

# 一种基于预留-重用联合的C-V2X通信Q学习型半持续调度算法

王萍\* 陆岩 王帅 姚汪鼎  
(东华大学信息科学与技术学院 上海 201600)

**摘要:** 第3代合作伙伴计划(3GPP)模式4提供了一种直连通信的接入方式, 以支持车联网C-V2X应用。然而V2X网络动态变化的业务负载和车辆移动性导致信道质量不稳定, 加剧了传输过程中分组碰撞问题。为了满足高可靠低时延的V2X通信需求, 该文针对动态负载环境下的高效分布式资源分配策略, 提出一种预留-重用联合的Q学习型半持续调度(RRC-QSPS)算法。该文首先对模式4中现有的半持续调度(SPS)算法分组碰撞问题进行理论建模, 分析了影响碰撞概率的关键参数, 继而提出了动态业务环境下车辆智能体的强化Q学习模型, 建立包括预留-重用联合的动作空间与Q目标函数, 并通过 $\epsilon$ -贪心算法求解, 智能决策动态负载环境下无线资源的预留与重用。仿真对比结果表明, 相比于已有的Lookahead-SPS优化算法, RRC-QSPS算法在高速高负载场景下分组接收率提高了7%, 数据包更新时延降低了10%。

**关键词:** V2X通信; 预留-重用联合; 半持续调度算法; 碰撞概率; Q学习

中图分类号: TN915.04

文献标识码: A

文章编号: 1009-5896(2022)08-2785-07

DOI: 10.11999/JEIT210543

## A Reservation and Reuse Combined Q-learning Semi Persistent Scheduling for C-V2X Communication

WANG Ping LU Yan WANG Shuai YAO Wangding

(School of Information Science and Technology, Donghua University, Shanghai 201600, China)

**Abstract:** The 3rd Generation Partnership Project (3GPP) mode 4 provides a direct communication mode to support the C-V2X applications. However, the dynamic traffic load and vehicle mobility lead to uncertainty of channel quality with serious packet collision problem. In order to meet the demand of ultra reliability and low latency of V2X communication. A Reservation-Reuse Combined Q-learning Semi Persistent Scheduling (RRC-QSPS) algorithm is proposed for efficient distributed resource allocation in dynamic load environment. Firstly, the theoretical model of the collision probability of Semi Persistent Scheduling (SPS) algorithm is built. Then, the Q-learning model of vehicle agent in dynamic load environment is proposed with the reservation-reuse combined action and Q function. By using  $\epsilon$ -greedy method, the optimal reservation and reuse of wireless resources in dynamic load environment can be solved. The simulation results show that compared with the existing Lookahead-SPS optimization algorithms, the packet reception ratio of RRC-QSPS is improved by 7% and the update packet delay is reduced by 10% in high speed and high load scenarios.

**Key words:** V2X Communication; Reservation-Reuse Combined (RRC); Semi Persistent Scheduling (SPS); Collision ratio; Q-learning

### 1 引言

为了支持5G高可靠低时延通信(Ultra Reliable Low Latency Communication, URLLC), 第3代合作伙伴计划(3rd Generation Partnership Project, 3GPP)为蜂窝车联网(Cellular Vehicle to Everything, C-V2X)定义了新的传输模式4(Mode 4)<sup>[1]</sup>。模式4使用PC5接口上的侧链通道, 支持直连通信<sup>[2]</sup>。C-V2X

可以利用蜂窝基础设施和表1中的新型5G网络技术来提升通信质量, 有望成为无人驾驶、自动驾驶的关键技术<sup>[3,4]</sup>。

C-V2X中车辆具有高速移动性, 会频繁地加入或离开车辆的通信范围。网络为了及时跟踪信道状态和车辆位置信息, 需要提高控制信令的发送速率。为了减少信令开销, 模式4允许车辆自主感知和资源重复使用, 利用基于感测的半持续调度(Semi Persistent Scheduling, SPS)算法进行资源分配<sup>[5]</sup>。然而SPS算法采用固定的资源选择方式, 在信道拥

表1 5G网络技术及优势

5G网络技术	优势
软件定义网络技术	提升数据网络灵活性
网络功能虚拟化技术	提升网络运维效率
网络切片技术	满足不同应用场景需求
大规模MIMO技术	提升网络的通信容量
设备到设备通信技术	支持直连通信,降低端到端时延

塞时会发生严重的分组碰撞问题,无法适应动态变化的网络拓扑和网络负载。分组碰撞是导致传输中断的主要原因<sup>[6]</sup>,严重影响C-V2X安全类应用的可靠性。本文重点研究高负载业务速率变化导致的V2X通信性能下降问题,针对动态业务流量环境提出灵活高效的资源分配算法,提升分组接收率和降低数据包更新时延。

