

# 基于 VANETs 修改的 K-means 分簇路由算法

许力文<sup>1,2</sup>, 乔丽娟<sup>1</sup>, 陈 杰<sup>1</sup>

(1. 海南热带海洋学院 海洋信息工程学院, 海南 三亚 572022;

2. 武汉大学 计算机学院, 湖北 武汉 430072)

**摘 要:** 基于 VANETs 分簇是指以最小能量开销将节点车辆通过集中形态选择出称为簇头的负责节点来管理其他节点的群集, 其结构相对简单, 易于管理, 使信息传输链路更加稳定、避免消息大量冗余导致严重的网络拥塞, 减少时延及有效提高了交通安全和通行效率。通过修改的 K-means 算法和 Floyd-Warshall 算法, 介绍了一种 MKCR 的分簇路由协议, 用于分簇形成和簇头 (CH) 选择。通过假设车辆流为截断的正态分布, 定义车辆速度的置信区间, 修改的 K-means 算法将车辆划分成在其速度置信范围内的三个簇, Floyd-Warshall 算法计算所有 VANETs 车辆的最短距离, 将具有到其余车辆的最小平均距离且具有最小速度方差的车辆选择为簇头 (CH)。经分析和模拟结果表明, MKCR 路由协议能够使簇集形状一致, 避免在新一轮 CH 重新选择, 从而形成稳定的车辆节点群集。

**关键词:** 车辆自组织网络; 分簇; 路由; 算法

中图分类号: TP393

文献标识码: A

文章编号: 1673-629X(2018)03-0015-05

doi: 10.3969/j.issn.1673-629X.2018.03.004

## A K-means Clustering Routing Algorithm Based on Modified VANETs

XU Li-wen<sup>1,2</sup>, QIAO Li-juan<sup>1</sup>, CHEN Jie<sup>1</sup>

(1. School of Ocean Information and Engineering, Hainan Tropical Ocean University, Sanya 572022, China;

2. School of Computer Science, Wuhan University, Wuhan 430072, China)

**Abstract:** VANETs-based clustering refers to the cluster with the minimum energy cost to select the node responsible for the cluster head to manage the nodes of the other nodes through the centralized form. Its structure is relatively simple and easy to manage, which makes the information transmission link more stable and avoids the network congestion caused by large number of messages redundancy, reducing latency and increasing traffic safety and efficiency. In this paper, according to a modified K-means algorithm and Floyd-Warshall algorithm we introduce a clustering routing protocol for MKCR for cluster formation and cluster head (CH) selection. The modified K-means algorithm divides the vehicle into three clusters within its confidence range by assuming a normal distribution of the vehicle flow as a truncated vehicle. The Floyd-Warshall algorithm calculates the shortest distance of all VANETs vehicles. The vehicle with the smallest average distance to the remaining vehicle and the smallest variance of the vehicle will be selected as the CH. The analysis and simulation shows that the MKCR routing protocol can make the cluster shape consistent and avoid the re-selection in the new round of CH, thus forming a fairly stable vehicle node cluster.

**Key words:** VANET; clustering; routing; algorithm

### 1 概 述

随着机动车保有量、路网结构、城市规模的不断增长与扩张, 交通拥堵、交通安全和节能减排引起了政府、学术和工业界的高度关注<sup>[1]</sup>。车辆自组织网络 (vehicle ad-hoc networks, VANETs) 是一种提高交通安全系数、提升乘车人员舒适度的网络系统, 是智能交

通系统 (intelligent transportation system, ITS) 的重要组成部分之一<sup>[2]</sup>。

美国联邦通信委员会 (federal communications commission, FCC) 在 1999 年 10 月就为专用短距离通信 (dedicated short range communication, DSRC) 服务在 5.9 GHz 频率处为 VANET 分配了宽 75 MHz 的专用

收稿日期: 2017-04-10

修回日期: 2017-08-16

网络出版时间: 2017-12-05

基金项目: 海南省高等学校科学研究项目 (Hnky2017-50); 海南省自然科学基金 (117150)

作者简介: 许力文 (1980-), 男, 硕士, 高级实验师, CCF 会员 (69629M), 研究方向为网络安全、车联网技术。

网络出版地址: <http://jns.cnki.net/kcms/detail/61.1450.TP.20171205.0904.032.html>

通信频段,实现车与车(vehicle to vehicle, V2V)、车与路边设施(vehicle to infrastructure, V2I)间的通信会话,并订制了 VANET 基于 IEEE802.11p 协议的规定和标准<sup>[3-4]</sup>。

