# Proximate Sharing of Downloaded Geo Data Using the Vehicular Social Network (VSN) Technique

*Chung-Ming* Huang* and *Yu-Fen* Chen

Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan

**Abstract.** This work proposed a clustering scheme for Point Of Interests' (POIs') geo data sharing in Vehicular Social Network (VSN). In the proposed scheme, some vehicles can form a cluster to downloaded POIs' geo data when they are approaching to a new set of POIs. One vehicle is chosen as the cluster head to download POIs' geo data using its cellular network and shares the downloaded POIs' geo data to its cluster members using IEEE 802.11p network. This work (1) uses GPS to get vehicles' locations to calculate the timing of triggering the clustering process, and (2) proposes a clustering method to organize a group of vehicles that are proximate with each other for a while during their touring to become a cluster. The simulation results show that the proposed scheme can have higher available distance and higher successful ratio of the complete sharing of downloaded POIs' geo data.

## 1 Introduction

Vehicular Social Network (VSN) is a special kind of Mobile Social Networks (MSN) that can connect nearby vehicles with common interest to perform some social interactions, e.g., downloading and sharing some geo data of POIs. An example of VSN's usage is as follows. Vehicles can download geo data of Point-Of-Interest (POI) such as restaurants, historical site, national parks, etc., according to vehicles' current locations and destinations during their journey. If proximate vehicles would like to download the same data and all of them download the same data individually around the same time, it would result in the bursty traffic in cellular network and each vehicle needs to pay the expense of using cellular network. Nearby vehicles can share the same interested geo data with each other using VSN. Let neighboring vehicles that belong to the same VSN organize a cluster [1-3]. Then the cluster head is in charge of downloading POIs' geo data using cellular network and then forwarding the downloaded data to other vehicles using IEEE 802.11p Dedicated Short Range Communication (DSRC) network. As a result, the network traffic and the expense of using cellular network is reduced to 1/n if the cluster contains n vehicles.

To tackle the issue for POIs' geo data sharing in VSN, a clustering scheme was proposed in this work. In this way, proximate vehicles that belong to the same VSN can form a cluster to share downloaded POIs' geo data, in which cluster head downloads POIs' geo data using its cellular network and shares the downloaded data to its members using IEEE 802.11p DSRC network. Two main issues that need to be tackled are as follows. (1) When to trigger POIs' geo data downloading? POIs' geo data downloading can be triggered when it is necessary. For example, a vehicle is approaching a set of new POIs and it needs to pre-download or refresh its cached POIs' geo data before it moves into the coverage of these POIs. The timing of triggering the corresponding geo data downloading is dependent on vehicle's position, speed and the cellular network available bandwidth. (2) When and how to form a cluster to share downloaded POIs' geo data? Depending on the cluster creating/join principle, the clustering process can be commenced.

The rest of this thesis is organized as follows: Section 2 presents related works. Section 3 introduces the architecture and the configuration of the proposed method. Section 4 presents the proposed scheme in details. Section 5 shows the performance analysis. Finally, the conclusion remarks are given in Section 6.

## 2 Related Work

This Section presents related works about (1) VSN and (2) Clustering for proximate sharing of downloaded POIs' geo data in the touring service.

### 2.1 Vehicular Social Network (VSN)

Vehicular Social Network (VSN) is an opportunistic network formed by vehicles on the road. Three main components of a VSN are (1) participants, (2) mobile devices and (3) network infrastructure that are widely used in VSNs [4-6]. In [4], the authors summarized the difference between the MSN and the VSN from the aspect of (1) participants, the participants in (i) MSN are mainly mobile devices held by people and (ii) VSN are mainly the vehicles that are equipped with On-Board

* Corresponding author: huangcm@locust.csie.ncku.edu.tw

Uints (OBUs); (2) MSNs need to consider the energy problem, while vehicles can provide the power resource in VSN. The authors also discussed the architecture of VSN, which can be classified into the following three categories. (i) Centralized VSN: The vehicles in a centralized VSN need to connect to a centralized server, which manages vehicles' interactions in the VSN all the time, even when the two vehicles are physically close with each other. The communication among vehicles is Vehicular-to-Infrastructure (V2I)-based. (ii) Decentralized VSN: In the decentralized VSN, no centralized server exists and the communication between vehicles is opportunistic. That is, data is delivered only when the two peered vehicles encounter with each other opportunistically. The communication among vehicles is Vehicular-to-Vehicular (V2V)-based. (iii) Hybrid VSN: In the hybrid VSN, the communication among vehicles are based on both V2I and V2V communications. In [5], the authors discussed some research challenges of VSN, such as (1) how messages are forwarded in the VSN, (2) how to collect data or handle the context in VSN, (3) how to encourage vehicles involved in the VSN to have sharing and (4) how to distinguish vehicles with different metric in order to provide the corresponding data, e.g., traffic warning messages to the vehicles that are approaching to the road having accident or traffic jam.

