International Journal on
"Technical and Physical Problems of Engineering"
(IJTPE)
Published by International Organization of IOTPE

ISSN 2077-3528
IJTPE Journal
www.iotpe.com
ijtpe@iotpe.com

December 2021 Issue 49 Volume 13 Number 4 Pages 7-16

NEW MACHINE LEARNING SOLUTION BASED ON CLUSTERING FOR DELAY-SENSITIVE APPLICATION IN VANET

S. Ouahou S. Bah Z. Bakkoury

AMIPS Laboratory, Mohammadia School of Engineering, Mohammed V University, Rabat, Morocco
soufiane.ouahou@research.emi.ac.ma, bah@emi.ac.ma, bakkoury@emi.ac.ma

Abstract- VANET or Vehicular Ad-hoc Networks have attracted huge considerations due to the offered road safety and driver comfort. Unsupervised machine learning especially clustering, is an effective solution of data dissemination in VANET, thanks to the legible data sharing during a travel. Nonetheless, the frequent links disconnection and topology changes caused by the high mobility of vehicular environment create many challenges to the efficient data delivery and QoS in VANET. To establish a stable and reliable communication between nodes, we propose in this paper a clustering scheme as an unsupervised machine learning solution with self-stabilisation approach for delay-sensitive application in VANET. The propose of this work is to deal with the delay of data sharing and offer high data availability as well as reduce packet loss for VANET multi-hop architecture. The simulation results show that our solution is perfectly adapted to the problems of delays and dissemination packets loss, as well as our self-stabilization approach improves the stability of the network.

Keywords: Vehicular Ad-hoc Network, Machine Learning, Clustering, QoS, Multi-Hop, Self-Stabilization.

1. INTRODUCTION

During the last years, Vehicular Ad-Hoc Network or VANET has attracted considerable industrial and academic attention because of the new generations of networks. The VANET network has become a practical innovation field, especially the cooperative applications like roads safety, congestion supervision and multimedia applications.

In VANET, nodes can be vehicles supplied with a smart device called on-board unit (OBU); or fixed nodes set on the road called Road Side Unit (RSU). The vehicles can reach each other by V2V links (vehicle to vehicle) or V2I links (vehicle to infrastructure).

The communication in VANET can support a range between 100 to 300m in urban environment, and a maximum range of 1000m in highway environment [1]. Many communication standards of VANET have been developed especially the dedicated short-range communication (DSRC) as a United States solution [2], and ITS-G5 as a Europe solution [3]. However, these solutions based on 802.11p technology, suffer from different issues, including insufficient high availability of data, competing channel access and low QoS.

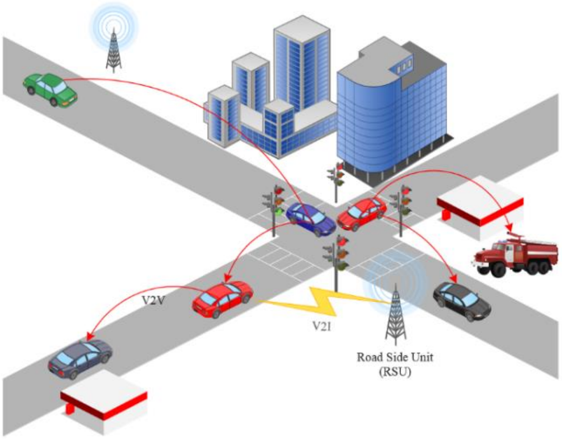

Figure 1. VANET network architecture

In VANET, multi-media data (exp: maps, streaming etc.) dissemination is complicated problem: indeed, these applications require huge stock resources (unavailable on vehicles), frequent internet access through the OBU or RSU (complicated in real life) and interesting bandwidth (high cost). For example, when a driver needs the map of his road, it is difficult to download the map of the whole road with precision in HD quality; as well as it is difficult to access the RSU in non-urban areas and download maps in parts as the road progresses.

Divers researches were conducted to design dissemination schemes based on different metrics like delay, density, distance etc., and other solutions based on probability data dissemination [4]. However, in such environment like VANETs, a blinded structure without a dissemination management portal will cause serious problems including hidden nodes, channel contentions, packets loss and packet delay delivery. Therefore, to fight against these problems, it is mandatory to set up a way that allows nodes to decide themselves whether to send data or not based on unsupervised machine learning solution especially clustering.

Given that VANETs is an ad-hoc network, clustering is an auspicious solution for this kind of structures by creating hierarchy topology without need of physical infrastructures. Furthermore, clustering allows vehicles to reach each other in a sufficient time even in no covered areas. Meanwhile, each vehicle will perfectly exchange data by request, which will allow us to avoid floods, making a data aggregation and save power.

However, there are as yet many challenges to conceive dissemination schemes by clustering, specifically: the high mobility of vehicles, the distributed topology of VANET network, the cost of managing the life of a cluster and finally the management of emissions and listening times between neighboring vehicles.

Based on the challenges above, in this work, we introduce a new clustering scheme for multimedia applications in vehicular ad-hoc network with self-stabilization approach. The propose of this work is to deal with the unavailable data problem and limited internet access in non-urban areas by making a data sharing solution between vehicles and offer comfort for drivers. The solution consists on the one hand, to build a multi-hop proactive clustering architecture ready to share multimedia data like (maps, streaming, service stations, etc…) and choose carefully the cluster head based on a new metric designed to satisfy QoS constraints. Or else, we propose a self-stabilization to improve network stability.

1.1. Related Work

Thanks to its high capability for enhancing road traffic efficiency and safety, several dissemination schemes have been developed of vehicular environment. the aim of scheme makes it possible to predict its design as well as its category. The remainder of this chapter, will present some solutions based on different metrics.

