

基于图神经网络和深度强化学习的V2X通信资源分配

马新骥, 吴琼, *Senior Member, IEEE*, 范平毅, *Senior Member, IEEE*, 程楠, *Senior Member, IEEE*, 陈文, *Senior Member, IEEE*, 王江舟, *Fellow, IEEE*, Khaled B. Letaief, *Fellow, IEEE*

5
2
0
2
n
u
J
6
1
G
L
s
c
[
2
v
8
1
5
6
0
7
0
4
2
:
v
i
X
r
a

Abstract 在车联网（IoV）技术快速发展的背景下，蜂窝车联网（C-V2X）通信因其在覆盖范围、时延和吞吐量方面的卓越性能而备受关注。C-V2X中的资源分配对于确保安全信息传输、满足车对车（V2V）通信超低时延和高可靠性严格要求至关重要。本文提出一种将图神经网络（GNN）与深度强化学习（DRL）相结合的方法来解决这一挑战。通过构建以通信链路为节点的动态图，并采用图采样与聚合（GraphSAGE）模型适应图结构变化，该模型旨在保证V2V通信高成功率的同时，最小化对车对基础设施（V2I）链路的干扰，从而确保V2V链路信息成功传输并维持V2I链路的高传输速率。所提方法保留了GNN的全局特征学习能力，支持分布式网络部署，使车辆能够基于本地观测从图网络中提取包含结构信息的低维特征，并独立做出资源分配决策。仿真结果表明，GNN的引入以适度增加计算负载为代价，有效提升了智能体的决策质量，展现出优于其他方法的性能。本研究不仅为V2V与V2I通信提供了理论高效的资源分配策略。

本研究部分得到了国家自然科学基金（批准号：61701197和62071296）、国家重点研发计划（批准号：2021YFA1000500(4)）、国家关键专项（批准号：2020YFB1807700）、上海市科委项目（批准号：22JC1404000）、研究资助局卓越学科领域计划（批准号：AoE/E-601/22-R）以及111计划（批准号：B23008）的资助。

马鑫基和吴琼就职于江南大学物联网工程学院，中国无锡 214122（电子邮件：maoxinji@stu.jiangnan.edu.cn, qiongwu@jiangnan.edu.cn）。

范平毅就职于清华大学电子工程系，北京信息科学与技术国家研究中心，北京 100084，中国（电子邮件：fpy@tsinghua.edu.cn）。

南成现就职于西安电子科技大学综合业务网理论及关键技术国家重点实验室和通信工程学院，西安 710071，中国（电子邮件：dr.nan.cheng@ieee.org）。

文晨就职于上海交通大学电子信息与电气工程学院，上海 200240（电子邮件：wenchen@sjtu.edu.cn）。

江州王（Jiangzhou Wang）就职于英国坎特伯雷CT2 7NT肯特大学工程学院（电子邮件：j.z.wang@kent.ac.uk）。

K. B. Letaief 就职于香港科技大学电子与计算机工程系，香港（电子邮件：eekhaled@ust.hk）。

版权所有 (c) 2024 IEEE。允许个人使用本材料。但如需将本材料用于任何其他目的，须通过向 pubs-permissions@ieee.org 发送请求获得 IEEE 许可。

同时也为实际车联网环境中的资源管理开辟了新的技术路径。

Index Terms—V2X、资源分配、图神经网络、强化学习

一、引言

在智慧城市的发展中，车联网（V2X）技术在智能交通系统中扮演着至关重要的角色[1]–[3]。V2X技术旨在实现车辆与其周围环境（包括其他车辆、交通基础设施、行人及网络资源）的全面通信[4]–[6]。随着汽车行业向自动驾驶、智能导航和自动泊车等前沿技术转型[7]–[11]，V2X的重要性日益凸显。然而，V2X在性能与安全性方面仍面临挑战。

鉴于这些挑战，各种V2X技术被开发出来以解决车辆通信的不同方面[12][13]。在众多V2X技术中，蜂窝车联网（C-V2X）被认为比IEEE 802.11p技术能提供更高的数据速率、更低的延迟和更高的可靠性。近期，第三代合作伙伴计划（3GPP）在其Release 16标准中正式规范了5G车联网（NR-V2X）技术[14][15]。

然而，随着这些先进技术的部署，新的问题也随之出现[16]–[20]。在实际部署中，资源分配问题成为车载网络技术的关键环节，需要满足V2X通信带来的巨大无线通信需求[21]–[27]。该问题通常属于NP难问题，导致在车载网络环境中难以同时满足车对车（V2V）链路的可靠性要求与车对基础设施（V2I）链路的速率要求[28][29]。此外，传统资源分配方法往往依赖精确的信道状态信息（CSI）[30]，这在车辆高速移动的环境中极难实现。

为了应对这些资源分配的挑战，研究人员转向了先进算法。随着深度学习（DL）和强化学习（RL）的发展，研究者开始利用强大的函数

深度强化学习（DRL）的近似特性被用于解决资源分配问题[31]–[36]。在信道状态信息（CSI）不准确的情况下，DRL能够通过持续试错积累经验，学习到在分布式资源分配场景中实现更优性能的通用策略。

然而，据观察，在正常情况下，集中式资源分配方案通常在效率方面优于分布式方案[37]。这种优势归因于集中式方法能够获取关于资源占用和每条链路具体需求的全局信息，这有助于制定更优的分配策略。

然而，在分布式环境中，强化学习中的智能体仅能基于局部观测做出决策，这些观测往往缺乏足够的全局系统信息。这种限制阻碍了车辆间的有效协作，严重影响了全局资源调度的质量。由于资源碰撞和拥塞的高发性，以及通信链路间相互干扰的管理困难，该场景进一步复杂化。此外，在车辆高速运动的环境中，基于局部车辆观测得到的CSI（信道状态信息）含有大量噪声[38]，难以准确表征信道条件，进而对DRL的性能产生不利影响。

为应对这些挑战，研究者们开始探索图神经网络（GNNs）的潜力。GNN在提取全局信息和抑制噪声方面具有独特优势，这体现在其在不同任务中的卓越性能，如链接预测[39]、节点分类[40]、图分类[41]和图生成。GNN的机制通过迭代聚合相邻节点特征，使信息能够沿着边逐层传播。因此，多层GNN能够捕获包含全局信息的节点特征，同时在聚合过程中有效抑制噪声。这种分层迭代的信息处理方式使GNN能够识别图结构数据中的复杂模式与依赖关系，从而增强其在广泛应用中的实用性。

事实上，在文献中，已有许多研究尝试将图神经网络（GNN）与深度强化学习（DRL）相结合以解决复杂问题。然而，由于GNN的独特结构，大多数结合GNN的方法都采用集中式框架运作。此外，传统GNN难以应对图结构的动态变化——一旦节点数量或节点间连接关系发生改变，就需要对图进行各种调整，甚至在某些情况下需要重新训练网络。这在车辆高速运动且车辆数量频繁波动的车载网络中尤为棘手，因为这会严重影响网络的适应性和效率。

Graph Sample and Aggregate（GraphSAGE）模型由William Hamilton、Rex Ying和Jure Leskovec于2017年提出[42]，是一种专为动态图设计的图神经网络变体。

图结构。它能够学习适用于所有节点的通用聚合函数，以处理节点数量变化的问题，同时通过每层采样节点来避免因大量邻居带来的高计算成本。然而，GraphSAGE主要专注于聚合节点特征，忽略了边权重的重要性。在车载网络环境中，V2V链路被表示为节点，这些链路之间的相互干扰被描绘为边，边的权重可以直观地表示干扰强度，因此成为一个关键参数。

此外，车载网络环境通常被表示为一个完全图，其中节点（V2V链路）通过干扰关系相互连接。随着环境中车辆数量的增加，图的复杂性升级，导致对计算资源的需求大幅增加。这种高资源消耗在低延迟是关键需求的车载网络环境中是不可行的。

本文为实现更高效的分布式资源分配，并解决车辆局部观测状态有限且不准确的影响，提出了一种图神经网络辅助的分布式深度强化学习（DRL）算法 $\{v^*\}$ 。本文的主要贡献如下：

- 1) 为丰富车辆自主选择资源时可用的状态信息，我们引入了能够适应动态图结构的Graph-SAGE框架，并充分考虑了边的特性。
- 2) 为构建GraphSAGE框架并实现资源分配方案，我们将图神经网络与双深度Q网络（DDQN）集成于分布式场景中。每个V2V链路在图中表示为节点，在DDQN中作为智能体。此外还分析了计算复杂度及对车辆规模增长的鲁棒性，最终通过仿真验证该方法。

本文的其余部分组织如下：第二节回顾了相关工作。第三节介绍了系统模型。第四节提出了所采用的图神经网络模型，并提供了一种新颖的图构建方法。在第五节中，我们利用先前构建的图神经网络模型，基于车辆的局部观测提取包含全局信息的低维特征，并与DDQN模型协同解决联合资源分配问题。第六节展示了仿真结果。最后，第七节对全文进行总结。为更好地理解所提出的方法，接下来我们将回顾相关方向的最新研究进展。

二、相关工作

在本节中，我们首先回顾一些基于传统方法的相关工作，然后回顾最新的工作

¹The source code has been released at: <https://github.com/qiongwu86/GNN-and-DRL-Based-Resource-Allocation-for-V2X-Communications>

利用DRL和GNN解决资源分配问题的方法。

A. Traditional Methods

传统的资源分配方法通常采用博弈论、拍卖理论和进化算法进行研究。在[43]中，研究了一种基于设备到设备（D2D）通信的无线信道资源管理方法。将V2V通信中对延迟和可靠性的要求转化为优化约束条件，这些约束条件可以直接利用慢变信道信息进行计算。将信道资源分配问题建模为优化问题，并采用启发式算法进行求解。在[44]中，将车辆服务接入类别划分为安全与非安全车辆用户设备，其目标是最大化系统总吞吐量。在[45]中，研究了非授权频谱中C-V2X用户与车辆自组织网络用户的共存问题，提出了基于能量感知的频谱共享方案以减少两类用户间的冲突，其中运用匹配理论研究无线资源分配问题，并给出了车辆动态资源分配算法。在[46]中，推导了考虑不完美CSI条件下单个VUE遍历容量的精确表达式，采用模拟退火（SA）算法获得良好的功率分配结果。在[47]中，基于802.11p和C-V2X技术提出了低延迟V2V通信资源分配方案，在混合架构中利用蜂窝基站确定信道选择，将延迟最小化问题建模为最大加权独立集问题，提出了贪婪V2V链路选择算法并推导了理论性能下界。

