Charoenchai and Siripongwutikorn
J Wireless Com Network (2022) 2022:67
https://doi.org/10.1186/s13638-022-02149-9

EURASIP Journal on Wireless Communications and Networking

RESEARCH

Open Access

Coalitional graph game for area maximization of multi-hop clustering in vehicular ad hoc networks

Siwapon Charoenchai* and Peerapon Siripongwutikorn*

*Correspondence:
siwapon.c@mail.kmutt.ac.th;
peerapon.sir@mail.kmutt.ac.th

Department of Computer Engineering, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Pracha-Uthit Rd., Bangkok 10140, Thailand

Abstract

Road traffic information can be utilized in many applications of intelligent transport systems. It can be collected from vehicles and sent over a vehicular ad hoc network (VANET) to roadside units (RSUs) acting as sink nodes. Due to rapid mobility and limited channel capacity in a VANET where vehicles must compete to access the RSUs to report their data, clustering is used to create a group of vehicles to facilitate data transfer to the RSUs. Unlike previous works that focus on cluster lifetime or throughput, we formulate a coalitional graph game for multi-hop clustering (CGG-MC) model to create a multi-hop cluster with the largest possible coverage area for a given transmission delay time constraint to economize on the number of RSUs installed. Vehicles cooperatively form a proper coalition with relation directed graphs among vehicles in a multi-hop cluster to collect, aggregate, and forward data to RSUs instead of individually competing to connect directly to RSUs. Vehicles decide to join or leave the coalition based on their individual utility, which is a weighted function of the coverage area, number of members in the cluster, relative velocities, distance to sink nodes, and transmission delay toward the sink nodes. The distributed-solution approach based on probabilistic greedy merging of coalitions is used to derive the grand coalition, and the probability of grand coalition formation is analyzed by using a discrete-time Markov chain. Our results show that the proposed solution approach yields a 95% confidence interval of the average utility between 61 and 68% relative to the maximum utility in the centralized-solutions. Additionally, our CGG-MC model outperforms the non-cooperation model by approximately 166% in terms of enlarging the coverage network area under a transmission delay time constraint.

Keywords: Clustering, Coalitional graph game, Game theory, Vehicular ad hoc network (VANET)

1 Introduction

A vehicular ad hoc network (VANET) is a group of vehicles with wireless transceivers that form a wireless ad hoc network, allowing for the exchange of data among vehicles not within single-hop communication ranges. Exchanging data through VANETs is utilized for implementing a wide variety of intelligent transportation systems (ITS) applications such as traffic monitoring [1], safety warnings, driving decisions, autonomous driving [2], transportation payments, secure database recordings on the blockchain [3], popular content advertisement [4], and vehicular cloud computing [5]. Vehicles in VANETs are usually equipped with IEEE 802.11p radio transceivers to provide wireless access in vehicular environments (WAVE), which allows data exchanges between vehicle-to-vehicle (V2V) and vehicle-to-roadside infrastructure (V2I). The radio transceiver installed in a vehicle is called an on-board unit (OBU), and a transceiver on a roadside infrastructure is called a roadside unit (RSU). In a typical ITS, vehicles must compete for limited wireless channel capacity to exchange data with the RSUs for reporting to the traffic control center, so the wireless channel can be overloaded when the number of competing vehicles is large. That load can be shared by installing more RSUs along the roadside, but RSU installation and maintenance can be expensive.

One of the challenging problems in VANETs is to achieve efficient routing in the presence of channel contention collisions and dynamic topological change due to the mobility of vehicles. To improve the routing scalability and reliability of VANETs, clustering techniques have been used to construct the tree-like hierarchical network structure, or a multi-hop cluster, where vehicles are grouped into overlapping clusters as shown in Fig. 1. In the cluster formation, vehicles are assigned one of the three roles—cluster head (CH), cluster member head (CMH), and cluster member (CM), and the hierarchical neighboring link relationship is established based on the assigned role. With multi-hop clustering, only CMHs act as intermediate nodes responsible for packet routing and can also aggregate data, which reduces the network traffic load and improves the overall network performance.

Several algorithms have been proposed to form a multi-hop cluster that is optimized for different performance objectives, such as cluster stability and delay. For instance, clustering for high cluster stability will attempt to form clusters whose neighboring vehicles have a low relative velocity to obtain stable links. If a low-delay cluster is desired, the

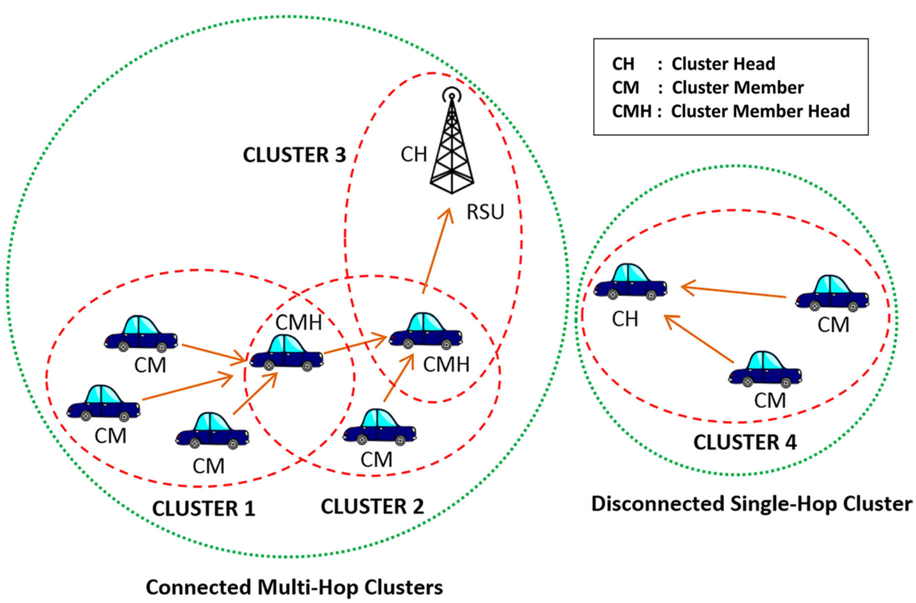


Fig. 1 Multi-hop clustering in a VANET with an RSU

network diameter in hops must be restricted. Therefore, a wide variety of metrics, such as relative velocity, relative position, signal strength, and signal-to-noise ratio, have been used for cluster formation depending on the performance objectives [6, 7]. The relative velocity of vehicles is commonly adopted to maximize link stability and hence the cluster lifetime. While most prior works focus on clustering with high link stability or with high throughput, we focus on expanding the coverage area of the connected multi-hop clusters to avoid the need for additional RSU installations, subject to the packet transmission delay time constraint.

The contributions of this paper are as follows:

- We propose a coalition graph game for multi-hop clustering (CGG-MC) model in which vehicles can cooperatively gather, aggregate, and forward data to a sink node or an RSU.
- The model attempts to maximize the cluster area of a network subject to the packet transmission delay constraint.
- The distributed-solution approach based on probabilistic greedy merging of coalitions is applied to derive the clustering results, and the probability of getting each clustering is analyzed by using a discrete-time Markov chain (DTMC).

Our CGG-MC model is a cooperative game clustering scheme in which all players in the game can cooperatively help each other in forming the probable coalitional structure and expanding the coverage area. Unlike the traditional and non-cooperative game clustering schemes where each vehicle or player selfishly competes to obtain the highest benefit, CGG-MC can fairly share the utility as a win-win. The higher-level nodes such as the RSU can gather more information from a larger coverage area, and the lower-level nodes can obtain the channels to access the RSU for reporting their data.

The remainder of this paper is organized as follows. In Sect. 2, we review related works on clustering techniques in VANETs and discuss their limitations and issues. In Sect. 3, we present the CGG-MC model for forming multi-hop clusters of vehicles. The cluster merging process and its complexity in terms of the number of exchanged messages are discussed. The probabilistic greedy merging of coalitions is then developed and analyzed by using DTMC. In Sect. 4, the performance evaluation results are presented to demonstrate the effectiveness of the proposed model under both static and dynamic scenarios. Finally, the conclusion is offered in Sect. 5.

2 Related work

2.1 Clustering schemes

Clustering for various types of wireless networks such as wireless sensor networks (WSN) and wireless mesh networks (WMN) has been intensively researched. However, these environments and clustering purposes are different from VANETs. In a WSN [8, 9], the sensor nodes are statically embedded for sensing and sending the data with a limited power supply. Energy management is the main requirement of WSN clustering for prolonging the network lifetime. The metrics commonly used in WSN clustering involve energy consumption, transmission power, and sleep and wake-up time schedules, whereas these energy metrics are not required for VANETs because typical vehicles can generate their own energy supply. On the other hand, concerns in a VANET include limited channel capacity and network stability because VANETs are larger than WSNs and the nodes have rapid mobility. In a WMN [10], nodes are connected in a mesh topology with robust connectivity and high reliability due to the natural self-healing of mesh connections. Typically, a WMN has two types of nodes [i.e., mesh router (MR) and mesh client (MC)], and two hierarchical levels (upper and lower). The upper level consists of static MR nodes, and the lower level consists of dynamic MC nodes that are movable and connected to their closest MR. The main purpose of a WMN is to continuously provide the same network access to all MC nodes even though they are moving. The signal strength is a very important metric for WMN clustering. WMNs may seem usable in the vehicular environment where an MR acts as an RSU and MCs act as vehicles. However, a WMN is most suitable for mobile devices that are used in an indoor environment, or within the building area, but not suitable for devices with rapid mobility.

