Computer Communications 152 (2020) 109–118

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Machine learning based code dissemination by selection of reliability mobile vehicles in 5G networks

Ting Li [a], Ming Zhao [a,*], Kelvin Kian Loong Wong [a,b,**]

[a] School of Computer Science and Engineering, Central South University, ChangSha 410083, China
[b] School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, SA 5005, Australia

ARTICLE INFO

Keywords:
Machine learning
Code disseminations
Safety degree
Coverage ratio
Reliability
5G networks

ABSTRACT

Recently, the evolving of 5G networks is foreseen as a major driver of future mobile vehicular social networks (VSNs), which can provide a novel method of code disseminations. Based on this concept, vehicles can be used as code disseminators. That is, infrastructures of a smart city can be upgraded by receiving updated program codes that are disseminated by vehicles in the VSNs. Specifically, vehicles in the 5G network are hard to be managed. Under this domain, safety of program codes is a key challenge. Meanwhile, improving coverage of program codes is also challenging. However, arranging plenty of vehicles as code disseminators will incur large costs of the ground control station (GCS). Therefore, by utilizing machine learning methods, this paper proposes a "Machine Learning based Code Dissemination by Selecting Reliability Mobile Vehicles in 5G Networks" (MLCD) scheme to choose vehicles with higher reliable degree and coverage ratio as code disseminators to deliver code with lower costs. Firstly, reliable degrees of vehicles are calculated and selected to improve safety degree of code disseminations. Secondly, vehicles with higher coverage ratio are preferred to promise code coverage. Thirdly, machine learning methods are utilized to select vehicles with both higher coverage ratios and reliable degrees as code disseminators with limited costs. Compared to random-selection and coverage-only scheme respectively, the MLCD scheme can improve safety degree of code dissemination process by 83.6% and 18.86% in 5G networks, and can improve coverage ratio of updated information by 23.16%. Comprehensive performances of the proposed scheme can be improved by 80.56% and 17.25% respectively. Future works focus on improving code security in 5G networks by more advanced and suitable machine learning methods.

1. Introduction

Recently, the increasing demands and requirements of vehicular social networks (VSNs) poses challenges for the 5G networks [1,2]. The 5G mobile system is complex and is hard to be managed. Therefore, a hot research aspect where a VSNs systems' reliability to enhance and promise a safety environment via the machine learning schemes (ML) becomes more prudent and significant. In the 5G mobile networks, because of the characteristics of large-number and mobility, the vehicles have been widely utilized in varies of situations in smart city. Large number of vehicles can act as mobile data mules to deliver data or to collect data from plenty of sensors that spread across virtually every smart city streets (garbage cans, streetlamps, bus stations, etc., with low costs [3]. Moreover, it is convenient for those mobile vehicle to transmit codes to hundreds of thousands of sensor derives in a smart city, especially in dynamic city networks, where there are plenty of temporarily deployed sensing devices that are not connected to the Internet. [4,5]. Researches above have developed and provided more convenient services to the social life.

The sensing devices, which are applied to public infrastructures such as streetlamps, garbage cans, advertising panels, etc., are wildly deployed in the 5G networks. Large amount of sensing devices can be upgraded through updating their program code by using software-defined technology and can radiate a new life. Therefore, updating program codes of those devices is an urgent but challenging issue. With expansions of mobile and sensing technologies, there is a tendency that vehicles are utilized as code disseminators, which can update program codes of sensing devices in the 5G mobile networks. Over code dissemination scenarios, vehicles have gained attentions because of the mobility feature. However, many issues raised in this process, such as the coverage issue [6,7], the safety issue [8,9], the reliability issue [10–12] and the cost saving issue [13–15], etc. Among those issues, both safety and coverage of vehicular disseminators are of great importance in the 5G mobile networks. Firstly, the safety issue occurs because of participations of malicious vehicles; malicious vehicles will lead to the phenomena such as code tampering, code losses, etc [16]. Secondly, coverage issue influences the performances of program code

* Corresponding author.
** Corresponding author at: School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, SA 5005, Australia.
E-mail addresses: meanzhao@mail.csu.edu.cn (M. Zhao), kelvin.wong@adelaide.edu.au (K.K.L. Wong).

https://doi.org/10.1016/j.comcom.2020.01.034
Received 1 November 2019; Received in revised form 15 December 2019; Accepted 17 January 2020
Available online 23 January 2020
0140-3664/© 2020 Published by Elsevier B.V.

disseminations, because the program codes of infrastructures in the uncovered regions cannot be updated. That will have direct influences on the results of the sensing devices. Therefore, selecting suitable vehicles that can improve both safety factor and coverage factor is the main target that has been investigated in this paper.

In practice, ML schemes are the processes that can automatically learn from previous experiences and studies, and reach unpredicted results with high precisions [17,18]. With popularizations of ML schemes, large number of modern fields [19] such as automotive and networks are exploiting machine learning to improve and enhance safety environment [20,21], which improves result precisions. Therefore, in this paper, the ML schemes are deployed in vehicular selection process to optimize both safety degree and coverage ratio of code disseminators in the 5G mobile networks.

In this paper, we propose an efficient code dissemination scheme named "Machine Learning based Code Dissemination by Selecting Reliability Mobile Vehicles in 5G Networks" (MLCD scheme), which can improve both safety ratio and coverage ratio via utilizing ML schemes in the 5G mobile networks. The researches of this paper mainly target on two aspects: (1) The methods to optimize safety factor in the dissemination processes of updated program codes and (2) The methods to improve the coverage factor of the vehicular disseminators. Malicious vehicles as code disseminators will affect code security. Coverage represents the area that the update code can cover. Larger coverage ratio means that more infrastructures can receive updated program codes. Infrastructures with sensing devices are distributed uniformly in a smart city, therefore, some in the suburb regions may cannot receive updated codes if vehicular disseminators do not pass those regions, which affects sensing results.

Therefore, the main contributions of the MLCD scheme are listed as follows:

(1) Based on historical trajectory datasets, the MLCD scheme selects vehicles as code disseminators by analyzing both trust ratio and coverage ratio mainly via the Genetic Algorithms (GA) in the machine learning schemes. In previous schemes, vehicular disseminators are selected randomly, which cannot promise security degree of program codes, and coverage ratio may not be satisfied. Thus, in this paper, with historical geographical data, GA is utilized in the MLCD scheme, which targets to select reliability-based vehicular disseminators to obtain larger coverage regions with higher accuracy rate.

(2) The MLCD scheme also evaluates costs of vehicular disseminators. In previous schemes, randomly selections of vehicular disseminators may generate unnecessary costs. Therefore, the MLCD scheme also take costs factor into considerations to optimize performances of code disseminations within limited costs.

(3) Performances of the MLCD scheme are evaluated by both theoretical analysis and experimental studies. Compared with previous approaches, the MLCD scheme can improve coverage ratio of code disseminators by 23.16%, and can improve data safety degree by 83.6% and 18.86% with limited costs. In general, performances of the MLCD scheme can be optimized by 80.65% and 17.25% approximately.

The remaining of this paper are organized as follows: in Section 2, related works are reviewed and presented. The system model and problem statements are introduced in Section 3. Detail designs of the MLCD scheme are presented in Section 4; experimental analysis and comparisons are shown in Section 5. Finally, we conclude this paper in Section 6.

2. Related works

In this section, the related works are presented. There are a number of manuscripts that research vehicular selection algorithm in VSNs [15,22,23] and usage of machine learning methods in VSNs [24,25]. There common goals are to improve coverage regions of selected code disseminators or to improve safety degree in code dissemination process.

In general, Siegel et al. [26] summarized current state of technologies in VSNs and introduced challenges and opportunities in the future. With connectivity improvements, vehicles could be used as development platform. Disseminating data to infrastructures is one usage of future VSNs. Over all challenges, privacy and safety have attracted plenty of attentions.

Braga et al. [27] proposed a scheme that used parked vehicles as alternatives to deployments of fixed units/applications in a smart city. Their paper introduced a method to park vehicles to form efficient vehicular networks, which could enhance the coverage ratio of information collections. Simulation results reflected that the parked vehicles could serve as alternatives to fixed units, which inspired us. However, the researchers did not consider the information security issue.