近来,不少研究者改进了SPS算法的资源选择方式。文献<sup>[7]</sup>提出基于SPS的资源交替选择(Resource Alternative Selection, RAS)算法,采用2个预留资源进行交替分配,能够缓解连续碰撞。文献<sup>[8]</sup>提出了使用Lookahead的半持续调度(Lookahead based Semi Persistent Scheduling, Lookahead-SPS)算法,在广播消息的控制字段增加车辆的SPS参数信息,减少由于缺少邻居车辆预留信息而产生的碰撞,但需要额外的信令开销<sup>[9,10]</sup>。文献<sup>[11]</sup>提出基于行车方向的SPS算法,可以减少反方向车辆的潜在干扰。此外,文献<sup>[12]</sup>提出一种基于深度强化学习的联合资源分配和功率控制方法,提升高密度场景下的分组接收率。在上述方法中,资源的预留和重用是固定的,不能有效反映信道质量动态变化的影响。当V2X网络业务流量动态变化时,SPS算法容易加剧分组冲突,无法满足高可靠低时延的要求。

本文提出一种预留-重用联合的Q学习型半持续调度(Reservation and Reuse Combined Q-learning Semi Persistent Scheduling, RRC-QSPS)算法,通

过建立动态预留和重用的联合Q学习模型,求解当前V2X网络环境下资源的最佳预留概率和重选计数器。本文主要贡献如下:(1)通过高负载情况下分组碰撞的理论建模,研究资源的预留概率和重选计数器对分组碰撞的影响;(2)引入强化学习方法,建立预留-重用动作和Q目标函数,通过 $\epsilon$ -贪心策略探索得到当前环境下的最优资源分配方案;(3)通过不同业务速率下的仿真,对比所提算法在分组接收率和更新时延方面获得的性能增益。

## 2 SPS资源调度模型

本节介绍了V2X中用于周期性消息传输的SPS算法模型,并建立了使用SPS算法进行资源分配的传输碰撞模型,得到了高负载情况下的数据包碰撞函数。

### 2.1 SPS算法模型

在车联网安全类业务中,协作感知消息(Cooperative Awareness Messages, CAM)是车辆之间周期性交换的一种基本信息<sup>[13]</sup>。为支持CAM通信,车辆采用SPS算法分配无线资源。图1为标准的SPS过程<sup>[14]</sup>,包括信道感测和资源重选两个步骤。由于没有基站集中调度,车辆需要先感知无线资源的质量,然后从可用资源列表 $L_A$ 中随机选择一个CAM资源(CAM Resource, CAMR)进行预留,并多次使用<sup>[5]</sup>。图1给出两个无线资源(红和蓝), $n$ 为当前子帧,资源选择窗口为 $[n+T1, n+T2]$ ,时长小于100 ms。RRI为资源预留间隔。SPS算法实现了资源的1次预留多次使用,从而减少控制信令开销,缓解端到端时延。

在感测过程中,车辆检测过去1 s内无线资源的接收信号强度指示(Received Signal Strength Indication, RSSI),并依据式(1)创建可用资源列表 $L_A$

$$L_A = \{CAMR_m \in L_{tot} | RSSI_m \leq P_{th}\} \quad (1)$$

其中,  $RSSI_m$ 为第 $m$ 个CAMR的RSSI值,  $L_{tot}$ 为选

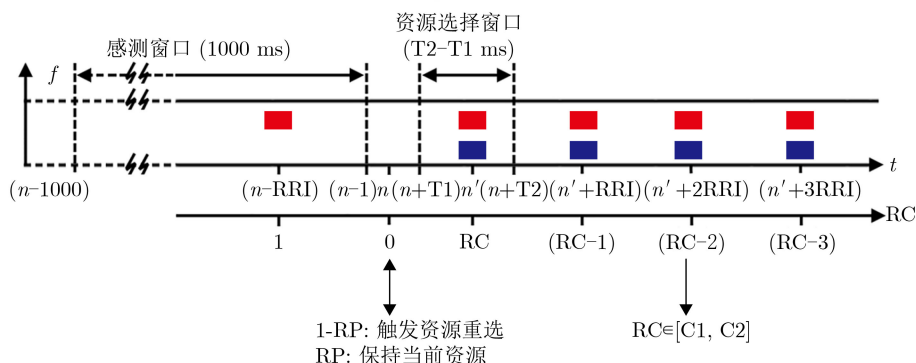


图1 标准SPS算法的资源预留过程

择窗口内的全部资源列表， $P_{th}$ 为功率阈值。按照RSSI值大小对 $L_A$ 中资源进行排序

$$[CAMR_1, CAMR_2, \dots, CAMR_{N_A}] = \arg \text{sort}(L_A) \quad (2)$$

其中， $N_A$ 是 $L_A$ 中CAMR数， $\text{sort}(\cdot)$ 表示升序排列操作。然后选择部分CAMR建立候选资源列表 $L_C$

$$L_C = \{CAMR_1, CAMR_2, \dots, CAMR_{N_C} | N_C = 20\%N_{tot}\} \quad (3)$$

其中， $N_{tot}$ 为 $L_{tot}$ 中的CAMR数， $N_C$ 为 $L_C$ 中CAMR数。依据3GPP标准， $N_C$ 相对于 $N_{tot}$ 的占比为20%。