VANET clustering can be classified into many categories depending on clustering metrics and grouping mechanisms, such as flat and hierarchical clustering, centralized and decentralized computation clustering, single-hop and multi-hop clustering, single-metric and multi-metric clustering, and uplink (vehicles-to-RSU) and downlink (RSU-to-vehicles) clustering. In flat clustering, vehicles independently join a cluster and connect as peer-to-peer, and data can be disseminated by broadcasting or flooding [11, 12]. Flat clustering minimizes the overhead of data exchanged, but it suffers a broadcast storm problem [13] in a dense flat network. In hierarchical clustering, a rank is assigned to every vehicle in the cluster to enable inherent tree-based routing. It requires one vehicle to be a cluster head (CH), who is responsible for relaying data packets between members of the intracluster and another vehicle in the intercluster. Hierarchy leads to successful dissemination by relaying data without the broadcast storm problem, but it requires more overhead and delay. The earliest work in hierarchical clustering is the lowest ID (LID) protocol [14, 15]. It uses the node identification number that is assigned in ascending order for the CH selection. This scheme is simple because the node ID is only a single-metric for clustering, but it is not robust or reliable for mobility. Mobility-based metric clustering (MOBIC) in [16] uses the received signal strength (RSS) power level as a single-metric to measure the relative mobility of nodes within a single-hop distance. Aggregate local mobility (ALM) in [17, 18] is similar to MOBIC but uses the location information from the global positioning system (GPS) instead of the RSS to derive the relative mobility of nodes. The node which has the smallest variance of relative mobility from its single-hop neighbors is elected as CH. These schemes improve cluster stability and prolong the cluster lifetime under mobility. Affinity propagation for vehicular network (APROVE) in [19] forms clusters by considering multi-metrics consisting of minimum distance and minimum relative velocity between each node and CH. Each node periodically finds and updates its new preferred CH. This process is similar to the traditional K-means algorithm for cluster stability convergence. Vehicular weighted clustering algorithm (VWCA) in [20] uses multi-metrics combined and weighted with vehicle direction, number of neighbors in the dynamic transmission range, entropy value calculated from relative mobility, and distrust value calculated from abnormal packet relaying rate. This weighted multi-metric scheme can balance the clustering factors well in order to obtain a longer duration of cluster lifetime. Hierarchical clustering techniques, as mentioned above, yield a clustering diameter that is limited to at most two hops where the relay node is the CH. However, it can be enlarged by spreading the multi-hop links.

A recent work in [21] applies clustering in the context of data dissemination. A single-hop cluster is formed by using arbitration based on pre-assigned cluster head indices and the probabilistic forwarding mechanism is used to broadcast packets, where the probability of broadcast attempts is a decreasing function of the number of transmission rounds. The maximum allowed cluster size and the maximum number of transmission rounds in the clustering and forwarding processes are determined by multi-objective optimization using the packet delivery ratio, end-to-end delay, and the number of dropped packets as the metrics. Multi-hop clustering using relative mobility to select CHs appears in the vehicular multi-hop algorithm for stable clustering (VMaSC) [22]. Each node attempts to connect with its neighbor, who is a CH or is already joined with a CH. The nodes affiliate in a multi-hop structure with a minimum number of hops to decrease the delay of packet transmission. This scheme combines the IEEE 802.11p VANET with the long-term evolution (LTE) cellular network. However, LTE is not mainly used for V2V communication. It just helps CHs to reduce clustering overhead and disseminate data to remote regions with high cost cellular communication between CHs and LTE base stations. The distributed multi-hop clustering algorithm for VANETs based on neighborhood follow (DMCNF) in [23] assumes that a vehicle can easily identify its most appropriate neighbor for following within a single-hop distance. To follow that neighbor node, it selects the same CH and joins that multi-hop cluster based on relative mobility, the current number of followers, and the historical cluster information metrics. Advanced multi-hop clustering (AMC) [24] is another technique that also uses neighborhood following to form a multi-hop cluster. With a pre-specified CH node, the nearby nodes can directly connect to the CH and get the CM status. The farther nodes follow only one of their single-hop CM neighbors who is the most suitable node to share the CH. The suitability is considered from the weighted summation of four important parameters, which are the packet transmission delay, the link lifetime, the distance to CH, and the neighborhood degree. These parameters are iteratively checked all the way to the CH node to maintain cluster stability and performance. The robust mobility-aware clustering (RMAC) in [25] selects CHs based on relative node speeds, locations, and directions. Two or more single-hop clusters, which have an overlapping transmission range, merge to form a larger multi-hop cluster. The CH of a smaller cluster simultaneously becomes a member of a larger cluster. This scheme constrains the number of connection links per node to avoid cluster instability, and produces a performant cluster lifetime. However, the degradation of throughput and packet delay as the cluster enlarges has not been investigated. Dynamic backbone-assisted medium access control (DBA-MAC) [26] and cluster-based location routing (CBLR) [27] are the routing approaches specifically applicable to multi-hop clustering for end-to-end communication. DBA-MAC forms a linear backbone topology among vehicles that resembles a snaking chain while CBLR merges clusters into a multi-hop cluster and assigns a node in an overlapping area of original clusters as the gateway to relay the information. Although both schemes route data packets from a source node to a destination node efficiently, they do not consider the delay time constraints involved in collecting data from all vehicles and routing them to a central CH node.

The computational approaches, such as centralized and decentralized computation, should be considered for clustering. Schemes that use centralized clustering commonly assign an RSU or a top-level CH as a central node to compute the clustering algorithms. Centralized clustering is simple, gives an exact solution, optimizes the clustering purpose, and seems appropriate for hierarchical topology. However, its computational time may increase quickly as the number of vehicles increases based on the complexity of the clustering algorithm. Conversely, decentralized clustering schemes distribute the computation to all cluster members. Each node discovers neighbors by exchanging hello messages which contain necessary information metrics. The node computes the best interaction with its neighbors by itself based on available information at that moment. Decentralized clustering is fast but does not guarantee the optimal solution. Distributed clustering algorithm (DCA) and distributive mobility adaptive clustering (DMAC) in [28] are examples of weighted multi-metric decentralized clustering schemes. The CH election power is distributed to all members. The node having the greatest weight-sum based on link quality and mobility metrics is elected to be a CH.

2.2 Game theoretical clustering

Another line of related work applies game theory to VANET clustering. Game theory not only considers the metrics as the previous explained works do, but can also make a decision based on the other nodes' actions. Vehicles are the players or competitors in the game. Various types of metrics are combined to form a payoff or utility function. Each player chooses the best action strategy, such as joining or ignoring the cluster, to obtain the best utility. Game patterns can be both non-cooperative and cooperative, but every game has the same purpose of finding the equilibrium point or the best strategy for all players. The multi-player game theoretic algorithm for intra-cluster data aggregation (MGADA) in [29] is a non-cooperative game for clustering. Players in this game selfishly compete to transmit the data to the CH. The only action strategies are transmitting or silencing. To obtain their highest utility, players solve for their optimal solution using the Nash equilibrium. In [30], an evolutionary game theoretic approach for stable and optimized clustering in VANETs (EGT) balances the population size of multiclusters by applying an iterative evolutionary process. The structure of this scheme has connection links from RSU to CHs, and each CH forms a single-hop cluster with its members. The RSU acts as a central node to compute the total throughput of the network. The population size of each cluster is raised or reduced depending on the available channel capacity resources of each cluster. In each iteration, each vehicle has a non-cooperative selection strategy to be a member of a cluster. It will move a membership from the current cluster to another cluster if it obtains higher utility until the utility converges. The metrics used in the utility calculation function are channel capacity and throughput. The evolutional dynamic replicator adjusts the population proportion between clusters to find the equilibrium point, and the cluster stability is proved by using the Lyapunov function.

In coalitional games [31, 32], the authors classify games into three types, such as canonical game, coalition formation game, and coalitional graph game. The canonical game denotes a game in which no group of players obtains worse utility than by acting alone. The coalition formation game denotes a group of players joining together to obtain the highest utility, but with a cost that each player must consider whether to leave or join the group. The coalitional graph game denotes a group of players forming with a relationship graph between players.

In the coalition formation game approach [33], vehicles are formed into coalitions. The vehicles in the same coalition cooperatively relay the downlink data packets from an RSU to a vehicle node. Instead of direct delivery, relaying data packets by vehicles in the same coalition yields higher efficiency than by the RSU acting alone. This scheme uses social network analysis (SNA) to filter out some less-utilized vehicles before forming a coalition. Nash bargaining is applied to find the utility of coalitions and the probability that vehicles are connectable to others in the same coalition. Vehicles can merge or split coalitions if they and the coalitions obtain higher utility. A coalitional structure is a combination of coalitions. There are many possible coalitional structures, but the most stable coalitional structure can be proved by using a discrete-time Markov chain (DTMC). This solution is exact. However, the possible solution space of this coalitional game grows as a combinatorial function. The computational time grows as the combinatorial complexity increases along with the number of vehicles. The approximate solution can be solved by another algorithm called the distributed iterative merge-and-split algorithm. This is another way to quickly find a fairly stable coalitional structure. In [34], the coalition formation game forms a connected automated vehicles (CAV) coalition at the multi-lane merging zone of a road. The coalition performs cooperative decision making by motion prediction. This scheme is interesting in that it can form a sub-coalition within a coalition, which may be applied to model a hierarchical multi-hop clustering.

In the coalitional graph game, the membership of the coalition affects not only the utility, but also the relationship between members. The first work introduced in this field is in [35]. The interconnection between players in the same group can be modeled by graphs and vertices. Relationship graphs strongly affect the utility of each player. Altering the link structure of a coalition will change the utility of each member. The non-transferable utility coalition graph game algorithm (NTU-CGGA) in [36] forms the interaction graph between players for improvement of the channel capacity resource allocation. The player with more available channels can help the player with less available channels to relay information. This scheme uses the throughput and the fairness index as the utilized metrics to form the coalition. The popular content distribution problem (PCD) can be solved by the coalitional graph game in [37]. It is the V2I communication that the RSU uses to broadcast the advertisement to vehicles. However, the data packets can be lost. The coalitional graph game forms a coalition among vehicles to exchange and complement the missing data packets. The missing packets and available channel slots are parameters to form the utility and graph. This scheme models the whole network as a grand graph and may take a long time to form a coalition when the vehicle density is high due to the algorithm complexity. The cooperative V2V-aided transmission coalitional game (CVCG) in [4] improves the issue of [37]. It uses the location-based blocking scheme to split a huge coalition into several small coalitions.

In summary, although there are already some works studying the coalitional graph game to form a stable coalition, forming a coalitional directed graph for a hierarchical multi-hop clustering to maximize the cluster coverage area still needs to be studied.

3 Method

3.1 CGG-MC model and coalition formation

We design a hierarchical multi-hop clustering model for a network of vehicles on a suburban or urban road with installed RSUs. Each vehicle is assumed to be equipped with GPS, sensors or cameras, and an ECU processor, so it can detect local information, such as position, velocity, movement direction, and surrounding events of interest. Its OBU is assumed to be an IEEE 802.11p standard radio transceiver that has an adjustable transmission range function. The communication bandwidth is separated into a cluster control channel and a data channel. Vehicles use the cluster control channel to organize the clustering process, and they transfer their collected data through the data channel. These two channels do not interfere with each other. To achieve our goal, the vehicles form a multi-hop cluster covering an area as large as possible, and cooperatively collect and transmit data via this cluster structure to a nearby RSU within a given transmission delay time constraint.

