Fair Self-Adaptive Clustering for Hybrid Cellular-Vehicular Networks

Julian Garbiso, Ada Diaconescu, Marceau Coupechoux and Bertrand Leroy

***Abstract*—Due to the increasing number of car-centered connected services, making efficient use of limited radio resources is critical in vehicular communications. Hybrid vehicular networks dispose of multiple Radio Access Technologies (RATs) like cellular and vehicle-to-vehicle (V2V) networks, with complementary characteristics that allow for developing smarter network traffic distribution methods. This paper proposes a self-adaptive clustering system for ensuring a suitable trade-off between data aggregation (over the cellular network) and communication congestion due to cluster management (within the V2V network). The systems algorithms use a distributive justice approach for selecting cluster heads, to improve fairness among car drivers and hence help the social acceptability of self-adaptive clustering. Simulation results show that this approach significantly improves fairness over time without affecting network performance. This solution can thus optimize the usage of radio resources, reducing cellular access costs, without the need for uniformization among different mobile operators access plans.**

***Index Terms*—ITS, Connected Vehicles, Hybrid Vehicular Networks, LTE, V2V, Distributive Justice, Clustering.**

I. Introduction

VEHICULAR connectivity has a major role in the development of smart cities, making travelling safer and providing real-time information that allows for an optimized traffic management, potentially reducing travel times, waiting times and the ecological footprint [1]. Every vehicle becomes a source of valuable information, while it takes advantage of the established communication to provide passengers with new infotainment services.

Several Radio Access Technologies (RATs) have been studied for vehicular communications. The IEEE 802.11p standard [2] is used in the IEEE 1609 family (WAVE) adopted in the US, and in the European ETSI ITS-G5. Future cellular networks will also be able to satisfy most latency and throughput requirements for vehicular applications [3]. However, the radio spectrum used for these communications is licensed to mobile network operators, hence its usage is paid. To balance the equation, compared to license-free alternatives like IEEE 802.11p, cellular communications offer better quality of service and reliability. Multiple economic models [4] can be considered for the payment of the cellular access fees, but they all eventually impact the consumer's budget. Conversely, protocols for V2V communication such as IEEE 802.11p are free to use, yet they do not necessarily provide access to the Internet without a substantial investment in deployment of Road Side Units, whereas such infrastructure is already available in cellular networks. Additionally, IEEE 802.11p is prone to congestion, and an increasing number of access demands can quickly lead to significant packet losses due to collisions [5], [6].

J. Garbiso and B. Leroy are with Institut Vedecom, 78000 Versailles, France. E-mail: { name.surname@vedecom.fr }.

A. Diaconescu and M. Coupechoux are with Telecom ParisTech, 75013 Paris, France. E-mail: { name.surname@telecom-paristech.fr }.

Manuscript received Month DD, YYYY; revised Month DD, YYYY.

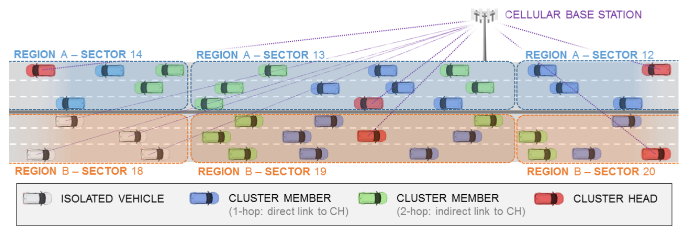


Fig. 1: Illustration of the proposed clustering system model.

The presented work aims to capitalize on each technology's advantages in the context of a hybrid vehicular network, in which each vehicle is equipped with access to both V2V and cellular networks. The proposed solution consists in partitioning the set of vehicles in a region covered by a cellular base station into clusters. In this system, only Cluster Heads (CH) can access the cellular network, while other Cluster Members (CM) communicate with the CH using IEEE 802.11p. This architecture decreases the communication costs thanks to data aggregation at the CHs. This approach aims to exploit the high local redundancy of position-based information, which is common to several services for connected vehicles (such as uploading Floating Car Data [7] for enhancing traffic management, or downloading maps). The CHs aggregate information transmitted by the vehicles in the cluster, leading to significant reductions in data volumes transferred through the cellular network.

New issues arise from this approach. First, IEEE 802.11p uses broadcast transmissions for position-based services. This induces network congestion and packet losses when clusters are large. On the other side, the larger the cluster, the more efficient the data aggregation. Secondly, at any given moment, only the CH bears the cost of cellular access. In a voluntary participation system where free-riders can appear, new solutions should be developed to ensure a *fair* distribution of the cellular usage.

To summarize, this paper proposes a fair self-adaptive and socially acceptable clustering system (Figure 1) that compromises between the communication costs of the cellular network and the performance of the V2V network, at runtime, depending on traffic conditions. The systems algorithms ensure an equitable distribution of communication costs amongst drivers, over the long term. This helps the social acceptability of the presented solution. Additionally, the proposed system does not need any further types of billing nor provider interoperability requirements.

In previous works we have presented different versions of the system's cluster head selection method and cluster-resizing procedure [6], [8], [4]. This paper presents a unified approach for the clustering system, and the complete set of algorithms, including the previously unpublished Fair Self-Adaptive (FSA) algorithm and its associated voting algorithm (an idea previously analyzed in [4] without algorithmic specifications). The presented approach could potentially be applied in other multi-RAT networks, or in traffic control problems with other resources to distribute among drivers.

This paper is organized as follows: Section II is a bibliography study. Section III formulates the problem and motivation, Section IV describes the proposed architecture and algorithms, and Section V evaluates their performance. Finally, we draw our conclusions and examine perspectives for future work in Section VI.

II. Related Work

Position-based services over vehicular networks have a great level of redundancy when considering vehicles individually. The only way to make a reasonable use of the available network resources is by eliminating redundancy by grouping vehicles by proximity. Clustering techniques have been used for a long time in network theory [9], and have since found new applications in wireless ad-hoc networks and, in particular, in vehicular networks [10]. Many early proposals focused on improving metrics that are intrinsic to clustering (e.g. maximizing cluster lifetime, minimizing cluster head changes), leaving aside the specificities of the application domain. More recently, new algorithms have started to focus on adapting clustering techniques to the particular characteristics of vehicular networks [11]. In this paper, we focus on *multi-hop* clustering algorithms [12], [13], [14] as they allow for building larger clusters by extending the communication range through packet forwarding, at the expense of an increased usage of the V2V network that can lead to the loss of communication packets if not handled properly.

Some of these algorithms were proposed to address the specificities of *hybrid vehicular networks* [15], [16], [17]. In most of them, cluster heads are selected in a decentralised manner based on observations of local metrics, often leading to an *overpopulation of CHs*. This phenomenon has an undesirable side effect in multi-hop clusters: When the number of hops is increased, the PLR increases naturally, but there is no improvement in data aggregation performance [6]. This is because vehicles try to join their nearest cluster head, and there are often many CHs in the communication range. A novelty of the work presented in this paper is that it has been designed to avoid CH overpopulation by delegating CH selection to the cellular base station, which leads to an improvement in cluster sizes, leading to better and more efficient information aggregation, which in turn leads to a more efficient network usage and cost reductions. Moreover, the clustering algorithms presented in this paper are self-adaptive: the number of hops changes dynamically in order to adapt to the traffic density, maximizing information aggregation while keeping V2V packet loss under control. To the best of our knowledge, no other work has taken this compromising approach for balancing performance.