1.2. Delay Sensitive Schemes

In order to deliver the message as soon as possible and deal with some network phenomena like storm, hidden nodes etc., many solutions of delay sensitive applications have been developed. These solutions are based on the need time of nodes to build the links to deliver data.

In [26], the authors propose a dynamic model based on the beacon delay for clustering architectures in VANETs. The idea of this work, is to elect cluster-head for one-hop cluster based on beacon daily. Then, the maintenance phase intervenes to build more stable clusters with a long life time.

In [5] the solution based on intersections vehicles by making them to transmit the message quickly according to a calculated priority. To improve the data forward, vehicles deploy two kinds of forwarding mechanisms, which allows to have a better packet arrival ratio and fewer redundant retransmissions. The mechanism deployed for the priority calculation is not treated in this paper.

In [27], the authors propose a multi-hop solution to disseminate the messages. The vehicle with the best (shortest) daily to exchange parquets with his neighbors is elected as CH. During the maintenance phase, the gateways are chosen based on the number of hops between the source and the destination and le cluster-head fitness.

In [6], the authors propose a backbone for each vehicle supervised by one timer. Each node decide which data to forward just base on the message head, its state and local information. As each node has its own timer, it can cause the total stopping of the dissemination if the management of the timers is not well managed in the concise protocol.

1.3. QoS Schemes

Quality of Service (QoS) is defined by the set of constraints to be satisfied by the network to transport data from a source to a destination.

Thus, it represents the overall performance of a network. The objective of QoS is to achieve better communication behavior, so that the content of the communication is correctly routed and the network resources are optimally utilized.

Dua, et al. [7] propose a new mechanism to keep a good QoS for data dissemination by selecting different vehicles in the network. The intelligent forwarding scheme uses different weights assigned to path. Therefore, different algorithms are designed in this scheme, depending on the specifications of the environment for route construction and maintenance.

In [8], Wahab et al. propose a new clustering algorithm by balancing between QoS requirements and high-speed mobility constraints. The end of this scheme is to build up stable clusters and conserve a good stability during data dissemination and links disconnection while satisfying a good QoS.

1.4. Probability Schemes

In order to reduce information redundancy, N. Wisitpongphan and al propose [9] a probability-based scheme that allows receivers nodes to rebroadcast packets with a given probability. The solution is based on broadcast suppression timer (slotted p-persistence, weighted p-persistence and slotted 1-persistence schemes) and probabilistic three. The results prove that schemes can reduce contention and reaching up to 70 percent of packet loss mitigation or then keep a satisfying end-to-end delay.

In [10], each node may directly rebroadcast each data packet that it receives. However, this unnecessary rebroadcasting of the same data leads to redundant retransmissions, which causes unneeded occupation of the transmission channel.

There is a new version of this scheme [11], in which the author extends the idea of probabilistic-based broadcast scheme. In this solution, the probability of rebroadcasting is determined based on environmental specifications of inter-vehicle exchanges, which effectively reduces the number of redundant retransmissions.

In [12], the authors propose a new packet transmission solution based on probabilistic diffusion to get good reliability and low collision. The proposed scheme is designed for distributed architectures on VANET where each node receiving a data packet rebroadcast it according to a calculated probability. Based on this probability that combines all of these factors, each vehicle can assess whether there is another vehicle that should receive this message and that it will receive the message if retransmitted.

In [30], authors propose a theory game approach. Using parameters like energy, distance between vehicles and the base station; each sensor calculates its probability of being a cluster head.

1.5. Push-Based Schemes

For local and public interest data, push data dissemination for some way is the best way to support them. In [13], the authors present an efficient push-based dissemination protocol for vehicular networks.

To address broadcast communication challenges for various network densities, the protocol employs an optimized broadcast solution that avoids the broadcast storm phenomenon in high density networks and it efficiently deals with frequent disconnected networks.

However, this protocol is not based on power control scheme to greatly reduce the network over-load in the dense environment. Liu and al propose in [14] a novel scheme, called RPB-MD. The aim of the model is to effectively define the intended recipients in order to guarantee reliable and efficient delivery of messages to vehicles in the targeted area. In order to ensure high packet delivery rate and short delivery time, the solution introduces Directional greedy broadcast routing (DGBR).

1.6. Pull-Based Schemes

In VANEs, there are often private applications that target a particular type of data. However, Pull-based packets dissemination is sufficient for this type of data where nodes can directly reach needed data.

[15] Proposes a pull-based solution of data dissemination in VANETs. The model broadcasts the information as response message after receiving all requests from the vehicle. Due to the relatively high waiting time, which is the average time required to send all response messages to vehicles, the proposed scheme is evaluated in terms of service availability and loss of messages. Unfortunately, this scheme will increase the computation overhead costs.

For fast dissemination of safety message within the critical area, [16] propose an efficient multi-hop broadcast scheme called VMP. The solution designates multiple trusted senders by transfer delay orders and a cooperative transfer mechanism. When the privileged sender fails to transmit the alert message, the next one takes over.

1.7. Cluster-Based Schemes

For overly dynamic networks like VANETS, clustering is an effective approach to improve scalability of networking protocols, so most of the solutions offered are self-organized to create dynamic clusters.

In [17], Zhang and al, propose for multiple target regions a Geocast parquets dissemination scheme, based on two main methods including a smart geocasting initialization and maintenance. The first one includes a path sharing and path splitting scheme in order to deliver to multiple regions the abnormal traffic information quickly and efficiently. In order to reduce the network maintenance cost and reduce the message redundancy, the second procedure divides the network into small areas in each region, the broadcast is repeated to newly arrived vehicles. The proposed scheme may reduce double reception of data, while trying to avoid missing important data.

In [18], Singh et al propose a clustering algorithm based on a hybrid backbone for VANETs. For the clusters construction and cluster heads selection, the model is based on the number of links and the mobility of the vehicles. Then, the vehicles form a leadership that will elect cluster head. The cluster-head is too responsible for reorganization based on aggregate relative velocities of vehicles in the direction.