The vehicles are grouped into clusters to form a coalition. Each cluster has one vehicle as a cluster head (CH) and at least one child node as a cluster member (CM). Data flows from CM to CH in a single-hop. The clusters merge into the same coalition to expand the network area as a connected multi-hop cluster as shown in Fig. 1. The CH of a cluster becomes a member of another cluster that it merged into. Its status changes to a dual status of CH and CM simultaneously, or it is called CMH. The CMH performs routing between its cluster members and its CH. Routing is automatically formed like a tree structure, so data can flow by multi-hop steps from leaf CM nodes and be aggregated while passing through each CMH node until reaching the root sink CH node. The root sink CH node is almost an RSU, but the vehicle that is the current top-level node will be the CH instead if no RSU exists. Thus, the coalition of any node is defined as the cooperation in which that node receives data from all its offspring to aggregate and send to its parent, as shown in Fig. 2.

We use a coalitional graph game with non-transferable utility (NTU) to maximize the coalition area A_i while satisfying the transmission delay time T_i from node i to the sink node. The number of clusters joining the coalition is limited by the transmission delay time constraint specified prior to the clustering. The coalitional graph game is a type of cooperation game theory, where the utility is affected by both a membership and a relationship structure between members of the coalition, as explained in the related work section, and can be analyzed by the relational link graph among players in the game. The NTU is a sub-type of the coalitional graph game, where a player cannot directly transfer its utility to another player in the same coalition, but they can take some actions to support the coalition to obtain higher utility.

We model the coalition formation of connected multi-hop clusters as shown in Fig. 2. The set of all vehicles and the RSU, denoted by $N = \{1, 2, 3, \ldots, |N|\}$, are the players in the game. Any vehicle node i has a 2D position (x_i, y_i). Its coalition is denoted by *family*(i) which is a set of its parent node, itself, and all its offspring nodes. Node i directly receives data from its child nodes and transmits data to its parent node. The Euclidean distance between the transmitter node t_x and the receiver node r_x is denoted by $d_{t_x \rightarrow r_x}$. The coalition area of node i covering the whole family of node i is calculated as

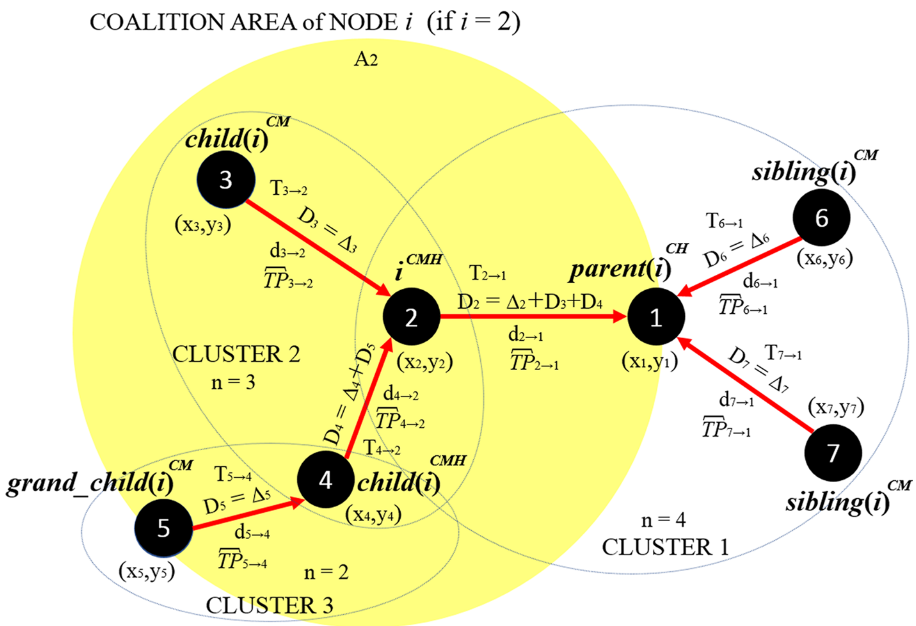


Fig. 2 Parameters of our proposed model

$$A_i = \begin{cases} 0 & ; \text{ node } i \text{ is alone (singleton-coalition)} \\ \pi R^2_{\text{cen}\rightarrow\text{bor}} & ; \text{ otherwise,} \end{cases} \tag{1}$$

where A_i is the coalition area of node i in the circular shape. This area shows the knowledge of node i because it can get information from all of its covered offspring to serve its parent. The radius $R_{\text{cen}\rightarrow\text{bor}}$ is the Euclidean distance from the centroid to the border of the coalition. The centroid is the average position of every node in the *family*(i) that is located at the position $(x_{\text{cen}}, y_{\text{cen}})$. The border can be referred to the position of the farthest node in the *family*(i). The maximum distance given from the centroid to any node in the *family*(i) is the coalition radius.

The *parent*(i) denotes the parent of node i, and *offspring*(i) denotes a set of offspring of node i. The T_i is the transmission delay time of node i that node i consumes while receiving data from all of *offspring*(i) and transmitting aggregated data to *parent*(i). However, the transmission delay time is zero if the node i is only one member in the coalition because it does not yet receive or transmit any data. Therefore, the transmission delay time of node i is calculated as

$$T_i = \begin{cases} 0 & ; \text{ node } i \text{ is alone (singleton-coalition)} \\ \left[\max\limits_{\forall j \in offspring(i)} (T_{j\rightarrow i})\right] + T_{i\rightarrow parent(i)} & ; \text{ otherwise,} \end{cases} \tag{2}$$

where $T_{i\rightarrow parent(i)}$ is the transmission time that node i successfully delivers the data packets to its parent node. Node j denotes any offspring of node i. The $T_{j\rightarrow i}$ is the transmission time that node j successfully delivers the data packets to node i. The data channel is fairly shared, so all of the offspring nodes can deliver data simultaneously. The offspring node that has a large data size or a bottleneck path will consume more time and be slower than others. We use the maximum function to pick up the time that offspring node consumes for successful delivery.

The transmission time $T_{i \to parent(i)}$ is calculated by the amount of aggregated data from node i that is successfully delivered to $parent(i)$ with the maximum throughput rate as

$$T_{i \to parent(i)} = \frac{D_i}{\widehat{TP}_{i \to parent(i)}}, \tag{3}$$

where D_i denotes the amount of aggregated data from node i. It consists of node i's self local data Δ_i and all data from all offspring nodes as depicted in Fig. 2.

The maximum throughput of the path between node i and its parent node, denoted by $\widehat{TP}_{i \to parent(i)}$, can be scaled from the data channel capacity and approximated based on Gupta–Kumar's model [38, 39] as

$$\widehat{TP}_{i \to parent(i)} \propto \frac{C_{i \to parent(i)}}{\sqrt{n \log n}}, \tag{4}$$

where n represents the number of nodes in the same cluster that consists of node i, $parent(i)$, and $siblings(i)$.

The data channel capacity, denoted by $C_{i \to parent(i)}$, can be approximated by the Shannon capacity formula [36] as

$$C_{i \to parent(i)} = B \log_2 \left(1 + \frac{P_{r_x}}{N_0}\right), \tag{5}$$

where B is the channel bandwidth, P_{r_x} is the received power at $parent(i)$ that has decayed along with the distance from the transmitter node, and N_0 is the noise power. From the equations of Gupta–Kumar and Shannon, the maximum throughput fundamentally depends on the distance between nodes and the number of nodes in the cluster.

Any coalition, denoted by S, can be regarded as the subset of all players in the game ($S \subseteq N$). Each player attempts to increase its coalition area by joining another coalition and aims to increase its utility as much as possible. However, a larger coalition area leads to throughput reduction, and thus more transmission delay time to the sink node due to the longer transmission distance and more interfering nodes. Thus, the utility of each player can be represented in terms of gain and cost, where the gain is the coalition coverage area, and the cost is the transmission delay time from its most distant offspring node to its parent node. Each player is unable to transfer its utility to another, but it can join or merge its coalition with another to achieve higher utility altogether.

The utility of node i, denoted by u_i, can be defined as a function of coalition area A_i and transmission delay time T_i as

$$u_i = (\alpha A_i - \tau T_i) + (\rho_i - \psi_i), \tag{6}$$

where the α and τ are the positive coefficients to weigh the importance of the coalition area and the transmission delay time. The parameters ρ_i and ψ_i are the reward and penalty, respectively. Node i will get a reward ρ_i if it connects to a coalition that has an RSU as a member. We assign a constant value for a reward, such as $\rho_i = 100$ for a coalition having an RSU, but $\rho_i = 0$ otherwise. Node i will get a penalty ψ_i if its velocity differs greatly from that of its parent node. In particular, the penalty ψ_i is defined as the magnitude of relative velocity between node i and $parent(i)$ as

$$\psi_i = \begin{cases} 0 & ;\ \text{node } i \text{ has no parent} \\ \left|\vec{v}_i - \vec{v}_{parent(i)}\right| & ;\ \text{otherwise.} \end{cases} \tag{7}$$

We define the coalitional structure, denoted by $\Gamma = \{S_1, S_2, \ldots, S_m\}$, as a set of m coalitions that occur at the same time. The coalition S that contains all nodes in N is called the grand coalition. However, the grand coalition may not be guaranteed to occur, and there may be many coalitions at the same time. If a coalition S_k and another coalition $S_{k'}$ occur at the same time, each coalition will not contain the same node ($S_k \cap S_{k'} = \emptyset$ for $k \neq k'$).

The rank precedence and relationship graph of the nodes in the coalition S are represented by the arrow sign ($\rightarrow$). The head of the arrow points at the higher-rank node, and the tail of the arrow is at the lower-rank node. The rank level status of each node can be any one in the set $L = \{\mathrm{UN}, \mathrm{CM}, \mathrm{CH}, \mathrm{CMH}\}$ which denotes un-cluster status, cluster member status, cluster head status, and dual status (both CM and CH), respectively. Therefore, the coalition S can be written in the coalitional graph form, such as (1^{UN}), $(1^{\mathrm{CM}} \rightarrow 2^{\mathrm{CH}})$, $(1^{\mathrm{CM}} \rightarrow 2^{\mathrm{CMH}} \rightarrow 3^{\mathrm{CH}})$, $(1^{\mathrm{CM}} \rightarrow 2^{\mathrm{CH}} \leftarrow 3^{\mathrm{CM}})$, ..., where $1, 2, 3, \ldots$ are the ID numbers of nodes in the game. The coalitional structure Γ can be written as follows:

$$\Gamma = \begin{cases} \left\{\left(1^{\mathrm{UN}}\right),\ \left(2^{\mathrm{UN}}\right),\ \left(3^{\mathrm{UN}}\right)\right\} \\ \left\{\left(1^{\mathrm{CH}} \leftarrow 2^{\mathrm{CM}}\right),\ \left(3^{\mathrm{UN}}\right)\right\} \\ \left\{\left(1^{\mathrm{CH}} \leftarrow 2^{\mathrm{CMH}} \leftarrow 3^{\mathrm{CM}}\right)\right\} \\ \left\{\left(2^{\mathrm{CM}} \rightarrow 1^{\mathrm{CH}} \leftarrow 3^{\mathrm{CM}}\right)\right\} \\ \qquad\qquad \vdots \end{cases}.$$

The merge and split process is a critical part of our game model. The rank level status can be changed during the clustering process due to merging and splitting of nodes. For a given node, its status can change with the following merging and splitting:

- The UN changes to CM, or CH changes to CMH, when the node joins another coalition as a child node.
- The UN changes to CH, or CM changes to CMH, when another coalition joins the node as its offspring.
- The CM changes to UN, or CH changes to UN, when the node leaves its coalition to be a stand-alone node.
- The CMH changes to CM when all its children leave the node.
- The CMH changes to CH when its parent leaves the node.