Closer to our approach, Rémy et al. [18] delegate the entire cluster formation process to the cellular base station. The improved efficiency in cluster dimensioning comes in this case at the price of severely increasing cellular traffic and access cost. The impact of PLR is not assessed. Moreover, a critical issue is not solved by the authors: At any given moment in the proposed algorithms, only the CHs bear the cost of accessing the cellular network. However, for the system to be socially acceptable and be widely adopted in a model where free-riding is possible, fairness over time should be ensured. Therefore, in this paper, we propose a CH selection rule based on a theory of distributive justice, which provides a powerful tool for social acceptability of the system.

In the field of distributive justice, different authors identify various norms [19], [20], [21] for addressing competing claims. Rescher [20] has a vision of *social justice* that consists in determining each individual's legitimate claims, and treating all the legitimate claims of the population equally. This theory does not focus on the origins of those claims, but on how far an individual's claims should be met by evaluating competing claims and the limitations of the available resources. This theory has been adopted by [22] for linear public good games. We have followed Rescher's approach as well for distributing the costs inherent to the role of Cluster Head. To the best of our knowledge, there are no previous studies of the application of distributive justice theory on VANETs. The proposed approach leads to a significant improvement in several fairness metrics, without affecting network performance, and without the need of establishing any interoperability between mobile operators or new separate billing systems.

This article presents a fair and self-adaptive clustering system in hybrid cellular-vehicular networks. This system has been designed building up on the conclusions of previous works, concerning cluster head selection methods [6], hop number adaptation [8] and fairness in vehicular communications [4], reuniting these ideas in a comprehensive formalization. This paper takes these preliminary works further providing: i) a formal representation of the Fair Self-Adaptive (FSA) cluster head selection algorithm and its associated voting algorithm, which is essential for the algorithms reusability and results replicability; ii) further evaluation metrics and results providing more solid support for the viability of the proposed algorithm; iii) a comprehensive system formalization, including the data structures, state machines and operation between the different parts.

We can summarize the contributions of the paper to the state of the art as follows:

- New clustering problem formulation in hybrid vehicular networks: this paper is the first one to address the trade-off between data aggregation and network congestion.
- A new system to address this problem (sections IV-A and B) made of (i) a new way of segmenting the road space

into regions and sectors under the control of base stations, (ii) a cluster formation algorithm together with finite state machine for vehicles and the related data structures, (iii) a new definition of CAM messages with enhanced fields.

A self-adaptive (SA) clustering algorithm that dynamically adapts the maximum number of hops to the vehicle density. SA is the proposal that addresses the data aggregation vs. network congestion trade-off.

A fair self-adaptive (FSA) clustering algorithm that extends the previous algorithm by including a fair mechanism for selecting cluster heads. FSA is based on the theory of distributed justice of Rescher [20]. To the best of our knowledge, this is the first time that this theory is applied to vehicular networks.

Extensive simulation results using realistic and well-established models and simulator (SUMO [23]). They show the effectiveness of the proposed algorithms compared to the state-of-the-art (taking the VMaSC clustering algorithm [13] as a representative example).

III. Problem Formulation and Motivation

A. Problem Formulation

We consider a highway section of fixed length L_r, consisting of L lanes. A new vehicle arrives at each lane every T seconds (at the beginning of the highway section). These vehicles, which form a vehicular network, traverse the highway section at constant speed and leave the network afterwards. We assume that a single cellular base station (BS) covers the entire highway section. Every vehicle has a V2V and a cellular interfaces, and has to send information to the BS at a rate of packets/s[1].

In this vehicular network, a clustering algorithm is assumed to be implemented. This means that the packets generated by a vehicle can either be sent to the BS either directly or indirectly via a CH. An isolated vehicle (i.e., belonging to a cluster of size 1) necessarily sends its packets directly through the cellular network. A vehicle in a cluster c of size $N_c > 1$ sends its packets to the CH through the V2V interface. The CH aggregates the information received through the V2V link from CMs and sends it to the BS. The network traffic sent by the cluster c to the BS is then $(N_c)N_c$, where $(N_c) \quad 1$ is a compression function performed by the CH that may be a decreasing function of N_c. Without loss of generality, we can assume that $(1) = 1$. This function results from the inherent redundancy of position-based data. We do not define an explicit expression of this function in this paper as it may depend on the specific application and required accuracy. For instance, the CH can upload the average value of the CMs measurements, in which case $(N_c) = 1/N_c$. We define a *cluster partition* as a set of non-overlapping clusters that includes all the vehicles within the network. In the following, we consider only cluster partitions with clusters having a maximum of H hops between any CM and its CH. As a consequence, the total traffic generated by the vehicular network on the uplink of the cellular network for a given cluster partition C can be written as:

$$(C) = \sum_{c \in C} (N_c)N_c \ ; \tag{1}$$

where C is the set of all clusters, and N_c is the number of vehicles in cluster c. We denote $N = \sum_c N_c$ the total number of vehicles in the network. We now define the global compression ratio of the clustering partition C as:

$$(C) \triangleq 1 \quad \frac{(C)}{N} = 1 \quad \frac{\sum_{c \in C} (N_c)N_c}{N} \tag{2}$$

Note that is also the average compression ratio. For the specific example, where $(N_c) = 1/N_c$, we see that it is advantageous to have large clusters or even a single cluster to maximize the compression ratio. Now, for a cluster partition C and the considered traffic model, we can compute a Packet Loss Rate $PLR(C; \ )$, which is a function of the cluster partition and the amount of traffic. We define the Packet Loss Rate in the V2V network as the ratio between lost packets (due to collisions or decoding failure when the signal is too weak) and correctly received and decoded packets. We will see in Section V that the PLR is increasing with the cluster size because more users are contending for the channel using the contention-based medium access V2V protocol. Our problem is thus for a given traffic condition to maximize the average compression ratio under the constraint of an acceptable packet loss rate:

$$\max_C \ (C) \tag{3}$$

$$\text{s.t. } PLR(C; \ ) \quad PLR_{max}; \tag{4}$$

where PLR_{max} is an application-specific constraint.

In this problem formulation, the difficulty lies in the fact that the PLR is only implicitly defined. To the best of our knowledge, there is no closed-form expression of the PLR for a dynamic multi-hop cluster of moving vehicles communicating with IEEE 802.11p. Hence, the PLR can only be obtained by simulations or measurements. In this case, heuristic methods have been traditionally employed as viable alternatives to optimality, see e.g. [13] or [18] in our context. From a combinatorial point of view, our problem is similar to the Minimum d-hop dominating set problem in a graph [24] if some simplifying assumptions are made in the model. This problem consists in finding a set S of vertices of minimum cardinality such that any vertex of the graph is either in S or at a distance at most d from S. However, this problem is known to be NP-complete even in planar unit disk graphs [24].

