

基于路口时延下界的车联网数据分发方案

彭鑫^{1,2}, 李仁发², 李哲涛^{2,3}, 谢国琪²

(1. 湖南理工学院信息与通信工程学院复杂系统优化与控制湖南省高校重点实验室, 湖南 岳阳 414000;

2. 湖南大学嵌入式与网络计算湖南省重点实验室, 湖南 长沙 410082; 3. 湘潭大学信息工程学院, 湖南 湘潭 411105)

摘要: 提出一种车联网数据分发机制。该方案利用车辆分布情况, 求出各交通路段数据传输时延, 并采用压缩感知方法构造交通网络各节点间的数据传输时延下界, 以此辅助车辆进行数据转发决策。然后通过车辆和数据分组的路径匹配度选择下一跳转发车辆。此外, 通过马尔可夫模型推导出交通路口的数据分组转发概率。仿真实验表明, 与现有的数据转发方案相比, 提出的方案具有较低的数据转发时延和较高的可靠性。

关键词: 车联网; 压缩感知; 马尔可夫过程; 数据分发

中图分类号: TP393

文献标识码: A

Data dissemination of VANET based on lower bound of road delay

PENG Xin^{1,2}, LI Ren-fa², LI Zhe-tao^{2,3}, XIE Guo-qi²

(1. Key Laboratory on Complex Systems Optimization and Controlling of Hunan High Education Institutions,

College of Information and Communication Engineering, Hunan Institute of Science and Technology, Yueyang 414000, China;

2. Key Laboratory for Embedded and Network Computing of Hunan Province, Hunan University, Changsha 410082, China;

3. College of Information Engineering, Xiangtan University, Xiangtan 411105, China)

Abstract: A reliable VANET data dissemination method was proposed. The method used the traffic information to estimate data transmission delay in each road. Compressive sensing method was used to deduce the lower bound of data transmission delay among each intersection. These information could assist carrier to choose forward routing. In the process of data forward, the vehicle that its route more similar with the forwarding data will be chosen as the next carrier. Furthermore, data forward probability at intersection was deduced based on Markov model. The simulation results demonstrate that presented method achieves lower-delay and higher reliable performance than existed packet forward protocols.

Key words: VANET, compressive sensing, Markov process, data forward

1 引言

车联网(VANET)是由配备无线通信接口的智能车辆和路边通信单元以自组织方式形成的移动自组织网络(MANET), 已成为智能交通的重要支撑技术。车联网可以对车辆的运行状态和交通路况实时感知, 从而实现交通信息发布、事故预警、辅助驾驶等服务^[1,2]。

车联网数据可靠转发是车联网各项应用服务的

技术基础。由于车联网以移动自组织网络为原型, 车辆间(V2V, vehicle to vehicle)和车辆与路边单元间(V2I, vehicle to infrastructure)的通信均以多跳“携带—转发”的方式实现。这样既降低了大规模部署基站的成本, 又避免了路边基础设施不完善对网络功能的影响。但是, 无线网络信道本身具有动态和时变特性, 链路的稳定性较差; 车辆行驶速度较快, 导致车联网拓扑动态改变, 使网络呈现间歇连通特性。因此, 上述问题给车联网的可靠数据分发带来了严峻挑

收稿日期: 2016-07-14; 修回日期: 2016-11-14

基金项目: 国家自然科学基金资助项目(No.61300039, No.61173036); 中国博士后面基金资助项目(No.2013M542109); 湖南省自然科学基金资助项目(No.14JJ3130); 湖南省教育厅重点基金资助项目(No.12A057)

Foundation Items: The National Natural Science Foundation of China (No.61300039, No.61173036), China Postdoctoral Science Foundation (No.2013M542109), The Natural Science Foundation of Hunan Province (No.14JJ3130), Research Foundation of Education Bureau of Hunan Province (No.12A057)

战^[3]。而传统多跳无线网络中的研究工作不能直接解决车联网数据分发面临的问题，仍需提出新的解决方案。

通常车联网数据分发协议主要有 2 种实现方式，包括路由表决策和交通流量统计决策。因车联网自身的特性导致路由表维持开销和交通流量信息统计开销过大，而且难以应对网络状态的改变。针对这一问题，本文提出一种基于压缩感知的数据分发方案，该方案能根据交通流量反映出的路段时延特性，实现数据的可靠转发。提出的方案利用各路段的交通流量统计信息，求得各路段的分组传输时延，并通过压缩感知方法得到各个路口间数据分组转发时延的下界。然后利用车辆的运行轨迹协助数据分组完成转发决策。由于车辆在预定交叉路口与其他车道的车辆交换数据，使数据分组向特定路段转发，所以本文根据马尔可夫排队模型，分析了数据分组在路口的转发概率模型。

2 相关研究

由于车辆的运动呈机会连通性，现有的车联网研究工作针对该问题，采用多种方式进行转发决策，下面分类介绍。

2.1 利用车辆微观运动信息进行转发决策

MOVE^[4]协议利用车辆速度信息，将靠近目标区域的车辆作为下一跳转发车辆，实现转发成功率与开销的平衡。文献[5]提出的 MaxProp 协议根据历史行车数据进行转发决策，将最可能接近目的区域的车辆作为下一跳转发车辆。文献[6]则针对车联网隐藏终端问题，选择离发送车辆最远的候选车辆转发数据分组。文献[7]通过广播路径发现数据分组建立转发路径，但是路径维持开销过大，容易产生广播风暴。文献[8]提出了分布式广播协议，使用回传机制确保数据转发。

2.2 利用宏观交通信息进行转发决策

文献[9]通过周期性广播嗅探分组获取交通信息，并将这些信息定向广播到特定路段。文献[10]利用历史交通流量统计数据，进行最小时延转发。SADV^[11]协议通过路边单元提供路口总体交通统计数据转发决策，并具备缓存功能。有些研究使用蜂窝自动机模拟车辆运动模型^[12,13]。文献[14]用流体动力学模型建立了城市交通环境下车辆连通度分析模型。文献[15,16]通过提取出租车 GPS 轨迹信息中车辆的运动相关性进行下一跳转发决策。

2.3 利用中继辅助进行转发决策

文献[17]利用城市公交车提供数据中继服务，采用马尔可夫模型估计出期望转发时延最短的路径。文献[18]提出的 CARD 方案利用路口设置的中继节点采集交通流量信息，辅助生成连通度延迟模型。文献[19]提出的 NICDM 方案委托行使方向上的最近 2 个 AP 进行协助下载，并选择合适的协助车辆将数据传送给下载车辆，实现在信号盲区也能进行下载服务。