The research by Zhou et al. [28] combined both the social and physical layer information for realizing rapid dissemination in the VSNs. This paper proposed a price-rising-based algorithm that could solve problem in physical layer. Experimental results proved efficiency. However, this method did not consider security issue. Vehicles might be malicious, which have bad influences.

Then, Liu et al. [29] proposed a reliable and stable communication scheme by using clustering and probabilistic broadcasting. This scheme was based on multi-vehicle communications. With this method, a vehicle could disseminate data to other vehicles within connection time. Moreover, this scheme could also improve coverage ratio. However, during vehicle-to-vehicle communications, this scheme could not figure out malicious vehicles, which might lead to data insecurity.

Kaewpuang et al. [30] introduced a method that could improve efficiency and reduce costs by sharing problem in the cooperative logistics environment. Considering cost factor, vehicles could choose whether to cooperate with others or not. With routing method and cooperation among vehicles, simulation indicated that the coverage could be improved and costs could be reduced. In this scheme, vehicles in a vehicle pool were assumed to be reliable and trusted. Therefore, it could not figure out whether a vehicle is malicious or not.

Recently, there is a tendency that machine learning methods are widely used in the VSNs. Radhakrishna et al. [31] proposed an approach for temporal association pattern prevalence values and a temporal fuzzy similarity measure. With the data mining algorithm proposed in this paper, data could be analyzed with lower complexity.

The research by Anjomshoaa et al. [32] utilized vehicles to monitor the urban environment efficiently. It discussed and researched drive-by approach and coverage performances of vehicles with scheduled trajectories. Moreover, sensor devices were applied to garbage trucks and collected drive-by data. Based on real datasets, it presented potential of using vehicles to optimize data collections, which is a theoretical basis of the MLCD scheme proposed in this paper. All these works did not consider both security degree and coverage ratio.

In our paper, we investigate the vehicular selection problem in the VSNs based by using machine learning algorithm that fully takes into account the influences of coverage and safety of vehicular disseminators. Based on previous studies, the VSNs have brought a novel method in disseminating program codes (a kind of data).

3. System model and problem statements

In this section, system model of the MLCD scheme is described and problem statements are introduced.

3.1. System model

Infrastructures (such as streetlamps and garbage cans) with sensing devices are randomly distributed in a smart city, which are utilized to sense surrounding information. Infrastructures can be upgraded by updating their program codes and have new functions. The ground control station (GCS) is in the center of a smart city, which is responsible for selecting reliability-based and coverage-based vehicles as

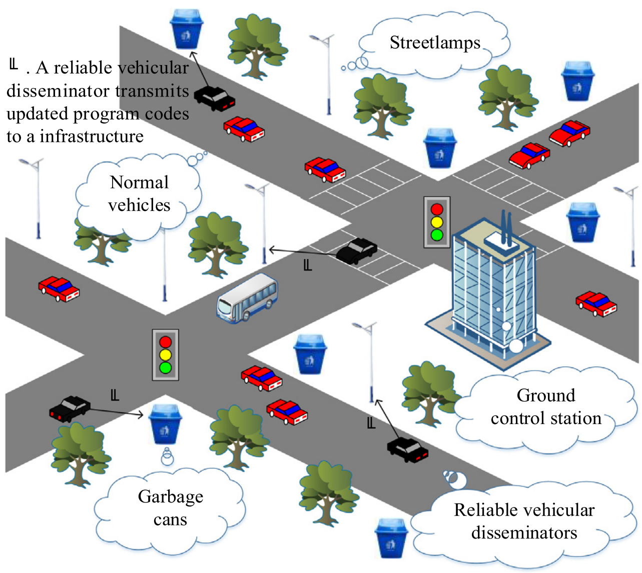

Fig. 1. Dissemination processes of updated program codes.

Table 1
Main parameters and notations.

Parameter	Descriptions
$\mathcal{C}_i$	Costs of a vehicle v_i
$\mathcal{V}_{\mathfrak{s}}$	Set of vehicular code disseminators
$\mathcal{V}$	Vehicular set in a city
ξ	Vehicular costs of an hour
$\mathcal{C}$	Costs of set $\mathcal{V}_{\mathfrak{s}}$
$\mathcal{S}_i$	The coverage ratio of v_i
ℓ	Number of landmarks
r_i	Reliable degree of vehicle v_i
ϕ_i	Comprehensive evaluation of vehicle v_i
$G_{\hbar}$	A gene set of vehicular code disseminators
n	The number of vehicles
$D(v_i)$	Total days in the dataset of vehicle v_i
α	Influence factor of coverage ratio
$\mathcal{F}_{G_{\hbar}}$	Value of gene set $G_{\hbar}$

disseminators, as well as transmitting updated program codes to the selected vehicular disseminators by using unmanned aerial vehicles (UAVs).

After the GCS decided a set of vehicular code disseminators, it needs to pay for them. The GCS wants pay less and reach more coverage ratio. Then, vehicular disseminators will transmit program codes to infrastructures. In this way, program codes of infrastructures can be updated. In general, Fig. 1 shows the dissemination processes of updated program codes in the VSNs.

Assume that the size of a set of updated program codes is small, thus, a reliable vehicular disseminator can transmit it to an infrastructure at once, instead of multiple transmissions. In the VSNs, program codes are transmitted by opportunistic communications of reliable vehicles. Therefore, some infrastructures in suburban regions cannot be covered by reliable vehicles, and they cannot update their program codes. According to analysis, infrastructures in downtown regions can easily receive updated program codes that are disseminated by reliable vehicles.

In order to ease understanding of our concept, main parameters and notations are shown in Table 1.

3.2. Problem statements

In the VSNs, malicious code disseminators influence the safety degree of code transmission processes. They can tamper or even capture program codes, which has a bad impact on dissemination processes of program codes. Reliable vehicles as code disseminators can improve security degree of program codes, thus can enhance safety dissemination environment. Therefore, the way to select reliability-based vehicles as code disseminators is a significant problem to be solved in the VSNs.

Moreover, due to uneven distributions of infrastructures with sensing devices, some infrastructures in the suburban regions cannot reach updated program codes that are transmitted by vehicular disseminators, which will affect code updating of those infrastructures. Therefore, based on reliability factor, it is important to select vehicles with high coverage ratio as code disseminators to guarantee code updating of infrastructures in suburban regions in the VSNs.

Therefore, the main idea of this paper is to design a new effective MLCD scheme to improve coverage ratio of vehicular disseminators, as well as improve safety degree in code dissemination processes. The goal of the MLCD scheme can be categorized in following aspects:

(1) Improve safety degree of code disseminations

Selection of reliable vehicular disseminators is an effective method in improving safety degree of updated program codes, which can also be regarded as improving safety degree of code dissemination processes. Safety degree improvements can be evaluated by number of reliable vehicular disseminators, which is calculated by Eq. (1). The more reliable vehicular disseminators there are, the higher the safety degree of code dissemination processes will be.

$$\mathfrak{S} = \frac{\mathbb{N}_{OCRV-MLMs} - \mathbb{N}_{pre}}{\mathbb{N}_{OCRV-MLMs}} \tag{1}$$

where $\mathbb{N}_{OCRV-MLMs}$ represents the number of reliable vehicular disseminators based on the MLCD scheme, and $\mathbb{N}_{pre}$ represents the number of reliable vehicular disseminators based on other schemes. Our target is to improve the safety degree in code dissemination process, which is to improve value of $\mathfrak{S}$ in Eq. (1).

(2) Improve coverage ratio of program codes

In the VSNs, distributions of infrastructures are uneven. According to analysis, compared to infrastructures in the suburban regions, infrastructures in the downtown regions are more likely to receive updated program codes. And some infrastructures in the suburban regions may cannot receive updated codes, which affects dissemination performances. Therefore, it is significant to improve coverage ratio of updated program codes (also known as coverage ratio of vehicular disseminators), especially in the suburban regions. Compare to other schemes, improvements of coverage ratio can be derived by Eq. (2).

$$\mathfrak{C} = \frac{Cov_{OCRV-MLMs} - Cov_{pre}}{Cov_{OCRV-MLMs}} \tag{2}$$

In this equation, $Cov_{OCRV-MLMs}$ is the coverage ratio of program codes based on the MLCD scheme and Cov_{pre} is coverage ratio of program codes based on other schemes. Our target is to improve coverage ratio of updated program codes, which is to improve the value of $\mathfrak{C}$ in Eq. (2).

