

无线移动多信道感知网络上的数据聚集传输规划

冯 诚 李治军 姜守旭

(哈尔滨工业大学计算机科学与技术系 哈尔滨 150001)

摘 要 智能交通系统上的实时导航、交通监控、安全保障等应用需要实时地收集大量的交通流量、车辆位置以及道路通行时间等信息。车上很多传感器(如 GPS)可以实时采集这些数据,而由车及路边设备构成的车联网可以实时地传输这些数据。面对大量感知数据的实时传输,无线通信能力远远不能满足传输需求,而车移动导致的链路变化又使得无线信道资源很难被高效利用。该文综合了 3 个方面内容来解决这一问题:考虑基于 802.11p 的多信道数据传输提高无线传输能力;利用数据聚集降低数据传输量;利用网络动态变化结构规划高效的传输。该文提出的数据聚集规划包括两个步骤:首先是构建一棵带信道个数约束的优化聚集路由树。基本思想是通过选择连通时刻较多的链路作为树边优化传输时刻可规划的空间。同时建树时需要满足干扰半径内的树间冲突数量不大于信道数量,实现多信道无冲突并行传输;然后规划优化的传输时刻。其核心是在避免树内冲突的前提下利用有限的无线网络资源将尽可能多的数据聚集到 AP 节点。该文给出了一个动态规划算法实现了最优传输规划。该文在真实轨迹数据集和模拟轨迹数据集上进行了大量的对比实验。结果表明该文的方法在数据收集率上比现有算法提高了近 1/4,同时也减少了数据平均收集延迟。

关键词 车联网;数据聚集;动态规划;多信道;数据收集;物联网

中图法分类号 TP301 **DOI 号** 10.11897/SP.J.1016.2016.00931

Data Aggregation Scheduling in Wireless Mobile Multichannel Sensor Networks

FENG Cheng LI Zhi-Jun JIANG Shou-Xu

(Department of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001)

Abstract Real-time navigation, traffic monitoring, warning broadcasts applications in Intelligent Transportation Systems need a large number of real-time traffic volume, vehicle locations travel time through the roads and so on. Many kinds of sensors on the vehicles (e. g. GPS) can sense these data. Vehicular networks composed by vehicles and access points can transmit these data. A mass of real-time data need be transmitted. But the capacity of wireless communication cannot satisfy the transmission requirements. Moreover, the frequent change of wireless links caused by the mobility of vehicles makes wireless resource difficult to be utilized. This paper synthesizes three aspects to solve this problem: utilize multi-channel transmission of 802.11p protocol to improve the capacity of wireless communication; utilize data aggregation to decrease the amount of data transmission; utilize dynamic communication topology to schedule effective transmissions. This paper proposes data aggregation transmission scheduling, which includes two steps. The first step is to create an optimal aggregation routing tree with channel constraints. The basic idea is to choose the links with more connected timeslots as edges on the tree to optimize the space of transmission time scheduling. At the same time, the algorithm needs to satisfy that the amount of inter-tree conflicts in the interference range is less than the amount of channels in order to realize multi-channel transmission without conflicts. The second step is to schedule the optimal

收稿日期:2014-12-28;在线出版日期:2015-07-23。本课题得到国家自然科学基金(61370214,60803148)资助。冯 诚,女,1986 年生,博士研究生,主要研究方向为移动感知网络上的数据收集、数据聚集技术。E-mail: fengcheng7@gmail.com。李治军,男,1977 年生,博士,副教授,主要研究方向为 P2P 网络、普适计算、传感器网络、操作系统。姜守旭(通信作者),男,1968 年生,博士,教授,博士生导师,主要研究领域为 P2P 网络、数据库、传感器网络、普适计算。E-mail: jsx@hit.edu.cn。

transmission time. The key point is to utilize limited wireless resource to aggregate data to AP as much as possible without transmission conflicts. In this part the paper proposes a dynamic programming algorithm to get optimal transmission scheduling. Finally, we conduct a mass of simulation experiments on a real taxi trajectory data set. The result demonstrates that compared with existing data aggregation algorithms, our algorithm improves nearly 1/4 in the data collection rate, and decreases the average delay.

Keywords VANET; data aggregation; dynamic programming; multi-channel; data collection; Internet of Things

1 引 言

近年来,随着无线通信技术和感知技术的发展,车用自组织网络(Vehicular Adhoc NETworks^[1]或 Vehicular Sensor Networks^[2])成为无线移动通信领域的新的热点研究,受到工业界和学术界的极大关注.车用自组织网络是由移动的节点和静态的路边接入点 AP (Access Point)形成的无线移动感知网络.车用自组织网络为智能交通系统(Intelligent Transportation Systems,ITS)^[3]的应用提供了实现的平台.

实时导航^[4]、安全警示广播^[5]等许多 ITS 上的

应用都需要实时车辆的速度位置等信息.因此,如何在 VANET 上高效地收集移动的车辆节点的实时数据成为车用自组织网络上的一个核心问题.

车用自组织网络场景如图 1 所示.随着道路上车辆节点不断移动,车辆节点之间以及车辆节点与路边接入点 AP 之间的无线通信链路不断变化.在 1 时刻,节点 a 与节点 b ,节点 b 与节点 d 以及节点 d 与节点 c 之间都存在无线链路,这时,节点 a 向节点 b 传数据,节点 c 向节点 d 传数据,当移动节点采用 802.11p 协议进行多信道无线通信时,在该时刻节点 a 和节点 c 通过控制信道协调分别采用信道 2 和信道 1 传输数据,即实现了该时刻两对节点对的无干扰数据传输.在时刻 2,随着车辆节点的移动,

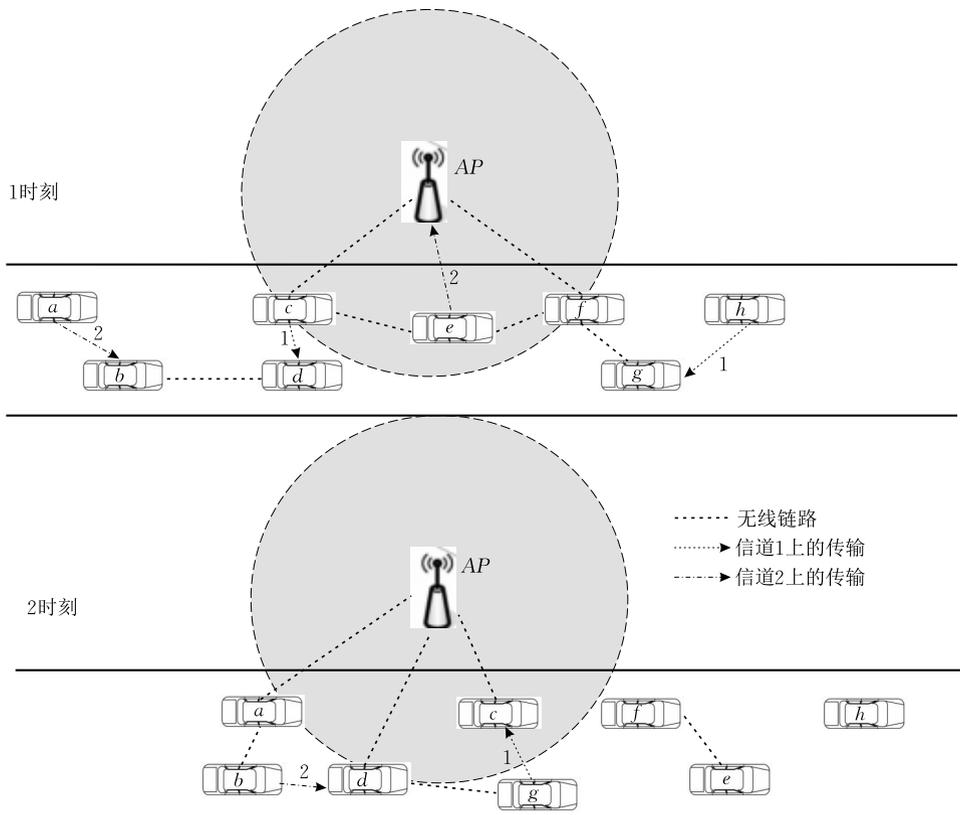


图 1 VANET 上的数据收集场景图

节点之间的通信拓扑结构也随之发生变化, 这时节点 b 和节点 d , 节点 d 和节点 g , 节点 g 和节点 c 之间存在无线链路, 此时, 节点 b 将上一时刻收到的节点 a 的数据与自身的数据进行聚集, 得到聚集数据, 将该聚集数据通过信道 2 传给节点 d , 节点 g 用信道 1 将数据传给节点 c . 从上述场景我们可知 VANET 上数据采集和处理的过程包括以下 3 个部分: 首先移动节点感知获得实时的速度、位置等信息, 然后移动节点之间通过多信道无线通信进行数据路由和聚集到 AP 节点, 最后在 AP 节点再进行解压缩、数据近似等处理分析. 由于移动节点产生的实时数据的信息量是非常大的, 在数据的路由过程中传输这些数据需要消耗大量的无线宽带资源, 同时移动节点之间的可用无线宽带资源是非常有限的, 这就出现了大量的数据传输与有限的无线通信资源之间的矛盾. 因此我们要合理地规划数据路由和聚集的路径和时刻.

