

基于运行轨迹特征分析的车辆自组织网路由算法

陶桦^{1,2}, 冯富琴^{1,2}, 肖鹏^{1,2}, 谭诚伟^{1,2}, 陶军^{1,2}

(1. 东南大学计算机科学与工程学院, 江苏 南京 210096; 2. 东南大学教育部计算机网络和信息集成重点实验室, 江苏 南京 210096)

摘要: 首先基于车辆 trace 数据提取了粗粒度的车辆移动信息, 在此基础上, 继续研究了 trace 数据的细粒度的车辆移动模型; 然后基于移动模型提出了车辆自组织网的路由算法 (RPT-D), 根据车辆移动特征将报文更快地传输到目的地; 接着将对传输的 QoS 需求放入报文选路目标中, 得到扩展性和选路结果更好的 RPT-GA 算法; 最后通过仿真实验, 分别从传输时延、投递成功率、跳数和辅助报文数量等 4 个性能参数角度, 基于车辆 trace 数据将所提出的路由算法与经典的车辆自组织网路由算法 (IGRP 和 GPSR) 进行比较, 实验结果验证了所提算法的有效性。

关键词: 车辆自组网; 车辆 trace; 路由算法; 传输时延; 投递成功率

中图分类号: TP393

文献标识码: A

Routing algorithm based on characteristics analysis of vehicle trace in vehicular ad hoc network

TAO Hua^{1,2}, FENG Fu-qin^{1,2}, XIAO Peng^{1,2}, TAN Cheng-wei^{1,2}, TAO Jun^{1,2}

(1. School of Computer Science and Engineering, Southeast University, Nanjing 210096, China;

2. Key Laboratory of Computer Network and Information Integration, MOE, Southeast University, Nanjing 210096, China)

Abstract: The coarse granularity vehicle mobility information is extracted from the vehicle trace data. Then a fine granularity mobility model was presented based on the coarse-grained mobility information. Based on the mobility model, a VANET routing algorithm, RPT-D, was proposed to quickly deliver the packets to the destination according to the mobility attributes. The RPT-GA algorithm, which was integrated with the QoS demands in the path selection objective, was designed. Finally, through the extensive simulations, the proposed algorithms are compared with other typical VANET routing algorithms, IGRP and GPSR, in terms of the transmission latency, the delivery ratio, the hop count and the extra package number. The simulation results verify the performance of the proposed algorithms.

Key words: vehicular ad hoc network, vehicle trace, routing algorithm, transmission latency, delivery ratio

1 引言

车辆自组织网简称车载网 (VANET, vehicular ad hoc network), 是一种特殊的移动自组织网络, 其利用车辆间的数据传递为驾乘人员提供交通行驶数据或娱乐信息的散播, 对智能交通系统 (ITS, intelligent transmutation system) 的实施起到了有力的支撑。车载网在 ITS 系统中的应用包括: 交互式

交通管制、实时路况分析、路线推荐、周边信息服务和车祸预警等。这些应用的开展都离不开车辆间数据传输方法的支持, 高效的数据传输依赖于底层路由。

由于车辆相对高速的移动使车载网具有拓扑变化快、车辆 (节点) 的计算能力和缓存容量较高、车辆行驶在已知道路上、车辆的行驶信息和位置信息可知等特点。正是由于这些特点, 传统移动自组

收稿日期: 2015-10-26; 修回日期: 2016-03-10

通信作者: 陶军, juntao@seu.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61272532, No.61370206, No.61370209); 教育部中国移动联合基金资助项目 (No.MCM20150502); 江苏省自然科学基金资助项目 (No.BK20151416)

Foundation Items: The National Natural Science Foundation of China (No.61272532, No.61370206, No.61370209), The MOE-CMCC Joint Foundation (No.MCM20150502), The Natural Science Foundation of Jiangsu Province (No.BK20151416)

织网络 (MANET, mobile ad hoc network) 的路由算法难以应用于车载网。现有的车辆自组织网路由在路由决策时, 缺乏对车辆行驶和道路信息的全局了解, 基于局部拓扑实施的优化操作难以获得最优的效果。实验表明, 在车载网中使用传统 MANET 路由, 数据投递率低, 且传输时延和辅助报文数较大。另一方面, 由于车辆能配备可充电装置 (如光伏或动能回收等设备), 故车载网通常不考虑诸如 MANET 中的节点能量受限问题。因此, 车载网应用需要符合其网络特点的路由算法的支持^[1]。为解决车载网中的数据运输问题, 向上层应用提供支撑, 车载网路由的研究吸引了学术界和工业界大量的关注。

为全面地了解车载网全局拓扑和车辆移动特性, 以获取较为全面的网络信息, 对现有车辆的行驶轨迹 (trace) 进行分析显得尤为必要。目前, 许多城市的出租车和公交车都已经配备 GPS, 其生成的诸如车辆速度、方向、位置、状态等的车辆轨迹信息 (车辆 trace 数据), 为车辆的运营和行驶提供了决策依据。同时出租车和公交车是城市公共交通的基础, 是构成车载网的骨干节点, 通过分析它们的轨迹信息, 有助于勾勒车载网拓扑的全局信息, 并将这些信息融入到车载网路由的制定中, 克服现有路由的局限。现有的许多研究开始重视车辆 trace 数据的研究和利用, 但这方面的研究尚不够全面, 难以体现车载网的特有特征。为此, 本文将挖掘车辆 trace 中更多的属性, 提炼天气信息与车辆分布的关系、时段信息与车辆等待时间的关系、节假日拓扑、高峰期拓扑、载客和非载客状态的行为等属性, 并将其融入到路由设计中。在此基础上, 根据车辆的 trace 数据对相关路由算法进行实际的轨迹测试和验证, 评估其效果。

2 研究现状

2.1 车载自组织网路由算法

当前, 车载网路由算法可分为单播路由和广播路由 2 类。在广播路由算法的研究中, REAR 算法根据节点的接收概率选择下一跳路由 (节点)^[1]。文献 [2] 通过改进 REAR 算法中等待时间的计算方法, 提出一种广播路由算法 RPR (receipt-based probabilistic routing)。相对于 REAR 算法, RPR 算法在报文使用数量、数据传输时延和覆盖率等方面都进行了优化。SB (smart broadcast)^[3] 是一种基于位置的分布式广

播路由算法, SB 中加入了 RTB (request-to-broadcast) 和 CTB (clear-to-broadcast) 报文, 通过增加传输的报文数量和减少二次广播时延, 获得更好的广播效果。UMB (urban multi-hop broadcast) 算法同样使用了 RTB 和 CTB 来选择广播节点^[4]。