### 2.2 Clustering

The clustering scheme is a technique that divides vehicles with similar spatial relationship, e.g., similar relative velocity, the same moving direction and in the same road, into several groups to improve the routing, save the network resource and stability in VANETs [7-8]. In [7], the authors proposed a mobility and stability based cluster scheme to form a stable cluster with low overhead. The mobility evaluation considers vehicle's position, moving direction and velocity. The authors also defined the link lifetime estimation (LLT) to evaluate the connection time in order to promise the cluster stability. Vehicles with the similar mobility pattern gains more chance to stay in the same cluster.

In [8], the authors proposed a clustering algorithm based on the utility function, which consists of the history information and current states of vehicles for selecting a cluster head. The credit history function considers the average available bandwidth, queue length and the accumulated time that the vehicle acted as cluster head in the past. The current state function takes node degree, velocity and distance into account. When it comes to the cluster formation process, the vehicle has the largest utility function is selected as the cluster head.

## 3 Architecture and the Functional Configuration

Figure 1 depicts the architecture and the functional configuration of the proposed method. Each vehicle is equipped an OBU, which consists of an IEEE 802.11p network interface and a 3G/4G cellular network interface, and the GPS system, which can provide vehicle's current location. Additionally, each vehicle periodically (1) transmits a HELLO message, which includes the current time and its vehicle ID, to its proximate vehicles and (2) reports its context, which includes (i) the current time, (ii) its current location, i.e., latitude and longitude, (iii) speed, (iv) direction and (v) IDs of proximate vehicles, i.e., those neighboring vehicles whose transmitted HELLO messages can be received, to the VSN server. The VSN server collects vehicles' reported contexts and manages the cluster creation/join/leave procedure for sharing the downloaded POIs' geo data, which are stored in the POIs' geo data server.

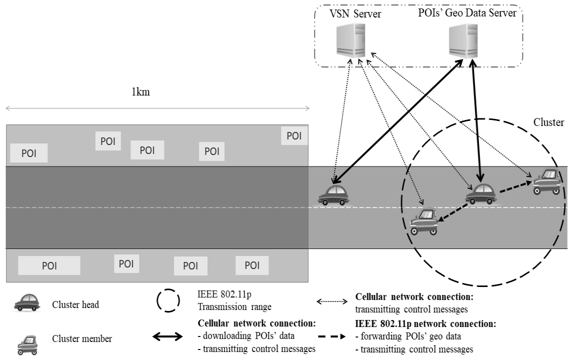


**Fig. 1.** The abstract architecture and the functional configuration of the proposed method.

Referencing to Figure 1, the abstract functional scenario is as follows. Vehicles are moving from the right to the left of the road. To have extra time for viewing POIs' geo data, each vehicle can pre-download or refresh its' cached POIs' geo data before it drives into a set of the ahead POIs. Hence it needs to reserve enough time for downloading POIs' geo data through the cellular network before the vehicle enters into the ahead POI's coverage, i.e., it needs to calculate when/where the vehicle needs to start to download POIs' geo data based on vehicle's speed, location and 3G/4G's networking situation.

For convenient processing, the functional configuration is defined as follows.

(1) It is assumed that POI's geo coverage is a rectangle along the road.

(2) Let POI's data contain some text and at most 5 pictures, for which the maximum size of each picture is 2MB [a]. So, the maximum data volume without considering the text content of one POI is 2MB*5pictures*8bits= 80Mbits.

[a] Our lab has developed a mobile digital culture heritage (M-DCH) exploring platform, which is called Demodulating and Encoding Heritage (DEH). The DEH platform (http://deh.csie.ncku.edu.tw) contains a website and several APPs in Google Play and APPLE Store. The aforementioned specification of a POI is the requirement in the DEH platform, which is for fast downloading of POIs during movement using the wireless mobile network.

(3) According to the report “Mobile Internet Speed Measurement“, which is from Telecom Technology Center (TTC), Taiwan, the bandwidth of 3G/4G while moving is 5.45~6.88(Mbps)/24.04~31.89(Mbps). According to the Traffic Regulation of Taiwan, the speed of a vehicle should be less than 40km/hour, which is 11.11meters/second, in the urban area. Thus, it is assumed that the speed is less than 40km/hour for the sightseeing vehicles, including buses, sedans, tram, etc. Since the proposed scheme aims to have a vehicle to download POIs’ geo data before it enters into POIs’ coverage, it considers the two worse cases for the vehicle to download a POI. The cellular network only has the 3G/4G signal and the bandwidth is 5.45/24.04 Mbps. Then corresponding vehicle will move at least (80Mb/5.45Mbps)*11.1meters/second = 162.94 meters for the 3G case and (80Mb/24.04Mbps) * 11.1 meters/second = 36.94 meters for the 4G case when the downloading is finished. Let the probability of vehicle’s connecting to 3G and 4G be 50% and 50% respectively. A POIs' range should be at least (162.94 + 36.94)/2 = 99.79 meters. For calculation convenience, let POIs' range be 100 meters wide along the road.