(3) Comprehensive evaluations

Based on above evaluations (both safety degree and coverage ratio), $\mathfrak{E}$ is defined to evaluate performances of the MLCD scheme comprehensively, defined by Eq. (3).

$$\mathfrak{E} = \frac{2 \times \mathfrak{S} \times \mathfrak{C}}{\mathfrak{S} + \mathfrak{C}} \tag{3}$$

With this equation, performance of the MLCD scheme can be evaluated. Compared with previous schemes, value of $\mathfrak{E}$ shows advancements of the proposed scheme.

4. Design of the MLCD scheme

In this section, detail designs of the MLCD scheme are introduced. Targets of the MLCD scheme are to build a safety-based environment of code disseminations, as well as improve coverage ratio of updated program codes within limited vehicular costs.

The MLCD scheme consists three main steps: firstly, function of vehicular costs is built to calculate the costs of gene sets; secondly, both coverage ratios and reliable degrees of vehicular code disseminators are derived to calculate the value of gene sets; finally, genetic algorithm is utilized to select reliability-based and coverage-based vehicles as vehicular code disseminators in the VSNs.

4.1. Calculation methods of vehicular costs

In this subsection, calculation methods of costs for vehicular code disseminators are determined. In the VSNs, set of vehicles is defined as $\mathcal{V} = (v_1, v_2, \ldots\ldots, v_{n-1}, v_n)$, where n is number of vehicles. Set of vehicular code disseminators is defined as $\mathcal{V}_{\delta} = (v_1^{\delta}, v_2^{\delta}, \ldots\ldots, v_{m-1}^{\delta}, v_m^{\delta})$, where m represents number of vehicular code disseminators and $m < n$. In the VSNs, the GCS need to pay them. The rewards of vehicles are costs of GCS. In the MLCD scheme, we define that costs of a vehicle are related to the time it can serve. Thus, costs of a vehicle v_i can be derived by Eq. (4).

$$\mathfrak{C}_i = \xi \times (L_i - F_i), 1 \le i \le n \tag{4}$$

Here, L_i represents the termination time of v_i services and F_i is the start time of v_i services. ξ is the costs of an hour, which is a constant value. $\mathfrak{C}_i$ is the costs of v_i, which is used in Section 4.3. Thus, costs of set $\mathcal{V}_{\delta}$ can be obtained, shown in Eq. (5).

$$\mathfrak{C} = \sum_{j \in [1,m]} \mathfrak{C}_j = \sum_{j \in [1,m]} \xi \times (L_j - F_j) \tag{5}$$

After determining vehicular code disseminators, costs calculation of set $\mathcal{V}_{\delta}$ can be derived, shown in Eq. (5).

In general, calculations of vehicular costs are shown in Algorithm 1.

Algorithm 1: Designs of vehicular costs

Input: $\mathcal{V}, \xi, i, L_i, F_i, c[0 \ldots n]$
Output: Costs of each vehicle in the set $\mathcal{V}$
Initial $c[0 \ldots n] = \{0\}$;
Initial all the variables;
For v_i in the vehicular set $\mathcal{V}$ and $0 \le i \le n$;
  Get the initial time F_{v_i} of v_i;
  Get the ending time L_{v_i} of v_i;
  $c[i] = \xi \times (L_{v_i} - F_{v_i})$; // costs of vehicle v_i
  $i++$;
End for
Output $c[0 \ldots n]$
End Algorithm 1

4.2. Calculation methods of vehicular coverage ratio and reliable degree

In this subsection, calculation methods of both coverage ratio and reliable degree of each vehicle are introduced. Firstly, calculation of vehicular coverage ratio is described; secondly, calculation of vehicular reliability degree is introduced.

Obviously, the GCS needs vehicles with higher coverage ratio to disseminate updated program codes. The higher, the better. There, several GPS points are randomly selected, which are utilized to evaluate the coverage number of each vehicle. Those GPS points are called "landmarks". The vehicle that can cover more landmarks has a greater coverage ratio. The number of landmarks is defined as ℓ. In the MLCD scheme, coverage ratio of a vehicle v_i is evaluated according to the number of landmarks it covers, which is shown in Eq. (6).

$$\mathcal{S}_i = \frac{N(v_i)}{\ell}, 1 \le i \le n \text{ and } \mathcal{S}_i \in [0, 1] \tag{6}$$

whereby, $N(v_i)$ represents the number of landmarks that vehicle v_i covers. $\mathcal{S}_i$ is the coverage ratio of v_i. With Eq. (6), coverage ratio of each vehicle can be obtained in the VSNs, which is utilized in following subsection. Obviously, the higher the value of $\mathcal{S}_i$ is, the better. Generally speaking, the vehicles with higher costs have long traveling distances. Therefore, they can reach more regions, which have a greater coverage ratio.

To reduce participants of malicious vehicles, reliable degrees of vehicles are evaluated by analyzing historical trajectory datasets. According to analysis, compared to vehicles with uncertain trajectories, vehicles with regular trajectories are more reliable. Selections of reliable code disseminators can improve safety degree of code dissemination processes. One feature of regular vehicular trajectories of a vehicle v_i is whether it has a regular parking space or not. By analyzing historical datasets, when vehicle v_i is driving, GPS points of it always change. And when v_i parks in a parking space, the GPS points of it are static. That is, based on time interval of GPS collection is the same, whether a vehicle v_i has a regular parking space or not can be derived by finding the maximum number of GPS points, which is defined by Max_{v_i}. The coordinate of Max_{v_i} is defined by $(x_{v_i}^{max}, y_{v_i}^{max})$.

Therefore, reliable degree of v_i can be obtained by calculating number of times that vehicle v_i parked in the same parking lot, which is the spaces around $(x_{v_i}^{max}, y_{v_i}^{max})$. Therefore, in the MLCD scheme, reliable degree of a vehicle v_i can be calculated, shown in Eq. (7).

$$r_i = \frac{N(I_{Max_{v_i}})}{D(v_i)}, r_i \in (0, 1] \tag{7}$$

Here, $I_{Max_{v_i}}$ represents a time interval that vehicle v_i stops in a parking space Max_{v_i}, for example, from 11:00 pm to 6: 00 am. And $N(I_{Max_{v_i}})$ is number of time interval. $D(v_i)$ represents day number in the historical datasets of v_i. With Eq. (7), reliable degree of vehicles in set $\mathcal{V}$ can be derived.

Both calculations of reliable degree and coverage ratio are presented in Algorithm 2.

Algorithm 2: Designs of both reliable degree and coverage ratio

Input: $\mathcal{V}, r[0 \ldots n], \mathcal{S}[0 \ldots n], \ell, i, k, p, \vartheta, z$
Output: Reliable degree and coverage ratio of vehicles in the set $\mathcal{V}$
Initial $r[0 \ldots n] = \{0\}$;
Initial $\mathcal{S}[0 \ldots n] = \{0\}$;
Initial all the variables
For v_i in the vehicular set $\mathcal{V}$ and $0 \le i \le n$;
  Search the largest number of GPS points
  Record $(x_{v_i}^{max}, y_{v_i}^{max})$
  Record the day number $D(v_i)$ of v_i
  $r_i = \frac{N(I_{Max_{v_i}})}{D(v_i)}$ // calculate reliable degree of v_i
  Record r_i in $r[n]$ //$r[n]$ is reliable degree of $\mathcal{V}$
End for
For v_i in the vehicular set $\mathcal{V}$ and $0 \le i \le n$;
  Initial $z = 0$;
  For all the coordinates $(x_{v_i}^k, y_{v_i}^k)$ in vehicle v_i
    For all the coordinates (x_ℓ^p, y_ℓ^p) in set ℓ
      If distance between $(x_{v_i}^k, y_{v_i}^k)$ and (x_ℓ^p, y_ℓ^p) is less than threshold value ϑ
        $z = z + 1$;
      End if
    End for
  End for
  $\mathcal{S}_i = \frac{z}{\ell}$ //calculate coverage ratio of vehicle v_i
  Record $\mathcal{S}_i$ in $\mathcal{S}[n]$ // $\mathcal{S}[n]$ is coverage ratio of $\mathcal{V}$
End for
Output $r[0 \ldots n]$;
Output $\mathcal{S}[0 \ldots n]$;
End Algorithm 2

4.3. Selection methods of both reliability-based and coverage-based code disseminators

Based on our above calculation methods, our targets are to find vehicles with higher reliability and larger coverage ratio to be code disseminators in the VSNs, within limited costs. In the MLCD scheme, genetic algorithm is used to select reliability-based and coverage-based vehicles to disseminate updated program codes to infrastructures, which is introduced as follows.