从上述分析发现，车联网数据分发协议利用局部车辆微观信息难以获取优化的路径转发决策。利用部分车辆的宏观交通历史统计信息难以反映整体交通流量的起伏对网络连通度的影响。部分工作通过路边单元 (RSU) 的辅助，避免车联网多跳转发，提高传输的可靠性；或利用路边单元的大量部署，对特定路段的车辆队列长度进行统计，进一步优化转发决策的依据。由于 RSU 可以进行交通数据的实时采集或数据缓存，甚至还可以直接利用运营商网络进行数据转发，从而具有可靠性高的特点，而且数据传输时延稳定。不足在于需要部署完善的通信基础设施，提高了网络成本。但是车联网通信低成本的先天优势主要来自其不依赖于 RSU 的多跳转发特性，如果大量部署路边单元，则提高了车联网的部署成本。在信息基础设施建设不完善的地区，不依赖基础设施的数据分发方案仍有重要的应用价值，因此，本文仍然从多跳转发的角度，提出基于压缩感知的车联网数据分发机制。

3 数据分发策略

假设每台联网的车辆都配备有 GPS 和地理信息系统，并在地图上映射自身的行驶路径信息，同时可通过地理信息系统获取各个路段的分时交通统计信息，如 Google Map 已开始提供部分城市的交通流量统计信息。发送方传输数据前会在分组头部插入相关控制信息，分组头部格式如图 1 所示。

数据头部					数据
源节点 ID	源节点位置	目的地位置	生成时间	失效时间	

图 1 数据分组头部格式

车辆均配备有无线通信接口，通信过程为“携带—转发”方式，转发决策遵循最小转发时延原则。图 2 为车联网数据转发示意，数据源需要向路边基站单元 (BS) 发送数据。虽然从传输距离上看，路

径 R_1 比路径 R_2 更短，但 R_2 比 R_1 有更高的车流密度，数据分组沿 R_1 传输只能以车辆“携带”的方式抵达 BS，而沿 R_2 传输则通过多跳无线转发，与车辆“携带”方式相比有更小的传输时延。

如图 2 所示，车联网数据转发遵循的基本原则并非简单地以最短地图路径为转发依据，还要考虑车辆密度，尽量以多跳转发的方式进行数据传输。转发车辆根据行驶区域的不同有 3 种转发状态，分别为直路模式、路口模式和终点模式。数据分组的整个传输过程将在这 3 种状态间进行切换。车辆在 3 种状态下的转发行为并不完全一致，其中，路口模式和直路模式下要进行下一跳车辆的选择，转发机制较为复杂。

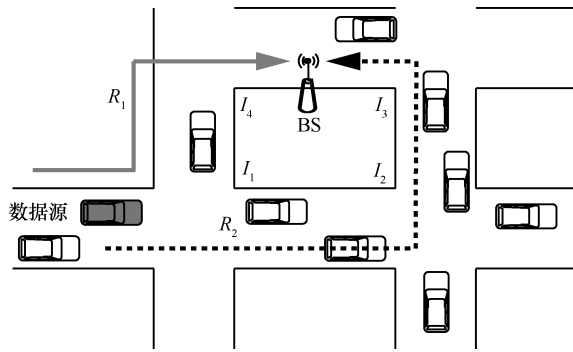


图 2 车联网数据转发示意

3.1 网络模型描述

先定义下文使用到的变量，如表 1 所示。

表 1 变量定义	
变量	定义
r_{ij}	从路口 I_i 到 I_j 的路径
d_{ij}	路径 r_{ij} 的长度
ρ_{ij}	路径 r_{ij} 上的车辆平均密度
v_{ij}	路径 r_{ij} 上车辆平均速度
t_{ij}	路径 r_{ij} 上的分组传输时延
R	车辆通信半径
τ	数据分组的单跳传输时延

令 $GR(V, E)$ 表示道路的拓扑结构，其中， V 为节点集合，代表路口； E 为边集，代表路段； $path_{i,j}$ 表示节点 i 和节点 j 间的路径， $r_{ij} = \langle i, j \rangle$ 且 $\langle i, j \rangle \in E$ 表示节点 i 和节点 j 是有道路直连的相邻路口。

假设车辆间的距离服从指数分布^[20]，即

$$f(x) = \begin{cases} \frac{1}{\rho_{ij}} e^{-\frac{1}{\rho_{ij}}x}, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (1)$$

则在路径 r_{ij} 上分组传输的平均时延为

$$t_{ij} = \left(1 - e^{-\frac{R}{\rho_{ij}}}\right)^{d_{ij}\rho_{ij}-1} \frac{d_{ij}}{R} \tau + \left[1 - \left(1 - e^{-\frac{R}{\rho_{ij}}}\right)^{d_{ij}\rho_{ij}-1}\right] \frac{d_{ij}}{2v_{ij}} \quad (2)$$

其中， $\left(1 - e^{-\frac{R}{\rho_{ij}}}\right)^{d_{ij}\rho_{ij}-1}$ 表示路径 r_{ij} 上的车辆密度较大，当间距均小于车辆通信半径 R 时数据分组以多跳传输的概率， $1 - \left(1 - e^{-\frac{R}{\rho_{ij}}}\right)^{d_{ij}\rho_{ij}-1}$ 表示路径 r_{ij} 上至少有一辆车与相邻车辆的间距大于 R 的概率，此时由于无法与相邻车辆通信，使数据分组只能采用携带的方式传输。本文假设链路断开的位置在道路上服从均匀分布，从而传输时延为 $\frac{d_{ij}}{2v_{ij}}$ 。传输时延 t_{ij} 将

作为路段 r_{ij} 的边权值 $w(path_{i,j})$ 。

根据交通信息确定 $GR(V, E)$ 的边权值 $w(path_{i,j})$ 后，最理想的数据分组转发策略是确定每个交通节点到目的地的最短路径，并使数据分组在每个路口沿最短路径传递，但这样的策略实际上并不完全可行。因为当携带数据分组的车辆到达预定的转发路口时，如果附近没有下一跳转发对象，则无法完成转发，只能由当前车辆继续沿自身的行驶路径携带。因此，转发过程，必须由携带数据分组的车辆在当前所处路段上进行连续的转发决策。为解决这一问题，本文提出一种时延下界估计策略，并依此进行转发决策。

首先，基于交通网络模型 $GR(V, E)$ 、边权值 $w(path_{i,j})$ 确定每个节点间的最短时延转发路径。通常采用 Dijkstra 算法计算节点间的最短距离。从计算复杂度方面考虑，利用 Dijkstra 算法计算网络模型中全部节点的距离至少需要 $O(n^3)$ 的时间复杂度。计算节点的权值最短路径，主要用于最短路径的数据转发决策，只要知道节点间权值的相对大小即可，无需知道确切的转发路径和转发时延。本文采用压缩感知方法，根据交通网络模型和边权值直接确定各交通节点间的最短权值路径下界，以此辅助实现最短路径时延的转发决策。由 $GR(V, E)$ 建立路径权值矩阵 $W = [w(path_{i,j})]_{n \times n}$ ，其中

$$w(path_{i,j}) = \begin{cases} t_{ij}^2, & r_{ij} \in E \\ \text{null}, & r_{ij} \notin E \end{cases} \quad (3)$$