目前在 VANET 上有一些关于数据聚集的研究^[6-27]. 这些研究主要是通过静态参数控制单节点上的聚集操作和数据传输^[14-16]. 但这些方法没有考虑到动态链路的细节信息. 数据路由聚集到聚集节点需要一定的时间, 同时由于节点的移动, 节点与 AP 之间的无线链路也存在时效性, 因此节点在与 AP 节点的链路失效之前要将收到的所有的可聚集数据转发给 AP 节点, 这样才能使得 AP 节点收集到移动节点的数据, 同时有效地降低网络传输的带宽资源消耗. 现有的研究没有考虑利用 802.11p 协议的多服务信道的特征进行传输规划. 在数据传输的过程中, 节点之间的通信采用无线通信方式. 同一信道上的无线通信会有无线干扰的问题. 无线干扰会导致传输效率下降. 802.11p 协议规定 VANET 上节点间的无线通信有 7 个信道: 1 个控制信道、6 个服务信道. 当在干扰范围内的多个节点对进行数据传输时, 他们可以通过控制信道进行协调使得不同的节点对利用不同的服务信道, 从而达到同一时刻干扰半径内的多节点对之间进行无干扰数据通信. 综上, 我们研究基于动态网络拓扑信息利用多信道规划移动节点的数据路由和数据聚集传输时刻的问题.

针对该问题, 本文分两个步骤进行求解. 首先确定数据路由结构. 本文提出了构造 k 约束优化聚集路由树的 CreateMRT 算法. 然后规划无冲突聚集传输时刻. 本文提出了基于动态规划思想的 DAS-MRT

算法. 本文将上述两个算法组合形成 DAC-WMMN 算法, 用以求解无线多信道移动感知网络下的数据聚集传输规划问题. 最后, 我们通过真实数据集和人工数据集上的模拟实验, 分析了该算法的数据聚集平均时延、收集数据量等性能. 实验结果表明本文的算法比现有的算法时延更小、效率更高.

本文第 2 节介绍现有的 VANET 上的数据聚集的研究工作; 第 3 节描述本文采用的系统模型; 第 4 节形式化了 k 约束优化聚集路由树的构造问题, 并提出构造 k 约束优化聚集路由树的算法, 同时分析该算法的时间复杂度; 第 5 节给出树上的数据聚集传输规划问题的形式化定义, 提出基于动态规划思想的数据聚集传输规划算法, 并证明该算法的优化性, 同时给出算法的时间复杂度的分析; 第 6 节通过大量真实轨迹数据集和模拟轨迹数据集上的实验分析本文算法在收集数据的数据量和时延上的性能, 并与现有的聚集算法进行比较. 最后总结全文.

2 相关工作

近年来, 国内外对 VANET 上数据聚集已经进行了一些研究^[6-27]. 这些研究主要集中在 3 个方面: (1) 考虑聚集路由的结构; (2) 考虑聚集数据的精度; (3) 考虑数据聚集的延迟.

文献[6-11]是研究数据聚集中的路由结构. SOTIS^[6]和 TrafficView^[7]采用无结构广播方式进行单节点上的数据聚集和分发. 采用无结构数据聚集的优点是没有额外的维护路由结构的开销, 缺点是同一数据的数据会出现多个版本, 如何从收集到的多个版本中选择最优的是核心问题. CGP^[8]采用的算法的基本思想是分簇. 它根据地理划分的路段对路面上的车进行分簇, 选择簇头, 将数据收集到簇头再进行聚集. 簇方法的优点是会产生多版本的不完整数据, 同时节省了节点的计算资源, 缺点是需要数据路由前建立分簇, 但节点快速移动时, 维护簇结构的代价较高. 文献[9-11]采用基于层次结构进行聚集. 基于层次结构的聚集是先基于地图建立多层次聚集区域, 然后逐层进行聚集. 文献[11]的思想是利用分层界标对车辆通行时间进行聚集. 层次性聚集的优点是聚集过程中传输的数据量是可控的, 传输数据量随层级呈指数级下降, 缺点是聚集引入的误差是不可控的.

现有的一些研究^[12-16]是考虑聚集数据精度, 他们

解决的问题主要是权衡聚集数据精度损失和传输数据需要的网络带宽的消耗. TrafficView^[7]考虑单节点上数据聚集算法,使得聚集数据的精度满足给定阈值的前提下聚集出最少的聚集数据. CASCADE^[12]算法进行无损数据聚集,主要思想是存储聚集数据和原始数据与聚集数据的方差.文献[13]对 CASCADE 算法进行改进,主要是对簇的大小进行了优化.文献[14]基于模糊推理进行数据聚集,它研究的是在高速公路场景下对数据进行聚集,最后得到的聚集数据能反映车流速度的实际分布情况.这种聚集方法考虑数据真实分布,但聚集后的数据量不容易控制.

还有一些研究是考虑数据聚集延迟. Catch-up^[17-18]是控制数据路由的算法.该文中采用马尔可夫模型对节点上数据的延迟进行优化,但没有考虑无线通信时出现冲突时如何调度节点进行传输的问题和移动节点之间无线链接的时效性的问题.文献[19]提出了一个查询响应协议 DB-VDG,该协议设计的主要思想是数据如果能在限定的延迟内被携带到目的地,就尽量少进行转发,这样可以最大限度地节省带宽.

还有一些其他的有关 VANET 数据聚集的研究. Dietzel 等人^[22]给出了 VANET 上数据聚集模型,对不同类型的数据不同的聚集操作等建立统一的聚集过程模型.文献[23-24]研究可扩展的数据聚集模式需要满足的理论性能.文献[25]的研究是针对停车位这类离散信息的应用,他利用的方法只适用于离散性数据的收集,不能用于收集速度、位置等连续性数据.

3 系统模型与问题描述

3.1 动态网络模型

我们考虑基于时间片的时钟同步系统.首先我们对车辆节点的移动进行建模.假设车辆节点 i 在时刻 j 处于 $(x_i^{(j)}, y_i^{(j)})$ 处,随着车辆节点的移动,车辆 i 在各个时间槽所处位置序对形成车辆 i 的移动轨迹 $Trad_i = ((x_i^{(0)}, y_i^{(0)}), (x_i^{(1)}, y_i^{(1)}), \dots, (x_i^{(D)}, y_i^{(D)}))$. 对于节点之间的无线通信我们采用圆盘通信模型.车辆 i 与车辆 j 在 k 时刻可以进行通信当且仅当在 k 时刻节点 i 和节点 j 之间的欧氏距离小于无线通信半径 R .假设已知车辆的移动轨迹,根据移动轨迹集以及圆盘通信模型,可以得到 D 个时刻的通信拓扑图序列.然后我们利用通信拓扑图序列构造移动通信拓扑.移动通信拓扑涵盖了 D 个时刻的节点之

间的连通关系.下面给出移动通信拓扑的定义.

定义 1(移动通信拓扑). 移动通信拓扑定义为 1 个三元组 $G=(V, E, LT)$, 其中 V 为节点集合, E 为链路集合. $\langle i, j \rangle$ (其中 $i, j \in V$) 是 E 中的一条链路当且仅当存在一个时刻, i 和 j 的距离小于通信半径 R . 移动造成的动态拓扑,其核心是使得链接 $\langle i, j \rangle$ 可以出现在多个时刻上, $LT(\langle i, j \rangle)$ 表示了链接 $\langle i, j \rangle$ 出现的时刻集合.

图 2 所示的是 3 个时刻的通信拓扑图序列,其中节点是 $\{a, b, c, d, e\}$, 虚线表示在该时刻节点之间存在无线链路.通过该图序列,我们可以构造移动通信拓扑 G , 如图 3 所示.该移动通信拓扑 G 共有 5 个节点,虚线为无线链路边,其中边上标识的集合为该链路的连通时刻集,例如,节点 a 和节点 b 之间的链路 l_{ab} , 该链路在 1, 2 这两个时刻是连通的,其他时刻不连通.

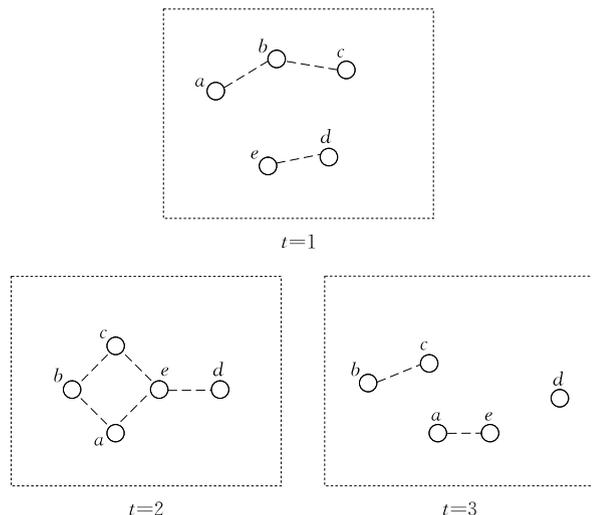


图 2 3 个时刻的通信拓扑图序列

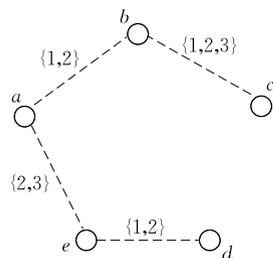


图 3 移动通信拓扑 G 示例图

3.2 树状路由模型

在给定的移动通信拓扑上移动节点进行聚集传输数据时,我们采用以 AP 节点为根的树型结构进行数据路由和聚集,这样使得每个节点到 AP 的路径唯一,即可以保证每个节点的聚集数据单版本.下面我们给出移动聚集路由结构的定义.

定义 2(移动聚集路由结构). 移动聚集路由是以 AP 节点为根节点, 所有移动节点到 AP 节点的唯一路由路径构成的树型结构, 表示为 $T = (V, E^T, LT)$, 其中 V 为节点集, E^T 为树 T 的边集, $LT(\langle i, j \rangle)$ 表示了链接 $\langle i, j \rangle$ 出现的时刻集合.