(4) Since vehicles are moving along the road and the composed topology of vehicles is changed frequently considering different speeds of different vehicles, it is unnecessary caching too many POIs’ geo data for one time. Hence, each vehicle is assumed to download POIs’ geo data in the area that is in the next 1km, for which one POI’s range is ≥100 meters according to the derivation of (3) and thus it has at most 1km/100m=10 POIs.

(5) When it is the time to refresh the cached POIs’ geo data in vehicle *X*, *X* can have one of the following three choices: (i) *X* tries to find a cluster to get POIs’ geo data. (ii) *X* creates a cluster and becomes the cluster head. *X* then sends the CLUSTER_HEAD_HELLO message to invite proximate vehicles, which also need to refresh their cached POIs’ geo data, to join *X*’s cluster. Then *X* downloads POIs’ geo data using its cellular network and forwards the downloaded POIs’ geo data to its cluster members using the IEEE 802.11p network. (iii) *X* downloads POIs’ geo data to refresh its cached POIs’ geo data using its own cellular network and neither shares its downloaded POIs’ geo data with others nor gets POIs’ geo data from other vehicles. That is, since *X* cannot find any proximate vehicle or none of its proximate vehicles is in the situation of refreshing their cached POIs’ geo data.

## 4 The Proposed Scheme

This Section presents the functional scenario and then details of the proposed scheme, including when a vehicle triggers POI’s geo data downloading and when to do related actions of the clustering process.

Referring to Figure 2, it shows the geo points of triggering clustering process. Current Point ($CP_c$) denotes the current position of a vehicle. Search Point ($SP_s$) denotes the vehicle starts searching an existed cluster to join and tries to become a cluster member. Create Point ($SP_c$) denotes the vehicle creates a cluster, becomes a cluster head and invites other nearby vehicles to join. Download Point ($SP_d$) denotes the cluster head starts to download POIs’ geo data and forwards POIs’ geo data to its members.

After the cluster head starts to download and forward POIs’ geo data, (i) cluster head can disband the cluster when there is no cluster member, i.e., all of the cluster members have left; (ii) cluster member can leave the cluster when it is going to be out of the transmission range of the cluster head, e.g., cluster member and its cluster head are driving to different roads in the road intersection.

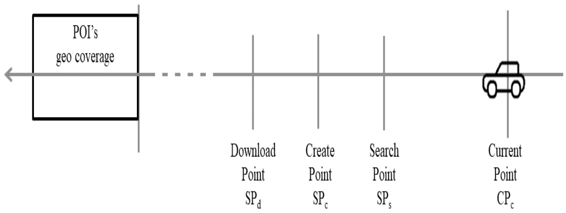


**Fig. 2.** The geo points of triggering the clustering process.

When the VSN server receives a context report from a vehicle, the VSN server calculates to see whether it is the due time or not to inform the vehicle, which does not belong to any cluster, to search an existed nearby cluster to join; however, when it cannot find a cluster to join, it itself still needs to create a cluster and then (i) invites other vehicles to join, (ii) downloads POIs’ geo data using its subscribed 3G/4G cellular network and (iii) shares the downloaded POIs’ geo data with its cluster members.

In the remaining part of this Section, details of the proposed scheme are presented.

### 4.1 The due time to start download POI’s data

One of the goals of the proposed scheme is to pre-download POI's geo data before vehicle *X* enters into the geo coverage of the ahead POIs. Since the speed of vehicle *X* and its 3G/4G cellular network’s available bandwidth vary with the time, the VSN server can use the following formula to derive the due time for downloading the ahead POIs’ geo data.

$$Sampling\ Times\ (ST) = \left\lceil \frac{\frac{D_c}{v}}{T_s} \right\rceil \quad (1)$$

$$Download\ Time\ of\ POIs'\ geo\ data\ (DT_{POI}) = \frac{\sum_{i=1}^{k} D_i}{B_{LTE}} \quad (2)$$

In formula (1), *ST* denotes how many GPS sampling times from vehicle’s current location to the borderline of the ahead POI’s coverage, where $D_c$ denotes the distance from vehicle’s current location to the borderline of the ahead POI’s geo coverage, *v* denotes vehicle’s current velocity and $T_s$ denotes GPS sampling period.

In formula (2), *k* denotes number of ahead POIs that need to download in the range of 1km, $DT_{POI}$ denotes the downloading time of the geo data of the ahead *k* POIs using 3G/4G cellular network, $D_i$ denotes the downloaded geo data volume of the ahead $POI_i$, where *i* = 1 to *k*, $B_{LTE}$ denotes available bandwidth of the cellular network.

