

## Research Article

# Message Relaying and Collaboration Motivating for Mobile Crowdsensing Service: An Edge-Assisted Approach

Shu Yang , Jinglin Li , Quan Yuan , Zhihan Liu , and Fangchun Yang

State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China

Correspondence should be addressed to Jinglin Li; [jlli@bupt.edu.cn](mailto:jlli@bupt.edu.cn)

Received 11 April 2018; Accepted 24 June 2018; Published 29 July 2018

Academic Editor: Kok-Seng Wong

Copyright © 2018 Shu Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Group sensing is a kind of crowdsensing service where HD map producers motivate private cars in a local region to collect data from real world. Group sensing needs vehicles to communicate physically and drivers to collaborate strategically in a mobile or edge-assisted environment. First, we consider collaboration module that motivates drivers to be participants; centralized and distributed motivating methods are discussed. Secondly, we consider communication module; two VANET-based methods are proposed to achieve message relaying in edge infrastructure. To accomplish participants' selection, three combinations of two modules are proposed and simulated based on a flexible framework. The results show that centralized selection could motivate collaboration at a low price but brings heavy communication overhead. Clustered selection requires more incentives and less communication overhead than centralized selection. Distributed selection is usually the first class choice because of its fine performances on both communicating and motivating.

## 1. Introduction

The rise of big data fuels the research of machine learning and data mining algorithms. Besides algorithms, the method of collecting data is also of great importance [1]. Being intellectualized and networked, smart vehicles are able to sense and communicate in urban area as edge infrastructures [2–4]. By means of crowdsensing, their intrinsic mobility can be leveraged to dynamically collect urban data in different time and areas [5].

A promising service of crowdsensing is collecting environmental data for building HD (High Definition) map [6]. The HD map producers, such as Here, TomTom, and Baidu, need lidar/camera/IMU data to build a live map for autonomous driving [7, 8]. The huge volume of information as well as fast updating requirement to build a “live” map challenges map producers because their own devices undoubtedly could not meet these requirements. Some researchers therefore proposed crowdsensing, in which private cars are incentivized to accomplish a sensing task and upload data to map producer. The incentives could be either real or virtual money.

This paper focuses on how to call up a group of smart vehicles to complete a sensing task in an edge environment,

where vehicles and RSUs work together to provide support for reliable mobile services [9]. Map producer launches a task by putting it on a “seed vehicle” and will propagate task information to others in the neighborhood. Then, nearby vehicles communicate with each other by VANET (Vehicular Ad hoc NETWORKS), forming a group and sensing one segment of road together. We do not concentrate on specific data collecting method (*e.g.*, how to elicit depth from camera data). Instead, we stand in a perspective of crowdsensing mechanism and we are interested in methods that enable vehicles to communicate and collaborate.

As Figure 1 illustrated, we design a basic system composed of message relaying module and collaboration motivating module to enable vehicles to communicate physically and collaborate strategically.

One feature of our system is its assumption where vehicles only communicate by VANET rather than cellular networks. Admittedly, allocating sensing task by cellular networks enables map producers to directly know and control their “sensors” from servers on the cloud. This scheme needs drivers to report their fine-grained trajectories so that map producers could allocate or recommend an appropriate task. However, a majority of drivers are conservative rather than

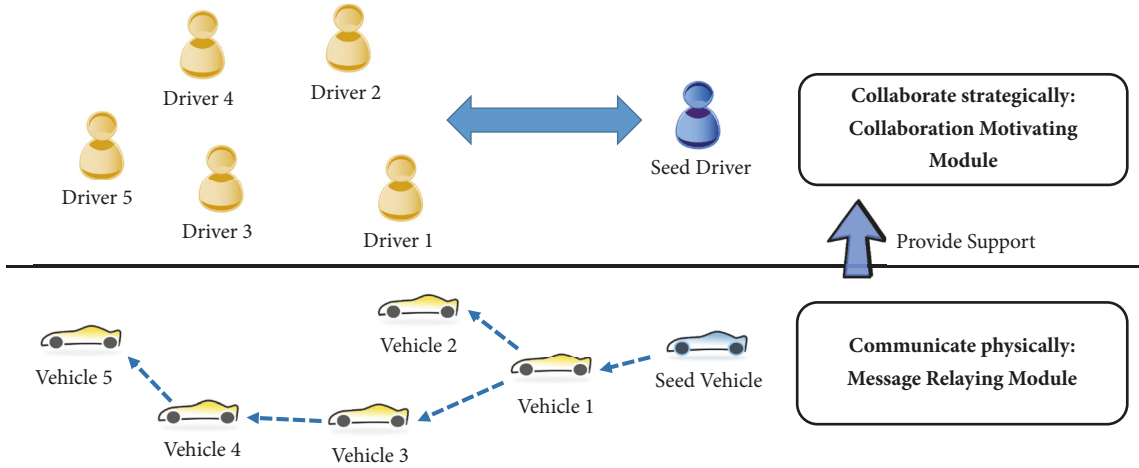


FIGURE 1: A schematic of group sensing problem. The problem is tackled by two modules.

collaborative if it comes to their private trajectories [10]. The cellular network based scheme therefore may not be fully accepted by most drivers. By contrast, our VANET based local participants selection could alleviate this problem by propagating task information from a seed vehicle and allows drivers to choose whether to join the task with no need of reporting fine-grained trajectories.

Another feature of our system is its game theory model. Vehicles are actually controlled by drivers whose behaviors are self-interest [11, 12]. The collaboration among drivers is modeled as a threshold-based game. We will see how vehicles' sensing costs, the numbers of vehicles, and the task incentives influence the collaboration's equilibrium.

Based on the basic system model, we mainly introduce and discuss the two following questions about group sensing:

(i) *How does task information propagated via VANET?*

Message relaying module is designed to consist of two processes, spread process and back process. In spread process, a message that contains basic task information is firstly given out by seed vehicle and subsequently relayed by other vehicles until it arrives at a region border specified by map producers. We call this spread process because it aims to spread task information to all vehicles within the region. Then it comes to back process; the message is modified and relayed back to seed vehicle. During back process, vehicles may add new information to the message.

(ii) *How could two modules work together?*

Collaboration motivating module helps drivers form a sensing group. Collaboration is relying on message relaying module underpinned to achieve information acquiring and propagating. Existing researches usually discuss relaying message and motivating collaboration separately while an integrated framework is discussed in this paper. The aim of module combination is to endow each networked vehicle ability to make appropriate decision. This combination should meet two requirements: first, the vehicle is automatic

enough at control or communication level so that one meta-action could be executed fast enough to support high-level collaboration; second, to maintain the high-level collaboration, vehicle should work smoothly with human driver. Upon these requirements, we have actually added human factor to the collaborating process among crowdsensing.

This paper tries to answer these questions based upon the following assumptions and simplifications:

- (i) We assume that each vehicle is equipped with Dedicated Short Range Communication (DSRC) device so that it could send or receive messages within a certain distance. We also assume that a vehicle could well coordinate with driver by advanced HMI technology, which means that two modules can seamlessly cooperate.
- (ii) Each driver is assumed to be rational and self-interest. He makes decision to maximize his profits. For each driver, the cost of participating is assumed to conform normal distribution.
- (iii) To concentrate on a perspective of map producers, we slightly oversimplify the physical layer by assuming perfect channel without collision and loss in simulation. This is detailed in Section 5.

The rest of this paper is organized as follows. The discussion of related work is presented in Section 2. Section 3 illustrates a framework of two modules. Section 4 describes how two modules work interactively to achieve local participants' selection. Section 5 presents simulation results and analysis. The paper ends with discussions in Section 6 and conclusions in Section 7.

## 2. Related Work

Leveraging vehicles or sensor networks for environmental sensing has been a hot topic in the recent years. Zhu [13] proposed PUS (Pervasive Urban Sensing) framework and

used probe cars to sense traffic density. The authors designed a compressive sensing algorithm to tackle sparsity of data. Chen [14] considered the distributed estimation problem over relay-assisted wireless sensor networks. Kalman filtering and consensus estimation are leveraged to improve the estimation efficiency and accuracy in large monitoring system. Yuan [15] observed that the distribution of probe vehicles is uneven over space and time. He therefore proposed an adaptive and compressive data gathering scheme based on matrix completion theory. The scheme could largely improve sensing efficiency and quality. Likewise, Du [16] employed matrix completion to estimate unsampled urban data. Besides data processing, there are some works focusing on optimal participants' selection of crowdsensing. Song [17] aims to select the most appropriate participants with different budget constraints. Considering different incentive requirements, associated sensing capabilities, and uncontrollable mobility, a multitask-oriented QoI (Quality of Information) optimization problem is discussed and converted to a nonlinear knapsack problem. Zhang [18, 19] proposed an event-driven QoI-aware participatory sensing framework with energy and budget constraints, where the main method is boundary detection. He [20] devised an efficient local ratio based algorithm and designed a motivating mechanism that decides the fair prices of sensing tasks. In most papers, multitask crowdsensing is modeled as a global optimization problem with need for participants' fine-grained trajectories. By contrast, our paper tries to call up plenty of participants for a single task in a specified region without reporting participants' trajectories.

