

利用机器学习和双平面博弈模型的车联网拥挤感知路由算法

谭晓芳¹ 张 攀² 付凡成¹

¹(南昌理工学院计算机信息工程系 江西 南昌 330044)

²(郑州轻工业大学计算机与通信工程学院 河南 郑州 450002)

摘 要 针对现有车载自组织网络(Vehicular Ad-Hoc Network, VANET)技术在既有自组织无线网络中很难通过直接调整以达到较高的可靠性,设计了新一代 VANET 管理方案以满足新的需求。设计用于 VANET 的新型纵向和横向博弈模型,在纵向博弈模型中,网络代理和车辆一起朝着合适的拥挤感知过程工作;在横向博弈中, VANET 为车辆选择最佳路线。通过横纵向模型,利用机器学习的自主学习和优化决策为车联网拥挤感知提供了一种有效的双平面控制机制。基于分布式在线方法实现拥挤感知的自适应性和实时性,并通过控制决策权衡最优性和适用性。仿真和性能分析结果验证了所提方案的有效性。与现有方案相比,该方案在 RSU 的任务成功率、规范化路由吞吐量和端到端数据包延迟方面有较大改善。

关键词 车联网 机器学习 双平面博弈模型 分布式在线计算 移动拥挤感知

中图分类号 TP393

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2020.12.024

A CONGESTION SENSING ROUTING ALGORITHM FOR VEHICLE NETWORKS USING MACHINE LEARNING AND BIPLANE GAME MODEL

Tan Xiaofang¹ Zhang Qian² Fu Fancheng¹

¹(Department of Computer Information Engineering, Nanchang Institute of Technology, Nanchang 330044, Jiangxi, China)

²(School of Computer and Communication Engineering, University of Light Industry, Zhengzhou 450002, Henan, China)

Abstract In view of the difficulty of direct adjustment of Vehicular Ad-Hoc Network(VANET) technology in existing ad-hoc wireless networks to achieve higher reliability, a new generation of VANET management scheme is designed to meet the new requirements. A new vertical and horizontal game model for VANET was designed. In the vertical game model, the network agent and the vehicle worked together towards the appropriate congestion sensing process. In the horizontal game, VANET choosed the best route for the vehicle. Through the horizontal-longitudinal model, the autonomous learning and optimization decision-making of machine learning was used to provide an effective two-plane control mechanism for the detection of vehicle network congestion. Based on the distributed on-line method, the self-adaptability and real-time performance of congestion sensing were realized, and the optimality and applicability were weighed by the control decision-making. The simulation and performance analysis results verify the effectiveness of the proposed scheme. Compared with the existing schemes, the proposed scheme has a great improvement on the success rate of RSU tasks, normalized routing throughput and end-to-end packet delay.

Keywords Vehicle networking Machine learning Two-plane game model Distributed on-line computing Mobile crowd sensing

0 引言

车载自组织网络作为一种特殊类型的移动自组织网络(Mobile Ad-Hoc Network, MANET),已成为智能交通系统(Intelligent Transportation System, ITS)的重要组成部分^[1]。在VANET中,车辆配备有车载单元(On-Board Units, OBU),这些单元能够用于事件数据记录器、全球定位系统、前向和后向雷达、传感设备和短程无线接口。使用OBU,每个车辆充当发送器、接收器和路由器,以将信息传递给其他车辆。此外,这些车辆还可以与沿路边部署的其他固定基础设施通信,这些基础设施被称为路侧单元^[2](Road Side Units, RSU)。作为互联网接入点,RSU在蜂窝网络中扮演基站的角色。通常,VANET由RSU和车辆的OBU组成,它们在动态变化的VANET环境中反复交互^[3]。

在VANET的运营期间,有两种基本技术可以传播车辆应用的数据:车辆与基础设施(Vehicle-to-Infrastructure, V2I)和车辆与车辆(Vehicle-to-Vehicle; V2V)之间的通信。V2I通信可以帮助车辆从/向RSU下载/上传内容,并且V2V通信主要依赖于车辆之间的协作来交换数据分组,而无需RSU帮助。V2I和V2V通信必须共存并相互补充,以满足VANET的各种通信要求^[4]。为了确保VANET性能,通过V2I和V2V通信有效地传递信息对于为移动车辆提供满意的服务变得重要。大多数研究都考虑如何有效地管理涉及V2I和V2V通信的控制问题^[5]。随着VANET的蓬勃发展,车载拥挤感知(Vehicular Crowd Sensing, VCS)成为一种新的范例,它利用普遍存在的传感器嵌入式OBU来有效地收集数据。文献[6]使用基于语义的信息网络新范式,解决了车联网中信息传播的问题,并对所提出模型的流量模型和网络架构进行了表征。文献[7]提出了两种新的路由算法用于群体传输车载网络,所提算法能够在较低开销下改善数据传输,而且开销可忽略不计。文献[8]从云计算和车辆互联网(Internet of Vehicles, IoV)之间的服务关系角度讨论了云辅助的分类,提出了一种利用来自联网车辆的数据来提高运输效率的新方法。文献[9]通过拥挤感知(Efficient Data Dissemination by Crowdsensing, EDDC)方案有效解决

了数据传播问题,并在车载网络中提出了一种混合路由机制。

本文提出了用于VANET管理的新型V2I和V2V算法。对于V2I通信,基于激励的VCS算法开发了一种新型垂直博弈模型。在垂直博弈中,车辆和RSU是博弈玩家,车辆充当移动传感器以收集所需信息,RSU充当系统服务器以共享和处理多种类型的收集信息。对于V2V通信,开发了一种新的水平博弈模型,以实现基于学习的车辆间通信协议。在横向博弈中,车辆是博弈玩家,以提高路由服务的质量和可靠性。在动态变化的VANET环境下,本文方案被设计为双平面垂直—水平博弈模型。通过考虑现实世界的VANET环境,提出的算法交互式顺序博弈过程的方法是合适的,同时也确保了系统的实用性。为了研究每个博弈中博弈玩家之间的战略互动,采用强化学习方法。强化学习涉及系统代理如何采取行动以最大化其奖励。本文主要创新如下:

- 1) 针对VANET设计了用于V2I和V2V通信的新型双平面模型,在纵向博弈模型中网络代理和车辆将向着拥挤感知过程工作;在横向博弈模型中VANET为车辆选择最佳路线。

- 2) 基于横纵向模型,利用强化学习的自主学习和优化决策为车联网拥挤感知提供了一种有效的双平面控制机制,能有效提升RSU任务成功率。

- 3) 不同于现有方案的非自适应性和离线模式,所提算法基于分布式在线方法实现拥挤感知的自适应性和实时性,并通过控制决策权衡最优性和适用性。

本文算法将动态VANET环境的响应以及博弈论和强化学习算法的相互组合,与现有方案相比,在RSU的任务成功率、规范化路由吞吐量和端到端数据包延迟方面有较大改善。

1 方案设计

本文方案主要由基于强化学习的VCS策略和基于双平面博弈模型的路由算法组成,其中双平面博弈模型和强化学习是VANET控制方案的理论基础。通过采用基于强化学习的博弈方法,并结合分布式在线计算和有效的路由算法,以适应快速变化的VANET环境。本文方案整体思路如图1所示。

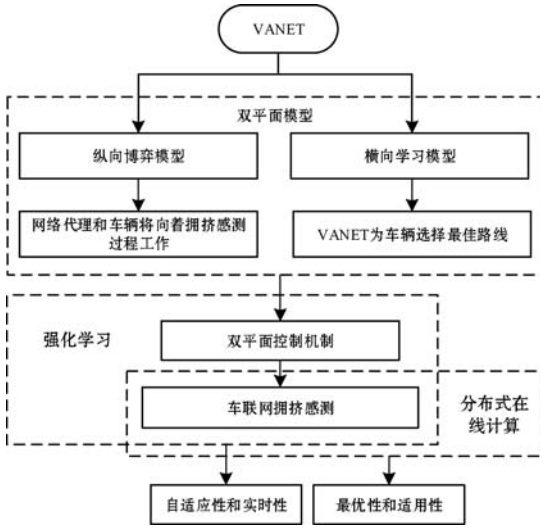


图1 所提方案的整体架构

1.1 新型纵向和横向博弈模型

在 VANET 系统的操作期间,RSU 和车辆通过考虑动态网络条件来作出单独的控制决策。以分布式在线方式,尝试根据学习方法最大化收益。基于此假设,开发了一种双平面博弈模型。RSU 和一组车辆之间的交互被公式化为 V2I 通信的垂直博弈,并且一组车辆之间的交互被公式化为 V2V 通信的水平博弈模型^[11]。形式上在博弈玩法的每个时间段定义垂直博弈为 $G^{V2I} = \{N, X, \{S^R, S_{v_i \in V}^v\}, \{u^R, u_{v_i \in V}^v\}, Z^{B_i^{X_S}}, T\}$ 。

博弈玩家 $N = \{R, V\}$ 的有限集合,其中: R 代表 RSU, $V = \{V_1, V_2, \dots, V_i, \dots, V_k\}$ 是多个车辆的集合, V_i 是第 i 个车辆, k 是 G^{V2I} 博弈的车辆数量;车辆则被假定为 VCS 过程中的移动节点。

RSU 中有限的传感任务集合 $X = \{X_1, X_2, \dots, X_S\}$, S 是总传感任务的数量。

$S^R = [\eta_R^{X_1}, \eta_R^{X_2}, \dots, \eta_R^{X_j}, \dots, \eta_R^{X_S}]$ 是 RSU 策略的向量,其中 $\eta_R^{X_j}$ 代表 X_j 的策略。在 G^{V2I} 博弈中, $\eta_R^{X_j}$ 表示单位感应时间的价格水平,如:

$$\eta_R^{X_j} = \left\{ \prod_{X_j \in N} B_i^{X_j} \mid B_i^{X_j} \in [B_{\min}^{X_j}, \dots, B_i^{X_j}, \dots, B_{\max}^{X_j}] \right\} \quad (1)$$

$S_{V_i}^v = [\mu_{V_i}^{X_1}, \mu_{V_i}^{X_2}, \dots, \mu_{V_i}^{X_j}, \dots, \mu_{V_i}^{X_S}]$, i 是 V_i 战略的载体, $\mu_{V_i}^{X_j}$ 表示 X_j 的策略。在 G^{V2I} 博弈中, $\mu_{V_i}^{X_j}$ 表示 V_i 在单位感应时间对 X_j 的贡献指示。

u^R 是 RSU 收到的收益, $u_{V_i \in V}^v$ 是 V_i 收到的收益。

$Z^{B_i^{X_S}}$ 是 RSU 中 $B_i^{X_S}$ 的学习值,用于估计下一个 RSU 策略选择的概率分布。

$T = \{H_1, H_2, \dots, H_t, H_{t+1}, \dots\}$ 表示时间,其由具有用于 V2I 通信的不完全信息的一系列时间步骤表示。

对于 V2V 通信,在博弈玩法的每个时间段定义水平博弈模型 $G^{V2V} = \{V, N_{V_i}, A_{V_i \in V}, U_{V_i \in V}, L_{V_i}^{a_{V_i}^{V_m}}, T\}$ 。