For example, suppose at time step t, the coalitional structure is

$$\Gamma_t = \{S_1, S_2, S_3\} = \{(1^{\mathrm{UN}}), (2^{\mathrm{UN}}), (3^{\mathrm{UN}})\}.$$

At time step $t+1$, if nodes 1 and 2 merge such that $S_1^{\mathrm{new}} = (1^{\mathrm{CH}} \leftarrow 2^{\mathrm{CM}})$, the new coalitional structure will be

$$\Gamma_{t+1}^{\text{new}} = \{S_1^{\text{new}}, S_3\} = \left\{ \left(1^{\text{CH}} \leftarrow 2^{\text{CM}}\right), \left(3^{\text{UN}}\right) \right\}.$$

3.2 Merging process

3.2.1 Messages exchanged in coalition forming

The UN or CH node who is the highest ranked node in the coalition acts as the requester to join another coalition. It broadcasts the join-request message over the cluster control channel. The join-request message contains the current node's position, velocity, utility, and coalitional graph. Any node of another coalition acts as the acceptor who can accept the request if both coalitions can improve their utility. The merging of these coalitions successfully forms a new coalition when the acceptor returns the join-accept message to the requester, and then the requester returns the join-confirm message to the acceptor. Then, the requester and acceptor change their rank status as explained above. Moreover, when the requester changes its status to CM or CMH, it cannot be the requester until it leaves the coalition.

3.2.2 Complexity of messages exchanged in CGG-MC

We consider the worst-case scenario where all N vehicles are in transmission range of each other (all nodes can connect together with a single-hop range), and only a pair of nodes successfully merges in each merging step. The complexity to form and expand the coalition depends on the number of messages exchanged between nodes. To merge two coalitions, three messages are transmitted: join-request, join-accept, and join-confirm. If all N nodes broadcast the join-request message, the number of messages exchanged is no more than $3N$. Initially, there are N singleton-coalitions. The worst-case coalition merging occurs when a singleton-coalition merges with another coalition so that the number of singleton-coalitions is reduced by one after each merging. Therefore, there are $N-1$ merging steps at most to finally obtain the grand coalition of size N, and the total number of messages exchanged is upper-bounded by $3N(N-1)$ or $O(N^2)$.

3.3 Probabilistic Greedy merging

The utility of a grand coalition depends on the coalition merging sequence. The sequence of coalitional structures obtained from the merging in a three-node case is shown in Fig. 3. To obtain the grand coalition with the highest utility, all possible coalitional structures must be elaborated to identify the one with the highest utility. However, this approach is not feasible in practice due to the explosive growth in the number of possible coalitional structures as the number of vehicles gets large. Under the sequential merging process discussed in Sect. 3.2, coalitions in Γ_1 can be changed to any coalition in Γ_2 to Γ_7 depending on which node initiates the join-request message first, and whether the new coalition improves the utility. The grand coalition can be any one in Γ_8–Γ_{16}, one of which has the maximum utility.

If we let the merging process occur without any intervention, the grand coalition may end up with a poor utility. In particular, any node can broadcast the join-request message, and the next coalitional structure is formed only if the utility is improved. As such, the next coalitional structure in the sequence can be any structure having a higher utility. For instance, from Fig. 3, if Node 2 is the first node that broadcasts the join-request,

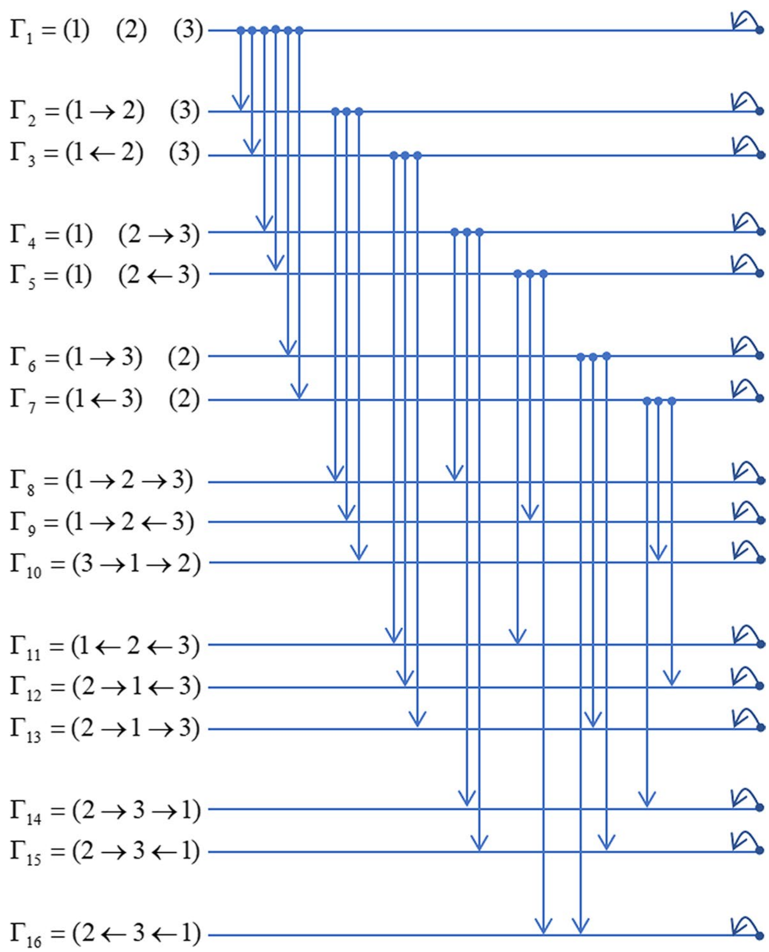


Fig. 3 Transition of possible coalitional structures for a case of 3 nodes

and Node 3 is the first node that accepts and returns the join-accept message, the next coalitional structure to form will be Γ_4.

A better approach, that improves the chance of obtaining the grand coalition with high utility, is to give higher preferences to coalitions with higher utilities in each merging stage, which we refer to as probabilistic greedy merging. Each requester node broadcasts the join-request message and waits for join accept messages from all acceptor nodes in its transmission range to determine possible coalitional structures in the next stage and their utility. All requester nodes exchange their potential coalitions in the next stage, and each candidate coalitional structure is assigned the probability of being selected that depends on its utility relative to the others. Therefore, the candidate coalitional structure with the highest utility has the largest probability of being selected.

3.4 Stability analysis of coalition formation

In this section, we show that the grand coalition obtained from probabilistic greedy merging (described in Sect. 3.3) is stable, and we derive the probability of reaching each of the possible grand coalitions. A coalition S_k is called internally stable if no member can improve its utility by leaving or splitting out of the coalition, and it is called externally stable if it cannot improve its utility by joining or merging with another coalition $S_{k'}$. So, the coalition is stable when its structure no longer changes.

Let $\Omega = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_\gamma\}$ denotes a set of all possible γ coalitional structures for a given number of nodes with fixed topology. An example of Ω for a case of three vehicles was shown earlier in Fig. 3. The probability that the merging of coalitions described in Sect. 3.2 will result in a given coalition structure Γ_j can be analyzed by using a discrete-time Markov chain (DTMC) where Γ_j represents any jth state in the DTMC state space Ω.

Let $\Gamma_j = \{S_1, S_2, \ldots, S_m\}$ be an arbitrary coalitional structure with m coalitions, and let the coalitional structure $\Gamma_1 = \{\{1\}, \{2\}, \ldots, \{|N|\}\}$ contain all singleton-coalitions. Let $v(S_k)$ be the sum of utilities from all the nodes in the coalition S_k of coalitional structure Γ_j, i.e., $v(S_k) = \sum_{\forall i \in S_k \in \Gamma_j} u_i$. It can be proved that there will be at least one absorbing state in the Markov chain if the utility summation of the singleton-coalitions is less than that of the non-singleton coalition [40]. In other words, the players attempt to form a coalition that improves their utility, because the utility of players is just equal to zero when they act alone $\left(\sum_{\forall \{i\} \in \Gamma_1} v(\{i\}) = 0\right)$. This property implies that the coalition merging always converges to one of the coalitional structures with a grand coalition, and remains there. Those coalitional structures are thus absorbing states in the DTMC.

Let P be the transition matrix whose entries are the probabilities $P_{\Gamma_j \Rightarrow \Gamma_{j'}}$ that the coalitional structure changes from the current structure Γ_j to the new structure $\Gamma_{j'}$ in a time step [41, 42], denoted by

$$P = \begin{bmatrix} P_{\Gamma_1 \Rightarrow \Gamma_1} & \cdots & P_{\Gamma_1 \Rightarrow \Gamma_j} & \cdots & P_{\Gamma_1 \Rightarrow \Gamma_\gamma} \\ \vdots & \ddots & \vdots & & \vdots \\ P_{\Gamma_j \Rightarrow \Gamma_1} & \cdots & P_{\Gamma_j \Rightarrow \Gamma_j} & \cdots & P_{\Gamma_j \Rightarrow \Gamma_\gamma} \\ \vdots & & \vdots & \ddots & \vdots \\ P_{\Gamma_\gamma \Rightarrow \Gamma_1} & \cdots & P_{\Gamma_\gamma \Rightarrow \Gamma_j} & \cdots & P_{\Gamma_\gamma \Rightarrow \Gamma_\gamma} \end{bmatrix}. \tag{8}$$