VANETs 其信息可靠传输性是通信会话的重要保障。文献[5]通过动态地更改相邻车辆之间的传输速率,使得数据更易于被接收和聚合;提出的基于 DSRC 控制信道的分布式跨层设计方法提高了 VANETs 中安全信息的快速传输可靠性;采用了基于车辆探测信息的功率控制算法,周期性地给车辆节点发送探测信息,实时对整个网络中的车辆数目和连通性状态进行分析,自适应地更改发射功率从而降低网络拥塞状况。

这些方法很好地实现了 V2V 和 V2I 间的通信会话,其优点为既定拓扑、持续能量、高效计算和车辆定位等<sup>[6]</sup>。由于车辆高速移动的特征,造成了动态网络拓扑和多径衰落,因此控制信道带宽和通信距离存在受限的问题,导致了 V2V 通信链路随机中断,使得车间接通信会话不稳定和不可靠<sup>[4]</sup>。主要存在的不足是:车辆簇头产生时对节点的状态考虑不够,使得能量极少的节点担任车辆簇头,导致节点车辆能量快速耗尽而提前死亡,缩短了簇群生存时间;存在极大簇或极小簇不均匀的较大区别;存在簇头在监测区域聚集分布不均匀的情况。

分簇形成的集群数过多则计算复杂度增大,集群体过大则簇间过于变化频繁簇头能耗过多易于更换。因此设计出优化的分簇路由算法至关重要。K-means 是一种常见的均值聚类方法,优点是快速、简单,对大数据有较高效率,时间复杂度近于线性<sup>[7]</sup>,为此提出基于 K-means 均值算法对车辆节点执行分簇形成,从而有效提高 VANETs 车辆节点的连通性和稳定性。

文中提出一种基于 VANETs 修改的 K-means 分簇路由协议(modified K-means clustering routing, MKCR)。协议设计分三步完成。第一步是车流形成簇群,即采用修改的 K-means 算法将有效传输范围内的车辆分簇,形成三个分簇区域;第二步是采用 Floyd-Warshall 算法来获得有关独自分簇区域内每个车辆节点的平均速度、位置和方向等信息;第三步是根据指定的集中位置和速度方差两个参数选择车分簇的簇头车辆(VH)。

在该方案中,根据车辆速度的截断正态分布,设置了车辆的速度置信区间,每辆车节点的速度在置信区间内被指定为分簇。具有中心位置和 VAR 的小聚集的车辆节点将选为 CH(cluster head)。MKCR 不仅考虑了簇的数量、簇的密度和簇的结构,构造出一致的车辆节点集群形状,同时还考虑簇头的寿命时间,能长时间维持簇的稳定,从而增加车辆之间的互连的寿命,并

以较少的迭代重新选择 CH。

## 2 相关工作

近年来,为了解决 VANETs 应用问题中通信连接的稳定性,研究人员提出了很多解决方案。广播算法的泛洪机制具有简单、快速和覆盖率广的特点,因此应用最为广泛,但却会引起恶劣的广播风暴,造成广播信息过大的冗余、过多能耗和恶劣的资源竞争及冲突。由 Tzay-Farn Shih 等<sup>[8]</sup>提出的核心定位辅助集群路由协议(CLACR),将通信区域划分为正方形簇,然后通过使用簇头算法为每个正方形选择簇头,减少了广播风暴和路由开销,但 CLACR 不适合于稀疏或复杂的 VANETs 场景。An Huiyao 等<sup>[9]</sup>提出基于集群的多路径动态源路由(cluster-based multipath dynamic source routing in MANET, CMDSR)是一种三级层次结构方案属于基于邻居的聚类集群,其中每个级别具有自己的簇头并且其余节点是一跳远距离的邻居,基于邻居的协议是混合路由协议(HRP)<sup>[10]</sup>;Bilal M 等<sup>[11]</sup>提出 FMHR 算法,是基于贪婪算法把最后下一跳的车辆节点做为簇头,提高了链路的相对稳定性,但同时考虑了车辆节点的运动速度和方向,算法增加了计算负载和额外开销。Cross-CBRP<sup>[12]</sup>算法在物理层使用功率信号信息进行路由以增加群集的寿命,有效提高网络吞吐量。胡升泽等<sup>[13]</sup>提出的 CMCH 算法能很好地降低单个簇头的负担,周期性地选择簇头,但由于节点移动频繁而不断更换簇头,增加了网络开销。Saleet H 等<sup>[14]</sup>提出的 IGRP 协议,Naumov V 等<sup>[15]</sup>提的 CAR 都是基于锚点的路由算法,通过路边锚节点进行数据分组转发,能够改善稀疏场景中由于网络分离导致数据不可达的问题,但对锚节点依赖性强,在区域中若有锚节点损坏或阻塞,则将导致整个网络瘫痪。COIN 算法适应车辆间距的振荡性质<sup>[16]</sup>,COIN 的主要目的是提高簇的稳定性。