The application of VANET has been wildly discussed. Papers in research community simply fall into two categories: safety-oriented and non-safety-oriented. The safety-oriented applications intend to avoid collision by broadcasting beacons. Taha [21] proposed a novel protocol for reliable broadcasting of life safety messages in VANET. In case of any dramatic change of speed or moving direction, the vehicle is regarded abnormal and hence it transmits an emergency warning message over the control channel of DSRC. Kumar [22] proposed CarSpeak, a communication system for autonomous driving modified from VANET. By changing the MAC protocol in a content-centric approach, CarSpeak scales the amount of data with the available bandwidth and improves safety of autonomous driving. The non-safety-oriented applications try to relay traffic or other non-real-time information to improve traffic efficiency or driving entertainment. To support high throughput in vehicular networks, Zhou [23] used a geolocation database assisted approach to jointly utilize DSRC and TVWS spectrum. It not only satisfied the dynamic vehicular access requirements but also has adaptive data piping performance. Yuan [24] used VANET for detecting traffic congestion on urban expressways. The scheme develops a spatiotemporal effectiveness model based on the potential energy theory to control the dissemination area and survival time of the congestion information. Similarly, Luo [25, 26] used a SDN-inspired approach to improve network capacity. Wang [27], Zhou [28], Cheng [29], and Yuan [30] use "hard" or "soft"

method to expand the communication capacity of Internet of vehicles, respectively.

To improve trustworthiness of VANET information, Chen [31] presented a trust-based message propagation and evaluation framework in VANET, where peers share information regarding road condition or safety and others provide opinions about whether the information can be trusted. Chang [32] proposed a Sybil attack detection mechanism, where footprint of vehicle is employed. Vehicles can generate location-hidden trajectory for location-privacy-preserved identification. PTRS [33] is a robust trust-based relay scheme with an objective to distinguish trust levels of the vehicles.

Our work is also related to MAS (Multiagent System). Supported by today's advancing cloud computing technology, the vehicle could be represented by its virtual agent [34] in Cyber-Physical System. One goal of MAS is to find algorithms to allocate tasks and to design mechanisms to motivate collaboration among rational agents. Based on Stackelberg game, Duan [35] analyzed and compared different incentives mechanisms for a master to motivate the collaboration of smartphone users on both data acquisition and distributed computing applications. More specifically, he proposed a reward-based collaboration mechanism, where the master announces a total reward to be shared among collaborators, and the collaboration is successful if there are enough users wanting to collaborate. The mechanism is novel and practical in many cases including participants' selection of group sensing. Walid [36] proposed a game theoretical approach to tackle the problem of distributed formation of the uplink tree structure among the relay stations in a WiMAX network. From a MAS perspective, the authors modeled the problem as a network formation game, where each relay station aims to maximize its utility from cooperation. Using the proposed dynamics, the relay stations can self-organize into the tree structure. Talal [37] proposed an algorithm enabling multiagent cooperative calculation. Each agent is assigned some part of the calculation such that the agents' shares are exhaustive and disjoint. The algorithm is decentralized and requires no communication between the agents. Jiang [38] considered task allocation and load balance in social networked MAS, where agents always need to negotiate with other agents about their resources. In the presented task allocation model, while a task comes to the system, it is first assigned to a principal agent that has high contextual enrichment factor for the required resources. Then the principal agent will negotiate with its contextual agents to execute the assigned task. The authors also considered load balancing to avoid overconvergence of tasks at certain agents that are rich in contextual resources. Motivated by above papers, this paper has discussed smart vehicle collaboration from a MAS perspective.

### 3. Framework

Group sensing aims to collect data in a specified region, thus focusing on local participants' selection. The participants' selection of group sensing involves two modules. Collaboration motivating module aims to motivate plenty of collaborators to form a sensing group. The math of motivating

mechanism will be illustrated. Another module, namely, message relaying module, provides VANET-based media to execute motivating mechanism.

**3.1. Collaboration Motivating Module.** HD map data need to be batch-processed to reduce noise and correct error. One single vehicle's data are less valued than a batch of data from a sensing group. Therefore, a sensing group usually consists of many vehicles that share a similar traverse in both space and time. Map producers prefer to harvest a batch of data by calling up a group of vehicles.

Based on above facts, the essence of group sensing is threshold based game [35]. Map producers are only willing to give incentives when a certain number of vehicles (or drivers), denoted as  $n_0$ , participate in the sensing task. Consider that a set of  $N$  vehicles  $\mathbb{N} = \{1, 2, \dots, N\}$  have been notified of the task and each vehicle knows  $N$  and  $n_0$ . *vehicle<sub>i</sub>* ( $i \in \mathbb{N}$ ) has a collaboration cost  $C_i$ , if it collaborates (*i.e.*, joins the sensing group). We assume that all collaboration costs conform normal distribution, known as  $C \sim N(\mu, \sigma)$ .

Since map producers need a sensing group, the sensing task will not be approved unless at least  $n_0$  of  $N$  vehicles become participants. Moreover, each driver decides whether to participate by predicting others' decisions. These facts make a threshold-based game, which consists of two phases. In phase 1, seed vehicle gives out task information  $(R, n_0)$ .  $R$  are total incentives given by map producers. In phase 2, after  $(R, n_0)$  is propagated and known by  $N$  vehicles, each driver decides whether to participate. If the number of participants, denoted by  $n$ , is bigger than  $n_0$ , then the task is approved; otherwise the task is abandoned by map producer. If the task is approved, incentives  $R$  are distributed equally to participants because each participant works in the similar road segment and time. As a result, participating *vehicle<sub>i</sub>* ( $i \in \mathbb{N}$ ) undertakes cost  $C_i$  and gets incentives of  $R/n$ , with its overall profits being

$$\left(\frac{R}{n} - C_i\right) \text{Pos}\{n \geq n_0\} \quad (1)$$

where  $\text{Pos}\{X\}$  is the indicator function (equals 1 when event  $X$  is true and equals 0 when event  $X$  is false). One could make negative profits even though he joins the collaboration, if  $R/n$  is smaller than  $C_i$ . If collaboration is not approved, profit is 0.

It is actually the driver who decides whether the vehicle collaborates or not. Driver is self-interest and makes own decision with limited information. All drivers' rational behaviors make a threshold-based game. The result of this game is subtle; for example, when  $n$  ( $n \gg n_0$ ) drivers decide to collaborate, they may even gain a negative profit because  $R/n$  is so small that it even could not meet sensing cost. On the other hand, if only  $n$  ( $n \ll n_0$ ) drivers collaborate, it is hard to persuade map producers to approve such task because the map producers could not harvest enough data. Therefore, how to motivate an appropriate number of drivers to collaborate is a critical problem in group sensing.

**3.2. Message Relaying Module.** Message relaying module executes two physical processes: spread process and back

TABLE 1: Content of task/collaboration message.

Task message	Collaboration message
Current relaying vehicle's location $(x, y)$	
Seed vehicle's location $(x_0, y_0)$	
Potential participants list $\mathbb{L}$	
Reward $R$	Participants list $\mathbb{L}'$
Threshold of participants' number $n_0$	/
Sensing segment: a rectangle by $(x_1^s, y_1^s), (x_2^s, y_2^s)$	/
Preparing segment: a rectangle by $(x_1^p, y_1^p), (x_2^p, y_2^p)$	/

process. Map producers have owned a few seed vehicles. Reporting their fine-grained trajectories, seed vehicles are well connected with servers and controlled by map producer. Server monitors seed vehicles' states and launch a task package to one specific seed vehicle. As Table 1 shows, the package contains four kinds of information: (1) The first kind is sensing segment. Participants only need to collect data within a specified region. The region is the sensing segment of task. (2) The second kind is preparing segment. Before entering sensing segment, task information is propagated and some participants are generated from normal vehicles. Such process, including message relaying and collaboration motivating, takes place on preparing segment. As shown in Figure 2, preparing segment should be upstream of sensing segment because vehicles would have already traversed a length of road segment when preparing work is done. Since our topic is communicating and collaborating mechanism rather than sensing techniques, this paper concentrates on preparing segment. (3) The third kind is incentive. Map producer is willing to pay incentives  $R$  if he approves the task. (4) The fourth kind is threshold of participants. The task is successfully approved only when number of participants  $n$  meets the threshold  $n_0$ .

