

文章编号: 2095-2163(2024)02-0199-05

中图分类号: U121

文献标志码: A

# 软件定义车辆网络中自适应负载均衡路由算法

王文翠, 张 剑

(上海工程技术大学 航空运输学院(飞行学院), 上海 201620)

**摘要:** 城市交通车辆密度高,为解决车辆通信过程中,数据包转发时中继节点负载分配不均衡、限制车联网中吞吐量等性能问题,本文提出在基于软件定义的移动自组网络架构中引入强化路由,来自适应学习负载分配决策,根据邻居节点的带宽状态学习负载分配收益;通过强化学习构建状态-策略表,使节点在不同状态下进行带宽分配决策,最终实现 SDN 数据层内的车辆相互协调,寻找最优路径。仿真结果表明,该算法可实现网络负载的均衡分配。与传统的路由算法相比,当车辆数为 300 辆时,该算法的丢包率可低至 20% 以下,端到端时延低于 4 s,网络能量消耗更加均衡。

**关键词:** 软件定义; 车辆通信; 负载分配; 强化学习

## Adaptive load balancing routing algorithm in software defined vehicle networks

WANG Wencui, ZHANG Jian

(College of Air Transport (Flight Academy), Shanghai University of Engineering Science, Shanghai 201620, China)

**Abstract:** Urban traffic is characterized by high vehicle density, which poses challenges for vehicular communication such as unbalanced load distribution among relay nodes during data packet forwarding and limited throughput within the vehicular networks. To address these performance issues, this paper proposes the introduction of reinforced routing within a Software Defined Mobile Ad hoc Network (SDN-MANET) architecture to adaptively learn load distribution decisions. This is achieved by learning the load distribution benefits based on the bandwidth status of neighboring nodes. Through reinforcement learning, a state-policy table is constructed, allowing nodes to make bandwidth distribution decisions under varying states, ultimately enabling vehicles within the SDN data layer to coordinate with each other to find the optimal path. Simulation results indicate that the algorithm can achieve balanced network load distribution; compared to traditional routing algorithms, with a vehicular count of 300, the proposed algorithm can reduce the packet loss rate to below 20%, achieve end-to-end latency of less than 4 seconds, and result in more balanced network energy consumption.

**Key words:** software-defined; vehicle communication; load distribution; reinforcement learning

## 0 引言

随着道路交通量的增长,车辆移动自组网络(VANETs)发展也更加迅速<sup>[1]</sup>。从获取娱乐信息到对外紧急安全信息传输,VANETs可以广泛服务很多应用领域。VANETs以自组织的分布式网络为基础,由道路上行驶的车辆节点形成。VANETs的典型结构是由连续移动的车辆组成,彼此之间以及与道路沿途固定的路侧单元(RSU)进行通信<sup>[2]</sup>。车辆配备的内置车载单元(OBU)可用于车辆与车辆或者车辆与RSU之间通信。以RSU为网关的拓扑网络中,每一个车辆都可以看成该网域内的网络节点,可以通过RSU连接外部因特网或者云计算网

络。在高速公路巡逻或者移动监控应用中,移动节点利用各种传感器定期收集地面数据,将收集的数据上传到RSU<sup>[3]</sup>。为了节省能源,节点不直接发送数据包,而是通过某些路由协议算法选择中继节点转发数据包到RSU<sup>[4]</sup>。在中继节点选择过程中,某些节点可能同时被多个源节点选择进行转发到达目的节点的工作,从而造成某个中继节点出现网络拥塞情况。

## 1 研究综述

Nguyen<sup>[5]</sup>提出了GPSRMA-LA协议,将负载均衡算法引入到GPSR协议中,通过公布节点的负载、位置、速度和方向等信息到附近的其他节点,使