在[48]中，提出了一种多状态协作方法。通过几何和网络分解方法，能量在节点间按状态逐级分配。每个状态的持续时间经过优化，以实现节点效用最大化。在[49]中，提出了一种基于分组正交频分复用（OFDM）系统的跨层自适应资源分配算法。该算法考虑了随机流量到达和用户公平性问题，提升了系统频谱效率并改善了队列性能。在[50]中，研究了具有协作组的无线供能通信网络。采用半定松弛方法联合优化时间分配、波束成形向量和功率分配，并在可用功率和服务质量要求的约束下针对两个通信组进行优化，保证了每个问题的全局最优解。在[51]中，研究了传感器网络中移动控制中心通过射频信号为传感器供电并收集其信息的场景，提供了两种提高能源效率的解决方案。在[52]中，探讨了基于网络编码的多播路由问题，旨在实现自组织网络中的最大流多播路由。编码节点的统计特性

基于随机图理论分析了自组织网络中的最大流问题。在[53]中考虑了节点移动性，提出了一种能够跟踪动态拓扑变化的方法。该方法允许路径交叉，显著提升了端到端时延和数据包投递率的性能。

B. Methods Based on DRL and GNN

近年来，DRL与GNN越来越多地被应用于解决资源分配问题。文献[54]首次将强化学习引入资源分配领域，提出了分布式决策方案。文献[55]采用DRL研究上行非正交多址接入（NOMA）系统中的资源分配，设计了离散化的多深度Q学习（DQN）结构以降低输出维度，同时引入深度确定性策略梯度（DDPG）实现连续功率分配的决策。文献[56]考虑用户隐私保护，将联邦学习（FL）引入资源分配：在较大时间尺度上，通过谱聚类将信道条件相似的车辆分组，组内车辆通过FL训练模型；在较小时刻尺度上，车辆将模型上传至中央实体进行聚合后分发给各车辆。文献[57]在Actor-Critic框架中应用DQN与DDPG方法，使用DQN进行离散信道选择，DDPG进行连续功率选择，最后通过元强化学习获得良好的网络初始化方法，增强其在动态变化环境中的快速适应能力。文献[58]采用图嵌入方法实现D2D网络链路调度，通过图表示学习提取节点特征以进行功率选择，并分析了监督式与非监督式训练对网络性能的影响。文献[59]将D2D通信链路视为图中的节点，链路间干扰表示为边，将无线网络建模为有向图以进行联合信道与功率分配，仅需少量样本即可获得接近最优的结果，且执行时间仅需几毫秒。文献[60]通过构建异质图并采用异质GNN学习策略，为深度神经网络（DNN）进行信道与功率选择提供先验知识，降低环境变化时需重新训练DNN的复杂度，同时提出参数共享策略确保GNN学习的先验知识有利于DNN决策。

总而言之，虽然现有研究大多聚焦于优化网络架构以实现更高效的资源分配策略，但明显缺乏对丰富车辆局部观测信息的考量。这一研究空白促使我们开展本项研究工作。

三、系统模型

在本节中，我们将依次介绍系统的基本结构、干涉计算方

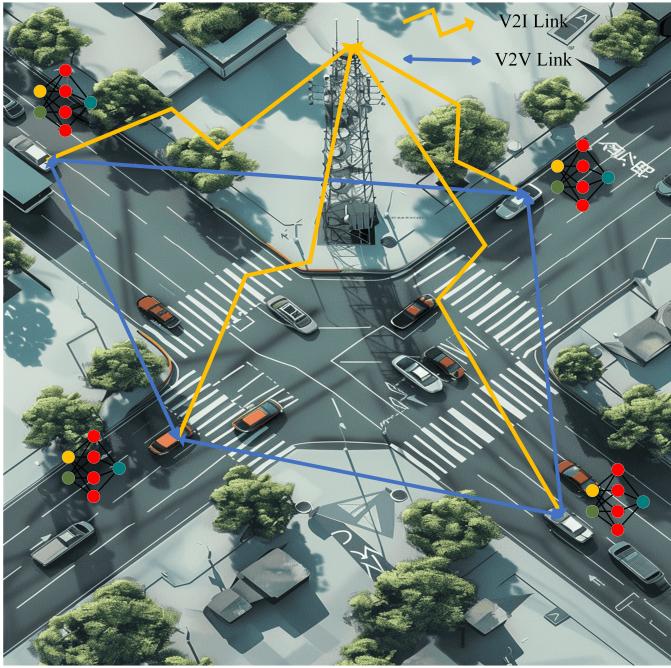


图1：系统模型。

因此，资源选择用于阐明本文提出的系统模型。

A. Basic Structure

本文主要关注V2X范式中的V2I和V2V通信。如图1所示，我们考虑交叉路口的车辆交通模型，其中基站（BS）被战略性地部署在交叉路口中心。车辆以随机分布进入道路，并具有预定义的速度范围；每辆车随机选择速度并保持恒定速率。在此环境中，车辆通过V2I链路与基站通信，以传输高速率的娱乐和生活信息。此外，车辆利用V2V链路互相通信，在驾驶过程中传输需要高可靠性的关键安全信息。

B. Interference Calculation Method

假设有 m 个蜂窝用户（CUE）与基站通信，记为 $M = [1, 2, 3, \dots, m]$ ，以及 k 对V2V用户（VUE），记为 $K = [1, 2, 3, \dots, k]$ 。考虑到上行链路资源利用率相对稀疏，为进一步优化频谱效率，我们假设V2V和V2I链路共享正交分配的上行链路频谱[61]。因此，第*i*个CUE的信号与干扰加噪声比（SINR）可表示为：

$$\gamma^c[i] = \frac{P_i^c h_i}{\sigma^2 + \sum_{j \in K} \rho_j[i] P_j^v \tilde{h}_j} \quad (1)$$

其中 $\sum_{j \in K} \rho_j[i] P_j^v \tilde{h}_j$ 表示使用相同信道的VUE对第*i*个CUE造成的干扰。在此情况下， $\rho_j[i] = 1$ 表示第*j*个VUE与第*i*个CUE存在

利用同一信道，而 $\rho_j[i] = 0$ 否则。 P_j^v 表示第*j*个VUE的发射功率， \tilde{h}_j 表示第*j*个VUE的功率增益。 σ^2 表示噪声功率， P_i^c 表示第*i*个CUE的发射功率， h_i 表示信道增益。根据香农公式，第*i*个CUE的通信容量可表示为：

$$C^c[i] = B \cdot \log(1 + \gamma^c[i]) \quad (2)$$

其中 B 代表信道带宽。

类似地，第*j*个VUE的SINR可表示为：

$$\gamma^v[j] = \frac{P_j^v \cdot g_j}{\sigma^2 + G_{V2I} + G_{V2V}} \quad (3)$$

其中：

$$G_{V2I} = \sum_{i \in M} \rho_j[i] P_i^c g_{i,j} \quad (4)$$

表示同一资源块内V2I链路产生的干扰，且

$$G_{V2V} = \sum_{i \in M} \sum_{j' \in K, j \neq j'} \rho_j[i] \rho_{j'}[i] P_{j'}^v g_{j',j}^v \quad (5)$$

表示同一资源块内其他V2V链路的干扰，其中 g_j 表示第*j*个VUE的功率增益，而 $g_{i,j}$ 和 $g_{j',j}^v$ 分别表示来自第*i*个CUE和第*j'*个VUE的干扰功率增益。第*j*个VUE的信道容量可表示如下：

$$C^v[j] = B \cdot \log(1 + \gamma^v[j]) \quad (6)$$

C. Resource Selection Method

在本文中，我们将车辆的资源配置细化为信道选择和功率级别选择。

- 对于车联网（IoV）中的通信网络而言，信道资源的高效利用是实现效率最大化的先决条件。必须避免某些信道拥塞而其他信道处于闲置状态的情况。尤其在车辆自主选择信道资源的背景下，车辆不仅需要考虑自身信息传输的成功率，还需考量是否在极端情况下占用了其他车辆唯一可用的通信选项。因此，仅依据各车辆自身采集的数据为其选择最优资源无法实现全局优化，整体效益才是关键所在。

- 功率级别的选择通常在信道资源选择之后进行。较小的传输功率可以减少对其他链路的干扰，但由于所选链路上的信噪比和干扰比较低，可能导致传输失败。相反，选择较高的传输功率会对其他链路造成更大的干扰，并增加能耗。因此，选择合适的传输功率水平对整个系统性能至关重要。

鉴于V2I链路的资源通常可由基站直接分配，我们考虑采用分布式算法来完成V2X资源分配任务。

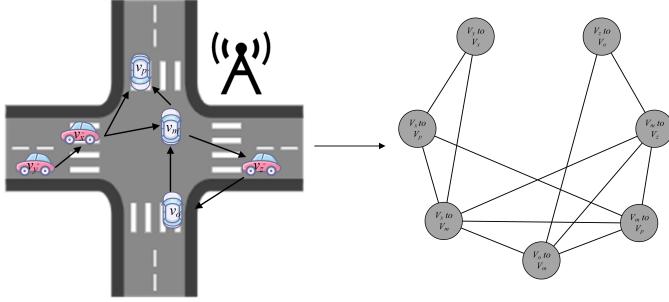


图2：车辆网络图构建。

在预定的V2I资源分配条件下，假设子信道数量等于V2I链路数量，记为 m ，且V2V通信功率列表中的功率水平数量为 n 。每个V2V链路共有 $m \cdot n$ 个资源选择选项。目标是在满足V2V链路时延要求和可靠性的同时，最小化对V2I链路的干扰，从而实现V2I链路传输速率的最大化。为更好评估V2V链路性能，我们将V2V链路的可靠性要求转化为中断概率[43]。

模型与环境之间的交互过程可以描述如下：

在更大的时间尺度上，车辆根据距离关系确定邻居信息，隐式地构建了一个图。在更小的时间尺度上，车辆获取环境的局部观测并收集邻居传输的信息。它们通过模型聚合这些输入，以获得封装全局信息的低维特征。随后，通过深度强化学习（DRL）学习的策略被用于选择信道和传输功率。需要注意的是，虽然车辆是系统模型的参与者，但在算法中，V2V链路作为图中的节点和DRL框架中的智能体。由算法决策得出的最终动作由车辆自身执行。这种去中心化的方法利用局部交互和学习策略来优化车辆网络中的资源分配。