Figure 2 also illustrates spread process and back process. Spread process starts with a TM (task message) broadcast by seed vehicle. A TM is relayed one-by-one until it reaches the border of preparing segment; thus all vehicles in preparing segment are notified of the task. The border guard, a vehicle located at the border, modifies TM to CM (collaboration message). He also starts back process by reversing relaying direction and sending CM back to seed vehicle. During back process, the vehicle that receives CM decides whether to join the group sensing and adds its decisions to CM. Seed vehicle receives the final CM and knows how many participants it has motivated. The server then could decide if the task is approved.

Note that, in our threshold-based game, all drivers make decisions by predicting others' decisions. To give participants list  $\mathbb{P}$  to seed vehicle, decisions are recorded in  $\mathbb{P}$ , which is relayed by vehicles. This means that some vehicles could "peep" decisions of former vehicles, which could lead to more complex models (*e.g.*, social learning [39]) and fails our threshold-based game model. To prevent this cheating, seed



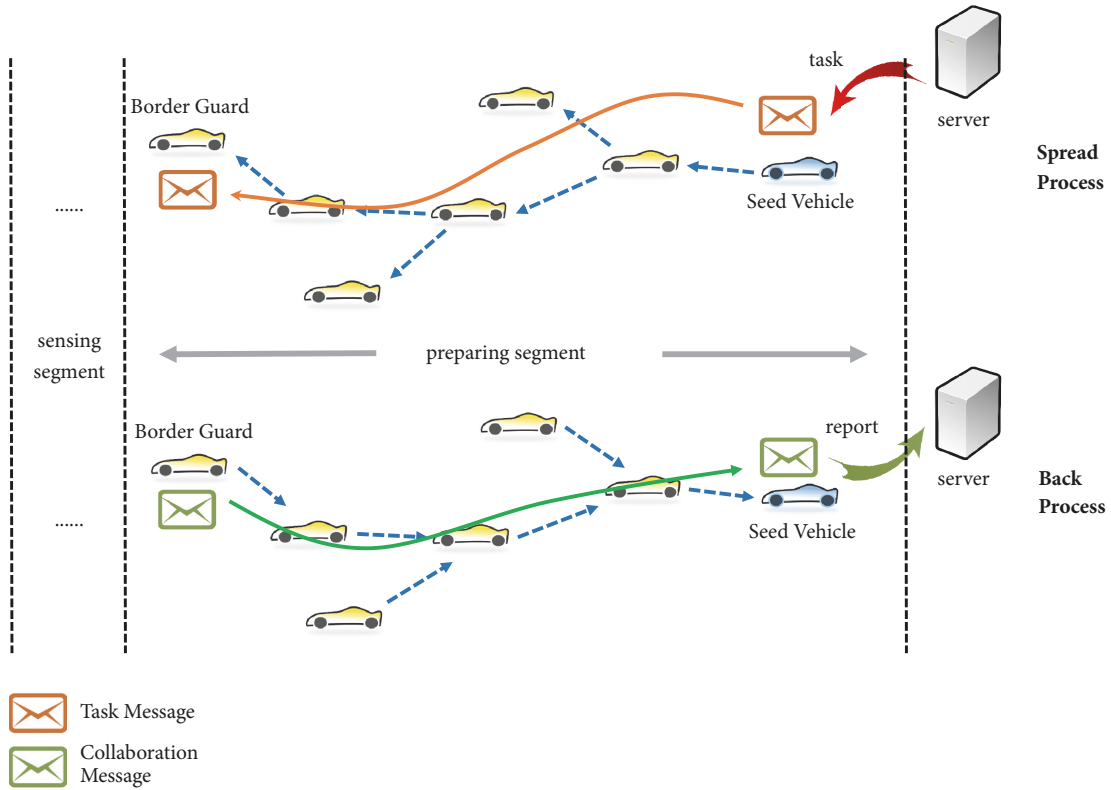


FIGURE 2: Illustration of spread process and back process.

and normal vehicles can use public-key cryptography [40] to tackle this problem. Normal vehicles encrypt their decisions with public key, while only seed vehicle can decipher the decisions with private key.

**3.3. Combining Modules with Flexibility.** Participants' selection needs two modules, namely, message relaying module and collaboration motivating module, to work together. The former module is in charge of propagating information among vehicles, and the latter module is responsible for coordinating drivers' behaviors.

In VANET, messages could be propagated by contention-based relaying. That is, only the periodical winner of contention could be relaying node. Another relaying method is built upon cluster structure, which helps vehicles reduce transmission data volume and improve communication efficiency. Supported by message relaying module, group formation still needs a mechanism to motivate and select appropriate participants. The motivating method could be realized in either a centralized or distributed way. Overall, our participants' selection framework is composed of two modules, and there are two alternative methods for each module, respectively.

Decoupling participants' selection into two modules endows the system flexibility, enabling map producers to choose optimal combination from several alternatives according to service requirements. We will discuss three combinations: centralized selection, distributed selection, and clustered selection. See Figure 3 for understanding how

the combination of different modules can make three distinct participants' selection schemes.

#### 4. Participants' Selection of Group Sensing

This section presents participants' selection schemes derived from one framework. Before diving into three schemes, some preliminaries on relaying methods should be introduced.

We only consider message relaying in "simple road condition." Simple road condition is referred to (1) the road being straight or having small curvature so that relaying method would not fail by extreme road shapes, such as zigzag or circle, and (2) neglecting intersection in preparing segment. In this way, we could simplify simulation by leaving out the situations where lots of vehicles join and leave. Furthermore, we could simplify simulation by decomposing complex road condition (zigzag, circle, etc.) into several simple road conditions. This divide-and-conquer strategy enables our proposed method to be adjustable in most transportation situations. Two examples of simple road condition are in Figure 4.

Contention-based relaying is motivated by geocasting [41], where message is relayed from original location  $P_0$  to  $P_{n+1}$  by a node chain. As described in Figure 5(a), seed vehicle is located at  $P_0(x_0, y_0)$ , previous relaying node's location is  $P_n(x_n, y_n)$ , and current receiver's location is  $P_{n+1}(x_{n+1}, y_{n+1})$ . Then we have  $L_{0,n} = \|P_0P_n\|$  and  $L_{n,n+1} = \|P_nP_{n+1}\|$ . The angle  $\theta = \angle P_0P_nP_{n+1}$  is calculated by the following equation:

$$\cos \theta = \frac{L_{0,n}^2 + L_{n,n+1}^2 - L_{0,n+1}^2}{2L_{0,n}L_{n,n+1}} \quad (2)$$

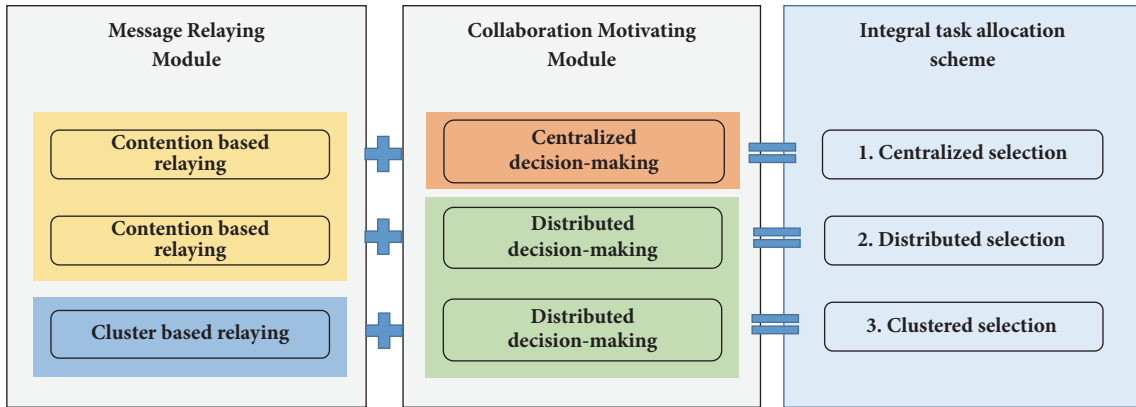


FIGURE 3: Three participants' selection schemes. Each scheme is a combination of two modules. Each module includes two methods.