Let $Z_{\Gamma_j \Rightarrow \Gamma_{j'}}$ be the set of coalitions whose members can change when the transition from the current coalitional structure Γ_j to the new coalitional structure $\Gamma_{j'}$ occurs. For the probabilistic greedy merging, the transition probability $P_{\Gamma_j \Rightarrow \Gamma_{j'}}$ in (8) can be calculated by

$$P_{\Gamma_j \Rightarrow \Gamma_{j'}} = \frac{\left[\sigma_{\Gamma_j \Rightarrow \Gamma_{j'}}\right]\left[u_{\mathrm{CH}}\left(S_k^{\mathrm{new}}\right)\right]}{\omega_{\Gamma_j \Rightarrow \Gamma_{j'}}} \; ; \; \begin{array}{l} S_k^{\mathrm{new}} \in \Gamma_{j'}, \\ \mathrm{CH} \in S_k \in Z_{\Gamma_j \Rightarrow \Gamma_{j'}}, \end{array} \tag{9}$$

where $u_{\mathrm{CH}}(S_k^{\mathrm{new}})$ is the utility of a node i that used to be the cluster head node of the coalition in $Z_{\Gamma_j \Rightarrow \Gamma_{j'}}$, and it still has a CH rank when changes to the new coalition S_k^{new}. The utility is normalized by dividing by $\omega_{\Gamma_j \Rightarrow \Gamma_{j'}}$ so that $0 \le P_{\Gamma_j \Rightarrow \Gamma_{j'}} \le 1$. The normalization term $\omega_{\Gamma_j \Rightarrow \Gamma_{j'}}$ is the sum of utilities obtained from CH node in every possible coalitional structure, which is calculated by

$$\omega_{\Gamma_j \Rightarrow \Gamma_{j'}} = \sum_{j'=1}^{\gamma} \left[\sigma_{\Gamma_j \Rightarrow \Gamma_{j'}}\right]\left[u_{\mathrm{CH}}\left(S_k^{\mathrm{new}}\right)\right] \; ; \; \begin{array}{l} S_k^{\mathrm{new}} \in \Gamma_{j'}, \\ \mathrm{CH} \in S_k \in Z_{\Gamma_j \Rightarrow \Gamma_{j'}}. \end{array} \tag{10}$$

The variable $\sigma_{\Gamma_j \Rightarrow \Gamma_{j'}}$ takes on a binary values (0 or 1) to enforce the condition that only the transition to a new coalitional structure with better utility is allowed. That is

$$\sigma_{\Gamma_j \Rightarrow \Gamma_{j'}} = \begin{cases} 1 \; ; \; u_{\mathrm{CH}}\left(S_k^{\mathrm{new}}\right) \geq u_{\mathrm{CH}}(S_k) \\ 0 \; ; \; \text{otherwise.} \end{cases} \tag{11}$$

Therefore, we can solve the probability of absorption into coalitional structures with a grand coalition by rearranging and partitioning the transition matrix P as

$$P = \left[\begin{array}{c|c} Q & R \\ \hline 0 & I \end{array}\right], \tag{12}$$

where Q is a matrix of transition probabilities between the transient states, R is a matrix of transition probabilities from the transient states to the absorbing states, and I is an identity matrix [41, 42]. For an absorbing DTMC, $(I - Q)^{-1}$ is the fundamental matrix that shows the expected time to absorption, and $(I - Q)^{-1}R$ is the matrix whose elements are the probabilities that each absorbing state is reached. The state which has the highest probability of absorption becomes the most probable coalitional structure.

3.5 Splitting process

In the non-singleton-coalition, any node whose utility has decreased is becoming unstable, and is ready to leave the current coalition to find a new one that can improve its utility. The condition for leaving is when its utility is less than zero, meaning that splitting itself off to form a new singleton-coalition is better than staying in the current coalition. Let the current coalitional structure be $\Gamma_j = \{S_k\}$. The members of the current coalition S_k can split to multiple new coalitions $S_k^{\mathrm{new}}, S_{k'}^{\mathrm{new}}$ where $S_k^{\mathrm{new}} \cap S_{k'}^{\mathrm{new}} = \emptyset$ if they obtain higher utility than the current coalition, i.e., $u_i(S_k^{\mathrm{new}}) \geq 0 \geq u_i(S_k)$ and $u_i(S_{k'}^{\mathrm{new}}) \geq 0 \geq u_i(S_k)$. The new coalitional structure after splitting is $\Gamma_{j'}^{\mathrm{new}} = \{S_k^{\mathrm{new}}, S_{k'}^{\mathrm{new}}\}$.

4 Performance evaluation results

We use Python and MATLAB for coding the simulations and have separated the experiments into two parts, static and dynamic scenarios. The static scenarios are used to evaluate the utilities of grand coalitions achieved by the merging process described in Sect. 3.2 in snapshots of networks based on the DTMC analysis. The weight coefficient ratio ($\alpha : \tau$) for balancing the importance of coalition area and transmission delay time is varied to investigate how this ratio affects the coalitional structures. The effect of the reward term ρ that an RSU attracts vehicles to join its coalition is investigated by setting the RSU as a sink node.

In dynamic scenarios, vehicles randomly move and their positions are updated every 0.1 s. The penalty term ψ determined from the relative velocity between nodes will be added to calculate the utility. In each snapshot, the corresponding DTMC is simulated to determine the coalitional structure and its utility to demonstrate how the coverage area evolves over the sequence of network snapshots and its relationship to the utility achieved. The performance comparison between our CGG-MC model and the non-cooperation model is also evaluated. We collect the coalition area results and the successful packet delivery rate obtained from each network snapshot and plot them against the number of vehicles in the network.

4.1 Static scenarios

The map of $50\,\mathrm{m} \times 50$ m with three vehicles is considered. The vehicles are represented by a set of three players $N = \{1, 2, 3\}$ located at the positions $(x_1, y_1) = (10, 25)$, $(x_2, y_2) = (20, 30)$, and $(x_3, y_3) = (30, 20)$. Initially, each player acts alone with un-cluster (UN) status. The initial coalitional structure is $\Gamma_1 = \{S_1 = (1), S_2 = (2), S_3 = (3)\}$. The initial utility of each player is zero because they each act alone. Shown in Fig. 3, the initial coalitional structure can converge to the grand coalition by passing through any of the six transient states (Γ_2 to Γ_7) and finishing in any of the nine absorbing states (Γ_8 to Γ_{16}).

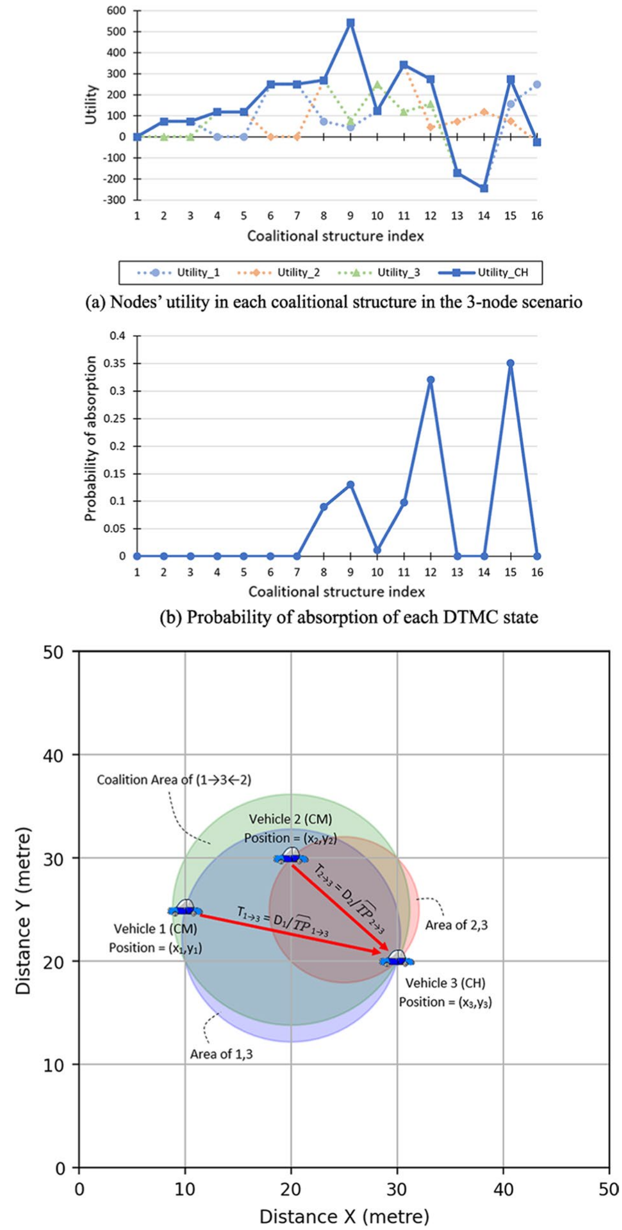
With the weight coefficients $\alpha = 2$ and $\tau = 1$ in the utility function, the utilities of all 16 coalitional structures are shown in Fig. 4a. Based on the probabilistic greedy merging, these utilities are used to compute the transition matrix in (12) and the probability of absorption for each coalitional structure is calculated as shown in Fig. 4b. The result shows that the most probable coalitional structure is $\Gamma_{15} = \{(2 \rightarrow 3 \leftarrow 1)\}$, with an absorption probability of 0.35 or a 35% chance of being formed. However, it does not have the maximum utility ($u_{\mathrm{CH}}^{\Gamma_{15}} = 275$). In this case, the maximum utility structure is Γ_9 ($u_{\mathrm{CH}}^{\Gamma_9} = 545$), with an absorption probability of 0.13 or a 13% chance of being formed.

To justify the utility of the grand coalition achieved by the probabilistic greedy merging, we compute the average utility weighted by the absorption probabilities and convert it to the percentage of the maximum utility. Letting Π_{Γ_j} be the state absorption probability of a coalitional structure Γ_j obtained by solving the corresponding DTMC, the max-normalized weighted average utility for this 3-node case is given by

$$\begin{aligned} u_{\mathrm{CH}}^{\mathrm{avg}} &= \frac{\sum_{j=8}^{16} \left(\Pi_{\Gamma_j}\right)\left(u_{\mathrm{CH}}^{\Gamma_j}\right)}{\left(u_{\mathrm{CH}}^{\Gamma_9}\right)} \times 100\% \\ &= \frac{(0.09)(269) + (0.13)(545) + \cdots + (0.35)(275)}{545} \times 100\% \\ &= 58\% \end{aligned}$$

Ten 3-node networks are generated with different node positions and the max-normalized weighted average utility of each case is computed as shown in Fig. 5, with the 95% confidence interval of $64.7 \pm 3.58 = [61.12,\ 68.28]$. Therefore, the probabilistic greedy merging can achieve between 61 and 68% of the maximum utility on average.