四、图神经网络模型的设计

A. Graph Construction

为了构建车辆网络的图表示，我们将每个V2V对视为节点，记为 $N_g = [v_1, v_2, v_3, \dots, v_k]$ ，并将链路间的干扰关系作为边。对于节点 v ，它包含初始特征向量 x_v 和存储其相邻节点索引的列表 $N(v)$ 。节点的初始特征封装了车辆对信道和干扰信息的本地观测。基于子信道数量等于CUE数量的假设（记为 m ），我们为第 i 个子信道记录：V2V链路的瞬时信道功率增益（表示为 $G_t[i], i \in M$ ）、V2I链路中发射端到接收端的子信道功率增益（记为 $H_t[i], i \in M$ ），以及来自

上一个时间槽，记为 $I_{t-1}[i], i \in M$ 。因此，节点 v 的特征可以表示为

$$x_v = \{G_t || H_t || I_{t-1}\} \quad (7)$$

其中 $||$ 表示向量的拼接。

现有方法主要将车辆网络构建为完全图，以模拟所有链路间的相互干扰关系。然而当环境中车辆数量较大时，图结构会变得异常复杂，节点特征聚合的计算负载急剧增加，导致决策时间延长并可能引入额外延迟。为解决该问题，我们提出基于车辆间通信关系构建图结构。由于每个V2V链路包含发射车辆和接收车辆，假设环境中中共有 s 辆车，表示为 $v = [V_1, V_2, V_3, \dots, V_s]$ 。我们为每个V2V链路分配唯一标签，记为 $V_t \text{ to } V_r$ ，其中 $V_t, V_r \in V$ 。此处 V_t 代表发射车辆， V_r 代表接收车辆。图2展示了我们提出的基于通信关系的图构建方法。假设车辆 V_x 有三个目标车辆： V_m 、 V_p 和 V_o ；同时 V_x 也是车辆 V_y 和 V_z 的目标车辆。这些车辆形成的V2V链路节点标签如下： V_x 至 V_m 、 V_x 至 V_p 、 V_x 至 V_o 、 V_y 至 V_x 以及 V_z 至 V_x 。以节点 V_x 至 V_m 为例，我们将所有包含 V_x 或 V_m 的标签节点视为该节点的邻居，包括标签为 V_x 至 Z 、 Z 至 V_x 、 V_m 至 Z 和 Z 至 V_m 的所有节点，其中 Z 代表符合条件的任意车辆。

我们规定每辆车从附近的相邻车辆列表中选择三个传输目的地，这意味着每辆车发起三个V2V链接。因此，环境中V2V链接的总数为 $3s$ ，其中 s 为车辆数量。对于每辆车，其相邻车辆的数量可以近似为12左右。这个近似基于以下事实：每辆车 (V_x, V_m) 有三个出向链接($V_x \text{ to } Z, V_m \text{ to } Z$)和不确定数量的入向链接($Z \text{ to } V_x, Z \text{ to } V_m$)。但由于链接总数是车辆数量的三倍，车辆作为接收者的平均次数也是3。因此，每辆车的邻居数量大约为12，并且不随环境中车辆数量的增加而变化。相比之下，在完全图中，每个节点的邻居数量是 $3s - 1$ ，这会随着车辆数量的增加而持续增长，导致复杂度随车辆数量增加而升高。

总之，我们提出的基于通信关系的图构建方法可以将每个节点的邻居数量限制在相对稳定的范围内。尽管该方法忽略了链路间的部分干扰关系，但本质上保留了来自邻居的最强干扰。当车辆密度较高时，仍能以相对较小的计算负载提取图中的特征。此外，由于特征

在图中逐层传播，即使节点之间没有直接连接，经过多次迭代后，它们仍然会受到远端节点的影响。此外，选择基于通信关系而非物理距离构建图的原因有两个方面：一方面，这种方法直观反映了链路间的通信关系，节点之间的边反映了链路间的干扰；另一方面，该方法降低了车辆在收集决策所需数据时的通信需求，从而有利于在车辆间隐式构建图。若基于物理距离构建图，由于车辆移动会导致图的拓扑结构变化过快，这可能导致精确保持图结构的复杂度较高，同时图中物理距离偏差的变化也会导致对不同链路干扰估计的最大偏差。对于由 $V_x \rightarrow Z$ 和 $V_m \rightarrow Z$ 表示的六个邻居节点，车辆 V_x 和 V_m 拥有所有相关信息。至于由 $Z \rightarrow V_x$ 和 $Z \rightarrow V_m$ 表示的邻居节点，它们在向 V_x 和 V_m 传输时可以携带节点特征及其他相关信息。因此，通过在 V_x 和 V_m 之间建立双向连接，信息可以在图中传输而不会产生额外的通信开销。

我们对完整图与非完整图场景下图网络的计算负载进行了对比分析。在环境中有 s 辆车的情况下，该图包含 $3s$ 个节点。假设每个节点具有 $D_{input} = 60$ 的特征维度和 $D_{output} = 20$ 的输出维度，并考虑单层聚合模型。每个节点的邻居数量为 N_{nei} ，因此所需乘法运算可计算如下：

$$N = D_{input} \cdot D_{output} \cdot 3s \cdot N_{nei} \quad (8)$$

在构建完整图时， $N_{nei} = 3s - 1$ ，而对于我们提出的图构建方法， $N_{nei} = 12$ 。本质上，这两种图构建方法之间乘法运算次数的差异由邻居数量决定，其计算负载之比也是邻居数量之比。因此，我们提出的方法中邻居数量不随车辆数量增加的特性显得尤为重要。

为了更好地描述车辆网络环境，我们为图中的每条边分配一个权重。通常，距离较远的链路对当前链路的干扰较小。因此，我们记录发射器之间的距离，表示为方阵 $D = [d_{11}, d_{12}, \dots, d_{1s}; d_{21}, d_{22}, \dots, d_{2s}; \dots; d_{s1}, d_{s2}, \dots, d_{ss}]$ 。假设链路 V_m 到 V_n 与链路 V_x 到 V_y 之间的干扰对应于节点 v_p 和 v_q 之间的边，其权重可表示为：

$$\delta_{pq} = 1 - \frac{d_{mx}}{\max(D)} \quad (9)$$

B. GNN Network

在构建图之后，我们将介绍GNN模型。由于车载网络中车辆数量及车辆间通信关系频繁变化，传统GNN模型难以适应节点数量和邻接矩阵的快速变化，需要频繁调整图结构（如修剪或添加边），甚至需要重新训练模型。因此我们引入擅长处理动态大规模图结构的GraphSAGE模型，并通过在节点特征聚合过程中加入边权重对其进行增强。

GraphSAGE模型的操作可分为三个主要步骤：邻居采样、特征聚合和特征更新。邻居采样适用于大规模图处理，其中各层每个节点需要采样的邻居数量是预定义的，允许随机选择这些邻居。这种随机采样机制限制了每个节点特征聚合的计算需求，确保即使面对极其复杂的图结构时，模型仍保持计算可行性。尽管邻居采样函数和我们提出的图构建方法都用于管理计算资源消耗，但我们的方法不同于在完整图中对所有节点进行随机采样。在我们构建的图中，每层采样的邻居被限制在定义范围内，并且它们对当前节点的干扰随着层深的增加而减小。这种层级式干扰衰减有助于提取更优质的特征，因为第二层邻居对当前节点的影响小于第一层邻居。

假设采样的邻居数量为 S ，对于节点 v ，采样邻居集合表示为 $N_s(v)$ ，聚合后的节点特征表示为 z_v 。特征聚合过程可表述如下：

$$z_v = f_{aggregate}(\{x_u | u \in N_s(v)\}) \quad (10)$$

其中， $f_{aggregate}$ 表示模型的聚合函数。我们采用均值聚合函数，但存在一个关键区别：我们引入边权重以反映每个相邻节点特征的重要性。这可以表示为：

$$z_v = \sigma(W_a \cdot \sum_{u \in N_s(v)} \frac{X_u \cdot \delta_{uv}}{|N_s(v)|} + b_a) \quad (11)$$

其中， W_a 表示聚合函数的权重， b_a 表示聚合函数的偏置项，而 σ 表示激活函数。

对于特征更新过程，我们将邻居特征的聚合结果与节点自身的初始特征相结合，以获得最终的节点特征嵌入：

$$h_v = f_{update}(z_v, x_v) \quad (12)$$

在这种方法中，我们选择求和作为将邻居特征的聚合结果与初始节点特征相整合的方式。整合后的结果随后被处理

Algorithm 1: GraphSAGE

```

1 Input: Graph network model, Node features;
2 Output: Aggregation result;
3 Define: Maximum iteration count max_iter and
    convergence threshold  $\epsilon$ ;
4 Initialize: Initialize model, iter  $\leftarrow 0$ ,
    converged  $\leftarrow False$ ;
5 Large Time Scale: Construct global graph;
6 Small Time Scale: Update and Aggregate Node
    Features;
7 while iter  $< max\_iter and converged  $= False$  do
8   for each node v from  $N_g$  do
9     Node Neighbor Sampling:
10    Randomly select five neighbors of node v and
11    store them in  $N_s(v)$ 
12    for each neighbor u  $\in N_s(v)$  do
13      Randomly select five neighbors of node u
14      and store them in  $N_s(u)$ 
15    end
16    Feature Aggregation and Update:
17    for each neighbor u  $\in N_s(v)$  do
18      // Update feature of neighbor
19      u
20       $h_u = f_{update}(x_u, f_{aggregate}(x_n, n \in$ 
21       $N_s(u)))$ 
22    end
23    // Update feature of node v
24     $h_v = f_{update}(x_v, f_{aggregate}(h_u, u \in N_s(v)))$ 
25  end
26  // Check for convergence
27  converged  $\leftarrow check\_convergence()$ 
28  iter  $\leftarrow iter + 1$ 
29 end$ 
```