**作者简介:** 王文翠(1997-),女,硕士研究生,主要研究方向:智能交通、移动自组网络。

**通讯作者:** 张 剑(1974-),男,博士,教授,主要研究方向:信息与通信工程、通信与信息系统。Email:08120003@sues.edu.cn

收稿日期: 2023-01-16

中间节点选择最佳下一跳来帮助决定路由; Wu D<sup>[6]</sup>将无线网络路由和 Ad Hoc 网络结合,设计了混合 VANETs 架构的地理负载平衡路由协议,即 GLRV,根据网络状态切换路由策略以保证服务质量,并为骨干节点提供拥堵控制机制,但该协议只为 Mesh 网络实现了负载平衡机制,没有为车辆节点执行负载平衡。VLBR 协议通过从网络中获取的拥堵反馈来均衡所有潜在连接路径之间的流量,并使用 k-shortest 路径算法将连接切换到较低拥堵的路线,即确定源节点到目的节点的多条最优路径,并选择其中一条进行数据包转发,同时进行实时拥塞检测,若网络发生拥塞,就将数据包传输路径转换为第二条最优路径,进而实现网络负载均衡<sup>[7]</sup>。虽然上述研究方法能够显著的缓解负载量过大的问题,但是,在城市道路中,车辆节点具有规律性,利用交通道路上车辆行驶的规律可以更好地优化路由算法,实现自适应学习负载分配决策。本文提出将强化学习算法  $Q$ -learning 和软件定义移动自组网络(SDN)的核心思想应用到 VANETs 路由中。

通过解耦 VANETs 中的控制和数据转发平面,将负载均衡算法部署到一个逻辑层面上集中控制。引入强化路由,通过  $Q$ -learning 表的构建来自适应学习负载分配决策。本文提出的软件定义车辆网络中自适应负载均衡路由算法(RL-LB-SDVN)具有高灵活性、自适应学习和可编程性。为了降低车辆节点的能耗,本文引入 SDN 架构,全局管理控制数据层车辆节点的路由信息,并在 SDN 控制层的全局管理控制中心使用强化学习算法计算全局路由路径,然后根据全局路由路径分析制定路由决策,最终通过 SDN 南向接口下发路由转发决策到数据层,以便在选择转发的中继节点时,实现自适应负载均衡,避免网络拥塞。在网络层,将每个节点的剩余带宽和到邻居节点的距离取权值构建状态空间,实现自适应负载均衡路由选择,优化中继节点的选择和数据包的分配,实现从源节点到目的节点信息传输的高效性和稳定性。

## 2 道路车辆转发模型

构建有 12 辆车 2 个区域网的道路车辆贪婪转发模型,如图 1 所示。本文基于 SDN 架构建立从车辆 n0 到车辆 n11 的路由转发路径。基于 SDN 的 VANETs 路由架构如图 2 所示。建立路由转发路径时,首先,每个节点都会将自身车辆信息包含速度、方向和剩余带宽通过邻居节点发送到网域内的

RSU 中,RSU 将收到的信息汇总成车辆信息表,该信息表主要包括车辆 ID、邻居节点 ID、目的节点 ID、车速、方向、位置信息、总带宽量等;其次,RSU 将汇总的区域车辆信息表上传至该路段的 SDN 控制层,控制层将所有区域内的车辆信息表通过车辆 ID 进行汇总管理,建立总车辆信息,并根据特定路由算法下发数据包转发策略。