在重选过程中，节点采用重选计数器(Reselection Counter, RC)确定CAMR的重复使用次数， $RC \in [C1, C2]$ <sup>[9]</sup>。RC越大，节点连续使用该CAMR的次数越多。每传输1个CAMR， $RC-1$ 。当 $RC=0$ 时，节点将以预留概率(Reservation Probability, RP)保留CAMR为后续传输继续使用。RP越大，当前CAMR被保留的概率越大， $RP \in [0, 0.8]$ 。

下面通过对SPS算法中碰撞概率的理论建模，研究参数RP和RC对分组碰撞的影响。

## 2.2 分组碰撞模型

分组碰撞是影响可靠通信的主要原因<sup>[6]</sup>。在资源选择过程中，如果多个节点同时选择了相同的CAMR，将会导致分组碰撞问题。同时满足 $RC=0$ 的车辆越多，则碰撞概率越大<sup>[14]</sup>。

定义信道繁忙比(Channel Busy Ratio, CBR)为信道的利用率<sup>[15]</sup>

$$CBR = \frac{\sum_{r=1}^{N_{tot}} \text{num}(RSSI_m \geq RSSI_{th})}{N_{tot}} \quad (4)$$

其中， $RSSI_{th}$ 为CBR门限值， $\sum_{r=1}^{N_{tot}} \text{num}(RSSI_m \geq RSSI_{th})$ 为RSSI值超过 $RSSI_{th}$ 的CAMR数量。CBR越大，数据包碰撞概率越高<sup>[8]</sup>。

CAMR的RSSI值超过 $RSSI_{th}$ 就表示有车辆在用该资源进行传输，则每个子帧的平均传输车辆数为

$$V_{sf} = N_{sf} \cdot CBR \quad (5)$$

其中， $N_{sf}$ 为单个子帧包含的CAMR数。

同时满足 $RC=0$ 的车辆将会同时启动分组传输，则同时发送数据包的车辆数 $V_0$ 为

$$V_0 = V_{sf} \cdot RP / C_{avg} \quad (6)$$

其中， $C_{avg} = (C1 + C2)/2$ 是数据包周期传输的平均长度。由式(6)可见， $V_0$ 受参数RP和RC的综合影响。

在时长为 $T$ 的选择窗口中，资源总数 $N_{tot}$ 为

$$N_{tot} = T \cdot N_{sf} \quad (7)$$

根据当前时刻的信道繁忙比CBR，可以计算出 $L_A$ 中可用资源数 $N_A$

$$N_A = N_{tot} \times (1 - CBR) \quad (8)$$

下面讨论V2X网络分别为高负载和低负载两种情况。低负载时， $L_A$ 中资源数充足， $N_A > N_C$ ；而高负载时， $L_A$ 中资源数目不足，需降低 $P_{th}$ ，使得 $N_A = N_C$ ，此时相当于在 $L_A$ 中选择一个资源用于传输。由于网络的吞吐量有限，车辆在高负载时更容易发生分组冲突。

高负载情况下可用资源较少，不同车辆的 $L_A$ 容易重叠，车辆就可能选择同一个CAMR，从而导致分组碰撞。车辆所具有的重叠资源数 $N_{lap}$ 为

$$N_{lap} = N_A \cdot (RRI - s) / RRI \quad (9)$$

其中， $s$ 为两个重选车辆之间的间隔子帧数，且 $0 \leq s < RRI$ 。某一重叠资源被车辆选择的概率 $p_s$ 为

$$p_s = RRI / N_A \cdot (RRI - s) \quad (10)$$

于是，主车辆 $k$ 与至少1个 $RC=0$ 的邻居车辆发生碰撞的概率 $P_{col,k}^{0 \leq s < RRI}$ 为

$$P_{col,k}^{0 \leq s < RRI} = 1 - (1 - p_s)^{V_0} \quad (11)$$

其中， $s=0$ 时为半双工通信造成的碰撞， $V_0$ 为同时满足 $RC=0$ 进行发送的车辆数目。

考虑 $s$ 范围内的所有竞争车辆造成的碰撞，主车辆 $k$ 的总碰撞概率 $P_{col,k}$ 为

$$\begin{aligned} P_{col,k} &= 1 - \prod_{s=0}^{RRI} (1 - P_{col,k}^{0 \leq s < RRI}) \\ &= 1 - \prod_{s=0}^{RRI} (1 - p_s)^{V_0} \\ &= 1 - \prod_{s=0}^{RRI} \left( 1 - \frac{1}{N_{sf}(1 - CBR)(RRI - s)} \right)^{V_0} \end{aligned} \quad (12)$$