依据路由测度的不同, 车载网路由可分为: 基于网络连通性和基于地理位置等信息的方法。在基于网络连通性的路由方面, 文献 [5] 改进了 AODV (ad hoc on-demand distance vector routing), 以适应车载网拓扑变化快、链路不稳定的特点, 同时利用同向行驶车辆的移动保持连通时间, 寻找更稳定的链接。PAODV (prior AODV) 在选择下一跳节点时, 过滤连接不稳定的节点, 解决了 AODV 中控制报文过多的问题^[6]。文献 [7] 将同方向、可一跳或多跳传输数据的车辆节点划分为簇。通过分析现有高速公路数据, 发现高速公路的车辆趋于成簇, 将车载网车辆行驶建模为分簇模型, 并简化路由的复杂性。文献 [8] 分析了簇内最后一个成员的概率、簇成员间的平均距离、簇间的平均距离、簇内平均成员数、簇的平均长度等簇特征, 然后计算了车辆稀疏环境下车辆间的连通概率, 并设计了相关的路由算法。文献 [9] 和文献 [10] 则分别使用分簇方法提出了适用于不同场景的路由算法。CAR (connectivity-aware routing) 算法通过优化之后的 PGB 算法找到目标车辆的位置, 然后设置“锚点”记录车辆的行驶轨迹, 以提高投递率^[11]。在基于地理位置的路由方面, PBR (position-based routing) 算法将基于位置的路由和多播技术相结合, 以同时保障路由算法的顽健性和高效性^[12]。文献 [13] 在 MOPR 方法中增加了邻居节点移动轨迹的预测, 提出了 MORA (movement-based routing algorithm), 延长节点间的连通时间。GPSR 算法把车载网抽象成一张地图, 每次转发数据时都将其转发给离终点最近的节点, 若无这样的节点, 则使用右手法则来拓展搜索的范围^[14]。由于 GPSR 是为广义的无线移动网络设计的贪婪转发策略, 并没有考虑车载网的特性, 因此文献 [15] 和文献 [16] 对 GPSR 在车载网中的应用做了相应的优化。文献 [17] 提出 IGRP (intersection-based geographical routing protocol) 算法, 通过分析报文发送轨迹上的十字路口信息, 评估报文投递路线, 基于每条路上的时延、跳数、CP (connectivity probability) 和 BER (bit error rate), 结合 QoS 路由判据, 计算相关路径的参数, 利用遗

传算法求解最优路径。此外，在本课题组前期的工作中，利用路边单元（RSU）的放置来增加网络拓扑的连通性，进而提高高速公路上车辆间数据散发的性能^[18]。

2.2 车辆 trace 数据集

目前，记录车辆轨迹的数据集主要包括：出租车 trace 数据和公交车 trace 数据。这 2 种车辆也正是车载自组织网络中对车流量贡献相对稳定的，且最有可能进行轨迹预测的节点。由于公交车是按照固定线路行驶，相对出租车的 trace 记录，各公交站点都可以成为公交车的网络接入点，并提供一些辅助信息，因此其 trace 的记录更加便捷，数据内容也更丰富。当前车载网中对于 trace 数据的研究主要体现在通过对车载网建模，为上层路由、数据传播、移动模拟、交互模型和节点检测等提供更为精确的网络模型和节点移动模型^[19-23]。

文献[19]设计了一种区域划分法的建模方法，首先通过计算 VKT (vehicle kilometers traveled)和 ART (accumulative residence time)等信息来识别热点，即 VKT 和 ART 达到一定数量或排名比较靠前区域，将上海交通图分为 12 个包含热点区域，计算车辆在相邻区间的转移概率和在某一区域的等待时间。车辆一旦进入某一区域，会有 2 个选择：1) 在该区域逗留一段时间；2) 离开去下一区域。依据这些移动属性来预测车辆的运行轨迹。

有些路由选择算法让车辆按照概率在多个可能的方向中选取移动方向，且赋予各方向不同概率的计算方法^[20,21]。这些方法需要以城市地图作为输入参数，将地图抽象为数据结构存储起来，并以此提取可能的路径，计算各个路径的使用概率。通常源和目的节点间的最短路径就是车辆理性行驶而选择的路径。这样车辆的行驶被抽象为：从一个点移动到另一个，停留平均等待时间后，移动到下一个，重复这样的运动。本文在前期的工作中，利用对 trace 数据的分析，探索基于位置的数据转发机制，并将车辆的 trace 数据应用于方案的测试和分析中^[24]。

本文通过对车辆 trace 的分析，提取相关的轨迹特征，并基于这些特征设计相应的车载网路由算法，以提高其性能。

3 基于 trace 数据的车辆移动模型

3.1 单位面积车辆出现次数

本文使用的车辆 trace 数据是基于上海超

过 4 000 辆出租车移动产生的 GPS 信息，记录时间为 2007 年 1 月~2 月，数据内容包括：车辆的经度、纬度、速度、方向、时间戳、是否载客等信息，车辆信息采集周期不固定（在 30 ~61 s）。此外，还将上海市地图用 100 m×100 m 的正方形方格进行划分和编号，用 m 和 n 分别表示划分网格中的横向编号和纵向编号，并基于 trace 数据统计每个方格在指定的时间 T 内车辆出现的累计次数 UAVAT (unit area vehicle appearing time)，整个网络的 UAVAT 值按照行列形成的矩阵则称为 UAVAT 矩阵。

定义 1 (UAVAT 相似度) 已知 $U_1, U_2, U_3, \dots, U_n$ 为 n 个 UAVAT 矩阵，则 $S(U_1, U_2, U_3, \dots, U_n)$ 为这 n 个矩阵的相似度

$$S(U_1, U_2, U_3, \dots, U_n) = \frac{\text{tr} \sum_{i \neq j}^s U_i U_j^T}{(n-1) \text{tr} \left(\sum_{i=1}^s U_i U_i^T \right)} \quad (1)$$

由定义 1 可知， $S(U_1, U_2, U_3, \dots, U_n)$ 反映了 $U_1, U_2, U_3, \dots, U_n$ 间相似关系，显然 $S(U_1, U_2, U_3, \dots, U_n)$ 越接近于 1，矩阵间的差异越小，反之差异越大。

首先，比较了不同时间段的 UAVAT 矩阵结果，通过式 (1)，可计算出 UAVAT 矩阵间的相似度。表 1 为上海 31 天内各个时间段的 UAVAT 矩阵相似度。由表 1 可知，每个时间段与相邻时间段的相似程度都较高，而与其他时间段的相似程度则降低。可以看出，城市的 UAVAT 矩阵随着时间不断变化，不能单纯以天为单位来研究相关车辆的属性，因此，进一步将一天划分为 12 个时间段，计算各时间段的 UAVAT 矩阵，来描述车载网拓扑的时间变化。