There, selections of vehicular code disseminators are a NP complete problem of combinatorial optimizations and can be analogize to the knapsack problem. That is, vehicles in set $\mathcal{V}$ are items in the knapsack problem. Costs of vehicles are weights of items in the knapsack problem. Comprehensive evaluations of coverage ratio and reliable degree for vehicles are values of items in the knapsack problem. Based on reliable degrees and coverage ratios, comprehensive evaluations of vehicle v_i are derived by Eq. (8).

$$\phi_i = \alpha \times \mathcal{S}_i + (1-\alpha) \times r_i, \phi_i \in (0,1] \tag{8}$$

whereby, ϕ_i represents comprehensive evaluation of vehicle v_i, which considers reliable degree and coverage ratio comprehensively. α is a influence factor, ranges from 0 to 1. With Eq. (8), comprehensive evaluations of vehicles in set $\mathcal{V}$ can be obtained, which can be regarded as item values.

According to above definitions, length of set $\mathcal{V}$ is n, which means that length of gene is n in the knapsack problem. Number of original genes is defined as ω. For the ith sequence $\mathcal{I}_i$ of a gene set $G_\hbar$, 0 represents that this vehicle v_i is not selected as a code disseminator and 1 represents that v_i is selected as a code disseminator, which is described in Eq. (9). Meanwhile, because the value of n will be very large in the VSNs, therefore, number of 1 need to be limited for a set of original gene, which is defined as k.

$$\mathcal{I}_i = \begin{cases} 1, & v_i \; is\, a\, code\, disseminator \\ 0, & v_i \; is\, not\, a\, code\, disseminator \end{cases} \tag{9}$$

Fitness calculations require a fitness function as ruler. There, total value of a gene set $G_\hbar$ is defined as the first fitness function, which is defined in Eq. (10).

$$\mathcal{F}_{G_\hbar} = \sum_{i \in 0 \to n} \phi_i \times \mathcal{I}_i \tag{10}$$

Here, $\mathcal{F}_{G_\hbar}$ is individual fitness of gene set $G_\hbar$. Eq. (10) is used to calculate the value of each gene set (each vehicular set). The bigger the value of $\mathcal{F}_{G_\hbar}$ is, the better.

Then, vehicular costs can be regarded as item weights in the knapsack problem. For each gene set like $G_\hbar$, total costs of items are calculated by the second fitness function, which is defined by Eq. (11).

$$\begin{cases} if\; \mathcal{K}_{G_\hbar} \le \mathcal{E}, \quad \mathcal{K}_{G_\hbar} = \sum_{i \in 0 \to n} \mathcal{C}_i \times \mathcal{I}_i \\ if\; \mathcal{K}_{G_\hbar} > \mathcal{E}, \; abandon\, gene\, set\, G_\hbar \end{cases} \tag{11}$$

wherein, $\mathcal{K}_{G_\hbar}$ represents total costs of vehicular disseminators (total weight of items) for a gene set. And $\mathcal{E}$ represents vehicular cost limitation, which is a threshold of total costs. If costs of a gene set are less than $\mathcal{E}$, then this gene set can be extended to the next generation. Otherwise, this gene set will be abandoned. Along with the value, the costs of each gene set can be obtained by the Eq. (11).

Suppose that a generation $\mathcal{G}_\tau$ has $\mathcal{N}$ gene sets, and there are μ number of gene sets that satisfy the Eq. (11), then total costs of $\mathcal{G}_\tau$ can be derived by Eq. (12).

$$C_{\mathcal{G}_\tau} = \sum_{0<\hbar\le\mu} \mathcal{K}_{G_\hbar} \tag{12}$$

Based on the Eqs. (10) and (12), the probability that a gene set $G_\hbar$ in a generation $\mathcal{G}_\tau$ can be extended to next generation $\mathcal{G}_{\tau+1}$ is derived by Eq. (13), according to individual differences.

$$P_{G_\hbar} = \frac{\mathcal{F}_{G_\hbar}}{C_{\mathcal{G}_\tau}} = \frac{\sum_{i\in 0\to n} \phi_i \times \mathcal{I}_i}{\sum_{0<\hbar\le\mu} \mathcal{K}_{G_\hbar}} \tag{13}$$

Obviously, if the value of $\mathcal{F}_{G_\hbar}$ is larger, gene set $G_\hbar$ has more probability to be extended to the next generation. Through gene recombination and mutations of gene set in generation $\mathcal{G}_\tau$, number of gene sets in next $\mathcal{G}_{\tau+1}$ is generated randomly. There, mutation probability is defined as ρ. Then repeat Eqs. (9) to (13) and find the gene set G_ψ with the maximum individual fitness value $MAX(\mathcal{K}_{G_\psi})$, based on limited costs $\mathcal{E}$. In gene sequence of G_ψ, 1 represents that vehicle in this position is selected as a code disseminator and 0 represents that vehicle in this position is not a code disseminator. With the iteration processes of Eqs. (9) to (13), values of individual fitness in generation $\mathcal{G}_{\tau+1}$ are better than those in generation $\mathcal{G}_\tau$ in general. During iteration processes, if fitness function values of successive generations do not increase or even decrease, then the function has converged. Finally, by using genetic algorithm, the MLCD scheme can find vehicles with high reliability degrees and coverage ratios to be code disseminators.

In general, based on genetic algorithm, selections of reliability-based and coverage-based code disseminators are introduced in Algorithm 3.

Algorithm 3: Selections of vehicular code disseminators based on genetic algorithm

Input: $\mathcal{V}, r[n], \mathcal{S}[n], \mathcal{V}_s, c[n]$
Output: Set of reliability-based and coverage-based vehicular code disseminators

```
1:  Initial V_s[m] = {0};
2:  Initial all the variables;
3:  Step 1: initialization process
4:      Define population size, hybridization probability,
        catastrophe probability, gene length and generation
        number
5:      Generate gene set  // randomly generating a 0-1 sequence
6:      If this gene set satisfies fitness function Eq. (11)
7:          Adopt this gene set as a member
8:      Else
9:          Return back to 5
10:     Calculate fitness value of all the gene set according to Eq. (10)
11: Step 2: gene selection
12:     According to Eq. (13), use roulette method to
        choose next generation
13: Step 3: hybrid and mutation
14:     Hybrid:
15:         Randomly generate two integers for hybrid
16:         Generate a new gene set
17:         If this new gene set satisfy Eq. (11)
18:             This gene set is a result of hybrid
19:         Else
20:             Return back to 14
21:     Mutation:
22:         Random choose a position and change the value
        of it
23:         Same as 16-20
24: Step 4: ending  // reach the generation number
25:     Select the gene set with maximum fitness value
        based on Eq. (10)
26:     V_s[m] = this gene set
27: Output V_s[m]
28: End Algorithm 3
```

5. Experiments and simulations

In this section, firstly, experimental settings are introduced; secondly, to evaluated the performance of our proposed scheme, simulation results of the MLCD scheme are conducted.