上述方法是依据簇进行节点管理和数据转发,较好地完成分簇并选择出簇头,降低路由发现开销和广播风暴,增加网络有效吞吐量,延长网络生存期,提高网络的稳定性。在 VANETs 基于分簇的路由算法中,簇划分是研究方向之一,本节通过模拟测试并对比 COIN 算法,分析和讨论在开放式车辆间通信网络中 MKCR 算法的优越性及分簇车头(VH)节点的稳定性。

## 3 系统模型

VANET 的车辆节点具有高速移动的动态性质,所以设计出一款最优迭代算法的路由机制对研究人员提出了更高的要求。结合十字路口场景,文中提出了一种 MKCR 协议,可以有效地扩展车辆之间的通信会

话。在十字路口场景中,所有车辆以非均匀速度行驶,由于交通管制要求在一定速度范围内行驶,车辆可能在道路上造成持续的交通模式。设计集中式路由协议的要求是流量模式的一致性,文中为一致的分簇定义了两个参数,即车辆速度的截断正态分布和簇内车辆的位置。

3.1 簇的形成

簇的形成通过修改的 K-mean 算法进行, K-means 的主要思想是识别用于车辆节点分类的质心或分割点或聚类点( cluster points,CP)。文中根据 DSRC 的传输范围有三个分割区域,每个分割区域将定义一个独立的分簇。在指定范围内,三个分割点 C<sub>1</sub>、C<sub>2</sub>和 C<sub>3</sub>被认为是三个车辆群的中心点。C<sub>1</sub>和 C<sub>3</sub>是范围的最远分割点,C<sub>2</sub>居于 C<sub>1</sub>和 C<sub>3</sub>的中间,如图 1 所示。

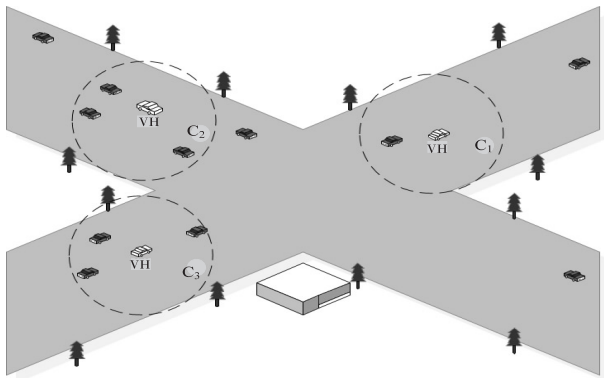


图 1 分簇形成及车辆节点分布结构

区域中的车辆节点在分类成簇之前,首先设定了车辆速度的置信区间。截断期间的正态分布模型  $N(\mu, \sigma)$  用来定义速度建模分布,其中  $\mu$  是车辆速度的平均值,  $\sigma$  是标准偏差。通过截断分布模型定义了平均值周围的速度的置信区间,如式(1):

$$E - \alpha_n < E < E + \alpha_n \tag{1}$$

其中,  $E$  表示车辆速度的期望值;  $\alpha_n$  表示速度间隔的百分比置信度。只要车辆节点的速度在式(1)区间内,此车辆均是分簇的考虑对象。

程序 1 代码描述了用修改的 K-means 方法形成三个分簇,即将道路划分为三个独立区域,其中每个区域形成图 1 所示的簇结构。

程序 1:Modified K-means clustering.

```
Void Cluster_Formation(N)
Begin
for i<----0 to N do
if distance1[i]<distance2[i] && distance1[i] <distance3[i]
then
Cluster1[ C1 ] <----i
C1++
else if distance2[i]<distance1[i] &&
distance2[i]<distance3[i]
```

```
then
Cluster2[ C2 ]
C2++
else
Cluster3[ C3 ]
C3++
End
```