在此基础上，以天为单位，计算了一周内相同时间段的 UAVAT 矩阵，由表 2 可见，城市中出租车的行驶状态不仅与时间段相关，还与日期的特殊性有关：周一到周五的 UAVAT 矩阵相似度较大；而周末（周六与周日）较为相似，但彼此间 UAVAT 矩阵的相似度较小。

3.2 单位面积内车辆停留时间

单位面积内车辆停留时间 UAVRT (unit area vehicle residence time)，在文献[16]中首次被提出，作为判断某个单位区域热度的依据。本文也将采用该尺度来衡量各单元格的相关属性，UAVRT 的计算方法与 UAVAT 基本相似，只需将车辆停留次数替换为车辆停留时间。

表 1 同一天不同时间段的 *UAVAT* 矩阵相似度

时间段	0~1	2~3	4~5	6~7	8~9	10~11	12~13	14~15	16~17	18~19	20~21	22~23
0~1	1.000 0	0.823 4	0.717 8	0.833 3	0.861 9	0.873 4	0.849 0	0.839 7	0.854 7	0.841 4	0.835 8	0.848 2
2~3	0.823 4	1.000 0	0.864 8	0.661 5	0.654 8	0.641 0	0.599 4	0.585 1	0.609 9	0.589 7	0.577 3	0.588 6
4~5	0.717 8	0.864 8	1.000 0	0.664 8	0.583 5	0.562 3	0.535 7	0.516 8	0.538 3	0.520 1	0.509 1	0.520 9
6~7	0.833 3	0.661 5	0.664 8	1.000 0	0.858 7	0.849 3	0.856 3	0.835 3	0.855 6	0.845 5	0.818 2	0.817 0
8~9	0.861 9	0.654 8	0.583 5	0.858 7	1.000 0	0.943 8	0.931 9	0.926 0	0.930 2	0.911 6	0.885 2	0.885 4
10~11	0.873 4	0.641 0	0.562 3	0.849 3	0.943 8	1.000 0	0.964 3	0.963 2	0.957 4	0.942 1	0.929 8	0.929 8
12~13	0.849 0	0.599 4	0.535 7	0.856 3	0.931 9	0.964 3	1.000 0	0.973 7	0.963 0	0.953 0	0.944 7	0.945 5
14~15	0.839 7	0.585 1	0.516 8	0.835 3	0.926 0	0.963 2	0.973 7	1.000 0	0.966 8	0.959 7	0.952 0	0.950 4
16~17	0.854 7	0.609 9	0.538 3	0.855 6	0.930 2	0.957 4	0.963 0	0.966 8	1.000 0	0.970 1	0.953 8	0.947 6
18~19	0.841 4	0.589 7	0.520 1	0.845 5	0.911 6	0.942 1	0.953 0	0.959 7	0.970 1	1.000 0	0.963 6	0.953 9
20~21	0.835 8	0.577 3	0.509 1	0.818 2	0.885 2	0.929 8	0.944 7	0.952 0	0.953 8	0.963 6	1.000 0	0.973 5
22~23	0.848 2	0.588 6	0.520 9	0.817 0	0.885 4	0.929 8	0.945 5	0.950 4	0.947 6	0.953 9	0.973 5	1.000 0

表 2 不同日期相同时间段的 *UAVAT* 矩阵比较

日期	周一	周二	周三	周四	周五	周六	周日
周一	1.000 0	0.974 0	0.973 8	0.970 9	0.972 8	0.922 6	0.884 7
周二	0.974 0	1.000 0	0.972 1	0.969 5	0.971 0	0.914 9	0.878 5
周三	0.973 8	0.972 1	1.000 0	0.974 6	0.974 1	0.920 3	0.879 9
周四	0.970 9	0.969 5	0.974 6	1.000 0	0.974 3	0.930 0	0.893 2
周五	0.972 8	0.971 0	0.974 1	0.974 3	1.000 0	0.921 8	0.880 7
周六	0.922 6	0.914 9	0.920 3	0.930 0	0.921 8	1.000 0	0.957 8
周日	0.884 7	0.878 5	0.879 9	0.893 2	0.880 7	0.957 8	1.000 0

首先将一天分为 12 个时间段，计算每个时间段的 *UAVRT* 矩阵以及彼此之间的相似度。与 *UAVAT* 相同，时间段不同的 *UAVRT* 矩阵之间的相似度比较低，因此与 *UAVAT* 矩阵的计算一样，为了减少实验值的误差，这里不使用统一的 *UAVRT* 矩阵来表示所有的情况。

表 3 为一周时间内相同时间段的 *UAVRT* 矩阵的比较，可以看出 5 个工作日期间车辆的行驶状态比较接近，周末 2 天车辆的行驶状态比较接近，但

是彼此间的状态差距较大。

3.3 车辆的移动模型

通过上文的分析，证明城市的 *UAVAT* 和 *UAVRT* 矩阵不仅与时间段相关，而且与周末及工作日相关。如果把所有时间下的车辆轨迹数据都抽象为一个 *UAVAT* 矩阵和 *UAVRT* 矩阵是不合理的，不能够正确反映车载网的拓扑和车辆的密度。因此，将历史 trace 数据分为 2 种场景：工作日和周末。同时，将所有的历史 trace 数据分为 12 个时间

表 3 不同日期相同时间段的 *UAVRT* 矩阵比较

日期	周一	周二	周三	周四	周五	周六	周日
周一	1.000 0	0.951 0	0.951 0	0.937 2	0.929 8	0.830 0	0.753 7
周二	0.951 0	1.000 0	0.949 9	0.939 8	0.931 0	0.838 5	0.769 6
周三	0.951 0	0.949 9	1.000 0	0.944 2	0.932 1	0.833 2	0.759 8
周四	0.937 2	0.939 8	0.944 2	1.000 0	0.937 3	0.830 1	0.764 4
周五	0.929 8	0.931 0	0.932 1	0.937 3	1.000 0	0.822 6	0.763 8
周六	0.830 0	0.838 5	0.833 2	0.830 1	0.822 6	1.000 0	0.922 2
周日	0.753 7	0.769 6	0.759 8	0.764 4	0.763 8	0.922 2	1.000 0

段, 00:00:00 ~ 01:59:59、02:00:00~03:59:59, ..., 22:00:00 ~ 23:59:59。然后, 为每个时间段计算 *UAVAT* 矩阵和 *UAVRT* 矩阵, 求出各个时段 *UAVAT* 矩阵和 *UAVRT* 矩阵的平均值, 这些平均值可以作为车辆的移动模型计算的依据。