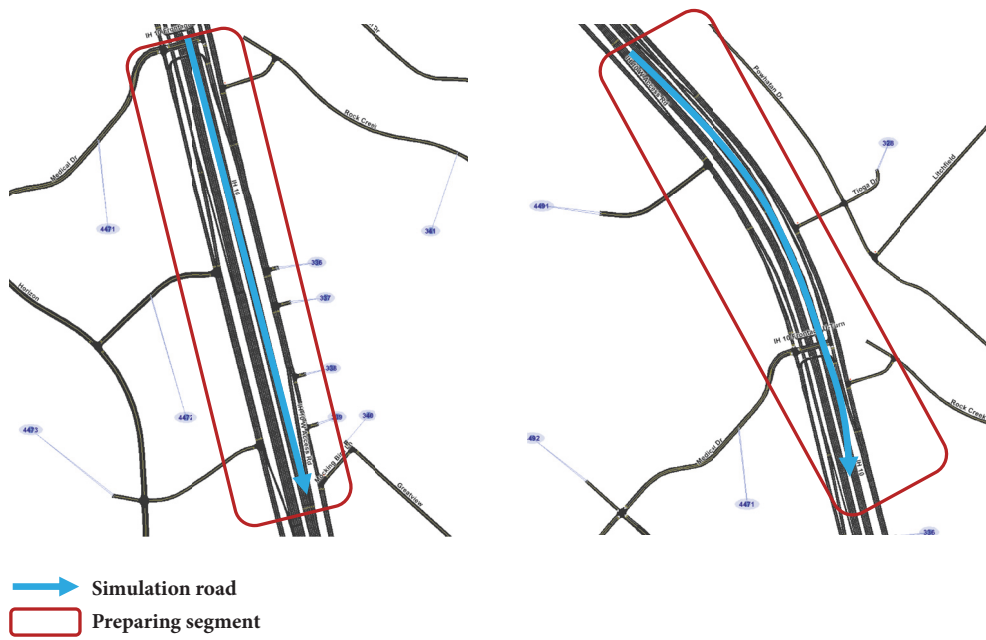


FIGURE 4: Simple road condition. Left: straight road. Right: road with small curvature. Both examples do not consider intersections.

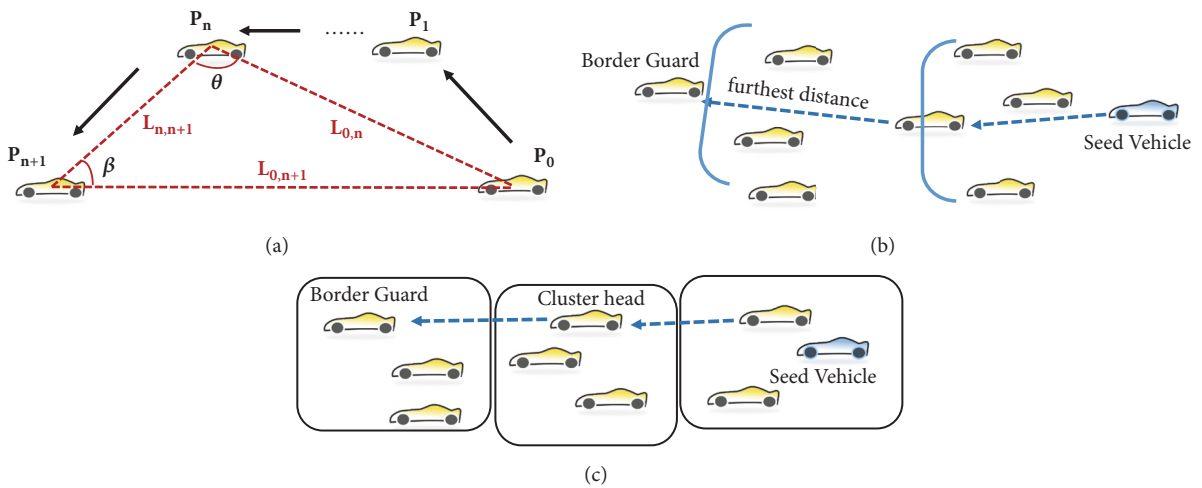


FIGURE 5: Message relaying methods. (a) Node chain. (b) Contention-based relay. (c) Cluster-based relay.

- (1) Seed vehicle broadcasts TM.
- (2) After receiving TM,  $vehicle_i$  calculates  $\cos \theta$ .  
If  $\cos \theta > 0$ , goto (6).
- (3)  $vehicle_i$  calculates  $L'$ , then contends to be relaying node. If fails to be, goto (6).
- (4)  $vehicle_i$  claims himself the new relaying node.
- (5) If  $vehicle_i$  is not on the border, it relays TM to downstream.
- (6) end.

PROCESS 1: Contention-based relaying.

It indicates that message is relayed away from original node  $P_0$  if  $\cos \theta < 0$ . The next relaying node is generated by downstream nodes' contention. The projection of  $L_{n,n+1}$  on  $L_{0,n}$  is  $L' = L_{n,n+1} \times \|\cos \beta\|$ . Each downstream node calculates its back off time  $t_{notifying}$  by [42]

$$t_{notifying} = T_{max} \cdot \left(1 - \frac{L'}{L_{max}}\right) \quad (3)$$

where  $T_{max}$  and  $L_{max}$  are predefined values and  $\cos \beta$  could be calculated in a similar way to  $\cos \theta$ . According to (3), the most downstream node has the minimum  $t_{notifying}$ , thus having the highest priority to be next relaying node. Figure 5(b) illustrates how messages are delivered by winners of contention. Usually, a message could reach the destination by several hops.

Another prevailing relaying method is cluster-based relaying. A virtual network infrastructure should be created through the clustering of nodes in order to provide stability and scalability. See Figure 5(c) for an illustration of cluster-based relaying. Each cluster has a cluster head, which is responsible for intra- and intercluster coordination in the network management functions. Nodes inside a cluster communicate by direct links.

**4.1. Distributed Selection.** Distributed selection is a combination of contention-based relaying and distributed decision-making. As Process 1 illustrated, TM is propagated by contention-based relaying to avoid broadcast storm and ensures that all drivers in assigned local region receive TM.

Then TM is modified to CM by border guard; all drivers decide whether to be participants in a distributed manner; the process of distributed decision-making is a game, where all players have the following equilibrium:

- (i) No collaboration: if  $R < n_0\mu$ , the reward is not enough to satisfy driver's expectation. No driver is willing to join collaboration
- (ii) Mix strategy equilibrium: if  $n_0\mu \leq R \leq N\mu$ , each driver joins collaboration by probability  $p^*$ , where  $p^*$  is the unique solution of

$$E_m \left( \left( \frac{R}{m+1} - \mu \right) Pos(m+1 \geq n_0) \right) = 0 \quad (4)$$

$m$  follows binomial distribution  $B(N-1, p)$ . Equation (4) indicates that  $p^*$  is an equilibrium where each driver has an expected payoff of 0.

Furthermore, we also consider the scenario where each player knows his own cost  $c_i$ . Then there exists a stronger constraint than (5); a driver will only collaborate when  $C_i \leq \gamma^*(R)$ .  $\gamma^*(R)$  is unique solution of the equation

$$E_m \left( \left( \frac{R}{m+1} - \gamma \right) Pos(m+1 \geq n_0) \right) = 0 \quad (5)$$

$m$  follows binomial distribution  $B(N-1, F(\gamma))$ . Note that  $F(\gamma) = P(C_i \leq \gamma)$ . We can learn that all members have the same information and estimation about others. Therefore, all members have the same  $\gamma^*(R)$ . However, different members will still make different decisions according to their private cost.

Finally, if a driver decides to collaborate, he sends a decision message to seed vehicle by contention-based relaying again.

**4.2. Centralized Selection.** Centralized selection is composed of contention-based relaying and centralized decision-making. It means that all potential participants send their messages into seed vehicle that collects all messages and selects participants in a centralized manner.

First, task message is propagated by process 1. After that, willing driver simply sends a message that includes his private cost and other information to seed vehicle.