通过使用从训练中获得的权重参数进行特征提取。得到的节点嵌入表示为 h_v 。因此，更新函数 f_{update} 可以表示如下：

$$f_{update}(z_v, x_v) = \sigma(W_u \cdot (z_v + x_v) + b_u) \quad (13)$$

其中， W_u 和 b_u 表示特征更新函数中经过训练的权重和偏置向量。因此，单轮完整聚合过程可表示为：

$$h_v = f_{update}(x_v, f_{aggregate}(\{x_u | u \in N_s(v)\})) \quad (14)$$

类似地，两层聚合过程可以表示为：

$$h_v = f_{update}(x_v, f_{aggregate}(\{f_{update}(x_u, f_{aggregate}(\{x_n | n \in N_s(u)\}) | u \in N_s(v)\})) \quad (15)$$

算法1总结了GraphSAGE模型的特征提取过程。

为确保GraphSAGE模型提取的特征有利于智能体的决策，我们为其设计了独特的更新机制。由于GraphSAGE模型采用全局部署，而DDQN

模型仅影响单个智能体，因此将两个网络整合进行联合训练具有挑战性。为此，我们选择单独训练GraphSAGE模型。

图神经网络的训练通常分为监督学习、半监督学习或无监督学习。无监督学习在训练阶段可能产生显著的计算开销，而监督和半监督学习则受限于节点标签的获取成本。因此，我们致力于寻找一种高性价比的方法，既能提供网络学习方向，又能增强DDQN模型的决策能力。

事实上，我们利用GraphSAGE模型的关键作用来聚合节点特征，帮助智能体辨别每个信道的质量。DDQN网络采取行动后获得的奖励为信道提供了可靠的评估。因此，它可以将智能体为每个子信道获取的奖励信息存储在一个与子信道数量对应的矩阵 $R_g = [r_1, r_2, r_3, \dots, r_m]$ 中。该矩阵作为相应节点的标签。

然而，存储的奖励值存在一定延迟，使其目前不适合作为绝对标签。但它们仍能为复杂环境中的网络学习提供模糊的方向性指导。针对该问题并尝试有效利用还原的奖励值，我们提出了一种更温和的更新策略。该策略通过使用滞后网络聚合节点特征，并按特定比例用标签对其进行加权。这种方法有两个主要优势：

- 1) 它削弱了标签的绝对影响力，保持了学习过程的稳定性。
- 2) 它逐步引导网络的预测值向理想标签值靠拢。

通过采用这种适度的更新策略，可以在学习稳定性和方向性指导之间取得平衡，最终提升DDQN模型基于GraphSAGE模型聚合特征进行决策的性能。

具体来说，这里采用均方误差（MSE）函数作为网络的损失函数，其定义如下：

$$Loss(\theta) = \sum_{v \in N_g} (y_v - h_v)^2 \quad (16)$$

其中：

$$y_v = \kappa h_v^{old} + (1 - \kappa) R_g^v \quad (17)$$

其中 θ 代表Graph-SAGE模型中的权重参数， y_v 表示用于网络更新的平滑标签， κ 指代聚合结果的权重因子， R_g^v 表示节点 v 的标签，而 h_v^{old} 代表滞后网络的聚合结果。

五、面向资源分配问题的GNN-DDQN模型

在本节中，我们主要介绍基于GNN-DDQN模型解决V2X通信资源分配问题的方法。首先，我们给出强化学习的关键方程 $\{v^*\}$ 及其核心原理

DDQN模型的优势。接着，我们讨论如何将DDQN模型与之前提到的GraphSAGE模型相结合，以解决资源分配问题。

A. Key Formulas in RL and DDQN

强化学习通常被建模为马尔可夫决策过程 (MDP)，其表现为智能体与环境之间的交互。智能体从环境中获取状态信息，根据策略 π 做出决策并执行动作。环境则向智能体提供奖励并更新当前状态。

具体而言，在每个时间步 t ，智能体从环境状态空间中获取当前状态 s_t ，并根据现有策略 π 从动作空间 A 中选择合适的动作 a_t ，从而为V2V链路选择传输信道和传输功率。该策略 π 由状态-动作函数（也称为Q函数，记为 $Q(s_t, a_t)$ ）决定。在深度强化学习中，通常使用深度学习来近似Q函数，以适应复杂的环境变化。智能体执行动作后，环境将转换到新状态 s_{t+1} ，并根据智能体选择的结果提供奖励 r_t 以评估动作质量。本文中的奖励源自V2V链路的中断概率和V2I链路的速率。

接下来，我们将依次介绍DDQN网络中状态、动作和奖励的细节：

1) *State space*: 在我们考虑的V2X环境中，真实状态信息主要包括车辆对环境的观测值 x_v 以及通过Graph-SAGE模型从这些观测值中提取的低维特征 h_v 。为帮助智能体做出更优决策，每辆车会将其信道选择信息发送给目标车辆。基于此，我们收集每个智能体相邻智能体先前的信道选择信息 N_{t-1} ，计算车辆待发送剩余比特与需发送总比特的比率 L_t ，以及在延迟约束下的剩余传输时间 U_t 。综合这些信息，智能体获取的状态表示如下：

$$S_v^t = \{h_v^t || x_v^t || N_{t-1}^v || L_t^v || U_t^v\} \quad (18)$$

2) *Action space*: 根据收集和观测到的状态信息，DDQN网络依据策略 π 选择动作 $a_t \in A$ 。由于智能体需要同时选择子信道和功率级别，我们将这两类动作组合成一个复合动作。智能体选择的复合动作会映射到两个维度上，分别代表子信道和功率级别的选择。本文考虑了一个相对简单的场景，包含三个功率级别和 m 个资源块，因此总共有 $3 \times m$ 种可能的动作。假设智能体选择动作 a_t ，我们将其分解如下：

$$a_r^t = a_t \% m \quad (19)$$

和

$$a_p^t = a_t / m \quad (20)$$

其中%表示模运算，/代表向下取整的除法运算， a_r^t 表示分解后的子信道选择动作， a_p^t 代表功率级别的选择。

3) *Reward function*: 强化学习的目标是使环境中的V2V链路在最大程度上满足V2V通信的低延迟和高可靠性要求，同时最小化对V2I链路的干扰，以最大化V2I链路的传输速率。先前，我们将V2V链路的延迟和可靠性要求转化为对中断概率的要求。因此，在设计奖励时，我们只需考虑V2V链路的中断概率和V2I链路的传输速率。此外，我们还根据链路传输已耗费的时间设置了惩罚项。奖励函数的表达式如下：

$$r_t = \lambda_c \sum_{i \in M} C^c[i] + (1 - \lambda_c) \sum_{j \in K} C^v[j] - \lambda_p (T_0 - U_t) \quad (21)$$

其中， λ_c 代表V2I链路的权重， λ_p 表示所用传输时间的权重。 U_t 表示剩余时间， T_0 代表传输延迟限制。因此，($T_0 - U_t$)表示用于传输的时间。智能体的长期折扣回报可表示为：

$$R_t = E \left[\sum_{n=0}^{\infty} \beta^n r_{t+n} \right] \quad (22)$$

其中， $\beta \in [0, 1]$ 表示奖励的折扣因子。 β 越大表明智能体具有更长远的视角，而较小的 β 意味着智能体更关注即时奖励。

在深度强化学习中，主要目标是学习一个能够最大化长期折扣奖励的最优策略 π^* ，并开发一个能够预测状态-动作对对应Q值的深度网络模型。对于给定的状态-动作对 (s_t, a_t) ，Q值 $Q(s_t, a_t)$ 表示根据策略 π 采取动作 $a_t \in A$ 后获得的预期累积折扣奖励。因此，Q值可用于评估给定状态下某个动作的质量。一旦获得准确的Q值估计，即可根据以下方程选择动作：

$$a_t = \arg \max_{a \in A} Q(s_t, a) \quad (23)$$

这意味着我们选择具有最大Q值的动作。在DDQN中，对应于最优策略的Q值（表示为 $\{v^*\}$ ）可以通过以下更新方程获得：

$$\begin{aligned} Q_{new}(s_t, a_t) = & Q_{old}(s_t, a_t) + \alpha [r_{t+1} + \\ & \beta Q_{old}^{target}(s_{t+1}, \arg \max_{a \in A} Q_{old}(s_{t+1}, a)) \\ & - Q_{old}(s_t, a_t)] \end{aligned} \quad (24)$$

其中， α 是学习率，方程右边的第二项是时序差分

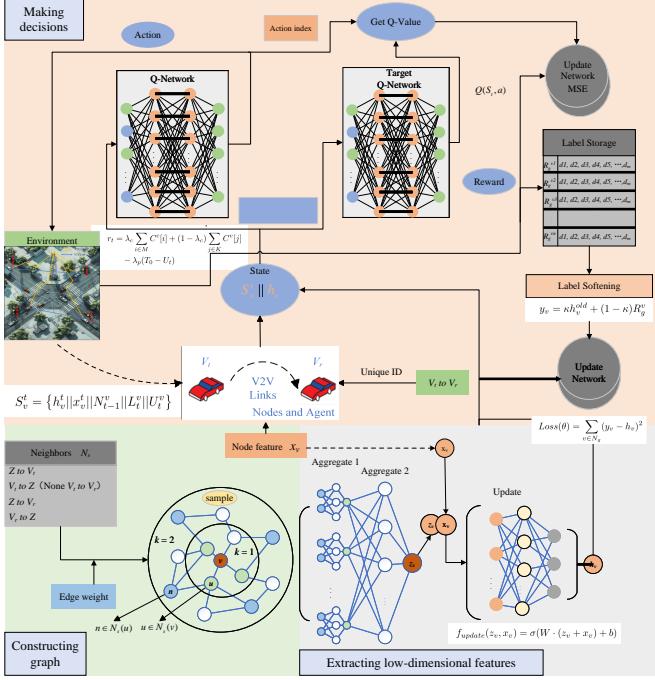


图3：GNN-DDQN框架。

(TD)误差用于更新Q值。折扣因子用 γ 表示。 Q_{old} 代表Q网络预测的Q值，而 Q_{old}^{target} 代表目标Q网络预测的Q值。

假设Q网络的权重由 φ 表示，当输入是智能体观测到的状态时，输出将是每个动作对应的Q值。我们使用TD误差来优化Q网络的参数 φ ，其表达式如下：

$$Loss(\varphi) = \sum_{(s_t, a_t) \in D} (y_Q - Q(s_t, a_t, \varphi))^2 \quad (25)$$