B. Example Application: FCD Aggregation

For the performance evaluation of the algorithms (Section V) we use the example application of Floating Car Data aggregation, where vehicles need to send periodic updates on their position, speed and heading to a server (e.g. Google Maps' trip history, Waze). By considering position and speed within a cluster to be similar enough for these services, we

[1]For the given example application (constant uploading of Floating Car Data), every vehicle will be constantly sending its position information through the Base Station to a distant server. This is exactly what very well-known mobile map applications do when collecting information about an user's daily trips.

consider $(N_c) = 1=N_c$: the CH aggregates the information of all its members (received on V2V) and sends a single average value through the cellular network. This approach can be used for several applications from map downloading to video streaming (uplink and downlink).

IV. Proposed System and Algorithms

A. Architecture and Delegated CH Selection

We assume that every vehicle is equipped with two radio transceivers: one for *V2V* communications (e.g. IEEE 802.11p), and another one for accessing the *cellular* network (e.g. 4G LTE)[2]. Every circulation direction of a road in a cellular BS's coverage area becomes a **clustering region**. The region is divided into **clustering sectors**. The sector length is calculated as the product of the estimated V2V communication range (fixed input) times the maximum number of hops (which changes dynamically in our algorithm). Hence, each sector has the approximate size of a cluster. The CH selection is delegated to the BS. At regular intervals of T seconds, it verifies the presence of a CH in each sector. If there is none, it selects a new CH as determined by the **CH Selection Algorithm** (see Figure 1). In this work we study two variants: the *Self-Adaptive (SA)* algorithm (Section IV-C), which appoints the vehicle that is closest to the center of the sector as CH, and the *Fair Self-Adaptive (FSA)* algorithm (Section IV-D), which analyzes several criteria of distributive justice. The rest of the cluster formation process takes place locally through the V2V ad-hoc network, as described in the next section.

B. Cluster Formation

1) Vehicle State Diagram: Figure 2 depicts the states of a vehicle and the transitions proposed in our model. A new vehicle enters the **Discovery** state, and listens to the advertisement messages to find a CH at the smallest possible number of hops. It also starts building its *Neighbour Information Table* (**NIT**). The Discovery state terminates when a timer expires. The default transition is to the **Isolated** state. If *CH Advertisements* were received, the vehicle attempts to join the CH at the minimum hop distance (or physical distance in case of a tie). A vehicle in the **Isolated** state immediately tries to join the first CH that it detects. If it succeeds, it transitions to the **Cluster Member** state. Any vehicle in the **Isolated** or **Cluster Member** state immediately transitions to the **Cluster Head** state upon appointment message from the Base Station. There are two conditions under which a vehicle in the **Cluster Head** state can transition back to the **Isolated** state: (1) Absence of cluster members during a specified time period; or (2) Reception of a request to cease from the base station (triggered when two or more CHs are present in the same sector).

[2]If we take a look at the European standards for ITS communications, the ETSI ITS-G5 standard implies the existence of a V2V connection, while the eCall directive implies that there is a mandatory cellular connection. Besides, different manufacturers already have incorporated cellular-based services in their cars, and it is becoming a mainstream feature.

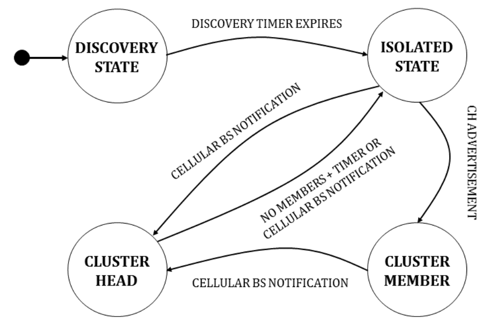


Fig. 2: Proposed vehicle state diagram.

2) Enhanced Cooperative Awareness Messages (CAMs): The main source of control information overhead in clustering algorithms is often the exchange of information between members during the cluster formation process. On the other hand, ITS standards such as ETSI ITS-G5 propose Cooperative Awareness (CA) services to enhance security. These services rely on a periodical, frequent broadcasting of beacons in the V2V network. Coincidentally, most of the information usually required for cluster formation (like sender ID, position, speed and timestamp) is available in these Cooperative Awareness Messages (CAM). For the implementation of our approach, we want to avoid unnecessary redundant signalling and take advantage of the standardized message exchange. We thus propose to add a few fields in the CAM body in order to have all the necessary information for implementing multi-hop forwarding (see Figure 3). With this minimum increment of around 3% of the CAM payload, the only control overhead of the proposed algorithm is limited to occasional *CH Advertisement*, *Join Request* and *Join Response* messages.

3) Cluster Head Discovery and Join Procedure: When a vehicle receives a notification from the BS to become CH, it starts broadcasting *CH Advertisement* messages regularly. Even though the CH selection is centralized within the coverage area of a base station, the rest of the cluster formation process is completely decentralized. After listening for *CH Advertisement* messages, a vehicle that decides to join a specific cluster sends an unicast *Join Request* message to the CH (through the specified intermediate nodes if it is a multi-hop communication), and the CH completes the handshake process with a *Join Response* unicast message following the same path.

4) The Neighbour Information Table (NIT): Every vehicle creates and maintains its own *Neighbour Information Table* (**NIT**) with the information received from *CH Advertisements* and multi-hop *CAM*s. It can be compared with the upper layer of the *Local Dynamic Map* [26]. For each neighbour detected in a specific time frame (*freshness threshold*), the following information is stored in the NIT: neighbour ID, parent CH ID, timestamp of the last message, direction, speed, state, hop distance, and the IDs of the intermediary hops if any.

C. Self-Adaptive (SA) Clustering Algorithm

The most common approach for CH selection in the literature is called *CH self-appointment*, in which every vehicle

	Extended Payload					
CAM Payload Sender ID, Position, Timestamp, Heading, Speed, ...	**State** (CH, CM, ...)	**Hop count**	**2nd hop** ID	**2nd hop** State	**3rd hop** ID	**3rd hop** State
100 bytes	*2 bits*	*2 bits*	*1 byte*	*2 bits*	*1 byte*	*2 bits*

Fig. 3: Enhanced CAM packet including extended payload for clustering information [25].

decides individually whether or not it becomes CH by observing specific metrics of its vicinity (like distance or relative speed). After having tested this approach by implementing a representative example [13], we observed that an important flaw of CH self-appointment is that it often generates too many CHs. This hinders the possibility of increasing the number of hops for increasing cluster size. To address these issues, we propose a BS-based CH Selection approach that has proven to be useful to overcome this issue (see Algorithm 1): If there is no CH in a sector, the BS selects a vehicle that is the closest to the sector center. In this algorithm, the number of hops H is an input.

Algorithm 1 Self-Adaptive CH Selection Algorithm (Base Station)

Initialisation:
Set maintenance period T and adaptation period $T_{adaptation} = mT$, $m \in \mathbb{N}\setminus\{0\}$.
Set IEEE 802.11p radio range R.
Set clustering region length L_r.
Set default values, upper and lower bounds for maximum number of hops $H_{default}$, H_u, H_l.
Set the Predictive Analysis Zone as the first L_{PAZ} = **min**($\frac{L_r}{4}$, $4 \cdot R$) meters of the clustering region.
Set triggering thresholds $\rho_{k\ hop_{min}}$ and $\rho_{k\ hop_{max}}$ for $k = H_l, H_l + 1, \ldots, H_u$ hops.
For $t = nT$, $n = 1, 2, \ldots$, **do**
  If $n == 0 \mod m$ **then**
    $H \leftarrow$ Algorithm 2
    Compute the clustering diameter $D = 2R \cdot H$.
    Divide the highway section into $S = L_r/D$
    sectors.
  EndIf
  For $s = 1, 2, \ldots, S$, **do**
    If there is no CH in s **then**
    Select as CH the vehicle that is the closest
    to the center of s.
    Endif
  Endfor
Endfor

Now, every BS dynamically adapts the number of hops H in its coverage area to the observed vehicular density following Algorithm 2. The decision to increase or decrease the number of hops is taken by comparing the observed vehicular density with two thresholds (for each possible current number of hops), $\rho_{H\ hop_{min}}$ and $\rho_{H\ hop_{max}}$ that have a hysteresis margin to avoid instability.

