\right)^2 \tag{6}$$

其中，η_1表示本地计算中TU的CPU电容系数。考虑到TU的能量限制，本文假设TU的计算和通信所消耗的总能量不超过最大能量限制$E_i^{\max}$。

在远程执行中，假设HAP在成功接收到来自TU的所有输入数据时才会开始处理计算任务。因此，第i个TU的安全卸载延迟可以表示为

$$t_i^{\mathrm{off}} = \frac{x_i S_i}{R_i^{\mathrm{sec}}}, \forall i \in \mathcal{I} \tag{7}$$

第i个TU的安全卸载能耗为

$$E_i^{\mathrm{off}} = p_i^{\mathrm{u}} t_i^{\mathrm{off}}, \forall i \in \mathcal{I} \tag{8}$$

由于MEC服务器强大的任务处理能力可以为每个TU分配足够的计算资源。因此，假设f_i^{com}为分配给第i个TU的计算资源，满足如下约束

$$f_i^{\mathrm{com}} \geqslant 0, \forall i \in \mathcal{I} \tag{9}$$

$$\sum_{i=1}^{I} f_i^{\mathrm{com}} \leqslant F^{\max}, \forall i \in \mathcal{I} \tag{10}$$

在不失一般性的情况下，计算结果的返回值较小可以忽略不计。因此，第i个TU到HAP的计算延迟可以由式(11)给出

$$t_i^{\mathrm{com}} = \frac{x_i c_i S_i}{f_i^{\mathrm{com}}}, \forall i \in \mathcal{I} \tag{11}$$

由此，任务远程执行部分的安全计算卸载延迟可以表示为

$$t_i^{\mathrm{rem}} = t_i^{\mathrm{off}} + t_i^{\mathrm{com}}, \forall i \in \mathcal{I} \tag{12}$$

综上所述，第i个TU完成计算任务的延迟是本地和边缘计算延迟的最大值，如下所示

$$t_i = \max\{t_i^{\mathrm{loc}}, t_i^{\mathrm{rem}}\}, \forall i \in \mathcal{I} \tag{13}$$

1.3　问题建模

本文的目标是最小化系统完成计算任务的延迟，因此在保证系统计算卸载安全的情况下联合优化TU的卸载决策$\boldsymbol{X}$、TU的传输功率$\boldsymbol{P} = \{p_i^{\mathrm{u}}\}_{i \in \mathcal{I}}$、NOMA集群的选择$\boldsymbol{M}$和HAP的计算资源分配$\boldsymbol{F} = \{f_i^{\mathrm{com}}\}_{i \in \mathcal{I}}$，问题建模为

$$\mathrm{P1}: \min_{\boldsymbol{X},\boldsymbol{P},\boldsymbol{M},\boldsymbol{F}} \max_{\forall i \in \mathcal{I}} t_i \tag{14}$$

$$\text{s.t. } 0 \leqslant x_i \leqslant 1, \forall i \in \mathcal{I} \tag{14a}$$

$$\mu_{ij} \in \{0,1\}, \forall i \in \mathcal{I}, \forall j \in \mathcal{J} \tag{14b}$$

$$\sum\nolimits_{i=1}^{I} \mu_{ij} \leqslant 1, \forall j \in \mathcal{J} \tag{14c}$$

$$0 \leqslant p_i^{\mathrm{u}} \leqslant P_i^{\max}, \forall i \in \mathcal{I} \tag{14d}$$

$$E_i^{\mathrm{loc}} + E_i^{\mathrm{off}} \leqslant E_i^{\max}, \forall i \in \mathcal{I} \tag{14e}$$

$$f_i^{\mathrm{com}} \geqslant 0, \forall i \in \mathcal{I} \tag{14f}$$

$$\sum\nolimits_{i=1}^{I} f_i^{\mathrm{com}} \leqslant F^{\max} \tag{14g}$$

$$R_i^{\mathrm{sec}} \geqslant R_i^{\min}, \forall i \in \mathcal{I} \tag{14h}$$

其中，式(14a)为TU的计算任务卸载到HAP的比率；式(14b)和式(14c)是NOMA集群的选择约束，表示一个JU只能与一个TU组成NOMA组；式

(14d)表示TU传输功率的范围，不能超过最大功率限制$P_i^{\max}$；式(14e)限制了系统的能耗不能超过最大能耗$E_i^{\max}$；式(14f)和式(14g)表明分配给TU的计算资源总和的限定范围；约束(14h)表示TU的安全计算速率必须超过其阈值以确保安全传输。

2　优化问题求解

由于P1是一个混合整数非凸优化问题，包含二进制变量和多个耦合变量，采用传统的求解方法具有很大的复杂度。因此，本文将优化问题转化为两个子问题，即NOMA集群选择子问题和计算卸载与资源分配子问题，并通过交替求解两个子问题来获得P1的最优解。