G^{V2V} 可以制定用于路线操作的车辆的相互作用。

有限的一组博弈玩家 $V = \{V_1, V_2, \dots, V_n\} N_{V_i}$, 其中 n 是 G^{V2V} 博弈的车辆数量。

N_{V_i} 是 V_i 的邻近车辆。

$A_{V_i \in V} = \{a_{V_i}^{V_m} \mid V_m \in N_{V_i}\}$ 是 V_i 的可用策略的有限集合,其中 $a_{V_i}^{V_m}$ 表示用于中继路由分组的邻近车辆 V_m 的选择。

$U_{V_i \in V}$ 是 V_i 收到的回报。

$L_{V_i}^{a_{V_i}^{V_m}} \in A_{V_i}$ 是 V_i 中 $a_{V_i}^{V_m}$ 的学习值; L 用于估计下一个 V_i 战略选择的概率分布。

$T = \{H_1, H_2, \dots, H_t, H_{t+1}, \dots\}$ 表示时间,其由具有用于 V2V 通信的不完全信息的一系列时间步骤表示。

1.2 利用强化学习的 VCS 策略

强化模型^[12] 能选择执行使系统状态得到更多增益的行为,这主要由该模型通过对系统现有状态的感知,并对下一步的行为所能带来的增益进行计算和评估来实现的;若节点执行状态的决策发生了变化,便可以得到执行此行为的系统增益值。由图 2 可知^[13],在路由发现过程 t 时刻,网络中的每个节点都会选择不同的动作策略 $a_i(t)$ 来共同对网络产生影响,节点从状态 $s_i(t)$ 切换到 $s_i(t+1)$ 后,便得到 $r_i(t+1)$ 即环境反馈的相应回报值,然后各个节点对自身的价值函数 $V_i(s)$ 再次进行新的评估,并对邻居节点广播该函数。

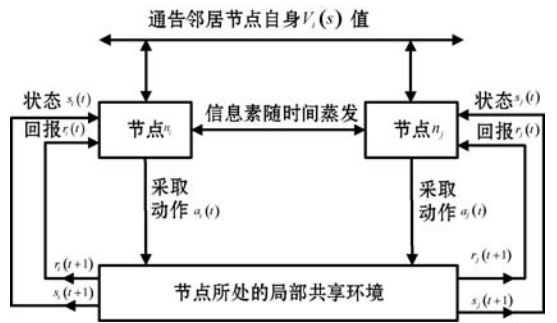


图2 强化学习决策过程

对于单一节点的系统状态评估值,可以使用无限视野价值评估模型来对其进行计算:

$$E \left\{ \sum_{t=0}^{\infty} \gamma \cdot r_t \right\} \quad (2)$$

并且将未来有限步内的系统状态转换所带来的增益 r_t 设置为拥有同样的权重 γ 。对于系统的优化问题,可以将其转化为对由源节点 S 到目的节点 D 中经过的所有数据传输节点进行调整的决策过程,从而得到价值函数 $V(s)$ 的最大值,以下为具体的求解过程。

将系统状态集记作 S , 当前状态能执行的动作集记作 A , 并且 S 和 A 的某种运算结果都可以使用系统增益 R 以及状态转移分布函数 $T(\cdot)$ 来表示。此处从

状态 s 执行动作 a 带来的系统增益由 $R(s, a)$ 表示, s 经过动作 a 后转移到 s' 的概率由 $T(s, a, s')$ 表示。对贝尔曼方程进行求解可得系统最优解 $V^*(s)$:

$$V^*(s) = \max_a [R(s, a) + \gamma \sum_{s' \in S} T(s, a, s') \cdot V^*(s')] \quad (3)$$

式中:状态转移增益权重记作 γ ;在状态集 S 中所对应的每个状态 s' 的当前价值评估值记作 $V^*(s')$;将价值评估函数 $V^*(s)$ 能得到最大值的动作记为 a , 并选择动作 a 作为下一步的状态转移操作。

将 $T(s, a, s')$ 作为每一条通信链路数据投递成功和投递失败的比例, 并对所有节点对尝试发送的单播包数 N_A 、单播传输失败包数 N_F 、接收到的单播包数 N_R 、接收到的广播包数 N_B 和混杂接收单播包数 N_P 等统计量进行记录, 从而对 $T(s, a, s')$ 进行计算。对投递率进行计算时, 需要在接收数据包统计量前面加上置信参数 σ , 并把 ρ 当作节点没有发送数据包时的投递率估计值, 得到最终投递率表示为:

$$p(\text{success} | \{N_i\}) = \frac{N_A - N_F + \rho\sigma(N_R + N_B + N_P)}{N_A + \sigma(N_R + N_B + N_P)} \quad (4)$$

式(4)表示的是系统接下来的状态转换为投递成功 $s' = S$ 时的转移概率 $T(s, a, S)$ 。因为指定了 r_S 和 r_F 为固定值 -1 和 -7 , 因此 $T(s, a, s')$ 的简单组合函数为:

$$R(s, a) = r_S \cdot T(s, a, S) + r_F \cdot T(s, a, F) \quad (5)$$

由此, $T(s, a, s')$ 、 $R(s, a)$ 评估模型得到了确定, 再对贝尔曼方程进行求解可得 $V(s)$ 的最优值:

$$V(s) = \max_a \left\{ \sum_{s'} T(s, a, s') \cdot [R(s, a, s') + V(s')] \right\} = \max_a Q(s, a) \quad (6)$$

通过分析可知, 每个动作集 a 只会让当前状态 s 向两种可能的状态进行转换, 若 P 表示为下一跳的节点, 那么系统状态就会转变为两种可能: 1) 发生传输成功事件 S , 系统会从当前状态 $s = N$ 转化到 $s' = P$; 2) 发生传输失败事件 F , 系统当前状态 $s = N$ 就会保持停留使得 $s' = N$ 。由此可得系统 Q 值的计算过程为:

$$Q(N, P) = p_S \cdot [r_S + V(P)] + p_F \cdot [r_F + V(N)] \quad (7)$$

式中:数据包成功传输到节点 P 的概率由 p_S 表示; 传输失败的概率由 p_F 表示。转化成求解问题:

$$V(N) = \max_a \left\{ \frac{p_S \cdot [r_S + V(P)] + p_F \cdot r_F}{1 - p_F} \right\} = \max_a \left\{ V(P) + r_S + \frac{p_F}{p_S} \cdot r_F \right\} \quad (8)$$