式(12)表明，高负载情况下，CBR越大， $P_{col}$ 越大，且 $P_{col}$ 与 $V_0$ 有关。结合式(6)可知，在高负载环境下减少分组碰撞的关键是对参数RP和RC进行控制。

## 2.3 问题描述

本节提出将最小化碰撞概率作为V2X网络可靠性的优化目标。假设网络车辆数目为 $K$ ，使用SPS算法为周期性CAM消息调度资源。结合式(12)，优化问题可以表示为

$$\begin{aligned} \min \sum_{k=1}^K P_{col,k}, \quad k &= \{1, 2, \dots, K\}, \\ \text{s.t. } 0 \leq RP &\leq 0.8, 10 \leq C_{avg} \leq 50 \end{aligned} \quad (13)$$

其中，RP代表CAMR的预留概率， $C_{avg}$ 代表数据包重复传输的平均次数，由RC决定。参数RP和RC的取值范围参考3GPP标准。由于V2X网络状态时刻变化，式(13)定义的优化问题很难直接求取最优解。下面引入强化学习方法对最优解进行全局搜索，使结果逼近最优解。

### 3 基于强化学习的半持续调度算法

本节提出RRC-QSPS算法,将强化Q学习用于V2X动态业务流量环境下的CAMR分配问题。车辆与环境实时交互,智能决策当前环境下最优的RP和RC。该算法能使车辆适应动态流量的网络环境,并降低碰撞概率。

#### 3.1 强化学习原理

强化学习具备自主决策且学习速度快等特点。 $\langle A, S, R, P \rangle$ 是强化学习中经典的元组, $A$ 代表代理的动作空间; $S$ 为代理所能感知的状态空间; $R$ 代表奖励或惩罚。代理根据与环境交互所获得的奖励或惩罚,不断地学习知识调整动作,使自身更加适应环境。 $P$ 为状态转移概率矩阵。依据马尔可夫决策过程理论, $P$ 仅与当前状态和动作有关。强化学习通过试错学习获得最优策略,无需先验知识,因此适用于解决V2X网络复杂环境中分布式资源分配问题。

V2X网络智能车辆的强化Q学习原理如图2所示。代理为车辆,环境为车辆周围的一切<sup>[16]</sup>,如动态变化的流量负载。车辆在时间 $t$ 从 $S$ 观察到一个状态 $s_t$ ,并从 $A$ 选择一个动作 $a_t$ 。这个过程相当于依据策略 $\pi(a_t|s_t)$ 选择了资源分配参数。执行动作 $a_t$ 后,状态从 $s_t$ 转移到 $s_{t+1}$ 的一步转移概率为 $p(s_{t+1}, r_t|s_t, a_t)$ ,并且车辆获得瞬时回报 $r_t$ 。对于一个各态遍历的马尔可夫决策过程,无论初始策略 $\pi(a_0|s_0)$ 取何值,随着转移步数的增大,该过程最终会达到稳态。车辆的目标是最大化长期期望回报 $s_t \in S$ 。通过不断的状态感知、决策调度和奖励反馈,车辆就可以优化动态环境下的资源分配策略,达到低延迟高可靠的通信质量目标。接下来,本文提出RRC-QSPS算法,定义相关Q函数,利用 $\varepsilon$ -贪心算法求解最大值 $Q^*$ 。

#### 3.2 RRC-QSPS算法模型

本系统中每个车辆状态包括历史碰撞情况和当前信道拥塞情况, $t$ 时刻系统的状态 $s_t \in S$ 为

$$s_t = (\text{COL}_{t-1}, \text{CBR}_t) \quad (14)$$

其中, COL为资源冲突标志。为了降低算法的复杂度和状态空间大小,本文将信道繁忙比CBR离散化<sup>[17]</sup>。

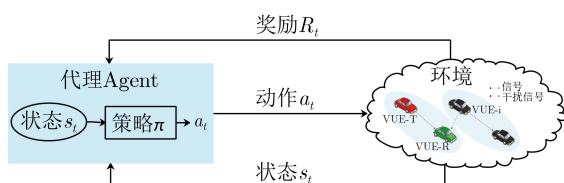


图2 强化学习原理图

根据式(13)中影响碰撞概率的可调参数,本文设计了一组离散的动作空间

$$A = \{a_t^i \in \text{RP}, a_t^j \in \text{RC}\} \quad (15)$$

其中,  $\text{RP} = \{0, 0.1, \dots, 0.8\}$ 和 $\text{RC} = \{5, 6, \dots, 15\}$ 。RP和RC的取值范围参考3GPP标准,决定了动作空间的大小。本文对动作RP和RC分别进行训练学习。