其中, distance1[ i ]表示监测区域内任意车辆节点 i 到独立分割区 C<sub>1</sub> 的距离,同样 distance2[ i ]和 distance3[ i ]分别表示到 C<sub>2</sub> 和 C<sub>3</sub> 的距离。

3.2 簇头的选择

簇头负责本簇内部的统一管理,簇头的选择最为关键。本簇内簇成员仅上传本地消息给簇头,并接收簇头的广播消息。簇头直接连接两个相邻簇或路边设备,实现了簇头的消息中继。在 VANET 中簇的稳定性定义为簇持续生存的时间,由于 VANET 的随机性质,在簇内每个车辆彼此交换 HELLO 消息,消息中包括速度、位置和方向等信息。设置簇头选择的两个参数,即簇群内的集中位置和速度方差。

程序 2:Algorithm for centralized vehicle.

```
void Centralize_Vehicle( )
Begin
for i<----0 to C1 do
for j<----0 to C1 do
Distance[ i ] [ j ]sqrt( pow( ( final_xcorl[ i ]-final_xcorl[ j ] ),
2)+pow( ( final_xcorl[ i ]-final_xcorl[ j ] ), 2) )
count<----0
counter<----0
sum<----0.0
for i<----0 to C1 do
for j<----0 to C1 do
sum<----sum+Distance[ i ] [ j ]
counter++
Center_Set[ count ] sum/counter
Center<----Center_Set[0]
for I<----0 to C1 do
if Center>Center_Set[ i ]
then
Center<----Center_Set[ i ]
Cluster_Center_Vehicle<----i
End
```

以上算法描述了车辆集中选择,如图 1 所示,其中白色车辆分别表示簇 C<sub>1</sub>、C<sub>2</sub>和 C<sub>3</sub>的集中式车辆。

接着,使用动态规划算法中的 Floyd-Warshall 算法来选择集中式车辆。其计算所有车辆的最短距离,任两车间距离为欧几里德距离。程序 2 中的伪代码表示集中车辆选择 C<sub>1</sub> 簇群。

最后阶段设定两个参数来计算每个车辆节点的累积分数( cumulative score, CS ),这两个参数是围绕均

值 $\mu$ 的集中位置和速度方差( $\text{Var}(V_i)$ )。CS 为两个参数的累积值,即:

$$\text{CS}_i = \text{Aggregate}(\text{VAR}(V_i) + \text{ADV}_i) \tag{2}$$

其中, $\text{Var}(V_i)$ 是车辆 $V_i$ 速度的方差平均值,其与群集中的 VH 的寿命成反比。在 VH 选择标准中给 $\text{VAR}(V_i)$ 分配 60% 的权重。ADV(平均距离值)是每个单独的车辆与集群内的其他车辆的平均距离。每个节点的 ADV 由集中式算法计算。在 MKCR 中,ADV 的权重是 40%。具有最低 CS 的任何节点将被选择为 CH 的候选。

4 性能分析与仿真

本节通过模拟测试并对比 COIN 算法,分析和讨论了在开放式车辆间通信网络中 MKCR 算法的优越性及分簇车头(VH)节点的稳定性。

4.1 环境设置

模拟环境根据 DSRC 的规范进行设置。每个车辆的运动范围设置为 1 000 m。车辆的速度被认为是非均匀的,即为 60~120 km/h。车辆密度从 10~100 辆不等<sup>[17]</sup>。表 1 为 MKCR 和 COIN<sup>[16]</sup>做模拟测试比较时所用的参数设置值。

表 1 MKCR 和 COIN 对比的参数值

参数	值
Number of vehicles	10
Average distance from neighbors	1~100 m
Estimated mean $\mu$	100 km/h
Standard deviation $\sigma$	30
Selected clusters	1
Protocols	COIN, MKCR

