A Green Cluster-based Traffic Information Acquisition Method in Vehicular Networks

GAO-FENG HONG[1], WEI SU[1], XIN-DI HOU[1], LI-SHENG MA[2]
[1] *School of Electronic and Information Engineering, Beijing Jiaotong University, China*
[2] *School of Computer and Information Engineering, Chuzhou University, China*
E-mail: {honggf; wsu; freezerburn}@bjtu.edu.cn, {mls}@chzu.edu.cn

The Smart Vehicle (SV) with multiple onboard sensors is one of the indispensable tools to collect dynamic traffic information in the Intelligent Transport System (ITS). In this paper, we aims at jointly relieving the cellular network load and the transmission energy consumption during the SVs' information uploading. A Green Cluster-based Traffic Information Acquisition Method named GCTIAM is proposed, where an edge clustering management architecture is designed to control the traffic information upload cycles of each SV. On this basis, We build an optimal clustering mechanism to further improve the energy utilization while relieve the cellular network load. The mechanism optimizes the cluster iteration radius and divides SVs into different competitive areas according to their movement characteristic. An adaptive multi-parameter iteration among the same competitive area will be executed to decide the data transmission mode and build optimal cluster structures by considering dynamic parameters such as link duration time, uploading data volume etc.. The simulation results show that the proposed scheme can reduce the transmission energy consumption by 81% while cut down the cellular access rate up to 70%.

Keywords: Vehicular Network, Information Collection, Cluster, Edge Cloud, Energy Efficiency, Cellular Network Load Offloading

1. INTRODUCTION

The operation of emerging ITS services, e.g., active turning movements at intersections, vehicle trajectory prediction, queue spillover management, always relies on ample and real-time traffic information [1]. Traffic information collected by the traditional road side infrastructure can hardly meet the information dimension requirements of these services. With the rapid development of vehicular networks, SVs equipped with intelligent sensors are considered to be available real-time traffic information collector [2] [3]. An energy-efficient transmission method can stimulate private SVs to participate in traffic information uploading, which can further extend the dimensions of the collected traffic information. Meanwhile, the cellular network load brought by the data uploading may increase the possibility of the network congestion. Therefore, during the SV traffic information uploading process, how to reduce the transmission energy consumption while keep the cellular network load at a relative low level deserves to be investigated.

There are two main communication technologies, say Dedicated Short Range Communication (DSRC) and Cellular-Vehicle-to-Everything (C-V2X), in current vehicular networks [4]. DSRC occupies a frequency band from 5.795 to 5.815GHz and provides data transmission rates ranging from 6 to 27 Mbps within a relative limited transmission

range [5]. Compared with DSRC technology, C-V2X can realize long-range transmission with higher transmission rate. However, when extensive terminals access the cellular access point (C-AP) at the same time, the channel preemption problem may largely increase C-AP's overhead on configuring its radio resources [6]. Meanwhile, the long transmission distance between the SV and the cellular AP will bring higher transmission energy consumption, which may cause a greater mutual interference in turn. When the interference reaches a certain value, it will cause frequent retransmissions which will consume much more energy [7]. Therefore, the utilization of cellular communication should be restricted in vehicular traffic information uploading process only when necessary.

Numerous optimal data transmission works in vehicular networks have been done. In [8], the authors proposed a message dissemination scheme with low control overhead to satisfy the transmission delay demands. To achieve an efficient data collection in Vehicular Ad hoc Network (VANET), Zhu et al. [9] built a data aggregation tree. Based on dynamic programming, their scheme obtained a better QoS performance with less transmission cost. Articles [10] [11] aims at utilizing cluster-based model to enhance the communication stability. In [12], the authors proposed an energy efficient and reliable information dissemination mechanism by utilizing the cluster theory. All of the above works are based on a single communication mode, which usually applied in the VANET. Meanwhile, some optimal transmission solutions have been raised for the heterogeneous vehicular network [13]. In [14], the authors proposed an optimal data transmission scheduling scheme based on the deep Q-learning to minimize transmission costs. Article [15] proposed a transmission relay selection mechanism to reduce the cellular cost. In [16], the authors investigated a collaborative content delivery mechanism through different access entities to improve the network efficiency. Articles [17] [18] investigate the optimal resource allocation mechanisms to satisfy the latency requirement and leverage the bandwidth cost. Authors in [19] proposed a self-adaptive clustering mechanism to achieve a trade-off between data aggregation and communication congestion in the cellular system. Article [20] proposed a cluster-based traffic management to relieve the traffic loads in cellular networks while keep a relative low transmission delay. Ucar et al. [21] first proposed a multi-hop-based clustering mechanism used in heterogeneous vehicular network to reduce cellular network overhead. In [22], authors proposed a centralized clustering method by using the modified K-means algorithm and some mobility metrics to achieve cluster head (CH) election. It can realize the transmission traffic offloading from cellular networks to DSRC. Memedi et al. [23] first proposed the power control concept in heterogeneous vehicular network basing on the clustering theory. Dong et al. [7] formulated the CH election as a p-median problem to realize the energy-efficient transmission in cellular vehicle networks. Moreover, the optimization in the backbone link transmission has been considered in [24].

Majority of the above works optimize the vehicular data transmission by utilizing clustering theory [10] [11] [12] [19]- [23] [7], which show the efficient use of clustering approaches in achieving collaborative and energy efficient transmission. However, none of these works considered the information uploading cycle management, so that the performance of these works in traffic information acquisition remains to be explored. Meanwhile, the competitive radius of each node during the CH election is always considered to be static [19]- [23], which may affect the stability of the cluster formation in dynamic environment. In addition, the uploading data volume of each SV plays an important role in the CH election process to reduce the total transmission energy consumption, which is rarely considered by the existing clustering algorithms. All the above factors put forward higher requirements for the clustering management in the traffic information acquisition.