Next, we experiment with different weight coefficient ratios ($\alpha : \tau$) by varying the weight coefficient $\alpha = \{1, 2, \ldots, 10\}$ while keeping τ equal to 1. Increasing α places more importance on the coalition area than the transmission delay time. When the weight coefficient ratio $(\alpha : \tau) = (1{:}1)$, Table 1 and Fig. 6 show that the probable coalitional structures are $\Gamma_2 = \{(1 \rightarrow 2)\ (3)\}$ and $\Gamma_3 = \{(1 \leftarrow 2)\ (3)\}$, which only group nearby nodes in the same coalition. If the value of α increases, other coalitional structures in the absorbing states become more probable. Therefore, the coalitional structures that have a larger distance between nodes can be formed without being restricted by transmission



(a) Nodes' utility in each coalitional structure in the 3-node scenario

(b) Probability of absorption of each DTMC state

(c) The most probable structure Γ_{15}

Fig. 4 DTMC results of three vehicles with weight coefficients $\alpha = 2, \tau = 1$

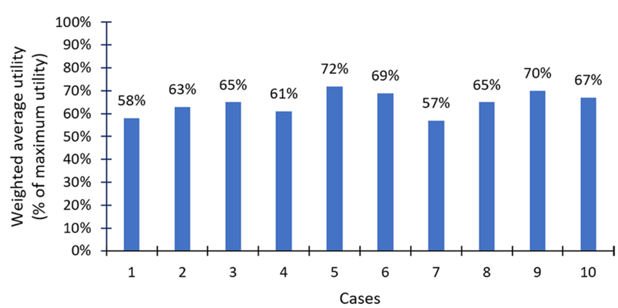


Fig. 5 Average utility weighted by probability of absorption and normalized by maximum utility

Table 1 Probability of absorption of each coalitional structure for weight coefficient $\alpha = 1, 2, \ldots, 10$ and $\tau = 1$

Coalitional structure	Probability of absorption									
1	0	0	0	0	0	0	0	0	0	0
2	0.50	0	0	0	0	0	0	0	0	0
3	0.50	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0
8	0	0.09	0.09	0.08	0.08	0.08	0.08	0.07	0.07	0.07
9	0	0.13	0.10	0.09	0.08	0.08	0.08	0.08	0.08	0.08
10	0	0.01	0.02	0.02	0.16	0.16	0.16	0.16	0.16	0.16
11	0	0.10	0.09	0.08	0.08	0.08	0.08	0.08	0.07	0.07
12	0	0.32	0.32	0.31	0.18	0.18	0.18	0.18	0.17	0.17
13	0	0	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.11
14	0	0	0	0.03	0.03	0.04	0.04	0.04	0.04	0.04
15	0	0.35	0.35	0.34	0.33	0.33	0.20	0.20	0.20	0.15
16	0	0	0.02	0.03	0.03	0.04	0.17	0.17	0.17	0.13
	1:1	**2:1**	**3:1**	**4:1**	**5:1**	**6:1**	**7:1**	**8:1**	**9:1**	**10:1**
	($\alpha : \tau$)									

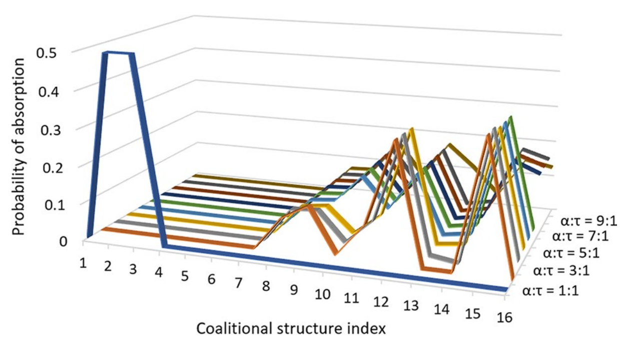
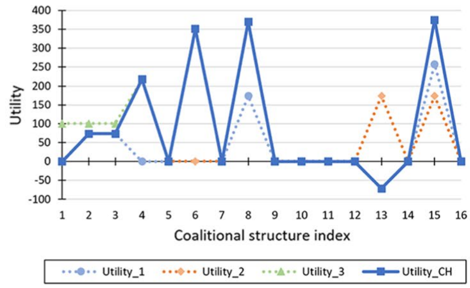


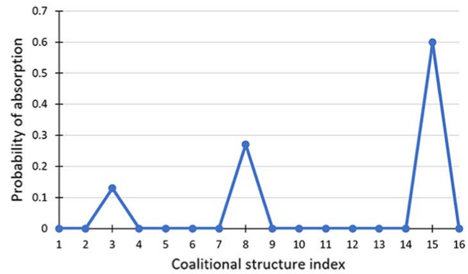
Fig. 6 Probability of absorption of each coalitional structure for weight coefficient $\alpha = 1, 2, \ldots, 10$ and $\tau = 1$

delay time, especially when the value of α is significantly greater than τ. On the other hand, when there is a small value of α (e.g., $\alpha = 1$), the coalitional structures with large distances between nodes are not probable, and thus cannot be formed due to the restriction of the transmission delay time.

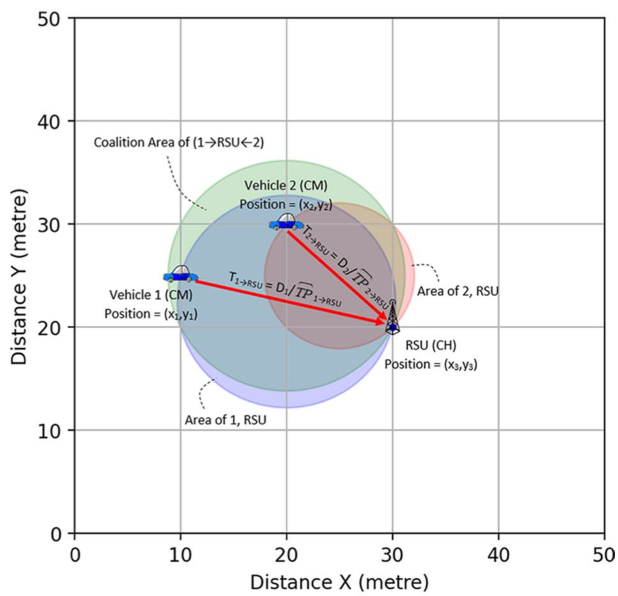
Next, we replace vehicle 3 with an RSU to observe how the absorption probability of each coalitional structure changes. In this case, the coalition where the data is flowing out of the RSU is not allowed and the transition probability to such a state becomes zero. Any node i in the same coalition with the RSU has the additional reward $\rho_i = 100$ to its utility. The utility of each node for weight coefficients $\alpha = 2$ and $\tau = 1$ is shown in Fig. 7a. The probability of absorption obtained from DTMC is shown in Fig. 7b. The most probable coalitional structure is still $\Gamma_{15} = \{(2 \rightarrow \text{RSU} \leftarrow 1)\}$ as shown in Fig. 7c. This structure has an RSU as a CH node, and vehicle 1 and 2 as the CM nodes. This



(a) Nodes' utility in each coalitional structure in the 3-node scenario



(b) Probability of absorption of each DTMC state



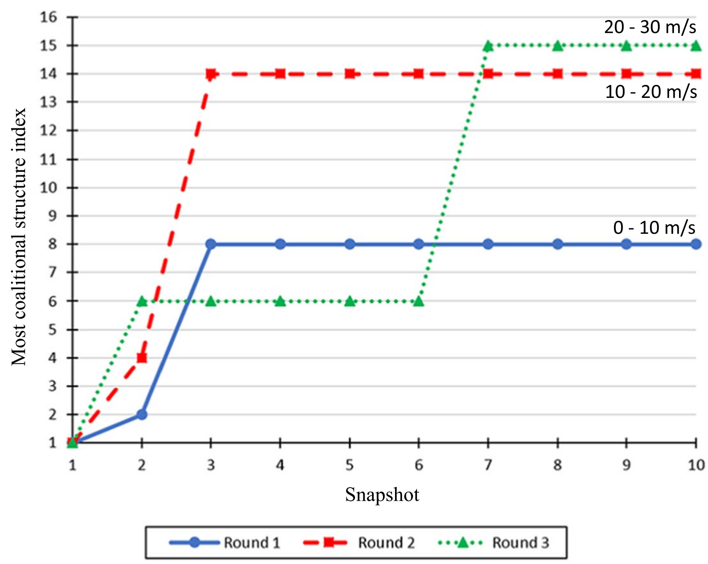
(c) The most probable structure Γ_{15}

Fig. 7 DTMC results while vehicle 3 is replaced by RSU with weight coefficients $\alpha = 2, \tau = 1$

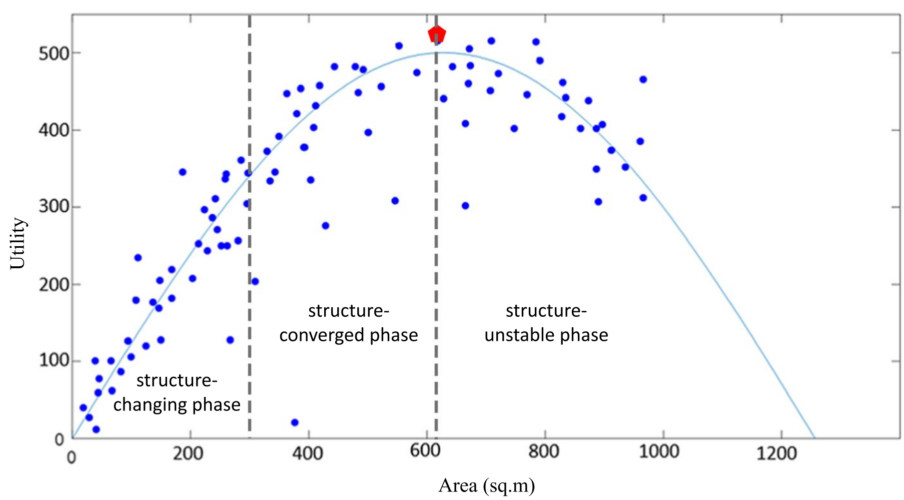
experiment demonstrates that the reward term helps the RSU attract vehicles to join its coalition. Without the reward term, vehicles may disconnect from the RSU as demonstrated in the coalitional structure $\Gamma_3 = \{(1 \leftarrow 2)\ (3)\}$.

4.2 Dynamic scenarios

In the dynamic scenarios, every parameter setup is the same as in the static scenario, except vehicle movement is added. We apply a random velocity $\vec{v}_i$ within a given range to each vehicle. The position of each vehicle is updated every 0.1 s as $(x_i, y_i)^{\text{new}} = (x_i, y_i)^{\text{current}} + \vec{v}_i t$ where $t = 0.1$, and the utility of each vehicle is also updated. With vehicle movement, the utility function has a penalty value calculated from the relative velocity between a node and its parent. The coalitional structure in each network snapshot is analyzed by using DTMC. We sample and record the most probable coalitional structure over a sequence of network snapshots in these velocity range settings: 0–10 m/s, 10–20 m/s, and 20–30 m/s. Figure 8a shows the most probable coalitional structure over a sequence of network snapshots. For the velocity ranges of 0–10 m/s and 10–20 m/s, the most probable coalitional structures remain stable for many snapshots while the structures change more often as the velocity range increases. This experiment demonstrates that high relative velocity reduces the stability of the coalitional structure. The penalty term prevents the nodes with



(a) Most probable coalitional structure under different velocity range



(b) Scatter plot between coalitional area vs. utility value of CH node over 100 network snapshots

Fig. 8 Most probable coalitional structures under vehicles' movement and the highest coalition area from utility convergence high relative velocity from joining a coalition and helps the coalition avoid frequent reconstruction.