其中， D 表示状态-动作对的集合， y_Q 代表期望目标Q值，可通过以下公式得出：

$$y_Q = r_t + \beta Q_{old}^{target}(s_{t+1}, \arg \max_{a \in A} Q_{old}(s_{t+1}, a), \varphi') \quad (26)$$

其中， φ' 表示目标Q网络的参数。

B. GNN-DDQN Framework and Training-Testing Process

图3详细展示了算法的结构和操作步骤。GNN-DDQN模型的操作可分为三个步骤：构建图、提取关于全局信息的低维特征，以及根据Q网络预测的Q值做出决策。请注意，GraphSAGE模型的训练过程可参考先前介绍的内容。此处我们将重点关注DDQN模型的训练和GNN-DDQN模型的测试。

在强化学习中，智能体所需的所有数据都是通过智能体与环境之间的交互产生的。我们使用元组 (s_t, a_t, r_t, s_{t+1}) 作为单个

训练样本并将其存储在回放缓冲区中，定期从中抽取样本来更新DDQN网络。环境模拟器包含VUE、CUE及其信道模型。环境中的车辆根据均匀分布随机放置在交叉口模型中。根据车辆间的距离关系，建立V2V和V2I通信链路。智能体根据环境中的CSI选择信道和传输功率。随后环境更新衰落信息和干扰信息，生成新状态 s_{t+1} 。同时，智能体会根据所采取动作的质量获得奖励 r_t 。在训练阶段，我们采用 ϵ -贪婪策略来平衡探索与利用。在测试阶段，根据预测的Q值选择动作，具体选择具有最高Q值的动作。

在初始化阶段，智能体基于较大的时间尺度，通过车与车之间的通信链路识别其邻居，隐式构建了一个边权由邻近度决定的图结构。在较小的时间尺度上，智能体收集表示为 S_v^t 的局部观测值，并将其与相邻智能体的特征进行聚合以执行特征更新，从而生成低维表示 h_v 。该表示与局部观测值 S_v^t 共同构成智能体的状态。随后智能体通过Q网络预测具有最高Q值的动作，并从环境中获取奖励。训练过程中，为每个动作预测的Q值将作为基线评估值存储，经软化处理后得到 y_v ，用于指导GNN模型的更新。智能体策略通过目标Q网络预测的Q值进行优化，该过程同时也会更新Q网络本身。在测试阶段，直接执行Q网络预测的最大Q值动作作为决策。

算法2明确表达了GNN-DDQN模型的训练过程，而算法3则阐述了GNN-DDQN模型的测试过程。

六、仿真结果与分析

A. Simulation settings

本研究中的仿真采用Python 3.8和TensorFlow 2.6实现，沿用了文献[54]中的类似参数设置。我们考虑在2 GHz载波频率下运行的单蜂窝系统，并依据3GPP TR 36.885 [62]所述的曼哈顿场景进行仿真，涵盖视距(LOS)与非视距(NLOS)信道条件。车辆分布采用空间泊松过程建模，随机分布于各车道。假设每辆车与其三个最近邻车辆建立V2V通信链路，使得V2V链路数量为车辆数量的三倍。

我们采用深度为2的GraphSAGE模型，每层采样5个邻居节点，并整合节点自身特征进行更新。图网络提取的特征维度设定为20。每个节点输入图网络的特征维度为60，并采用平均聚合方式。

Algorithm 2: Training Process for GNN-DDQN

Input: GraphSAGE model, Q-network model, Simulation environment

Output: Q-network, GraphSAGE-network

- 1 **Define:** Maximum iteration count max_iter and convergence threshold ϵ ;
- 2 **Initialize:** Randomly initialize policy π , initialize model, start environment simulation and add vehicles, CUEs, and VUEs, $iter \leftarrow 0$, $converged \leftarrow False$;
- 3 **Large Time Scale:** Construct global graph;
- 4 **Small Time Scale:** Joint GNN-DDQN model selects actions;
- 5 **while** $iter < max_iter$ and $converged = False$ **do**
 - // Execute feature extraction and aggregation
- 6 **Execute Algorithm 1**
 - // Run the DDQN Network
- 7 For node v , the agent selects subchannels and transmission power based on policy π using h_v (Eq. 18) and locally observed states
- 8 The environment simulator generates a new state s_{t+1} and reward r_t based on the agent's actions (Eq. 21)
- 9 Store the selected subchannels and corresponding rewards in R_g
- 10 Collect and store the tuple (s_t, a_t, r_t, s_{t+1}) in the experience replay buffer
- 11 Sample a mini-batch from the experience replay buffer
- 12 Train the DDQN network using the mini-batch (Eq. 24 and Eq. 25)
- 13 Update the GNN network using labels derived from R_g
 - // Update policy π
- 14 Update policy π : select the action with the highest Q-value (Eq. 23)
- 15 $converged \leftarrow check_convergence()$
 - $iter \leftarrow iter + 1$

采用函数。对于DDQN模型，状态维度输入为102，采用三层神经网络模型，每层分别包含500、250和120个神经元。最终输出60维动作的Q值。层间使用线性整流函数（ReLU）作为非线性激活函数。该激活函数可表示为：

$$f_{ReLU}(x) = \max(0, x) \quad (27)$$

学习率设置为在训练期间逐渐降低。图网络的初始学习率选为0.01，而DDQN网络的初始学习率选为0.005，两者最低值均为0.0001。更详细的参数设置见表I。

Algorithm 3: Testing Process for GNN-DDQN

Input: Trained Q-network model, Trained GraphSAGE model, Simulation environment

Output: Evaluation results

- 1 **Define:** Maximum iteration count max_iter ;
- 2 **Initialize:** Load the Q-network model and GraphSAGE model, start the environment simulation, and add vehicles, CUEs, and VUEs, $iter \leftarrow 0$;
- 3 **Large Time Scale:** Construct global graph;
- 4 **Small Time Scale:** Joint GNN-DDQN model selects actions;
- 5 **while** $iter < max_iter$ **do**
 - // Execute feature extraction and aggregation
- 6 **Execute Algorithm 1**
 - // Run the DDQN Network
- 7 For node v , the agent selects the action with the highest Q-value based on h_v (Eq. 18) and locally observed states
- 8 The environment simulator updates the environment based on the agent's actions (Eq. 21)
- 9 Update the evaluation results, which include the average V2I capacity and the success rate of V2V communication
- 10 $iter \leftarrow iter + 1$

表I：系统参数

Parameters of System Model	
Description	Specification
Carrier Frequency	2 GHz
Height of Vehicle Antenna	1.5 meters
Bandwidth of Single Subchannel	1.5 MHz
Gain of Vehicle Antenna	3 dBi
Height of Base Station Antenna	25 meters
Noise Figure of Vehicle Receiver	9 dB
Gain of Base Station Antenna	8 dBi
Vehicle Speed	36 km/h to 54 km/h
Noise Figure of Base Station Receiver	5 dB
Distance Threshold for Neighbor Vehicles	150 meters
Number of Lanes in Environment	4 per direction, total 16 lanes
Maximum Delay for V2V Link	100 ms
Transmission Power Levels	[23, 10, 5] dBm
Noise Power	-114 dBm
Weight Coefficients $[\lambda_c, \lambda_p]$	[0.3, 1]
First Layer Neighbor Sampling Count	5
Second Layer Neighbor Sampling Count	5
Depth of GraphSAGE	2
Dimension of Node Feature Output	20

注意：诸如“载波频率”、“车辆天线高度”、“单子信道带宽”、“车辆天线增益”、“基站天线高度”等参数均改编自[54]。

B. Model Training Status

为了分析网络的训练过程，我们监控了训练进度。由于我们的GraphSAGE模型和DDQN网络是分开训练的，我们将分别分析这两个网络的训练性能。

1) GraphSAGE模型的训练损失：图4展示了GraphSAGE模型在5000次训练迭代中的损失性能，其中该模型

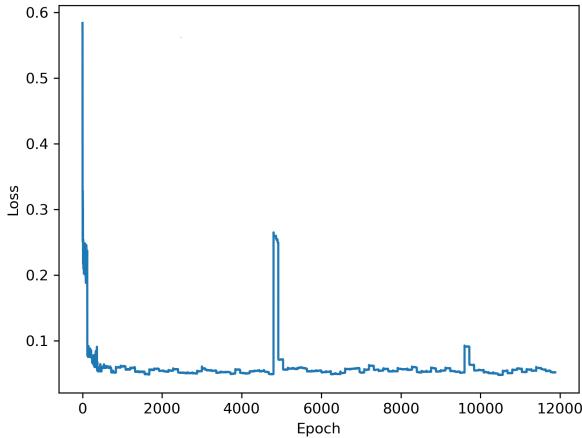


图4：GraphSAGE的训练损失。

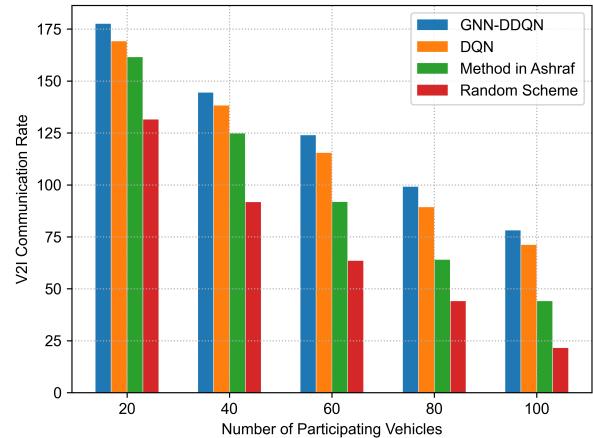


图6：平均V2I速率与车辆数量的关系。

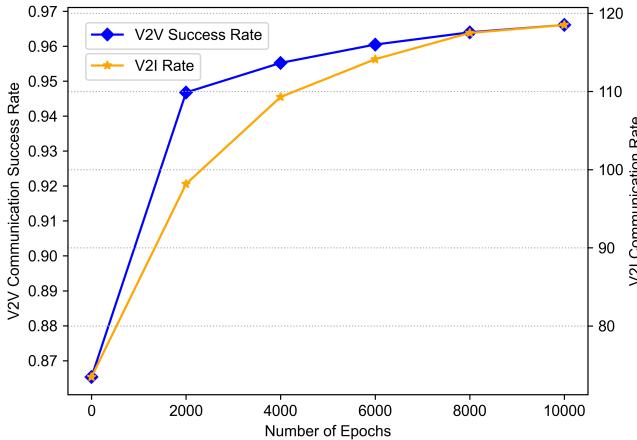


图5：GNN-DDQN的训练效果。

每50次迭代更新120次。可以观察到损失迅速收敛。由于每2000次迭代重置一次环境，每次重置环境时损失会出现显著波动。然而，随着迭代次数的增加，环境重置带来的波动逐渐减弱，模型开始适应不同的环境。