2.1　计算卸载与资源分配子问题求解

假设已知NOMA集群选择，则优化$(\boldsymbol{X},\boldsymbol{P},\boldsymbol{F})$的计算卸载与资源分配问题P2可表示为

$$\text{P2:}\min_{\boldsymbol{X},\boldsymbol{P},\boldsymbol{F}}\ \hat{t} \tag{15}$$

s. t. 式(14a), 式(14d), 式(14e), 式(14f), 式(14g), 式(14h)

其中，$\hat{t}=\max_{i\in\mathcal{I}}t_i$。

可以看出，由于动态时变的无线通信环境和资源，包括无线信道状态和MEC服务器的计算能力，整个优化过程实际上是参与者与基于动态变化的无线通信资源环境之间的动态交互。因此，P2可以被建模为马尔可夫决策过程（MDP, Markov decision process）并通过深度强化学习（DRL, deep reinforcement learning）DDPG算法逼近最优解。DDPG算法是一种无模型策略算法，结合了策略梯度和价值函数的方法，利用深度神经网络（DNN, deep neural network）在连续动作空间中学习策略[28-29]，DDPG算法框架如图2所示。

MDP在时隙k下定义为元组$<s_k,a_k,r_k,s_{k+1}>$，其中，s_k表示状态，a_k表示动作，r表示在状态s_k下动作的即时奖励，s_{k+1}表示下一步状态。详细表述如下。

1) 状态空间$\mathcal{S}$：时隙k中的状态$s_k\in\mathcal{S}$由所有USV到HAP和ED的信道增益组成，表示为

$$s_k=\left\{\left\{h_{i,\mathrm{H}}^{\mathrm{u}}(k)\right\}_{i\in\mathcal{I}},\left\{h_{j,\mathrm{H}}^{\mathrm{v}}(k)\right\}_{j\in\mathcal{J}},\left\{h_{i,n}^{\mathrm{u}}(k)\right\}_{i\in\mathcal{I},j\in\mathcal{J}},\left\{h_{j,n}^{\mathrm{v}}(k)\right\}_{j\in\mathcal{J},n\in\mathcal{N}}\right\} \tag{16}$$

2) 动作空间$\mathcal{A}$：动作空间由3个部分组成，卸载决策、传输功率和计算资源分配。时隙k中的动作$a_k\in\mathcal{A}$定义为

$$a_k=\left\{\left\{x_i(k)\right\}_{i\in\mathcal{I}},\left\{p_i^{\mathrm{u}}(k)\right\}_{i\in\mathcal{I}},\left\{f_i^{\mathrm{com}}(k)\right\}_{i\in\mathcal{I}}\right\} \tag{17}$$

3) 奖励函数$\mathcal{R}$：由于本文的优化目标是最小化最大的任务处理延迟，我们将即时奖励$r_k\in\mathcal{R}$设置为负的最大任务处理延迟，也就是说，如果采用增