若通过评估模型计算得到最优值函数, 则选取能让每个状态能够获得最大 Q 值的动作就是最佳的决策过程, 并将该过程称为系统的开发策略。

1.3 双平面博弈模型结合强化学习的 V2V 路由算法

V2V 通信代表了支持使用端到端多跳路由的各种上下文感知应用程序的最理想技术之一。当日常道路旅行的安全性、效率和舒适性需要提高时, 车辆间通信是实现这一目标的关键。然而, 不同的车辆移动性经常导致不可靠的连接性和相应不稳定的服务质量。因此, MANET 的传统路由协议表明它们在 VANET 中的性能很差; VANET 的路由协议应该能够克服车辆的高移动性。

为 V2V 通信开发了一种新的水平博弈模型, 作为 G^{V2V} 中的博弈玩家, 车辆动态地决定其路线。为了配置路由拓扑, 每个车辆保持两个路由信息: 到达目的地车辆的路径成本 PC 和可用路径的学习值 L 。 PC 被计算为从当前车辆到源车辆的所有中继链路成本的总和, 并且获得 L 用于自组织和独立路由控制。由于高度动态的拓扑和通信环境, V2V 通信中的无线链路不断变化。为了处理动态 VANET 条件, 通过考虑几个因素来估计当前链路状态是很重要的。本文算法在时间 H_t 定义 V_i 和 V_j 之间的无线链路状态 ($LC_{V_i}^{V_j}(H_t)$) 为:

$$LC_{V_i}^{V_j}(H_t) = (\alpha \times (1 - ((D_{V_i}^M - d_{V_i}^{V_j}) / D_{V_i}^M))) + ((1 - \alpha) \times (\Psi_{V_i}^{V_j}(H_t))) \quad (9)$$

$$\alpha = \left(\frac{F(M_{V_j}^{H_t})}{\max_{V_r \in M_{V_j}^{H_t}} \{F(M_{V_r}^{H_t})\}} \right)$$

式中: $D_{V_i}^M$ 是 V_i 的最大覆盖范围; $d_{V_i}^{V_j}$ 是 V_i 和 V_j 之间的相对距离。在无线通信中, 能量消耗与传播距离的平方成比例地减小。因此, 更近的邻近车辆更适合于路由。在式(9)中, 第一项捕获相对距离, 第二项捕获车辆的稳定性。参数 α 控制给予两个术语的相对权重。 $M_{V_j}^{H_t}$ 表示在时间 H_t 的 V_j 的相邻车辆的集合, 并且 $F(M_{V_j}^{H_t})$ 是集合 $M_{V_j}^{H_t}$ 的基数(度)。通常, $F(M_{V_j}^{H_t})$ 表示 V_j 的密度水平是从 V_i 的角度看时间 H_t 处的 V_j 的熵。一般而言, 熵是系统中无序的不确定性和度量, 它代表了拓扑变化, 它是车辆移动性对 VANET 连通性服务影响的自然量化^[7, 25-26]。本文采用熵的基本概念来支持和评估稳定的路由。 $\Psi_{V_i}^{V_j}(H_t)$ 的计算如下^[10]:

$$\Psi_{V_i}^{V_j}(H_t) = - \sum_{V_k \in M_{V_j}^{H_t}} \left(\frac{F_{V_k}(H_t, \Delta_y) \times \log F_{V_k}(H_t, \Delta_y)}{\log(F(M_{V_j}^{H_t}))} \right) \quad (10)$$

$$F_{V_k}(H_t, \Delta_y) = \frac{1}{y} \times \left(\frac{\sum_{H_x=H_{t-y}}^{H_t} (|V_j(H_x) - V_i(H_x)|)}{\sum_{H_x=H_{t-y}}^{H_t} \sum_{V_k \in M_{V_j}^{H_t}} (|V_j(H_x) - V_i(H_x)|)} \right)$$

式中: Δ_y 是从 H_{t-y} 到 H_t 的时间间隔, 可以在时间间隔

Δ_y 内计算并传播移动信息给其他相邻车辆; $V_i(H_t)$ 、 $V_j(H_x)$ 分别是时间 H_t 处的 V_i 和 V_j 的速度矢量。熵 $\Psi_{V_i}^{V_j}(H_t)$ 被归一化为 $0 \leq \Psi_{V_i}^{V_j}(H_t) \leq 1$ 。如果 $\Psi_{V_i}^{V_j}(H_t)$ 的值接近 1, 则表示路径的链路部分与中间车辆 V_j 是稳定关联的。如果 $\Psi_{V_i}^{V_j}(H_t)$ 的值接近 0, 则本地路由 $\Psi_{V_i}^{V_j}(H_t)$ 为不稳定值^[14]。

对于自适应 $LC_{V_i}^{V_j}(H_t)$ 估计, 参数 α 控制给予相应车辆的距离和熵的相对权重。在不同的 VANET 环境下, V_i 动态地修改时间 H_t 。当密度水平较轻时, 可以更加重视车辆的稳定性。在这种情况下, 较低的 α 值更合适。但是如果密度水平很高, 链路成本应该很大程度上取决于相对距离。在这种情况下, 较高的 α 值更合适。因此, 必须根据相应车辆的当前密度水平来决定 α 的值^[15]。