定义 2 A_{wd} 为工作日的 *UAVAT* 矩阵平均值数组

$$A_{wd} = \{A_{wd}[i] | i = 1, 2, L, 12\}$$

其中, $A_{wd}[i]$ 表示第 i 个时间段的 *UAVAT* 矩阵平均值。

定义 3 R_{wd} 为工作日的 *UAVRT* 矩阵平均值数组

$$R_{wd} = \{R_{wd}[i] | i = 1, 2, L, 12\}$$

其中, $R_{wd}[i]$ 为时间段 i 的 *UAVRT* 矩阵平均值。

定义 4 A_{we} 为周末的 *UAVAT* 矩阵平均值数组

$$A_{we} = \{A_{we}[i] | i = 1, 2, L, 12\}$$

定义 5 R_{we} 为周末的 *UAVRT* 矩阵平均值数组

$$R_{we} = \{R_{we}[i] | i = 1, 2, L, 12\}。$$

本文使用四元组 $\langle A_{wd}, R_{wd}, A_{we}, R_{we} \rangle$ 表示车流量模型。下面对本文提出的模型 A_{wd} 、 R_{wd} 进行验证, 从

已有的 *trace* 中选出 12 个工作日, 其中, 10 个工作日用来计算 A_{wd} 、 R_{wd} , 另外 2 天 (1 月 26 日和 27 日) 用来验证计算出来的 A_{wd} 和 R_{wd} 参数的准确性。

表 4 给出了 26 日、27 日与 A_{wd} 的 *UAVAT* 矩阵的相似度对比, 将这 2 天对应时间段的 *UAVAT* 矩阵与 A_{wd} 进行对比。可以看出, 它们的 *UAVAT* 矩阵相似度很高, 平均相似度为 0.98, 最小相似度为 0.96。同样, 表 5 也给出相应 *UAVRT* 相似度对比, 其相似度也较高, 说明采用此方法对后续工作日的 *UAVAT* 矩阵和 *UAVRT* 矩阵进行预测是可行的。

最后, 本文使用同样的方法对 A_{we} 、 R_{we} 进行验证, 选出 6 个周末, 其中, 4 个数据用来计算 A_{we} , 其余 2 个日期 (24 日和 25 日) 用来验证模型的准确性。表 6 和表 7 为对比结果。可以看出 A_{we} 、 R_{we} 与这 2 个验证日期的矩阵相似度较高, 使用这样的方式对周末的 *UAVAT* 矩阵和 *UAVRT* 矩阵进行预测是较为可信的。

4 基于 *trace* 的车载网路由的研究

下面, 本文将提出一种基于 *trace* 的路由协议 (RPT, routing protocol based on trace data), RPT 算法是属于基于地理位置信息的路由算法, 以车辆 *trace* 得到其移动模型, 通过遗传算法 (genetic algorithm) 和 Dijkstra 算法计算最优路径。

表 4 A_{wd} 与 1 月 26 日和 27 日 *UAVAT* 矩阵相似度

日期	$A_{wd}[1]$	$A_{wd}[2]$	$A_{wd}[3]$	$A_{wd}[4]$	$A_{wd}[5]$	$A_{wd}[6]$	$A_{wd}[7]$	$A_{wd}[8]$	$A_{wd}[9]$	$A_{wd}[10]$	$A_{wd}[11]$	$A_{wd}[12]$
1 月 26 日	0.972 7	0.959 4	0.960 4	0.983 0	0.985 3	0.983 7	0.986 1	0.985 2	0.985 4	0.983 2	0.983 5	0.982 0
1 月 27 日	0.972 4	0.955 1	0.959 7	0.974 6	0.986 2	0.985 7	0.983 8	0.984 9	0.986 3	0.986 5	0.981 2	0.978 8

表 5 R_{wd} 与 1 月 26 日和 27 日 *UAVRT* 矩阵相似度

日期	$R_{wd}[1]$	$R_{wd}[2]$	$R_{wd}[3]$	$R_{wd}[4]$	$R_{wd}[5]$	$R_{wd}[6]$	$R_{wd}[7]$	$R_{wd}[8]$	$R_{wd}[9]$	$R_{wd}[10]$	$R_{wd}[11]$	$R_{wd}[12]$
1 月 26 日	0.935 7	0.949 0	0.936 2	0.958 5	0.974 3	0.972 7	0.971 2	0.971 6	0.968 2	0.969 2	0.970 7	0.963 5
1 月 27 日	0.937 5	0.938 1	0.942 3	0.946 7	0.970 0	0.967 4	0.965 7	0.970 7	0.967 0	0.966 2	0.969 4	0.966 7

表 6 A_{we} 与 1 月 24 日和 25 日 *UAVAT* 矩阵相似度

日期	$A_{we}[1]$	$A_{we}[2]$	$A_{we}[3]$	$A_{we}[4]$	$A_{we}[5]$	$A_{we}[6]$	$A_{we}[7]$	$A_{we}[8]$	$A_{we}[9]$	$A_{we}[10]$	$A_{we}[11]$	$A_{we}[12]$
1 月 24 日	0.979 1	0.978 2	0.975 5	0.979 4	0.982 2	0.985 4	0.986 5	0.987 2	0.988 0	0.983 8	0.983 0	0.982 7
1 月 25 日	0.978 3	0.976 1	0.975 4	0.979 5	0.983 5	0.987 6	0.986 0	0.987 7	0.987 0	0.985 4	0.986 8	0.978 7

表 7 R_{we} 与 1 月 24 日和 25 日 *UAVRT* 矩阵相似度

日期	$R_{we}[1]$	$R_{we}[2]$	$R_{we}[3]$	$R_{we}[4]$	$R_{we}[5]$	$R_{we}[6]$	$R_{we}[7]$	$R_{we}[8]$	$R_{we}[9]$	$R_{we}[10]$	$R_{we}[11]$	$R_{we}[12]$
1 月 24 日	0.963 0	0.946 0	0.916 5	0.947 7	0.967 7	0.970 3	0.972 5	0.975 6	0.974 7	0.973 5	0.967 6	0.968 2
1 月 25 日	0.956 6	0.938 4	0.907 0	0.938 1	0.966 5	0.972 9	0.973 5	0.975 8	0.978 3	0.973 8	0.971 0	0.964 3

4.1 数据传输时间

数据的传输有 2 种方式：1) 车辆周围没有适合的节点，车辆携带数据分组；2) 车辆将数据通过无线方式传递给下一跳。第 2 种方式所需的传输时间可忽略不计。假设车辆通过当前单元格到达目标区域（单元格）的平均时间为 t ，目标区域内有节点的概率为 p ，则数据通过单元格的平均传输时间 $T=t(1-p)$ 。