By now, the edge cloud server (ECS) is considered to be a powerful decentralized clustering management entity [25]. The location-dependent characteristic of the ECS can well support the control of the information uploading cycle while the road side unit (RSU) can act as a control signaling interaction entity without occupying cellular resources. Utilizing the advantages of the edge clustering management provides a more reasonable data aggregation cycle during the traffic information uploading, which can largely reduce the unnecessary transmission energy consumption while select a suitable transmission mode for the SVs to relieve the cellular resource.

In this paper, to reduce the transmission energy consumption while keep a relative low cellular network load during the SV traffic information acquisition process, a cluster-based data uploading mechanism called GCTIAM is proposed. Firstly, a regional green clustering management architecture is proposed to control the data uploading cycle of the SVs. Secondly, an optimal clustering mechanism is designed by jointly considering the multiple factors such as: link maintenance time, uploading data volume, distance between the SV and the C-AP etc., which can achieve the data aggregation by minimizing the transmission energy consumption and reduce the C-AP access rate. Moreover, we obtain the optimal values of the variable parameters in the proposed cluster algorithm through simulations. The main contributions are shown as follows:

- *To support the implementation of GCTIAM and ensure the transmission reliability, a regional green clustering management architecture is proposed. The proposed green clustering management architecture can well control the uploading cycle of traffic information collected by each SV. Meanwhile, the concept of the transmission identifier and relevant communication interaction process under the proposed architecture are described in details. This architecture can reduce the unnecessary communication overhead while ensure the reliability of data transmission.*

- *To realize a low energy consumption data transmission and relieve the cellular network load, the optimal clustering mechanism operating on the above architecture is proposed. It concludes four main phases: Initial broadcast phase, CH election phase, Clustering phase and Traffic information aggregation phase. The competitive radius of each SV is set to be dynamic, the value of which is depended on the average link duration time of the SV's neighbors. CHs are selected on the basis of the adoptive multi-parameter iteration, so as to minimize the transmission energy consumption while reducing the cellular APs load and ensuring the cluster stability.*

- *We build the relevant simulation environment on the NS-2 simulator and evaluate the performance of GCTIAM with different adjustment factors and get the optimal values of these factors. Basing on the optimal parameters, we verify the feasibility and reliability of GCTIAM and compare the energy consumption performance of GCTIAM with the baseline methods.*

2. System architecture of GCTIAM

We propose a green clustering management architecture as shown in Fig.1, which can effectively support the implementation of the subsequent clustering mechanism. This architecture mainly includes four types of communication nodes: SV, ECS, C-AP and RSU. Table 1 summarizes all the important notations used in our paper.

SV: Acting as a traffic information collector, and uploading the traffic information at a specific time.

ECS: To gather and analyze traffic information near the user side [26], and reduce relevant data transmission delay.

C-AP: Working as a relay for information transmission between the SV and the ECS, providing SVs with a wide network connection for data uploading.

RSU: To monitor SV movement within the coverage area of its master C-AP.

Table 1: Key Notations

Notation	Definition
v_i	A smart vehicle i in the road segment.
U_i	The information acquisition identifier of v_i.
S_i	Speed of v_i.
R	The maximum communication range of SVs using DSRC communication mode.
P_n	A cellular access point n in the road segment.
C_{ap}	The coverage size of cellular access points.
N_n	The number of SVs under the coverage of cellular access point n.
N_{opt}	The number of SVs which can realize a better network performance within a cluster.
N_{i-N}	The number of SVs within the coverage range of v_i.
I_i	The amount of data v_i ready to be uploaded.
d_{ij}	The distance between v_i and v_j.
d_{v_i}	The distance between v_i and its current cellular access point.
$d_{P_n}^{max}$	The maximum distance between each smart vehicle and P_n.
$d_{P_n}^{min}$	The minimum distance between each smart vehicle and P_n.
E_{elec}	The energy consumed by an OBU to process unit data.
ε_{fs}	The energy consumed by an OBU to transmit unit data under the free space model.
ε_{mp}	The energy consumed by an OBU to transmit unit data under the multi-path fading model.
H_r	The height of the receiving antenna.
H_t	The height of the transmitting antenna.

Assuming SVs move on a two directions straight road. Each C-AP has an ECS attached to it. SVs under the same cellular network coverage will be divided into several clusters based on the subsequent clustering mechanism. In each cluster, DSRC technology is used for the CH to gather the traffic information collected by CMs. After aggregating the cluster information, CHs transmit the aggregated data to the attached C-AP through C-V2X. These data will be uploaded to the ECS for the relevant traffic situation analysis. To reduce the redundancy of the collected traffic information, each SV will only upload traffic information once under the coverage of each C-AP. An information acquisition identifier U_i is built for each SV v_i. This identifier is used to indicate whether the SV has completed the uploading of the collected traffic information under the communication range of the current C-AP. If the information has not been uploaded successfully, the value of this identifier will be 1, otherwise it will be set to 0. For SVs under the coverage of a C-AP, only when U_i=1 will the SV v_i participate in the subsequent clustering process. Once an RSU detects a v_i enters a new C-AP coverage area, it will send a beacon message to v_i notifying it to set the value of U_i to 1. Meanwhile, a SV information base (SVIB) is built for each SV to record the state information of the SV and its neighbors within the specified communication range. SVIB is used to support for subsequent CH elections.