Figure 8b shows the scatter plot of utilities of CH nodes in various coalitional structures from network snapshots versus the coalition areas. The result is separated into 3 phases: structure-changing phase, structure-converged phase, and structure-unstable phase. In the structure-changing phase, the nodes form and change their coalitional structure to obtain better utilities while moving and covering larger coalition areas (0–315 sq.m). In the structure-converged phase, the coalitional structures remain stable although the vehicles are moving. The utility reaches its maximum at $u_{\text{CH}} = 512$ at the coverage area of 627 sq.m. After this phase, the coalitional structure becomes unstable as soon as the transmission delay time begins to diminish the utility. The vehicles will abandon their current coalitional structure and reconstruct a new one in the next snapshot. This experiment reveals that the highest coverage area is achieved when the CH's utility reaches its highest value.

4.3 Comparison to non-cooperation scheme

We compare the performance of our CGG-MC model to the non-cooperation scheme by setting the dynamic scenario on a large size map of 1 km $\times$ 1 km, with 100 vehicles and an RSU as shown in Fig. 9. The expiration time T_{expire} is set to 1 s before the RSU rejects data packets. With this constrained time, the RSU selects only nodes that can serve the channel to receive the data packets in time. In Fig. 9a, vehicles non-cooperatively compete to connect to the RSU, and only a few nodes located within the transmission range of the RSU can successfully send data to the RSU. Figure 9b shows that our model cooperatively forms a hierarchy multi-hop cluster with a coalition that has an average coverage area larger than that of the non-cooperation model.

Figure 10a shows the average coverage area obtained from each network snapshot when the number of vehicles increases. The average coverage area result shows that the non-cooperation model and the CGG-MC model both perform equally well with a few vehicles, but the CGG-MC attains approximately 166% more of the average coverage area when the number of vehicles increases beyond 30 nodes, because the nodes in the coalition are able to cooperatively relay the data packets to the RSU. Figure 10b shows

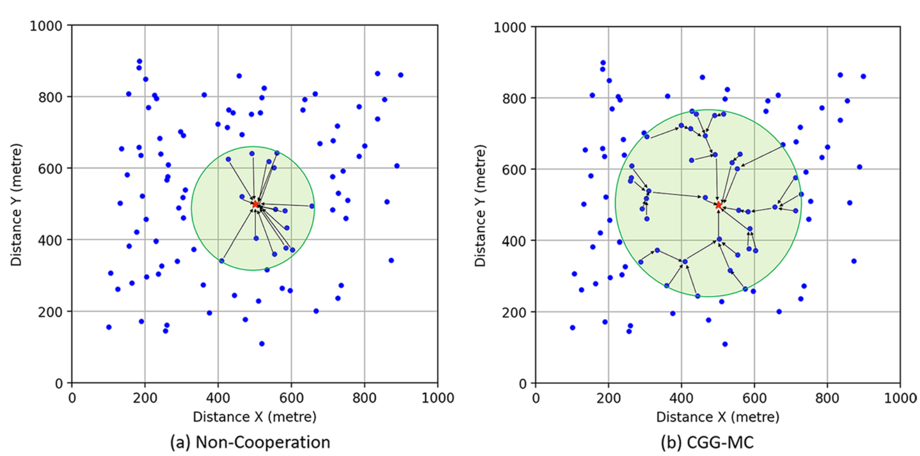


Fig. 9 Coverage area of non-cooperation model versus CGG-MC model

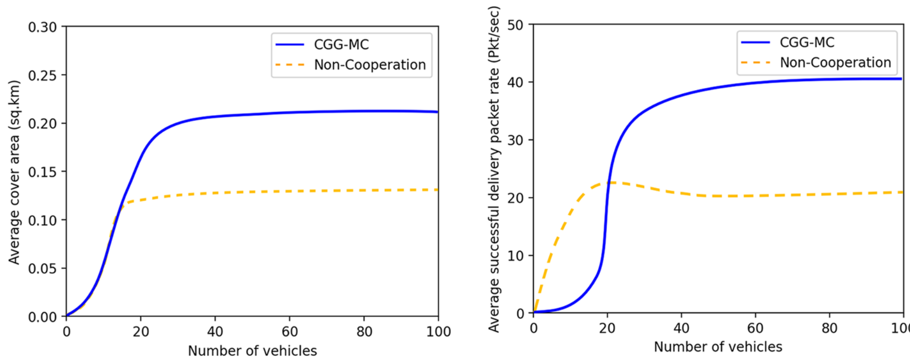


Fig. 10 Average coverage area and average successful packet delivery rate

the average successful packet delivery rate calculated from (4) and (5). The non-cooperation model outperforms the CGG-MC when a few vehicles are available at the beginning, however, the CGG-MC attains approximately 195% more of the delivery rate when the number of vehicles increases beyond 50 nodes.

5 Conclusion

In this paper, we propose the coalitional graph game to form multi-hop clustering (CGG-MC) among vehicles and RSUs, where the objective is to maximize the network coverage area subject to the transmission delay time constraint. The solution is obtained by using the probabilistic greedy merging approach to construct a sequence of coalitions in a distributed manner. A DTMC is used to analyze the probability of the resulting coalitional structures. The results from a three-node network case show that the grand coalition can attain the utility around 61–68% of the maximum utility from the centralized solutions. The effects of weight coefficients α and τ for balancing the importance of the coverage area and the packet transmission delay time penalty in the utility function on the stable coalition structure solution are also explored. The results suggest that we should increase α to form a larger coalition coverage area, and increase τ to reduce the transmission delay time. When an RSU is a sink node, vehicles may not connect and form a coalition with the RSU, and the reward term helps the RSU attract vehicles to join its coalition.

In dynamic scenarios, with vehicle movement, the DTMC analysis shows that the duration where the coalitional structure remains stable decreases as the velocity increases to a higher range. The relative velocity between nodes may be high and reduce the stability of the coalitional structure if each vehicle increases its velocity. The penalty term prevents the coalition-joining of the nodes with high relative velocity, helps the coalition to avoid frequent reconstruction, and prolongs the lifetime of the coalitional structure. The relationship between the utility value and the coalition area is concave. When the coalition area increases, the utility of the CH node also increases until reaching the maximum point. After that point, the utility reduces, and the coalitional structure becomes unstable.

By using the probabilistic merging approach to obtain the solution for the coalitional graph game, our CGG-MC model can handle a large problem size with a large number of vehicles. The comparisons between the CGG-MC model and the non-cooperation model show that our CGG-MC model outperforms the non-cooperation model by covering approximately 166% more of the coverage area. The CGG-MC achieves an expansion of the network area by cooperatively forming multi-hop connections to more distant nodes with the stable coalitional structure.

Abbreviations

ALM	Aggregate local mobility
AMC	Advanced multi-hop clustering
APROVE	Affinity propagation for vehicular network
BOR	Border
CAV	Connected automated vehicles
CBLR	Cluster-based location routing
CEN	Centroid
CGG-MC	Coalitional graph game of multi-hop clustering
CH	Cluster head
CM	Cluster member
CMH	Cluster member head (CH and CM simultaneously)
CVCG	Cooperative V2V-aided transmission coalitional game
DBA-MAC	Dynamic backbone-assisted medium access control
DCA	Distributed clustering algorithm
DMAC	Distributive mobility adaptive clustering
DMCNF	Distributed multi-hop clustering based on neighborhood follow
DTMC	Discrete-time Markov chain
EGT	Evolutionary game theory
GPS	Global positioning system
ID	Identification
ITS	Intelligent transport system
LID	Lowest-ID
LTE	Long-term evolution
MC	Mesh client
MGADA	Multi-player game theoretic algorithm for intra-cluster data aggregation
MOBIC	Mobility based metric clustering
MR	Mesh router
NTU	Non-transferable utility
NTU-CGGA	NTU-Coalition graph game algorithm
OBU	On-board unit
PCD	Popular content distribution
RMAC	Robust mobility-aware clustering
RSS	Received signal strength
RSU	Roadside unit
SNA	Social network analysis
UN	Un-cluster
V2I	Vehicle-to-infrastructure communication
V2V	Vehicle-to-vehicle communication
VANET	Vehicular ad hoc network
VMaSC	Vehicular multi-hop algorithm for stable clustering
VWCA	Vehicular weighted clustering algorithm
WAVE	Wireless access in vehicular environments
WMN	Wireless mesh network
WSN	Wireless sensor network

Acknowledgements
The authors would like to express gratefulness for the Petchra Pra Jom Klao research scholarship and King Mongkut's University of Technology Thonburi for the support.

Author informations

Siwapon Charoenchai received the B.Eng. degree in electrical communication engineering from Mahidol University, Thailand, in 2010. He is currently studying for the Ph.D. degree in electrical and computer engineering at King Mongkut's University of Technology Thonburi, Thailand. His current research interests include vehicular ad hoc networks, coalitional game theory, and data aggregation.

Peerapon Siripongwutikorn received the M.S. and Ph.D. degrees in telecommunications from the University of Pittsburgh, Pittsburgh, PA, in 1998 and 2003, respectively. He is an Associate Professor in the Department of Computer Engineering, King Mongkut's University of Technology Thonburi, Thailand. His current research interests include ad hoc networks, dynamic resource allocation and control in communication networks, and network performance analysis.

Author contributions

SC and PS conceived the idea. SC designed the model, programmed the simulation codes, collected and analyzed the experimental results, and wrote the manuscript. PS gave overall advice, improved and corrected the model, reviewed and edited the manuscript. All authors read and approved the final manuscript.

Funding

This research is supported by the Petchra Pra Jom Klao research scholarship Ph.D. program, funded by King Mongkut's University of Technology Thonburi.