本文考虑多信道无线通信模式. 这里假设可用信道的个数为 k . 节点可以选择任一信道进行数据传输, 但一个节点在一个时刻只能发送数据或接收数据, 当接收数据时, 只能接收一个节点发送的数据. 当多对节点在同一信道中进行数据传输时, 本文考虑一跳冲突模型, 即一个节点在接收数据时在它的通信半径内不能有其他的节点对进行通信.

3.3 传输规划及数据聚集

在数据聚集传输的过程中, 节点发送数据的时刻决定传输是否存在冲突以及 AP 节点最终能收集到的聚集信息量. 因此规划数据聚集传输时, 即规划节点发送数据的时刻. 每个节点 i 被规划的聚集传输时刻记为 W_i . 因此一个聚集传输规划 W 是一个从移动节点集到时刻集的映射. 下面我们给出聚集传输规划的定义.

定义 3(聚集传输规划). 一个聚集传输规划是映射 $W: V/\{AP\} \rightarrow LT$, 其中 $LT = \bigcup \{LT_e | e \in E^T\}$, E^T 为移动聚集路由结构 T 的边集. 节点 i 的 W_i 表示通过聚集传输规划 W 得到的发送数据的时刻, 当 W_i 为 0 时, 表示规划节点 i 不发送数据.

在发送数据之前节点要执行数据聚集操作. 我们假设这里的聚集操作为可分函数. 可分函数指的是可以分治计算的函数, 例如 MIN, MAX, SUM 等聚集操作. 这些操作也是数据聚集中常常出现的操作. 同时假设在节点上进行聚集计算的时间是可以忽略的, 即不占用单独的时间片.

在初始时刻, 每个移动节点产生一个感知数据, 即为原始数据. 一个原始数据的信息量为 1. 然后节点在传输数据的过程中通过聚集操作得到聚集数据. 一个聚集数据包含的信息量是参与该聚集数据计算的原始数据的个数, 即该数据包含的移动节点的信息的个数. 一个节点上累积的聚集信息量为该节点感知数据的信息量和该节点接收到的所有聚集数据的信息量之和. 我们下面给出节点上的累积聚集信息量的定义.

定义 4(节点上的累积聚集信息量). 该值是节点 i 感知数据的信息量和节点 i 截止到 t 时刻接收的聚集数据的聚集信息量之和, 记为 $M(i, t)$. $M(i, t)$ 定义如下:

$$M(i, t) = \begin{cases} 1, & i \in V^L \\ 1 + \sum_{j \in C_i} W_j M(j, W_j), & i \in V/V^L \end{cases} \quad (1)$$

其中: V^L 为移动聚集路由结构 T 中的叶节点集合; C_i 为节点 i 的儿子节点集.

4 k 约束优化聚集路由结构及其构造算法

4.1 k 约束优化聚集路由结构

构造移动聚集路由结构即确定了数据路由的路径, 也就是确定了节点发送的数据的接收节点. 根据移动聚集路由结构将冲突情况分为两种类型: 一种是节点不能同时接收多个数据, 即在移动聚集路由结构上同一父节点的多个儿子节点不能同一时刻向父节点发送数据, 同时父节点发送数据的时刻不能和儿子节点发送数据的时刻相同, 这种冲突被称为树内冲突; 另一种是节点在接收数据时不能受到其他发送节点的影响, 即接收节点不能在其他发送节点的通信范围内, 这种冲突被称为树间冲突. 图 4(a) 所示的情况即是树内冲突, 如图所示节点 u 和节点 v 有共同的父节点 x , 如果节点 u 和节点 v 同时给节点 x 发送数据, 那么在节点 x 处会发生冲突, 节点 x 无法同时接受节点 u 和节点 v 的数据. 图 4(b) 所示的是树间冲突的情况. 节点 u 给父节点 x 发送数据, 节点 v 给父节点 y 发送数据, 但节点 x 在节点 v 的通信半径内, 如果节点 u 和节点 v 同时发送数据, 那么在节点 x 处会受到节点 v 传输数据的干扰, 无法接收到节点 u 的数据.

我们考虑在建树时加入节点在树上的度的约束, 使得树间冲突可以采用不同信道进行避免, 也就是在构造路由树时, 使得每个节点的树间边的个数不超过 k , 这也就保证了当一个节点发送数据的时刻与它的除父节点和儿子节点的其他邻居节点接收数据的时刻相同时, 他们可以采用 k 个信道中的不同信道, 使得树间冲突不会存在. 因此建树问题的约束条件是树上的任一节点 i 要满足 $\Delta_i - \delta_i < k$, 其中 Δ_i 为节点 i 在移动通信拓扑 G 上的度数, δ_i 为节点 i 在移动聚集路由结构上的度数, k 为可用信道数量, 即任一节点 i 在通信半径范围内可能被规划同时传输的节点个数不超过信道个数, 则当这些节点被规划同时传输时, 则采用不同信道进行传输.

该树结构是传输时刻规划问题的输入, 因此该

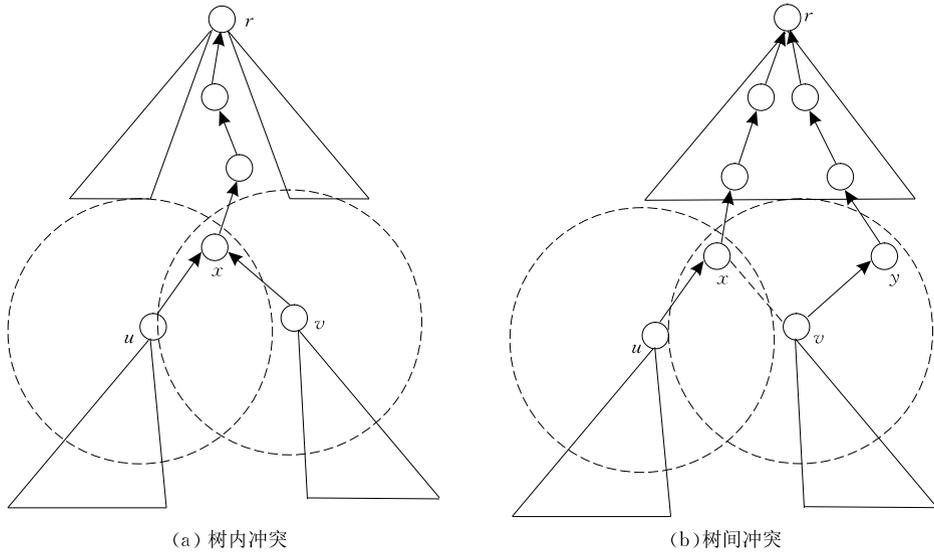


图 4 两种冲突情况

树的结构影响传输规划的性能. 建树时考虑连通时刻较多的边作为树上的边, 这样在规划传输时刻时, 可选择的空间更大, 也就是可规划的空间更大, 因此更利于最大化收集数据量这一目标的实现. 综上, 我们给出 k 约束优化聚集路由结构的定义如下.

定义 5 (k 约束优化聚集路由结构). 树 $T=(V, E^T, LT)$ 为移动通信拓扑 G 上的 k 约束优化聚集路由结构当且仅当在该树上的边的连通时刻集的基数之和最大, 且每个节点满足 $\Delta_i - \delta_i < k$ 限制, 其中 Δ_i 为节点 i 在移动通信拓扑 G 上的度数, δ_i 为节点 i 在移动聚集路由结构上的度数, k 为信道数量.

如图 5 所示, 图 5(a) 为 5 个节点的移动通信拓扑, 图 5(b) 为该拓扑的 2 约束优化路由结构, 其中 2 约束体现在对于节点 b 来说, 在路由结构中虚线连接的节点个数小于 2, 这里与节点 b 虚线相连的只有节点 e . 优化体现在对于节点 e , 它可以选择节点 a 作为父节点, 也可以选择节点 b 作为父节点, 但链路 ae 的连通时刻集为 $\{1, 2, 3\}$, 链路 be 的连通时刻集为 $\{1, 2\}$, 链路 ae 的连通时刻更多, 因此节点 e 选择节点 a 作为父节点.

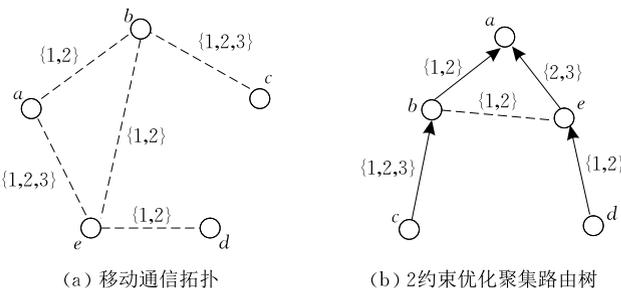


图 5 2 约束优化聚集路由结构

我们将构造 k 约束优化聚集路由结构的问题形式化为如下形式, 记为 k -MRT 问题:

k -MRT 问题:

已知: 移动通信拓扑 $G=(V, E, LT)$, 其中 $AP \in V$, 无线通信可用的信道个数为 k .

求: 以 AP 为根的 k 约束优化聚集路由结构 T .

4.2 k 约束优化聚集路由树算法

k -MRT 问题是基于移动通信拓扑求以 AP 为根节点的 k 约束优化聚集路由树. 我们给出 CreateMRT 算法. 该算法首先保证 k 约束, 即每个节点的至少 $\min\{\Delta_i - k, 0\}$ 个邻居节点会被构建为儿子节点, 这就保证了每个节点的树间边的个数不超过 k , 在选择哪些节点加入时, 根据链路的连通时刻集的大小进行选择. 然后, 当一个节点选择哪个节点作为父节点时, 我们仍考虑它与该节点的链路连通时刻集, 选择这个集合基数最大的链路的节点作为父节点. 这里 N_i 表示节点 i 在 G 上的邻居节点集. 该算法的伪码如算法 1.