After the VSN server calculates vehicle *X*'s current *ST*, it knows whether the driving time length from X's current location to the borderline of the ahead POIs' coverage is enough to download the ahead POIs' geo data or not: if vehicle *X*'s current $ST \times T_s \gg DT_{POI}$, it means that vehicle *X* is still far away from POIs' geo coverage; if $ST \times T_s > DT_{POI} \geq (ST-1) \times T_s$, then it means that if the downloading of the ahead POIs' geo data is started from the next GPS sampling time point, the time length is not enough to finish downloading before vehicle *X* enters into the ahead POIs' geo coverage based on vehicle *X*'s currently attached 3G/4G cellular network's bandwidth and the vehicle's speed. Thus, the VSN server should inform vehicle *X* to start downloading right away.

### 4.2 The time point to trigger the clustering procedure

The next problem is (i) which vehicle should be the cluster head and which vehicles can be cluster members and (ii) when to trigger the clustering process. With the help of the *ST* depicted in Equation (1), it can clearly define when a vehicle should start downloading POIs' geo data. However, it needs to have some time to let other vehicles to join a cluster. Let the due time for triggering the downloading of the ahead POIs' geo data be the last $SP_d^{th}$ GPS sampling point before entering into the borderline of the ahead POIs' geo coverage, i.e., the vehicle will enter into the geo coverage of the ahead POIs after $SP_d \times T_s$ seconds if the speed remains the same. Let the due time of creating a cluster be the last $SP_c^{th}$ GPS sampling point before entering into the borderline of the ahead POIs' geo coverage. Let the due time of searching a nearby cluster head be the last $SP_s^{th}$ GPS sampling time point before entering into the borderline of the ahead POIs' coverage. Then, referencing to Figure 2, $SP_s = SP_d + 2$ and $SP_c = SP_d + 1$.

### 4.3 Clustering Process

The clustering process includes (1) cluster's searching, (2) cluster's creation, (3) member join, (4) member leave, and (5) cluster disband.

#### *4.3.1 Cluster's Searching*

To make sure that a vehicle can stay in the cluster before the cluster head finishing downloading and forwarding POIs' geo data, a vehicle *M* that wants to join an existed cluster *C* needs to evaluate the connected time between itself and the cluster head *H* at first. The connected time depends on both relative position and relative speed of *M* and *H*. Let *CT (H,M)* denote the connected time between cluster head *H* and vehicle *M*. Vehicle *M* sends the JOIN_REQUEST message to cluster head *H* if its connected time with *H* is equal to or greater than the time for downloading POIs' geo data, which means that the corresponding vehicle will not leave the cluster before downloading the ahead POIs' geo data is finished. The procedure of cluster's searching is as follows:

(1) When vehicle *M* receives a notification of starting to search a neighboring cluster to join from the VSN server, vehicle *M* listens to all CLUSTER_HEAD_HELLO messages broadcasted from its proximate cluster heads before receiving a JOIN_OK message.

(2) When vehicle *M* receives a CLUSTER_HEAD_HELLO message sent from cluster head *H*, *M* checks whether *H*'s Road ID is the same as *M*'s Road ID or not; if it is, then *M* calculates the connected time between itself and cluster head *H*.

(3) If $CT\ (H,M) \geq DT_{POI}\ (H)$, then, *M* sends a JOIN_REQUEST (*M, H*) to cluster head *H*, and continues listening to CLUSTER_HEAD_HELLO messages sent from other cluster heads; if $CT\ (H,M) < DT_{POI}\ (H)$ then, *M* keeps listening to other CLUSTER_HEAD_HELLO messages sent from other cluster heads. Since it may have racing situation, i.e., *M* sends the JOIN_REQUEST ($M,H_1$) to $H_1$ earlier than *M* sends the JOIN_REQUEST ($M,H_2$) to $H_2$, but *M* receives $H_2$'s reply message earlier than $H_1$' reply message, or some cluster heads may reject requests for some reasons, *M* still needs to keep listen to CLUSTER_HEAD_HELLO messages sent from other cluster heads and reply the JOIN_REQUEST *(M,* $H_i$*)* message to the corresponding cluster head $H_i$, *i = 3 or 4 or .. m.*

(4) *M* stops listening to CLUSTER_HEAD_HELLO messages until (i) *M* receives a JOIN_OK *(M, Cluster_ID)* message or (ii) *M* reaches the $SP_c$ point. For (i), *M* sends a JOIN_CONFIRM *(M, Cluster_ID)* message to the corresponding cluster head $H_i$, *i = 3 or 4 or .. m,* from which *M* receives the first JOIN_OK (*M, Cluster_ID)* message; then *M* becomes a cluster member and waits for the forwarded POIs' geo data from the corresponding cluster. For (ii), *M* creates a cluster by itself.

#### *4.3.2 Cluster's Creation*

Cluster's creation begins when the VSN server finds that vehicle *X* reaches the $SP_c$ point and vehicle *X* does not belong to any cluster.

The procedure of cluster creation is as follows:

(1) The VSN server assigns a unique Cluster_ID for vehicle *X* and sends the CLUSTER_CREATE (Cluster_ID) message to vehicle *X.*

(2) After vehicle *X* receives the CLUSTER_CREATE (Cluster_ID) message, vehicle *X* can broadcast the CLUSTER_HEAD_HELLO message, which includes current time, Cluster_ID, X's speed, *X*'s location, estimated $DT_{POI}$ (*X*), and *X*'s Road_ID, periodically, e.g., ever one second, to invite other vehicles to join the cluster before reaching the next GPS sampling time point, i.e., the $SP_d$ time point depicted in Figure 2.