不难看出，矩阵 W 是交通网络中相邻路口的边权值矩阵，而非相邻路口之间的路径权值矩阵。本文拟对 W 应用压缩感知确定非相邻路口节点间的

相对时延下界矩阵。但是应用压缩感知方法的前提是只有时延下界矩阵具备良好的低秩稀疏特性的情况下，才能由矩阵 W 确定。本文对矩阵 W 应用压缩感知方法的可行性进行分析。假设矩阵 $D = (\delta_{ij}^2)_{n \times n}$ 表示全部交通路口的时延下界矩阵，如果 $\langle i, j \rangle \in E$ ，则 $\delta_{ij} = t_{ij}$ 。矩阵 W 为矩阵 D 的稀疏采样矩阵。对于 D 有如下定理。

定理 1 有交通路网 $GR(V, E)$ ， V 表示路口集合，且 $|V|=n$ ， E 表示路段集合。 δ_{ij} 表示相邻路口 i 、 j 之间路段的时延，那么由 δ_{ij} 构成的交通路网时延下界矩阵 $D = (\delta_{ij}^2)_{n \times n}$ 具有稀疏低秩特性。

证明 网络中的节点代表交通路口，而路口通常为道路的交叉点，交叉路口的连通度至少为 3。矩阵 $D = (\delta_{ij}^2)_{n \times n}$ 为交通网络 $GR(V, E)$ 的时延矩阵，而 $GR(V, E)$ 是二维空间的拓扑模型，在已知坐标 $\{x_i | i=1, \dots, n\}$ 的情况下，可知

$$\delta_{ij}^2 = \frac{\|x_i - x_j\|^2}{v_{ij}^2} \quad (4)$$

其中， $\|\cdot\|$ 表示矩阵范数， v_{ij} 表示路口间的数据分组传输速度。令 Γ 表示时延向量， $I=(1,1,\dots,1)$ ，则矩阵 D 可以表示为

$$D = \Gamma^T I - 2\Gamma^T \Gamma + I^T \Gamma \quad (5)$$

其中， $(\cdot)^T$ 表示转置运算。若用 $\mathfrak{R}(\cdot)$ 表示矩阵的取秩运算符，而式(5)的 Γ 和 I 均为 n 维向量，则 $\mathfrak{R}(\Gamma^T I) = \mathfrak{R}(I^T \Gamma) = 1$ ， $\mathfrak{R}(\Gamma^T \Gamma) = 2$ ，则 $\mathfrak{R}(D) \leq 4$ 。注意到矩阵 D 的阶数即节点的个数 $n \gg 4$ ，表明矩阵 D 具有显著的稀疏低秩特征。证毕。

考虑到交通网络 $GR(V, E)$ 的节点表示路口，即 2 条道路的交点，所以，常见的十字路口连通度为 4。则根据定理 1，可以通过压缩感知方法，由稀疏矩阵 W 唯一确定矩阵 D 。矩阵 D 的压缩感知构造过程目标函数可表示为

$$\begin{aligned} & \min \sum_k \sigma_k \\ & \text{s.t. } \delta_{ij}^2 = \frac{\|x_i - x_j\|^2}{v_{ij}^2}, \langle i, j \rangle \in E \end{aligned} \quad (6)$$

其中， σ_k 表示矩阵 D 的第 k 个奇异值。根据定理 1，构造过程的秩取 4 即可。因为秩取值太高会引入不必要的开销，取值太低会造成高维信息的丢失。由于稀疏矩阵重建属于 NP-难问题，所以本文在式(6)的基础上给出基于奇异值分解的启发式算法。

截断奇异值分解是压缩感知中常用的降维运算方法。由定理 1 可知， $\mathfrak{R}(D) \leq 4$ ，可以采用秩为 4 的截断奇异值分解。对稀疏采样阵 W 有奇异值分解 $W = U\Omega V^T$ ，其中， $\Omega = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ ，并且 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ ， U 、 V 均为 n 阶正交矩阵。秩 4 重建过程可表示为

$$W_4 = \frac{n^2}{|E|} U\Omega_4 V^T = \frac{n^2}{|E|} \sum_{k=1}^4 \sigma_k u_k v_k^T \quad (7)$$

其中， $\Omega_4 = \text{diag}(\sigma_1, \dots, \sigma_4, 0, \dots, 0)$ ， u 、 v 为对应的奇异值向量，系数 $\frac{n^2}{|E|}$ 为数值补偿因子。在实际交通网络中，少数路口为三叉路口，路口连通度仅为 3。对于这部分节点，可以补充其余两跳邻居路口间的时延值，从而将该路口连通度补偿为 4，然后完成截断奇异值分解，取 $D = W_4$ 构造出矩阵 D 。算法 1 如下所示。

算法 1 时延下界矩阵 D 的构造过程
Const_D($GR(V, E)$)

- 1) for each $v \in V$;
- 2) if ($Connect(v) < 4$) Then
- 3) {if ($v_1 \in Neigh(v)$) & ($v_2 \in Neigh(v_1)$) & (v_2, v)
Then
- 4) $E = E + \langle v, v_2 \rangle$;
- 5) $d_{v,v_2} = d_{v,v_1} + d_{v_1,v_2}$;
- 6) end if;}
- 7) else break;
- 8) end for;
- 9) for $i=1:|V|$; each $\langle i, j \rangle \in E$;
- 10) for $j=1:|V|$;
- 11) if ($\langle i, j \rangle \in E$) Then
- 12) $\omega_{ij} = \delta_{ij}$;
- 13) else $\omega_{ij} = 0$;
- 14) end for;
- 15) end for;
- 16) $W = [\omega_{ij}]_{n \times n}$
- 17) $W = U\Omega V^T$

$$18) W_4 = \frac{n^2}{|E|} U\Omega_4 V^T = \frac{n^2}{|E|} \sum_{k=1}^4 u_k v_k^T$$

通过压缩感知构造的矩阵 D 可反映交通网络各节点间的数据传输时延下界。虽然利用矩阵 D 中的时延数据可以建立从当前所在路口到目的地的最短时延转发路径，但是以此作为转发决策依据并不可行，原因是携带数据分组的车辆在特定的转发路口并不能保证一定遇见驶向转发路段的车辆。因此，数据分组的转发决策无法在转发之前完成，只

能在转发过程中进行动态决策。针对这一问题，本文建立了以时延下界矩阵 D 为基础，在转发过程中对路口进行时延估计的决策模型。对于即将到达路口 I_k 的数据，通过路段 r_{kl} 到达目的地 I_l 的期望转发时延下界为

$$\varepsilon_{kl} = t_{kl} + \sum_{m \in \frac{Neigh(l)}{k}} P_{lm} (t_{lm} + t_{ml}) \quad (8)$$

其中， P_{lm} 表示数据经道路 r_{lm} 传输的概率。将式(8)应用于图2所示的交通区域。数据即将到达路口 I_1 ，那么数据通过路段 r_{12} ，经路口 I_2 抵达路口 I_3 附近目的地时延下界为

$$\varepsilon_{12} = t_{12} + P_{23}(t_{23} + t_{3r}) + P_{25}(t_{25} + t_{5r}) + P_{27}(t_{27} + t_{7r}) \quad (9)$$