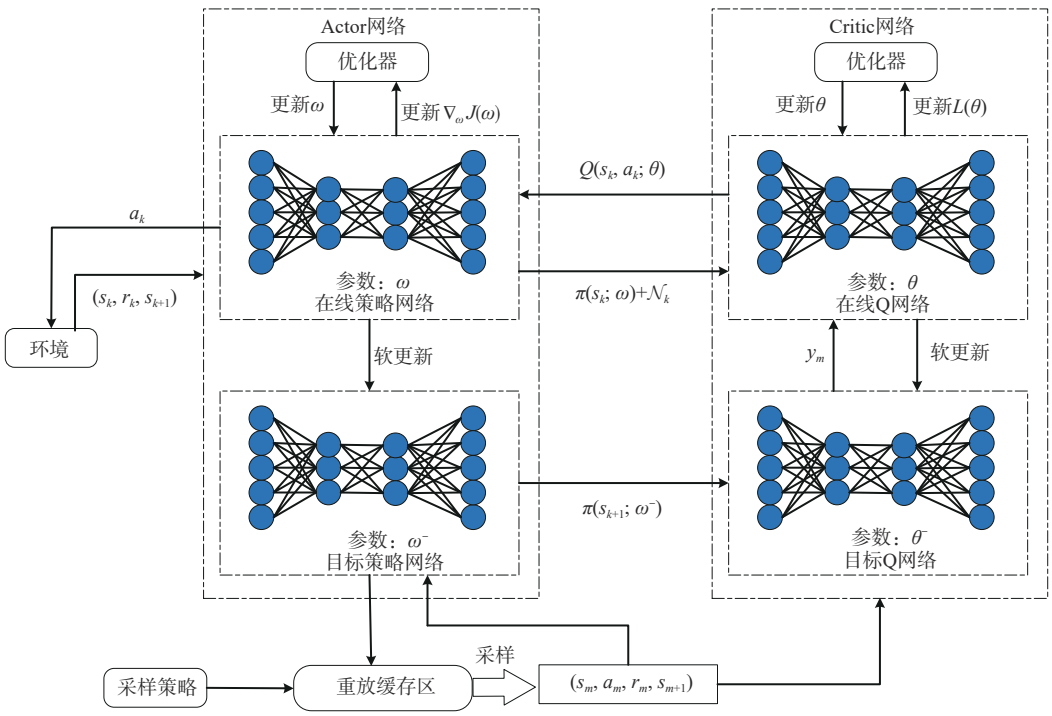

图2　DDPG算法框架

加奖励的行动，最大任务处理延迟将减少，定义为

$$r_k(s_k, a_k) = -\hat{t} \tag{18}$$

DDPG网络旨在寻找一种最优策略，该策略可以最大化长期折扣奖励R_k，定义为

$$R_k = \sum_{k=0}^{K} \gamma^k r_k \tag{19}$$

其中，$\gamma \in [0,1]$是决定当前和未来奖励平衡的折扣因子。如果γ很小，则专注于最大化当前奖励。当γ增加时，则更倾向于选择能够最大化未来回报的行动。本文中，各时隙之间的动作是相互独立的，因为每个动作的约束仅针对每个时隙设置，因此，本文更倾向于选择一个相对较小的γ来关注当前的奖励。

DDPG算法使用Actor-Critic算法作为基本框架，其中包含了一个Actor网络和一个Critic网络，双重神经网络的架构使得算法的学习过程更加稳定，收敛的速度加快。此外，每个网络都有对应的目标网络。

在Actor网络中，将状态s_k作为输入，根据权重ω和随机噪声将即时动作输出到环境中，即

$$a_k = \pi(s_k; \omega) + \mathcal{N}_k \tag{20}$$

其中，$\mathcal{N}_k$为探索噪声，可以平衡对新动作的探索和对先前动作的利用，可以防止神经网络陷入局部最优决策的情况；ω表示Actor网络的权重。首先，动作a_k确定后，Actor网络的策略$\pi(s_k; \omega)$输入Critic网络进行评估，通过Critic网络的权重θ输出估计Q值$Q(s_k,\pi(s_k; \omega); \theta)$。其次，利用策略梯度原理最大化累计期望回报$J(\omega) = Q(s_k, \pi(s_k; \omega); \theta)$，从而对Actor网络的参数进行更新

$$\nabla_\omega J(\omega) = \frac{1}{M}\sum_{m=1}^{M}\left[\nabla_a Q(s, a; \theta)|_{s_m,\pi(s_m)} \nabla_\omega \pi(s; \omega)|_{s_m}\right] \tag{21}$$

在Critic网络中，利用目标网络计算出(s,a)的目标Q值

$$y_m = r_m + \gamma Q\left(s_{m+1}, \pi(s_{m+1}; \omega^-); \theta^-\right) \tag{22}$$

其中，m是随机样本参数；ω^-和θ^-是Actor和Critic目标网络的权重。因此，可以利用梯度下降法最小化Critic网络的损失函数，定义为

$$L(\theta) = \left(y_m - Q(s_m, a_m; \theta)\right)^2 \tag{23}$$

对于目标网络，DDPG算法采用软更新方式使输出更加稳定，从而进一步保证Critic网络的学习过程更平稳，权重更新过程为

$$\omega^- \leftarrow \tau\omega + (1-\tau)\omega^- \tag{24}$$

$$\theta^- \leftarrow \tau\theta + (1-\tau)\theta^- \tag{25}$$

其中，$0 < \tau < 1$是超参数。综上所述，基于DDPG的计算卸载和资源分配算法如算法1所示。

算法1　基于DDPG的计算卸载和资源分配算法

初始化　随机初始化Actor网络和Critic网络的权重ω、θ，并赋值给相应的目标网络。初始化重放缓存区D，并将批量大小设置为M。初始化折扣系数γ，软更新参数τ，以及噪声分布以进行探索。

for episode = 1, 2, ⋯, N_{ep} **do**

重置算法的仿真参数，获得初始状态s_1；

for step k = 1, 2, ⋯, K **do**

输入状态s_k到Actor网络中，根据式(20)获得动作a_k；

由式(18)获得即时奖励r_t，并观察下一个状态s_{k+1}；

如果重放缓存区未满，则将经验元组$< s_k, a_k, r_k, s_{k+1} >$存储在重放缓存区$D$中；

否则用元组$< s_k, a_k, r_k, s_{k+1} >$在缓存区$D$中进行随意替换；

随机从缓存区D中抽取一个小批量元组$< s_m, a_m, r_m, s_{m+1} >$；

由式(22)计算目标Q值；

根据式(21)和式(23)更新Actor网络和Critic网络的权重；

根据式(24)和式(25)更新目标网络的权重；

end for

end for

2.2　NOMA集群选择子问题求解

在给定$\boldsymbol{M}$的情况下，基于算法1可以得到优化变量$(\hat{\boldsymbol{X}}, \hat{\boldsymbol{P}}, \hat{\boldsymbol{F}})_{(\boldsymbol{M})}$，我们继续求解$\boldsymbol{M}$以最小化系统的最大任务处理延迟$\hat{t}$，从而得到P1的最优解。优化$\boldsymbol{M}$的具体问题P3为

$$\text{P3: } \min_{\boldsymbol{M}} \ \hat{t} \tag{26}$$

s.t. 式(14b), 式(14c), 式(14e), 式(14h)