#### *4.3.3 Member Join*

Since vehicle *Y,* which is searching for a nearby cluster to join, may send the JOIN_REQUEST message to more than one cluster head, when cluster head *H* does not receive the JOIN_CONFIRM message from vehicle *Y* after *H* sending the JOIN_OK message to *Y* for a while, it means that *Y* decides to join the other cluster and does not join *H*'s cluster.

The procedure of member joining a cluster is as follows:

(1) Cluster heads $H_i$, *i=1..k*, broadcasts the CLUSTER_HEAD_HELLO message periodically, e.g., every one second, to their proximate vehicles during one GPS sampling time period.
(2) When cluster head $H_i$, *i=1..d,* receives the JOIN_REQUEST(*Y*, $H_i$) message from vehicle *Y*, it sends the JOIN_OK (*Y, Cluster_ID*) message to vehicle *Y.*
(3) If cluster head $H_m$, *m=1 or 2 or..d,* receives the JOIN_CONFIRM (*Y, Cluster_ID*) message sent from vehicle *Y*, then cluster head $H_m$, *m=1 or 2 or … m* includes vehicle *Y* as a member of its cluster and sends the UPDATE_REQUEST *(Y, Cluster_ID)* message to the VSN server.
If cluster head $H_i$, i=1… m-1, m-2, .. d, reaches its next GPS sampling time point, i.e., the $SP_d$ time point depicted in Figure 4, and still does not receive the JOIN_CONFIRM (*Y, Cluster_ID*) message sent from vehicle *Y*, it means that *Y* does not join $H_i$'s cluster, i=1… m-1, m-2 … d.

#### *4.3.4 Member Leave*

To avoid the situation that (1) cluster members and cluster head drive to different roads, e.g., cluster head or some cluster members drive to the other road in a road intersection, or (2) the connected time between cluster head and it cluster member is not enough due to the sudden speed up/down of vehicles, the VSN server periodically checks whether the topology of each cluster keeps the same or not. If any of the aforementioned situation happens, the VSN server informs cluster members to leave.

The procedure of member leave is as follows:

(1) While cluster member *M* reports its current context to the VSN server periodically, the VSN server checks whether the corresponding $CT$ with *M*'s current cluster head is enough or not and checks whether the current Road ID is still the same as that of the cluster head or not.
(2) (a) When the VSN server finds that *M* 's current Road_ID is not the same as cluster head *H*'s Road_ID, i.e., cluster member *M* or cluster head detours to the other road in a road intersection, the VSN server sends a LEAVE_INDICATION (*Cluster_ID, M*) to cluster member *M* to notify *M* to leave the cluster.
(b) When the VSN server finds that the corresponding $CT$ is not enough, i.e., cluster head or *M* changes its speed too much, it sends a LEAVE_INDICATION (*Cluster_ID, M*) to cluster member *M* to notify *M* to leave the cluster.
(3) Once cluster member *M* is informed to leave the cluster by the VSN server, *M* sends the LEAVE_CONFIRM (*Cluster_ID, M*) message, which contains the volume of received POIs' geo data to the VSN server.
(4) The VSN server sends the UPDATE_INDICATION (*M, Cluster_ID*) message to cluster head *H* to update its member list.

#### *4.3.5 Cluster Disband*

When there is no cluster member left before the cluster head finishes downloading and forwarding POIs' geo data, the cluster head can disband its cluster.

The procedure of disbanding a cluster is as follows:

(1) If all of the cluster members have left, i.e., there is no cluster member, cluster head *H* sends a DISBAND_REQUEST message to the VSN server.
(2) After the VSN server receives the DISBAND_REQUEST message, the VSN server removes the corresponding cluster's information and then the VSN server sends back the DISBAND_CONFIRM message to cluster head *H*.

### 4.4 The procedures of clustering process executed in the VSN server

The VSN server collects vehicles' reported context and executes the cluster creation/join/leave/disband procedure.

When the VSN server receives (i) a context report from a vehicle that does not belong to any cluster, (ii) a context report from a vehicle that is a cluster member and (iii) a message from a vehicle that is cluster head, it needs to do related actions accordingly.

For (i), the VSN server calculates whether it is the due time for the vehicle *X* searching for a cluster to join. For (ii), when the vehicle *X* is a cluster member, the VSN server checks whether (a) the connected time length between *X* and *X*'s cluster head is enough to receive the POIs' geo data from *X*'s cluster head or (b) *X* and *X*'s cluster head are still in the same road or not. If the VSN server finds that the connected time length is not enough or cluster member *X* and *X*'s cluster head are not in the same road, the VSN server informs cluster member *X* to leave and also informs *X*'s cluster head to update its cluster member list. For (iii), when the vehicle *X* is a cluster head, the VSN server updates its recorded *X's* member list according to the received *X*'s context. If there is no more cluster member, the cluster head informs the VSN server to disband the cluster.