车辆执行两个动作后,从环境中学习得到瞬时回报 $R_t$

$$R_t = \begin{cases} \lambda_U(T_0 - U_t) - \lambda_C V_0, & \text{COL} = 0 \\ -1, & \text{COL} = 1 \end{cases} \quad (16)$$

其中, $T_0$ 是约束时间, $U_t$ 是当前时延, $\lambda_U$ 和 $\lambda_C$ 分别是两个部分的权重,且 $\lambda_U + \lambda_C = 1$ 。COL=0表示没有发生冲突,此时, $R_t > 0$ ,且由剩余可用时间 $(T_0 - U_t)$ 和同时传输的车辆数 $V_0$ 决定。 $U_t$ 越小或者 $V_0$ 越小,车辆得到的回报值越高。COL=1表示资源冲突,则回报函数为-1。

下面建立资源预留-重用联合的Q学习模型。Q值采用时间差分法进行更新,即利用后继状态的Q值估计当前Q值。在不断迭代后,可以找到具有最优策略的Q值

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[R_{t+1} + \eta \max_{s \in S} Q(s_{t+1}, a_t) - Q(s_t, a_t)] \quad (17)$$

其中, $\alpha$ 为学习速率, $\eta$ 是折扣系数, $R$ 为回报值。 $\eta$ 越大,未来的回报越重要。通过重复学习最大化Q值,车辆得到最优动作 $a^* \in A$ 。已经证明,在马尔可夫决策过程中,如果 $(s_t, a_t)$ 离散且每个动作在每种状态下执行无数次,则Q值将会收敛到最优值 $Q^*$ <sup>[18]</sup>。

车辆在优化搜索最优动作时,需要兼顾所学知识的利用和探索两个过程,以保证每个动作都有被选中的可能。本文采用 $\varepsilon$ -贪心策略来平衡利用-探索过程:

$$a_t = \begin{cases} \arg \max_{a_t \in A} Q(s_t, a_t), & \text{with prob. } 1 - \varepsilon \\ \text{random}, & \text{with prob. } \varepsilon \end{cases} \quad (18)$$

其中,车辆以 $1 - \varepsilon$ 的概率选择最优动作,这种情况下的Q最大;以概率 $\varepsilon$ 随机选择其他动作进行探索。概率 $\varepsilon$ 的表达式为

$$\varepsilon = 1 - Q_{\max} \quad (19)$$

其中, $Q_{\max}$ 为车辆在当前 $(s_t, a_t)$ 下的最大Q值。在初始学习阶段,车辆还未学得有效的策略, $Q_{\max}$ 较小,车辆将以较高的 $\varepsilon$ 随机选择其他动作进行探索。随着学习的进行,车辆利用现有学习结果的概率也相应提升。

### 3.3 RRC-QSPS算法实现

RRC-QSPS算法使用Q学习来获得最优的RP和RC值，使得资源选择适应网络负载的动态变化。算法实现如图3所示，CAM在网络中周期生成，由SPS调度器为其分配时频资源。SPS调度器以概率RP保持当前使用资源，以概率1-RP重新选择资源，且选中的资源将重复使用RC次。车辆观察到 $t$ 时刻的一个状态 $s_t$ ，并选择一个动作 $a_t$ ，即选择一个合适的RP和RC值。车辆根据动作的执行从状态 $s_t$ 更新到 $s_{t+1}$ 。最后，车辆将当前状态采取动作所得到的回报记录在Q表中，以便再次达到相同状态时能采取最优的动作。RRC-QSPS算法的伪代码在表2中进行了描述。

## 4 仿真结果与分析

本节在仿真中模拟动态变化的网络负载，对比

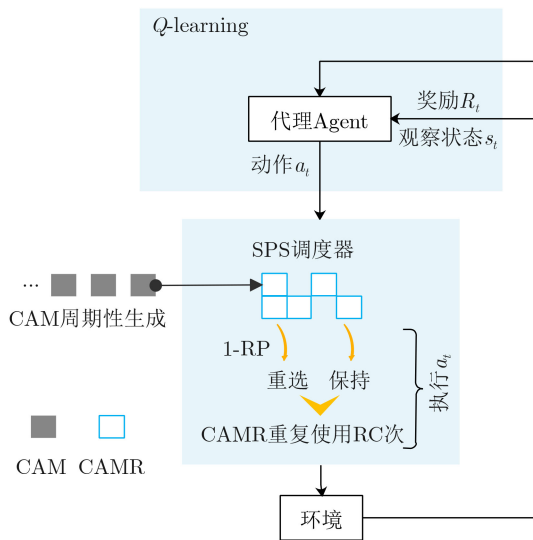


图3 RRC-QSPS算法结构图

表2 算法1 RRC-QSPS

输入: $C1, C2, NCAMRs, Q(s_t, a_t)$
输出: $CAMR_{id}$
(1) 初始化Q学习参数和SPS算法参数 $RC \leftarrow \text{random}(C1, C2)$ , $CAMR_{id} \leftarrow \text{random}(1, NCAMRs)$
(2) 观察当前状态 $s_t$
(3) LOOP
(4) IF $RC \neq 0$ THEN
(5) $RC \leftarrow RC - 1$
(6) ELSE
(7) 由式(18)选择动作 $a_t$ 并执行，即更新RC和RP值，保持或重选 $CAMR_{id}$
(8) ELSE IF
(9) 观察后续状态 $s_{t+1}$ ，并由式(16)计算回报函数 $R_t$
(10) 根据式(17)更新 $Q(s_t, a_t)$ 值
(11) END LOOP