During the clustering process, different SVs can be in different states. In addition

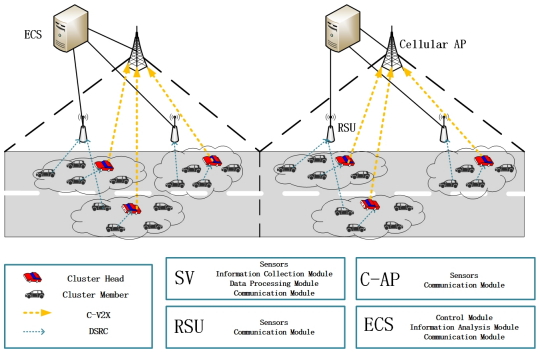


Fig. 1: The green clustering management architecture.

to the SVs which become CHs and CMs, some SVs may become isolated nodes for they haven't received any announcement beacons from CHs during a specific time period. Parts of CMs may fail to complete data uploading during a CH data aggregation round. For the sake of subsequent explanation, we divide the SV states in GCTIAM into five categories, each state and its corresponding SV behaviors are described as follows:

Orphan1: An SV enters a new C-AP coverage area and receive the beacon from the corresponding RSU.

Orphan2: An SV hasn't received CH announcement beacons during T_{BC} or a CM hasn't finished data uploading during a CH data aggregation round.

CH: An SV which has been elected as a cluster head.

CM: An SV becomes a member of a cluster.

Free: An SV which has finished its data uploading to the corresponding target receiver.

3. Clustering mechanism formulation

To reduce transmission energy consumption and relieve network load, an optimal clustering mechanism executing on the above architecture is presented in this section. We describe and analysis the four main phases in the clustering mechanism: Initial broadcast phase, CH election phase, Clustering phase and Traffic information aggregation phase.

3.1 Initial broadcast phase

Considering CH elections for SVs under the coverage of the same C-AP, the number of CHs can directly affect the efficiency of the traffic information uploading process. Inadequate CHs will cause the number of CMs within a single cluster to be too large, the difficulty of CHs to aggregate traffic information within the cluster will increase dramatically, which bring huge communication overhead to the CH. Excessive number of CHs may lead to more occupations of cellular resource, which can not achieve the purpose of reducing the cellular network load. The broadcast radius of each SV in the initial broadcast phase directly affects the number of neighbor SVs to which the SV can communicate, which in turn affects the number of CHs. So the broadcast radius is also known as the competitive radius from the point of view of CH election, which is important for the calculation of the optimal number of CHs in a specific area.

In the proposed scenario, each C-AP covers an equal area of the street. The size of

the coverage area of a C-AP is C_{ap}. SVs are evenly distributed on the street and travelling at a constant speed S_i. Assuming the final selected CHs are evenly distributed, all circular competitive areas of SVs can completely cover the coverage area of the current C-AP. Then the average competitive radius R_{ave} should satisfy the following condition:

$$\frac{N_n}{N_{opt}} R_{ave}^2 \pi = \alpha^2 C_{ap} \Longleftrightarrow R_{ave} = \alpha \sqrt{\frac{C_{ap} N_{opt}}{\pi N_n}} \tag{1}$$

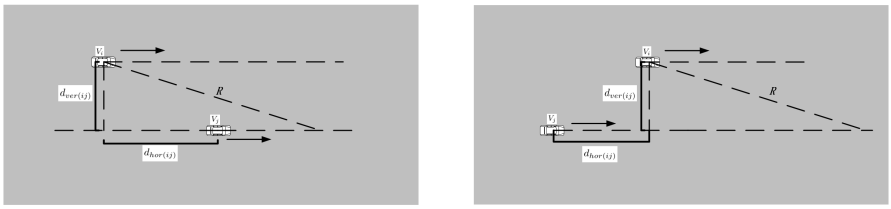
where α is the control coefficient under different types of network coverage areas, N_n is the number of SVs covered by the C-AP P_n, N_{opt} is the number of SVs which can make the network performance best within a cluster. According to [27], the value of N_{opt} can be set as 6 in the proposed mechanism.

Due to mobility features of SVs, the link duration time between v_i to its neighbor v_j affects the competitive radius of v_i. Two SVs moving in the same direction have two movement trends due to their speeds: (1)Driving away from each other as shown in Fig.2(a). (2)Driving close to each other as shown in Fig.2(b). The calculation methods of link duration time in both cases are shown in Formula (2) and (3) respectively:

$$T_{ij}^{dur} = \frac{(\sqrt{R^2 - d_{ver(ij)}^2} - d_{hor(ij)})}{|S_i - vel_j|} \tag{2}$$

$$T_{ij}^{dur} = \frac{(\sqrt{R^2 - d_{ver(ij)}^2} + d_{hor(ij)})}{|S_i - vel_j|} \tag{3}$$

where R is the maximum communication distance of SVs under DSRC communication mode, $d_{hor(ij)}$ and $d_{ver(ij)}$ are distances between v_i and v_j in x, y directions.



(a) v_j driving away from $v_i(S_j > S_i)$. (b) v_j driving close to $v_i(S_j > S_i)$.

Fig. 2: Vehicular movement trends.