## 5 Performance Analysis

This Section presents performance analysis of the proposed scheme. The simulation was performed using network simulator NS-3 (Network Simulation 3.25) [9]. The traffic trace mobility model was generated using SUMO (Simulation of Urban MObility) [10]. Multicast is used to forward the downloaded POIs' geo data to cluster members.

### 5.1 The Simulation Environment

The performance analysis is based on the vehicles moving through the urban touring scenario. Table 1 shows related simulation parameters.

**Table 1.** Simulation Configuration Parameters

| Parameter | Value |
|---|---|
| Velocity | 0.0 m/s - 11.1 m/s |
| Transmission range | 300m |
| Duration of simulation | 1870s |
| GPS sampling period | 10s |

Referring to Table 2, vehicles enter into the road with different periods, which denote how many seconds a vehicle is generated, based on different vehicle's density. Let $V_n$ denote the total number of vehicles on the urban scenario, e.g., in the situation of high vehicle's density, a new vehicle is generated every 5 seconds and the total number of vehicles on the urban scenario is 40. Each vehicle's available network bandwidth becomes less when vehicle's density increases.

**Table 2.** Vehicle density parameter

| Vehicle Density | Period (Seconds) | $V_n$ |
|---|---|---|
| Low | 10 | 20 |
| Middle | 6.66 | 30 |
| High | 5 | 40 |

Our proposed scheme is compared with the following three methods. Method A is to have each vehicle to download POIs' geo data using its own cellular network individually. That is, when the vehicle reaches the $SP_d$ point, each vehicle downloads its needed POIs' geo data individually. Method A can be divided into the following two sub-methods. In sub-method A1, the downloading point is the last 1st GPS sampling point before entering into the geo coverage of the ahead POIs. In sub-method A2, the downloading point is the last $SP_d^{th}$ GPS sampling point before entering into the geo coverage of the ahead POIs, i.e., the vehicle will enter into the geo coverage of the ahead POIs after $SP_d \times T_s$ seconds if the speed remains the same. The $SP_d$ point is determined using the equations and the principles depicted in Section 4.1 and Section 4.2. Method B is to have each vehicle starts the clustering process from cluster search when it reaches $SP_s$ point. Method B can be divided into the following two sub-methods. In sub-method B1, the $SP_s/SP_c/SP_d$ points are the last 3rd/2nd/1st GPS sampling point before entering into the borderline of the ahead POIs' geo coverage. Sub-method B2 is our proposed scheme. For method B, each vehicle starts the clustering process from cluster search when it reaches $SP_s$ point. When the vehicle reaches the $SP_c$ point and it still cannot find a nearby cluster head, it creates a new cluster and then invites other vehicles to join until it reaches $SP_d$ point. When the vehicle reaches the $SP_d$ point, it starts to download POIs' geo data. The $SP_s/SP_c/SP_d$ point is determined using the equations and the principles depicted in Section 4.1 and Section 4.2.

### 5.2 Experiment Results

The metrics that are used to evaluate the performance of each method are as follows: (1) the downloaded data volume, (2) the available distance, (3) sharing portion and (4) successful ratio of complete downloading. To get the general trend of the above (1)~(4) metrics in each method for different situations of vehicle's density, each method is executed for 10 times based on the same touring scenario. Then the sum of these 10 times for each metrics is divided by 10 respectively.

#### *5.2.1 Downloaded data volume*

Figure 3 depicts the average downloaded data volume using the 3G/4G cellular network of each vehicle in the situations of the low, middle, high vehicle's density. The average downloaded data volume using the 3G/4G cellular network for each method is calculated as follows. The sum of the downloaded data volume using the 3G/4G cellular network is divided by the total number of vehicles.

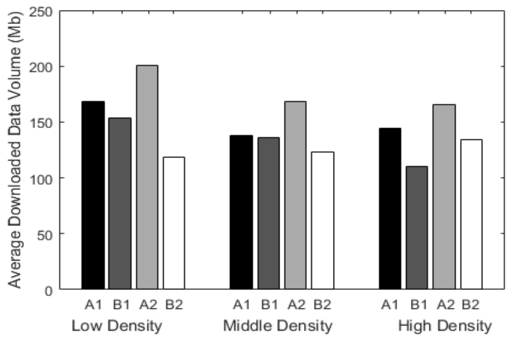


**Fig. 3.** Average Downloaded Data volume in the situations of different vehicle's densities.

Sub-method A1 and sub-method A2 have higher average downloaded data volume than the other two methods. It is because each vehicle in sub-method A1 and sub-method A2 needs to download the needed POIs' geo data individually. The difference between sub-method B1's average downloaded data volume and sub-method B2's average downloaded data volume (i) is similar in the situations of low and middle vehicle's density, but (ii) is more different in the situations of high vehicle's density. In the situation of low vehicle's density, sub-method B1 may not have the chance to organize clusters to share downloaded POIs' geo data because there are not so many proximate vehicles comparing with that in other two situations of vehicle's density. Thus, sub-method B1 may need to download POIs' geo data individually more often and have higher downloaded data volume. In the situation of middle vehicle's density, both sub-methods B1 and B2 have more cluster members and thus sub-method B1 and B2 have similar average downloaded data volume. In the situation of high vehicle's density, the network becomes much congested such that vehicles in sub-method B1 cannot have enough time to download the complete POIs' geo data and thus sub-method B1's average downloaded data volume < sub-method B2's average downloaded data volume.