2) GNN-DDQN网络在不同训练迭代次数下的性能比较：图5展示了在模拟环境中经过10,000次训练迭代后所得策略的测试性能。实际应用中，为获得收敛策略，我们进行了近40,000次训练迭代。此处展示前10,000次训练迭代的结果。可以观察到，随着训练迭代次数的增加，模型测试中的评估V2I通信速率和平均V2V通信成功率逐渐提升，但增长速率逐渐放缓。

C. Static Environment Testing

三种基线方法被用于比较。第一种是随机资源分配方法，智能体随机选择信道和功率水平，作为性能下限基准。第二种是[63]中提及的方法，该方法将具有相似信道条件的车辆分组，并迭代分配子信道。第三种是[54]采用的方法，使用通用DQN模型进行资源分配，这有助于我们分析引入GNN带来的性能增益。

我们将所有智能体的决策过程分为十个批次，以解决因信息延迟导致智能体同时选择信道时产生的资源冲突问题。一旦智能体做出决策，其行动将保持固定直至下一轮重新分配。为了获得更具泛化能力的策略，我们在训练阶段会定期重置环境。

在获取测试结果并绘制图表时，我们将环境重置100次，每次环境取200个样本的平均值，再对100个环境的数据进行整体平均。我们观察到智能体学习到的策略在变化环境中会表现出波动性，但波动范围是可控的。鉴于车辆密度等因素在不同环境中可能存在显著差异，并不存在恒定的最优解，因此出现波动是合理的。

1) V2I通信速率：图6展示了所有子信道的总V2I速率与车辆数量之间的关系。从图中可以明显看出，随着车辆数量的增加，环境中的V2V链路数量也随之增加。因此，V2V通信链路对V2I链路的干扰增加，导致V2I链路速率下降。虽然DQN方法和我们提出的GNN-DDQN方法都基于强化学习，但GNN的引入使智能体能够获取更丰富、更全面的信息——

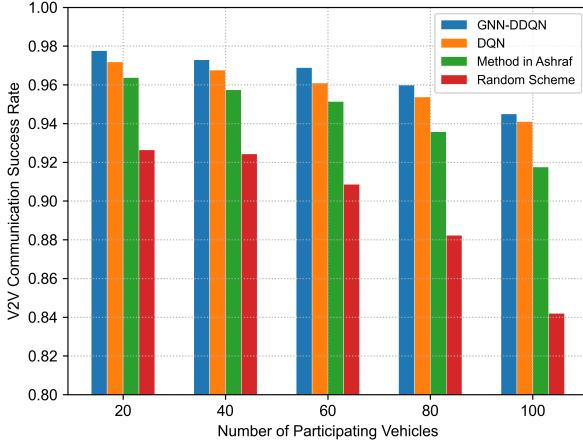


图7：平均V2V成功率与车辆数量之间的关系。

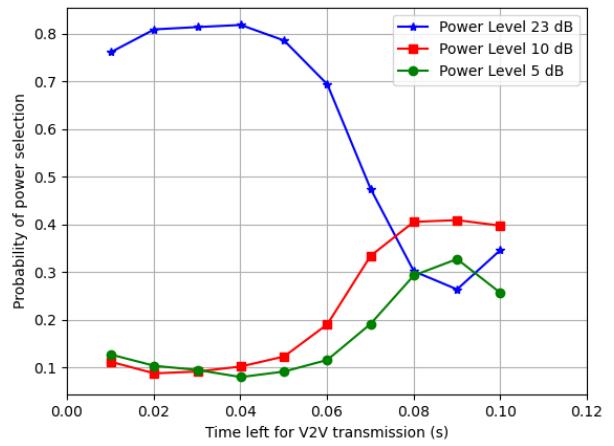


图8：剩余时间与功率选择的关系。

由于缺乏全局状态信息，因此能够实现更优越的性能。基于传统方法和随机策略的方法受到特别严重的干扰，因为强化学习方法能够学习最大化长期奖励的策略，从而在高密度通信环境中表现更佳。

2) V2V通信成功率：图7展示了环境中所有V2V链路的传输成功率与车辆数量的关系。如前所述，环境中的V2V链路数量是车辆数量的三倍。可以观察到，随着车辆数量的增加，整体通信成功率呈现下降趋势，这与常识一致。采用随机策略的V2V通信成功率下降速度最快。其他方法的下降速度较慢，表明其具有提升资源利用效率的能力。显然，我们提出的GNN-DDQN方法优于传统DQN网络，进一步证明GNN可以辅助DRL提升系统性能。

D. Strategy Analysis

如图8所示，为分析GNN对智能体学习策略的影响，我们遵循[54]的方法，收集了智能体在不同剩余传输时间下做出的决策，并分析每个功率级别被选择的统计概率。与[54]不同的是，当V2V链路传输失败时，我们对智能体施加了显著惩罚。因此，当剩余传输时间仅为0.01秒时，智能体倾向于选择更高功率级别。如图所示，当传输时间充足时，智能体倾向于选择较低功率级别以减少对V2I链路的干扰；而当传输时间紧张时，则倾向于选择更高功率以确保V2V链路传输的成功。此外，在传输开始时，智能体也

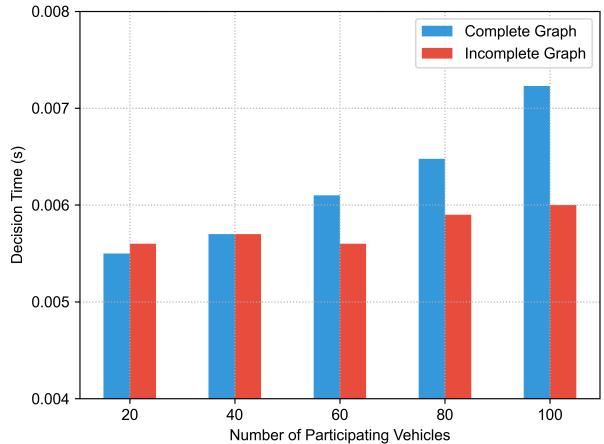


图9：完整图与不完整图的单次决策时间对比。

倾向于选择更高的功率，可能是因为智能体并非在每个时隙都重新选择信道，而初始设置较高的功率水平是一种审慎的选择。这表明智能体学习到的策略捕捉到了某种程度的经验，有助于分析策略的有效性。

图9展示了使用完整图与不完整图构建GNN时的性能差异。当采用完整图时，环境中的节点数量随车辆数线性增长，导致所需训练时间略有增加。可以观察到使用不完整图所需时间几乎保持恒定，而完整图所需时间则随着节点数量增加逐渐增长。为确保对比的可靠性，此处省略了不完整图场景的数据预处理步骤，因为不完整图需要额外步骤从图中获取邻居信息。我们未采用经过优化的高效API——

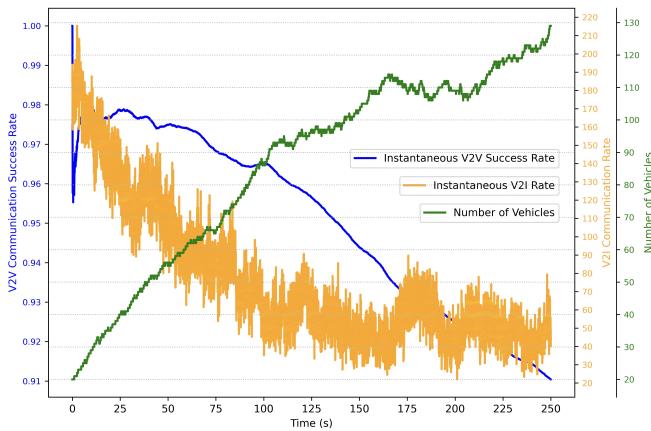


图10：动态环境中车辆数量、V2I速率和V2V成功率的时间动态变化。

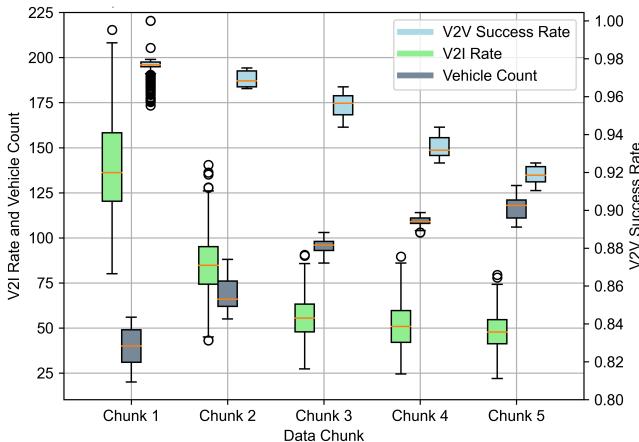


图11：算法性能箱线图：五个时间段内的车辆数量、V2I通信速率和V2V成功率。

为了加速这一过程，GNN已进行了优化[64], [65]；但我们仍采用原始代码实现。因此，在CPU上的处理可能会消耗一些时间，包括从CPU传输数据到GPU所造成的延迟。我们的比较主要关注GPU上的计算速度。对于不完整图曲线，由于邻居数量较少且不随车辆数量增加，单个决策所需时间保持相对较小，且受环境中车辆数量的影响较小。

E. Dynamic Environment Testing

最后，为了验证我们提出的方法在车辆数量动态变化环境中的性能，我们提供了动态条件下的测试结果。一方面，动态环境要求算法具有高鲁棒性，因为车辆分布会随时问频繁变化；另一方面，环境中快速变化的信道占用情况对算法的性能提出了更高且更稳健的要求。为获取丰富的测试数据，

我们设计了一个动态环境，使用字典来存储车辆，当车辆离开环境时删除相应的车辆条目。每次更新车辆位置时，都有一定概率添加新车辆；如果没有添加车辆，则概率增加。一旦添加了车辆，概率将重置为零。为了确保对更多车辆进行彻底探索，我们采用概率模式添加车辆，确保环境中车辆数量稳步增加。