Availability of data and materials

The simulation codes and result datasets used for the evaluation and analysis in this research are available from the corresponding authors on reasonable requests.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 17 April 2021 Accepted: 9 July 2022
Published online: 09 August 2022

References

1. S. Wongdeethai, P. Siripongwutikorn, Collecting road traffic information using vehicular ad hoc networks. EURASIP J. Wirel. Commun. Netw. **2016**(1), 9 (2016). https://doi.org/10.1186/s13638-015-0513-0
2. C. Vitale, N. Piperigkos, C. Laoudias, G. Ellinas, J. Casademont, J. Escrig, A. Kloukiniotis, A.S. Lalos, K. Moustakas, R. Diaz Rodriguez, D. Baños, G. Roqueta Crusats, P. Kapsalas, K.-P. Hofmann, P.S. Khodashenas, Caramel: results on a secure architecture for connected and autonomous vehicles detecting GPS spoofing attacks. EURASIP J. Wirel. Commun. Netw. **2021**(1), 115 (2021). https://doi.org/10.1186/s13638-021-01971-x
3. C. Wang, X. Cheng, J. Li, Y. He, K. Xiao, A survey: applications of blockchain in the internet of vehicles. EURASIP J. Wirel. Commun. Netw. **2021**(1), 77 (2021). https://doi.org/10.1186/s13638-021-01958-8
4. C. Chen, J. Hu, T. Qiu, M. Atiquzzaman, Z. Ren, CVCG: cooperative V2V-aided transmission scheme based on coalitional game for popular content distribution in vehicular ad-hoc networks. IEEE Trans. Mob. Comput. **18**(12), 2811–2828 (2019). https://doi.org/10.1109/TMC.2018.2883312
5. B. Ahmed, A.W. Malik, T. Hafeez, N. Ahmed, Services and simulation frameworks for vehicular cloud computing: a contemporary survey. EURASIP J. Wirel. Commun. Netw. **2019**(1), 4 (2019). https://doi.org/10.1186/s13638-018-1315-y
6. A. Katiyar, D. Singh, R.S. Yadav, State-of-the-art approach to clustering protocols in VANET: a survey. Wirel. Netw. **26**(7), 5307–5336 (2020). https://doi.org/10.1007/s11276-020-02392-2
7. C. Cooper, D. Franklin, M. Ros, F. Safaei, M. Abolhasan, A comparative survey of VANET clustering techniques. IEEE Commun. Surv. Tutor. **19**(1), 657–681 (2017). https://doi.org/10.1109/COMST.2016.2611524
8. S. Shukry, Stable routing and energy-conserved data transmission over wireless sensor networks. EURASIP J. Wirel. Commun. Netw. **2021**(1), 36 (2021). https://doi.org/10.1186/s13638-021-01925-3
9. M.A. Alharbi, M. Kolberg, M. Zeeshan, Towards improved clustering and routing protocol for wireless sensor networks. EURASIP J. Wirel. Commun. Netw. **2021**(1), 46 (2021). https://doi.org/10.1186/s13638-021-01911-9
10. Y. Li, I.-R. Chen, Dynamic agent-based hierarchical multicast for wireless mesh networks. Ad Hoc Netw. **11**(6), 1683–1698 (2013). https://doi.org/10.1016/j.adhoc.2013.03.004
11. L. Wischhof, A. Ebner, H. Rohling, Information dissemination in self-organizing intervehicle networks. IEEE Trans. Intell. Transp. Syst. **6**(1), 90–101 (2005). https://doi.org/10.1109/TITS.2004.842407
12. F. Cuckov, M. Song, Geocast-driven structureless information dissemination scheme for vehicular ad hoc networks, in *2010 IEEE Fifth International Conference on Networking, Architecture, and Storage* (2010), pp. 325–332. https://doi.org/10.1109/NAS.2010.51
13. S. Latif, S. Mahfooz, B. Jan, N. Ahmad, Y. Cao, M. Asif, A comparative study of scenario-driven multi-hop broadcast protocols for VANETs. Veh. Commun. **12**, 88–109 (2018). https://doi.org/10.1016/j.vehcom.2018.01.009
14. P. Fan, J.G. Haran, J. Dillenburg, P.C. Nelson, Cluster-based framework in vehicular ad-hoc networks, in *Ad-Hoc, Mobile, and Wireless Networks*. ed. by V.R. Syrotiuk, E. Chávez (Springer, Berlin, Heidelberg, 2005), pp. 32–42
15. M. Gerla, J. Tzu-Chieh Tsai, Multicluster, mobile, multimedia radio network. Wirel. Netw. **1**(3), 255–265 (1995). https://doi.org/10.1007/BF01200845
16. P. Basu, N. Khan, T.D.C. Little, A mobility based metric for clustering in mobile ad hoc networks, in *Proceedings 21st International Conference on Distributed Computing Systems Workshops* (2001), pp. 413–418. https://doi.org/10.1109/CDCS.2001.918738
17. E. Souza, I. Nikolaidis, P. Gburzynski, A new aggregate local mobility (ALM) clustering algorithm for VANETS, in *2010 IEEE International Conference on Communications* (2010), pp. 1–5. https://doi.org/10.1109/ICC.2010.5501789
18. Z.Y. Rawashdeh, S.M. Mahmud, A novel algorithm to form stable clusters in vehicular ad hoc networks on highways. EURASIP J. Wirel. Commun. Netw. **2012**(1), 15 (2012). https://doi.org/10.1186/1687-1499-2012-15
19. C. Shea, B. Hassanabadi, S. Valaee, Mobility-based clustering in VANETs using affinity propagation, in *GLOBECOM 2009—2009 IEEE Global Telecommunications Conference* (2009), pp. 1–6. https://doi.org/10.1109/GLOCOM.2009.5425236

20. A. Daeinabi, A.G. Pour Rahbar, A. Khademzadeh, VWCA: An efficient clustering algorithm in vehicular ad hoc networks. J. Netw. Comput. Appl. **34**(1), 207–222 (2011). https://doi.org/10.1016/j.jnca.2010.07.016
21. M.M. Hamdi, L. Audah, S.A. Rashid, Data dissemination in VANETs using clustering and probabilistic forwarding based on adaptive jumping multi-objective firefly optimization. IEEE Access **10**, 14624–14642 (2022). https://doi.org/10.1109/ACCESS.2022.3147498
22. S. Ucar, S.C. Ergen, O. Ozkasap, Multihop-cluster-based IEEE 802.11p and LTE hybrid architecture for VANET safety message dissemination. IEEE Trans. Veh. Technol. **65**(4), 2621–2636 (2016). https://doi.org/10.1109/TVT.2015.2421277
23. Y. Chen, M. Fang, S. Shi, W. Guo, X. Zheng, Distributed multi-hop clustering algorithm for VANETs based on neighborhood follow. EURASIP J. Wirel. Commun. Netw. **2015**(1), 98 (2015). https://doi.org/10.1186/s13638-015-0327-0
24. A. Katiyar, D. Singh, R.S. Yadav, Advanced multi-hop clustering (AMC) in vehicular ad-hoc network. Wirel. Netw. **28**(1), 45–68 (2022). https://doi.org/10.1007/s11276-021-02822-9
25. R.T. Goonewardene, F.H. Ali, E. Stipidis, Robust mobility adaptive clustering scheme with support for geographic routing for vehicular ad hoc networks. IET Intell. Transp. Syst. **3**, 148–15810 (2009)
26. L. Bononi, M. di Felice, A cross layered mac and clustering scheme for efficient broadcast in vanets, in *2007 IEEE International Conference on Mobile Adhoc and Sensor Systems* (2007), pp. 1–8. https://doi.org/10.1109/MOBHOC.2007.4428735
27. R.A. Santos, R.M. Edwards, A. Edwards, Cluster-based location routing algorithm for inter-vehicle communication, in *IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall*, vol. 2 (2004), pp. 914–9182. https://doi.org/10.1109/VETECF.2004.1400154
28. S. Basagni, Distributed clustering for ad hoc networks, in *Proceedings Fourth International Symposium on Parallel Architectures, Algorithms, and Networks (I-SPAN'99)* (1999), pp. 310–315. https://doi.org/10.1109/ISPAN.1999.778957
29. Y. Chen, S. Weng, W. Guo, N. Xiong, A game theory algorithm for intra-cluster data aggregation in a vehicular ad hoc network. Sensors (2016). https://doi.org/10.3390/s16020245
30. A.A. Khan, M. Abolhasan, W. Ni, An evolutionary game theoretic approach for stable and optimized clustering in VANETs. IEEE Trans. Veh. Technol. **67**(5), 4501–4513 (2018). https://doi.org/10.1109/TVT.2018.2790391
31. Z. Han, D. Niyato, W. Saad, T. Başar, A. Hjørungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications* (Cambridge University Press, Cambridge, 2011), p. 188. https://doi.org/10.1017/CBO9780511895043
32. W. Saad, Z. Han, M. Debbah, A. Hjorungnes, T. Basar, Coalitional game theory for communication networks. IEEE Signal Process. Mag. **26**(5), 77–97 (2009). https://doi.org/10.1109/MSP.2009.000000
33. K. Akkarajitsakul, E. Hossain, D. Niyato, Cooperative packet delivery in hybrid wireless mobile networks: a coalitional game approach. IEEE Trans. Mob. Comput. **12**(5), 840–854 (2013). https://doi.org/10.1109/TMC.2012.46
34. P. Hang, C. Lv, C. Huang, Y. Xing, Z. Hu, Cooperative decision making of connected automated vehicles at multi-lane merging zone: a coalitional game approach. IEEE Trans. Intell. Transp. Syst. (2021). https://doi.org/10.1109/TITS.2021.3069463
35. R. Myerson, Graphs and cooperation in games. Math. Oper. Res. **2**(3), 225–229 (1977)
36. L. Zhai, H. Ji, X. Li, Y. Tang, Coalition graph game for joint relay selection and resource allocation in cooperative cognitive radio networks, in *2012 IEEE Global Communications Conference (GLOBECOM)* (2012), pp. 1818–1823. https://doi.org/10.1109/GLOCOM.2012.6503379
37. T. Wang, L. Song, Z. Han, Coalitional graph games for popular content distribution in cognitive radio VANETs. IEEE Trans. Veh. Technol. **62**(8), 4010–4019 (2013). https://doi.org/10.1109/TVT.2013.2245353
38. P. Gupta, P.R. Kumar, The capacity of wireless networks. IEEE Trans. Inf. Theory **46**(2), 388–404 (2000). https://doi.org/10.1109/18.825799
39. W. Liu, X. He, Z. Huang, Y. Ji, Transmission capacity characterization in VANETs with enhanced distributed channel access. Electronics (2019). https://doi.org/10.3390/electronics8030340
40. T. Arnold, U. Schwalbe, Dynamic coalition formation and the core. J. Econ. Behav. Organ. **49**(3), 363–380 (2002)
41. J.G. Kemeny, J.L. Snell, *Finite Markov Chains* (Springer, New York, 1983)
42. H.A. Taha, *Operations Research: An Introduction*, 8th edn. (Prentice-Hall Inc, Upper Saddle River, 2006)

Publisher's Note