在路由路径形成期间, 每个单独的车辆估计其相邻车辆的 LC 值。目标车辆一跳内的车辆的 PC 值设置为从它们自己到目的地节点的单跳 LC 值。其余车辆, 即 V_j , LC 值设定如下:

$$PC_{V_i}^{V_j} = \min_{V_v} \{PC_{V_v} + LC_{V_i}^{V_v}\} \quad \text{s.t.} \quad V \in M_{V_i} \quad (11)$$

基于邻近车辆的 PC 值, 可以估计 V2V 通信的路由成本。通常, 大多数 V2V 路由算法是通过依赖于车辆协作的假设来设计的, 假设车辆愿意充当路由路径中的中继节点, 而充当中继节点的车辆必须牺牲其能量和带宽。因此, 路由算法需要刺激 VANET 车辆之间的合作行为才能有效^[16-17]。

在 V2V 路由算法中, 源车辆支付中继车辆的成本。在垂直博弈 G^{V21} 中, 每辆车通过 VCS 过程获得虚拟货币 Γ ; Γ 根据 RSU 的奖励金累积。在水平博弈 G^{V2V} 中, Γ 用于发送路由数据包。在 G^{V2V} 博弈模型中, 源车辆的效用函数可以被解释为接收的利益减去所产生的成本。作为博弈玩家, 源媒体的主要目标是最大化其收益。如果 V_i 是源车辆并且选择 V_j 作为下一个中继节点 (即 $a_{V_i}^{V_j}$), 则其在时间 H_t 处的效用函数可以被定义为:

$$H_t(U_{V_i}(V_i, H_t)) = \begin{cases} \Psi_{V_i} - J \times (Q_{V_j}^{H_t-1} \times (LC_{V_i}^{V_j} + PC_{V_i}^{V_j}))^{-1} & \Gamma^{V_i} \geq J \times (LC_{V_i}^{V_j} + PC_{V_i}^{V_j}) \\ 0 & \Gamma^{V_i} < J \times (LC_{V_i}^{V_j} + PC_{V_i}^{V_j}) \end{cases} \quad (12)$$

式中: $Q_{V_j}^{H_t-1}$ 是 V_j 在时间 H_{t+1} 时数据包传输的成功率, 如果在时间 H_{t+1} 没有选择 V_j 作为中继节点, 则 $Q_{V_j}^{H_t-1}$ 是 1; Γ^{V_i} 是 G^{V21} 期间 V_i 的累积虚拟货币; J 是分组转发成本的系数因子。源车辆 V_i 寻求选择路线路径以最大化其收益; $A_{V_i \in V} = \{a_{V_i}^{V_j} \mid V_j \in N_{V_i}\}$ 是 V_i 的可用策略集。在

G^{V21} 中, 基于强化学习方法来选择有效的路由路径。 $a_{V_i}^{V_j}$ 在时间 H_{t+1} 的 $L_{H_{t+1}}^{V_j}$ 学习值定义如下:

$$L_{H_{t+1}}^{V_j} = \begin{cases} (1 - \xi) \times L_{H_t}^{V_j} + \xi \times (u_{H_t}^{V_j} + \{\gamma^V \times (u_{H_t}^{V_j} - \sum_{\substack{V_x \in N_{V_i} \\ V_x \neq V_j}} \frac{w_{H_t}^{V_x}}{|N_{V_i}| - 1}\})\}) & H_t \text{ 时刻选择 } V_j \\ (1 - \xi) \times L_{H_t}^{V_j} + \xi \times (u_{H_t}^{V_j} + \{\gamma^V \times (u_{H_t}^{V_j} - \sum_{\substack{V_x \in N_{V_i} \\ V_x \neq V_j}} \frac{w_{H_t}^{V_x}}{|N_{V_i}| - 1}\})\}) & \text{其他} \end{cases} \quad (13)$$

即: $u_{H_t}^{V_j} = Q_{V_j}^{H_t-1} \times (U_{V_i}(V_i, H_t))$; $w_{H_t}^{V_j} = Q_{V_x}^{H_t-1} \times (LC_{V_i}^{V_j} + PC_{V_x})$ 。

式(13)的第一项捕获了 L 的历史; 第二项捕获了所选策略的相对影响。参数 ξ 控制 V_i 决策中给出两个项的相对权重。在 G^{V2V} 博弈过程中, 源车辆学习当前的 VANET 状态, 并选择具有最高 L 值的策略, 这导致分组通过最佳路线被中继到目的地车辆。为了实现有效的路径学习机制, 以与垂直博弈模型相同的方式定义战略选择分布 P^V 。假设在时间 H_t 处存在 V_i 的相邻车辆 $\{V_1, V_2, \dots, V_m\}$ 的有限集合, 并且 V_i 维持这些车辆的个体 L 值。基于该信息, 可以获得 $P^{V_i}(H_{t+1}) = \{P_{H_{t+1}}^{V_1}, P_{H_{t+1}}^{V_2}, \dots, P_{H_{t+1}}^{V_m}\}$, 它可以是由 V_i 选择的策略的概率分布。时间 H_{t+1} ($P_{H_{t+1}}^{V_i}$) 处的 $a_{V_i}^{V_j}$ 策略选择概率定义为:

$$P_{H_{t+1}}^{V_i} = \frac{\exp(L_{H_{t+1}}^{V_j})}{\sum_{V_k \in N_{V_i}} \exp(L_{H_{t+1}}^{V_k})} \quad (14)$$

V2I 和 V2V 算法的伪代码如算法 1 所示。

算法 1 V2I 和 V2V 算法

函数 V2I () {

1. 初始化参数
2. 使用式(1), 制定车辆选取策略 S^v 使收益 u^v 最大化
3. 使用式(2), 确定价值评估模型
4. 使用式(3)和(4), 确定目标函数和投递概率
5. 使用式(5)进行函数组合
6. Call V2V ()