算法 1. CreateMRT 算法.

输入: 移动通信拓扑 $G=(V, E, LT)$, 根节点为 AP 节点, 信道个数为 k

输出: k 约束优化聚集路由树 $T=(V', E^T, LT')$

1. 初始化: $V' = \text{NULL}$, $E^T = \text{NULL}$
2. 将 AP 加入 V' , 将 AP 加入 $tempV$
3. WHILE ($tempV$ 不为空)
4. 取节点 $i \in tempV$, 设 $tempN = N_i \setminus V'$
5. IF ($\Delta_i > k$) {
6. 从 $tempN$ 中取前 $\Delta_i - k$ 个与节点 i 的链路连通时刻集的基数最大的节点加入 $tempV$, 加入 V' ,

从 $tempN$ 中删除, 相应的与节点 i 的链路加入 E^T , 将节点 i 从 $tempV$ 中删除,

7. 将 $tempF$ 加入 V' 中, 相应的边加入 E^T 中
8. WHILE ($tempN$ 不为空)
9. IF (链路 $\langle i, j \rangle$ 的连通时刻集的大小是与节点 j 相连的所有链路中最大的) {
10. 将节点 j 加入 V' , 从 $tempN$ 中删除, 加入 $tempV$ 中, 将链路 $\langle i, j \rangle$ 加入 E^T }

首先初始化树的节点集和边集为空. 将 AP 节点加入节点集, 作为根节点. 然后对于每个新加入的节点 i , 判断该节点在图 G 上的度 Δ_i 与信道个数 k 的关系, 如果大于, 则要选择 $\Delta_i - k$ 个节点添加为节点 i 的儿子节点, 这时要将节点 i 的未加入的邻居节点与节点 i 的链路的连通时刻集进行排序, 从中选择最大的前 $\Delta_i - k$ 个, 将这些节点加入新加节点集 $tempV$, 加入树节点集 V' , 相应的边加入边集. 接下来我们要考虑节点 i 的未加入的邻居节点集中仍没处理的节点, 设其中的一个节点 j , 判断它与节点 i 的链路连通时刻集的大小是否为它的所有相连链路中最大的, 如果是, 节点 j 选择节点 i 作为父节点, 相应地做些加入和删除操作. 循环上述过程, 直至新加节点集为空, 算法终止.

4.3 算法时间复杂度分析

设移动通信拓扑 G 的节点个数为 n , 边数为 m , 节点在 G 上的最大度为 Δ .

定理 1. CreateMRT 算法的时间复杂度为 $O(n\Delta\log\Delta + \Delta^2)$.

证明. 该算法的主体部分为步 3~10. 对于每个新加入的节点, 要对该节点的所有未添加的邻居节点进行排序, 时间复杂度为 $O(\Delta\log\Delta)$, 其他的加入删除等操作时间复杂度为 $O(1)$, 因此对于所有节点进行上述操作的时间复杂度为 $O(n\Delta\log\Delta)$. 对于步 9~10, 要找出一个节点的所有链路的连通时刻集基数最大的, 时间复杂度为 $O(\Delta)$, 对每个节点来说, 做这个操作的次数不超过 Δ , 因此这部分总时间复杂度为 $O(\Delta^2)$.

综上, CreateMRT 算法的时间复杂度为 $O(n\Delta\log\Delta + \Delta^2)$. 证毕.

5 移动聚集路由结构上的聚集传输规划问题及其算法

5.1 移动聚集路由结构上的聚集传输规划问题

k 约束优化聚集路由由树确定了数据的路由, 下面就是在该结构上规划数据聚集传输时刻. 已知 k

约束优化聚集路由结构 T , 在 T_0 时刻移动节点 i 产生感知数据 d_i , 求解无冲突的聚集传输规划 W 使得 AP 节点收集的聚集信息量 $M(AP, W_{AP})$ 最多, 其中 $W_{AP} = 1 + \operatorname{argmax}\{t | t \in LT_{i, AP}, i \in C_{AP}\}$. 我们将移动聚集路由结构上的数据聚集传输规划问题进行形式化定义, 记为 DAS-MRT 问题.

DAS-MRT 问题:

Maximize $M(AP, W_{AP})$

Subject to $W_i \in LT_{i, p_i}, \forall i \in V$

$$W_i < W_{p_i}, \forall i \in V, \text{ if } W_i, W_{p_i} \neq 0 \quad (2)$$

$$W_i \neq W_j, \forall i, j \in C_k, k \in V/V^L, \text{ if}$$

$$W_i, W_{p_i} \neq 0 \quad (3)$$

其中式(2)约束表示的是当儿子节点和父节点都传输数据时, 儿子节点传输数据的时刻要小于父节点传输数据的时刻, 即儿子节点要在父节点发送数据之前将数据发送给父节点, 式(3)约束表示的是同一父节点的多个儿子节点都传输数据时, 它们传输数据的时刻不能相同.

5.2 移动聚集路由由结构上的传输规划算法

通过对 DAS-MRT 问题的分析, 我们给出了基于动态规划思想的 DAS-MRT 算法. 我们首先进行分层, 将叶子节点作为第 0 层, 父节点的层数为儿子节点的层数的最大值加 1. 按照层数由低到高的顺序, 对每个节点与父节点的链路的连通时刻计算该时刻传输可以收集的最大数据量. 设 $X[i, W]$ 表示节点 i 在传输时刻为 W 时发送的聚集数据包的最大信息量. 设节点 i 的父节点为 p_i . 设节点 i 的儿子节点集为 C_i , 有 k 个儿子节点 c_1, c_2, \dots, c_k , 对应传输时刻为 $W_{c_1}, W_{c_2}, \dots, W_{c_k}$. 该算法的伪码如算法 2.

算法 2. DAS-MRT 算法.

输入: k 约束优化聚集路由结构 $T = (V', E^T, LT)$

输出: 每个节点的传输时刻 W_i

1. 初始化 $X[i, W] = 1, \forall i \in V^L, \forall t \in LT_{i, p_i}$
2. 初始化 $X[i, 0] = 0$, IF $1 \in LT_{i, p_i}$, 并且 $X[i, 1] = 1$
3. WHILE ($i \in V \setminus (V^L \cup \{AP\})$)
4. WHILE ($w \in LT_{i, p_i}$)
5. 设临时变量 $NC = C_i, m = 1$
6. WHILE ($(w - m) > 1$)
7. IF $|\{c_j | (w - m) \in LT_{c_j, i}\}| = 0$, 则 m 加 1
8. IF $|\{c_j | (w - m) \in LT_{c_j, i}\}| = 1$
9. 则 $W_{c_j} = w - m, NC = NC \setminus \{c_j\}$
10. IF $|\{c_j | (w - m) \in LT_{c_j, i}\}| > 1$, 跳出
11. 从 $LT_{c_j, i} \cap \{1, \dots, W - m\}$ 中选择最大的前 $|NC|$ 个, 如果不足则全选

12. 用 NC 中的节点和可选时刻集构造最大权值匹配问题,调用 KM 算法
13. 设 $W_{AP} = 1 + \arg \max \{t | t \in LT_{i,AP}, i \in C_{AP}\}$, 计算 $X[AP, W_{AP}]$

在该算法中,前两步是初始化的步骤,从步 3~12 是计算除叶节点和根节点的所有其他节点在所有可传输时刻的 $X[i, W]$,步 13 是计算根节点 AP 节点的 $X[AP, W_{AP}]$. 这里从步 3~12 是 DAS-MRT 算法的核心. 步 7~10 是一个优化处理,即将没有冲突的时刻直接分配给节点,步 11~12 是确定节点的候选时刻集,用候选时刻集和儿子节点集构造二部图的最大权值匹配问题,然后利用 KM 算法^[28]进行求解.

5.3 分析算法的优化性

首先我们很容易可以得到以下结论:对于任一节点 $i, t \in LT_{i,p_i}, M(i, t)$ 是递增的. 这是由于在节点 i 处数据不会被丢弃,因此,累积的数据量一定是随着时间的增加递增的. 后面的证明中我们将用到这一结论. 为证明 DAS-MRT 算法的优化性,我们首先证明该问题有最优子结构,然后证明 DAS-MRT 算法中选择的候选时刻集一定包含最优解在其中,最后我们给出 DAS-MRT 算法能得到最优解. 定理 2 即证明 DAS-MRT 问题存在优化子结构.

定理 2. 设 AP 节点有 k 个儿子节点 c_1, c_2, \dots, c_k , 设这 k 个儿子节点相应的优化传输时刻为 $W_{c_1}^*, W_{c_2}^*, \dots, W_{c_k}^*, M_{c_i}^*$ 为节点 c_i 在 $W_{c_i}^*$ 传输的数据包含最多的节点个数. $M(AP, W_{AP})$ 达到最大值当且仅当任一儿子节点 c_i 的数据传输时刻为 $W_{c_i}^*$, 数据包含的信息量为 $M(c_i, W_{c_i}^*)$.

证明. 这里需要证明两个部分:一部分是任一儿子节点 c_i 的数据传输时刻为 $W_{c_i}^*$; 另一部分是传输的数据包含的信息量为 $M(c_i, W_{c_i}^*)$.