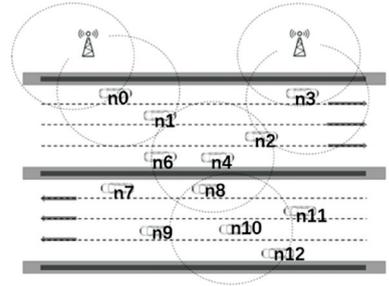


图 1 道路车辆贪婪转发模型

Fig. 1 Road vehicle greedy forwarding model

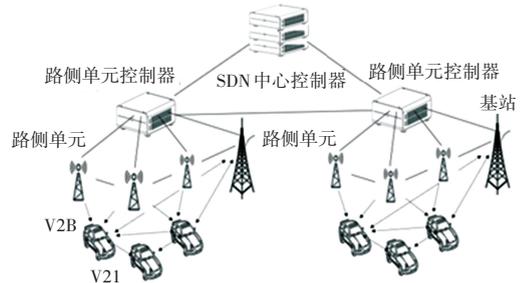


图 2 基于 SDN 的 VANETs 路由框架

Fig. 2 SDN-based VANETs routing framework

## 3 基于强化学习的负载均衡路由

在城市交通中,车辆密度过高,从源节点到目的节点建立的一条最优路径往往存在负载过大的问题,因此环境中每辆车的目标是选择一个或者多个邻居节点来传输数据信息,并最终建立起一条或多条路由路径,方便彼此之间稳定高效的通信,本文所需要解决的问题是在得知一个节点的邻居节点的剩余带宽情况后,如何使用合理的比例进行负载分配。

### 3.1 自适应强化路由在 VANETs 中的应用

与其他机器学习相比,强化学习中智能体在当前场景下做出的决策,会对未来的数据序列产生直接的影响,更符合车辆移动自组网络的现实要求<sup>[8]</sup>。因此,本文将 VANETs 路由选择问题表述为强化学习问题。

强化学习主要包括策略、收益信号和价值函数 3 个要素。策略表示智能体在某一时间的行为;在

强化学习中可以根据特定的环境所反映出的收益信号来调整对智能体的决策;与收益信号相比,价值函数表示一个智能体从这个状态开始,对将来累积的总收益的期望。

在 VANETs 环境下离散时间  $t = 0, 1, 2, \dots$  代理的目标是找到一条或者多条从原车辆到目的车辆的最佳负载均衡路由路径。假设  $N = \{0, 1, 2, \dots\}$  表示 VANETs 网络中的节点集合。在 VANETs 网络中,有一条路线从车辆  $n_i$  到  $n_d$  要经过  $n_j$ ,  $R(d)$  表示节点  $n_i$  的邻居节点集合。由于时间序列下,多个节点不同的分配决策会导致邻居节点的剩余带宽发生变化,即分配决策会影响下一跳邻居节点的负载分配比例,因此该节点的负载分配决策问题具有马尔科夫性质。同时,强化学习中需要使用马尔科夫决策过程(MDPs)来形式化描述强化学习中的环境<sup>[9]</sup>。在 MDPs 中,当前状态下执行某一个动作后,当前状态会按照一定的概率转移到另一个状态,将转移到未来状态的概率分布记为状态转移矩阵,可以记为式(1):

$$P_{ss'}^a = P[S_{t+1} = s' | S_t = s, A_t = a] \quad (1)$$

其中,  $S_t$  为节点即时的状态;  $S_{t+1}$  表示下一时刻节点状态;  $A_t$  表示节点的动作;  $P$  为状态的转移概率。

当状态转移概率不确定时,通过参考智能体在不同分配策略下的收益期望值来确定状态转移概率。针对在未知转移概率下获取收益期望值这一问题,  $Q$ -learning 可以利用状态-动作函数构建  $Q$  来很好的表示收益期望值,从而确定状态转移概率。因此本文采用  $Q$ -learning 算法来解决负载均衡问题。定义 MDPs 问题中的 3 个要素:状态空间  $S_i$ 、动作空间  $A_i$  和奖励反馈。

### 3.1.1 状态空间

智能体的状态空间表示  $S_i = \{S_i^j, S_i^k, S_i^l, \dots\}$ , 即网络中的节点可连接状态。 $S_i^j$  表示节点  $i$  的邻居节点  $j$  的可连接度,可连接度是以数值形式对剩余带宽和邻居节点到源节点的距离进行表示,定义为式(2):

$$S_i^j = \sqrt{(\psi W_i^2 - (1 - \psi) D_i^2) / 2} \quad (2)$$

其中,  $W_i$  表示节点  $i$  的剩余带宽;  $D_i$  表示邻居节点  $i$  与源节点的曼哈顿距离;  $\psi$  为剩余带宽与距离权重比例。