Jin and al propose a clustering algorithm that guarantees efficient clustering in [19]. The main idea of this model is similar to our solution, given that each vehicle calculates its weight based on a set of metrics. The vehicle with the smallest weight is elected as cluster head. The model uses an efficient mechanism that shows how the emergency messages are routed in clustering structure.

In [31], authors propose an algorithm to reduce clustering error based on MRI image segmentation.

2. SYSTEM ASSUMPTIONS

In this part of the paper, we will present the assumptions of our solution and then present our system model. The main Vocabulary and Definitions of this paper are listed in Table 1.

Each node can know at any time its geographical position by GPS, its speed and exchange them by periodic messages (BEACON message).

Each vehicle in the network can send and receive information with its neighbors, including speed, position and our metric Fitness by periodically broadcasting beacons.

Vehicles can exchange data without direction constraint. Max hops number of cluster is five hops ≈ 1000m, distance that is considered sufficient to receive the data and make the correct decision.

Table 1. Vocabulary and Definitions

Notation	Description
cv	Current vehicle executing the algorithm
STATE(*cv*)	The state of *cv*: Free, Cluster Member, Gateway, and Cluster Head
CS(*cv*)	The current speed of vehicle *cv*
P(*cv*)	The position of vehicle *cv*
CT(*cv*)	Current time on the vehicle *cv*
NS(*cv*)	able of the speed of neighboring vehicles of *cv* Each element of *NS* set has the following form: (id; speed; delay)

F(*cv*)	The Fitness metric of the vehicle *cv*
BF(*cv*)	The best fitness that *cv* has computed BF has the following form: (fitness; distance; id)
CH(*cv*)	Cluster Head of *cv*, *CH* has the following form: (distance; fitness; IDCH; IDGT)
GT(*cv*)	selected gateway to reach the cluster Head of *cv*
ChP(*cv*)	Table of received proposals from clusters in surroundings Each element of *CP* set is a *CH*: (distance; fitness; IDCH; IDGT)
ClusterMes	message having the form: (*ID*(*cv*); *CS*(*cv*); *F*(*cv*); *BF*(*cv*); *CH*(*cv*)
Child(*cv*)	Children IDs set of vehicle *cv*

2.1. Solution Aim

The goal of our solution is to build a proactive and stable architecture as best as possible for application requiring reliability and high data availability with certain constraints relating to the dissemination time, which must be short to allow drivers to react at the right time. Different kind of these applications are:

- *Comfort Applications:* In this type of applications, drivers can get data from vehicles and transport center that may help driver during the travel. Such application type: weather updates, road services and leisure places, tourist monuments, accessible parking area at a stopping place, global assistance handover, road congestion and city information, road navigation.
- *Efficiency Applications:* This family of applications needs a high data availability in order to offer the information as fast as possible and allow the driver to make a good choice at the right time. Such application type: Carrefour and intersections management application (traffic light controller) and road congestion management.
- *Interactive Entertainment:* This kind of applications aim to deliver multimedia and entertainment data for drivers and passengers. The major constraints of this type of application are connectivity and availability. The communication portal can be a V2V or V2I communication. Such application type: Internet access, social games, road chatting, music downloads, file sharing, E-commerce, house control, etc.

All the applications cited above need a good reliability and high data availability with certain constraints relating to the dissemination time, which must be short to allow the drivers to react at the right time.

2.2. Fitness Description

In order to exchange messages in a short time with a successful reception and based on the constraints above, we design one new metric named Fitness to measure the link performance and speed between vehicles to elect appropriate cluster Head:

$$F(cv)=\left[CS(cv)-\frac{\sum_{cm\in NS(cv)} cm.speed}{|NS(cv)|}\right]\cdot\left[\frac{\sum_{cm\in NS(cv)} cm.delay}{|NS(cv)|}\right] \tag{1}$$

The performance parameters taken into account are: the transmission delay and the transmission efficiency of the messages. The performance parameters which represent are speed as well as the number of direct neighbors (vehicles in coverage area). The speed is introduced as the vehicle with a speed close to the average speed of its neighbors has more chances to become a cluster head; in other words, if most of the neighboring vehicles are slow the elected cluster head should be slow too.

The second parameter for stability is the number of direct neighbors, which means that the vehicle with many neighbors has more chances to become a cluster head. The end for choosing such parameter is to avoid frequent disconnections caused by getaways leaving the cluster.

$$CS(cv)-\frac{\sum_{cm\in NS(cv)} cm.speed}{|NS(cv)|} \tag{2}$$

To make suitable clusters, this value means that the vehicle with a speed close to the average speed of the neighbors deserves to be the cluster head.

$$\frac{\sum_{cm\in NS(cv)} cm.delay}{|NS(cv)|} \tag{3}$$

In order to construct enough fast and efficient links with cluster head, this value is the average waiting time between the emission of a beacon and the reception of all the neighbor's beacons. Having a good beacon delay means that we have an efficient links, short distance with neighbors and enough calculation resource.