由于目的地在 I_3 附近，所以式(9)中 $t_{3r}=0$ 。通过上述分析，数据分组由于只能随车流动，所以在数据源和目的地之间有若干可能的转发路径。但是，基于矩阵 D ，采用上述方式无需考察全部可能的转发路口，便可直接得到下一路口的转发时延下界。通过式(8)，携带数据分组的车辆可依次判断所经过路口的转发时延，从而实现最小转发时延下界的转发决策。如果在路口没有合适的下一跳转发车辆，则数据分组继续随车驶向下一个可能的转发路口，继续进行转发决策。

3.2 下一跳车辆的选择机制

本节主要阐述携带数据分组的车辆如何选择下一跳转发车辆。数据分组的转发，根据其所处的地理位置可分为2种状态：直路模式和路口模式。2种模式通过转发事件所处的地理位置和转发动作进行区分。数据分组在直路转发模式下，车辆采用地理转发策略，当转发窗口出现时，让数据分组迅速沿道路逼近具有最小转发时延的路口；在路口模式下，车辆选择驶向转发时延更小路段的车辆进行转发，如果没有遇到这样的车辆则自身继续携带数据分组。本文下一跳车辆选择的基本原则为，按照前文提出的路径决策方案，数据分组以位置优先的原则尽量沿最小转发时延路径传递。但是在转发过程中，下一跳车辆可能存在多个候选，下一跳转发车辆的选择机制是数据传输面临的重要问题。本文采用路径相似度匹配法甄选下一跳转发车辆。

定义1 路径 $Path_{i,j}$ 。运动对象(数据分组或车辆)在道路拓扑结构图 $GR(V, E)$ 上从节点 i 到 j 的移动轨迹，以 V 中的有序节点表示。

定义2 最长公共子序列 $LCS(A, B)$ 。如果序列

A 和序列 B 有一个公共子序列 L ，且对于 A 和 B 的任意公共子序列 K ，均有 $|L| \geq |K|$ ，则称 L 为序列 A 和 B 的最长公共子序列。

不难看出，序列的最长公共子序列可能并不唯一，也许存在多个长度一致的公共子序列。但是并不影响本文的转发决策。这里，采用动态规划法^[21]求解 $LCS(A, B)$ 。对于长度分别为 n 和 m 的序列 A 和序列 B ，构建 $n \times m$ 矩阵 C ，依据动态规划，可构造递归过程为

$$C[i, j] = \begin{cases} 0, i = 0 \text{ 或 } j = 0 \\ C[i-1, j-1] + 1, i, j > 0 \text{ 且 } A_i = B_j \\ \max(C[i, j-1], C[i-1, j]), i, j > 0 \text{ 且 } A_i \neq B_j \end{cases} \quad (10)$$

式(10)为回溯递归过程，从 $A[n]$ 和 $B[m]$ 开始依次比对，每出现一次相同的元素， $C[n-1, m-1]$ 保存的值便递增一次。比对完成以后最长公共子序列的长度便保存在 $C[0,0]$ 中。由于比对过程会遍历矩阵 C 的全部元素，所以时间复杂度为 $O(mn)$ 。比对过程中，序列 A 的每个元素均要与序列 B 的元素一次对比，同时要访问前一个元素对比的结果，所以空间复杂度为 $O(2\min(m, n))$ 。根据上述动态规划过程，建立如下求解最长公共子序列的算法2。

算法2 求解最长公共子序列

LCS(Char *A, Char *B)

- 1) int $a=\text{length}(A)$, $b=\text{length}(B)$, $C[a,b]$;
- 2) for($i=0$; $i < a$; $i++$)
- 3) for($j=0$; $j < b$; $j++$)
- 4) { if ($A[i] == B[j]$)
- 5) $C[i+1, j+1] = C[i, j] + 1$;
- 6) else $C[i+1, j+1] = \max(C[i, j+1], C[i+1, j])$;
- 7) end for;
- 8) end for ;
- 9) int $e=C[0,0]$, $k=0$, $P_1[e]$, $P_2[e,2]$;
- 10) for ($i=a$; $j >= 0$; $j--$)
- 11) for ($j=b$; $j >= 0$; $j--$)
- 12) if ($C[i, j] == C[a, b] - 1$)
- 13) { $P_1[k] = C[a, b]$, $P_2[k, 1] = i$, $P_2[k, 2] = j$;
- 14) $k = k + 1$;
- 15) break;}
- 16) end for ;
- 17) end for ;
- 18) return P_1 ;

定义3 序列相似度。序列 B 和序列 A 的相似度定义为 $Sim(A, B) = \frac{(LCS(A, B))^2}{N(A, B)}$ 。其中， $N(A, B)$

表示 B 与 A 的最长公共子序列在序列 B 中占有的最大序号值。

上述定义表示 2 个序列的相似程度并非仅取决于公共子序列的长度，还关系到公共子序列在序列 B 中的分散程度。候选车辆在接收到数据分组后，读取数据分组的源目的地，根据 $GR(V, E)$ 和矩阵 D 建立数据分组到目的地的最短时延路径，并将该路径与自身的行驶路径进行比对，计算序列的相似度。然后以地理位置优先的原则，候选车辆根据自身计算得到的序列相似度，对当前车辆发送持续时间与相似度成正比的 ACK 确认信号，形成同步 ACK 确认机制，从而让其他候选节点获知有路径更为匹配的候选车辆^[22]。如果车辆 u 接收到数据分组 C_i ，过程如下。

1) 接收到数据分组后， u 根据当前位置和 C_i 的源目的地址字段，利用矩阵 D 建立最短时延路径 $path(C_i)$ 。

2) 车辆 u 根据自身在 $GR(V, E)$ 上的行驶轨迹 $tr(u)$ 计算与 $path(C_i)$ 的路径相似度 $Sim(tr(u), path(C_i))$ 。

3) 车辆 u 广播包含自身 ID 的 ACK，并且持续时间与 $Sim(tr(u)$ 和 $path(C_i))$ 成正比。

采用上述候选车辆甄选策略，可避免贪婪地理转发策略带来的转发环路问题。

定理 2 基于地理位置的路径相似度下一跳车辆甄选策略，可以避免出现转发环路。

证明 利用反证法。根据本文转发规则，如果在转发过程中，在车辆 $Vehi_1$ 和 $Vehi_2$ 之间出现环路，意味着至少有一个同状态的数据分组从 $Vehi_1$ 发往 $Vehi_2$ ，然后又回到了 $Vehi_1$ 。在直路模式下，依据地理转发机制，数据分组从 $Vehi_1$ 发往 $Vehi_2$ 表示 $Vehi_2$ 离预定路口更近，或 $Vehi_2$ 与数据分组的路径相似度比 $Vehi_1$ 更大。而再次回到 $Vehi_1$ 表示 $Vehi_1$ 比 $Vehi_2$ 离预定路口更近，或与数据分组的路径相似度更大，显然矛盾。在路口模式下， $Vehi_1$ 将数据分组转发给 $Vehi_2$ 意味着 $Vehi_2$ 的行驶路径与数据分组的最短时延转发路径相似度 $Sim(A, B)$ 更大。如果数据分组又回到了 $Vehi_1$ ，则说明 $Vehi_1$ 的行驶路径与数据分组最佳传输路径更相似，显然矛盾。定理得证。