In order to enable each SV to automatically adjust the size of competition radius in each information uploading period according to the average link duration time of its neighbors, an adjustment factor is added in Equation (1). Considering in the urban environment, the link duration time of SVs based on DSRC communication is usually around 30 seconds when SVs travel at a speed of 50km/h [28], the competitive radius of each SV is calculated as follows:

$$R_i = tanh(\theta \overline{T}_{i-dur}) R_{ave} \tag{4}$$

$\overline{T}_{i-dur}$ is calculated as shown in formula (5), N_{i-N} is the number of neighbor SVs. When the value of $\overline{T}_{i-dur}$ is low, the competitive radius will be reduced accordingly to reduce

the number of SVs competing with v_i and increase the number of CHs selected in the unit area, maintain the stability of intra-cluster communication.

$$\overline{T}_{i-dur} = \frac{\sum_{j=1}^{N_{i-N}} T_{ij}^{dur}}{N_{i-N}} \tag{5}$$

In the initial broadcast phase, the SV will broadcast its own state information such as: ID number(v_i), size of data to be uploaded(I_i), information acquisition identifier, speed, driving direction etc. After receiving broadcast messages from other SVs, v_i will update its SVIB, count the number of its neighbor SVs and calculate the average link duration time of its neighbors. Before the CH election phase begins, SVs will calculate their respective competitive radius based on formula (4).

3.2 CH election phase

After the initial broadcast phase, an SV in state Orphan1 starts the CH election phase. In order to achieve the objectives of of reducing transmission energy consumption and alleviating cellular network communication overhead, the factors referenced in the CH election should be comprehensive.

For v_i, the data transmission power consumption can be derived from the RF module model [29]:

$$E_{Tx}^{i}(I_i, d_{ij}) = \begin{cases} E_{elec}I_i + \varepsilon_{fs}I_i d_{ij}^2 & d_{ij} < d_{crossover} \\ E_{elec}I_i + \varepsilon_{mp}I_i d_{ij}^4 & d_{ij} \geq d_{crossover} \end{cases} \tag{6}$$

We suppose that E_{elec} is the energy consumed by an OBU to process unit data, I_i is the amount of data v_i ready to be sent, ε_{fs} is the energy consumed by an OBU to transmit unit data under the free space model, ε_{mp} is the energy consumed by an OBU to transmit unit data under the multipath fading model, d_{ij} is the distance between v_i and its neighbor v_j, and $d_{crossover}$ is the critical value of distance between free space model and multipath fading model which can be calculated as:

$$d_{crossover} = \frac{4\pi\sqrt{L}H_r H_t}{\lambda} \tag{7}$$

where H_r and H_t are heights of the receiving antenna and the transmitting antenna, L is the system loss factor and λ is the carrier wavelength.

According to formula (6), the amount of data transferred and the distance of transmission are two vital factors that determine the transmission energy consumption. Large amounts of data and a long transmission distance mean that more energy will be consumed during the process of transmission. Therefore, the SV with bigger data volume is selected as a CH to reduce transmission energy consumption caused by the increase number of transmission hops. On the other hand, the selection of CHs also considers the distribution and the movement trend of SVs. SV-distributed areas tend to select more CHs which can balance the number of SVs within the cluster and reduce the pressure on CHs to process aggregated data. Areas closer to the C-AP tend to select more CHs to balance the energy consumed by forwarding data from clusters farther away from the C-AP. In this paper, d_{v_i} represents the distance between v_i and its current connected C-AP, D_i represents the density of SVs near v_i which is defined as follows:

$$D_i = \frac{\sum_{j=1}^{N_{i-N}} d_{ij}}{N_{i-N}} \tag{8}$$

A small value of D_i means that there is a high density of SVs near v_i. Meanwhile, selecting SVs with long link duration time to be CHs is beneficial to the stability of clusters.

We divide the above factors into two categories: Data section(I_i) and SV distribution section(d_{v_i}, D_i, $\overline{T}_{i-dur}$). The CH election factor function is presented as follows:

$$Sco_i = e^{\beta \frac{|I_i - \mu_i|}{\sigma_i}} + e^{(1-\beta)W_i} \tag{9}$$

$$W_i = \frac{d_{P_n}^{max} - d_{v_i}}{d_{P_n}^{max} - d_{P_n}^{min}} - \frac{D_i}{R_{ave}} + tanh(\theta \overline{T}_{i-dur}) \tag{10}$$

where $d_{P_n}^{max}$ is the maximum distance between each SV and the C-AP under the coverage of P_n, $d_{P_n}^{min}$ is the minimum distance between each SV and the C-AP under the coverage of P_n, μ_i and σ_i can be calculated as:

$$\begin{cases} \mu_i = \frac{\sum_{i=1}^{N_{i-N}} I_i}{N_{i-N}} \\ \sigma_i = \sqrt{\frac{\sum_{i=1}^{N_{i-N}} (I_i - \mu_i)^2}{N_{i-N}}} \end{cases} \tag{11}$$

In the CH election phase, SVs in Orphan1 state iteratively compare the values of their respective election factors, SVs with higher election score are elected as CHs, and the rest SVs prepare for adding clusters.

3.3 Clustering phase

In the clustering phase, CHs will broadcast their own identity information, and SVs in state Orphan1 will start to add a cluster after receiving the broadcast information from CHs. The size of the broadcast radius of the CH directly affects the coverage of its cluster and the number of its CMs. From the perspective of energy balance, the CH near the C-AP should have a smaller broadcast radius to reduce energy consumed by the CH in processing aggregated information within the cluster. Moreover, when the density value D_i of the CH is high, its broadcast radius size should be reduced to prevent the time slot conflict caused by too many CMs in the cluster.