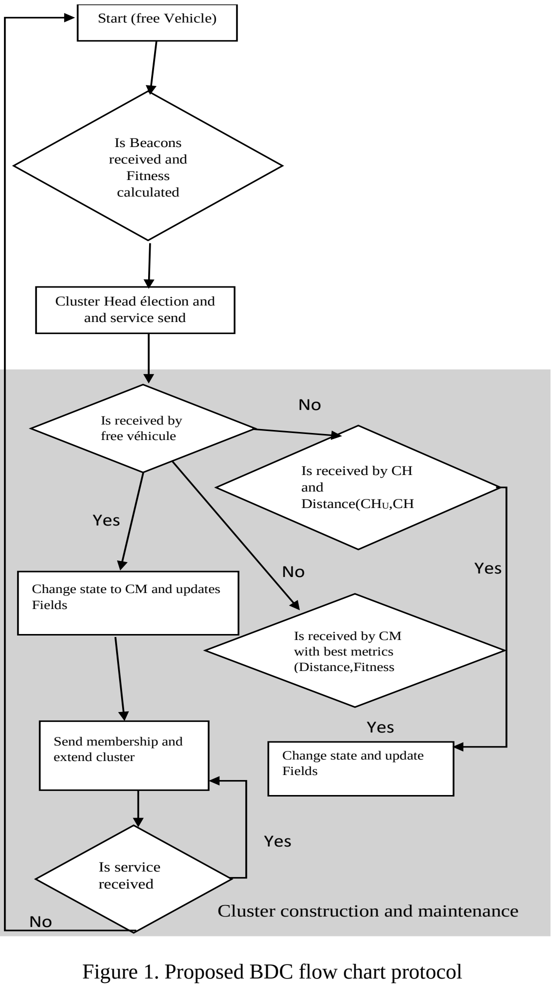

Figure 1. Proposed BDC flow chart protocol

2.3. Fitness Calculation and Cluster Head Election

In the Initial state of the network, all the vehicles are free. As the solution is a proactive algorithm, once they start communicating with each other by sending and receiving beacons; each vehicle computes its own fitness *F* and broadcast it with the neighbors in one-hop. If the received fitness is less than the stored best fitness BF, then the vehicle updates the stored best fitness BF and it rebroadcast it. Once all neighbor's fitness is received; if the best fitness is current vehicle fitness, the last elect itself as a cluster head. Otherwise, it becomes cluster member.

2.4. Description of Round Clusters Formation and Maintenance

Once the cluster head is elected, the construction phase begins. In the first step, the service is proposed by the cluster head itself to direct neighbors in the reach. Once the links are established with the cluster members in the coverage, these are responsible for the cluster extension by offering the cluster service of free vehicles in their coverage including: cluster head ID, cluster head fitness as well as the number of hops between it and the cluster head called distance. If the sender it's a cluster-head, the receiver (free vehicle) updates its cluster head ID and its fitness, then it starts to propose the service for other free vehicles in the reach of cluster.

If the sender it's a cluster member, the receiver updates its cluster head ID, fitness and the gateway ID.

In some cases, a free vehicle may receive several offers at the same time: in this case, the free vehicle chooses the cluster with small distance (fewer hops) to achieve the cluster head. If there are more than one service proposals with the same distance (same number of hops), the vehicle chooses the cluster head with the best fitness (BF). If there are several offers with the same best fitness, the vehicle chooses the cluster head with the smallest ID.

If two cluster heads are in the same coverage directly, the cluster head with the biggest fitness change its state to a cluster member and it becomes a gateway to all its cluster members. It's possible that a gateway loses the link with a cluster but still having the links with cluster members (current gateway children's). In this case, to ensure the data transfer without wasting time to elect and forming new cluster, the gateway at the head of our tree becomes the cluster head of the cluster, and we leave the optimization to the maintenance phase.

Figure 2. Example of solution case study

Algorithm 1: State Definition and Cluster Formation; Algorithm on cv

```
Initialisation
STATE(cv) ← Free; NS(cv) ← Ø; ChP(cv) ← Ø;
Child(cv) ← Ø;
Round of Fitness calculation(Round 1+i) with 1 < i < X + 1
X is the number of repeated rounds as long as
|NS(cv)| = Ø;
Broadcast in its neighboring clusterMes(IDcv, CScv, -, -, -);
t0 ← CT(cv);
On receive clusterMes(IDu, CSu, -, -, -) :
Add (idu, CSu, CT(cv) − t0) to NS(cv);
if |NS(cv)| = Ø then
        F(cv) ← ⊥
else
        $F(cv)=\left[CS(cv)-\frac{\sum_{cm\in NS(cv)} cm.speed}{|NS(cv)|}\right].\left[\frac{\sum_{cm\in NS(cv)} cm.delay}{|NS(cv)|}\right]$
        BF(cv) ← (F(cv), 0, IDcv);
end if
Round of Best fitness spread and state calculation
(Round X+1) with X + 1 < i < Y + 1
Y is the number of repeated round as long as
t < broadcastInterval
Broadcast in its neighboring clusterMes(IDcv, CScv, F(cv), BF(cv), ⊥);
On receive clusterMes(IDu, CSu, BFu, -, -):
BF(cv) ← (F(cv), 0, IDcv);
if BF(u) < BF(cv) then
    BF(cv) ← (BF(u).fitness, BF(u).distance + 1, BF(u).id);
    if BF(u).id = IDu then
        GT(cv) ← IDu;
    else
        GT(cv) ← ⊥;
    end if
end if
/******When i = Y+1******/
if BF(cv) < (F(cv), 0, IDcv) then
    CH(cv) ← (BFcv.distance, BFcv.fitness, BFcv.id, GT(cv);
    STATE(cv) = ClusterMember;
else
    CH(cv) ← (0, IDcv, F(cv), ⊥);
    STATE(cv) = ClusterHead;
end if
```