首先,我们假设第一部分成立,证明第二部分. 假设 c_i 节点传输的数据包含的信息量为 M_{c_i} , 则 $M_{c_i} < M_{c_i}^*$, 这时以 c_i 为根的子树的节点数据传输时刻规划为 s_1 . 当 c_i 节点传输的数据包含的信息量为 $M_{c_i}^*$ 时,以 c_i 为根的子树的节点数据传输时刻规划为 s_2 .

根据式(1)可知, $M(AP, W_{AP}) = \sum_{c_i \in C_{AP}} M(c_i, W_{c_i})$, 用 s_2 映射可以使得 c_i 节点传输的数据包含信息量为 $M_{c_i}^*$, 也就使得 $M_{AP} > M_{AP}^*$, 这与 M_{AP} 是最优解矛盾, 因此假设不成立.

下面我们证明第一部分. 假设 k 个儿子节点的

传输时刻为 $W_{c_1}, W_{c_2}, \dots, W_{c_k}$, 根据式(1)计算根节点 AP 的 $M(AP, W_{AP}) = \sum_{c_i \in C_{AP}} M(c_i, W_{c_i})$. 对于 k 个儿子节点的传输规划为 $W_{c_1}^*, W_{c_2}^*, \dots, W_{c_k}^*$, 根据式(1)计算根节点 AP 的 $M'(AP, W_{AP}) = \sum_{c_i \in C_{AP}} M(c_i, W_{c_i}^*)$. 根据最大权匹配问题可知 $M'(AP, W_{AP}) > M(AP, W_{AP})$, 因此与 $M(AP, W_{AP})$ 是最优解矛盾, 因此假设不成立.

综上所述可知 DAS-MRT 问题有优化子结构.

证毕.

下面我们给出算法在选择传输时刻的候选集时必定包含了优化解在内,即在优化过程中构造的候选集是正确的.

定理 3. 考虑节点 i , 它的儿子节点为 k 个, c_1, c_2, \dots, c_k , 如果节点 i 的传输时刻为 W , 对于任一儿子节点 c_i 取 $LT_{c_i,i} \cap \{1, 2, \dots, W-1\}$, 如果 $|LT_{c_i,i} \cap \{1, 2, \dots, W-1\}| \geq k$, 则将集合 $LT_{c_i,i} \cap \{1, 2, \dots, W-1\}$ 中元素进行从大到小进行排序, 选取前 k 个元素构成的集合 RT_{c_i} , 否则设 $RT_{c_i} = LT_{c_i,i} \cap \{1, 2, \dots, W-1\} \cup \{0\}$, 儿子节点的传输规划候选时刻集为 $RT_{c_1}, RT_{c_2}, \dots, RT_{c_k}$. 如果所有儿子节点满足 $|LT_{c_i,i} \cap \{1, 2, \dots, W-1\}| \geq k$, 则 k 个儿子节点都能分配到传输时刻, 儿子节点的传输规划候选时刻集为 $RT_{c_1}, RT_{c_2}, \dots, RT_{c_k}$.

证明. 我们根据 $LT_{c_i,i} \cap \{1, 2, \dots, W-1\}$ 的大小进行分情况讨论.

情况 1. 所有儿子节点的 $|RT|$ 为 k , 且不含 0 时刻. 假设存在儿子节点 c_i , 当它的父节点 i 的传输时刻为 W , 它的优化传输时刻为 $W_{c_i}^*$ 不属于 $RT_{c_i,i}$, 则 $W_{c_i}^*$ 小于 RT_{c_i} 中的任意元素. 由于所有儿子节点的 $|RT|$ 为 k , 且不含 0 元素, 则 RT_{c_i} 中一定存在一个时刻是被空出来的, 即没有节点在该时刻传输数据. 由于 $W_{c_i}^*$ 小于 RT_{c_i} 中的任意元素, 根据 $M(i, t)$ 函数的单调递增性, 截止到该时刻节点 c_i 收集的数据量一定大于等于截止 $W_{c_i}^*$ 时刻节点 c_i 收集的数据量. 因此, 儿子节点 c_i 在该时刻传输的收益优于在 $W_{c_i}^*$ 时刻的传输数据, 出现矛盾, 因此上述假设不成立.

情况 2. 存在儿子节点 c_i 的 $|RT| \leq k$, 且含 0 时刻. 在这种情况下所有的儿子节点分成了两类: A 类是 RT 中不含 0 时刻; B 类是 RT 中含有 0 时刻. A 类中的节点的优化传输时刻并在 RT 中, 证明过程与情况 1 中的证明过程类似. 对于 B 类中的节

点 i , $|RT| \leq k$, 因此当 RT 中的除 0 以外的元素与其他节点的 RT 交集不为空, 且其他节点的优化时刻选取在 RT 中的所有除 0 时刻时, 节点 i 没有可规划的空闲时刻, 因此, 节点 i 不会被传输, 即传输时刻为 0. 综上, B 类中节点的优化传输时刻在 RT 中. 证毕.

下面我们给出 DAS-MRT 算法得到的解为 DAS-MRT 问题最优解的证明.

定理 4. DAS-MRT 算法能得到 DAS-MRT 问题的最优解.

证明. 根据定理 3, 我们已经找到了节点 i 的儿子节点 c_1, c_2, \dots, c_k 的优化传输时刻的候选集, 现在我们只需要从这些候选集中分配传输时刻使得在不冲突的前提下 $M(i, W_i)$ 最大. 这里 $X[i, W]$ 表示 i 的传输时刻为 W 时发送数据包含的节点个数的最大值. 求解最优的 $M(i, W_i)$ 是一个二部图上的最大权匹配问题. 我们构造带权二部图 $G = (V_1 \cup V_2, E)$. 其中 V_1 由 k 个儿子节点 c_1, c_2, \dots, c_k 组成, V_2 由每个儿子节点的传输规划候选时刻集组成, 这里要去掉 0 时刻, 即 $URT_j \setminus \{0\}$. 将节点 c_j 与所有 RT_{c_j} 中的节点 t_j 进行连接, 边上的权值为 $X[c_i, t_j]$. 限制条件是多个节点不能有相同的传输时刻且每个节点最多有一个传输时刻. 目标是给儿子节点们分配传输时刻使得被分配的边上的权值之和最大. 这是一个最大权匹配问题, 在一步冲突下可以用 KM 算法在多项式时间得到最优解.

MRT 的高度为 h , 我们根据节点距离 AP 节点的跳数对节点进行分层, AP 节点作为第 0 层, 叶子节点作为第 h 层. 我们从第 $h-1$ 层的节点开始, 给定任一节点的传输时刻, 能够通过 DAS-MRT 算法确定儿子节点的优化传输时刻, 将该节点的所有可传输时刻都进行计算, 将该层所有节点计算完后, 计算 $h-2$ 层节点, 递归下去, 即可以计算 AP 节点在 W_{AP} 时刻的 X 值. 因此, DAS-MRT 算法得到了 DAS-MRT 问题的最优解. 证毕.

5.4 分析算法的时间复杂度

在这一部分我们分析 DAS-MRT 算法的时间复杂度. 设移动聚集路由结构 T 的高度为 h , 树上节点的最大度为 δ , 节点总数为 n , 边的连通时刻集的基数最大为 m . 二分图最优匹配问题的 KM 算法^[34]的复杂度为 $O(q^3)$, 其中 q 为二分图的边数.

定理 5. DAS-MRT 算法的时间复杂度为 $O(m + (h-1)m\delta^6)$.

证明. 在每个节点 i , $\forall W \in LT_{i, p_i}$, 我们需要计算 $X[i, W]$.

对于叶子节点来说, $\forall i \in V^L, \forall W \in LT_{i, p_i}$, $X[i, W]$ 均设为 1, 因此计算需要的时间为 $O(m)$.

设第 k 层的节点集为 $V^{(k)}$. 对于第 k 层的节点来说, 对于 $\forall i \in V^{(k)}, \forall W \in LT_{i, p_i}$, 计算 $X[i, W]$ 需要的时间取决于 KM 算法的时间复杂度, 这里二部图的边数为 $O(\delta^2)$, 由于 KM 算法的时间复杂度为 $O(q^3)$, q 为边数, 因此计算 $X[i, W]$ 的时间复杂度为 $O(\delta^6)$, 从而计算第 k 层一个节点的所有可传输时刻的 X 值的时间复杂度为 $O(m\delta^6)$. 由于同一层的节点的 X 值相互不影响, 可以并行计算, 因此第 k 层的节点的 X 值计算的时间复杂度为 $O(m\delta^6)$.

第 k 层节点的 X 值计算取决于第 $k+1$ 层节点的 X 值. 第 h 层为叶子节点, 从 $h-1$ 层开始逐层计算, 计算到第 0 层的根节点的时间复杂度为 $O(m + (h-1)m\delta^6)$.

因此, DAS-MRT 算法的时间复杂度为 $O(m + (h-1)m\delta^6)$.

6 实验及其结果分析

我们将 CreateMRT 算法和 DAS-MRT 算法合并形成 DAC-WMMN 算法. 通过模拟实验来评估 DAC-WMMN 算法的性能, 将它与 CatchUp^[17] 算法和 DB-VDG^[19] 算法进行比较. 需要说明的是: CatchUp 中需要设定的两个参数, WALK 速度为 100 m/s, RUN 速度为 1000 m/s. DB-VDG 中有两种模式: SBSS 和 DBSS, 实验表明 DBSS 的性能更好, 因此我们选择 DBSS 模式. 为便于比较, 我们将两点之间的距离定义为最短路径的边长之和.

6.1 实验数据以及参数设置

我们采用两个数据集进行模拟实验: 一个是人工轨迹数据集 SMD; 另一个是真实的出租车移动轨迹数据集 CSGPS.