To solve the above problem, the zero point of the exponential function is used here to realize the two-way adjustment of the broadcast radius size. By jointly considering the density of SVs and the distance between the cluster and the C-AP , we optimize the cluster structure from energy efficiency. CH can independently calculate its broadcast radius as:

$$R_B^i = (e^{\gamma \frac{d_{v_i} - E(d_{P_n})}{d_{P_n}^{max} - d_{P_n}^{min}} + (1-\gamma)\frac{D_i}{R_{ave}}}) * \overline{R}_B \tag{12}$$

where $E(d_{P_n})$ is the average distance of SVs to P_n under its communication coverage, $\overline{R}_B$ is the adjustable base broadcast radius of CHs, γ is the adjustment factor.

For a certain SV v_i in state Orphan1, it may receive broadcast announcement beacons from multiple CHs. When this happens, v_i will select a CH with the highest election score for cluster addition. If the SV does not receive any broadcast information from CHs during a specific time period T_{BC} after the CH election phase, it will upload the collected traffic information directly through C-V2X.

3.4 Traffic information aggregation phase

Here, we describe the traffic information aggregation within the cluster in detail. After the CH receives the clustering information from SVs, it will record each SV in its own CM list and generate a TDMA time slot coordination rule for all CMs to match the corresponding transmission time slot. The SVs in the cluster only send their own data to the CH in their own time slots. When the CH receives the data from one of its CM, it will return a confirmation message to the CM and the CM will set its U_i to 0. The CH compresses the data and prepares for data uploading after collecting the traffic information of all members in its cluster. The CM which hasn't received the confirmation message from the CH during the intra-cluster data aggregated round will be considered as a failure to upload its data, it will then upload its data to the server directly through C-V2X.

The description in Fig.3 gives a global perspective of all the possible SV states (Orphan1, Orphan2, CM, CH and free) and the relevant actions in GCTIAM. The complete traffic information uploading process in GCTIAM is presented in Algorithm 1.

Algorithm 1: : GCTIAM Execution Process

Initialization: SV enters a new C-AP coverage area, information acquisition identifier resets.

for All SVs in state Orphan1 **do**
    SV adjusts its competitive radius according to formula (1)(2)(3)(4);
end for
Generating the set of v_i's competitive nodes:
for All SVs in v_i's communication range **do**
    $Scom = \{j | if(d_{ij} < R_i)\}$;
end for
Electing a CH within SVs in *Scom*;
For each CH, calculating its broadcasting radius according to formula (12);
CHs broadcast their CH announcement, SVs in state Orphan1 will join the cluster according to the received CH announcement beacons;
if An SV in state Orphan1 hasn't received any CH announcements during T_{BC} **then**
    Change its state to Orphan2;
end if
CMs upload their collected traffic information to corresponding CHs through DSRC;

CH uploads the aggregated information to the ECS through C-V2X;
if A CM in a specific cluster hasn't finished data uploading during the intra-cluster data aggregated round **then**
    Change its state to Orphan2;
end if
The SVs in state Orphan2 upload their collected traffic information to the ESC directly through C-V2X.

3.5 Algorithm complexity analysis

The complexity of the GCTIAM mainly depends on the CH election phase of the clustering mechanism. For a specific C-AP, assuming that the number of SVs which are ready to upload their collected information is N. Considering that the CH election in the

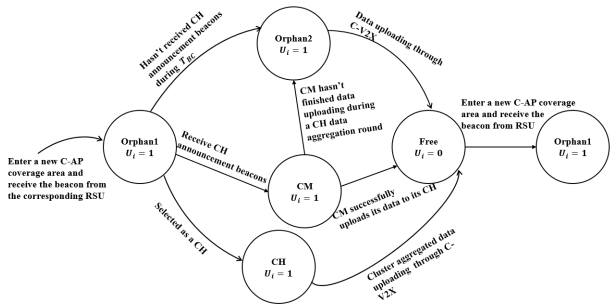


Fig. 3: States of a smart vehicle.

GCTIAM is based on the comparison of the Sco_i value, so that the time complexity of the GCTIAM can be approximated by $O(NlogN)$ per traffic information uploading round.

4. Simulations and analysis

To simulate and test the performance of GCTIAM, we utilize the NS-2 simulator to build the simulation environment where SVs evenly move on a straight four-lanes road with a random speed, each C-AP covers a road segment of 800 meters long and 20 meters wide. The velocity and the collected data volume of each SV will take random values in the specified range respectively. The relevant simulation parameters are shown in Table 2, where the relevant settings of the data transmission power consumption are referred to [29] as the classical values in the wireless communication.

Table 2: Simulation parameters

Parameters	Values
Number of C-APs	18
Number of SVs within the range of a C-AP	100
SV velocity	40~50km/h
Intra-cluster data transfer protocol	802.11p
The amount of data to upload per SV	45~55MB
CH data compression ratio	80%
C-AP coverage area	800m*20m
E_{elec}	50nJ/bit [29]
ε_{fs}	$10pJ/bit/m^2$ [29]
ε_{mp}	$100pJ/bit/m^2$ [29]

4.1 Analysis of measurement parameters

In our experiment, we mainly measure the following four types of parameters to evaluate the performance of our proposed mechanism.

Optimal values of the adjustment factors in the clustering mechanism: In the proposed clustering mechanism, β, γ and $\overline{R}_B$ are the vital adjustment factors which largely affect the performance of the traffic acquisition method. Through the simulation, we can get the optimal values of the adjustment factors to achieve the best performance of the proposed method.