显然，P3属于整数非凸优化问题，包含一个二进制离散变量。为了解决P3，本文提出了CEtsL算法。具体来说，该算法主要是基于自适应概率学习的方法，利用交叉熵理论逐步改进样本分布来寻找$\boldsymbol{M}$的最优解，以最小化系统的最大任务处理延迟。通过将$\boldsymbol{M} = \{\mu_{ij} | \mu_{ij} \in \{0,1\}, \forall i \in \mathcal{I}, \forall j \in \mathcal{J}\}$建

模为随机的二元变量，P3 可以被视为学习这些变量的最优概率分布问题。为了解决这个问题，最直接的方法就是生成随机样本，并选择表现良好的样本。通过对这些样本进行概率学习，可以确定最小的系统最大任务处理延迟的概率分布，具体如下。

1) 将每个μ_{ij}建模为遵循伯努利分布的随机变量，概率为ϕ_{ij}，即

$$\Pr\left(\mu_{ij}\right)=\phi_{ij}^{\mu_{ij}}\left(1-\phi_{ij}\right)^{1-\mu_{ij}},\forall i\in\mathcal{I},\forall j\in\mathcal{J} \quad (27)$$

2) 根据假设的概率分布式(27)和式(14c)，随机生成$\mu_{11}^{(l)}\mu_{12}^{(l)}\cdots\mu_{1J}^{(l)}\cdots\mu_{ij}^{(l)}\cdots\mu_{IJ}^{(l)},\forall l\in\mathcal{L}=\{1,2,\cdots,L\}$的$L$个采样轮廓。

3) 对于第l个采样轮廓，通过DDPG算法得到相应的$\hat{t}$。然后，我们根据$\hat{t}$的结果按降序对采样轮廓进行排序，并选择前$\hat{L}$个最佳采样轮廓更新下一轮随机采样的$\{\phi_{ij}\}_{\forall i\in\mathcal{I},\forall j\in\mathcal{J}}$。

为了避免潜在的震荡，在每次迭代中更新概率分布参数为

$$\phi_{ij}^{\text{update}}=(1-\upsilon)\phi_{ij}+\upsilon\phi_{ij}^{*},\forall i\in\mathcal{I},\forall j\in\mathcal{J} \quad (28)$$

其中，υ表示权重参数，每个ϕ_{ij}^{*}满足

$$\phi_{ij}^{*}=\frac{1}{\hat{L}}\sum_{l=1}^{L}\mu_{ij}^{(l)},\forall i\in\mathcal{I},\forall j\in\mathcal{J} \quad (29)$$

4) 检查变化，并在变化小于阈值时终止迭代。当满足停止条件时，我们得到了P1的最优解。

基于CEtsL的NOMA集群选择算法如算法2所示。

算法2　基于CEtsL的NOMA集群选择算法

初始化　$\phi_{ij}=0.5$，$\forall i\in\mathcal{I}$，$\forall j\in\mathcal{J}$，并将间隙δ设置为一个较小的正数。

while 1 do

通过式(27)随机生成$\{\phi_{ij}\}_{\forall i\in\mathcal{I},\forall j\in\mathcal{J}}$的$L$个采样轮廓，且采样轮廓在式(14c)的约束下是可行的；

for $\mu_{11}^{(l)}\mu_{12}^{(l)}\cdots\mu_{1J}^{(l)}\cdots\mu_{ij}^{(l)}\cdots\mu_{IJ}^{(l)}$，$\forall l\in\mathcal{L}$ **do**