3.3 马尔可夫路口转发概率模型

本节主要推导道路转发概率 P_{ij} 的计算模型。根据交通基础设施，路口通常有 2 种情形：交通灯控制和高架桥。这 2 种情况下，车辆的通行方式有较

大不同，因而转发概率的计算也有所区别。根据文献[23]可知，路口车辆的到达率服从泊松分布。车辆在路口 I_i 的逗留时间用 T_i 表示，交通灯间隔时长为 I_i ，道路的宽度为 W_i ，车辆通信半径为 R 。由此推出，在高架路口车辆无需等待可直接驶过，因而逗留时长为

$$T_{ij} = \frac{w_{ij} + 2R}{v_{ij}} \quad (11)$$

但是在交通灯路口，由于有排队等待时长，所以，逗留时间的确定较为复杂。本文将逗留时长分为 2 个部分：一部分是红灯等待时长，另一部分为通过路口所用时长。后一部分不仅取决于路口宽度，还涉及路口排队长度。这里，采用马尔可夫链预测车辆在路口的平均排队长度。假设路口车辆的到达率服从参数为 λ 的泊松分布。在任意时刻 t ，路口的排队长度 $Q(t)$ 与 $t-\tau$ 时刻排队长度 $Q(t-\tau)$ 的关系可表示为

$$Q(t) = Q(t-\tau) + \alpha(A(\tau) - B(\tau)) \quad (12)$$

其中， $A(\tau)$ 表示在时长 τ 内到达路口的车辆， $B(\tau)$ 表示在时长 τ 内离开路口的车辆。显然，时刻 t 的排队长度与 $t-\tau$ 时刻的排队长度以及车辆到达率和离开率有关。所以，通过一个递归过程计算排队长度，本文将时间分为时长较短的时隙 $\Delta\tau$ 。由于时隙 $\Delta\tau$ 较短，在 $\Delta\tau$ 内队列状态仅能改变一次。那么，在 t 时刻队列长度达到 n 的概率为

$$p_n(t) = p_{n,n-1}(\Delta\tau)p_{n-1}(t-\Delta\tau) + p_{n,n}(\Delta\tau)p_n(t-\Delta\tau) + p_{n,n+1}(\Delta\tau)p_{n+1}(t-\Delta\tau) \quad (13)$$

路口车辆到达率为 λ ，离开率为 η 。则有

$$p_n(t) = \lambda\Delta\tau p_{n-1}(t-\Delta\tau) + (1-\lambda\Delta\tau-\eta\Delta\tau) \cdot p_n(t-\Delta\tau) + \eta\Delta\tau p_{n+1}(t-\Delta\tau) \quad (14)$$

整理之后可得

$$\frac{d p_n(t)}{dt} = \lambda p_{n-1}(t-\Delta\tau) - (\lambda + \eta) p_n(t-\Delta\tau) + \eta p_{n+1}(t-\Delta\tau) \quad (15)$$

在实际的车辆排队过程中，时隙内排队长度的概率可近似为常数，所以有

$$0 = \lambda p_{n-1} - (\lambda + \eta) p_n + \eta p_{n+1} \quad (16)$$

考虑到当排队长度为 0 时，离开率为 0，而且 $p_{n-1} = 0$ ，由式(16)可得

$$0 = \eta p_1 - \lambda p_0 \quad (17)$$

即

$$p_1 = \frac{\lambda}{\eta} p_0 \quad (18)$$

将式(18)代入式(16)可得

$$p_n = \left(\frac{\lambda}{\eta}\right)^n p_0 \quad (19)$$

对于概率 p_n , 满足 $\sum_{n=0}^{\infty} p_n = 1$, 所以由式(19)可得

$$p_n = \left(\frac{\lambda}{\eta}\right)^n \left(1 - \frac{\lambda}{\eta}\right) \quad (20)$$

由此可以确定路口排队长度的均值为 $E(Q(t)) = \sum_{n=0}^{\infty} np_n$ 。则车辆在交通灯路口的停留时间为

$$T_{ij} = \frac{l_i}{2} + \frac{w_{ij} + E(Q(t))z}{v_{ij}} \quad (21)$$

其中, z 为车辆平均长度。数据分组能在预定路口 I_i 向路段 r_{ij} 转发的前提是在该路口遇到至少一台驶往 r_{ij} 的车辆。令 $\zeta(T_{ij})$ 表示在 T_{ij} 时长内, 从路口 I_i 驶向路段 r_{ij} 的车辆数。由于车辆在路口的到达率服从泊松分布, 不难推出, 数据分组在路口 I_i 能够成功转发至路段 r_{ij} 的概率 PT_{ij} 为

$$\begin{aligned} PT_{ij} &= P(\zeta(T_{ij}) \geq 1) \\ &= 1 - e^{-\lambda_y T_{ij}} \frac{(\lambda_y T_{ij})^0}{0!} \\ &= 1 - e^{-\lambda_y T_{ij}} \end{aligned} \quad (22)$$

如果路口 I_i 有 n 个出度, 即 $\{r_{ij}|j=1, \dots, n\}$, 并且其转发时延满足 $\{D_{ij} \leq D_{ig}|g \geq f \& g, f=1, \dots, n\}$, 其中, g, f 表示节点。根据本文转发规则, 数据分组从路段 r_{ix} 转发的概率 $P_{ix} = PT_{ix} \prod_{i=1}^{x-1} (1 - PT_i)$ 。由于数据分组

在路口会尽量沿着有更小期望时延的道路转发, 如果车辆通过路口 I_i 后进入 r_{ix_1} 路段, 所有期望转发时延比 r_{ix_1} 大的路段其转发概率将为 0, 路口 I_i 的车辆也可能不转发而是自己继续携带数据分组。假设车辆驶向为 r_{ix_1} , 数据分组经 r_{ix_2} 转发的概率为 $P(r_{ix_2}|r_{ix_1})$, 可建立如下条件概率

$$P(r_{ix_2} | r_{ix_1}) = \begin{cases} P_{ix_2} & , D_{ix_2} < D_{ix_1} \\ \prod_{D_{ix_a} < D_{ix_1}} (1 - P_{ix_a}) & , D_{ix_2} = D_{ix_1} \\ 0 & , D_{ix_2} > D_{ix_1} \end{cases} \quad (23)$$