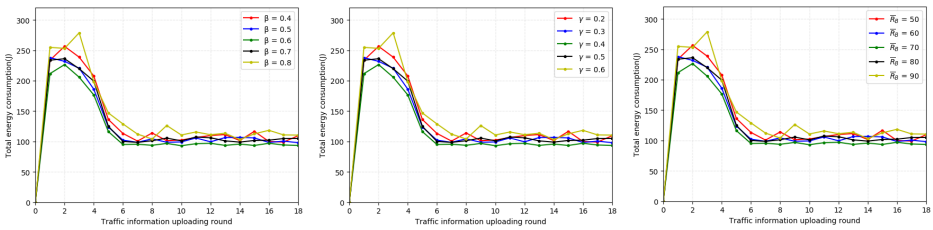
Ratio of total collected data received by ECSs: We expect that more SVs upload information through clustering to save energy and reduce cellular network load. This parameter reflects the proportion of SVs which upload their collected information through cluster.

C-AP access rate: To verify that our method can efficiently reduce the cellular network load while numerous SVs upload data simultaneously, we measured the C-AP access rate under each information uploading round. This parameter indicates the change in relevant performance compared with no optimization measures.

The total energy consumption: The energy consumption produced by the SV in the process of information uploading is the decisive factor affecting whether a private car is willing to participate in the information uploading. Therefore, the total energy consumption under each information uploading round is the main performance index in our scheme.

4.2 Simulation result

Fig.4(a)~4(c) show the influence of β, γ and $\overline{R}_B$ on the energy consumption performance of GCTIAM. It can be seen from the Fig.4(a)~4(c), the energy consumption is relatively high at the beginning of the GCTIAM execution. This can be explained as follow. More signaling interaction is required when the cluster is initialized and built, and most of the consumed energy is used for broadcasting relevant signaling. Subsequently, as the driving state of the SV is relatively stable during a straight line driving, the signaling cost for maintaining cluster stability is much less than that in the early stage of cluster construction, which resulting in a decrease in the total energy consumption for data uploading. Meanwhile, we find out that the different value of β, γ and $\overline{R}_B$ can affect the energy consumption performance of the GCTIAM to a certain extent. It can be concluded that when $\beta = 0.6$, $\gamma = 0.4$, $\overline{R}_B = 70$, the GCTIAM presents the lowest energy consumption performance respectively.

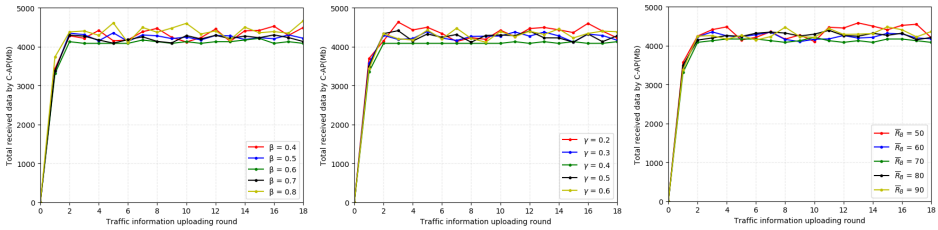


(a) Influence of β. (b) Influence of γ. (c) Influence of $\overline{R}_B$.

Fig. 4: The total energy consumption under different adjustment factors.

Fig.5(a)~5(c) show the influence of β, γ and $\overline{R}_B$ on the total data size transmitted by C-V2X by utilizing GCTIAM. Less data uploading volume through C-V2X means a more reasonable clustering division, where the CHs can better compress the data in an allowable range. Meanwhile, less occupation of cellular network resources and less energy consumption can be achieved when the total data size transmitted by C-V2X becomes smaller. It can be concluded that when $\beta = 0.6$, $\gamma = 0.4$, $\overline{R}_B = 70$, the total data size transmitted by C-V2X under each C-AP is the lowest.

Based on the above simulation results, we can conclude that when $\beta = 0.6$, $\gamma = 0.4$ and $\overline{R}_B = 70$, the overall performance of GCTIAM is optimal. In the subsequent simu-



(a) Influence of β. (b) Influence of γ. (c) Influence of $\overline{R}_B$.

Fig. 5: The total data volume transmitted by C-V2X under different adjustment factors.

lation, we determine the values of the three parameters as the optimal values for further performance testing.

Fig.6 shows the ratio of total collected data received by ECSs after each round of data uploading by utilizing GCTIAM. Considering that each CH will compress 80% of the intra-cluster information during the simulation, the result reveals that the data received by the C-AP accounts for 80%~85% of the total data volume collected by 100 SVs, which verifies the stability of GCTIAM. We can see that most of the SVs upload data through intra-cluster information aggregation to realize the load reduction on the cellular network.

Fig.7 shows the proportion of SVs uploading data through C-V2X after adopting GCTIAM. It can be seen that, GCTIAM can largely relieve the load on C-APs which reduce the access rate of the C-AP to around 20%. It can maintain similar performance during each round of data uploading, combined with the information acquisition identifier and the dynamic competitive radius, GCTIAM can minimize the delay caused by CH election iterations and ensure the efficiency of information upload while occupying less cellular network resources.

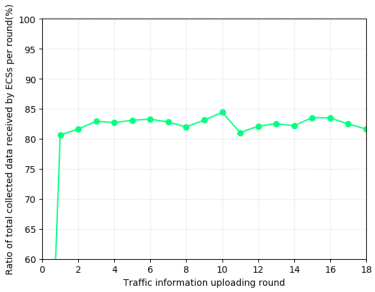


Fig. 6: Ratio of total collected data.

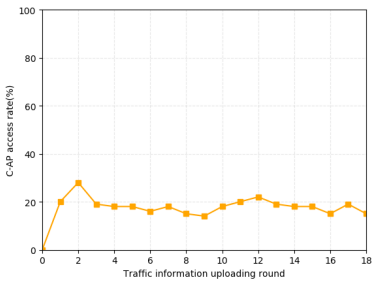


Fig. 7: C-AP access rate per round.