通过算法1求解P2，得到解$(\hat{\boldsymbol{X}}^{*},\hat{\boldsymbol{P}}^{*},\hat{\boldsymbol{F}}^{*},\hat{t}^{*})$；

end for

按降序对相应的$\hat{t}^{*}$的采样轮廓重新排序；

选择前$\hat{L}$个最佳采样轮廓并根据式(29)更新$\{\phi_{ij}^{\text{update}}\}_{\forall i\in\mathcal{I},\forall j\in\mathcal{J}}$；

if $\max_{\forall i\in\mathcal{I},\forall j\in\mathcal{J}}|\phi_{ij}^{\text{update}}-\phi_{ij}|<\delta$ **then**

break；

end if

设置$\phi_{ij}=\phi_{ij}^{\text{update}}$，$\forall i\in\mathcal{I}$，$\forall j\in\mathcal{J}$；

end while

2.3　计算复杂度分析

最后，本文讨论了所提算法的复杂度。本文的方法复杂度包括两个部分，一部分是解决P1的DDPG算法的复杂度，另一部分是CEtsL算法的复杂度。对于DDPG算法，假设Actor网络和Critic网络的层数分别为L_{A}和L_{C}，每层的单元数量分别为U_{A}和U_{C}。在训练过程中，假设DDPG神经网络参数更新W个时段，每个时段执行K次，则计算复杂度约为$O(WK(L_{\text{A}}U_{\text{A}}^{2}+L_{\text{C}}U_{\text{C}}^{2}))$。假设基于交叉熵的CEtsL算法每次生成$L$个样本，迭代步数为$F$，则计算复杂度为$O(LF)$。因此，所提算法的总计算复杂度为$O(WK(L_{\text{A}}U_{\text{A}}^{2}+L_{\text{C}}U_{\text{C}}^{2}+LF))$。

3　仿真结果与分析

本文通过一系列仿真来验证在M-IoT中USV的安全计算卸载场景下所提算法的性能优势。假设USV和ED随机分布在以(0 m, 0 m)为中心，半径为800 m的圆内，而HAP分布在高度为100 m、以(0 m, 0 m)为中心、半径为500 m的平面内。瑞利衰落的路径损耗指数设置为3.8[30]，TU的数量$I=4$，JU的数量$J=6$，ED的数量$N=4$。除有特殊说明，仿真详细参数设置见表1[31]。

表1　仿真详细参数设置

参数	参数值
带宽w_i	8 MHz
参考距离d=1 m的信道增益ζ_0	0.001 W
背景噪声σ_0^2,σ_n^2	1.0×10^{-13} W
处理每比特数据需要的CPU周期数c_i	1 000 cycle/s
TU的工作负载S_i	6 Mbit
HAP最大计算能力$F^{\max}$	50 GHz
TU的计算能力f_i^{loc}	1.2 GHz
TU的CPU电容系数η_1	1.0×10^{-27}
TU的最大功率$P_i^{\max}$	1 W
JU的发射功率p_j^{v}	0.5 W
TU的最大能量$E_i^{\max}$	5 J
安全传输的最小阈值$R_i^{\min}$	0.05
Actor网络的学习率α	0.001
Critic网络的学习率β	0.001
折扣系数γ	0.1
软更新参数τ	0.001

不同算法的收敛性能对比如图3所示。与全局优化求解器LINGO相比，所提算法达到了96%的最优性，另外，所提算法在计算时间表现上远优于

LINGO，可以快速适应参数的变化，有效地降低计算的复杂度。在相同的条件下，与DQN算法和PPO算法相比，所提算法的系统延迟在大约100次的训练后趋于平稳，是3种算法中最低的。而DQN算法在150次的训练后趋于稳定，由于DQN算法的连续动作离散化会限制可选择的动作，算法无法寻找到最优解。PPO算法在大约200次的训练时趋于平稳，这是因为PPO算法用于混合空间时，使用相对保守的更新策略，导致在高维状态空间中收敛速度较慢，短期奖励值较低。本文所提算法提供了比其他基线方法更高的回报，证明了其优越性。

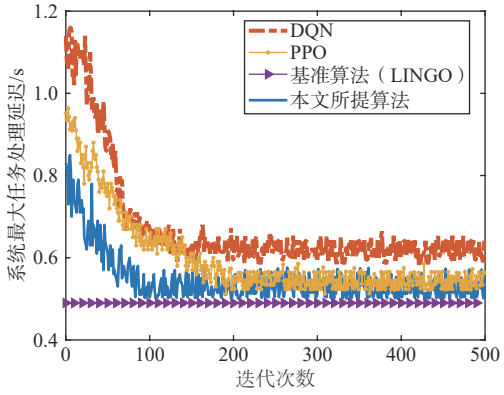

图3　不同算法的收敛性能对比

下面通过将所提算法与完全本地计算、完全边缘计算以及PPO算法进行对比，从而验证所提算法的性能。TU的工作负载与系统任务处理延迟的关系如图4所示，随着工作负载的增加，系统的延迟也逐渐增加。与其他基准算法相比，所提算法可以分别有效地将延迟降低34.6%、21.5%和3.8%。