### *5.2.2 Available Distance*

Available distance denotes the distance to the exit borderline of the pre-downloaded area when a vehicle finishes downloading/receiving the complete POIs' geo data. If a vehicle does not finish downloading or receiving the complete POIs' geo data when it reaches the exit borderline of the pre-downloaded area, then the distance value is 0. The higher distance means that vehicles have more sufficient time to watch downloaded POIs' geo data.

Figure 4 depicts the available distance to the exit borderline of the pre-downloaded area when a vehicle finishes downloading/receiving the complete POIs' geo data for the situations of different vehicle's densities.

In the situations of low, middle and high vehicle's density, sub-method A1 has smaller available distance than sub-method A2, sub-method B1 has smaller available distance than sub-method B2, sub-methods A1 and B1 have smaller available distance than sub-methods A2 and B2. The reason is the same, sub-methods A1 and B1 start to download the POIs' geo data later than sub-methods A2 and B2. Thus sub-methods A1 and B1 have smaller available distance than sub-methods A2 and B2 have, respectively.

Generally speaking, the available distance of each method becomes smaller when the vehicle's density increases. The reason is that higher density makes network more congested and thus vehicles need more time to or even cannot finish downloading and forwarding POIs' geo data. Thus, it would be more near the exit borderline of the pre-downloaded area when the downloading and forwarding of POIs' geo data are finished.

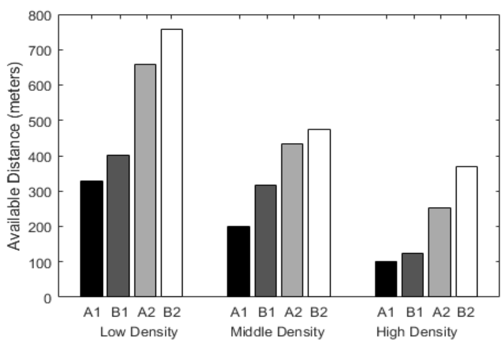


**Fig. 4.** Available Distance in the situations of different vehicle's densities.

### *5.2.3 Sharing portion*

The sharing portion is derived from dividing the total volume of received POIs' geo data using IEEE 802.11p network by the total volume of downloaded POIs' geo data, which include the ones that were downloaded by the vehicles individually and the ones that were downloaded by cluster heads, using cellular network. Figure 5 depicts the sharing portion of different vehicle's densities using different methods. The lower value of sharing portion indicates that lower POIs' geo data were shared among vehicles through the IEEE 802.11p network.

Sub-methods B1 and B2 have higher sharing portion in the situation of middle vehicle's density than that in the situation of low and high vehicle's density. The reason is that in the situation of high vehicle's density, the network is too congested such that sub-methods B1 and B2 may not have enough time to finish downloading in time and the data volumes that were forwarded from IEEE 802.11p decrease and thus they have lower sharing portion; in the situation of middle vehicle's density, there are more vehicles in the approximate location than that in the situation of low vehicle's density such that each cluster in sub-methods B1 and B2 can have more cluster members in the situation of middle vehicle's density and thus have higher sharing portion.

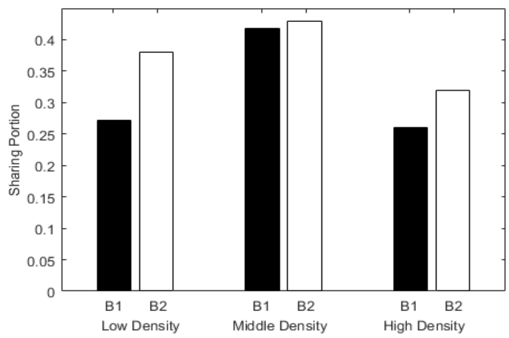


**Fig. 5.** Sharing Portion in the situations of different vehicle's densities.

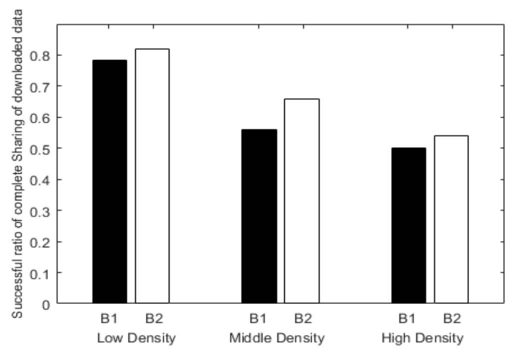


**Fig. 6.** Successful ratio of complete sharing of downloaded POIs' geo data in the situations of different vehicle's densities.