As shown in Fig.8, we make the energy consumption performance of GCTIAM with the all-C-V2X method where all SVs directly upload their data to ECSs through C-V2X and the EECM proposed in [7]. The value of p in [7] is set to 10. It can be seen from Fig.8 that the average energy consumption of the GCTIAM is reduced by 81% and 62% compared with the all C-V2X method and EECM respectively. It can be explained as follow. The dynamic cluster broadcast radius design and the modified muti-parameter CH election method of GCTIAM, optimize the intra-cluster information aggregation procedure and make the data uploading more energy-efficient.

As shown in Fig.9, we make the C-AP access rate performance of GCTIAM with the all-C-V2X method the EECM proposed in [7]. The value of p in [7] is set to 10. It can be seen from Fig.9 that compared with the EECM, the GCTIAM can also reduce the C-AP access rate to a certain extent. It is because the dynamic broadcast radius in the GCTIAM can achieve an adjustment in the number of CM, which reduces the number of the isolated node in the clustering process.

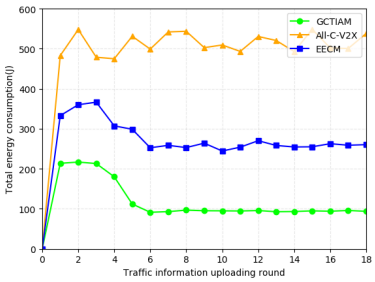


Fig. 8: Energy consumption comparison.

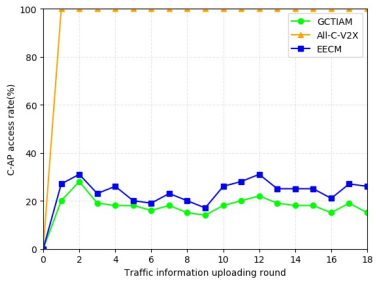


Fig. 9: C-AP access rate comparison.

5. Conclusion

In this paper, we have focused on chasing a reliable and economical method to upload the traffic information collected by private SVs. We proposed an optimal pervasive traffic information acquisition method called GCTIAM. To control the data uploading cycle and ensure the reliability of data transmission, a green clustering management architecture is constructed. An energy-efficient clustering mechanism is implemented on the above architecture to release the cellular network load and reduce the average network energy consumption. Simulation results show the optimal values of the variable parameters β, γ and $\overline{R}_B$ in GCTIAM, and verify the performance of GCTIAM in reducing energy consumption, relieving cellular network load and ensuring data transmission reliability.

Future work will focus on improve the performance of GCTIAM in more complex traffic environment.

ACKNOWLEDGMENT

This paper is supported in part by the National Key Research and Development Project of China (No. 2019YFF0303101), the Academic Funding Project for Top Talents of Disciplines (Majors) in Universities of Anhui Province of China (No. gxbjZD2021080), and the Key Project of Science Research in Universities of Anhui Province of China (No. KJ2020A0718).

REFERENCES

1. H. Li, G. Zhao, L. Qin, and Y. Yang, “Design and optimization of a hybrid sensor network for traffic information acquisition,” *IEEE Sensors Journal*, Vol. 20, no. 4,

2020, pp. 2132–2144.
2. K. Jurczenia and J. Rak, "A survey of vehicular network systems for road traffic management," *IEEE Access*, Vol. 10, 2022, pp. 42 365–42 385.
3. J. Wang, C. Jiang, Z. Han, Y. Ren, and L. Hanzo, "Internet of vehicles: Sensing-aided transportation information collection and diffusion," *IEEE Transactions on Vehicular Technology*, Vol. 67, no. 5, 2018, pp. 3813–3825.
4. Z. H. Mir, J. Toutouh, F. Filali, and Y.-B. Ko, "Enabling dsrc and c-v2x integrated hybrid vehicular networks: Architecture and protocol," *IEEE Access*, Vol. 8, 2020, pp. 180 909–180 927.
5. K. Z. Ghafoor, M. Guizani, L. Kong, H. S. Maghdid, and K. F. Jasim, "Enabling efficient coexistence of dsrc and c-v2x in vehicular networks," *IEEE Wireless Communications*, Vol. 27, no. 2, 2020, pp. 134–140.
6. Y. Mao, E. Piovano, and B. Clerckx, "Rate-splitting multiple access for overloaded cellular internet of things," *IEEE Transactions on Communications*, Vol. 69, no. 7, 2021, pp. 4504–4519.
7. P. Dong, X. Du, J. Sun, and H. Zhang, "Energy-efficient cluster management in heterogeneous vehicular networks," in *2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2016, pp. 644–649.
8. W. Benrhaiem, A. Hafid, and P. K. Sahu, "Reliable emergency message dissemination scheme for urban vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 21, no. 3, 2020, pp. 1154–1166.
9. Y. Zhu, Q. Zhao, and Q. Zhang, "Delay-constrained data aggregation in vanets," *IEEE Transactions on Vehicular Technology*, Vol. 64, no. 5, 2015, pp. 2097–2107.
10. G. Liu, N. Qi, J. Chen, C. Dong, and Z. Huang, "Enhancing clustering stability in vanet: A spectral clustering based approach," Vol. 17, no. 4, 2020, pp. 140–151.
11. J. Cheng, G. Yuan, M. Zhou, S. Gao, Z. Huang, and C. Liu, "A connectivity-prediction-based dynamic clustering model for vanet in an urban scene," Vol. 7, no. 9, 2020, pp. 8410–8418.
12. B. Paramasivan, M. Bhuvaneswari, and K. M. Pitchai, "Augmented dtn based energy efficient routing protocol for vehicular ad hoc networks," in *2015 IEEE SENSORS*, 2015, pp. 1–4.
13. H. Zhou, W. Xu, J. Chen, and W. Wang, "Evolutionary v2x technologies toward the internet of vehicles: Challenges and opportunities," *Proceedings of the IEEE*, Vol. 108, no. 2, 2020, pp. 308–323.
14. K. Zhang, S. Leng, X. Peng, L. Pan, S. Maharjan, and Y. Zhang, "Artificial intelligence inspired transmission scheduling in cognitive vehicular communications and networks," *IEEE Internet of Things Journal*, Vol. 6, no. 2, 2019, pp. 1987–1997.
15. Y. Hui, Z. Su, T. H. Luan, and J. Cai, "Content in motion: An edge computing based relay scheme for content dissemination in urban vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 20, no. 8, 2019, pp. 3115–3128.
16. Y. Hui, Z. Su, and T. H. Luan, "Collaborative content delivery in software-defined heterogeneous vehicular networks," *IEEE/ACM Transactions on Networking*, Vol. 28, no. 2, 2020, pp. 575–587.
17. F. Abbas, P. Fan, and Z. Khan, "A novel low-latency v2v resource allocation scheme based on cellular v2x communications," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 20, no. 6, 2019, pp. 2185–2197.
18. C.-H. Fang, K.-T. Feng, and L.-L. Yang, "Resource allocation for urllc service in in-band full-duplex-based v2i networks," *IEEE Transactions on Communications*, Vol. 70, no. 5, 2022, pp. 3266–3281.