1) 车辆到目标区域的平均时间

由车辆 trace 可知车辆各时间段内在单元格的平均速度。假设车辆到达目标区域需要经过 k 个单元格，那么可得车辆在其中行驶的距离 s_k ，结合各单元格对应的平均速度 v_k ，可求得到达目标区域的平均时间为 $t_{travel} = \sum_{k=1}^n \frac{s_k}{v_k}$ 。

2) 单元格内存在车辆节点的概率

每个单元格在过去的 T 时间段内通过车辆节点的数量为 $A_{we}[i][m,n]$ 或者 $A_{wd}[i][m,n]$ ，其中， m 和 n 表示划分网格中的横向编号和纵向编号，文献[16]和文献[25]证明了车辆节点到达符合泊松分布，那么单元格每间隔 T 的到达强度为

$$\lambda = \frac{A_{wd}[i][m,n]}{R_{wd}[i][m,n]} = \frac{R_{wd}[i][m,n]}{T}$$

车辆到达的概率可计算为

$$P(x=k) = \frac{e^{-\lambda} \lambda^k}{k!} = \frac{e^{-\frac{R_{ij}[m,n]}{T}} \left(\frac{R_{wd}[m,n]}{T}\right)^k}{k!}$$

这样，单元格内存在节点的概率为

$$p=1-P(0)=1-e^{-\frac{R_{wd}[i][m,n]}{T}}$$

根据车辆通过单元格的平均时间 t_{travel} 和单元格内有一个或多个车辆节点的概率 p ，数据分组通过单元格的平均时间 t_{pkg}

$$t_{pkg} = e^{-\frac{R_{wd}[i][m,n]}{T}} t_{travel} = e^{-R_{wd}[i][m,n]} \sum_{k=1}^n \frac{s_k}{v_k} \quad (2)$$

4.2 RPT-D 算法

由于本文将城市地图划分成 $100\text{ m} \times 100\text{ m}$ 的单元格，车辆节点位于相应的单元格内，报文可以传递到源节点传输范围内的相关节点，其延迟为 t_{pkg} ，如式(2)所示。此时车载网中的路由问题可以建模为在网络拓扑中寻找目的地代价最小的路径，

本文通过 Dijkstra 算法来求解，并将算法记为 RPT-D。

1) RPT-D 算法分析

通过斐波那契堆和最小优先级队列优化之后实现的 Dijkstra 算法时间复杂度是 $O(|E|+|V|\log|V|)$ ，其中， $|E|$ 是拓扑中边的数量。考虑到由矩阵抽象出的拓扑边数较小，使用矩阵来优化拓扑的边数。例如，拓扑边的数量为 $\frac{25 \times 220\,000}{2} = 2\,750\,000$ 时，优化后的算法

复杂度为 $O(2\,750\,000+220\,000)=O(10^6)$ 。根据当前计算机的计算能力，RPT-D 算法所需时间不超过 0.01 s。

2) RPT-D 算法的问题

使用 RPT-D 算法得到路径的时延是最短的，但是如果需要路由考虑其他限制条件，如跳数、误码率等参数时，RPT-D 算法往往难以同时满足多个优化条件。

4.3 RPT-GA 算法

为了进行多目标优化，本文采用遗传算法进行具有多个优化目标参数的最优路径的求解。这里，增加一个路由判据：跳数，故现在路由判据为源与目的节点间的时延和跳数。本文将求解算法记为 RPT-GA(routing protocol based on trace using genetic algorithm)，RPT-GA 算法通过如下步骤完成。

Step 1 确定编码方式和初始化族群。

路径使用一系列二元组表示，单个二元组标明了相关矩阵中位置，因此，一系列二元组便标记出从源到目的的路径，然后使用广度优先的方法来初始化族群，族群的初始化大小为 N 。

Step 2 选择操作。

假设路径的编号为 $1, 2, 3, \dots, N$ ，则每条路径的跳数记为 HC_i ，车辆能够通过多跳的方式将报文传输出去的条件是该车辆下一跳区域内存在车辆，依据单元格内存在节点的概率 p ，则

$$HC_i = \sum p = \sum (1 - e^{-\frac{R_{ij}[i][m,n]}{T}})$$

首先，依据 HC_i 对路径进行筛选，将族群中所有 HC 值大于 HC_{th} 的路径全部剔除。这样剩下的路径都是满足 HC 小于 HC_{th} 的路径。其中， HC_{th} 为假定阈值，所有路径须满足跳数小于 HC_{th} 。然后，将路径时延记为 D_i ，则

$$D_i = \sum t_{pkg}$$

所有路径的时延之和记为 D_{sum} ，则

$$D_{\text{sum}} = \sum_1^N D_i$$

此时，路径 i 被排除的概率为

$$p_{\text{reject}} = \frac{D_i}{D_{\text{sum}}}$$

因此，时延大的路径被排除的可能性就越大，循环此操作，直到群组数量为 M 。

Step 3 基于概率的交叉优化。

接下来，将从路径族群中任选 2 条路径，依据概率 p_{cross} 进行交叉操作，即从 2 条路径中提取公共节点，然后将 2 条路径上位于公共节点后的部分交换。如存在多个公共节点，则随机选择一个作为路径交换的起始点。通过路径交叉的方式得到更多的路径，并增加路径优化的可能性，从而优化最终的族群。

Step 4 执行变异。

与 Step3 相似，进行变异操作基于概率 p_{mutation} 进行。与 Step 3 不同的是变异的对象为一条路径。具体的操作方式有 2 种：1) 在传输范围内将 2 跳传输合并为一跳；2) 将一跳传输随机地拆分为 2 跳。这 2 种变异方式按照 50% 的概率进行，即如果选择对一条路径进行变异，则变异的时候以 50% 的概率按照方式 1 变异，50% 的概率按照方式 2 变异，这 2 种变异操作在实际中都有意义。由于有可能路径上的节点很多，因此路径微小的改变便可能影响整体的优化结果。

与 RPT-D 算法相比，RPT-GA 算法的结果可能不是时延最短的，但所得路径的累积跳数却能满足相应的要求和约定，使算法能够兼顾两方面的优化效果。同时，RPT-GA 算法还表现出较强的扩展性。例如，可以考虑除本文提到的跳数和时延以外的约束条件，可加入相应的 QoS 约束条件（误码率、连通概率等）。

5 仿真实验和实验结果分析

下面采用 NS-2 工具对 RPT-D 算法和 RPT-GA 算法进行仿真实验。为了更好地评估 2 种算法的性能，还对经典的基于位置的路由算法 GPSR 算法和 IGRP 算法进行实现，在传输时延、跳数、辅助报文数量等性能指标上对它们进行横向对比和分析。