Experiments are conducted based on C++ and MATLAB platform. Historical datasets of T-Driver provided by MSRA are used to evaluate performances of the MLCD scheme. The datasets that are recorded by GPS include 10,357 vehicles, which is suitable platform for making simulations. Periods of the dataset ranges from Feb.2 to Feb. 8. Total

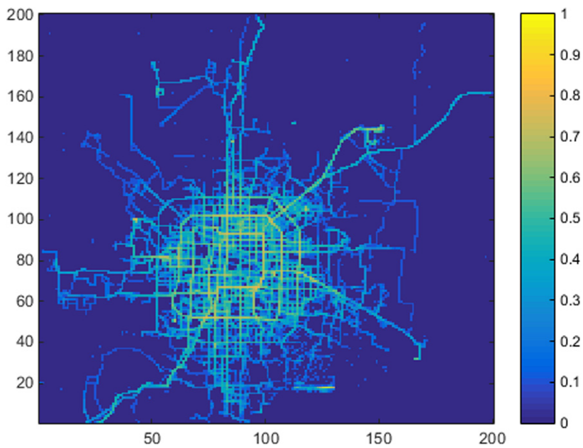

Fig. 2. Total trajectories of vehicular datasets.

length of trajectories reaches 9 million kilometers approximately. Fig. 2 shows the total trajectories of this T-Driver dataset. After processing this dataset, there are 1126 vehicles that do not have historical GPS data in the datasets. Therefore, the rest of vehicles are used to do the following experiments.

To verify performances of the MLCD scheme, both safety degree and coverage ratio need to be evaluated. With the MLCD scheme, vehicles with higher reliable degree and coverage ratio are chosen as code disseminators whose number is defined as ϱ. Therefore, improvements of safety degree can be derived by calculating number of vehicles with higher reliable ratios. There, we set a threshold value θ.

Number of vehicles whose reliable degrees are greater than θ or equal to θ is defined as γ in the MLCD scheme. Based on previous scheme, it is defined as ϖ. Improvement calculations of safety degree are evaluated in Eq. (14).

$$\mathcal{X}_s(\theta) = \frac{\gamma(\theta) - \varpi(\theta)}{\gamma(\theta)} \tag{14}$$

Different value of θ will result in different value of $\mathcal{X}_s$, which is introduced in the following comparisons.

Then to evaluate improvements of coverage degree, within limited costs, the performances of total coverage ratio for code disseminator set are calculated by Eq. (15).

$$\mathcal{Y}_c = \frac{\mathcal{S}^{OCRV}_{\mathcal{V}_s} - \mathcal{S}^{pre}_{\mathcal{V}_s}}{\mathcal{S}^{OCRV}_{\mathcal{V}_s}} \tag{15}$$

Here, within limited costs $\mathcal{E}$ of the GCS, $\mathcal{S}^{OCRV}_{\mathcal{V}_s}$ is the coverage ratio of $\mathcal{V}_s$ in the MLCD scheme and $\mathcal{S}^{pre}_{\mathcal{V}_s}$ is the coverage ratio of $\mathcal{V}_s$ in previous scheme. Value of $\mathcal{Y}_c$ reflects improvement coverage ratio based on MLCD scheme. Cost limitation $\mathcal{E}$ of code disseminator set can affect total coverage ratio, which is discussed in the followings.

To evaluate performances of the MLCD scheme comprehensively, $\mathcal{Z}$ is set based on safety degree in Eq. (14) and coverage ratio in Eq. (15). Calculation of $\mathcal{Z}$ is defined in Eq. (16).

$$\mathcal{Z} = \frac{2 \times \mathcal{U}(\theta) \times \mathcal{S}_{\mathcal{V}_s}}{\mathcal{U}(\theta) + \mathcal{S}_{\mathcal{V}_s}} \tag{16}$$

Consider both coverage ratio and reliable degree, Eq. (16) can be used to evaluate performances of the MLCD scheme comprehensively. In Eq. (16), $\mathcal{U}(\theta)$ is the safety degree of a code disseminator set, which is $\mathcal{U}(\theta) = \gamma(\theta)/\varrho$. And $\mathcal{S}_{\mathcal{V}_s}$ represents the coverage ratio of code disseminator set $\mathcal{V}_s$. Value of $\mathcal{Z}$ can reflect performances of the MLCD scheme comprehensively.

In our simulations, 180 GPS points are randomly chosen as landmarks, which is $\ell = 180$. Coordinates of part of the 180 landmarks are shown in Table 2.

Table 2
Coordinates of landmarks.

Number	X value	Y value
1	116.4003	39.91668
2	116.3931	39.92692
3	116.3898	39.93863
4	116.3466	39.97332
5	116.34827	39.93638
6	116.2257	39.87014
7	116.215	39.84733
8	116.15339	39.761
9	116.1	39.94
…	…	…
171	116.613	40.11841
172	116.6	40.144977
173	116.63821	40.21988
174	116.47608	40.04116
175	116.38014	39.85706
176	116.34001	39.8501
177	116.26823	39.94867
178	116.52281	39.95845
179	116.34857	40.21004
180	116.11034	39.84356

With coordinates in Table 2, coverage ratio of vehicles can be derived.

Then, experimental results are evaluated in the following. Firstly, vehicular costs are calculated, which is shown in Fig. 3(a).

In simulations, vehicles with empty historical datasets are filtered out. Thus, vehicular number is 9231. Costs of the 9231 vehicles are derived according to Eq. (4), which are used in the following genetic algorithms. Costs of a gene set are the second fitness function in genetic algorithm. In the simulations, threshold of costs is set as 25 000.

Secondly, reliable degrees of the 9231 vehicles are evaluated, which is shown in Fig. 3(b).

In Fig. 3(b), values of vehicular reliable degrees are calculated according to Eq. (7). Thirdly, based on above analysis, number of landmarks covered by each vehicle is shown in the Fig. 3(c).

Based on the 180 landmarks in the Table 2, Fig. 3(c) shows the number of landmarks covered by each vehicle. Then, coverage ratios of vehicles can be evaluated, which is shown in Fig. 3(d).

With Eq. (6), coverage ratio of each vehicle can be derived, the tendency of Fig. 3(d) is the same as that of the Fig. 3(c). Coverage ratios of vehicles in Fig. 3(d) are utilized in the following vehicular selections.

According to Eqs. (8) to (10), value of ϕ_i is calculated based on both coverage ratio and reliable degree, which is utilized in the first fitness function. There, value of α is set as 0.5. Based on Eq. (8), value of ϕ_i for a vehicle v_i is evaluated comprehensively in Fig. 3(e).

Fig. 3(e) shows the value of ϕ_i for vehicle v_i in the simulation set, which is used in the first fitness function. Based on value of ϕ_i and value of $\mathcal{C}_i$, genetic algorithm is utilized to select vehicles as code disseminators.

In simulations, the size of population is defined as 100 and the probability of crossover is 0.618. We also consider the mutation situation in real environments. The probability of mutation is set as $3e^{-4}$. Length of the chromosome is set as 9231 and number of generations is set as 1000. With analysis above, simulations are made to select both coverage-based and reliability-based code disseminators.

With method of genetic algorithm, simulation results show that the population's max fitness reaches to 16 941 in the 960th generation G_{960}. Total costs of this gene set are 24 994, which is in the threshold 25 000. Based on above analysis, in the gene G_{960}, sequences where value are 1 are selected as code disseminators. G_{960} selects 167 vehicles as code disseminators. Number of landmarks that G_{960} covers is 171. According to Eq. (6), coverage ratio of the G_{960} set is $\frac{171}{180} \times 100\% = 95\%$. To evaluate performances of the proposed scheme, coverage regions of the code disseminators based on MLCD scheme are compared to that based on previous scheme, which is shown in Fig. 4(a).