RCC-QSPS算法与现有的SPS算法和Lookahead-SPS算法的各个性能指标。

### 4.1 仿真参数

实验使用了LTEV2Vsim仿真器<sup>[19]</sup>，模拟高速公路场景下安全类业务CAM的传输。车辆移动模型被建模为泊松点过程，用于描述仿真中位置的动态性。实验中的路径损耗和阴影衰落都根据WINNER+信道模型进行计算。本文设置CAM业务的发包率(packet per second, pps)为5~100 pps。表3给出了仿真参数和配置。

为了评估不同网络负载下的系统性能，本文使用了如式(20)–式(23)指标

(1) 碰撞概率(Collision Ratio, CR)

$$CR = CAM_{col} / CAM_{tot} \quad (20)$$

其中， $CAM_{col}$ 代表发生碰撞的CAM数，碰撞数越多成功接收率越低。 $CAM_{tot}$ 为系统传输的CAM总数。

(2) 分组接收率(Packet Reception Ratio, PRR)

$$PRR = CAM_{suc} / CAM_{tot} \quad (21)$$

其中， $CAM_{suc}$ 为成功接收的CAM数。

(3) 更新时延(Update Delay, UD)

$$UD = (1 / CAM_{tot}) \sum_{i=1}^{CAM_{tot}} (T_i - T_{i-1}) \quad (22)$$

表示数据包两次连续成功接收所经过的时间，它与端到端时延相关<sup>[1,12]</sup>。

(4) 平均吞吐量(Average Throughput, AT)

$$AT = CAM_{suc} / T_{tra} \quad (23)$$

其中， $T_{tra}$ 为传输时间，即所有CAM从发射端到接收端所花时间。

### 4.2 仿真结果

首先，本文评估了碰撞概率CR在不同网络负

表3 仿真参数和配置

参数	参数值
仿真时间	50 s
仿真区域中的车辆数( $\rho$ )	100
平均车速	120 km/h(方差=3)
道路长度	1000 m
车道数	2(每个方向1个)
CAM大小	190 Bytes
发包率	5~50 pps
感测距离(raw)	150 m
发射功率	15 dBm
传播模型	WINNER+, B1
阴影衰落方差	3 dB
路径损耗指数( $\beta$ )	2.75
调制和编码方案	MCS 7 (QPSK)

载下的表现。如图4所示,随着业务速率的增加,RRC-QSPS的CR始终保持较低水平1%以内,而SPS和Lookahead-SPS的CR均呈指数增长。由于RRC-QSPS采用了Q学习模型,令车辆与变化的网络环境实时交互,并决策出与当前环境最匹配的SPS动作策略,重选CAMR和调整重复使用次数,从而避免了分组同时传输造成的碰撞。因此RRC-QSPS可以在高负载状态下实现低碰撞率。

图5和图6分别展示了分组接收率PRR和数据包更新时延UD随业务速率的变化情况。图5中3种算法的PRR均随发包率的增大而减小。RRC-QSPS的PRR下降幅度最小,基本保持在92%以上。与Lookahead-SPS相比,RRC-QSPS在高负载情况下的PRR提升了7%。图6中,UD随着发包速率的增加而减小,RRC-QSPS的UD最低。与Lookahead-SPS相比,RRC-QSPS在高负载情况下的UD降低了10%。RRC-QSPS在PRR和UD上具有明显优势,这是因为它利用Q学习方法有效抑制了数据包碰撞。

在图7中,3种调度算法的平均吞吐量随着业务速率的增加而上升,最后趋于饱和。与Lookahead-SPS相比,RRC-QSPS在高负载情况下的吞吐量提升了50%。虽然传统SPS的吞吐量最高,但高负载情况下PRR和UD性能明显恶化。而RRC-QSPS能在动态业务速率环境下明显改善PRR和UD性能,且保持较好的吞吐量水平,更加适用于URLLC场景,支持安全类车联网业务。