Seed vehicle makes decisions for all drivers, since it knows all drivers cost set  $\mathbb{C} = \{C_{i_1}, C_{i_2}, \dots, C_{i_N}\}$ , which could be contained in  $\mathbb{P}$ . Let us just assume that there is no same cost value among cost set. Then costs could be sorted in ascending order in  $\mathbb{C}' = \{C_1, C_2, \dots, C_N\}$ . Seed vehicle will choose  $1^{st}$  to  $n_0^{th}$  smallest cost in  $\mathbb{C}'$  to be participants. The collaboration has optimal result under centralized decision-making:

- (i) Optimal collaboration: each driver with  $C_i \leq C_{n_0}$  will be chosen and afterwards receives  $C_{n_0}$  incentives. The remaining  $N - n_0$  drivers do not collaborate.

Each selected participant will receive  $C_{n_0}$  incentives, since we distribute incentives equally. By centralized selection, no driver needs to send his decision, since seed vehicle has already known the collaboration results.

**4.3. Clustered Selection.** Clustered selection scheme has a different relaying method from above two schemes. As Figure 3 illustrates, the main difference between clustered selection and the other two lies in the question of how messages are relayed. Clustered selection maintains a hierarchical structure, including cluster head and cluster members. Cluster head takes responsibility for intercluster communication, thus reducing message collision and improving network efficiency.

In this case, decision-making is still distributed. The integral scheme is described in process 2. The formation of cluster has been widely researched [43] and thus is not discussed in this paper.

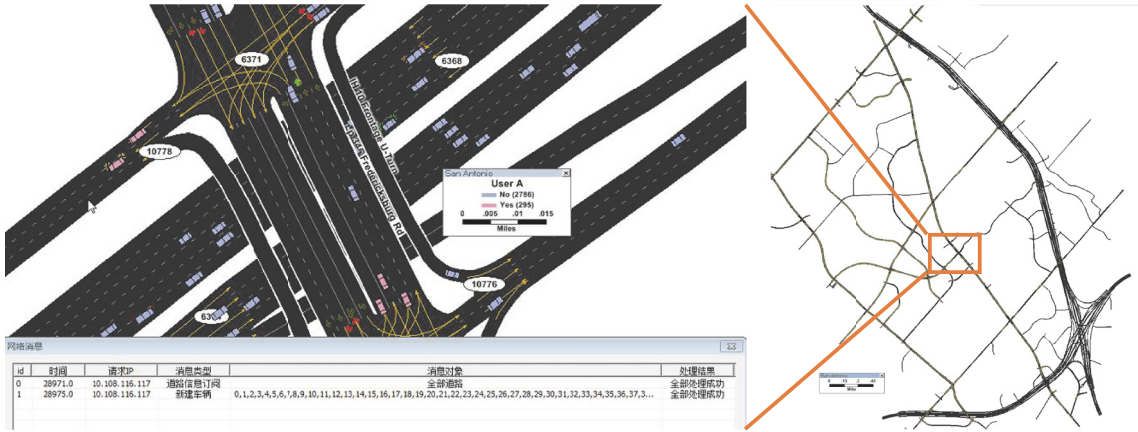


FIGURE 6: TransModeler provides a virtual transportation scenario.

- (1) Seed vehicle broadcasts TM.
- (2) After receiving TM, cluster head  $vehicle_i$  informs cluster members the TM. It also calculates  $\cos \theta$ . If  $\cos \theta > 0$ , goto (4).
- (3) If cluster head  $vehicle_i$  is not on the border, it relays TM to a cluster head of downstream.
- (4) end.

PROCESS 2: Cluster-based relaying.

Finally, all members are informed of task information. Being similar to Section 4.1, distributed decision-making is leveraged to select appropriate participants. The results of decisions are sent back to seed vehicle by cluster-based relaying again.

## 5. Simulation and Analysis

We simulate the framework with TransModeler, a transportation simulating environment. Figure 6 is an illustration of our simulator. Based on its API, we achieve logical communication framework over vehicles to support network layer simulation. Additionally, we assume that each relaying consumes the same time slot and assume perfect channel without collision and loss in simulation. With such framework, we could simulate network layer of VANET communication in a virtual transportation scenario. Since the framework that integrates message relaying and collaboration motivating has been little discussed previously, no previous work could be compared with it. However, we still make comparison among our self-proposed methods and try to make this paper self-contained.

**5.1. Effect Analysis of Motivating.** We first analyze two motivating methods separately. Parameters are explained in Table 2 and the simulation results are demonstrated in Figure 7.

**5.1.1. Result of Distributed Decision-Making.** Figure 7(a) shows  $R - \rho^*$  diagram depicting how driver's collaboration

TABLE 2: Parameters in motivating simulation.

$R$	Reward or incentive
$N$	The number of potential participants
$\mu$	Mean of all potential participants' cost
$\sigma$	Std. of potential participants' cost
$n_0$	Threshold of participants' number
$M$	The number of final participants

probability reacts with reward. When  $R < 20$ , reward is smaller than  $n_0\mu$ ; no driver collaborates. When  $20 \leq R < 35$ , reward is enough to motivate part of drivers. And it shows that the larger reward is, the higher probability is that drivers will collaborate. When  $R \geq 35$ , that is,  $R \geq N\mu$ , all drivers will collaborate; thus  $p^* = 1$ . Similarly, Figure 7(b) shows that larger  $\mu$  leads to lower collaboration probability. It means that, given the same reward  $R$ , fewer drivers will collaborate if the mean of costs goes higher. These findings in two figures are consistent with human intuition.

**5.1.2. Result of Centralized Decision-Making.** Figure 7(c) describes  $R$ 's influence on  $M$ . As  $R$  increases,  $M$  also increases. More reward could motivate more participants. In Figure 7(d), we could see that higher  $\mu$  leads to smaller  $M$ . Figure 7(e) shows that larger  $\sigma$  (which means that the costs values are more dispersed) leads to smaller  $M$ . In the case in Figure 7(e),  $n_0 = 40 > N/2$ , so it needs to motivate more than  $N/2$  drivers. With the increase of  $\sigma$ ,  $n_0^{\text{th}}(n_0 > N/2)$  smallest cost  $C_{n_0}(C_{n_0} - \mu > 0)$  becomes bigger, and the value of  $R/C_{n_0}$  gets smaller; thus participants' number  $\lfloor R/C_{n_0} \rfloor$  decreases.

**5.1.3. Comparison of Two Motivating Methods.** We set  $N = 70$ ,  $n_0 = 40$ ,  $R = 20$ ,  $\mu = 0.5$ , and  $\sigma = 0.1$  and use  $M$  as indicator of motivating effect. The simulation is repeated for 20 times. Figure 7(f) shows that centralized decision-making achieves larger  $M$  than distributed decision-making.

It should be mentioned that a task is approved by map producer only when  $M \geq n_0 = 40$ . However, as Figure 7(f) shows, no method could achieve  $M \geq n_0$ . For centralized decision-making, map producer should prepare  $R' = 22 \sim 23$