Algorithm 2: Cluster Formation and Maintenance

```
(Round W+1) with $Y + 1 < i < W + 1$
W is the number of repeated rounds as long as
$t$ < broadcastInterval
Broadcast in its neighboring
clusterMes($ID_{cv}$, $CS_{cv}$, $F(cv)$, $BF(cv)$, $CH(cv)$);
On receive clusterMes($ID_u$, -, -, -, $CH_u$):
if ($CH_{cv}.IDCH = CH_u.IDCH$) then
    if ($CH_u.IDGT = IDCV$) then
        Add ( $Id_u$ ) to Child(cv);
        $STATE(cv) \leftarrow$ Gateway;
    else
        Add ( $Id_u$ ) to Child(cv);
    end if
else
    Add ($ID_u$, -, -, -, $CH(u)$) to $ChP(cv)$;
end if
Cluster Head selection phase
if ($ChP = \emptyset$) then
    if $\{(CH(cv) > Min\{CH_u \mid CH_u \in ChP(cv)\} \vee (CH(cv).IDCH \notin \{CH_u.id \mid CH_u \in ChP(cv)\})\}$
    with $u = Min\,ChP(cv)$ then
        if ($CH_u.IDCH = ID_u$) then
            $CH(cv) \leftarrow (CH_u.distance + 1, CH_u.fitness, CH_u.id, ID_u)$;
        else if ($CH_u.IDCH = ID_u$) then
            $CH(cv) \leftarrow (CH_u.distance + 1, CH_u.fitness, CH_u.id, \perp)$;
        end if
    else if $\{(STATE(cv) = ClusterHead) \wedge (CH_u.distance = 0) \wedge (CH_{cv}.fitness > CH_u.fitness)\}$ then
        $STATE(cv) \leftarrow ClusterMember$;
        $CH(cv) \leftarrow (CH_u.distance + 1, CH_u.fitness, CH_u.id, ID_u)$;
    end if
else if ($Child(cv) = \emptyset$) then
    $CH(cv) \leftarrow (0, ID_{cv}, F(cv), \perp)$;
    $STATE(cv) = ClusterHead$
else
    $STATE(cv) \leftarrow free$; $NS(cv) \leftarrow \emptyset$; $t0 \leftarrow CT(cv)$; $ChP \leftarrow \emptyset$;
```

Figure 3. Our solution algorithm

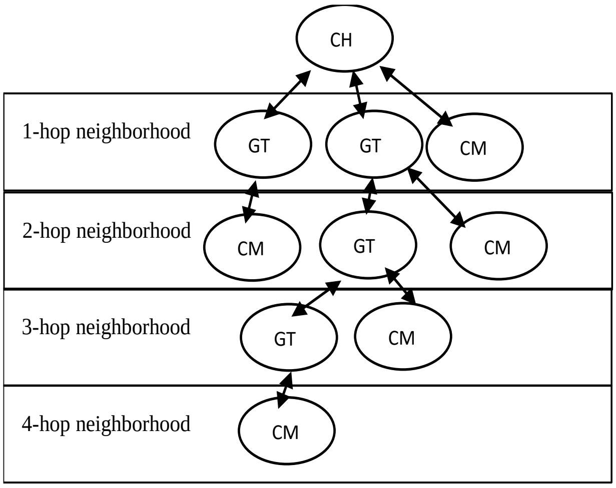

Figure 4. The tree architecture of the cluster

3. THEORETICAL ANALYSIS

3.1. Assumptions

We are assuming:

- Nodes are working in synchronous communication rounds.
- Messages of all vehicles are received at the same time t and processed together.
- In each round, each vehicle can successfully send one message to all its neighbors.
- Any necessary computation can be completed before the next round.

3.2. Communication Complexity

The cluster-message communication complexity is:

- $O(|E|)$
- E is set of communication edges in the network
- $|E|$ is the number of communication edges

As each node sends the message to each other nodes exactly once:

- The actual number of messages is $2|E|$.

4. PERFORMANCE EVALUATION

In this section, we evaluate our solution on Omnet simulator [20], VEINS (vehicles in network simulation) framework [21] and Sumo [22] as a road traffic mobility regenerator and emulator. We also implement the Obstacle Shadowing model proposed by veins framework to have a real scenario measure from the real world [23, 24]. Our simulation results are compared with the two schemes:

- Weighted p-persistence [25]; whose authors propose a distance-based probabilistic diffusion approach as a parameter to determine the rebroadcast probability of vehicles; the limit of this solution, often the furthest nodes still have the greatest probability of transmitting messages.
- Slotted 1-persistence [25]; in this solution, the authors assign vehicles at different time timeslot on their distances from the sender, so that the more distant vehicles receive the shortest delay before the replay.

we use this benchmark to compare our solution, as it is currently the best delay-based delivery method in terms of data delivery, broadcast overhead and delivery delay.

In the rest of this section, we describe our simulation scenario. Then, we define our performance metrics before analyzing our simulation results.

Table 2. Simulation Parameters

Parameters	Values
Band with	10 MHz
Vehicle density	20-100 /KM
Vehicle speed	<80 km/hr
Transmission Range	~330 m
Simulation Runs	10
Simulation Time	800 s
PHY Model	Nakagami-m
Mac Model	IEEE 802.11DCF
Broadcast Interval	0.05 s
CBR packet	512 Bytes
Queue Length	50 packets

4.1. Simulation Scenario

Our simulation scenario is a highway with a length of 5 km. Data messages are exchanged during the simulation between vehicles in both traffic directions. To define the physical layer and the MAC layer, we use the implementation of IEEE 802.11p available in the Veins [21] framework. As shown in the table, we use a bandwidth of 10 MHz, and a bit rate of 6 Mbps at the MAC layer as a default value. The transmission coverage is adapted over a transmission range of around 360 m (suburban environment). The transmission interval is 0.05s and the size of all messages is 100 bytes. For the vehicle density, is framed between 10 and 100 vehicles for each kilometer, which represents a light density increasing over time to reach a high density. For the number of messages, we keep the default value set at 50 messages. The mobility model used in the implementation is the freeway mobility model. The simulation parameters are summarized in Table 2.

4.2. Evaluation of Metrics Performance

The metrics used to evaluate our scheme performance are given below:

- Average transmission delay: the average time an emitting vehicle of beacon wait before it receives all the neighbor’s beacons:

$$D = \frac{\sum_{i=1}^{n} d_i}{|n|} \tag{4}$$

where, d_i is the average end-to-end transmission delay of beacon message of a node and its n neighbors.