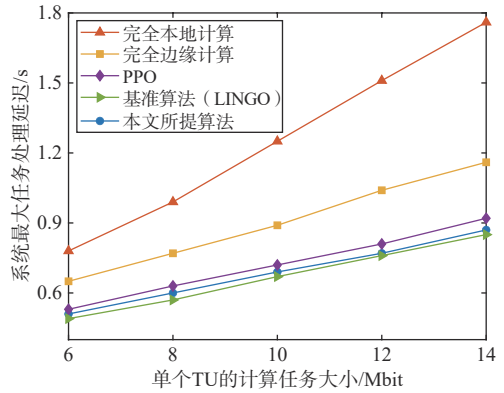

图4　TU的工作负载与系统任务处理延迟的关系

每个NOMA组的带宽与系统任务处理延迟的关系如图5所示。显然，系统最大任务处理延迟随着带宽的增加而减小，这是因为更多的带宽资源可以带来更多的通信自由度，从而缩短系统延迟。从图5可以看出，系统最大任务处理延迟在带宽较小时首先迅速下降，然后在带宽增加时缓慢下降，这是因为当带宽较小时，系统的性能受到带宽的限制，而当带宽逐渐增加时，系统性能的主要约束不再是带宽而是功率。

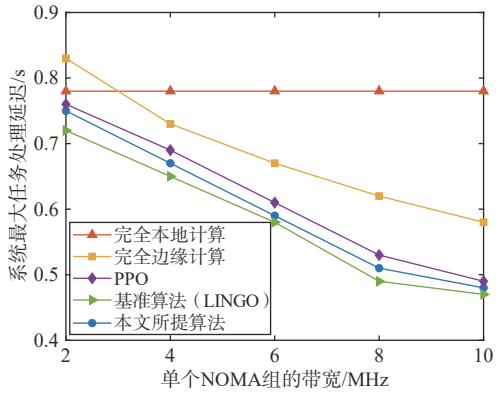

图5　每个NOMA组的带宽与系统任务处理延迟的关系

TU的传输功率上限与系统任务处理延迟的关系如图6所示。当TU的最大传输功率很小时，完全边缘计算比完全本地计算的传输延迟要大，这是由于每个TU的计算能力限制为1.2 GHz，完全可以保证本地计算的最大任务数。随着最大功率的增加，TU进行部分卸载可以有效地减少系统最大任务处理延迟。

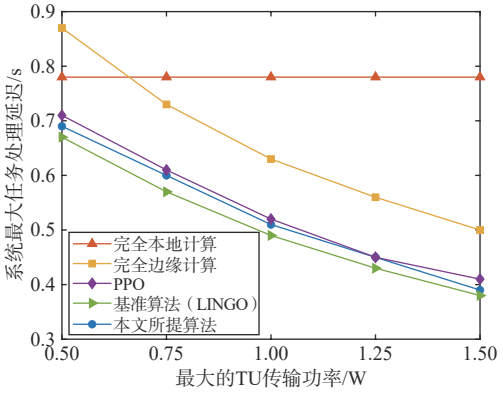

图6　TU的传输功率上限与系统任务处理延迟的关系

MEC计算能力的上限与系统任务处理延迟的关系如图7所示，系统最大任务处理延迟随着MEC计算能力的增加而降低。显然，随着MEC计算能力的增加，所提算法与完全边缘计算方案和PPO算法的系统最大任务处理延迟逐渐趋于一致。这是因为MEC计算能力的增加使系统优化卸载决策时偏向于将任务卸载到MEC服务器上进行计算。

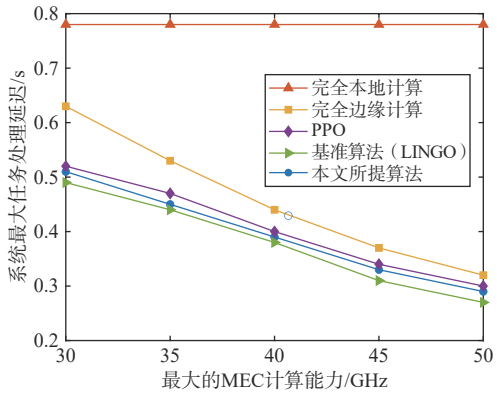

图7　MEC计算能力的上限与系统任务处理延迟的关系

4　结束语

本文研究了M-IoT中USV的安全计算卸载和资源分配问题。通过联合优化卸载决策、NOMA集群选择、传输功率和MEC计算资源分配来最小化系统的最大任务处理延迟。由于优化问题是一个混合整数非凸优化问题，本文提出了一种适用于混合动作空间的DDPG算法和交叉熵相结合的方法来解决所提出的问题。该方法的关键思想是将优化问题划分为计算卸载与资源分配和NOMA集群选择两个子问题，将两个子问题的解决算法相互迭代来逼近最优解。我们利用DDPG算法解决计算卸载与资源分配子问题，并使用交叉熵理论处理NOMA集群选择问题。仿真结果显示，所提算法可以快速收敛，并优于仅适用于离散动作的DQN算法和使用混合动作的PPO算法。在未来的工作中，将考虑USV的运动特性，在位置变化的情况下讨论USV的安全计算卸载和资源分配问题。

参考文献：

[1] AL-ZAIDI R, WOODS J, AL-KHALIDI M, et al. Next generation marine data networks in an IoT environment[C]//Proceedings of the 2017 Second International Conference on Fog and Mobile Edge Computing (FMEC). Piscataway: IEEE Press, 2017: 50-55.