Vehicle density can vary quickly, creating a heterogeneous distribution within a clustering region. This can lead to a performance degradation if the adaptation of the number of hops is not sufficiently fast. We define the **Predictive Analysis Zone** (PAZ) of a clustering region in Algorithm 2 as the first portion of the region, measuring in the sense of traffic. The decision to increase or decrease the number of hops is made by comparing the vehicular density in the PAZ (ρ_{PAZ}) and in its complement ($\rho_{\overline{PAZ}}$), and then comparing the maximum of those values with the aforementioned thresholds. This approach takes advantage of the particular movement pattern of vehicles in order to anticipate density changes, avoiding performance degradation due to sudden changes.

Algorithm 2 Hop adaptation algorithm (Base Station)

If $t == 0$ **then** Return $H_{default}$
Compute vehicular density $\rho = \max\{\rho_{PAZ}, \rho_{\overline{PAZ}}\}$.
If $\rho < \rho_{H\ hop_{min}}$ **then**
  $H \leftarrow \min\{H + 1, H_u\}$
Endif
If $\rho > \rho_{H\ hop_{max}}$ **then**
  $H \leftarrow \max\{H - 1, H_l\}$
Endif
Notify all vehicles in the clustering region
Return H.

D. *Fair Self-Adaptive (FSA) Clustering Algorithm*

At any given moment, only the CH bears the economic cost of using the cellular network. In order for the system to be socially acceptable, the system must ensure that a fair distribution of costs is achieved over time. The Fair Self-Adaptive clustering algorithm uses the same hop adaptation mechanism as the Fairness-agnostic approach, but it proposes a new CH selection methodology based on a *theory of distributive justice*. Our approach can be similarly applicable to other clustering algorithms.

1) Theory of Distributive Justice for CH Selection: We have chosen to base our method on Rescher's [20] vision of *social justice*, which aims to determine the different legitimate claims that participants may have, treating those claims impartially. Instead of analyzing their causes, this theory focuses on delivering *distributive justice*: How far an individual's claims should be met, taking into account other people's claims and the limited resources. In his work, Rescher surveys the way other authors have analyzed fairness and extract seven so called *canons of distributive justice* (or fairness requirements): 1) to be treated as equals, 2) according to one's needs, 3) to one's productive contributions, 4) to efforts and sacrifices, 5) to the social value of the services provided by the individual to the society, 6) to supply and demand, or 7) to merits and achievements. Rescher's proposal states that none of these canons, taken individually, could grant distributive justice. Instead, he proposes to analyze every participant's legitimate claims following each of these aspects, and focuses on how to balance them in case of conflict.

In order to apply Rescher's approach to our CH selection algorithm, we have mapped the seven *canons* of justice to our problem's variables. Vehicles in a clustering sector where a CH selection takes place vote for their preferred canon, perceived as the most convenient one according to their situation. The cellular BS takes this vote into account as a weighting factor for each of the *canons*, and tries to make the best choice for selecting the new CH in terms of fairness and performance.

We now show how Rescher's *canons* of distributive justice are mapped in the terms of our problem, which is to distribute the cost of the cellular access. The seventh canon (*merits and achievements*) rewards extraordinary actions that are not possible in our context; any form of contribution is always taken into account by other canons. Hence, we do not include this canon in our mapping. The fourth (*efforts*) and fifth (*social utility*) canons are merged into a single one because all efforts contribute to the social utility. This results in the five following canons. Each one is associated to a preference value that is used by vehicles to rank the canons and vote for their preferred one. Time is divided into billing periods of duration t_b and we define t_e the elapsed time of the current billing period. The billing period corresponds to the validity period of a cellular data quota Q_b (for example $Q_b = 30$ Gbits and $t_b = 1$ month). At time t_e, the remaining quota is denoted Q_e. In the following, all preference values are computed at t_e.

- Canon of equality: The preference value for this canon is a normalized amount of data sent as CM:

$$E_{d_{CM}} = \left[1 - \frac{d_{CM}}{t_e/t_b \cdot M_{d_{CM}}}\right]_0^1, \tag{5}$$

where d_{CM} is the data sent as a CM at t_e during the current billing period, $M_{d_{CM}}$ is the historic maximum value of sent data during any billing period and $[x]_0^1 \triangleq \max\{\min\{x, 1\}, 0\}$. The parameter $M_{d_{CM}}$ allows us to compare the current data consumption with respect to the history of this vehicle. The preference value $E_{d_{CM}}$ measures how extraordinary is the data consumption at t_e. With this definition, the higher the amount of data sent as CM, the higher the probability of being selected as CH.

- Canon of needs: The preference value is defined as the normalized remaining quota:

$$E_{Q_e} = \left[1 - \frac{Q_e}{Q_b}\right]_0^1. \tag{6}$$

The higher the amount of available cellular quota, the higher the probability of being selected as CH.

- Canon of productivity: The preference value is defined as the normalized amount of cellular data sent as CH:

$$E_{d_{CH}} = \left[\frac{d_{CH}}{t_e/t_b \cdot M_{d_{CH}}}\right]_0^1, \tag{7}$$

where d_{CH} is the amount of cellular data sent as CH in the current billing period and $M_{d_{CH}}$ is the historic maximum value of d_{CH} in any billing period. The higher the volume of data sent as CH, the lower the chance to be CH.

- Canon of effort/social utility: The preference value is defined as the normalized number of times having served as CH:

$$E_{n_{CH}} = \left[\frac{n_{CH}}{t_e/t_b \cdot M_{n_{CH}}}\right]_0^1, \tag{8}$$

where n_{CH} is the number of times the vehicle has been selected as CH in the current billing period and $M_{n_{CH}}$ is the historic maximum value of n_{CH} in any billing period. The more times a vehicle has been CH, the less likely it is to be selected again.

- Canon of supply and demand: Following [22], we interpret this canon as a measure of the compliance to the rules. An example of non-compliance is switching off the cellular connection while being CH. The preference value is defined as the normalized number of non-compliance events:

$$E_{n_{NC}} = \left[1 - \frac{n_{NC}}{t_e/t_b \cdot M_{n_{NC}}}\right]_0^1, \tag{9}$$

where n_{NC} is the number of non-compliance events and $M_{n_{NC}}$ is its historic maximum value of n_{NC} in any billing period. The more non-compliance events are detected, the more likely it is to be selected as CH.

Each vehicle uses the five preference values above to know which canon it will vote for. All the variables required to build the preference values are assumed to be known by the BS for all vehicles in a sector.