- Packet Delivery ratio: the ration of all the packets successfully received to the total sent by a node [28]:

$$PDR = \frac{\sum \text{Number of packet received}}{\sum \text{Number of packet send}} \tag{5}$$

- Control packet overhead is the rate of control messages used by the sender to discover the network including the fast and secure path to send data to all neighbors. This is denoted by O and calculated as follows [29]:

$$O = \frac{\left(\sum_{i=1}^{n} rrq_i + \sum_{i=1}^{n} rrt_i\right) \times cpsize}{\left(\sum_{i=1}^{n} N_i^s \times dpsize_i\right) + \left(\sum_{i=1}^{n} rrq_i + \sum_{i=1}^{n} rrt_i\right) \times cpsize} \times 100\% \tag{6}$$

where, rrq_i is the number of route requests sent by the sender, rrt_i is the number of route request retries done by the sender $cpsize$ is the size of the request packet, $dpsize_i$ is the size of the neighbor data packet in the ith neighbors, N_i^s is the number of data packets sent by the ith neighbor source and n is the number of neighbors.

4.3. Simulation Results and Performance Analysis

Figure 5 illustrates the average transmission delay as a function of network density. For the Slotted 1-persistence scheme, we see that many nodes participate in the transmission by producing a struggle to access the channel. As a result, data collisions easily occur, which will cause more retransmissions and therefore additional transmission delay.

For the weighted p-persistence model and the proposed BDC model, thanks to their architecture based on clustering, the two solutions manage the allocation of bandwidth and concurrent access to the channel. This mechanism can solve the problems of severe channel conflicts and optimize the use of scarce spectral resources, thereby reducing the time required for successful dissemination of a packet between nodes.

Although the 1-slotted 1-persistence model uses a clustering architecture to broadcast messages, it does not take into account the driving directions of vehicles or the size of the clusters when building the cluster network. Comparing the average transmission delay of BDC is the most optimal than that of other protocols. This is due to our protocol's self-stabilization mechanism, as retransmission is carefully assigned to previously chosen nodes, thus combating broadcast storm and packet collisions as well as channel access concurrency.

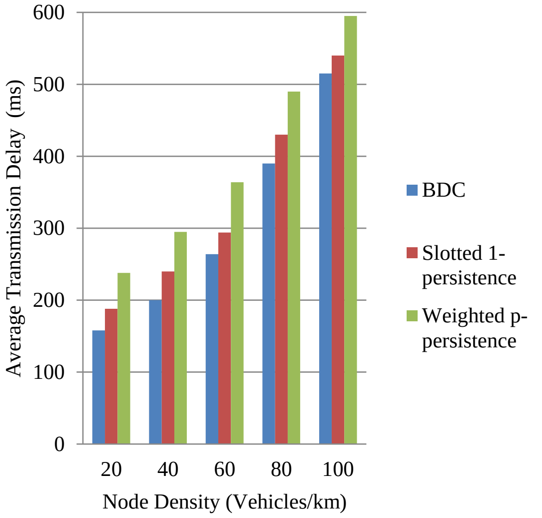

Figure 5. Average transmission delay vs density

As shown in Figure 6, the information coverage is evaluated between three protocols based on vehicle density. We notice in the figure that the information coverage in the network is low at the start of the scenario; this is because there is no stable end-to-end routing path between nodes that can guarantee coordination due to sparse network. With increasing network density, links between vehicles become more stable and information coverage improves. It is concluded that increasing the number of vehicles improves the probability of vehicles to reach each other by ensuring the dissemination links.

However, when the network density reaches a threshold of 60 vehicles/km, information coverage begins to decline. The cause is that in high-density networks, the phenomenon of broadcast storm will become more harmful, thus reducing information coverage. In the weighted p-persistence scheme, the farthest vehicle will forward the packet with the highest probability. However, it is found that the probability of transmission increases linearly with the distance of vehicles in this protocol. So high density causes broadcast storms and will result in more redundant communication.

On the other hand, the slotted 1-persistence model and our proposed BDC model both use a clustering scheme to organize the hierarchical network, to mitigate channel competition, thereby increasing spectrum feasibility and reducing network overload. Compared to the two solutions, BDC also offers an auto-stability algorithm to manage disconnections and ensure the link between vehicles. Auto-stability also allows for adaptation to traffic density and cluster size balancing, providing the best performance compared to other models.

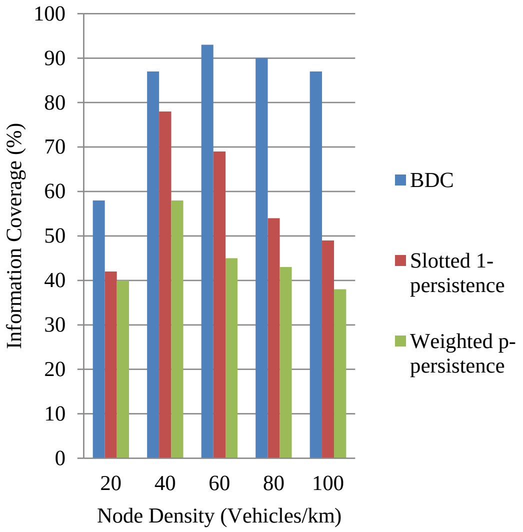

Figure 6. Information coverage vs density

Figure 7 shows the impact of network density on the packet delivery rate. From the Figure 7, we can see that the increase in the number of nodes affects the packet delivery rate. This is because at the start of the simulation the network is sparse with less overlap between nodes. For light densities, the number of nodes will improve network connectivity and reduce lost packets. Once the network density becomes high, channel access competition will lead to more data collisions, causing data transmission failures.