首先根据在上海市范围内车辆 trace 的数据对不同时间段拓扑变化进行模拟，使用不同时间段的 UAVRT 矩阵，网络节点数根据模拟时间段 trace 数据中的车辆数而确定（总数量约 4 000 左右）。仿真

中，默认车辆节点的传输半径是 250 m。车辆的移动轨迹将按照 trace 数据生成，各车辆节点根据其 trace 数据行驶。对网络拓扑进行多组实验，每组实验的数据传输距离，即报文发送位置（源节点）与接收端（目的地）间的距离在 [0, 25 000] m 随机变化，且每个传输将要传输报文 100 个，仿真持续时间为 300 s。

5.1 不同平均报文传输距离的传输性能

为排除其他因素的干扰，在对比报文不同的平均传输距离对传输性能的影响时，本文使用了相同时间段的 trace 进行分析，所涉及的 trace 数据为 2007-01-26 08:00:00 ~ 2007-01-30 08:05:00 时间段。

在图 1 中，传输时延随着传输距离的增大而增大，显然报文的传输距离越长，需要的时间也就越多。由图 1 可见，RPT-D 算法的时延是最小的。由于 RPT-D 是全局优化的路由选择，因此，时延是唯一的选择判据；IGRP 算法则能够对 RSU 覆盖范围内的节点路由进行局部优化，因此，IGRP 所获得的传输时延较 RPT-D 略高；RPT-GA 则使用遗传算法对路径依据时延和跳数 2 个参数进行优化，因此，时延不再是唯一的考虑，还需要兼顾跳数的优化，使算法的执行力最高，但 RPT-GA 算法的时延却相对较高。

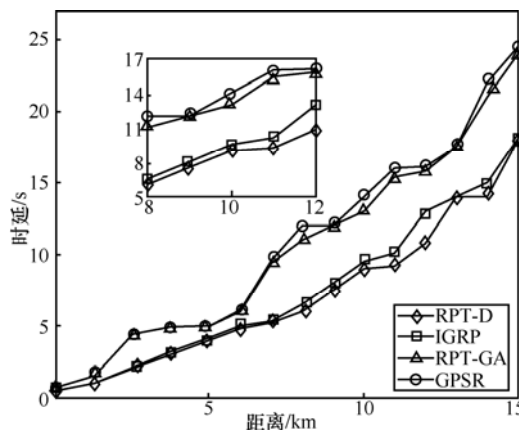


图 1 传输距离与时延之间的关系

在图 2 中，分析了距离对跳数的影响，显然跳数会随传输距离的增加而增大，且跳数与时延呈反比的趋势。RPT-GA 算法所获得的平均跳数最小，因为在 RPT-GA 算法中，跳数作为优化目标之一对遗传算法的相关路径进行筛选，将删除跳数大的路径，保证了路径的最大跳数值。GPSR 算法跳数较小的原因在于其贪婪地选择距离目的地最近的节点加入路径，此时，虽然路径的跳数较少，但有可

能将报文传递到节点较少的区域，此时将只能靠路过的车辆自己携带报文投递，传输延迟可能会倍增。RPT-D 算法和 IGRP 算法跳数较大的原因在于它们会为了寻找时延更短路径，放弃跳数更少的路径。

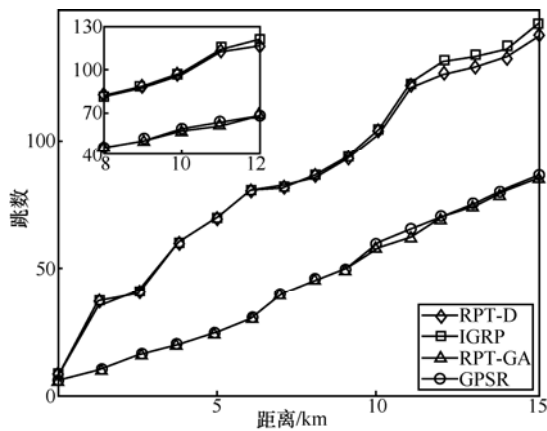


图 2 传输距离与跳数之间的关系

在图 3 中，RPT-D 和 RPT-GA 算法使用的辅助报文数较少，究其原因这是 2 种算法都对节点的辅助报文数量进行了限制。IGRP 算法使用的辅助报文数量较之其他算法高了很多，原因是 IGRP 算法常需要 RSU 的辅助，因此，车辆需将其数据中继到 RSU，再由 RSU 进行传递。

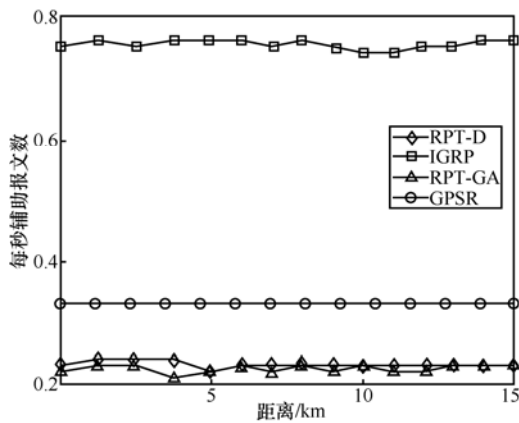


图 3 传输距离与辅助报文数之间的关系

由图 4 可知，报文投递率与传输时延的走势相似，这是由于时延较小的方法选取的路径往往跳数较大，即连通度较好的网络中的路径，这类路径经过的区域往往车辆相对密集，报文能较快地传向目的地。此时，不会造成报文长期滞留在某个车辆节点，又因难以找到合适的下一跳，造成报文超时或缓冲区的溢出而丢弃。因此，RPT-D 算法和 IGRP 算法的报文投递率较高，其他算法的投递率较低。

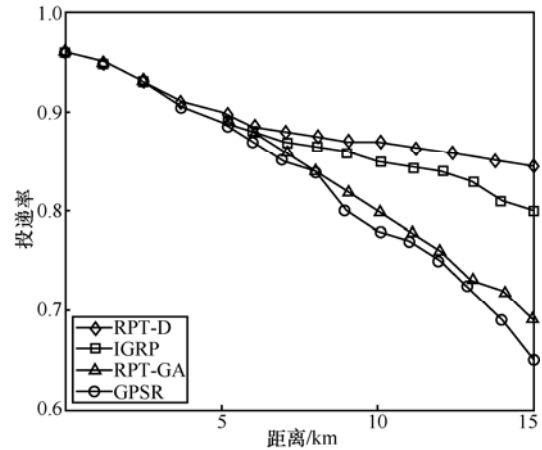


图 4 距离与投递率之间的关系

5.2 不同时段网络的传输效果

为了对比不同时段网络拓扑对传输的影响，本文对每一组 UAVAT 和 UAVRT 的 12 个不同网络拓扑，在数据传输距离均为 8 km 的场景中，对相关传输性能进行实验。