2) Algorithm: The BS runs Algorithm 3 for every sector s when there is no CH in s at time t_e (step 4). For every vehicle i, it retrieves the variables required for evaluating the canons and $\mathbf{v}_i$ (step 6). This vector is a vote to elect the preferred canon (see Algorithm 4). A weight is allocated to every canon. The weight is proportional to the number of votes received for that canon (step 11). They are used in a weighted Borda vote [27] where each canon acts as a Borda-type voter. Each of the five lists built in steps 15 to 19 becomes a prioritized voting ballot: for a list of n elements, the first vehicle in the list receives n points, the second one receives $n - 1$, and so on (step 21). The resulting score $_i$ for each vehicle is the sum of its five scores, weighed by the factors calculated in the distributed vote (step 22).

We also propose to take into account the distance to the center of the sector to balance fairness and efficiency. Distances of the vehicle to the center of s are retrieved (step 7) and normalized (step 10). A centrality score $_i$ is associated to every vehicle (step 24). The final score balances the Borda and the centrality scores via a parameter (step 26). The vehicle with the best score is selected as CH (step 27-28).

3) Computational complexity: The most computationally intensive part of the process is the ordering of the five lists, which can be considered to be *O(n log n)*, where *n* is the number of vehicles in the clustering sector. This cost is also alleviated by the distributed and scalable nature of the

Algorithm 3 Fair Self-Adaptive CH Selection Algorithm (run at Base Station in every sector s)

Inputs: : a parameter to balance fairness and efficiency.
V_s: the set of vehicles in sector s. $N_s \quad |V_s|$.
x_s: the central location of s.
Triggering condition: No CH in s (at $t_e \neq 0$).
% **Procedure:**
$\forall i \in V_s$, $d_{CM}[i]$, $Q_e[i]$, $d_{CH}[i]$, $n_{CH}[i]$, $n_{NC}[i]$, $\mathbf{v_i}$ Algorithm 4
$\forall i \in V_s$, d_i distance between i and x_s
$d_{min} \quad \min_{i \in V_s} d_i$; $d_{max} \quad \max_{i \in V_s} d_i$.
% normalized distance to the sector centre.
$\forall i \in V_s$, $D_i \quad \frac{d_i \quad d_{min}}{d_{max} \quad d_{min}}$
% **Compute** the weight of each canon:

$$w_j \quad \frac{1}{N_s} \sum_{i \in V_s} \mathbf{v_i}[j]; \; j = 1; \ldots; 5$$

% **Create** $L_1; \ldots; L_5$: ordered lists for each canon:
  Define list element: (value, vehicle ID)
  Order by: value
  $L_1 \quad (d_{CM}[i]; i)$, decreasing order
  $L_2 \quad (Q_e[i]; i)$, decreasing order
  $L_3 \quad (d_{CH}[i]; i)$, increasing order
  $L_4 \quad (n_{CH}[i]; i)$, increasing order
  $L_5 \quad (N_{NC}[i]; i)$, decreasing order
% **Compute** the weighted Borda score for every vehicle:
$\forall i \in V_s$, p_{ij} position of i in L_j, $j = 1; \ldots; 5$.
$\forall i \in V_s$, $_i \quad \sum_{j=1}^{5} (N_s + 1 \quad p_{ij}) w_j$
% **Calculate** the centrality score for every vehicle in s:
$\forall i \in V_s$, $_i \quad N_s D_i$
% **Calculate** final scores:
$\forall i \in V_s$, $score_i \quad {}_i + (1 \quad) \; {}_i$
$k \quad \arg\max_{i \in V_s} score_i$.
Select vehicle k as new CH and notify it.

Algorithm 4 Distributed Criteria Vote (run at Vehicle)

Init: $(\mathbf{v}[1]; \ldots; \mathbf{v}[5]) = (0; \ldots; 0)$
Retrieve from local database: n_{CH}, $M_{n_{CH}}$, d_{CH}, $M_{d_{CH}}$, d_{CM}, $M_{d_{CM}}$, Q_e, Q_b, n_{NC}, $M_{n_{NC}}$
Compute the vector of preference values for each canon: $\mathrm{E} \quad [E_{n_{CH}}; E_{d_{CH}}; E_{d_{CM}}; E_{Q_e}; E_{n_{NC}}]$ according to equations (5) to (9).
$i \quad \arg\max_{j=1;\ldots;5} \mathrm{E}[j]$
$\mathbf{v}[i] \quad 1$
Return n_{CH}, d_{CH}, d_{CM}, Q_e, n_{NC}, $\mathbf{v}$.

proposed solution. Firstly, in the coverage area of a base station (which is the unit performing the sorting operations), which can only be a few kilometers wide, there are only a *limited number of clustering sectors*, as the multi-hop cluster sizes can only go up to a few hundred meters. Secondly, within each sector (in the order of the hundred meters) there can only be a *limited number of vehicles* (this will typically be below one hundred). Finally, these sorting algorithms will only be called when a new cluster needs to be formed. The *average time between two cluster formation operations* is, due to the nature of the problem, significantly high compared to the time required for a server to perform a sorting algorithm over five lists of approximately a hundred elements each. Therefore, the computation time can be considered negligible in the problem's timescale.

4) Billing and interoperability: The proposed approach does not involve any supplementary billing, does not require the mobile operators to agree to a common credit/reimbursement/cost-sharing plan, and still delivers long-term fairness.

Alternative arrangements such as sharing the cost per-round among all users in a cluster cannot be done directly. Users can subscribe to different data plans with different access providers, and the implementation of a mechanism that directly shares the costs amongst all users would require interoperability among all existing data plans from all providers. As data plans may also change over time, the interoperability problem would have to be revisited accordingly. Forcing compatibility between different business strategies is far from evident.

Another ad-hoc cost distribution alternative is, then, to propose that users subscribe to an application in which they would have to constantly or periodically make (and receive) payments for the service usage, that would go on top of their mobile provider's bill. This would certainly produce a generalized unwillingness to adopt the service. In contrast, the FSA algorithm can be implemented without any further action from operators or users.

V. Performance Evaluation

The performance of the proposed approach is evaluated in three different parts. In Section V-C, is dedicated to the cluster formation process. Section V-D concerns the Self-Adaptive (SA) algorithm, and Section V-E focuses on the Fair Self-Adaptive (FSA) algorithm. The evaluation methodology is described in Section V-A and the common simulation settings are presented in Section V-B.

A. Evaluation Methodology

The algorithms are evaluated through simulations. The road traffic simulations are performed with SUMO [23], together with OMNET++ [28] for network protocols (from physical to application layers). Both simulators are references in the domain and implement well-established models (for example, we use the Krauss car-following model [29] in SUMO, and a two-ray interference model in OMNET++ for the physical layer of the V2V network [30]). We also use the Veins framework [31] to establish the scenario synchronization between SUMO and OMNET++. The test scenarios in these simulations are conceived for performing a stress test of the clustering algorithms in each evaluated aspect, covering both regular and extraordinary situations. The results presented in this section for every evaluated metric are the average values over twenty repetitions.