4.2 簇头开销

如在 MKCR 中,有三个固定的群集,其通过修改的 $K$ 均值计算。根据文中提出的方法,聚类开销(CO)的计算为:

$$\text{CO}_{\text{MKCR}} = \sum_{C=1}^N \sum_{K=1}^n \min_{\text{Euc\_Dist}}(\text{CP}_C, V_k) \tag{3}$$

其中, $N$ 和 $n$ 分别是分簇的数量和每个簇群的车辆数量;CP 是簇类分割点,并且在传输区域内有三个分割点。在模拟测试中,高速公路的端点和中点被指定为分割点; $V$ 是提名为分簇部分的车辆;Min 函数找到 CP 和正常车辆之间的最小欧氏距离。

在正常 $K$ 均值的情况下,CO 的时间复杂度为 $O(KnT_{\text{dist}})$ 。 $K$ 表示组的数量, $n$ 表示每组的对象数量。如在 MKCR 中,簇的数量是固定的,因此累积时间复杂度将是 $O(3nT_{\text{dist}})$ 。另一方面,在 COIN 中,车辆分为两组 $C_i^v$ 和 $V-C_i^v$ 。在选择的分断正态分布中,高速公路的车辆数据范围不太分散。由于车辆的随机

分散, $K$ 的值将大于 3。因此,COIN 簇形成的总时间复杂度将高于 MKCR。

4.3 适应的车辆速度置信区间

簇的形成是每种分簇算法的初始步骤,簇成员的稳定性趋于一致的分簇。COIN 协议随机地将该车辆节点分成两组,称为正常车辆和簇头。COIN 的主要焦点是选择适当的簇头,而不是适当的簇成员。而在 MKCR 中,类的稳定性从簇形成的初始阶段考虑。初始设置是定义车辆在高速公路上的速度的置信区间。定义 95% 的集群形成的置信区间。在图 2 中,阴影区域显示车辆速度的置信区间。根据定义的置信区间,只有中等车辆将成为集群的一部分,因此集群的整体一致性将增强。

4.4 簇头的稳定性

在动态的 VANET 中,VH 的稳定性更为突出。而分簇优化聚群的目的是通过构造一致的车辆组来延长 VH 的寿命,本节比较了 MKCR 和 COIN 的稳定性。该仿真是在 MKCR 和 COIN 相同车辆数量和规格下进行的。

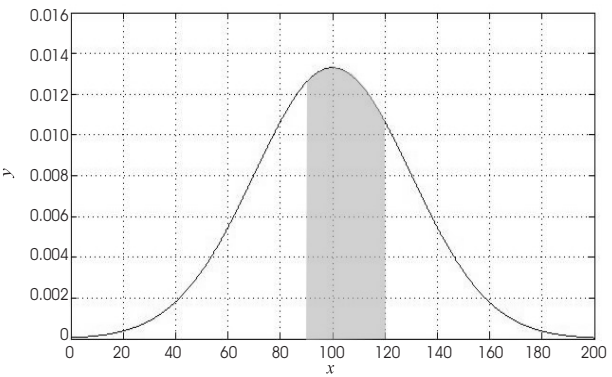


图 2 车辆速度的置信区间

COIN 的簇头选择是以簇头集合 $C_i^v$ 内在时间 $t$ 的最近的车辆 $V$ 。另一方面,MKCR 基于式(2)给出的两个参数来选择 CH。

由式(2)表明,CS 越小,VH 的稳定性越大。一些用于模拟的样品,以及 MKCR 车头 VH 选择的结果如表 2 所示。

表 2 MKCR 的车辆和车头选择

V_ID	ADV	VAR( $V_i$ )	CSi
1	11	68	54.8
2	14	51	63.8
3	43	64	44.4
4	16	43	67.8
5	44	87	30.2
6	9	94	40
7	96	19	50.2
8	57	27	61
9	63	56	41.2
10	44	69	41



表2中,车辆V\_ID4由TCRP选择为VH,因为其CS速率大于所有其他集群成员。然而,在COIN的情况下,具有V\_ID6的车辆将被选为簇头,因为簇内的最小ADV。选择COIN是最差的,因为V\_ID6与其他成员的速度差异相当大,并且簇头的存在将是非常短的时间间隔。速度差在VH的选择中非常重要。将保留高速的集中式车辆在集群中的时间非常短。距离不是测量稳定性的唯一标准,因为速度的高方差能快速的影响平均距离。