图 5 呈现的是 4 种路由算法在不同拓扑中网络时延的表现，X 轴为 trace 中的相关时间段，如 X 轴的 5 表示 04:00:00~05:00:00 时间段的传输时延。结果表明在早晚高峰时段，由于车辆密集、车速缓慢，报文不需借助 Carry-and-Forward 的方式传输，大部分可以通过车辆间多跳传输完成，故路由算法的传输时延均有所降低。

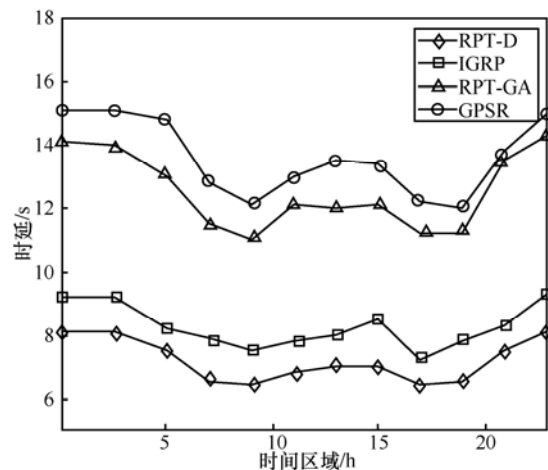


图 5 不同时段下拓扑对时延的影响

从图 6 中可看到，在早晚高峰时段，由于车辆数和车辆密度增加，多跳传输成为优选，相应的传输跳数不断增加。RPT-D 算法和 IGRP 算法则能有更多优化的多跳传输路径可供选择，在缩短传输延迟的同时，增加了一些传输跳数。由于 RPT-GA 算

法和 GPSR 算法更注重将报文及时地送到离目的地更近的节点，因此，车辆数和密度的增加对于这 2 种算法影响不大，选择的路径会更接近源节点和目的地位置间的最小跳数。另一方面，这 2 种算法在车辆密度较大的道路上会优先使用多跳的传输，因此，在车辆密度和数量增大时，这 2 种算法所需跳数会相应的增加。

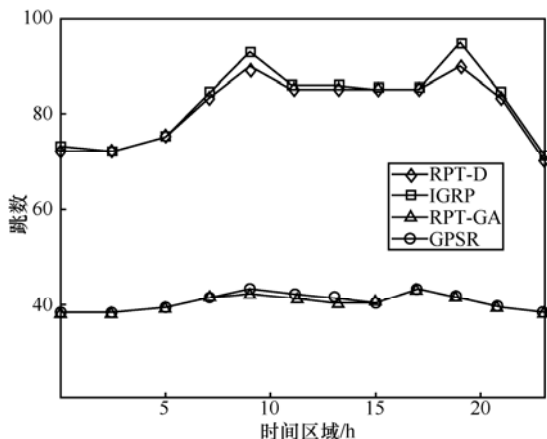


图 6 不同时段拓扑对跳数的影响

如图 7 所示，4 种路由算法中只有 IGRP 算法需要借助 RSU 进行通信，因此，IGRP 算法使用的辅助报文数会随着车辆的增多而上升；其他路由算法由于不需转发辅助报文，因此，辅助报文数与网络拓扑的其他节点关联较少。正如前面的分析，RPT-D 和 RPT-GA 算法与车辆行驶速度相关，在车流高峰期，车速较慢，车辆需要的辅助报文数较少；而车辆密度和数量的增加，使车辆间交互的 HelloBeacon 报文数也相应地上升，因此，RPT-D 和 RPT-GA 算法使用辅助报文数在高峰期和非高峰期差别不大。

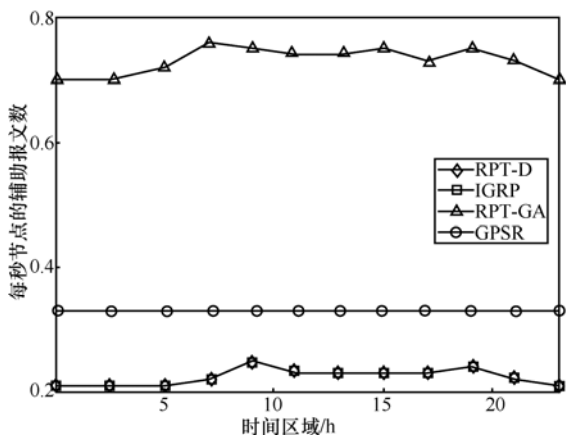


图 7 不同时段下拓扑对辅助报文数的影响

如图 8 所示，相关路由由算法的投递率与其时延的走势相吻合。在车辆高峰期，4 个算法的投递率都很高，同样的原因：车辆密度增大，车速降低，使报文的传递更快速稳定。

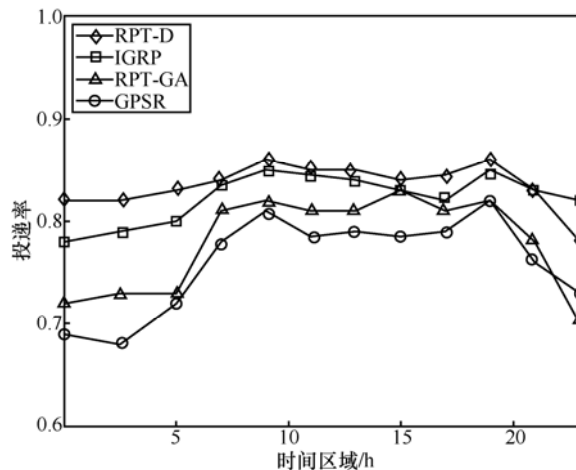


图 8 不同时段拓扑对投递率的影响

由于本实验中所使用的车辆 trace 数据均来自于实际环境中的车辆行驶轨迹，因此，在使用 NS-2 模拟时，车辆的方向、速度及行驶模式都能够再现真实的车联网环境，实验结果能够验证本文的相关路由算法的性能和可用性。

6 结束语

本文通过分析车辆的 trace 数据，将车联网的拓扑与实际车流量联系起来，分析了不同时段（工作日、双休日以及节假日）的车联网拓扑情况，进而更准确地预测未来的网络拓扑，为路由的设计提供支持。在此基础上，本文提出了 2 种基于车辆 trace 模型的路由算法：RPT-D 算法使用经典的 Dijkstra 算法计算下一跳节点，将 2 个区间的距离转化为报文通过这 2 个区间的预计时间，在计算传输时延时不仅考虑车辆在不同单元格内的速度，也兼顾了下一跳单元格内车辆存在的概率，以更精确地估算时间；而 RPT-GA 则在 RPT-D 基础上，增加了新的路由判据。