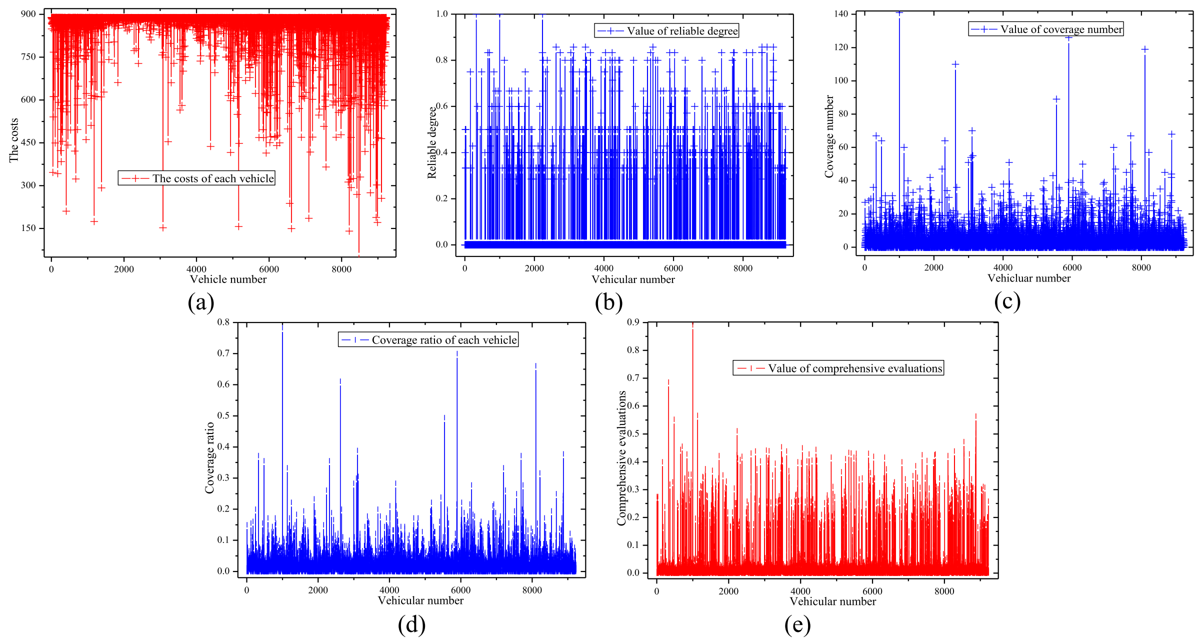

Fig. 3. (a) Vehicular costs; (b) Reliable degree of vehicles; (c) Number of landmarks covered by vehicles; (d) Coverage ratio of vehicles; (e) Value of comprehensive evaluation ϕ_i.

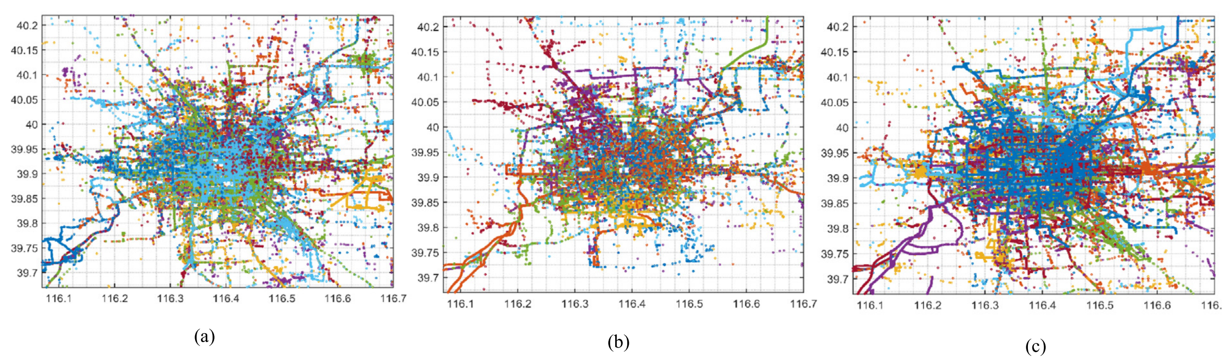

Fig. 4. (a) Coverage regions of code disseminators based on MLCD scheme; (b) Coverage regions of code disseminators based on random-selection scheme; (c) Coverage regions of code disseminators based on coverage-only scheme;.

And the coverage regions based on random selection scheme are shown in Fig. 4(b). In the random selection scheme, 167 vehicles are randomly chosen as code disseminators.

Clearly, coverage regions in Fig. 4(b) are smaller than those in Fig. 4(a). With evaluation, vehicles in Fig. 4(b) cover 132 regions. According to Eq. (6), coverage ratio of the random selection scheme is 73.3%. Therefore, based on the MLCD scheme, coverage ratio can be improved by 23.16%, which shows efficiency of the proposed scheme. It is because that the MLCD scheme considers coverage factor in vehicular selection processes by using genetic algorithm. Compared to MLCD scheme, coverage regions of each code disseminator based on the random-selection scheme are shown in Fig. 5. Fig. 5 reflects that the number of coverage regions for each vehicle based on the MLCD scheme is larger than that in the random-selection scheme. The coverage of vehicles is sorted from large to small. With genetic algorithm, the vehicles can cover more landmarks and reach a greater coverage ratio.

Then, to further evaluate performances of coverage ratio, coverage ratio of the MLCD scheme is compared to that of the coverage-only scheme, which is shown in Fig. 4(c).

Compared to Fig. 4(a), coverage regions based on coverage-only scheme are similar to those based on MLCD scheme. With evaluation, in Fig. 4 (c), the code disseminators can cover 177 landmarks of Table 2. Therefore, coverage ratio of coverage-only scheme is 98.3%. Coverage ratio based on the MLCD scheme is decreased to 3.39% approximately, which has little impact on the whole performances.

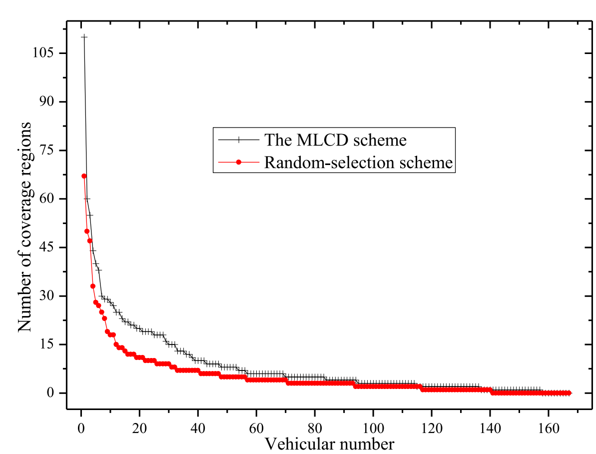

Fig. 5. Coverage regions of vehicles based on MLCD.

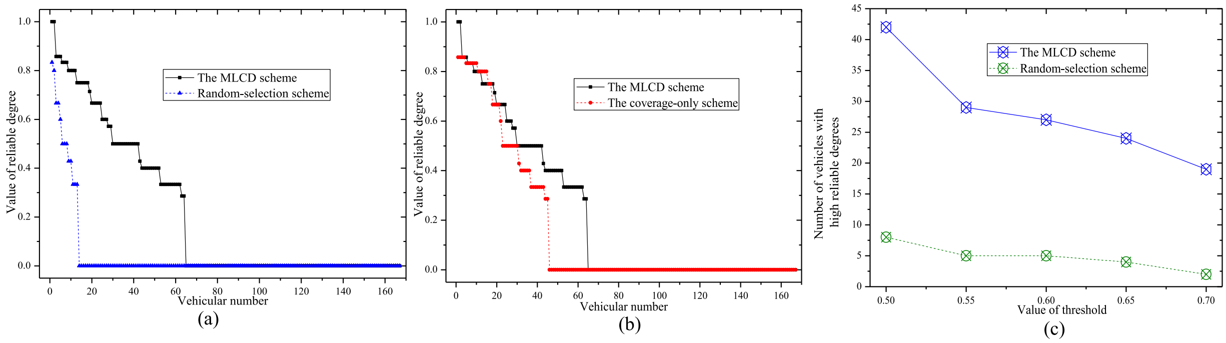

Fig. 6. (a) Comparisons of vehicular reliable degree based on MLCD scheme and random-selection scheme; (b) Comparisons of vehicular reliable degree based on MLCD scheme and coverage-only scheme; (c) Compared to random-selection scheme, number of vehicles with high reliable degrees based on $\theta = 0.5$, 0.55, 0.6, 0.65 and 0.7.