}

函数 V2V () {

1. 使用式(6), 获得最优值
2. 使用式(7), 获得系统 Q 值
3. 使用式(8), 在源车处获得到达目的地的车辆
4. 使用式(9), 进行估计
5. 使用式(10), 获得求和

Return

}

函数 Main () {

1. 初始化控制参数。
2. 学习值并平均分配。

```
For( ; ; ) {
```

```
3. Call V2I ( )
```

```
}
}
```

1.4 算法流程

基于强化学习的垂直和水平博弈模型,通过逐步交互式博弈过程设计了一种新型 V2I 和 V2V 通信算法。该算法允许 RSU 和车辆学习当前的 VANET 情况并确定他们的最佳策略。在 V2I 博弈模型中,每个车辆通过参与 VCS 服务获得激励。在 V2V 博弈模型中,该激励被用作虚拟货币以补偿中继车辆的成本。在实际操作方面,可以将计算负担从中央系统转移到分布式车辆。所提出的 V2I 和 V2V 算法的流程如图 3 所示。

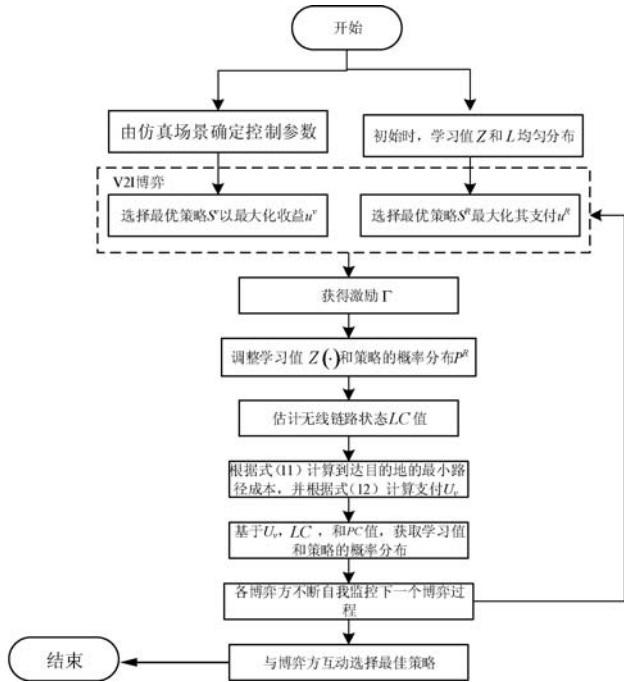


图 3 本文算法流程图

2 实验

基于 MATLAB/2016b 仿真环境搭建了用于 VANET 的仿真平台,该平台运行于 Windows 8 GB RAM 64 位 Inter(R) Core™ i5 - 7500K 2.7 GHz 台式计算机中。为验证本文方法的感知和路由性能,将其与文献[6]、文献[8]和文献[9]所提方法进行比较。文献[6]制定了一个移动众包博弈方案,其中每辆车根据传感和传输成本以及预期的支付选择其传感工作,而服务器根据其传感精度支付每辆车的成本。文献[8]中,车辆必须在数据更新之前检索所有数据单元,否则检索数据单元版本将不一致,车辆必须在规定的延迟要求范围内接收所有要求的数据单元,此外实验

方案侧重于异构车辆网络中协同状态下同态数据的分发效率。文献[9]侧重于车联网数据中的带宽效率与业务异构性,实验方案中车流量负载设置略有不同。EDDC 方案通过使用附近 RSU 的拥挤感知来设计混合路由,以保证车载网络中数据传播的质量。文献[9]评估了一种机会性拥挤感知方案,其中来自大量连接传感器的流量通过新兴的 NB-IoT 技术传输,基于仿真结果,确认了其优越性。为确保公平比较,根据文献[18],针对本文方法的实验,采用以下假设和对应的实验方案:

- 假设模拟系统是 VANET 的常见无线通信系统。
- 在 200 km 路段上随机分布了 1 200 辆车和 60 辆 RSU, 每辆移动车的速度随机选择为 40 km/h、80 km/h 或 100 km/h。
- 每辆车的最大无线覆盖范围设置为 500 m。
- VCS 的车辆成本 C 设置为 RSU 与相应车辆之间的相对距离与最大覆盖范围的比率。
- 随机选择源车辆和目的车辆。最初,每辆车的虚拟货币设定为 100。
- 数据包在源节点以 k (包/s) 的速率生成,并且提供的负载范围在 0 ~ 4.0 之间变化。因此仿真模型中的持续时间 H 为 1 s,模拟的总持续时间为 2 h。
- 假设车辆不受噪音或物理障碍的影响。
- 车辆带宽容量为 20 Mbit/s, 每条消息由 CBR 数据包组成。
- RSU 的任务和容量分别为 10 和 2 Gbit/s。
- 基于 200 次模拟运行获得的网络性能测量被绘制为每秒数据包生成的函数。

为了证明方法的有效性,测量了 RSU 的任务成功率,标准化路由吞吐量和端到端数据包延迟。表 1 为仿真实验中所使用的系统参数。

表 1 模拟实验中使用的系统参数

参数	值	描述
B_{\min}, B_{\max}	0.1, 1	预定义的最低和最高价格水平
$[B_{\min}, \dots, B_{\max}]$	[0.1, 0.3, 0.5, 0.7, 1]	RSU 任务的价格水平
L^1, L^2, L^3, L^4	11, 10, 11, 10	预定义的任务要求
β	0.4	控制参数以估计学习值 Z
γ	0.4	折扣因子来估计学习价值 Z
J	0.06	分组转发成本的系数因子
ξ	0.75	控制参数以估计学习值 L
γ	0.6	用于估计学习值 L 的折扣因子