用 G_{ix} 表示车辆从路口 I_i 进入 r_{ix} 路段的概率。那么数据分组在路口 I_i 经路段 r_{ij} 转发的概率为

$$P_{ij} = \sum_{x \in \text{Neigh}(I)} G_{ix} P(r_{ij} | r_{ix}) \quad (24)$$

如果交通路口个数为 n , 那么, 时延下界矩阵的构造过程需要生成 n 阶矩阵, 过程中将路口连通度统一为 4, 所以计算复杂度为 $O(4n^2)$ 。在甄选下一跳车辆时, 如果对比的数据分组路径长度和候选车辆行驶路径长度分别为 a 和 b , 那么计算复杂度为 $O(ab)$, 由于 $a, b < n$, 所以计算复杂度仍为 $O(4n^2)$ 。方案的消息复杂度主要是选择下一跳转发车辆, 如果车辆连通度为 c , 在路段 r_{ij} 上的消息复杂度为 $O\left[\frac{r_{ij}}{R} c(c-1)\right]$ 。

4 仿真分析

本节利用 NS-2 和 SUMO-0.12 对提出的数据分发方案进行仿真评价^[24]。考虑到本文提出的方案无需基础设施支持, 因而与同样无需基础设施的 VADD^[10]和 CMGR^[7]车联网经典数据分发协议进行了对比。本文协议采用同步 ACK 确认的 MAC 机制, 车辆通信距离使用 DSRC 的 500 m, 具体仿真参数如表 2 所示。

表 2 仿真参数	
参数变量	变量值
车辆通信半径/m	500
车辆速度/(km·h ⁻¹)	60
发送速率/(Mbit·s ⁻¹)	1~10
MAC 协议	802.11DCF/同步 ACK
ACK 帧尺寸/byte	20
分组丢失率	0.01
仿真时间/min	280
分组生存时间/s	300

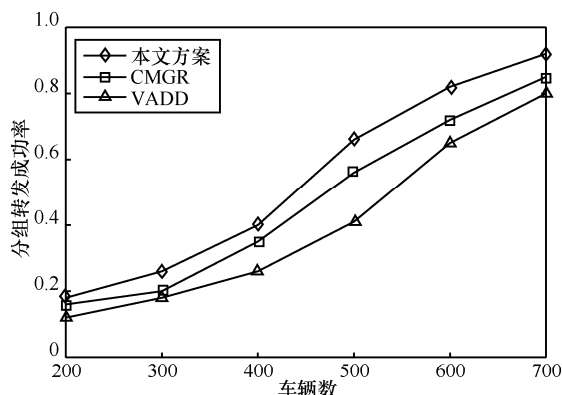
交通场景提取了长沙市星沙区 7 000 m×3 000 m 区域, 并通过 SUMO-0.26 建模为具有 34 个路口节点的仿真区域, 其中, 包含连通度为 3 和 4 的路口。所有路段均为双向路段, 路口包含信号灯控制。在交通区域的右上路口设置一个路边单元 RSU 作为目的节点, 随机选取部分车辆向目的节点发送数据分组。在模拟过程中, 车辆随机选择一个目的地, 然后以最短行驶路径到达目的地, 再随机选择下一个目的地。这样的建模方式主要考虑到车联网与传统的移动自组网的不同, 车辆在路网中并非随机移动, 而是具有自身的特定行驶轨迹, 车辆在特定的路口由于红绿灯的影响, 会停留一段时间。所以, 这样的车辆运动模型更具有现实的意义。

随机选取一定数目的车辆向 RSU 发送数据分组，数据发送速率在 1~10 Mbit/s 随机选择。

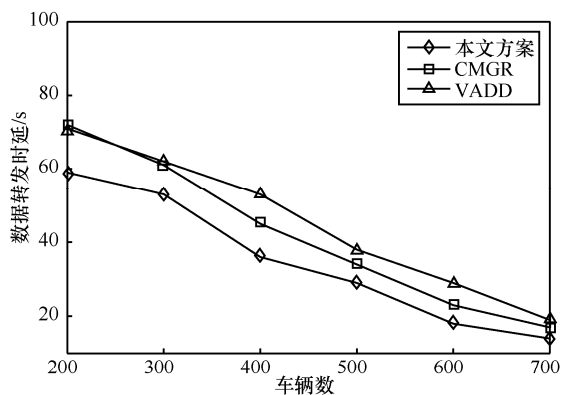
首先，改变仿真区域的车辆数目，在分别使用 30 台、50 台和 70 台会话车辆的情况下，分组转发成功率的实验结果如图 3 所示。该项实验改变车辆台数主要影响仿真区域的车辆密度，从而导致数据分组的转发率改变，也就是说当仿真区域车辆密度变小时，由于网络连通度降低，车辆自身携带数据的概率更大。从实验结果发现，随着车辆密度的增加，转发成功率逐渐提高，但是当车辆密度增加到一定程度时，转发成功率则逐渐趋于稳定。这是因

为当车辆密度较小时，网络连通度较低，车辆在更多时候选择自身携带数据，容易造成数据分组传输超时。随着车辆密度的增大，网络连通度增加，车辆在路口向下一跳车辆成功转发的概率也显著提升，而且数据分组更多采用多跳传输，提高了转发成功率。但是随着车辆密度的进一步增大，多跳传输造成链路负载更大，链路竞争激烈造成数据分组丢失。本文方案兼顾了贪婪转发和路径匹配转发，确保了在不同网络连通度下数据分组的转发成功率。

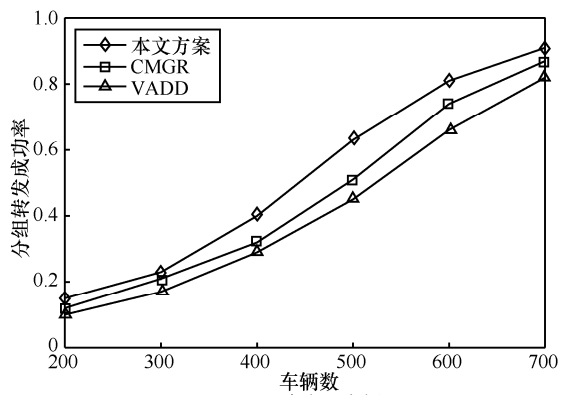
在相同实验条件下，图 4 给出了不同车辆数目时，数据转发时延的结果。实验结果显示，