对于人工数据集, 我们采用随机方法生成. 每个时刻的图为 Erdos Renyi 随机图. 给定节点的总数和任意两个节点之间边的生成概率, 随机生成静态图, 再在边上随机连通时刻集, 即形成一个移动通信拓扑. 人工数据集参数见表 1.

表 1 人工数据集默认参数

参数	缺省值
Node number	50
Rounds	20
Delay Constraint	10
Connected probability	0.2
Cardinal Number of Connected Time Set	7

CSGPS 数据集^①是由 Micheal Piorkowski, Natasa Sarafijanovic-Djukic, Matthias Grossglauser^[29] 在 2008 年于美国旧金山采集的出租车 GPS 移动轨迹数据. 该数据集的数据是由 Cabspotting^② 项目组提供的. 该数据集包含了在旧金山湾区的大约 500 辆出租车 30 天的 GPS 数据. 数据采集的开始时间为 2008-05-17, 结束时间为 2008-06-10. 数据集中每辆车的轨迹数据存储为以该车辆 ID 为名字的 txt 文件, 该文件中每一条记录为一个时刻该车辆的 GPS 数据.

我们对数据进行预处理, 首先整个数据采集时间段进行分片, 将每个时间片内出现的 GPS 数据整理为一个文件, 形成单时间片 GPS 数据文件, 然后根据给定的通信半径以及圆盘模型, 形成单时间片通信拓扑图文件. 数据集的具体参数见表 2.

表 2 CSGPS 数据集默认参数

参数	缺省值
Node number	[40, 50]
Rounds	20
Communication Range	500 m
MAC layer	802.11p
Cardinal Number of TS	10
Delay Constraint	10
Time Frame	100 s

我们采用 NS2 模拟器模拟无线网络传输, 这里 MAC 层采用 802.11p 协议. 我们将处理后的 GPS 轨迹导入 NS2, 在 NS2 上运行规划算法, 按照得到的结果进行实际数据传输.

6.2 性能指标

为了分析数据收集算法的性能, 我们定义 3 个指标评测数据收集的有效性、效率和时效性.

(1) 数据收集率是度量数据收集算法的有效性的指标, 也就是数据收集算法将相关的节点数据路由到 AP 的能力. 该度量是 AP 节点收集到的节点数据占节点总数的比例, 定义如下:

$$Effectiveness = \frac{N_{delivered}}{N_{nodes}} \% \quad (4)$$

这里 $N_{delivered}$ 是 AP 收集到的节点的数据总量, N_{nodes} 是节点总数.

(2) 平均传输数据率是评测传输数据时消耗的带宽量的指标, 即度量算法的效率. 该度量是节点发送的数据总量与 AP 收集到的节点的数据总量的比值, 定义如下:

$$Efficiency = \frac{\sum_{i \in delivered} |S_i|}{N_{delivered}} \quad (5)$$

这里 S_i 为节点 i 发送的数据量, $delivered$ 是 AP 节点收集到的数据对应的节点集, $N_{delivered}$ 是 AP 收集到的节点的数据总量.

(3) 平均延迟是度量数据收集算法的时间性能指标. 该度量是收集上来的数据的延迟的平均值, 定义如下:

$$AverageDelay = \sum_{i \in received} t_i / N_{received} \quad (6)$$

这里 t_i 为节点 i 的数据到达 AP 节点的时刻, $received$ 是 AP 节点收集到的数据对应的节点集, $N_{received}$ 为 AP 收集到的节点的数据总量.

6.3 性能对比

为了分析 DAC-WMMN 的性能, 我们将该算法与 Catch-Up 算法和 DB-VDG 算法从数据收集的有效性、效率和时效性 3 个方面进行对比.

默认参数下性能比较见表 3. 综合 3 个性能指标可以看出 DAC-WMMN 的性能最好. 它在 3 种方法中具有最高的数据收集率和最低的平均传输数据率, 同时平均延迟也是最小的. 总体上来说, 在人工数据集上运行的结果要优于实际 GPS 轨迹数据集上的运行结果. 这是因为人工数据集上节点之间的动态连通性更有规律, 而实际数据集上节点之间的连通变化更频繁. 从表 3 中可以看出 DAC-WMMN 的数据收集率高于 Catch-Up 和 DB-VDG, 这是因为 Catch-Up 不是以延迟界限驱动的, 前方节点通过等待聚集的效果更好, 但是有可能错过延迟界, 因此数据收集率会受到影响. DB-VDG 在节点选择路由模式时考虑了延迟界限, 但只是简单地根据线性模型计算, 没有考虑未来实际的节点移动情况和通信拓扑, 而 DAC-WMMN 是根据截止到延迟界的全部通信拓扑信息优化数据收集率, 因此能够在保证最大延迟界的情况下使得尽量多的节点数据被收集上来. 从表 3 中可以看出 DAC-WMMN 的平均传输率明显高于 Catch-Up 和 DB-VDG, 这是因为 DAC-WMMN 中每个节点只发送一次数据, 因此进行了最大程度地聚集, 而在 Catch-Up 和 DB-VDG 中节点数据经过两个模式进行路由, 因此聚集的数据在经过一段时间到达同一节点时才能进行聚集, 而之前他们是分别路由的, 因此 DAC-WMMN 算法本身决定了它是消耗传输带宽量最小的方式. 对于平均延迟方面, Catch-Up 的平均延迟要略小于 DAC-

① 该数据集的更详细的信息可在 <http://crawdad.cs.dartmouth.edu/epfl/mobility/CSGPS> 上获取.

② Cabspotting Project. <http://cabspotting.org/about.html>

WMMN 和 DB-VDG,这是因为 Catch-Up 中存在一部分节点还未进行聚集或只进行部分聚集就已经将数据传到目的节点,因此这部分的节点的延迟相比较来说更小,也就使得平均延迟更小,但 DAC-

WMMN 较之相差不大,因为 DAC-WMMN 中在节点的路由路径选择时考虑数据的最小可能延迟,从而使得数据收集过程中在优化数据收集率的同时也能使得数据尽可能快地到达 AP 节点.

表 3 默认参数下仿真结果的对比

数据集 参数	SMD			CSGPS		
	有效性/%	平均延迟/100 s	效率	有效性/%	平均延迟/100 s	效率
DAC-WMMN	98.8	7.27	1.00	81.1	6.71	1.00
DB-VDG	79.6	8.12	3.25	64.7	7.13	3.96
Catch-Up	68.8	6.95	2.67	58.2	6.96	3.25

6.3.1 节点之间的连通度对性能的影响

本组实验主要研究移动通信拓扑图中节点之间的连通情况对数据收集率的影响.其他默认参数固定,节点的连通情况变化通过改变边的连通概率和边的连通时刻集的基数实现.这里我们采用人工数据集进行实验.从图 6 中可以看出,随着连通概率和连通时刻集基数的增加,DAC-WMMN、Catch-Up 和 DB-VDG 算法的数据收集率都是增加的,DAC-

WMMN 的数据收集率高于另两个算法.从图 6 中,我们可以看出数据收集率随连通时刻集基数的变化更明显.这是由于连通时刻集的基数反映边的动态变化中的持续连通情况,当边上该值较大时,相应的节点的候选发送时刻集较大,该节点以及该节点收集的数据能够传输的机会也较多,因此整体的数据收集率会较大.从图 7 中可以看出随着连通概率和连通时刻集基数的增加,DAC-WMMN 算法得到平