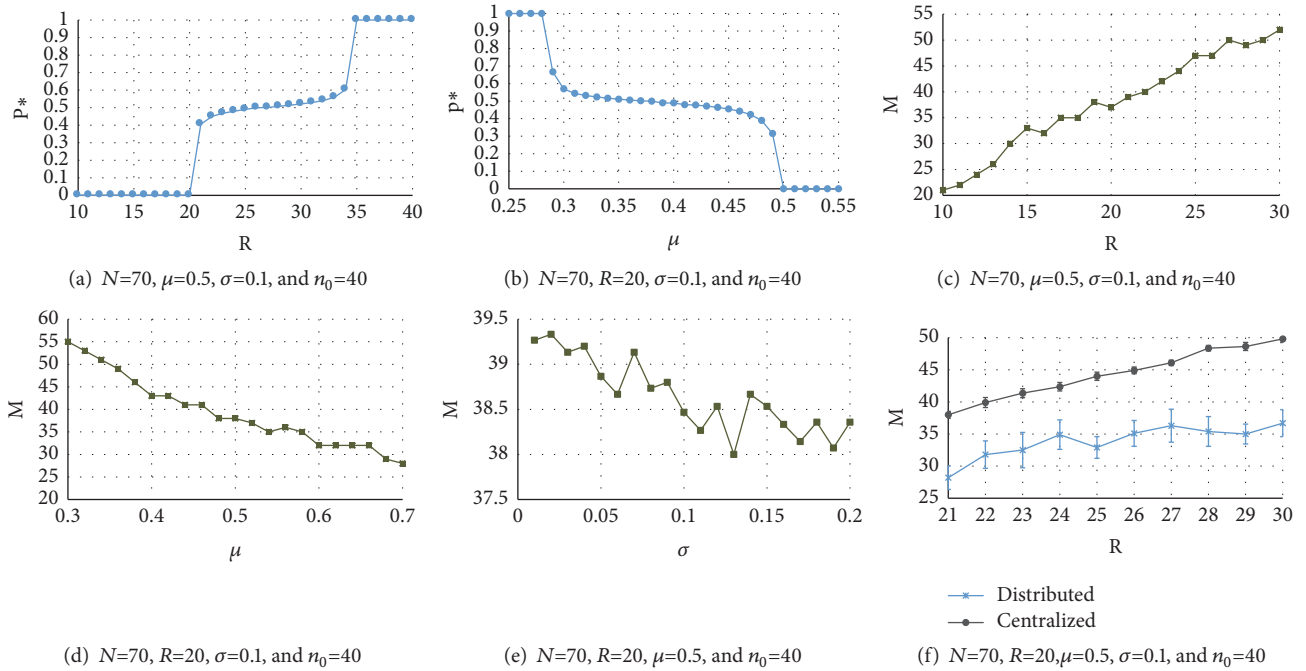


FIGURE 7: Parameters and their influences on motivating effect.

(while the planned reward is  $R = 20$ ) to successfully motivate  $n_0 = 40$  collaborators. Distributed decision-making could only motivate 3/4 participants of centralized method. To conclude, map producers should always give “extra reward” besides planned reward to reach expected  $M$ .

In this section, only two basic motivating methods are discussed. We did not discuss three participants’ selection schemes, because motivating effect is only depending on the collaboration motivating module and is irrelevant to message relaying module. However, in Section 5.2, we will find that the communication overhead is influenced by both modules; we therefore need to compare three integral schemes.

**5.2. Communication Overhead of Message Relaying.** The communication simulation scenario is 1000-meter road (in Figure 4). DSRC range is set as 150 m. Final results are averaged by 20 simulations. Performance analysis is based on three indicators. (1) First indicator is *Slots Duration*. We assume that each relay consumes one time slot. During whole process, the counted number of slots is slots duration. (2) Second indicator is *Total Transmission bytes*. The total bytes are gross of all messages bytes. (3) Third indicator is *Total Relaying Times*. It shows how many times of relay for all vehicles have taken place during the process.

**5.2.1. Communication Overhead on Straight Road.** Figure 8 shows communication performance of three schemes on straight road (in Figure 4 (left)). Figure 8(a) shows that, for three schemes, total transmission bytes go up with vehicle density. The reason is intuitive; more vehicles cause more transmitted content. Centralized selection has the largest total bytes because all vehicles send messages to seed vehicle.

In Figure 8(b), slots duration of each scheme is a constant regardless of vehicle density. Clustered selection consumes

more slots than centralized and distributed selection because it spares slots on cluster maintenance.

In Figure 8(c), centralized selection has the largest total relaying times. Because each potential participant generates a message and sends it to seed vehicle, centralized selection has larger total relaying times than the other two.

**5.2.2. Communication Overhead on Curve Road.** Robustness of scheme is reflected on communication performance in Figure 9. A robust scheme is able to work well on both straight (in Figure 4 (left)) and curve roads (in Figure 4 (right)). In all three schemes, the indicator in straight road is lower than that in curve road. The performance difference between two road conditions is at most 20%. This figure shows that, for each local participants’ selection scheme, communication overhead on curve road is a little higher than straight road, demonstrating that these schemes could still work fine on roads with small curvature.

## 6. Discussion

Given the same reward, centralized decision-making could motivate more participants than distributed decision-making. The reason is that distributed decision-making is executed by each potential participant who lacks a global view of all information. Derived from distributed decision-making, clustered selection scheme and distributed selection have the same performance on motivating. Centralized selection is better than the other two if only motivating effect is considered.

Besides motivating effect, communication overhead also deserves attention. If communication overhead is considered alone, distributed selection should be the first choice, since it performs well in three indicators. Centralized selection

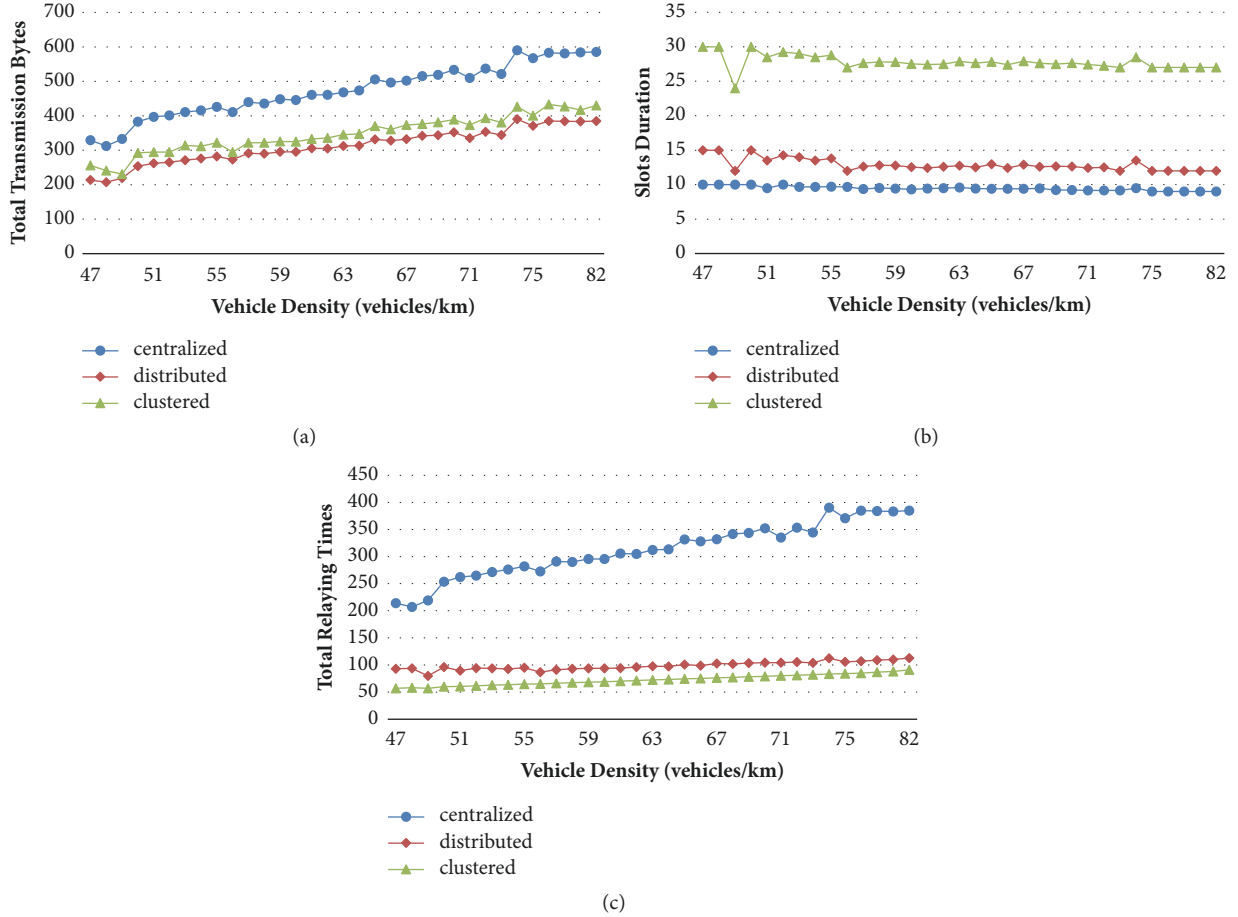


FIGURE 8: Communication overhead of three participants' selection. (a) Centralized selection has the largest bytes overhead because each vehicle sends its information to seed vehicle. (b) Clustered selection has the longest slots duration because of its stable structure. (c) Centralized selection consumes the largest total relaying times for the same reason as (a).

TABLE 3: Comprehensive comparison of three schemes.