B. Common simulation settings

In all the following simulations, the access protocol in the V2V network is IEEE 802.11p. The protocol for

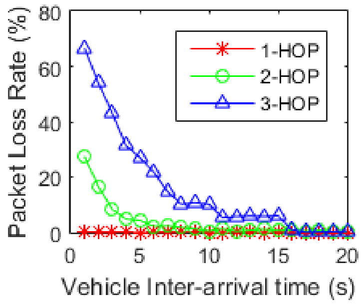


Fig. 4: IEEE 802.11p (V2V) PLR as a function of the vehicular inter-arrival time, for different numbers of hops.

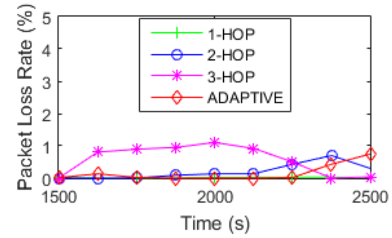
the network and transport layers is IEEE 1609.3. In the application layer we implement the Cooperative Awareness Messages (CAM) from the European ETSI ITS G5 standard, set at a frequency of 1 Hz (every vehicle broadcasts one message per second on V2V). The cellular network protocol is assumed to be LTE. The aggregation ratio performed by the CH is $(N_c) = 1{=}N_c$, where N_c is the cluster size. The path loss follows a Two-Ray Interference model [30]. We consider that to ensure the system's reliability, V2V PLR cannot he be higher than PLR_{max} = 10%. The IEEE 802.11p maximum range is set to $R = 800$ m. For the SA and FSA algorithms (Sections V-D and V-E), we set the hop number adaptation time $T = 40$ s, the maximum, minimum and default number of hops $(H_u; H_l; H_{default}) = (3; 1; 3)$, and the hop change density thresholds $(\ 1\ hop_{min};\ 2\ hop_{min};\ 2\ hop_{max};\ 3\ hop_{max}) = (17{:}5; 5{:}5; 22{:}0; 7{:}0)$ vehicles/km. These values have been set according to the findings published in [6].

C. Cluster Formation: Packet Loss Rate vs. Number of Hops

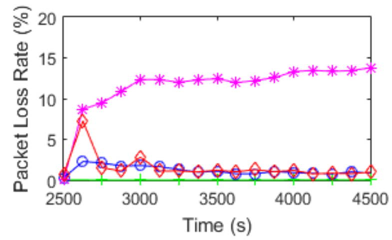
In the following, we show how the Packet Loss Rate (PLR) depends on the maximum allowed number of hops in a cluster. The results obtained were employed to determine the contexts (traffic densities) for which different cluster sizes were viable; and hence determine the triggering conditions for cluster size adaptations (see Algorithm 2 in IV-C).

1) Simulation Configuration: We consider a highway section which is $L_r = 5$ km long and has $L = 3$ lanes. $N = 60$ vehicles traverse the highway section at an average speed of $16{:}6$ m/s. Inter-arrival times (inversely proportional to density) vary between 1 s and 20 s. The maximum number of hops is set to H = 1, 2 and 3 in separate runs (it is fixed in every experiment).

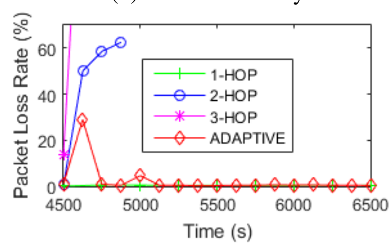
2) Simulation Results: Figure 4 shows the PLR in the V2V network as a function of the vehicular inter-arrival time for static hop configurations ($H = 1; 2$ and 3). Considering the inverse proportionality between inter-arrival time and vehicular density, we observe that for very high traffic densities, the V2V PLR worsens significantly for every supplementary allowed hop. This is due to the amplification of the *broadcast storm* effect when the number of nodes increases. We observe that the 2- and 3-hop configurations do not meet the PLR constraint of 10% for inter-arrival times shorter than 3 and 10 seconds respectively. Knowing that on the other hand, a higher number of hops improves the data aggregation performance, we conclude that there are three different vehicular density scenarios (low, medium and high) for which a different number of hops is optimal in terms of equations (3) and (4).



(a) Low density



(b) Medium density



(c) High density

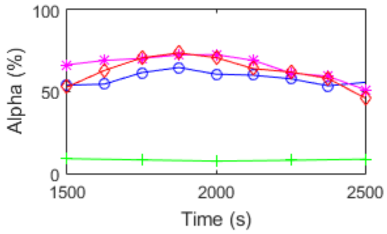
Fig. 5: Packet Loss Rate (PLR) in function of time for the tested scenario. Comparison between 1-,2- and 3-hop configurations vs. Self-Adaptive algorithm.

D. Self-Adaptive (SA) Clustering Algorithm

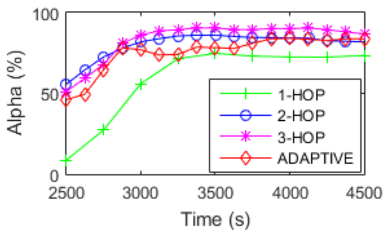
The proposed hop adaptation algorithm has been tested in order to verify its reactivity to changes in vehicular density and its convergence to the results in terms of PLR and (C) to the ideal hop configuration for the traffic situation.

1) Simulation Configuration: The traffic simulation consists of a 10 km long highway segment, divided into two *clustering regions* of equal length. In terms of vehicular traffic, the tested scenario consists of three consecutive flows of very different densities: from 0 s to 2500 s, 100 vehicles enter the highway section at an average inter-arrival time of 25 s. From 2500 s to 4500 s, a second flow of 200 vehicles will enter the highway section at an average inter-arrival time of 10 s. And finally, starting from 4500s, a flow of 1600 vehicles will enter at an average inter-arrival time of 1 second. A performance evaluation in terms of (C) and PLR is made, comparing the Self-Adaptive algorithm to the fixed-hop static configurations of the previous section, with 1, 2 and 3 hops.

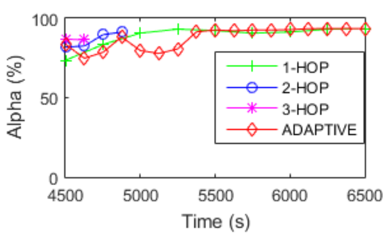
2) Simulation Results: The results are presented in separate figures for the different traffic flows for an improved reading



(a) Low density

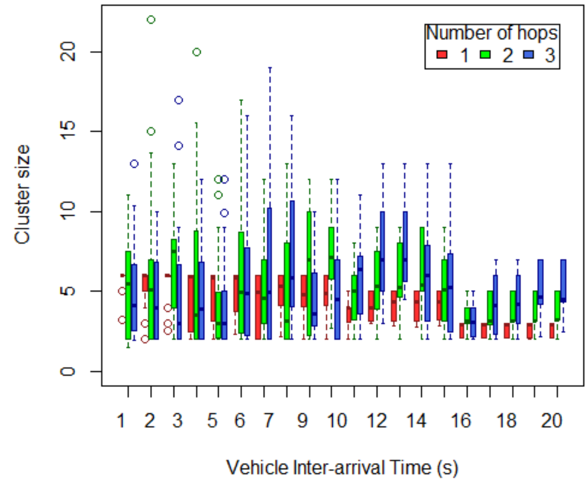


(b) Medium density

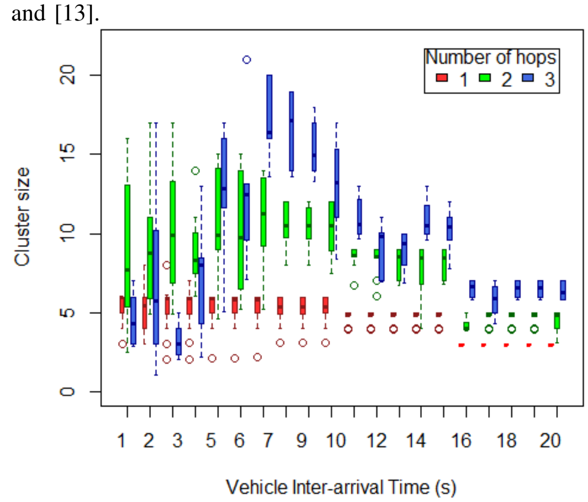


(c) High density

Fig. 6: Cellular data consumption reduction (Alpha) in function of time for the tested scenario. Comparison between 1-,2- and 3-hop configurations vs. Self-Adaptive algorithm.