[2] 蔡立鹏, 蒋海阳, 谢卓冉, 等. 海洋监测平台研究综述[J]. 电脑知识与技术, 2023, 19(36): 114-116.
CAI L P, JIANG H Y, XIE Z R, et al. Review on research of ocean monitoring platform[J]. Computer Knowledge and Technology, 2023, 19(36): 114-116.

[3] 瞿逢重, 付雁冰, 杨劭坚, 等. 应用于海洋物联网的水声通信技术发展综述[J]. 哈尔滨工程大学学报, 2023, 44(11): 1937-1949.
QU F Z, FU Y B, YANG S J, et al. An overview of the development status of underwater acoustic communication technology applied to ocean Internet-of-things[J]. Journal of Harbin Engineering University, 2023, 44(11): 1937-1949.

[4] LYU L, DAI Y P, CHENG N, et al. AoI-aware co-design of cooperative transmission and state estimation for marine IoT systems[J]. IEEE Internet of Things Journal, 2021, 8(10): 7889-7901.

[5] WANG Y, ZHENG Y, LIU J H. Secure task offloading and resource scheduling in maritime edge computing systems[C]//Proceedings of the 2023 IEEE/CIC International Conference on Communications in China (ICCC). Piscataway: IEEE Press, 2023: 1-6.

[6] HUANG P H, HSIEH F C, HSIEH W J, et al. Prioritized traffic shaping for low-latency MEC flows in MEC-enabled cellular networks[C]//Proceedings of the 2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC). Piscataway: IEEE Press, 2022: 120-125.

[7] 王翔. 移动边缘计算中分布式智能服务缓存和资源分配联合优化[J]. 重庆理工大学学报(自然科学), 2024, 38(8): 219-226.
WANG X. Joint optimization of distributed intelligent service caching and resource allocation in mobile edge computing[J]. Journal of Chongqing University of Technology (Natural Science), 2024, 38(8): 219-226.

[8] 杨守义, 陈怡航, 张双玲, 等. 面向未来移动通信的移动边缘计算研究综述[J]. 郑州大学学报(工学版), 2024, 45(4): 1-10, 29.
YANG S Y, CHEN Y H, ZHANG S L, et al. Research of mobile edge computing for future mobile communications: a review[J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(4): 1-10, 29.

[9] ZHAO J. Offloading selection based on heterogeneous utility in MEC networks: a coalition formation game-theoretic approach [C]//Proceedings of the 2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS), 2021: 469-472.

[10] GAN Y H, HE Y. Trajectory optimization and computing offloading strategy in UAV-assisted MEC system [C]//Proceedings of the 2021 Computing, Communications and IoT Applications (ComComAp), 2021: 132-137.

[11] 王婷. 无人机辅助海上MEC系统任务卸载与资源分配策略研究[D]. 长春：吉林大学, 2024.
WANG T. Research on task unloading and resource allocation strategy of UAV-assisted maritime MEC system [D]. Changchun: Jilin University, 2024.

[12] WU W, ZHOU F H, HU R Q, et al. Energy-efficient resource allocation for secure NOMA-enabled mobile edge computing networks[J]. IEEE Transactions on Communications, 2020, 68(1): 493-505.

[13] 薛建彬, 豆俊, 王涛, 等. 无人机辅助边缘计算安全通信能力最大化方案[J]. 计算机科学, 2024, 51(S1): 961-967.
XUE J B, DOU J, WANG T, et al. Scheme for maximizing security communication capability of UAV aided edge computing[J]. Computer Science, 2024, 51(S1): 961-967.

[14] WYNER A. The wire-tap channel[J]. The Bell System Technical Journal, 1975, 54: 1355-1387.

[15] 邓志祥, 戴陈庆, 张志威. 混合可重构智能表面和人工噪声辅助的物理层安全通信[J]. 电子与信息学报, 2024, 46(8): 3155-3164.
DENG Z X, DAI C Q, ZHANG Z W. Physical layer security for hybrid reconfigurable intelligent surface and artificial noise assisted communication[J]. Journal of Electronics & Information

Technology, 2024, 46(8): 3155-3164.

[16] YERRAPRAGADA A K, EISMAN T, KELLEY B. Physical-layer security for beyond 5G: ultra secure low latency communications [J]. IEEE Open Journal of the Communications Society, 2021, 2: 2232-2242.

[17] BAGHANI M, PARSAEEFARD S, DERAKHSHANI M, et al. Dynamic non-orthogonal multiple access and orthogonal multiple access in 5G wireless networks [J]. IEEE Transactions on Communications, 2019, 67(9): 6360-6373.