图 4 比较了每个方案的 RSU 任务成功率。任务成功率是以任务完成相对于总 RSU 任务的百分比来

衡量的,是 V2I 通信操作中的关键因素。随着车辆的交通负荷增加,车辆致力于其自己的 V2V 通信,导致任务成功率降低。所有方案都表现出类似的趋势,然而本文方案在从低到大的交通负荷情况下优于现有方法。RSU 在交互式垂直博弈过程中为车辆支付适应性激励,车辆自发地参与了众筹工作。这种情况可以保证比其他方案更高的 RSU 任务成功率。

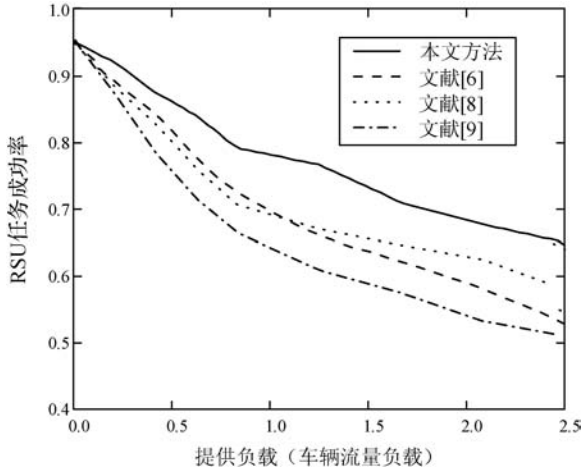


图4 RSU的任务成功率

图5比较了路由吞吐量。路由吞吐量被定义为在目的地车辆处接收的数据分组与在源车辆中生成的数据分组的总数的比率。通过本文方案实现的路由吞吐量的增益,是采用迭代学习模型的有效水平博弈范例的结果。特别地,本文方案捕获相对距离和车辆的稳定性,自适应地选择从源车辆到目的地车辆的最佳路径。该方法可以增加V2V通信中的路由吞吐量。因此,本文方案实现了优于现有方案的路由吞吐量性能,现有方案被设计为单侧协议并且不响应当前的VANET条件。

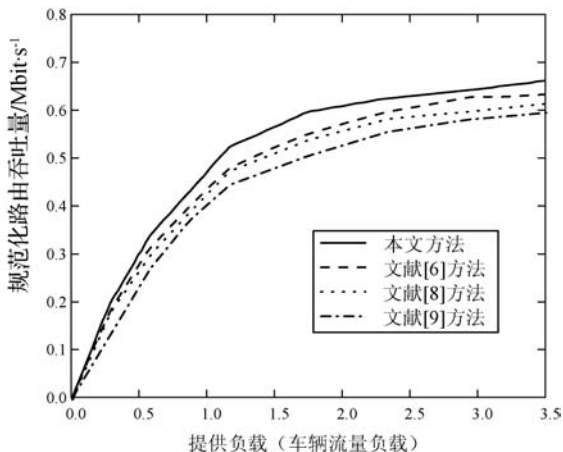


图5 规范化路由吞吐量

图6曲线表示归一化的端到端分组延迟。通常随着分组生成速率增加,分组延迟随着业务负载线性增加。由于本文方案以分布式学习方式建立路由路径,因此在每个水平博弈时段,提出的V2V通信算法中的

每个车辆通过反映VANET环境中的变化来自适应地做出路由决定。所以,本文方法比其他方案更有效地减少了分组延迟。

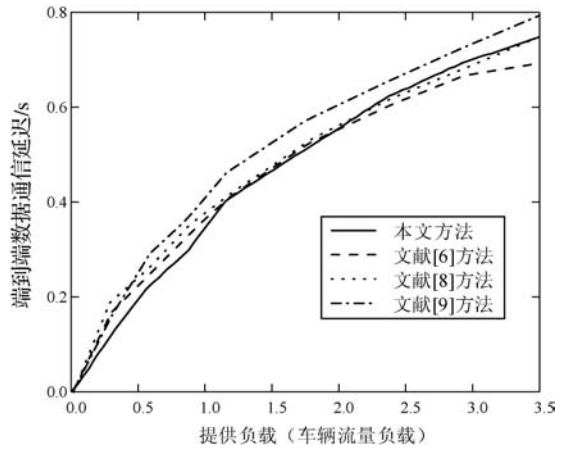


图6 规范化的端到端数据包延迟

仿真结果表明,本文方案使用基于学习的双平面博弈模型,可以监视当前的VANET条件并适应高度动态的环境。方法中的RSU和车辆从环境中获取信息,获得知识,并以自适应的方式做出明智的决策。仿真结果表明,该方案通常表现出比现有文献[6,8-9]方案更好的性能。尽管本文方案没有提供最佳解决方案,但比现有方案具有实质性优势。表1中的系统参数值是可变的,如果更改变系统参数值,则会更改模拟结果。在不同的模拟参数值下,性能趋势几乎相同。不同的仿真场景实验结果证实了本文方案优于其他现有方案。

3 结 语

在VANET中,针对V2I和V2V通信问题,本文提出了一种新型双平面博弈模型。基于VANET的特征,集成了交互式反馈机制,并基于强化学习和博弈论概念设计了纵向和横向博弈模型,开发了适用于不完备信息条件下的高效VCS和交互式路由算法。在所提出的双平面博弈模型中,单独的RSU和车辆是博弈玩家,并且为自适应的机会性VCS和路由动态地学习其策略。使用基于反馈的自我监控和分布式学习技术,博弈玩家可以动态地适应当前的VANET情况,并有效地最大化预期收益。通过仿真实验与现有方案进行对比,证明了本文算法在RSU任务成功率和吞吐量等指标的显著改善。

尽管本文算法已经实现了VANET管理中提高吞吐量和VCS任务完整性的目标,但仍有待进一步提高VANET系统效率。其主要问题包括在负载平衡、认知无线电和网络安全领域设计和验证新的VANET控制