(a) State-of-the-art: clustering method based on [17] and [13].



(b) Proposed cluster head selection and cluster formation algorithm.

Fig. 7: Resulting cluster sizes for 1, 2 and 3-hop configurations. Comparison between the implementation of a state-of-the-art clustering algorithm (a) and the approach presented in this paper (b).

and analysis, but the reader should keep in mind that they are part of a single, continuous simulation. PLR are shown in Figure 5 and compression ratios are shown in Figure 6. Our Self-Adaptive algorithm is compared to fixed-hop algorithms, that do not adapt the cluster sizes to traffic density.

During the first part of the simulation, a very light traffic density is introduced, so that PLRs are very low (Figure 5.a): all the algorithms remain below 1%. Vehicles are rather far away from each other and, as we can see in Figure 6, the 1-hop algorithm is unable to form big enough clusters, and is severely penalized in its aggregation performance when compared to the others. The best aggregation performance is achieved for the 3-hop algorithm closely followed by our Self-Adaptive algorithm (Figure 6.a). When the second flow of vehicles arrives at the mark of 2500 s, curves gradually change, and the 3-hop algorithm goes beyond the PLR acceptability threshold of 10% (Figure 5.b). The Self-Adaptive algorithm then changes the number of hops, from 3 to 2, and we can see a significant reduction of the PLR after the peak we get when the new flow starts. The Self-Adaptive algorithm's compression curve starts following the 2-hop curve (Figure 6.b). Finally, for the highest density (Figures 5.c and 6.c), the PLR curves of 2- and 3-hop go off-chart, leaving 1-hop as the only viable possibility. The Self-Adaptive algorithm triggers a hop change again, resolving another PLR peak, while its aggregation performance follows the curve of 1-hop. In the worst case, signalling represents only 0.17% of the savings in terms of number of messages. Even if messages have different lengths, this rough estimation shows that the signalling associated to our algorithm is negligible.

As an intermediate conclusion, we see that the Self-Adaptive algorithm correctly adapts to extreme density changes, achieving an important reduction in cellular network usage in every scenario (thus making important monetary savings at large scale), while respecting the imposed constraints of packet loss rate on the V2V network, guaranteeing that specific applications' requirements can be met. This method is efficient for preserving network performance, but the problem of fairly distributing the communication costs in which CHs incur still needs to be addressed. The next section evaluates our proposal to tackle this issue.

3) Comparison with state-of-the art clustering algorithms - Maximum number of hops vs. cluster size: Figure 7 shows the distributions of cluster sizes when the maximum number of hops is respectively 1, 2 or 3. On the left (a), we show the performance of VMaSC, a state-of-the-art clustering algorithm [13], [17]; on the right (b) we show the performance of SA. Cluster size is indeed a metric of utmost importance since it is directly related to the aggregation ratio.

The VMaSC algorithm that we compare to has been chosen because it is a paradigmatic example of the most common design strategy in the literature of clustering algorithms for vehicular networks. This design consists in Cluster Heads that take this role in function of their own observation of their environment (what we can call a *self-proclaimed CH*). In contrast, the approach we propose in this paper is to let the cellular base station select a CH, with the rest of the cluster formation process taking place through V2V interaction.

We observe from the figure[3] that at high inter-arrival times (say above 17 s), there is a clear correlation between the number of hops and the cluster sizes. For both algorithms indeed, increasing the number of hops logically increases the average cluster sizes. However, for low inter-arrival times (i.e. high vehicle densities), VMaSC is unable maintain this correlation. The issue comes from the fact that VMaSC allows too many nodes to self-proclaim as CHs. This results in smaller cluster sizes, even when more hops are permitted. In contrast, the region-based CH selection performed by the cellular base station in our approach yields clearly distinct cluster sizes when increasing the number of hops, even in higher vehicular density scenarios. This allows us to control the average cluster sizes by allowing more or less numbers of hops.

Thus, we conclude that the CH selection design choice of our proposal gives it a clear advantage with respect to VMaSC for the data aggregation-based cellular cost reduction problem.

E. Fair Self-Adaptive (FSA) Clustering Algorithm

In this section we compare the Fairness-agnostic and Fairness-aware approaches of our Self-Adaptive clustering algorithm in terms of network performance and fairness metrics.

1) Simulation Configuration: The simulation experiment consists of a group of 100 vehicles that pass through a 10 km highway segment, for approximately 100 times each, with randomized order of re-entrance. In the curves that follow, where time is represented in the horizontal axis, it refers to the simulated time, where the 100 vehicles randomly re-enter with a precise inter-arrival time. This is thus the equivalent of the time that it would take for 10,000 vehicles to traverse the 10 km highway segment, with an uniform inter-arrival time. We set the fairness vs. efficiency parameter $= 0:65$.In the simulations, we don't simulate events of non-compliance with the rules – the implementation of a behavioral model, as well as the analysis of the variation of is left for future work.

2) Vehicles Having Served as CH: A first relevant outcome of introducing distributive justice over the long term is that the role of CH, which seems like a temporary burden, will be taken by a greater proportion of the participants if we analyze successive samples over time. The box-plot in Figure 8 shows the distribution of the number of times that each vehicle has been designated as CH over time. Regardless of the time passing, the value range of the box (which represents the majority of the vehicles) for the *FSA Algorithm* remains small and doesn't increase. Their whiskers (representing slightly less than 25% of the values below and above the value range of the box each) are usually tiny or non-existent. This means that, even if the average number of times that vehicles in general serve as CH increases with time (which is necessary), all of the vehicles serve approximately the same amount of times as CH. On the other hand, for the *SA Algorithm's* boxes, their size increases, and the whiskers keep getting bigger, showing extreme disparities between the participants.