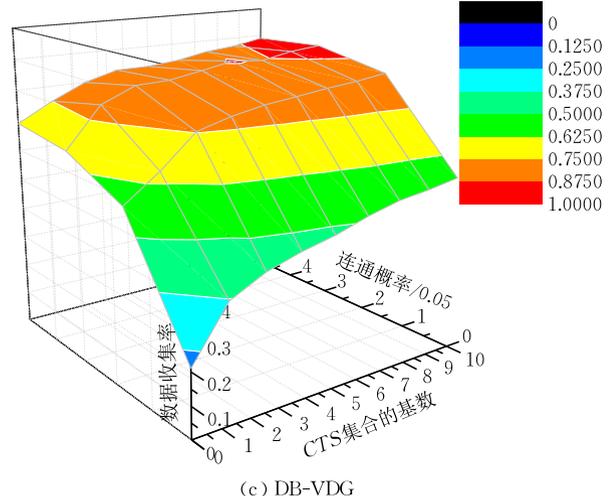
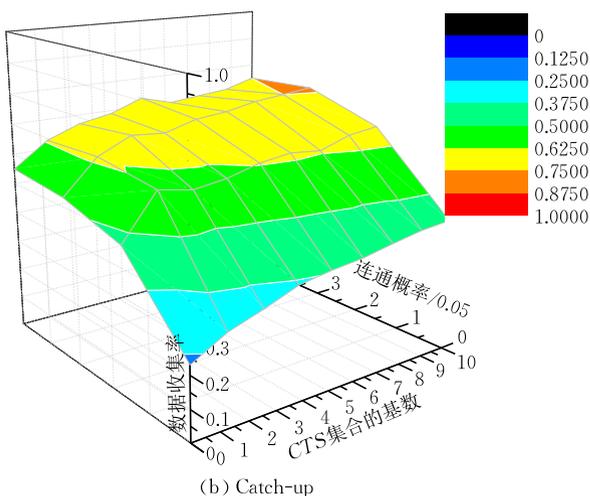
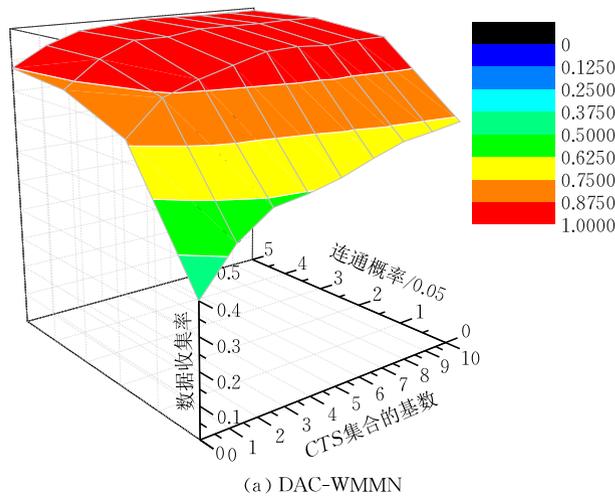
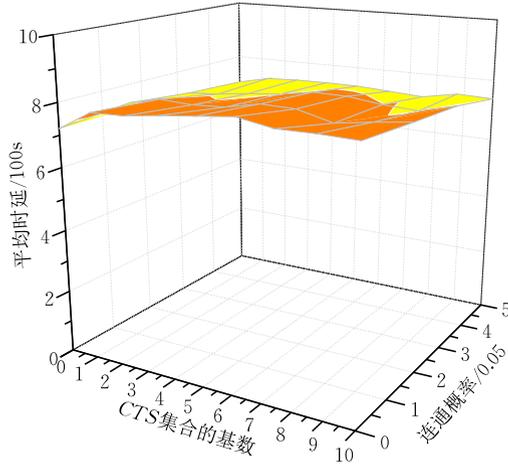
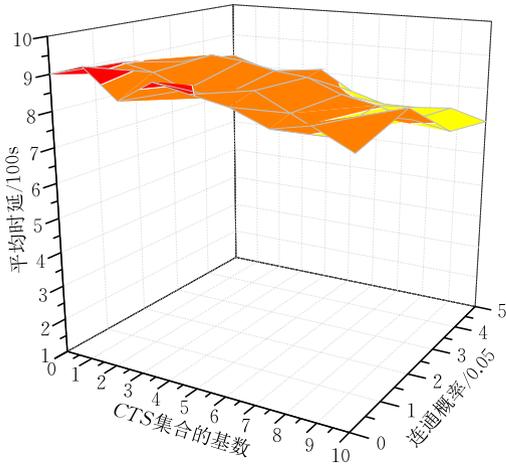


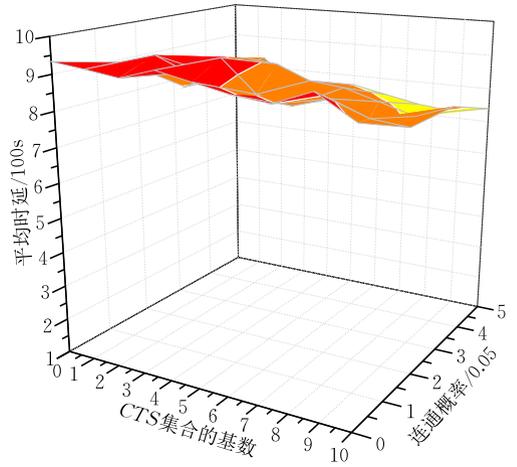
图 6 边连通概率和连通时间集基数对数据收集率的影响



(a) DAC-WMMN



(b) Catch-up



(c) DB-VDG

图 7 边连通概率和连通时间集基数对平均延迟的影响

均延迟变化不大,而 Catch-Up 和 DB-VDG 算法中平均延迟有明显的减小趋势.这是由于 DAC-WMMN 在选择节点的数据接收节点时考虑了最小可达时刻的优化,因此在限定延迟界的情况下,平均延迟不会受到通信网络动态连通情况太多的影响,保持了相对稳定.而 Catch-Up 和 DB-VDG 算法中是没有涉及接收节点的选择问题,在连通情况较好的情况下,数据的路由机会更多,流动性更大,延迟小的节点的比例会更大,因此,平均延迟会更小,因此延迟更多地会受到网络动态连通情况的影响.

6.3.2 节点数量对性能的影响

本组实验主要研究节点数量对性能的影响.对于 SMD 数据集,我们通过改变移动通信拓扑中点的个数来实现,对于 CSGPS 数据集,我们提取不同时间段的公共子序列图,对节点总数根据间隔为 5 划分移动通信拓扑类别,每个类别包含的图例均为 20 个.从图 8 中,我们可以看出对于 SMD 数据集,随着节点数量的增加,3 个算法的数据收集率都降

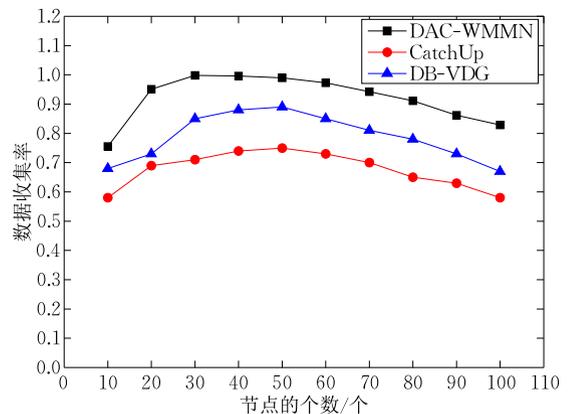


图 8 节点总数对数据收集率的影响(SMD 数据集)

低.从图 8 中可以看出,DAC-WMMN 算法的数据收集率下降较缓慢,并仍高于其他的两个算法的数据收集率. Catch-Up 算法和 DB-VDG 算法的数据收集率下降较快,这是由于当节点数量增多时,节点密度和连通情况不变,数据收集量基本不变,同时这两个算法收集的效率低于 DAC-WMMN 算法,因

此相对的数据收集率下降较快. 从图 9 中, 我们可以看出对于 CSGPS 数据集, 3 个算法的数据收集率都是随着节点个数的增加而下降. 从图中可以看出, CSGPS 数据集的实验结果相对 SMD 数据集的实验结果变化更缓慢, 这是由于出租车轨迹数据中, 当节

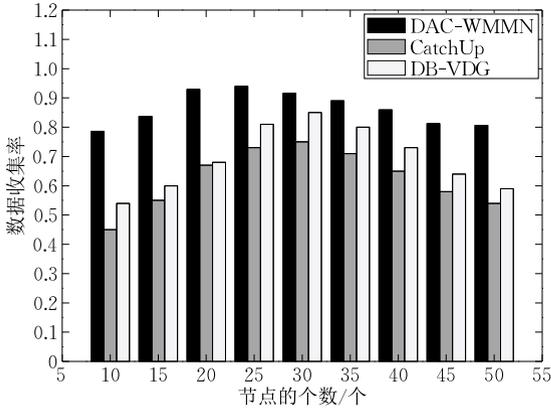
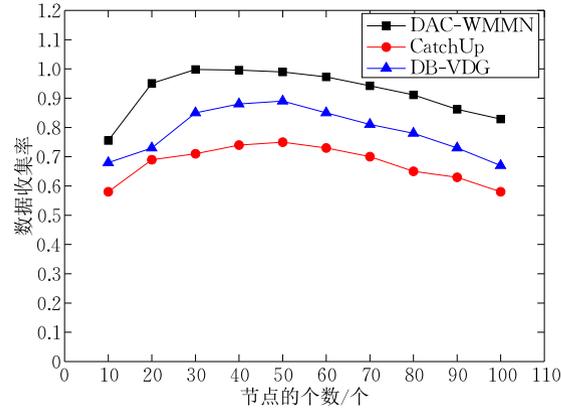


图 9 节点总数对数据收集率的影响(CSGPS 数据集)

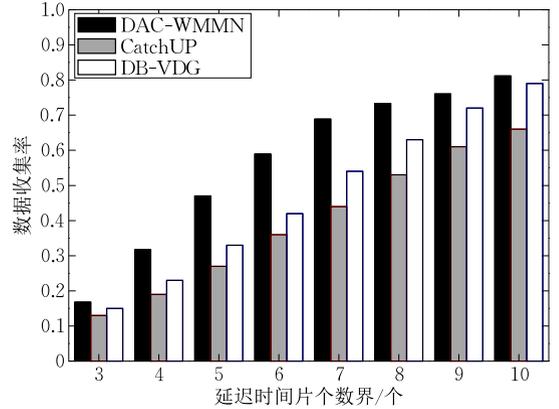


(a) SMD数据集

点个数较少时, 移动过程中通信的连通情况更弱, 当节点个数较多时, 连通情况更好, 因此, 数据收集的效率更高.

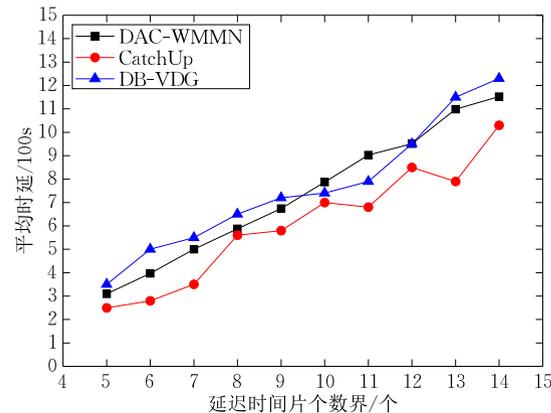
6.3.3 延迟界对性能的影响

本组实验主要研究延迟界对性能的影响. 我们在 SMD 数据集和 CSGPS 数据集上分别比较了 3 种算法的有效性和时效性. 图 10 表明, 随着延迟界限的增加, 3 种算法的数据收集率在两个数据集上都有所增加. 从图 10 中我们可以看出延迟界对数据收集率的影响是非常大的, 当延迟界较小时, 能收集上来的数据量是非常有限的, 随着延迟界的增加, DAC-WMMN 算法中的节点可规划机会更多, 因此, 数据收集率会迅速增加. 延迟界是影响数据收集的效率的重要参数. 从图 11 中可以看出 DAC-WMMN 算法的平均延迟几乎随着延迟界线性变化.