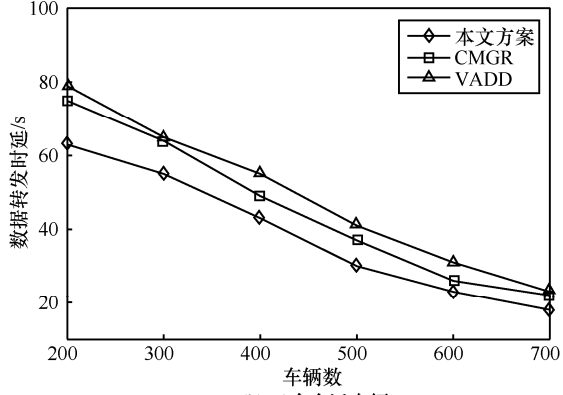
(a) 30台会话车辆



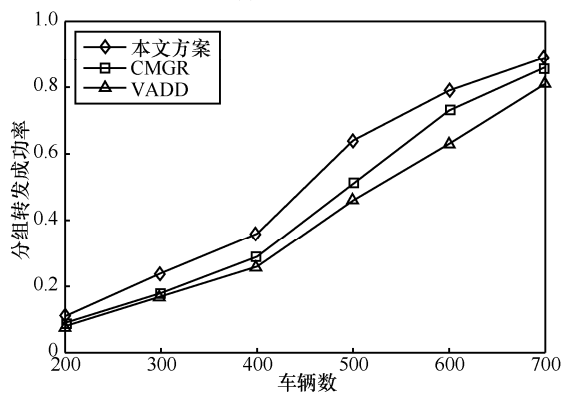
(a) 30台会话车辆



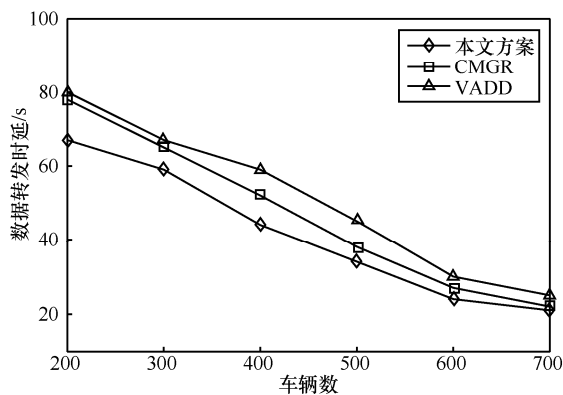
(b) 50台会话车辆



(b) 50台会话车辆



(c) 70台会话车辆



(c) 70台会话车辆

图 3 不同车辆密度条件下分组转发成功率

图 4 不同车辆密度条件下数据转发时延

随着车辆密度的增加，数据转发时延显著降低。但是随着车辆密度的进一步增加转发时延趋于稳定。这是因为车辆数目少，网络连通度低，在路口能遇到合适的下一跳车辆的概率也较低，车辆多采用携带的方式传输数据分组导致转发时延较大，甚至超时。而随着车辆数目的增大，网络连通率较好，可以降低数据转发时延。当车辆数目进一步上升，在仿真区域，尤其是交叉路口产生 MAC 冲突，影响了转发时延。

图 5 比较了 3 种方案在不用车辆密度情况下因分组生存时间和 MAC 冲突造成的分组丢失率，其中，图 5(a)中设定的分组生存时间为 300 s，图 5(b)中按跳数限定的分组生存时间为 40 跳。图 5(a)显示，随着车辆密度的增大，分组丢失率显著降低，这是由于当车辆密度较低时，网络连通度较小，数据分组主要依赖车辆携带传输，容易造成数据超时。图 5(b)结果显示，随着车辆密度增大，分组丢失率的变化没有图 5(a)显著，这是由于图 5(b)以跳数作为生存时间，车辆密度的增加可能会引起传输跳数增加，但在这种情况下，数据分组也更容易到达 RSU，故并不容易引起超时。

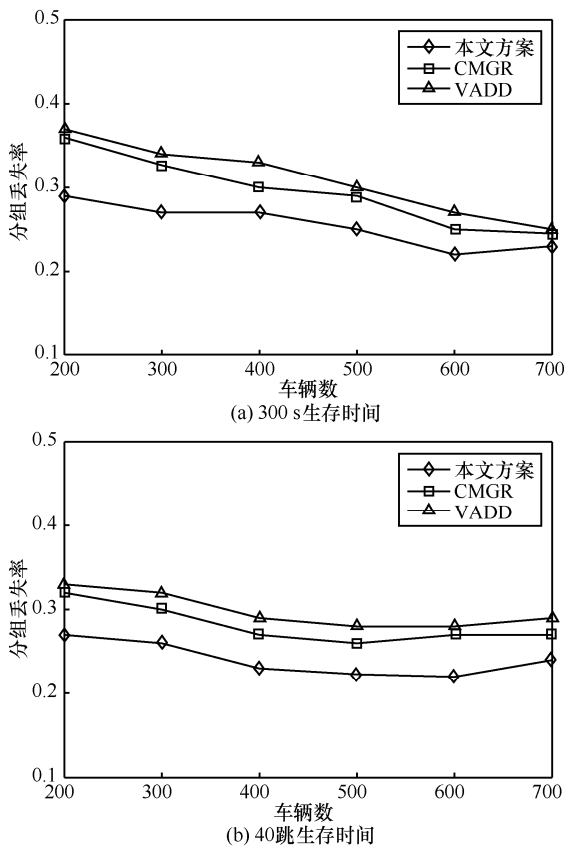


图 5 不同车辆密度条件下丢失分组率

图 6 和图 7 比较了 3 种方案在不同数据分组长度下的性能。数据分组长度的增大，需要更大的传输带宽或更长的通信窗口，也造成更多的信道访问冲突。图 6 显示，随着分组长度的增加，数据分组转发成功率显著降低，这是由于数据分组长度的增加，需要更大的通信窗口，一次传输成功率降低，车辆又处于不断运动中，因而增加了实际传输路径长度，导致分组丢失率的上升。图 7 显示，随着分组长度的增加数据分组转发时延逐渐上升。因为数据分组长度越长，信道竞争越激烈，车辆又处在运动状态，传输的失败会直接导致重新进行路径选择，导致时延的增大。在上述 2 种情况下，由于采用了路径相似度转发的原则，在重新进行路径选择时，依然保持选择更可能靠近目的节点的车辆转发数据，所以与另外 2 种方案相比，表现出更好的可靠性。

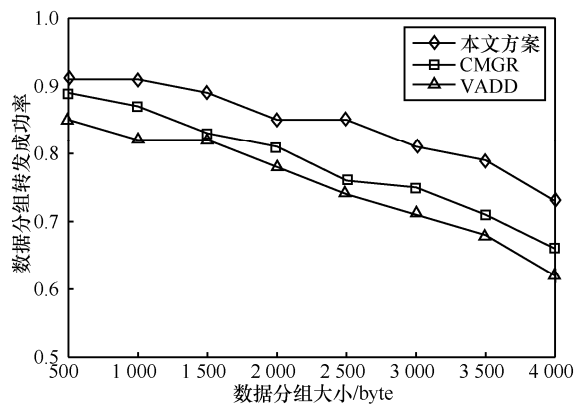


图 6 不同数据分组长度条件下数据分组转发成功率

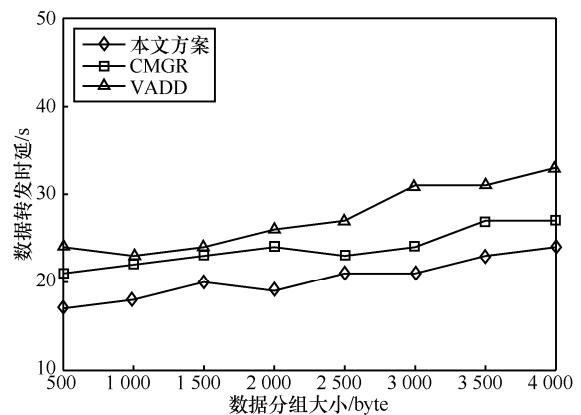


图 7 不同数据分组长度条件下数据转发时延

5 结束语

本文提出了一种车联网数据分发方案。方案根据车辆分布密度求出各路段的数据传输时延，并通

过压缩感知方法构造各交通路口间的数据传输时延下界,为数据分组的转发提供决策依据。在下一跳转发车辆选择过程中,采用以地理位置为基础的路径相似度比较的方式,有效确保数据转发的可靠性。此外,还推导了基于马尔可夫模型的交通路口数据分组转发概率。实验结果表明,本文方案与 CMGR、VADD 相比,有效提高了数据分组转发成功率,降低了传输时延。