方案。此外,当 RSU 旨在最大限度地减少对车辆的总支付时,可以扩展包括博弈模型中简单惩罚的设计机制,以优化 VANET 特征控制算法。

参 考 文 献

- [1] 张沪寅,王菁,唐星. 认知无线车载自组织网络中的联合路由调度[J]. 计算机研究与发展,2017,54(11):2445 - 2455.
- [2] 刘静,吴仲城,李芳,等. 基于 Boost. Asio 的智能车载终端数据采集系统[J]. 计算机应用与软件,2018,35(2): 248 - 255.
- [3] Yao L, Chen A L, Deng J, et al. A cooperative caching scheme based on mobility prediction in vehicular content centric networks[J]. IEEE Transactions on Vehicular Technology,2018,67(6):5435 - 5444.
- [4] 黎阳,王哲,张楚文,等. 面向车载自组织网络路由的轨迹预测算法[J]. 计算机研究与发展,2017,54(11):2421 - 2433.
- [5] Chen J Q, Mao G Q, Li C L, et al. Capacity of cooperative vehicular networks with infrastructure support: multiuser case [J]. IEEE Transactions on Vehicular Technology, 2018,67(2):1546 - 1560.
- [6] Campioni F, Choudhury S, Salomaa K, et al. Improved recruitment algorithms for vehicular crowdsensing networks [J]. IEEE Transactions on Vehicular Technology, 2019,68(2):1198 - 1207.
- [7] 任春江,降爱莲. 基于移动状态的车载自组织网络路由算法[J]. 计算机工程与科学,2018,40(3):464 - 471.
- [8] Dai P L, Liu K, Wu X, et al. Cooperative temporal data dissemination in SDN-based heterogeneous vehicular networks [J]. IEEE Internet of Things Journal, 2019,6(1):72 - 83.
- [9] Dai P L, Liu K, Wu X, et al. Bandwidth efficiency and service adaptiveness oriented data dissemination in heterogeneous vehicular networks[J]. IEEE Transactions on Vehicular Technology,2018,67(7): 6585 - 6598.
- [10] 徐哲鑫,蔡苓玲,郑丽丽,等. VANET 中基于协议序列及 TDMA 的混合信道接入机制[J]. 电子学报,2017,45(6):1408 - 1414.
- [11] 唐亚纯,乔玲玲,尹向东. 基于博弈论的 D2D 通信资源共享算法[J]. 吉林大学学报(理学版),2017,55(5):1239 - 1243.
- [12] 马成前,王利,刘畅. 基于机器学习的图像偏色检测[J]. 计算机应用与软件,2018,35(3):199 - 204.
- [13] Wang J J, Jiang C X, Zhang K, et al. Vehicular sensing networks in a smart city: Principles, technologies and applications[J]. IEEE Wireless Communications,2018,25(1): 122 - 132.
- [14] 张宏,吕悦晶. 车载自组织网络信息流量特性分析[J]. 工程数学学报,2017,34(5):449 - 457.
- [15] Wang M W, Cui Y, Wang X, et al. Machine learning for networking: Workflow, advances and opportunities [J]. IEEE Network,2018,32(2): 92 - 99.
- [16] 付彬,查理佳,李仁发,等. 软件定义的 VANET 下流表用量感知的 QoS 路由机制[J]. 计算机研究与发展,2017,54(11):213 - 223.
- [17] 杨羽琦,章国安,金喜龙. 车载自组织网络中基于车辆密度的双簇头路由协议[J]. 计算机科学,2018,45(4): 126 - 130.
- [18] Kim S. Effective crowdsensing and routing algorithms for next generation vehicular networks[J]. Wireless Networks, 2019,25:1815 - 1827.

(上接第 149 页)

- [8] Kuk J G, An J H, Ki H, et al. Fast lane detection & tracking based on Hough transform with reduced memory requirement [C] // The 13th International IEEE Conference on Intelligent Transportation Systems. IEEE, 2010.
- [9] Liu G L, Wörgötter F, Markelić I. Combining statistical hough transform and particle filter for robust lane detection and tracking [C] // 2010 IEEE Intelligent Vehicles Symposium. IEEE, 2010:993 - 997.
- [10] Aly M. Real time detection of lane markers in urban streets [C] // 2008 IEEE Intelligent Vehicles Symposium. IEEE, 2008:7 - 12.
- [11] Lin Q, Han Y, Hahn H. Real-time lane departure detection based on extended edge-linking algorithm [C] // 2010 Second International Conference on Computer Research and Development. IEEE, 2010:725 - 730.
- [12] 王科,黄智,钟志华. 基于不定 Bezier 变形模板的城市道路检测算法[J]. 机械工程学报,2013,49(8):143 - 150.
- [13] Chen Q, Wang H. A real-time lane detection algorithm based on a hyperbola-pair model [C] // 2006 IEEE Intelligent Vehicles Symposium. IEEE, 2006:510 - 515.
- [14] Bertozzi M, Broggi A. GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection [J]. IEEE Transactions on Image Processing, 1998,7(1):62 - 81.
- [15] Sánchez J, Perronnin F, Mensink T, et al. Image classification with the fisher vector: Theory and practice [J]. International journal of computer vision, 2013,105(3):222 - 245.
- [16] 钱基德,陈斌,钱基业,等. 基于感兴趣区域模型的车道线快速检测算法[J]. 电子科技大学学报,2018,47(3): 356 - 361.
- [17] 陈涵深,姚明海,陈志浩,等. 基于多帧叠加和窗口搜索的快速车道检测[J]. 计算机科学,2018,45(10):255 - 260.
- [18] Lu C T, Chou T C. Denoising of salt-and-pepper noise corrupted image using modified directional-weighted-median filter [J]. Pattern Recognition Letters, 2012,33(10):1287 - 1295.