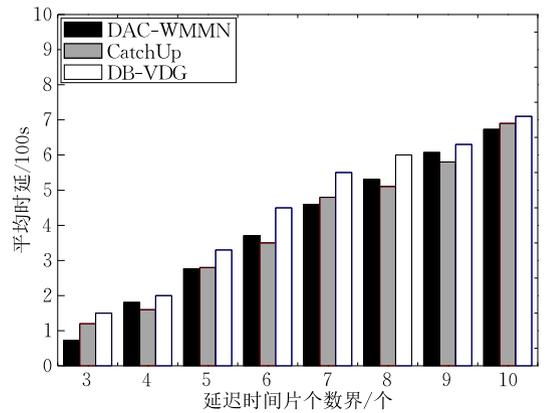


(b) CSGPS数据集

图 10 数据收集率随延迟界的变化情况



(a) SMD数据集



(b) CSGPS数据集

图 11 平均延迟随延迟界的变化情况

7 结 论

与传统的无线传感器网络相比, 无线多信道移

动感知网络具有一些特性, 使得其上的数据聚集收集过程面临许多困难和挑战. 为此, 我们引入移动通信拓扑模型, 提出了 DAC-WMMN 算法, 该算法包括两个部分: 首先根据 CreateMRT 算法建立移

动聚集路由结构,然后根据无冲突传输时刻规划 DAS-MRT 算法优化传输时刻。CreateMRT 算法构建 k 约束优化聚集路由结构,为后面的规划提供了更多可规划的空间,同时 DAS-MRT 算法是在移动聚集路由结构上无冲突传输规划的最优算法。通过在真实出租车移动轨迹数据集和人工模拟轨迹数据集上的大量实验,表明 DAC-WMMN 算法具有更优异的性能。在数据聚集收集过程中,消耗的传输带宽为最低值,同时有更高的数据收集率以及较稳定的平均延迟。

参 考 文 献

- [1] Hannes H, Kenneth P. VANET: Vehicular Applications and Inter-Networking Technologies. UK: Wiley Press, 2010
- [2] Lee U, Zhou B, Gerla M, et al. Mobeyes: Smart mobs for urban monitoring with a vehicular sensor network. *IEEE Wireless Communications*, 2006, 13(5): 52-57
- [3] Beresford A R, Bacon J. Intelligent transportation systems. *IEEE Pervasive Computing*, 2006, 5(4): 63-67
- [4] Zhu J, Roy S. MAC for dedicated short range communications in intelligent transport system. *IEEE Communications Magazine*, 2006, 41(12): 60-67
- [5] Sukumaran S, Ramachandran L, Sunny S R. Study of various geographic routing protocols in VANET. *International Journal of Research in Wireless Systems*, 2013, 2(1): 45-49
- [6] Wischoff L, Ebner A, Rohling H, Lott M. SOTIS—A self-organizing traffic information system//Proceedings of the Vehicular Technology Conference, 2003-Spring, Seoul, Korea, 2003: 2442-2446
- [7] Tamer N, Sasan D, Chunyuan L, Liviu I. TrafficView: Traffic data dissemination using car-to-car communication//Proceedings of the ACM Mobile Computing and Communications Review, Special Issue on MDM (MDM'04). Boston, USA, 2004: 6-19
- [8] Salhi I, Cherif M O, Senouci S M. A new architecture for data collection in vehicular networks//Proceedings of the IEEE International Conference on Communications (ICC'09). Dresden, Germany, 2009: 1-6
- [9] Tian J, Marrón P J, Rothermel K. Location-based hierarchical data aggregation in vehicular Ad Hoc networks//Müller P, Gotzhein R, Schmitt J B eds. *Kommunikation in Verteilten Systemen (KiVS)*. Berlin, Germany: Springer, 2005: 166-177
- [10] Murat C, Daniel G, Martin M. Decentralized discovery of free parking places//Proceedings of the 3rd ACM International Workshop on Vehicular Ad Hoc Networks. Los Angeles, USA, 2006: 30-40
- [11] Christian L, Scheuermann B, Christian W, et al. Data aggregation and roadside unit placement for a VANET traffic information system//Proceedings of the 5th ACM International Workshop on Vehicular Inter-NETworking. San Francisco, USA, 2008: 58-65
- [12] Khaled I, Michele C. CASCADE: Cluster-based accurate syntactic compression of aggregated data in VANETs//Proceedings of the Globe Communication Conference 2008. New Orleans, USA, 2008: 1-10
- [13] Khaled I, Michele C W. Optimizing CASCADE data aggregation for VANETs//Proceedings of the 5th IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS'08). Atlanta, USA, 2008: 724-729
- [14] Schoch E, Dietzel S, Bako B, Kargl F. A structure-free aggregation framework for vehicular Ad Hoc networks//Proceedings of the 6th International Workshop on Intelligent Transportation. Hamburg, Germany, 2009: 35-47
- [15] Dan W, Jianliang X, Feng W, Jiangchuan L. Mobile filter: Exploring filter migration for error-bounded continuous sensor data collection. *IEEE Transactions on Vehicular Technology*, 2010, 59(8): 4093-4104
- [16] Khaled I, Michele C. Accurate data aggregation for VANETs//Proceedings of the 4th ACM International Workshop on Vehicular Ad-Hoc Networks. New York, USA, 2007: 71-74
- [17] Bo Y, Jiayu G, Chengzhong X. Catch-Up: A data aggregation scheme for VANETs//Proceedings of the 5th ACM International Workshop on Vehicular Inter-NETworking. New York, USA, 2008: 49-57
- [18] Bo Y, Chengzhong X. Adaptive forwarding delay control for VANET data aggregation. *IEEE Transactions on Parallel and Distributed Systems*, 2012, 23(1): 11-18
- [19] Claudio E, Fabio P, Pedro M. Delay-bounded data gathering in urban vehicular sensor networks. *Journal on Pervasive and Mobile Computing*, 2012, 8(2): 180-193
- [20] van der Pim T. Data aggregation in V2V and V2I communication infrastructures//Proceedings of the 13th Twente Student Conference on IT. Enschede, Netherlands, 2010: 765-771
- [21] Rangari A B, Agrawai G, Bhanarkar P. Data aggregation control for VANET: A reviews. *International Journal of Engineering Research & Technology*, 2013, 2(4): 1-10
- [22] Dietzel S, Kargl F, Heijenk G, Schaub F. On the potential of generic Modeling for VANET data aggregation protocols//Proceedings of the Vehicular Networking Conference. Jersey City, USA, 2010: 78-85
- [23] Bjorn S, Christian L, Jędrzej R, Martin M. A fundamental scalability criterion for data aggregation in VANETs//Proceedings of the 15th Annual International Conference on Mobile Computing and Networking. Beijing, China, 2009: 285-296
- [24] Bjorn S, Christian L, Jędrzej R, Martin M. A probabilistic method for cooperative hierarchical aggregation of data in VANETs. *Journal on Ad Hoc Networks*, 2010, 8(5): 518-530

- [25] Suhas M, Sanjit K, Marco G, Wade T. ParkNet: A mobile sensor network for harvesting real time vehicular parking information//Proceedings of the 2009 Mobihoc S3 Workshop on Mobihoc S3. Beijing, China, 2009: 25-28
- [26] Mohammad N, Ana A. Data gathering for sensing applications in vehicular networks. 2012. arXiv: 1209.0684
- [27] Bruno D, Thierry D, Sergio I, et al. Data aggregation in VANETs: The VESPA approach//Proceedings of the 5th Annual International Conference on Mobile and Ubiquitous Systems; Computing, Networking and Services. Belgium, Ireland, 2008: 13-25
- [28] Thomas H, Char E, Ronald L, Stein C. Introduction to Algorithm. 2nd Edition. USA: The MIT Press, 2008
- [29] Piorowski M, Sarafijanovic-Djukic N, Grossglauser M. A parsimonious model of mobile partitioned networks with clustering//Proceedings of the COMSNETS 09. Bangalore, India, 2009: 156-163



FENG Cheng, born in 1986, Ph. D. candidate. Her research interests include data collection and data aggregation in mobile sensor networks.

LI Zhi-Jun, born in 1977, Ph. D. , associate professor. His research interests include P2P, pervasive computing, wireless sensor networks and operation system.

JIANG Shou-Xu, born in 1968, Ph. D. , professor, Ph. D. supervisor. His research interests include P2P, database, wireless sensor networks and pervasive computing.

Background

Real-time data gathering is the basic research work of Intelligent Traffic System. For ITS applications, the goal of data gathering is to get accurate real-time sensor data as much as possible with dynamic wireless resource. Data aggregation is a useful technology in the process of data gathering in order to use communication bandwidth effectively. Many research results have been gained in this area, but they did not consider multi channels and dynamic communication topology. We study data aggregation and routing scheduling in wireless multi-channel mobile sensor networks. Its main

idea is to utilize multi-channel to improve the capacity of wireless transmission and utilize dynamic communication topology to schedule transmission time without conflicts.

This research is supported by the National Natural Science Foundation of China (No. 61370214). These projects focus on the data processing in VANETs. Our group has been working on the in-network data processing methods and data routing and gathering in VANETs. Some research results have been accepted by VTC'12, WAIM'12 or other conferences and international journals.