[18] 李美玲, 王玉旻, 王思敬, 等. STAR-RIS辅助的NOMA系统物理层安全性能优化[J]. 通信学报, 2024, 45(5): 214-225.
LI M L, WANG Y M, WANG S J, et al. Performance optimization of physical layer security in STAR-RIS aided NOMA system[J]. Journal on Communications, 2024, 45(5): 214-225.

[19] SARASWAT S K, SINGH D. Analysis of optimization of rate in power domain NOMA schemes for MIMO[C]//Proceedings of the 2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC). Piscataway: IEEE Press, 2020: 481-484.

[20] CHAUHAN A, GHOSH S, JAISWAL A. RIS partition-assisted non-orthogonal multiple access (NOMA) and quadrature-NOMA with imperfect SIC[J]. IEEE Transactions on Wireless Communications, 2023, 22(7): 4371-4386.

[21] 阔永红, 曹琳, 吕璐, 等. 全双工主动窃听非正交多址接入系统智能超表面辅助物理层安全传输技术[J]. 电子与信息学报, 2024, 46(3): 798-807.
KUO Y H, CAO L, LYU L, et al. Reconfigurable intelligent surfaces-aided physical layer secure transmission in non-orthogonal multi-access systems against full-duplex active eavesdropping[J]. Journal of Electronics & Information Technology, 2024, 46(3): 798-807.

[22] GAO Y, GUO Y, WANG P, et al. Secure enhancement in NOMA-based UAV-MEC networks[C]//Proceedings of the IEEE INFOCOM 2022 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Piscataway: IEEE Press, 2022: 1-6.

[23] ZHOU Y, SUN H J, MA X, et al. Energy-efficient secure offloading for NOMA-enabled machine-type mobile-edge computing[C]//Proceedings of the 2023 IEEE International Conference on Industrial Technology (ICIT). Piscataway: IEEE Press, 2023: 1-6.

[24] 余雪勇, 傅新程, 朱洪波. 基于混合协作NOMA的安全MEC能耗优化[J]. 系统工程与电子技术, 2024, 46(3): 1116-1124.
YU X Y, FU X C, ZHU H B. Optimization of energy consumption with hybrid cooperative NOMA for secure MEC[J]. Systems Engineering and Electronics, 2024, 46(3): 1116-1124.

[25] NOMIKOS N, GKONIS P K, BITHAS P S, et al. A survey on UAV-aided maritime communications: deployment considerations, applications, and future challenges[J]. IEEE Open Journal of the Communications Society, 2023, 4: 56-78.

[26] CONSUL P, BUDHIRAJA I, GARG D. A hybrid secure resource allocation and trajectory optimization approach for mobile edge computing using federated learning based on WEB 3.0[J]. IEEE Transactions on Consumer Electronics, 2024, 70(1): 1167-1179.

[27] LI Y X, WANG W, LIU M Q, et al. Joint trajectory and power optimization for jamming-aided NOMA-UAV secure networks[J]. IEEE Systems Journal, 2023, 17(1): 732-743.

[28] ZHAO T T, LI F, HE L J. Secure video offloading in MEC-enabled IIoT networks: a multicell federated deep reinforcement learning approach[J]. IEEE Transactions on Industrial Informatics, 2024, 20(2): 1618-1629.

[29] 俱莹, 陈宇超, 田素恒, 等. 毫米波车联网多基站多用户下的安全传输方案[J]. 通信学报, 2024, 45(8): 84-99.
JU Y, CHEN Y C, TIAN S H, et al. Secure transmission scheme for millimeter-wave Internet of vehicles with multiple base stations and users[J]. Journal on Communications, 2024, 45(8): 84-99.

[30] QIN P, FU Y, ZHANG J, et al. DRL-based resource allocation and trajectory planning for NOMA-enabled multi-UAV collaborative caching 6G network[J]. IEEE Transactions on Vehicular Technology, 2024, 73(6): 8750-8764.

[31] SALEEM U, LIU Y, JANGSHER S, et al. Mobility-aware joint task scheduling and resource allocation for cooperative mobile edge computing[J]. IEEE Transactions on Wireless Communications, 2021, 20(1): 360-374.

[作者简介]

姜微（1991– ），女，博士，浙江工业大学网络空间安全研究院讲师，主要研究方向为移动通信网络、工业物联网、移动边缘计算等。

袁宵（1999– ），女，浙江工业大学网络空间安全研究院硕士生，主要研究方向为移动通信网络、物理层安全、移动边缘计算等。

王倩（1990– ），女，博士，浙江工业大学网络空间安全研究院讲师，主要研究方向为无线通信、工业物联网、空天地一体化等。

钱丽萍（1981– ），女，博士，浙江工业大学网络空间安全研究院教授，主要研究方向为无线通信、工业物联网、智能计算等。