最后，通过仿真实验评估了 RPT-D 和 RPT-GA 算法的有效性，与经典的路由算法 IGRP 和 GPSR 进行横向比较。仿真实验结果表明，RPT-D 和 RPT-GA 算法在传输时延、跳数、投递成功率和辅助报文数等方面均有较好的表现，验证了其有效性和可靠性。

参考文献:

- [1] JIANG H, GUO H, CHEN L. Reliable and efficient alarm message routing in VANET[C]// 28th International Conference on Distributed Computing Systems Workshops, ICDCS'08. c2008: 186-191.
- [2] TAO J, XIAO P, LIU Y, et al. A routing algorithm based on the probability of topology connectivity in vehicular ad-hoc networks[J]. Journal of Southeast University, 2013,(3): 1-4.
- [3] FASOLO E, ZANELLA A, ZORZI M. An effective broadcast scheme for alert message propagation in vehicular ad hoc networks[C]// IEEE International Conference on ICC'06. c2006: 3960-3965.
- [4] KORKMAZ G, EKICI E, ÖZGÜNER F, et al. Urban multi-hop broadcast protocol for inter-vehicle communication systems[C]// Proceedings of the 1st ACM International Workshop on Vehicular Ad Hoc Networks. ACM, c2004: 76-85.
- [5] DING B, CHEN Z, WANG Y, et al. An improved AODV routing protocol for VANET[C]// 2011 International Conference on Wireless Communications and Signal Processing (WCSP). c2011: 1-5.
- [6] ABEDI O, BERANGI R, AZGOMI M A. Improving route stability and overhead on AODV routing protocol and make it usable for VANET[C]// 29th IEEE International Conference on Distributed Computing Systems Workshops, ICDCS Workshops' 09. c2009: 464-467.
- [7] ABEDI O, FATHY M, TAGHILOO J. Enhancing AODV routing protocol using mobility parameters in VANET[C]// IEEE/ACS International Conference on Computer Systems and Applications. c2008: 229-235.
- [8] WISITPONGPHAN N, BAI F, MUDALIGE P, et al. Routing in sparse vehicular ad hoc wireless networks[J]. IEEE Journal on Selected Areas in Communications, 2007, 25(8): 1538-1556.
- [9] BLUM J, ESKANDARIAN A, HOFFMAN L. Mobility management in IVC networks[C]// Proceedings IEEE Intelligent Vehicles Symposium. c2003: 150-155.
- [10] JAAP S, BECHLER M, WOLF L. Evaluation of routing protocols for vehicular ad hoc networks in typical road traffic scenarios[C]//11th EUNICE Open European Summer School on Networked Applications, c2005: 584-602.
- [11] NAUMOV V, GROSS T R. Connectivity-aware routing (CAR) in vehicular ad-hoc networks[C]//INFOCOM 26th IEEE International Conference on Computer Communications, c2007: 1919-1927.
- [12] HARSCH C, FESTAG A, PAPADIMITRATOS P. Secure position-based routing for VANET[C]// IEEE 66th Vehicular Technology Conference. c2007: 26-30.
- [13] MENOVAR M, LENARDI M, FILALI F. A movement prediction based routing protocol for vehicle-to-vehicle communications[J]. Communications, 2005, 21: 07-2005.
- [14] KARP B, KUNG H T. GPSR: Greedy perimeter stateless routing for wireless networks[C]//The 6th Annual International Conference on Mobile Computing and Networking. ACM, c2000: 243-254.
- [15] LI F, WANG Y. Routing in vehicular ad hoc networks: a survey[J]. IEEE Vehicular Technology Magazine, 2007, 2(2): 12-22.
- [16] RAO S A, PAI M C, BOUSSEDJRA M, et al. GPSR-L: Greedy perimeter stateless routing with lifetime for VANETS[C]//8th International Conference on ITS Telecommunications, c2008: 299-304.
- [17] SALEET H, LANGAR R, NAIK K, et al. Intersection-based geographical routing protocol for VANETs: a proposal and analysis[J]. IEEE Transactions on Vehicular Technology, 2011, 60(9): 4560-4574.
- [18] TAO J, ZHU L, WANG X, et al. RSU deployment scheme with power control for highway message propagation in VANETs[C]. IEEE Global Communications Conference (GLOBECOM). c2014: 169-174.
- [19] LI X, PAN G, WU Z, et al. Prediction of urban human mobility using large-scale taxi traces and its applications[J]. Frontiers of Computer Science, 2012, 6(1): 111-121.
- [20] KIM M, KOTZ D, KIM S. Extracting a mobility model from real user traces[C]//INFOCOM. c2006: 1-13.
- [21] ZHANG X, KUROSE J, LEVINE B N, et al. Study of a bus-based disruption-tolerant network: mobility modeling and impact on routing[C]//The 13th Annual ACM International Conference on Mobile Computing and Networking. ACM, c2007: 195-206.
- [22] YOON J, NOBLE B D, LIU M, et al. Building realistic mobility models from coarse-grained traces[C]//The 4th International Conference on Mobile Systems, Applications and Services. ACM, c2006: 177-190.
- [23] ZHANG L, YU B, PAN J. GeoMob: a mobility-aware geocast scheme in metropolitans via taxicabs and buses[C]//IEEE INFOCOM, c2014: 1279-1287.
- [24] TAO J, XU Y, TAN C, et al. Location-aware opportunistic forwarding in mobile opportunistic networks[C]// IEEE Wireless Communications and Networking Conference (WCNC). c2015: 1847-1852.
- [25] PARK J S, LEE U, OH S Y, et al. Emergency related video streaming in VANET using network coding[C]//The 3rd International Workshop on Vehicular Ad Hoc Networks. ACM, c2006: 102-103.

作者简介:



陶桦 (1976-), 男, 江苏吴江人, 东南大学博士生, 主要研究方向为计算机网络技术、网络安全。



冯富琴 (1992-), 女, 江苏盐城人, 东南大学硕士生, 主要研究方向为延迟容忍网络中的缓存管理及路由策略。

肖鹏 (1988-), 男, 江苏海门人, 东南大学硕士生, 主要研究方向为车辆自组织网络中的路由。

谭诚伟 (1989-), 男, 浙江嘉兴人, 东南大学硕士生, 主要研究方向为机会社会网络、无线传感网络。

陶军 (1975-), 男, 江苏南京人, 东南大学副教授、博士生导师, 主要研究方向为无线网络、高性能网络、信息经济学及分布式计算与系统。