本文下一步工作将研究网络中存在少量固定 RSU 进行辅助转发的情况下,如何利用其优化数据转发性能,以及当接收方是动态目标的情况下,如何进行有效实时的转发决策。

参考文献:

- [1] YIN J, EIBATT T, YEUNG G, et al. Performance evaluation of safety applications over DSRC vehicular ad hoc networks[C]//The VANET. 2004: 1-9.
- [2] SUN J Y, ZHU X Y, ZHANG C, et al. RescueMe: location based secure and dependable VANETs for disaster rescue[J]. IEEE Journal on Selected Areas in Communications, 2011, 29(3): 659-669.
- [3] BLUM J, ESKANDARIAN A, HOFFMAN L. Challenges of intervehicle ad hoc networks[J]. IEEE Transactions on Intelligent Transportation Systems, 2004, 5(4): 347-351.
- [4] LEBRUN J, CHUAH C N, GHOSAL D, et al. Knowledge based opportunistic forwarding in vehicular wireless ad hoc networks[C]//The 61st IEEE Vehicular Technology Conference. Piscataway, NJ: IEEE, 2005: 2289-2293.
- [5] BURGESS J, GALLAGHER B, JENSEN D, et al. MaxProp: routing for vehicle-based disruption-tolerant networks[C]//The IEEE INFOCOM 06. 2006.
- [6] GÖKHAN K, EYLEM E, FÜSÜN Ö, et al. Urban multi-hop broadcast protocol for inter-vehicle communication systems[C]//ACM International Workshop on Vehicular Ad Hoc Networks. 2004: 76-85.
- [7] SHAFIIE K, LEUNG V. Connectivity-aware minimum-delay geographic routing with vehicle tracking in VANETs[J]. Ad Hoc Networks, 2011, 9(2): 131-141.
- [8] TONGUZ O, WISITPONGPHAN N, BAI F, et al. Broadcasting in VANET[C]//The Mobile Networking for Vehicular Environments. 2007: 7-12.
- [9] NZOONTA J, RAJGURE N, WANG G, et al. VANET routing on city roads using real-time vehicular traffic information[J]. IEEE Transactions on Vehicular Technology, 2009, 58(7): 3609-3626.
- [10] ZHAO J, CAO G H. VADD: Vehicle assisted data delivery in vehicular ad hoc networks[J]. IEEE Transactions on Vehicular Technology, 2008, 57(3):1910-1920.
- [11] DING Y, XIAO L. SADV: static-node-assisted adaptive data dissemination in vehicular networks[J]. IEEE Transactions on Vehicular Technology, 2010, 59(5):2445-2455.
- [12] VIRIYASITAVAT W, BAI F, TONGUZ O K. Dynamics of network connectivity in urban vehicular networks[J]. IEEE Journal on Selected Areas in Communications, 2011, 29(3): 515-533.
- [13] NZOONTA J, NAKAYAMA M K, BORCEA C. On deriving and incorporating multi-hop path duration estimates in VANET protocols[J]. ACM Transactions on Modeling & Computer Simulation, 2011, 21(2):297-306.
- [14] HO W H, LEUNG K K, POLAK J W. Stochastic model and connectivity dynamics for VANETs in signalized road systems[J]. IEEE/ACM Transactions on Networking, 2011, 19(1):195-208.
- [15] ZHANG L, YU B, PAN J. GeoMob: a mobility-aware geocast scheme in metropolitans via taxicabs and buses[J]. Proceedings-IEEE INFOCOM, 2014:1279-1787.
- [16] ZHU H, DONG M, CHANG S, et al. ZOOM: scaling the mobility for fast opportunistic forwarding in vehicular networks[J]. Proceedings IEEE Infocom, 2013, 12(11): 2832-2840.
- [17] 陈丽, 李治军, 姜守旭, 等. 车载 ad hoc 网络中基于移动网关的数据传输[J]. 计算机学报, 2012, 35(3):454-463.
CHEN L, LI Z J, JIANG S X, et al. MGF: mobile gateway based forwarding for infrastructure-to-vehicle data delivery in vehicular ad hoc networks[J]. Chinese Journal of Computers, 2012, 35(3):454-463.
- [18] 张利锋, 金蓓弘, 嵇伟. 车载自组网络中一种连通度感知的可靠数据分发机制[J]. 计算机学报, 2013(4):701-715.
ZHANG L F, JIN B H, ZHOU W. On connectivity-aware and reliable data dissemination in VANETs[J]. Chinese Journal of Computers, 2013,36(4):701-715.
- [19] 谢永, 吴黎兵, 何炎祥, 等. 无间隙的车联网协助下载方法[J]. 通信学报, 2016(1):180-190.
XIE Y, WU L B, HE Y X, et al. Non-intermittent cooperative downloading approach for VANET[J]. Journal on Communications, 2016,37(1): 180-190.
- [20] KHABAZIAN M, ALI M K M. A performance modeling of connectivity in vehicular ad hoc networks[J]. IEEE Transactions on Vehicular Technology, 2008, 57(4):2440-2450.
- [21] 王开云. 两种基于双向比较的最长公共子串算法[J]. 中国工程物理研究院科技年报, 2013:167-170.
WANG K Y, KONG S Q, FU Y S, et al. Two longest common substring algorithms based on bi-directional comparison[J]. Journal of Computer Research and Development, 2013,50(11):2444-2454.
- [22] FOH C H, CAI J, QURESHI J. Collision codes: decoding superimposed BPSK modulated wireless transmissions[C]//Consumer Communications and NETWORKING Conference. IEEE, 2010:1-5.
- [23] OLSZEWSKI P S. Modelling of queue probability distribution at traffic signals[C]//The 11th International Symposium on Transportation and Traffic Theory. 1990:569-588.
- [24] SAHA A K, JOHNSON D B. Modeling mobility for vehicular ad hoc networks[C]//The VANET. Philadelphia. 2004: 91-92.

作者简介:



彭鑫 (1981-), 男, 湖南岳阳人, 博士, 湖南理工大学副教授, 主要研究方向为无线传感器网络、车联网。

李仁发 (1957-), 男, 湖南郴州人, 湖南大学教授、博士生导师, 主要研究方向为嵌入式系统和无线传感器网络。

李哲涛 (1980-), 男, 湖南邵阳人, 博士, 湘潭大学副教授, 主要研究方向为物联网和 CPS。

谢国琪 (1983-), 男, 湖南宁乡人, 博士, 湖南大学讲师, 主要研究方向为嵌入式系统和 CPS。