### *5.2.4 The successful ratio of complete sharing of downloaded POIs' geo data*

If all of the cluster members in a cluster have received all of the POIs' geo data successfully before reaching the exit borderline of the pre-downloaded area, then the cluster's sharing of downloaded POIs' geo data is successful; otherwise, it is failed. Successful ratio of the complete sharing of the downloaded POIs' geo data is derived from dividing the number of successful clusters by the total number of clusters. Figure 6 depicts the successful ratio in different vehicle's densities.

Generally speaking, the successful ratio of the complete sharing of the downloaded POIs' geo data is decreased when the vehicle's density is increased for all methods. That is, the higher vehicle's density it is, the lower successful ratio of complete sharing of the downloaded POIs' geo data it has for each method. The reason is that the network becomes more congested when more vehicles are in the road. The cluster heads may not download and forward the complete POIs' geo data in time and thus it is possible that the complete forwarding to their cluster members is failed more often.

## 6 Conclusion

In this paper, a clustering scheme was proposed to enable a group of vehicles, which belong to the same VSN, moving on the same road with the same direction to form a cluster. Then one vehicle is chosen as the cluster head to download POIs' geo data through its cellular network and shared the downloaded POIs' geo data to other vehicles in the same cluster through IEEE 802.11p network. The results of performance analysis have shown that the proposed scheme can have vehicles in a VSN to form a cluster, have higher available distance and higher successful ratio of the complete sharing of downloaded POIs' geo data than other methods in the three situations of vehicle's density. The reason is that the proposed scheme does not need to download many duplicate POIs' geo data using cellular network. Additionally, the proposed sheme has higher downloaded data volume in the situation of high vehicle's density. The reason is that the network becomes much congested such that the proposed scheme tries to start clustering process eariler and thus the proposed scheme has enough time to download POIs' geo data and higher successful ratio, which results in more POIs' geo data can be downloaded. For the future work, it can take the credit scheme into consideration such that it can be more fair to share the downloading duty from the charged 4G cellular network

## Acknowledgment

This research was supported by the Ministry Of Science and Technology (MOST), Taiwan (R.O.C.) under the grant number MOST 106-2221-E-006-029.

## References

1. R. S. Bali, N. Kumar, and J. J. P. C. Rodrigues, "An intelligent clustering algorithm for VANETs," *Proceedings of 2014 International Conference on Connected Vehicles and Expo (ICCVE)*, pp. 974-979, 3-7 Nov. 2014 (2014).
2. C. Cooper, D. Franklin, M. Ros, F. Safaei, and M. Abolhasan, "A Comparative Survey of VANET Clustering Techniques," *IEEE Communications Surveys & Tutorials*, VOL. **19**, NO. 1, pp. 657-681, (2017).
3. S. M. AlMheiri and H. S. AlQamzi, "MANETs and VANETs clustering algorithms: a survey," *Proceedings of the 8th IEEE GCC Conference and Exhibition (GCCCE)*, pp. 1-6, 1-4 Feb. 2015 (2015).
4. F. Mezghani, R. Dhaou, M. Nogueira, and A. L. Beylot, "Content dissemination in vehicular social networks: taxonomy and user satisfaction," *IEEE Communications Magazine*, VOL. **52**, NO. 12, pp. 34-40, (2014).
5. Z. Ning, F. Xia, N. Ullah, X. Kong, and X. Hu, "Vehicular social networks: enabling smart mobility," *IEEE Communications Magazine*, VOL. **55**, NO. 5, pp. 16-55, (2017).
6. Q. Yang and H. Wang, "Toward trustworthy vehicular social networks," *IEEE Communications Magazine*, VOL. **53**, NO. 8, pp. 42-47, (2015).
7. M. Ren, L. Khoukhi, H. Labiod, J. Zhang, and V. Veque, "A new mobility-based clustering algorithm for vehicular ad hoc networks (VANETs)," *Proceedings of IEEE/IFIP Network Operations and Management Symposium (NOMS 2016)*, pp. 1203-1208, 25-29 April (2016).
8. R. Chai, X. Ge, X. Hu, and B. Yang, "Work in progress paper: Utility based clustering algorithm for VANET," *Proceedings of the 9th International Conference on Communications and Networking in China*, pp. 187-190, 14-16 Aug. (2014).
9. L. Katsikas, K. Chatzikokolakis, and N. Alonistioti, "Implementing clustering for vehicular ad-hoc networks in ns-3," *Proceedings of the 2015 Workshop on ns-3*, pp. 25-31, (2015).
10. D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO-Simulation of Urban MObility," *International Journal On Advances in Systems and Measurements*, VOL. **5**, NO. 3&4, pp. 128-138, (2012).