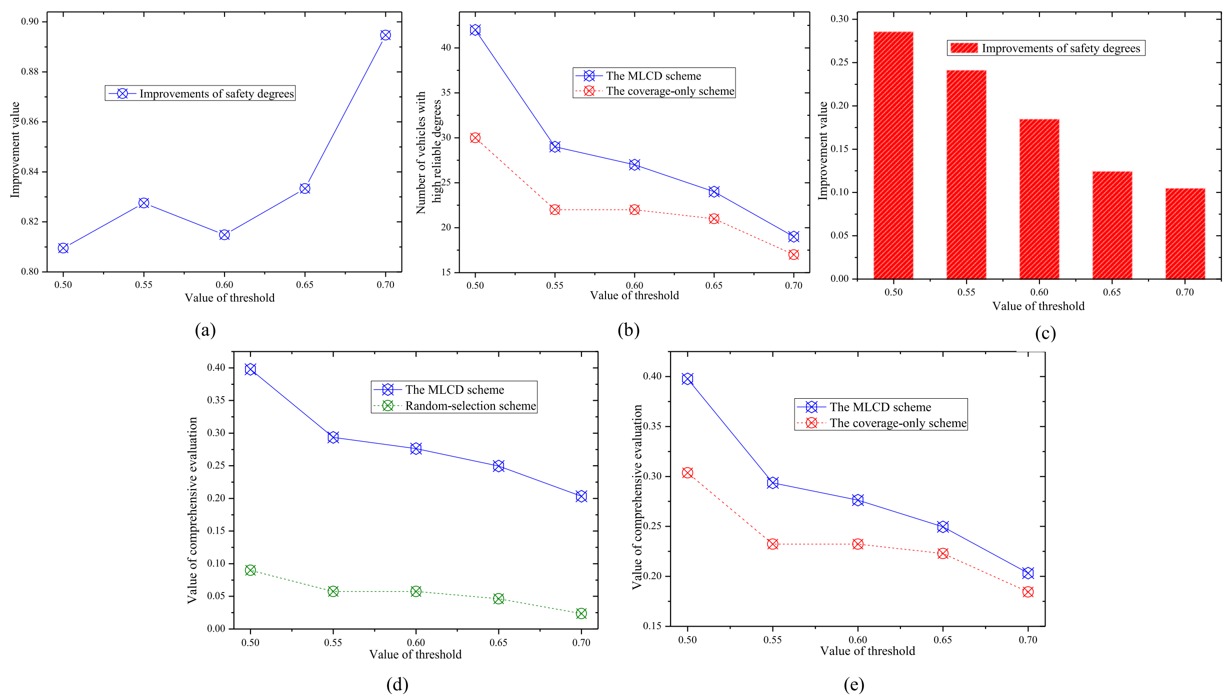

Fig. 7. (a) Compared to random-selection scheme, improvements of safety degrees based on $\theta = 0.5$, 0.55, 0.6, 0.65 and 0.7; (b) Compared to coverage-only scheme, number of vehicles with high reliable degrees based on $\theta = 0.5$, 0.55, 0.6, 0.65 and 0.7; (c) Compared to coverage-only scheme, improvements of safety degrees based on $\theta = 0.5$, 0.55, 0.6, 0.65 and 0.7; (d) Compared to random-selection scheme, value of $\mathcal{Z}$ based on $\theta = 0.5$, 0.55, 0.6, 0.65 and 0.7; (e) Compared to coverage-only scheme, value of $\mathcal{Z}$ based on $\theta = 0.5$, 0.55, 0.6, 0.65 and 0.7.

The MLCD scheme not only consider coverage factor, but also reliable degrees of code disseminators. Therefore, with the two influence factors, coverage ratio of the MLCD scheme is decreased minimally, but do not have much effect on the whole performances.

After evaluating performances of coverage ratio based on the MLCD scheme, we then evaluate the safety degree of code disseminations by evaluating reliable degrees of vehicles, and compare them to random-selection scheme and coverage-only scheme respectively. Firstly, reliable degrees of vehicles based on proposed scheme is compared to those based on random-selection scheme, which is shown in Fig. 6(a). Fig. 6(a) shows comparisons of reliable degree based on two schemes. Clearly, reliable degrees of code disseminators in the MLCD scheme are higher than those in the random-selection scheme. With selections of reliability, the proposed scheme can choose vehicles with higher reliable degrees as code disseminators, which enhances safety of code disseminations to a large scale.

To evaluate reliability performances, we then compare reliable degrees of code disseminators in the MLCD scheme to those in the coverage-only scheme. Comparisons are shown in Fig. 6(b).

In Fig. 6(b), compared to the MLCD scheme, values of reliable degrees in the coverage-only scheme decline faster. Clearly, there are more vehicles with higher reliable degrees in the MLCD scheme. With reliability-based selection method, the MLCD scheme can improve safety degree of code disseminations by increasing number of reliable vehicles as code disseminators, compared to the coverage-only scheme. Both Fig. 6(a) and (b) show advancements of the proposed scheme.

In the above analysis, to evaluate performances of safety degrees, we need to calculate number of code disseminators whose reliable degrees are bigger than threshold θ. Therefore, value of θ can influence improvements of safety degree. With Eq. (14), compared to random-selection scheme, number of vehicles with high reliable degrees are evaluated based on value of $\theta = 0.5$, 0.55, 0.6, 0.65 and 0.7 respectively, which are shown in Fig. 6(c).

In Fig. 7(a), clearly, compared to random selections scheme, there are more reliability-based vehicles in the MLCD scheme based on different value of threshold θ. It is because that the proposed scheme takes reliable degree of vehicles into considerations. Compared to random-selection scheme, the MLCD scheme can improve safety degree by 80.95% to 89.47%. In general, safety degree can be improved by 83.6%. The improvements of safety degrees are shown in Fig. 7(a).

Then, to further evaluate safety performances of the proposed scheme, compared to coverage-only scheme, number of vehicles with high reliable degrees are evaluated based on value of $\theta = 0.5$, 0.55, 0.6, 0.65 and 0.7 respectively, which is shown in Fig. 7(b).

In Fig. 7(b), compared to coverage-only scheme, number of vehicles with high reliable degrees is bigger in the MLCD scheme, which improves safety degrees of code disseminations. Compared to coverage-only scheme, improvements of safety degrees are shown in Fig. 7(c).

Based on different value of θ, compared to coverage-only scheme, safety improvements ranges from 10.53% to 28.57% in Fig. 7(c). In general, the MLCD scheme can improve safety degree by 18.86% comprehensively.

After evaluating coverage ratio and reliable degree of code disseminators in the three schemes respectively, performances of the MLCD scheme are evaluated comprehensively by utilizing Eq. (16). Firstly, considering both coverage ratio and safety degree, based on different value of θ, value of $\mathbb{Z}$ can be derived by Eq. (16). Comparisons between MLCD scheme and random-selection scheme are shown in Fig. 7(d).

Clearly, performances of the MLCD scheme are better than those of the random-selection scheme. Based on value of $\theta = 0.5$, 0.55, 0.6, 0.65 and 0.7, compared to random-selection scheme, comprehensive evaluation $\mathbb{Z}$ can be improved to 77.39% to 88.4% approximately. With evaluation in Fig. 7(d), average improvement is 80.65%.

Based on different value of θ, comparisons of $\mathbb{Z}$ between the MLCD scheme and coverage-only scheme are shown in Fig. 7(e).

In Fig. 7(e), based on different value of θ, compared to coverage-only scheme, improvements of comprehensive evaluation $\mathbb{Z}$ can reach to 23.66%. In general, the improvements of $\mathbb{Z}$ is 17.25%, which shows efficiency of the proposed scheme.

Based on above evaluations, compared to two previous schemes, the MLCD scheme can improve both coverage ratio and safety degree. Value of $\mathbb{Z}$ shows the efficiencies and advancements of the proposed scheme.

6. Conclusions

In this paper, we propose the MLCD scheme that aims to improve safety degree of code disseminations by calculating vehicular reliability, and to improve coverage ratio of program codes by calculating vehicular coverage. Firstly, vehicular reliability is evaluated by its historical trajectory dataset. Specifically, the vehicle with a regular trajectory is more reliable compared to vehicle without a regular trajectory. Secondly, the vehicle with higher coverage ratio is preferred to be code disseminators. Thirdly, considering vehicular costs of the GCS, based on those two factors (reliability and coverage) and limited costs, vehicular code disseminators are selected by using genetic algorithm. Simulations and experiments are made based on trajectory datasets of a real city. In addition, compared to two previous schemes, safety degree of code disseminations can be improved by 83.6% and 18.86%, and coverage ratio of updated program codes can be improved by 23.16%. In general, performances of the MLCD scheme can be improved by 80.65% and 17.25% respectively, which shows advancement and efficiency. Our future research will focus on improving information security of IoT in the 5G networks by utilizing more advanced machine learning models.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Ting Li: Methodology, Writing - original draft. **Ming Zhao:** Funding acquisition, Supervision. **Kelvin Kian Loong Wong:** Investigation, Writing - review & editing.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (No. 61572526, 61772554, 61572528).