19. J. Garbiso, A. Diaconescu, M. Coupechoux, and B. Leroy, "Fair self-adaptive clustering for hybrid cellular-vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 22, no. 2, 2021, pp. 1225–1236.
20. X. Wang, Z. Ning, X. Hu, L. Wang, B. Hu, J. Cheng, and V. C. M. Leung, "Optimizing content dissemination for real-time traffic management in large-scale internet of vehicle systems," *IEEE Transactions on Vehicular Technology*, Vol. 68, no. 2, 2019, pp. 1093–1105.
21. S. Ucar, S. C. Ergen, and O. Ozkasap, "Multihop-cluster-based ieee 802.11p and lte hybrid architecture for vanet safety message dissemination," *IEEE Transactions on Vehicular Technology*, Vol. 65, no. 4, 2016, pp. 2621–2636.
22. C. Shi, Y. Zhou, W. Li, H. Li, N. Lu, N. Cheng, and T. Yang, "A centralized clustering based hybrid vehicular networking architecture for safety data delivery," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–6.
23. A. Memedi, F. Hagenauer, F. Dressler, and C. Sommer, "Cluster-based transmit power control in heterogeneous vehicular networks," in *2015 IEEE Vehicular Networking Conference (VNC)*, 2015, pp. 60–63.
24. Y. Cao, R. Ji, L. Ji, G. Lei, H. Wang, and X. Shao, "l^2-mptcp: A learning-driven latency-aware multipath transport scheme for industrial internet applications," *IEEE Transactions on Industrial Informatics*, 2022.
25. W. Qi, B. Landfeldt, Q. Song, L. Guo, and A. Jamalipour, "Traffic differentiated clustering routing in dsrc and c-v2x hybrid vehicular networks," *IEEE Transactions on Vehicular Technology*, Vol. 69, no. 7, 2020, pp. 7723–7734.
26. C.-M. Huang and C.-F. Lai, "The delay-constrained and network-situation-aware v2v2i vanet data offloading based on the multi-access edge computing (mec) architecture," *IEEE Open Journal of Vehicular Technology*, Vol. 1, 2020, pp. 331–347.
27. F. Shabnam and A. Jamalipour, "An efficient coordinator selection method for geo-routing protocol in vehicular network," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, 2020, pp. 1–5.
28. F. Lyu, H. Zhu, N. Cheng, H. Zhou, W. Xu, M. Li, and X. Shen, "Characterizing urban vehicle-to-vehicle communications for reliable safety applications," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 21, no. 6, 2020, pp. 2586–2602.
29. W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, 2000.

Gao-Feng Hong , born in Nanchang, Jiangxi province in 1995, is now a doctoral student at Beijing Jiaotong university. He is mainly engaged in the research of the vehicular network and mobile edge cloud computing.

Wei Su , born in October 1978. He got the Ph.D. degrees in Communication and Information Systems from Beijing Jiaotong University in January 2008. Now he is a teacher in the School of Electronics and Information Engineering, Beijing Jiaotong University. He granted the title of professor in November 2015. Dr. Su has published more than 50 research papers and four monographs in areas of communications and computer networks. His research interests are next generation network and mobile Internet. He won the Second Prize of the National Technology Invention Award in 2014.

Xin-Di Hou , received the M.S. degree in information and communication engineering from Northeast Electric Power University, Jilin, China, in 2018. He is currently pursuing the Ph.D. degree in information and communications engineering at Beijing Jiaotong University. His specic areas of research interest mainly focus on network routing and addressing, network integration, and future Internet architecture.

Li-Sheng Ma , received the B.S., M.S. and Ph.D. degrees from Taiyuan Normal University, China in 2004, from Southwest University, China in 2007 and from Future University Hakodate, Japan in 2017, respectively. He is currently an associate professor of Chuzhou University, China. His research interests include network survivability, emergency communication and management.