在动态环境中的决策过程中，环境中十分之一的车辆在每个时间槽内做出决策。也就是说，车辆被分成十批依次进行决策，每0.005秒完成一批，这与我们先前提出的方法一致。

1) 动态环境中的车辆数量、瞬时V2I速率与瞬时V2V成功率：图10展示了车辆数量、V2V通信瞬时平均成功率及各信道瞬时V2I通信速率总和随时间的变化情况。车辆数量随时间持续增长。在环境初始阶段，V2V通信成功率存在显著振荡现象——这种振荡并非每次测试都会出现，源于环境随机性且会快速趋于稳定。从图中可见，V2V通信成功率保持高度稳定，不随车辆位置变化而波动，但随着环境中车辆数量增加而逐渐降低。对于V2I速率，由于策略是在静态环境中训练的，无法完全适应动态变化的环境。在保证V2V通信成功率的同时，未能有效最小化对V2I链路的干扰，导致V2I速率出现较大波动。总体而言，随着车辆数量增加，V2I速率呈现下降趋势。

2) 结果分析：在图11中，采样的5000个数据点被划分为五个时段，每个时段代表特定时间段内的数据。针对每个时段分析了车辆数量、V2I通信速率和V2V通信成功率的表现。在第一个时段，V2V通信成功率出现异常值的频率较高，这可以从瞬时值观测到的环境波动得到解释。虽然V2I通信速率偶尔出现异常值，但仍处于可控范围内。总体而言，随着车辆数量逐渐增加，V2I通信速率和V2V通信成功率呈现逐渐下降趋势。各时段数值保持相对稳定且异常值较少，证明了算法的鲁棒性。

七、结论

本文从集中式资源分配方法优于分布式方法的原因出发，并

尝试将图神经网络（GNN）与深度强化学习（DRL）相结合，使每个智能体能够从局部观测中获取更多信息。针对GNN通常部署在全局范围且难以适应动态图结构的问题，我们提出了一种特别适合车辆网络的新型图构建方法。该方法无需增加额外通信开销即可隐式构建图网络，并采用归纳式GraphSAGE模型来处理车辆数量的可变性。最后，我们将提出的GNN模型与双深度Q网络（DDQN）相结合来解决资源分配问题，同时构建了固定车辆数量的静态环境和车辆数量变化的动态环境进行测试。仿真结果表明，我们提出的GNN模型能有效改进智能体学习到的策略，其性能优于独立的DQN方法。

基于理论分析和仿真结果，可以总结出以下结论：

- 本文提出的图构建方法确保了分布式网络部署，并减少了所需的决策时间。
- GraphSAGE模型能够捕获包含图中全局结构信息的特征。
- 图神经网络的整合使智能体能够从局部观察中捕获更多信息，从而促进更好的决策。

尽管所提出的方法具有优势，但仍存在改进空间。GraphSAGE模型采用归纳式方法学习通用聚合函数，使其难以适应不同节点的个体特征，这限制了其性能表现。未来的工作可着重增强模型架构，确保图神经网络在提取更精确状态特征的同时能适应车辆数量的变化。这可能涉及开发更复杂的聚合函数，或将额外的上下文信息整合到模型中。通过解决这些挑战，所提出的方法可以实现更优性能，并进一步优化车辆网络中的资源分配。

参考文献

- [1] A. Zanella, N. Bui, A. Castellani, L. Vangelista 与 M. Zorzi, 《智慧城市中的物联网》，*IEEE Internet of Things Journal* 第1卷第1期，第22–32页，2014年。[2] Z. Feng, B. Wang, Z. Chang, T. Hämäläinen, Y. Zhao 与 F. Hu, 《基于可重构智能表面的车辆定位主动与被动波束成形联合方案》，收录于*IEEE Transactions on Intelligent Transportation Systems*, 2024年。DOI: 10.1109/TITS.2024.3408315。[3] M. Luan, B. Wang, Z. Chang, T. Hämäläinen 与 F. Hu, 《RIS辅助集成传感通信系统的鲁棒波束成形设计》，*IEEE Transactions on Intelligent Transportation Systems* 第24卷第6期，第6227–6243页，2023年6月。DOI: 10.1109/TITS.2023.3248145。[4] K. Abboud, H. A. Omar 与 W. Zhuang, 《DSRC与蜂窝网络技术在车联网通信中的互联互通：综述》，*IEEE Transactions on Vehicular Technology* 第65卷第12期，第9457–9470页，2016年。[5] J. B. Kenney, 《美国专用短程通信（DSRC）标准》，*Proceedings of the IEEE* 第99卷第7期，第1162–1182页，2011年。

- [6] F. Wu, F. Lyu, J. Ren, P. Yang, K. Qian, S. Gao, and Y. Zhang, “面向高效流失预测模型设计的互联网卡用户画像特征分析,” in *IEEE Transactions on Mobile Computing*, vol. 23, no. 2, pp. 1735–1752, 2024年2月. doi: 10.1109/TMC.2023.3241206。[7] N. Lu, N. Cheng, N. Zhang, X. Shen, and J. W. Mark, “网联车辆：解决方案与挑战,” *IEEE Internet of Things Journal*, vol. 1, no. 4, pp. 289–299, 2014。[8] J. E. Siegel, D. C. Erb, and S. E. Sarma, “网联车辆全景综述——架构、使能技术、应用与发展领域,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 8, pp. 2391–2406, 2018。[9] J. Wang, J. Liu, and N. Kato, “自动驾驶中的网络与通信技术综述,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1243–1274, 2019。[10] G. Luo, C. Shao, N. Cheng, H. Zhou, H. Zhang, Q. Yuan, and J. Li, “EdgeCooper：面向增强车辆感知的网络感知协同LiDAR感知,” in *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 1, pp. 207–222, 2024年1月. doi: 10.1109/JSAC.2023.3322764。[11] W. Wang, N. Cheng, M. Li, T. Yang, C. Zhou, C. Li, and F. Chen, “价值导向：一种新颖的基于信息价值的CAV资源调度方法,” in *IEEE Transactions on Vehicular Technology*, vol. 73, no. 6, pp. 8720–8735, 2024年6月. doi: 10.1109/TVT.2024.3355119。[12] Z. Shao, Q. Wu, P. Fan, N. Cheng, Q. Fan, and J. Wang, “面向5G-V2X异构网络的基于深度强化学习的语义感知资源分配,” *IEEE Communications Letters*, pp. 1–1, 2024。[13] Q. Wu, W. Wang, P. Fan, Q. Fan, J. Wang, and K. B. Letaief, “异构车辆边缘计算中面向URLLC的资源分配,” *IEEE Transactions on Vehicular Technology*, vol. 73, no. 8, pp. 11789–11805, 2024。[14] M. H. C. Garcia, A. Molina-Galan, M. Boban, J. Gozalvez, B. Coll-Perales, T. Şahin, and A. Kousaridas, “5G NR V2X通信教程,” *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1972–2026, 2021。[15] K. Sehla, T. M. T. Nguyen, G. Pujolle, and P. B. Velloso, “C-V2X中的资源分配模式：从LTE-V2X到5G-V2X,” *IEEE Internet of Things Journal*, vol. 9, no. 11, pp. 8291–8314, 2022。[16] C. Zhang, W. Zhang, Q. Wu, P. Fan, Q. Fan, J. Wang, and K. B. Letaief, “基于分布式深度强化学习的梯度量化技术赋能车辆边缘计算中的联邦学习,” *IEEE Internet of Things Journal*, pp. 1–1, 2024。[17] Q. Wu, W. Wang, P. Fan, Q. Fan, H. Zhu, and K. B. Letaief, “基于弹性联邦与多智能体深度强化学习的下一代网络协同边缘缓存,” *IEEE Transactions on Network and Service Management*, vol. 21, no. 4, pp. 4179–4196, 2024。[18] Q. Wang, D. O. Wu, and P. Fan, “无线传感器网络中时延约束下的最优链路调度,” *IEEE Transactions on Vehicular Technology*, vol. 59, no. 9, pp. 4564–4577, 2010。[19] R. Deng, Y. Zhang, H. Zhang, B. Di, H. Zhang, H. V. Poor, and L. Song, “面向6G超大规模MIMO的可重构全息表面：实用设计、优化与实现,” *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 8, pp. 2367–2379, 2023, doi: 10.1109/JSAC.2023.3288248。[20] Y. Lin, H. Zhao, X. Ma, Y. Tu, and M. Wang, “针对卷积神经网络调制识别的对抗攻击,” *IEEE Transactions on Reliability*, vol. 70, no. 1, pp. 389–401, 2021, doi: 10.1109/TR.2020.3032744。[21] S. A. Ashraf, R. Blasco, H. Do, G. Fodor, C. Zhang, and W. Sun, “5G新空口Release-16系统对车联网万物（V2X）服务的支持,” *IEEE Communications Standards Magazine*, vol. 4, no. 1, pp. 26–32, 2020。[22] R. Zhang, K. Xiong, H. Du, D. Niyato, J. Kang, X. Shen, and H. V. Poor, “生成式AI赋能的车辆网络：基础、框架与案例研究,” in *IEEE Network*, vol. 38, no. 4, pp. 259–267, 2024年7月. doi: 10.1109/MNET.2024.3391767。[23] F. Wu, F. Lyu, H. Wu, J. Ren, Y. Zhang, and X. Shen, “面向小基站边缘系统性能优化的用户关联模式特征分析,” in *IEEE Network*, vol. 37, no. 3, pp. 210–217, 2023年5/6月. doi: 10.1109/MNET.121.2200089。[24] W. Zhuang, Q. Ye, F. Lyu, N. Cheng, and J. Ren, “赋能未来车联网的SDN/NFV技术：增强通信、计算与缓存,” in *Proceedings of the IEEE*, vol. 108, no. 2, pp. 274–291, 2020年2月. doi: 10.1109/JPROC.2019.2951169。[25] Y. Chen, Z. Chang, G. Min, S. Mao, and T. Hämäläinen, “移动边缘计算中状态更新的感知与计算联合优化,”