\	Centralized selection	Distributed selection	Clustered selection
moti. perf.	Good	Normal	Normal
comm. perf.	Limited	Good	Normal

and clustered selection are second-class citizens and are only suitable for particular scenarios. For example, if a map producer hopes the participants' selection to be finished as soon as possible, then he would like to use centralized selection. Despite its shortcomings in total transmission bytes and total relaying times, centralized selection has the shortest slots duration.

In Table 3, we give a comprehensive comparison of motivating performance and communication performance. Centralized selection has a good performance in motivating but has a heavy communication overhead. Clustered selection requires more incentives and less communication overhead than centralized method. Distributed selection seems to be

the first-class choice because it works fine in both motivating and communicating. It should be noted that all three selection schemes could be realized from one single framework with few modifications of two modules. The flexibility of design framework enables map producers to select participants adaptively according to scenario requirements.

## 7. Conclusions

This paper has discussed local participants' selection of crowdsensing. We have proposed a flexible framework that could support three kinds of selection schemes with respect to different service requirements. The framework is consisted of two modules. For collaboration motivating module, we introduced two motivating methods: centralized decision-making and distributed decision-making. For message relaying module, we introduced two methods: contention-based relaying and cluster-based relaying. Three combinations of two modules are discussed and simulated. The result shows that distributed selection is usually the first-class choice because of its fine performances on communicating and motivating. Our research analyzed group sensing problem from both communication and collaboration perspectives,

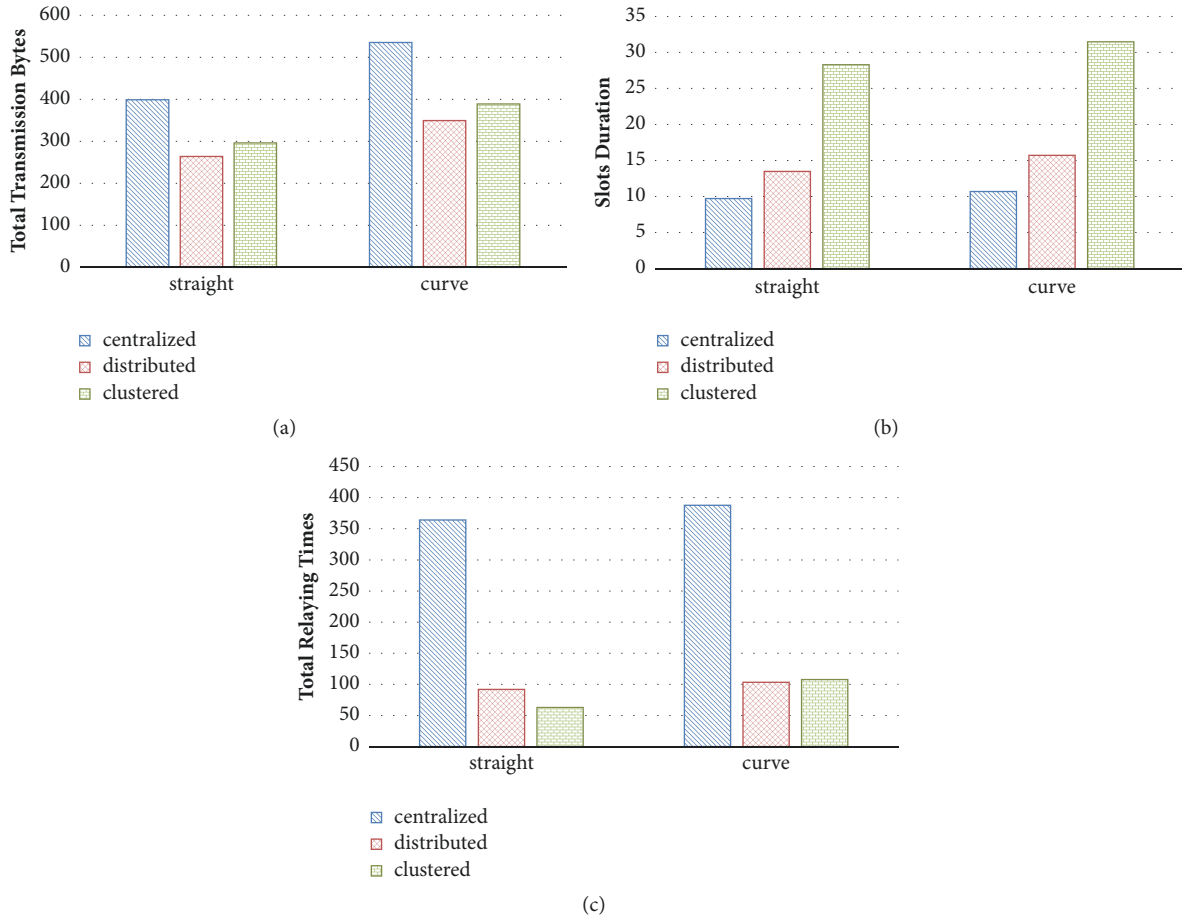


FIGURE 9: Communication overhead of straight and curve road: (a) comparison of total transmission bytes between straight and curve. (b) Comparison of slots duration between straight and curve. (c) Comparison of total relaying times between straight and curve. Three schemes still work fine in road with small curvature.

providing a primary but integrated reference for data service providers or crowdsensing operators, especially for map producers.

In the future, we will integrate ns-2 [44] into our framework to enable accurate time-simulation. We will also try to refine spread process and back process into “one-shot process.”

## Data Availability

No additional data are available.

## Disclosure

Part of this work has been published by 2017 IEEE International Conference on Services Computing [45].

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work is supported by National Science and Technology Major Project of the Ministry of Science and Technology of China (Grant no. 2016ZX03001025-003) and the Natural Science Foundation of Beijing (Grant no. 4181002).

## References

- [1] R. K. Ganti, F. Ye, and H. Lei, “Mobile crowdsensing: current state and future challenges,” *IEEE Communications Magazine*, vol. 49, no. 11, pp. 32–39, 2011.
- [2] X. Hou, Y. Li, M. Chen, D. Wu, D. Jin, and S. Chen, “Vehicular fog computing: a viewpoint of vehicles as the infrastructures,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, pp. 3860–3873, 2016.
- [3] N. Cheng, N. Lu, N. Zhang, T. Yang, X. S. Shen, and J. W. Mark, “Vehicle-assisted device-to-device data delivery for smart grid,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2325–2340, 2016.
- [4] T. Lei, S. Wang, J. Li, and F. Yang, “A cooperative route choice approach via virtual vehicle in IoV,” *Vehicular Communications*, 2017.