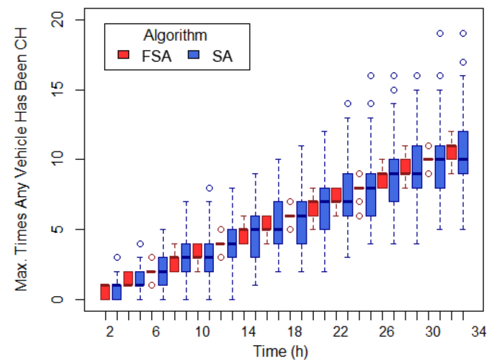
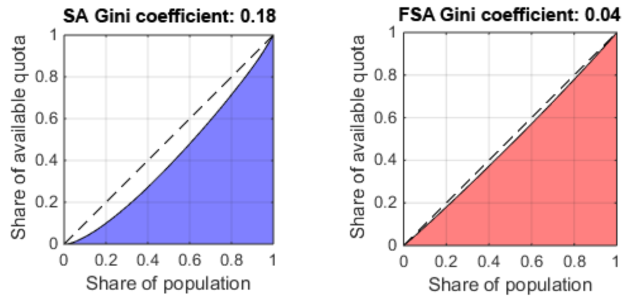


Fig. 8: Box plot showing the distribution of the number of times that every vehicle has served as CH, for both algorithms, over time.

3) Cellular Quota Consumption: The differences in the outcome of the CH selection that we discussed above, has a direct impact on cellular quota consumption. We now see some examples of the different distribution of the cellular quota usage with both algorithms, which translates almost directly into different economic costs. Since the final available quota can be seen as a remaining wealth that can be distributed more or less evenly among the participants depending on the algorithm's distribution policies (or lack thereof), we have analyzed the distribution of the remainder of this resource following two common fairness metrics: the Gini coefficient and the Lorenz curve [32] (see Figure 9). The Gini coefficient measures the statistical dispersion of the distribution of a resource among a certain population (like the distribution of wealth between a country's residents). The perfect distribution where every person has the same resources has a coefficient $G = 0$, whereas the case of one person having all the resources has a coefficient of $G = 1$. The Lorenz curve is a graphical representation of this distribution, which shows the proportion of resources held by the bottom x proportion of the population. Perfect equality is represented by the diagonal line. If we call A the area between the perfect equality line and the Lorenz curve, and we call B the area below the Lorenz curve, the Gini index can be defined as $G = \frac{A}{A+B}$. Results show that the Fair Self-Adaptive (FSA) algorithm reduces the Gini coefficient by 78%, with the Lorenz curve approaching the perfect equality

[3] In the box plots in this article, the solid band represents the median (second quartile), while the box is delimited by the first and third quartiles. The whiskers mark the lowest and highest datum still within 1.5 Inter-quartile Range (IQR) of the lower and upper quartiles, respectively. The outliers are marked as circles.



(a) Self-Adaptive (SA) algorithm (b) Fair Self-Adaptive (FSA) algorithm

Fig. 9: Lorenz curves and Gini indices analyzing the final distribution of available cellular quota in both algorithms.

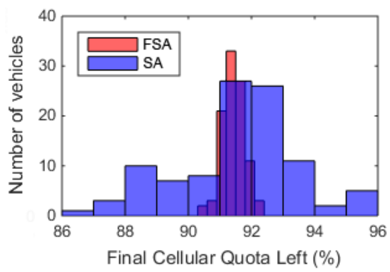


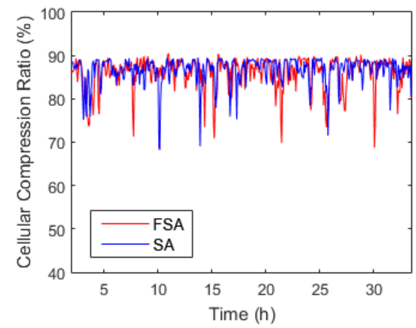
Fig. 10: Histogram of the final available cellular quotas among the vehicles having served as CH.

line. The clear final picture of the situation can be seen in the histogram of Figure 10, where we can see that for the Fair Self-Adaptive algorithm, the variance of the final cellular quota is much smaller than for the Self-Adaptive algorithm.

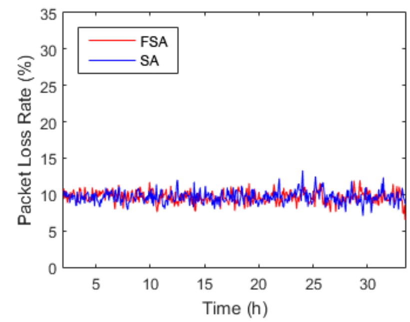
4) Network Performance: The original Self-Adaptive clustering algorithm aims to adapt the size of the clusters, finding compromises between cellular consumption reduction and PLR on the V2V network. We could expect that modifying the geometry and criteria of the election process could affect the metrics that this algorithm was designed for. Figures 11a and 11b compare both algorithms in terms of data compression and PLR, respectively. As we can clearly see, the improvements in distributive justice come at no cost in terms of network performance.

VI. Conclusion

In this paper, we studied and proposed a solution addressing the required trade-off between data aggregation over cellular networks and minimal performance over V2V networks. Innovative ITS applications will require big volumes of data to be uploaded from the vehicles to the cloud using the cellular network. The cost of access to this network can become a major problem and data aggregation at Cluster Head (CH) appears as a promising solution. This approach yields better results when cluster sizes increase. However, this comes with a higher Packet Loss Rate (PLR) in the V2V network because of the contention-based access protocol. We have thus proposed an Self-Adaptive clustering algorithm that dynamically adapts the cluster size to the traffic density. We have also proposed a Fair Self-Adaptive algorithm based on the theory of distributed justice to ensure a fair distribution over time of the responsibilities of being elected CH and hence a better social acceptability. Simulations show that the Self-Adaptive algorithm exhibits high compression ratios and low PLR. The Fair Self-Adpative alogorithm improves the Gini index by 78% in terms of remaining quota with respect to the baseline algorithm.



(a)



(b)

Fig. 11: Comparison between the FSA and SA algorithms, regarding: (a) Data compression ratio on the cellular network link over (simulated) time. (b) Packet Loss Rate (PLR) on the V2V network.

References

[1] C. Englund, L. Chen, A. Vinel, and S. Y. Lin, "Future applications of vanets," in *Vehicular ad hoc Networks*. Springer, 2015, pp. 525–544.

[2] "Part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications amendment 6: Wireless access in vehicular environments," IEEE 802.11p published standard, July 2010.

[3] S. A. A. Shah, E. Ahmed, M. Imran, and S. Zeadally, "5g for vehicular communications," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 111–117, 2018.

[4] J. P. Garbiso, A. Diaconescu, M. Coupechoux, J. Pitt, and B. Leroy, "Distributive justice for fair auto-adaptive clusters of connected vehicles," in *Foundations and Applications of Self* Systems (FAS* W), 2017 IEEE 2nd International Workshops on*. IEEE, 2017, pp. 79–84.

[5] C. Campolo, A. Vinel, A. Molinaro, and Y. Koucheryavy, "Modeling broadcasting in ieee 802.11 p/wave vehicular networks," *IEEE Communications letters*, vol. 15, no. 2, pp. 199–201, 2010.

[6] J. Garbiso, A. Diaconescu, M. Coupechoux, and B. Leroy, "Dynamic cluster size optimization in hybrid cellular-vehicular networks," in *Proceedings of the 19th IEEE International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2016, pp. 557–563.

[7] W. Vandenberghe, E. Vanhauwaert, S. Verbrugge, I. Moerman, and P. Demeester, "Feasibility of expanding traffic monitoring systems with floating car data technology," *IET Intelligent Transport Systems*, vol. 6, no. 4, pp. 347–354, 2012.