- 计算系统，”发表于《*IEEE Transactions on Wireless Communications*》，第22卷第11期，第8230-8243页，2023年11月。doi: 10.1109/TWC.2023.3261338。 [26] N. Cheng, F. Lyu, W. Quan, C. Zhou, H. He, W. Shi, X. Shen, 《空间/天基辅助物联网应用计算卸载：一种基于学习的方法》，《*IEEE Journal on Selected Areas in Communications*》，第37卷第5期，第1117-1129页，2019年。 [27] R. Zhang, K. Xiong, Y. Lu, B. Gao, P. Fan, K. B. Letaief, 《MU-MISO SWIPT赋能异构网络中联合协调波束成形与功率分配比优化：一种基于多智能体DDQN的方法》，《*IEEE Journal on Selected Areas in Communications*》，第40卷第2期，第677-693页，2022年。 [28] D. Zhai, R. Zhang, L. Cai, B. Li, Y. Jiang, 《基于NOMA的大规模物联网设备无线网络能效用户调度与功率分配》，《*IEEE Internet of Things Journal*》，第5卷第3期，第1857-1868页，2018年。 [29] R. Zhang, K. Xiong, Y. Lu, P. Fan, D. W. K. Ng, K. B. Letaief, 《RIS辅助SWIPT网络中RSMA能效最大化：一种基于PPO的方法》，收录于《*IEEE Journal on Selected Areas in Communications*》，第41卷第5期，第1413-1430页，2023年5月。doi: 10.1109/JSAC.2023.3240707。 [30] S. Yue, S. Zeng, L. Liu, Y. C. Eldar, B. Di, 《全息MIMO通信的混合远近场信道估计》，《*IEEE Transactions on Wireless Communications*》，2024年，doi: 10.1109/TWC.2024.3433491。 [31] Y. Xiao, J. Liu, J. Wu, N. Ansari, 《利用深度强化学习进行流量工程：综述》，《*IEEE Communications Surveys & Tutorials*》，第23卷第4期，第2064-2097页，2021年。 [32] Z. Zhao, S. Bu, T. Zhao, Z. Yin, M. Peng, Z. Ding, T. Q. S. Quek, 《雾无线接入网络中计算卸载的设计》，《*IEEE Transactions on Vehicular Technology*》，第68卷第7期，第7136-7149页，2019年。 [33] Y. Lin, M. Wang, X. Zhou, G. Ding, S. Mao, 《具有优先级的无人机编队通信动态频谱交互：一种深度强化学习方法》，收录于《*IEEE Transactions on Cognitive Communications and Networking*》，第6卷第3期，第892-903页，2020年9月。doi: 10.1109/TCCN.2020.2973376。 [34] Y. Lin, Y. Tu, Z. Dou, L. Chen, S. Mao, 《轮廓星图与深度学习在物理层信号识别中的应用》，收录于《*IEEE Transactions on Cognitive Communications and Networking*》，第7卷第1期，第34-46页，2021年3月。doi: 10.1109/TCCN.2020.3024610。 [35] R. Sun, N. Cheng, C. Li, F. Chen, W. Chen, 《6G无线网络优化中知识驱动的深度学习范式》，《*IEEE Network*》，第38卷第2期，第70-78页，2024年。 [36] K. Qi, Q. Wu, P. Fan, N. Cheng, Q. Fan, J. Wang, 《基于多智能体强化学习的可重构智能表面辅助VEC》，《*IEEE Communications Letters*》，2024年，doi: 10.1109/LCOMM.2024.3451182。 [37] J. Shen, N. Cheng, X. Wang, F. Lyu, W. Xu, Z. Liu, K. Aldubaihy, X. Shen, 《RingSFL：一种应对客户端异构性的自适应分裂联邦学习》，收录于《*IEEE Transactions on Mobile Computing*》，第23卷第5期，第5462-5478页，2024年5月。doi: 10.1109/TMC.2023.3309633。 [38] C. Guo, L. Liang, G. Y. Li, 《低延迟车载通信的资源分配：有效容量视角》，《*IEEE Journal on Selected Areas in Communications*》，第37卷第4期，第905-917页，2019年。 [39] M. Zhang, Y. Chen, 《基于图神经网络的链路预测》，收录于《*Neural Information Processing Systems*》，2018年。[在线]。可用：<https://api.semanticscholar.org/CorpusID:3573161> [40] T. Kipf, M. Welling, 《图卷积网络的半监督分类》，《*ArXiv*》，卷abs/1609.02907，2016年。[在线]。可用：<https://api.semanticscholar.org/CorpusID:3144218> [41] M. Zhang, Z. Cui, M. Neumann, Y. Chen, 《一种端到端的图分类深度学习架构》，收录于《*AAAI Conference on Artificial Intelligence*》，2018年。[在线]。可用：<https://api.semanticscholar.org/CorpusID:4770492> [42] W. L. Hamilton, Z. Ying, J. Leskovec, 《大规模图上的归纳表示学习》，收录于《*Neural Information Processing Systems*》，2017年。[在线]。可用：<https://api.semanticscholar.org/CorpusID:4755450> [43] W. Sun, E. G. Ström, F. Bränström, K. C. Sou, Y. Sui, 《基于D2D的V2V通信的无线资源管理》，《*IEEE Transactions on Vehicular Technology*》，第65卷第8期，第6636-6650页，2016年。 [44] Q. Wei, L. Wang, Z. Feng, Z. Ding, 《LTE-U驱动的异构V2X通信网络中的无线资源管理》，《*IEEE Transactions on Vehicular Technology*》，第67卷第8期，第7508-7522页，2018年。
- [45] 王平、狄博、张宏、边凯与宋令，“非授权频谱中的蜂窝车联网通信：与5G系统中VANET的和谐共存”，《*IEEE Transactions on Wireless Communications*》，第17卷第8期，第5212-5224页，2018年。 [46] 李晓、马龙、徐毅与R. Shankaran, “基于D2D的车联网通信在不完美CSI下的资源分配”，《*IEEE Internet of Things Journal*》，第7卷第4期，第3545-3558页，2020年。 [47] F. Abbas, 范平与Z. Khan, “一种基于蜂窝车联网通信的低延迟V2V资源分配新方案”，《*IEEE Transactions on Intelligent Transportation Systems*》，第20卷第6期，第2185-2197页，2019年。 [48] 陈伟、戴凌、Khaled B. Letaief与曹政，“协作网络中资源分配的统一跨层框架”，《*IEEE Transactions on Wireless Communications*》，第7卷第8期，第3000-3012页，2008年。 [49] 张延军与Khaled B. Letaief, “多用户分组OFDM网络的自适应资源分配与调度”，收录于《2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577)》第5卷，2004年，第2949-2953页。 [50] 熊凯、陈超、曲广、范平与Khaled B. Letaief, “无线供能通信网络中群组协作的最优资源分配”，《*IEEE Transactions on Wireless Communications*》，第16卷第6期，第3840-3853页，2017年。 [51] 李拓、范平、陈政与Khaled B. Letaief, “由移动控制中心供能的能量收集传感器网络最优传输策略”，《*IEEE Transactions on Wireless Communications*》，第15卷第9期，第6132-6145页，2016年。 [52] 张健、范平与Khaled B. Letaief, “无线自组网中高效多播路由的网络编码技术”，《*IEEE Transactions on Communications*》，第56卷第4期，第598-607页，2008年。 [53] 姚铮、蒋健、范平、曹政与Victor O. K. Li, “基于邻居表的Ad Hoc网络多路径路由”，收录于《Proc. 57th IEEE Semiannual Vehicular Technology Conference, 2003. VTC 2003-Spring》第3卷，2003年，第1739-1743页。 [54] 叶海、李桂源与B.-H. F. Juang, “基于深度强化学习的V2V通信资源分配”，《*IEEE Transactions on Vehicular Technology*》，第68卷第4期，第3163-3173页，2019年。 [55] 王欣、张宇、沈荣、徐毅与郑福春, “基于DRL的上行NOMA系统能效资源分配框架”，《*IEEE Internet of Things Journal*》，第7卷第8期，第7279-7294页，2020年。 [56] 张旭、彭木根、严帅与孙毅, “基于深度强化学习的蜂窝车联网通信模式选择与资源分配”，《*IEEE Internet of Things Journal*》，第7卷第7期，第6380-6391页，2020年。 [57] 袁煜、郑冠松、黄凯与Khaled B. Letaief, “基于元强化学习的动态车联网通信资源分配”，《*IEEE Transactions on Vehicular Technology*》，第70卷，第8964-8977页，2021年。[在线]。可用：<https://api.semanticscholar.org/CorpusID:237600245> [58] 李敏、余官定与李桂源, “基于图嵌入的小样本无线链路调度”，《*IEEE Transactions on Wireless Communications*》，第20卷第4期，第2282-2294页，2021年。 [59] 陈涛、张旭、游默、郑冠松与S. Lambotharan, “基于GNN的无线物联网网络资源分配监督学习框架”，《*IEEE Internet of Things Journal*》，第9卷第3期，第1712-1724页，2022年。 [60] 郭健与杨晨, “基于异质图神经网络的多小区多用户系统功率分配学习”，《*IEEE Transactions on Wireless Communications*》，第21卷第2期，第884-897页，2022年。 [61] 邵振、吴强、范平、程宁、陈伟、王军与Khaled B. Letaief, “基于深度强化学习的车联网语义感知频谱共享”，《*IEEE Internet of Things Journal*》，第1-1页，2024年。 [62] “第三代合作伙伴计划；无线接入网技术规范组；演进通用陆地无线接入（E-UTRA）；E-UTRA物理层方面的进一步增强（第9版）。”[在线]。可用：<https://api.semanticscholar.org/CorpusID:16652630> [63] M. I. Ashraf、M. Bennis、C. Perfecto与W. Saad, “车辆到车辆（V2V）通信中的动态邻近感知资源分配”，收录于《2016 IEEE Globecom Workshops (GC Wkshps)》，2016年，第1-6页。 [64] 郑政、石雪、何亮、金海、魏松、戴海涛与彭晓, “Feluca: GPU上以颜色为中心的两阶段图着色算法”，收录于《*IEEE Transactions on Parallel and Distributed Systems*》第32卷第1期，第160-173页，2021年1月1日。doi: 10.1109/TPDS.2020.3014173。 [65] 郑政、杜博、赵超与谢鹏, “延迟容忍网络中基于路径合并的介数中心性算法”，收录于《*IEEE Journal on Selected Areas in Communications*》第41卷第10期，第3133-3145页，2023年10月。doi: 10.1109/JSAC.2023.3310071。