### 3.1.2 动作空间

动作  $A_i = \{A_i^j, A_i^k, A_i^l, \dots\}$  表示智能体的活动空间。 $A_i^j$  表示节点  $i$  向节点  $j$  发送数据包的分配比例,即从当前节点的邻居节点中选择合适的一个或多个

节点作为路由路径的下一跳节点,数据包从当前节点转发到下一跳节点负载分配策略就是一个动作,本文的动作主要表示信息传输到连接度较高的多个邻居节点的负载分配比例,动作的值为一个哈希表,关键字为可连接度高的邻居节点,值为当前节点传输到指定邻居节点的负载分配比例,这与状态空间有关,从环境状态到动作空间的映射即策略是环境所在状态和智能体所采取的动作的随机函数。

### 3.1.3 奖励反馈

每一次进行负载分配后,邻居节点的带宽和链路状况都会发生变化,网络的状态空间也会随之发生转变。为均衡下一跳节点的负载量以及网络连接情况,通过该时刻网络中节点的状态可表示出邻居节点的可连接度,由此可以利用下一时刻节点的状态来获取相应的奖励反馈,作为当前时间负载分配动作所获得的奖励。本文将状态与奖励反馈之间进行如式(3)所示的转换:

$$R_i(t) = \begin{cases} -R, & \frac{N_i^+(t+1)}{N_i^+(t+1) + N_i^-(t+1)} < 0.5 \\ R, & \frac{N_i^+(t+1)}{N_i^+(t+1) + N_i^-(t+1)} \geq 0.5 \end{cases} \quad (3)$$

其中,  $N_i^+(t+1)$  表示下一时刻节点  $i$  的邻居节点中可连接度  $S_i^j(t+1)$  大于阈值  $C$  的节点个数,  $N_i^-(t+1)$  则表示可连接度小于阈值  $C$  的邻居节点的数量。

可连接度  $S_i^j(t+1)$ , 计算公式见式(4):

$$S_i^j(t+1) = \sqrt{(\psi W_i^2(t) - (1 - \psi) D_i^2(t)) / 2} \quad (4)$$

其中,  $W_i(t) = W_i - tf_i(t)$ ;  $W_i$  表示节点  $i$  的负载量;  $tf_i(t)$  表示节点在  $t$  时刻需要发送的数据包大小。

当大多数下一跳邻居节点的可连接度情况较低时,如 20% 的邻居节点连接度高而剩余 80% 的邻居节点可连接度较低,节点的状态将转换为负反馈,反之,大部分下一跳邻居节点的可连接度情况较高时,节点的状态转换为正反馈。

## 3.2 基于 $Q$ -learning 的路由决策算法

本文的路由算法是对  $Q$ -learning 收益期望  $Q$  函数的学习,学习收敛后可得到相应的负载分配比例。在 SDN 控制层,通过一定的探索策略和  $Q$ -learning 算法更新  $Q$  函数,生成该路段内的状态-动作映射表,即  $Q$  表,  $Q$  表中的值代表某一状态下采取不同动作的收益期望值,学习收敛后可以用来确定

负载分配比例,根据反馈情况制定决策并下发。根据负载均衡需求设计  $Q$ -learning 算法更新  $Q$  函数的状态-动作映射表,见表1。

表1  $Q$  函数的状态-动作映射表

Table 1 State-action mapping table for  $Q$ -functions

状态动作	$S_1$	$S_2$	...	$S_n$
$A_1$	$Q(S_1, A_1)$	$Q(S_2, A_1)$	...	$Q(S_n, A_1)$
$A_2$	$Q(S_1, A_2)$	$Q(S_2, A_2)$	...	$Q(S_n, A_2)$
...	...	...	...	...
$A_n$	$Q(S_1, A_n)$	$Q(S_2, A_n)$	...	$Q(S_n, A_n)$