For the p-weighted persistence solution, being the schema is based on a flat network structure and due to the complicated characteristics of the VANET network, the solution is not efficient in terms of exchange between vehicles. Indeed, the absence of a clustering architecture, unmanaged broadcasting can cause serious collisions due to the lack of coordinator. Comparing with the 1-sloted model and our proposed BDC model, both solutions can ensure stable and durable communication between nodes. As a result, good performance is observed compared to the weighted scheme. In the 1-slotted model, many nodes broadcast the packet almost simultaneously. As a result, serious data collisions are observed, thus easily causing data transmission failure. While our BDC model, we can further reduce the risk of data collision by introducing the self-stability mechanism, designating a few nodes in advance to participate in the transmission.

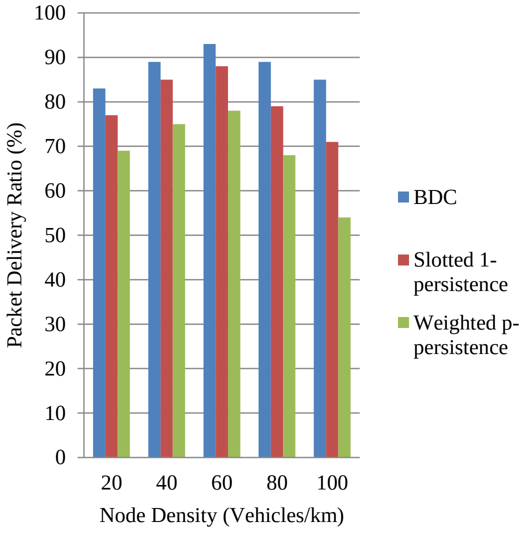

Figure 7. Packet delivery ratio vs density

5. CONCLUSION

In this contribution, we have addressed the problem of data sharing delay, packet loss as well as data availability in the vehicular environment. This scheme is based on a combination of the beacon delay, the speed as well as the distance between vehicles as a metric and the self-stability mechanism. This approach allows us to improve the stability of the clusters and consequently reduce the dissemination time, reduce the packet collision and finally extend the data coverage.

However, there are also some limitations in our work, which can be summarized as follows:

1) For data sharing applications, the cluster head is the central element in the architecture, which can cause additional time of data fusion and aggregation.

2) In this work, we share data for vehicles of all directions, which means that direction is not taking into account to build clusters.

In conclusion, the efficiency of our approach is justified by the optimization of the retransmission for each hop thus minimizing the duplication of messages and collisions.

REFERENCES

[1] O. Tonguz, W. Viriyasitavat, "Cars as roadside units: a self organizing network solution", IEEE Communications Magazine, vol. 51, pp. 112-120, December 2013.

[2] J.B. Kenney, "Dedicated short-range communications (DSRC) standards in the United States", IEEE Trans. Emerg. Sel. Topics Power Electron, vol. 99, no. 7, pp. 1162-1182, Jul. 2011.

[3] E.E. TR, "Intelligent Transport Systems (ITS); Cooperative ITS (C-ITS): Release 1, Technical Report Ver 1.1. 1", European Telecommunications Standards Institute, vol. 5, no. 23, p. 21, 2013.

[4] L. Liu, C. Chen, T. Qiu, M. Zhang, S. Li, B. Zhou, "A data dissemination scheme based on clustering and probabilistic broadcasting in VANETs", Vehicular Communications, no. 13, pp. 78-88, 2018.

[5] R. Akamatsu, M. Suzuki, T. Okamoto, K. Hara, H. Shigeno, "Adaptive delay-based GEOCAST protocol for data dissemination in urban VANET", IEEE 17th International Conference on Mobile Computing and Ubiquitous Networking (ICMU), pp. 141-146, 2014.

[6] P. Salvo, M. De Felice, A. BAIOCCHI, F. Cuomo, I. Rubin, "Timer-based distributed dissemination protocols for VANETs and their interaction with MAC layer", IEEE 77th Vehicular Technology Conference (VTC Spring), pp. 1-6, 2013.

[7] A. Dua, N. Kumar, S. Bawa, "QoS-aware data dissemination for dense urban regions in vehicular ad hoc networks", Mobile Networks and Applications, vol. 20, no. 6, 773-780, 2015.

[8] O.A. Wahab, H. Otrok, A. Mourad, "VANET QoS-OLSR: QoS-based clustering protocol for vehicular ad hoc networks", Computer Communications, vol. 36, no. 13, pp. 1422-1435, 2013.

[9] N. Wisitpongphan, O. K. Tonguz, J.S. Parikh, P. Mudalige, F. Bai, V. Sadekar, "Broadcast storm mitigation techniques in vehicular ad hoc networks", IEEE Wireless Communications, vol. 14, no. 6, pp. 84-94, 2007.

[10] S. Panichpapiboon, L. Cheng, "Irresponsible forwarding under real intervehicle spacing distributions", IEEE Transactions on Vehicular Technology, vol. 62, no. 5, pp. 2264-2272, 2013.

[11] S. Panichpapiboon, L. Cheng, "Irresponsible forwarding under real intervehicle spacing distributions", IEEE Transactions on Vehicular Technology, vol. 62, no. 5, pp. 2264-2272, 2013.

[12] A. Mostafa, A.M. Vegni, D.P. Agrawal, "A probabilistic routing by using multi-hop retransmission forecast with packet collision-aware constraints in vehicular networks", Ad Hoc Networks, vol. 14, pp. 118-129, 2014.

[13] R.S. Schwartz, R.R. Barbosa, N. Meratnia, G. Heijenk, H.A. Scholten, "A directional data dissemination protocol for vehicular environments", Computer Communications, vol. 34, no. 17, pp. 2057-2071, 2011.

[14] C. Liu, C. Chigan, "RPB-MD: Providing robust message dissemination for vehicular ad hoc networks", Ad Hoc Networks, vol. 10, no. 3, pp. 497-511, 2012.