- [5] S. S. Kanhere, "Participatory sensing: crowdsourcing data from mobile smartphones in urban spaces," in *Proceedings of the 12th IEEE International Conference on Mobile Data Management (MDM '11)*, vol. 2, pp. 3–6, June 2011.
- [6] Q. Yuan, H. Zhou, J. Li, Z. Liu, F. Yang, and X. S. Shen, "Toward Efficient Content Delivery for Automated Driving Services: An Edge Computing Solution," *IEEE Network*, vol. 32, no. 1, pp. 80–86, 2018.
- [7] J. Levinson, J. Askeland, J. Becker et al., "Towards fully autonomous driving: systems and algorithms," in *Proceedings of the IEEE Intelligent Vehicles Symposium (IV '11)*, pp. 163–168, June 2011.
- [8] M. Campbell, M. Egerstedt, J. P. How, and R. M. Murray, "Autonomous driving in urban environments: Approaches, lessons and challenges," *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 368, no. 1928, pp. 4649–4672, 2010.
- [9] A. Zhou, S. Wang, Z. Zheng, C.-H. Hsu, M. R. Lyu, and F. Yang, "On cloud service reliability enhancement with optimal resource usage," *IEEE Transactions on Cloud Computing*, vol. 4, no. 4, pp. 452–466, 2016.
- [10] B. Zhang, C. H. Liu, J. Lu et al., "Privacy-preserving QoI-aware participant coordination for mobile crowdsourcing," *Computer Networks*, vol. 101, pp. 29–41, 2016.
- [11] R. Cavallo, D. C. Parkes, and S. Singh, *Optimal coordinated planning amongst self-interested agents with private state*, 2012.
- [12] E. Ephrati and J. Rosenschein, "Distributed consensus mechanisms for self-interested heterogeneous agents," in *Proceedings of the [1993] International Conference on Intelligent and Cooperative Information Systems*, pp. 71–79, Rotterdam, Netherlands.
- [13] Y. Zhu, X. Liu, and Y. Wang, "Pervasive urban sensing with large-scale mobile probe vehicles," *International Journal of Distributed Sensor Networks*, vol. 2013, 2013.
- [14] C. Chen, J. Yan, N. Lu, Y. Wang, X. Yang, and X. Guan, "Ubiquitous monitoring for industrial cyber-physical systems over relay—assisted wireless sensor networks," *IEEE Transactions on Emerging Topics in Computing*, vol. 3, no. 3, pp. 352–362, 2015.
- [15] Q. Yuan, Z. Liu, J. Li, S. Yang, and F. Yang, "An adaptive and compressive data gathering scheme in vehicular sensor networks," in *Proceedings of the In Parallel and Distributed Systems (ICPADS), 2015 IEEE 21st International Conference*, pp. 207–215, Melbourne, VIC, December 2015.
- [16] R. Du, C. Chen, B. Yang, N. Lu, X. Guan, and X. Shen, "Effective urban traffic monitoring by vehicular sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 1, pp. 273–286, 2015.
- [17] Z. Song, C. H. Liu, J. Wu, J. Ma, and W. Wang, "QoI-aware multitask-oriented dynamic participant selection with budget constraints," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 9, pp. 4618–4632, 2014.
- [18] B. Zhang, Z. Song, C. H. Liu, J. Ma, and W. Wang, "An Event-Driven QoI-Aware Participatory Sensing Framework with Energy and Budget Constraints," *ACM Transactions on Intelligent Systems and Technology*, vol. 6, no. 3, pp. 1–19, 2015.
- [19] C. H. Liu, B. Zhang, X. Su, J. Ma, W. Wang, and K. K. Leung, "Energy-aware participant selection for smartphone-enabled mobile crowd sensing," *IEEE Systems Journal*, vol. PP, no. 99, 2015.
- [20] S. He, D.-H. Shin, J. Zhang, and J. Chen, "Toward optimal allocation of location dependent tasks in crowdsensing," in *Proceedings of the 33rd IEEE Conference on Computer Communications (INFOCOM '14)*, pp. 745–753, IEEE, Toronto, Canada, May 2014.
- [21] M. M. I. Taha and Y. M. Y. Hasan, "VANET-DSRC protocol for reliable broadcasting of life safety messages," in *Proceedings of the ISSPIT 2007 - 2007 IEEE International Symposium on Signal Processing and Information Technology*, pp. 104–109, Egypt, December 2007.
- [22] S. Kumar, L. Shi, N. Ahmed, S. Gil, D. Katabi, and D. Rus, "Car-Speak: A content-centric network for autonomous driving," in *Proceedings of the ACM SIGCOMM 2012 Conference Applications, Technologies, Architectures, and Protocols for Computer Communication, SIGCOMM 2012*, pp. 259–270, fin, August 2012.
- [23] H. Zhou, N. Cheng, Q. Yu, X. S. Shen, D. Shan, and F. Bai, "Toward multi-radio vehicular data piping for dynamic DSRC/TVWS spectrum sharing," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 10, pp. 2575–2588, 2016.
- [24] Q. Yuan, Z. Liu, J. Li, J. Zhang, and F. Yang, "A traffic congestion detection and information dissemination scheme for urban expressways using vehicular networks," *Transportation Research Part C: Emerging Technologies*, vol. 47, no. 2, pp. 114–127, 2014.
- [25] G. Luo, J. Li, L. Zhang, Q. Yuan, Z. Liu, and F. Yang, "sdn-MAC: A Software-Defined Network Inspired MAC Protocol for Cooperative Safety in VANETs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 6, pp. 2011–2024, 2018.
- [26] G. Luo, S. Jia, Z. Liu, K. Zhu, and L. Zhang, "SdnMAC: A software defined networking based MAC protocol in VANETs," in *Proceedings of the 24th IEEE/ACM International Symposium on Quality of Service, IWQoS 2016*, chn, June 2016.
- [27] S. Wang, C. Fan, C.-H. Hsu, Q. Sun, and F. Yang, "A vertical handoff method via self-selection decision tree for internet of vehicles," *IEEE Systems Journal*, 2014.
- [28] H. B. Zhou, B. Liu, T. H. Luan et al., "ChainCluster: engineering a cooperative content distribution framework for highway vehicular communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2644–2657, 2014.
- [29] N. Cheng, N. Lu, N. Zhang, X. S. Shen, and J. W. Mark, "Opportunistic WiFi offloading in vehicular environment: A queueing analysis," in *Proceedings of the 2014 IEEE Global Communications Conference, GLOBECOM 2014*, pp. 211–216, Austin, Tex, USA, December 2014.
- [30] Q. Yuan, J. Li, Z. Liu, and F. Yang, "Space and Time Constrained Data Offloading in Vehicular Networks," in *Proceedings of the 2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pp. 398–405, Sydney, Australia, December 2016.
- [31] C. Chen, J. Zhang, R. Cohen, and P.-H. Ho, "A trust modeling framework for message propagation and evaluation in VANETs," in *Proceedings of the 2nd International Conference on Information Technology Convergence and Services (ITCS '10)*, IEEE, August 2010.
- [32] S. Chang, Y. Qi, H. Zhu, J. Zhao, and X. Shen, "Footprint: detecting Sybil attacks in urban vehicular networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 6, pp. 1103–1114, 2012.
- [33] H. Hu, R. Lu, C. Huang, and Z. Zhang, "PTRS: A privacy-preserving trust-based relay selection scheme in VANETs," *Peer-to-Peer Networking and Applications*, vol. 10, no. 5, pp. 1204–1218, 2017.



- [34] F. Yang, S. Wang, J. Li, Z. Liu, and Q. Sun, "An overview of internet of vehicles," *China Communications*, vol. 11, no. 10, pp. 1–15, 2014.
- [35] L. Duan, T. Kubo, K. Sugiyama, J. Huang, T. Hasegawa, and J. Walrand, "Motivating smartphone collaboration in data acquisition and distributed computing," *IEEE Transactions on Mobile Computing*, vol. 13, no. 10, pp. 2320–2333, 2014.
- [36] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Basar, "A Game-Based Self-Organizing Uplink Tree for VoIP Services in IEEE 802.16j Networks," in *Proceedings of the ICC 2009 - 2009 IEEE International Conference on Communications*, pp. 1–5, Dresden, Germany, June 2009.
- [37] T. Rahwan and N. R. Jennings, "Distributing Coalitional Value Calculations Among Cooperating Agents," in *Proceedings of the 25th National Conference on AI (AAAI)*, pp. 152–157, Pittsburgh, Pa, USA, 2005.
- [38] Y. Jiang and J. Jiang, "Contextual resource negotiation-based task allocation and load balancing in complex software systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 5, pp. 641–653, 2009.
- [39] V. Krishnamurthy and H. Vincent Poor, "Social learning and Bayesian games in multiagent signal processing: How do local and global decision makers interact?" *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 43–57, 2013.
- [40] M. N. Mejri, J. Ben-Othman, and M. Hamdi, "Survey on VANET security challenges and possible cryptographic solutions," *Vehicular Communications*, vol. 1, no. 2, pp. 53–66, 2014.
- [41] F. Li and Y. Wang, "Routing in vehicular ad hoc networks: a survey," *IEEE Vehicular Technology Magazine*, vol. 2, no. 2, pp. 12–22, 2007.
- [42] H. Füßler, J. Widmer, M. Käsemann, M. Mauve, and H. Hartenstein, "Contention-based forwarding for mobile ad hoc networks," *Ad Hoc Networks*, vol. 1, no. 4, pp. 351–369, 2003.
- [43] S. Vodopivec, J. Bešter, and A. Kos, "A survey on clustering algorithms for vehicular ad-hoc networks," in *Proceedings of the 2012 35th International Conference on Telecommunications and Signal Processing, TSP 2012*, pp. 52–56, cze, July 2012.
- [44] Q. Chen, F. Schmidt-Eisenlohr, D. Jiang, M. Torrent-Moreno, L. Delgrossi, and H. Hartenstein, "Overhaul of IEEE 802.11 modeling and simulation in NS-2," in *Proceedings of the 10th ACM Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems (MSWiM '07)*, pp. 159–168, October 2007.
- [45] S. Yang, J. Li, Z. Liu, and Q. Yuan, "Relaying Message and Motivating Collaboration for VANET Data Service," in *Proceedings of the 14th IEEE International Conference on Services Computing, SCC 2017*, pp. 52–59, usa, June 2017.