References

[1] M. Chen, et al., Cognitive information measurements: A new perspective, Inform. Sci. 505 (2019) 487–497.

[2] D.T. Huyhnh, et al., Social-aware energy efficiency optimization for device-to-device communications in 5G networks, Comput. Commun. 120 (2018) 102–111.

[3] M. Chen, Y. Hao, C. Lai, D. Wu, Y. Li, et al., Opportunistic task scheduling over co-located clouds in mobile environment, IEEE Trans. Serv. Comput. 11 (3) (2018) 549–561.

[4] X. Liu, A. Liu, T. Wang, K. Ota, M. Dong, Y. Liu, Z. Cai, Adaptive data and verified message disjoint security routing for gathering big data in energy harvesting networks, J. Parallel Distrib. Comput. 135 (2019) 140–155.

[5] H. Teng, Y. Liu, A. Liu, N.N. Xiong, Z. Cai, T. Wang, X. Liu, A novel code data dissemination scheme for Internet of Things through mobile vehicle of smart cities, Future Gener. Comput. Syst. 94 (2019) 351–367.

[6] Y. Liu, A. Liu, T. Wang, X. Liu, N.N. Xiong, An intelligent incentive mechanism for coverage of data collection in cognitive internet of things, Future Gener. Comput. Syst. 100 (2019) 701–714.

[7] F. Tang, Z.M. Fadlullah, N. Kato, F. Ono, R. Miura, AC-POCA: Anticoordination game based partially overlapping channels assignment in combined UAV and D2D-based networks, IEEE Trans. Veh. Technol. 67 (2) (2018) 1672–1683.

[8] Y. Liu, M. Ma, X. Liu, N.N. Xiong, A. Liu, Y. Zhu, Design and analysis of probing route to defense sink-hole attacks for internet of things security, IEEE Trans. Netw. Sci. Eng. (2018) http://dx.doi.org/10.1109/TNSE.2018.2881152.

[9] Y. Liu, Z. Zeng, X. Liu, X. Zhu, M. Bhuiyan, A novel load balancing and low response delay framework for edge-cloud network based on SDN, IEEE Internet Things J. (2019) http://dx.doi.org/10.1109/JIOT.2019.2951857.

[10] Y. Chen, W. Liu, T. Wang, Q. Deng, A. Liu, H. Song, An adaptive retransmit mechanism for delay differentiated services in industrial WSNs, EURASIP J. Wireless Commun. Networking (2019) http://dx.doi.org/10.1186/s13638-019-1566-2.

[11] M. Huang, W. Liu, T. Wang, A. Liu, S. Zhang, A cloud-MEC collaborative task offloading scheme with service orchestration, IEEE Internet Things J. (2019) http://dx.doi.org/10.1109/JIOT.2019.2952767.

[12] Y. Liu, A. Liu, X. Liu, M. Ma, A trust-based active detection for cyber-physical security in industrial environments, IEEE Trans. Ind. Inf. 15 (12) (2019) 6593–6603.

[13] D. Kim, Y. Velasco, W. Wang, R. Uma, R. Hussain, S. Lee, A new comprehensive RSU installation strategy for cost-efficient VANET deployment, IEEE Trans. Veh. Technol. 66 (5) (2017) 4200–4211.

[14] S. Abdelhamid, H.S. Hassanein, G. Takahara. Reputation-Aware, Trajectory-based recruitment of smart vehicles for public sensing, IEEE Trans. Intell. Transp. Syst. 19 (5) (2018) 1387–1400.

[15] B. Jain, G. Brar, J. Malhotra, S. Rani, S.H. Ahmed, A cross layer protocol for traffic management in Social Internet of Vehicles, Future Gener. Comput. Syst. 82 (2018) 707–714.

[16] J.H. Lim, K. Naito, J.H. Yun, M. Gerla, Reliable safety message dissemination in NLOS intersections using TV white spectrum, IEEE Trans. Mob. Comput. 17 (1) (2018) 169–182.

[17] F. Tang, B. Mao, Z.M. Fadlullah, N. Kato, On a novel deep-learning-based intelligent partially overlapping channel assignment in SDN-IoT, IEEE Commun. Mag. 56 (9) (2018) 80–86.

[18] Penghao Sun, et al., TIDE: Time-relevant deep reinforcement learning for routing optimization, Future Gener. Comput. Syst. 99 (2019) 401–409.

[19] N. Arunkumar, M.A. Mohammed, M.K.A. Ghani, D.A. Ibrahim, E. Abdulhay, G. Ramirez-Gonzalez, V.H.C. Albuquerque, K-means clustering and neural network for object detecting and identifying abnormality of brain tumor, Soft Comput. (2018) 1–14.

[20] J. Oehling, D.J. Barry, Using machine learning methods in airline flight data monitoring to generate new operational safety knowledge from existing data, Saf. Sci. 114 (2019) 89–104.

[21] F. Tang, Z.M. Fadlullah, B. Mao, N. Kato, An intelligent traffic load prediction based adaptive channel assignment algorithm in SDN-IoT: A deep learning approach, IEEE Internet Things J. 5 (6) (2018) 5141–5154.

[22] T. Li, N.N. Xiong, J. Gao, H. Song, A. Liu, Z. Cai, Reliable code disseminations through opportunistic communication in vehicular wireless networks, IEEE Access 6 (2018) 55509–55527.
[23] G. Sun, Y. Zhang, D. Liao, H. Yu, X. Du, M. Guizani, Bus-trajectory-based street-centric routing for message delivery in urban vehicular ad hoc networks, IEEE Trans. Veh. Technol. 67 (8) (2018) 7550–7563.
[24] F. Tang, B. Mao, Z.M. Fadlullah, N. Kato, Q. Akashi, T. Inoue, K. Mizutani, On removing routing protocol from future wireless networks: A real-time deep learning approach for intelligent traffic control, IEEE Wirel. Commun. 25 (1) (2018) 154–160.
[25] L. Liang, H. Ye, G. Li, Towards intelligent vehicular networks: A machine learning framework, IEEE Internet Things J. 6 (1) (2019) 124–135.
[26] J.E. Siegel, D.C. Erb, S.E. Sarma, A survey of the connected vehicle landscape—Architectures, enabling technologies, applications, and development areas, IEEE Trans. Intell. Transp. Syst. 19 (8) (2018) 2391–2406.
[27] A. Braga, S. Sargento, O.K. Tonguz, Smarter cities with parked cars as roadside units, IEEE Trans. Intell. Transp. Syst. 19 (7) (2018) 1–15.
[28] Z. Zhou, C. Gao, C. Xu, Y. Zhang, S. Mumtaz, J. Rodriguez, Social big-data-based content dissemination in internet of vehicles, IEEE Trans. Ind. Inf. 4 (2) (2018) 768–777.
[29] L. Liu, C. Chen, T. Qiu, M. Zhang, S. Li, B. Zhou, A data dissemination scheme based on clustering and probabilistic broadcasting in VANETs, Veh. Commun. 13 (2018) 78–88.
[30] R. Kaewpuang, D. Niyato, P. Tan, P. Wang, Cooperative management in full-truckload and less-than-truckload vehicle system, IEEE Trans. Veh. Technol. 66 (7) (2017) 5707–5722.
[31] V. Radhakrishna, S.A. Aljawarneh, P.V. Kumar, V. Janaki, A novel fuzzy similarity measure and prevalence estimation approach for similarity profiled temporal association pattern mining, Future Gener. Comput. Syst. 83 (2018) 582–595.
[32] A. Anjomshoaa, F. Duarte, D. Rennings, T. Matarazzo, P. Souza, C. Ratti, City scanner: Building and scheduling a mobile sensing platform for smart city services, IEEE Internet Things J. 5 (6) (2018) 4567–4579.