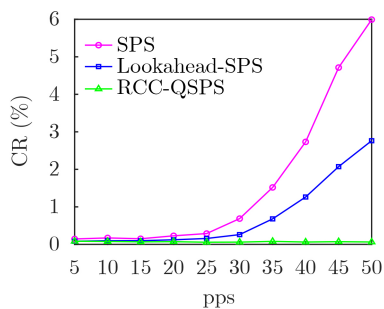


图4 碰撞概率与发包率的关系

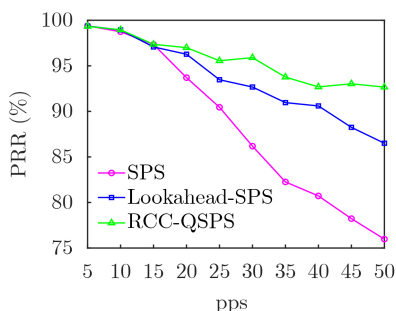


图5 分组接收率与发包率的关系

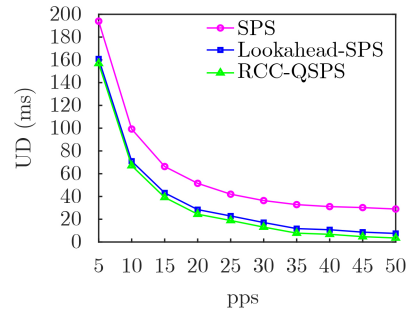


图6 数据包更新时延与发包率的关系

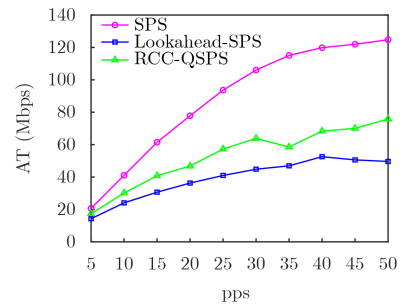


图7 平均吞吐量与发包率的关系

## 5 结束语

本文首先建立了动态C-V2X网络中使用SPS调度算法的分组碰撞模型,研究参数RP和RC对高负载下分组碰撞的影响。其次,提出了预留和重用联合的RRC-QSPS算法。该算法将时延和碰撞概率作为瞬时回报进行强化学习,智能决策RP和RC值,使得资源选择能适应网络的动态变化。最后,仿真对比了不同业务速率下的算法性能。结果表明RRC-QSPS算法在高速高负载网络下明显提高分组接收率,降低数据包更新时延,更加适用于动态网络下高可靠低时延的车联网应用。此外,与已有Lookahead-SPS优化算法相比,RRC-QSPS算法的吞吐量有所提升。未来的工作可以考虑更复杂的信道状态和干扰信息,结合SPS资源选择和功率控制,采用深度Q学习来处理高维复杂的映射关系,进一步改善吞吐量。

## 参考文献

- [1] JEON Y and KIM H. An explicit reservation-augmented resource allocation scheme for C-V2X Sidelink mode 4[J]. *IEEE Access*, 2020, 8: 147241–147255. doi: [10.1109/ACCESS.2020.3015549](https://doi.org/10.1109/ACCESS.2020.3015549).
- [2] BAZZI A, CECCHINI G, ZANELLA A, et al. Study of the impact of PHY and MAC parameters in 3GPP C-V2V mode 4[J]. *IEEE Access*, 2018, 6: 71685–71698. doi: [10.1109/ACCESS.2018.2883401](https://doi.org/10.1109/ACCESS.2018.2883401).
- [3] NAIK G, CHOUDHURY B, and PARK J M. IEEE 802.11 bd & 5G NR V2X: Evolution of radio access technologies for V2X communications[J]. *IEEE Access*, 2019, 7:

- 70169–70184. doi: [10.1109/ACCESS.2019.2919489](https://doi.org/10.1109/ACCESS.2019.2919489).
- [4] WANG Shuai, LIU Yan, ZHU Jie, *et al.* A novel collision supervision and avoidance algorithm for scalable MAC of vehicular networks[J]. *Chinese Journal of Electronics*, 2021, 30(1): 164–170. doi: [10.1049/cje.2020.12.001](https://doi.org/10.1049/cje.2020.12.001).
- [5] MOLINA-MASEGOSA R and GOZALVEZ J. LTE-V for sidelink 5G V2X vehicular communications: A new 5G technology for short-range vehicle-to-everything communications[J]. *IEEE Vehicular Technology Magazine*, 2017, 12(4): 30–39. doi: [10.1109/MVT.2017.2752798](https://doi.org/10.1109/MVT.2017.2752798).
- [6] GONZALEZ-MARTÍN M, SEPULCRE M, MOLINA-MASEGOSA R, *et al.* Analytical models of the performance of C-V2X mode 4 vehicular communications[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(2): 1155–1166. doi: [10.1109/TVT.2018.2888704](https://doi.org/10.1109/TVT.2018.2888704).
- [7] JUNG S Y, CHEON H R, and KIM J H. Reducing consecutive collisions in sensing based semi persistent scheduling for cellular-V2X[C]. 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall), Honolulu, USA, 2019: 1–5.
- [8] JEON Y, KUK S, and KIM H. Reducing message collisions in sensing-based semi-persistent scheduling (SPS) by using reselection lookaheads in cellular V2X[J]. *Sensors*, 2018, 18(12): 4388. doi: [10.3390/s18124388](https://doi.org/10.3390/s18124388).
- [9] BONJORN N, FOUKALAS F, and POP P. Enhanced 5G V2X services using sidelink device-to-device communications[C]. 2018 17th Annual Mediterranean ad HOC Networking Workshop (Med-Hoc-Net), Capri, Italy, 2018: 1–7.
- [10] HONNAIAH P J, MATURO N, and CHATZINOTAS S. Foreseeing semi-persistent scheduling in mode-4 for 5G enhanced V2X communication[C]. 2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, USA, 2020: 1–2.
- [11] 余翔, 陈晓东, 王政, 等. 基于LTE-V2X的车联网资源分配算法[J]. *计算机工程*, 2021, 47(2): 188–193.
- YU Xiang, CHEN Xiaodong, WANG Zheng, *et al.* Resource allocation algorithm for internet of vehicles based on LTE-V2X[J]. *Computer Engineering*, 2021, 47(2): 188–193.
- [12] 金久一, 邱恭安. C-V2X通信中资源分配与功率控制联合优化[J/OL]. *计算机工程*, 2020: 1–10. <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=JSJC20201014004&v=fpDLQvPDFGcn3qeXmuRIceO0Zyaaau3ClfB8VIRE3Gnyph%25mmd2FWjA8xWhfkjIUBdOBz>, 2020.
- JIN Jiuyi and QIU Gong'an. Joint optimization of resource allocation and power control in C-V2X communications[J/OL]. *Computer Engineering*, 2020: 1–10. <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&fil>ename=JSJC20201014004&v=fpDLQvPDFGcn3qeXmuRIceO0Zyaaau3ClfB8VIRE3Gnyph%25mmd2FWjA8xWhfkjIUBdOBz, 2020.
- [13] CAMPOLO C, MOLINARO A, ROMEO F, *et al.* 5G NR V2X: On the impact of a flexible numerology on the autonomous sidelink mode[C]. 2019 IEEE 2nd 5G World Forum (5GWF), Dresden, Germany, 2019: 102–107.
- [14] BONJORN N, FOUKALAS F, CAÑELLAS F, *et al.* Cooperative resource allocation and scheduling for 5G eV2X services[J]. *IEEE Access*, 2019, 7: 58212–58220. doi: [10.1109/ACCESS.2018.2889190](https://doi.org/10.1109/ACCESS.2018.2889190).
- [15] HAIDER A and HWANG S H. Adaptive transmit power control algorithm for sensing-based semi-persistent scheduling in C-V2X mode 4 communication[J]. *Electronics*, 2019, 8(8): 846. doi: [10.3390/electronics8080846](https://doi.org/10.3390/electronics8080846).
- [16] YE Hao, LI G Y, and JUANG B H F. Deep reinforcement learning based resource allocation for V2V communications[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(4): 3163–3173. doi: [10.1109/TVT.2019.2897134](https://doi.org/10.1109/TVT.2019.2897134).
- [17] 陈前斌, 管令进, 李子煜, 等. 基于深度强化学习的异构云无线接入网自适应无线资源分配算法[J]. *电子与信息学报*, 2020, 42(6): 1468–1477. doi: [10.11999/JEIT190511](https://doi.org/10.11999/JEIT190511).
- CHEN Qianbin, GUAN Lingjin, LI Ziyu, *et al.* Deep reinforcement learning-based adaptive wireless resource allocation algorithm for heterogeneous cloud wireless access network[J]. *Journal of Electronics & Information Technology*, 2020, 42(6): 1468–1477. doi: [10.11999/JEIT190511](https://doi.org/10.11999/JEIT190511).
- [18] SUTTON R S and BARTO A G. Reinforcement learning: An introduction[J]. *IEEE Transactions on Neural Networks*, 1998, 9(5): 1054.
- [19] CECCHINI G, BAZZI A, MASINI B M, *et al.* LTEV2Vsim: An LTE-V2V simulator for the investigation of resource allocation for cooperative awareness[C]. 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), Naples, Italy, 2017: 80–85.
- 王萍: 女, 1973年生, 教授, 博士生导师, 研究方向为5G无线网络关键技术、V2X通信、异构网络优化理论、移动物联网。
- 陆岩: 女, 1996年生, 硕士生, 研究方向为5G移动通信、V2X通信、无线资源管理。
- 王帅: 男, 1994年生, 博士生, 研究方向为V2X通信、空间调制、车载网络与5G移动通信。
- 姚汪鼎: 男, 1996年生, 硕士生, 研究方向为V2X通信、机会网络、机会路由与5G移动通信。