[15] A. Mondal, S. Mitra, "Secure data dissemination in VANET-a pull based approach", IEEE International Conference on Communication, Networks and Satellite (COMNESTAT), pp. 60-67, 2015.

[16] S. Bai, Z. Huang, D. Kwak, S. Lee, H. Oh, J. Jung, "Vehicular multi-hop broadcasting protocol for safety message dissemination in VANETs", IEEE 70th Vehicular Technology Conference Fall, pp. 1-5, 2009.

[17] L. Zhang, D. Gao, S. Gao, V.C. Leung, "Smartgeocast: Dynamic abnormal traffic information dissemination to multiple regions in VANET", IEEE 9th International Wireless Communications and Mobile Computing Conference (IWCMC), pp. 1750-1755, 2013.

[18] J.P. Singh, R.S. Bali, "A hybrid backbone based clustering algorithm for vehicular ad-hoc networks", Procedia Computer Science, vol. 46, pp. 1005-1013, 2015.

[19] D. Jin, F. Shi, J. Song, "Cluster based emergency message dissemination scheme for vehicular ad hoc networks", The 9th International Conference on Ubiquitous Information Management and Communication, pp. 1-8, 2015.

[20] S.G. Gupta, M.M. Ghonge, P.D. Thakare, P.M. Jawandhiya, "Open-source network simulation tools: An overview", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 2, no. 4, pp. 1629-1635, 2013.

[21] C. Sommer, R. German, F. Dressler, "Bidirectionally coupled network and road traffic simulation for improved IVC analysis", IEEE Transactions on Mobile Computing, vol. 10, no. 1, pp. 3-15, 2010.

[22] M. Behrisch, L. Bieker, J. Erdmann, D. Krajzewicz, "SUMO-simulation of urban mobility: an overview", The Third International Conference on Advances in System Simulation, ThinkMind (SIMUL 2011), 2011.

[23] C. Sommer, D. Eckhoff, R. German, F. Dressler, "A computationally inexpensive empirical model of IEEE 802.11 p radio shadowing in urban environments", IEEE Eighth International Conference on Wireless on-Demand Network Systems and Services, pp. 84-90, 2011.

[24] C. Sommer, D. Eckhoff, F. Dressler, "IVC in cities: Signal attenuation by buildings and how parked cars can improve the situation", IEEE Transactions on Mobile Computing, vol. 13, no. 8, pp. 1733-1745, 2013.

[25] N. Wisitpongphan, O.K. Tonguz, J.S. Parikh, P. Mudalige, F. Bai, V. Sadekar, "Broadcast storm mitigation techniques in vehicular ad hoc networks", IEEE Wireless Communications, vol. 14, no. 6, pp. 84-94, 2007.

[26] S. Ouahou, S. Bah, Z. Bakkoury, A. Hafid, "Dynamic clustering algorithm based on beacon delay", IEEE 11th International Conference on Intelligent Systems: Theories & Applications (SITA), pp. 1-6, 2016.

[27] S. Ouahou, S. Bah, Z. Bakkoury, A. Hafid, "Multi-hop Clustering Solution Based on Beacon Delay for Vehicular Ad-Hoc Networks", International Symposium on Ubiquitous Networking, pp. 357-367, Springer, Cham, 2017.

[28] N. Jiang, Y.Z. He, L. Guo, "Heuristic algorithm based traffic classified routing algorithm in distributed satellite cluster network", Applied Mechanics and Materials, vol. 738, pp. 1150-1159, Trans Tech Publications Ltd, 2015.

[29] J. Sucec, I. Marsic, "Hierarchical routing overhead in mobile ad hoc networks", IEEE Transactions on Mobile Computing, vol. 3, no. 1, pp. 46-56, 2004.

[30] N.B. Karimi, S.N. Razavi, H.S. Aghdasi, "Distributed Clustering in Wireless Sensor Networks Using a Game Theoretical Approach", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 18, vol. 6, no. 1, pp. 1-8, March 2014.

[31] M.A. Balafar, "Improving Clustering Results using Re-Evaluation of Boundary Data", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 14, vol. 5, no. 1, pp. 103-107, March 2013.

BIOGRAPHIES

Soufiane Ouahou holds a master’s degree in computer science at Hassan 1st University, Settat, Morocco. Currently, he is a data scientist and technical lead on behalf of Aeorospatial Clients, Toulouse, France. He has participated in international conferences and workshops where he has published research works around vehicular networks. He has also carried out research stays in research laboratories such as the LABRI laboratory in Brodeaux, France.

Slimane Bah is a professor at the Computer Science Department, Mohammadia Engineering School, Rabat, Morocco. He has an experience as adjunct professor at the Department of Computer Science, University of Moncton, Moncton, Canada. He holds a Ph.D. in computer networks from the Electrical and Computer Engineering Department of Concordia University, Montreal, Canada. During his Ph.D. research he was an intern at Ericsson Canada. He also holds an M.Sc. in Computer Networks from University of Montreal (Université de Montréal) and an engineering degree in computer science - Networking, from National School of Computer Science and Systems Analysis (ENSIAS), Morocco. His research interests include end-user services, self-organizing and challenging networks, services and protocols engineering. He is also a technical committee member for several international conferences and journals (SITA, NETYS, AFRICATEK, MEDCT, Elsevier Journal).

Zohra Bakkoury holds a Ph.D. in Engineering from Exeter University in UK and a degree of engineer in computer science from Mohammadia Engineering School (EMI), Rabat, Morocco. She is a Full Professor at the Computer Science Department of EMI and founder and director of AMIPS research group and Head of MASI research Laboratory in EMI. Her current research interests are focused on business process management, process mining, data mining and intelligent systems. She has published book chapters beside numerous articles in scientific journals and conferences. She was the co-chair of the eight and 12th SITA conference on Intelligent Systems in 2013 and 2018.