5 结束语

基于VANETs的车辆节点具有高速移动的动态性,导致车辆节点通信会话不稳定。文中提出的MKCR分簇路由算法是把车辆节点通过形成群集的形式,分簇的持续稳定性很好地控制和管理节点,达到消除消息泛洪的目的,减少不必要的路由开销。其中,CH负责管理其中的分簇车辆节点,同时也负责簇间互连,稳定的VH可以增加簇形状的持续寿命时间。首先采用修改的K-means算法形成分簇,然后利用Floyd-Warshall算法选择分簇的集中和稳定的CH。CH基于它们的相邻距离来选择,最佳的CH具有集中式的位置并且具有较小的速度方差。

通过分析和模拟结果表明,MKCR路由协议能够使群集形状一致,避免在新一轮CH重新选择,从而形成相当稳定的车辆节点群集,使网络具有较好的稳定性,有效提高信息传输。

参考文献:

[1] 吴黎兵,杜锦,聂雷,等.无线传感器网络动态k值簇头选择方法[J]. 华中科技大学学报:自然科学版,2015,43(10):37-41.

[2] 唐伦,王晨梦,陈前斌.车载自组织网络中基于时分复用的异步多信道MAC协议[J]. 计算机学报,2015,38(3):673-684.

[3] 邱恭安,包志华,章国安,等.高稳定被动群集车联网连通性研究[J]. 通信学报,2016,37(11):42-48.

[4] HARTENSTEIN H, LABERTEAUX K P. VANET: 车载网技术及应用[M]. 孙得民,译. 北京:清华大学出版社,2013.

[5] 默罕莫德·默森,许凯凯,夏玮玮,等.荒漠场景应用的车联网及其分簇路由算法[J]. 通信学报,2012,33(10):166-174.

[6] 程嘉朗,倪巍,吴维刚,等.车载自组织网络在智能交通

中的应用研究综述[J]. 计算机科学,2014,41(6A):1-10.

[7] 余秀雅,刘东平,杨军.基于K-means++的无线传感网分簇算法研究[J]. 计算机应用研究,2017,34(1):181-185.

[8] SHIH T F, YEN H C. Core location-aided cluster-based routing protocol for mobile ad hoc networks[C]//Proceedings of the 10th WSEAS international conference on communications. Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society,2006:223-228.

[9] AN Huiyao, ZHONG Ling, LU Xicheng, et al. A cluster-based multipath dynamic source routing in MANET[C]//IEEE international conference on wireless and mobile computing, networking and communications. [s. l.]: IEEE, 2005:369-376.

[10] AHN C W, RAMAKRISHNA R S, KANG C G. Efficient clustering-based routing protocol in mobile ad-hoc networks[C]//Vehicular technology conference. [s. l.]: [s. n.], 2002.

[11] BILAL M, CHAN P M L, PILLAI P. A fastest multi-hop routing scheme for information dissemination in vehicular communication systems[C]//International conference on software telecommunications and computer networks. [s. l.]: IEEE, 2010:35-41.

[12] DANA A, YADEGARI A M, HAJHOSSEINI M, et al. A robust cross-layer design of clustering-based routing protocol for MANET[C]//10th international conference on advanced communication technology. [s. l.]: [s. n.], 2008:1055-1059.

[13] 胡升泽,包卫东,王博等.无线传感器网络基于多元簇首的分簇数据收集算法[J]. 电子与信息学报, 2014,36(2):403-408.

[14] SALEET H, LANGAR R, NAIK K. Intersection-based geographical routing protocol for VANETs: a proposal and analysis[J]. IEEE Transactions on Vehicular Technology, 2011, 60(9):4560-4574.

[15] NAUMOV V, GROSS T R. Connectivity-aware routing (car) in vehicular ad-hoc networks[C]//IEEE international conference on computer communications. [s. l.]: [s. n.], 2007:1919-1927.

[16] BLUM J, ESKANDRIAN A, HOFFMAN L. Mobility management in IVC networks[C]//IEEE intelligent vehicles symposium. [s. l.]: IEEE, 2003:150-155.

[17] 吴怡,沈连丰,邵震洪,等.基于探测信息的车辆自组织网络功率控制算法[J]. 东南大学学报:自然科学版,2011,41(4):659-664.