初始状态  $Q$  表中每个状态-动作价值为0,通过探索策略与状态选取当前时间段的动作  $A_i(t)$  所获得的反馈  $R_i(t)$ ,并根据  $Q$ -learning 算法更新  $Q$  函数,计算公式见式(5):

$$Q_{t+1}(S_i(t), A_i(t)) = Q_{t+1}(S_i(t), A_i(t)) + \alpha [R + \gamma \max_{A_i} Q_i(S_i(t), A_i(t)) - Q_{t+1}(S_i(t), A_i(t))] \quad (5)$$

其中,  $Q_{t+1}(S_i(t), A_i(t))$  表示  $t+1$  时刻状态更新后状态  $S_i(t)$  和动作  $A_i(t)$  所对应的  $Q$  值;  $R$  表示奖励反馈;  $\max_{A_i} Q_i(S_i(t), A_i(t))$  即在状态  $S_i(t+1)$  能取到的最大  $Q$  函数值,表示该动作对将来收益的影响;  $\alpha$  表示将来的  $Q$  值对现在的影响;  $\gamma$  为学习率。

随着信息的传输,  $Q$  函数会不断的更新直至收敛稳定。负载分配路由算法具体流程:

输入 节点  $i$ , 初始时刻  $t$ , 设定训练轮数  $T$ , 初始化  $Q$  值表, VANETs 网络节点设备, 获取  $t$  时刻车辆节点的位置和剩余带宽

If 目的节点为邻居节点 then

数据包重传直接定位到该邻居节点

Else if 目的节点不是邻居节点 then

repeat

根据探索策略, 选取负载分配动作  $A_i(t)$ ,

并按照比例传输数据包;

更新  $Q$  表、车辆位置表和剩余带宽信息;

Until  $t > T$  or  $Q$  函数收敛

repeat

获取  $t$  时刻车辆节点的位置和剩余带宽

选择  $Q$  值最大的分配比例  $A_i$

$t = t + 1$

End

## 4 仿真实验与结果分析

为了验证本文提出的软件定义车辆网络中自适

应负载均衡路由算法能够有效实现网络负载的均衡分配,改善网络性能,本文将在相同环境状态下与传统的 GPSR(节点不考虑负载均衡,利用贪婪算法转发数据)和 VLBR(采用固定的负载分配策略)进行对比实验。仿真实验评价指标包括负载均衡实现效果和网络性能评价。负载均衡的实现效果是通过测量网络传输效率来评估的,反映网络的拥塞状况,网络传输效率是一个时间片内网络吞吐量和网络生成的总负载量的比率;性能评估指标包括丢包率、端到端延迟和吞吐量,表示发送端发送的数据包在传输中未成功到达接收端的比例、到达接收端的时间和网络传输的数据量。

### 4.1 仿真实验设计

本文在交通仿真软件(SUMO)上建立城市道路交通车辆模型,并在 NS3 网络模拟器上搭建实验环境。实验的主要仿真参数见表2。

表2 仿真参数

Table 2 Simulation parameters

仿真参数	值
仿真范围	2×2 km
车辆通信范围	150 m
车辆数	50-300
车速	40 m/s
数据包大小	512 Byte
仿真时间	1 000 s

### 4.2 结果分析

#### 4.2.1 网络传输效率

每个时间片内的网络传输效率如图3所示,可见 RL-LB-SDVN 算法在初期由于负载分配策略在探索阶段,负载分配策略不断变化,传输效率值有所波动,后因算法收敛,使用负载分配最佳的决策,最终提高传输效率。

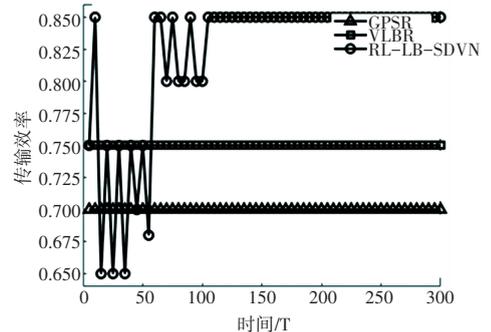


图3 每个时间片内的网络传输效率

Fig. 3 Network transmission efficiency within each time slot

#### 4.2.2 丢包率

相同车辆速度和车辆运动轨迹,数据传输中丢