IEEE Open Journal of
Vehicular Technology

Received 23 March 2020; accepted 15 April 2020. Date of publication 22 April 2020;
date of current version 26 May 2020. The review of this article was arranged by Associate Editor R. Lu.

Digital Object Identifier 10.1109/OJVT.2020.2988645

Dimensioning and Layout Planning of 5G-Based Vehicular Edge Computing Networks Towards Intelligent Transportation

BIN LIN (Member, IEEE), XIAN ZHOU, AND JIANLI DUAN

Department of Information Science and Technology, Dalian Maritime University, Liao Ning, China

CORRESPONDING AUTHOR: BIN LIN (e-mail: binlin@dlmu.edu.cn)

The work was supported in part by the National Natural Science Foundation of China under Grants 61971083 and 51939001 and in part by Dalian Science and Technology Innovation Fund under Grant 2019J11CY015.

ABSTRACT Fast-response communication is crucial for Vehicular Ad Hoc Network (VANET). In practice, the conventional VANETs, suffering from the high mobility of the vehicles and the ever-growing data to percept and process, cannot meet the demand of fast response currently. In this paper, we study the Dimensioning and Layout Planning (DLP) problem under 5G-based Vehicular Edge Computing Network (VECN) architecture which integrates the 5G Micro Base Station (gNB) and Edge Computing (EC) to reduce the response time. The DLP problem aims to minimize the total placement cost under the constraint of the full coverage. This paper formulates the DLP problem as an integer linear program (ILP) and then proposes a Greedy Algorithm (GA) and a Cost-Effective Heuristic Algorithm (CEHA) toimprove the computation efficiency. The case studies have verified the feasibility and scalability of the DLP formulation and showed that the proposed CEHA is fairly effective and efficient to solve the DLP problem.

INDEX TERMS Vehicular Ad Hoc network, edge computing, gNB, vehicular edge computing network, dimensioning and layout planning.

I. INTRODUCTION

As a vital component of the Intelligent Transportation System (ITS), Vehicular Ad Hoc Network (VANET) now has immersed itself in all walks of life to form an indivisible part of the new Information Age [1]–[5]. Each vehicle in VANETs has an On Board Unit (OBU) to communicate with the neighbor vehicles or the Roadside Unit (RSU), and the connected vehicles are envisioned to provide intelligent transportation services (e.g., high definition map, autonomous driving, smart roadside units) and offer various kinds of in-vehicle infotainment applications (e.g., 3-D virtual reality game and web browsing) [5]–[7].

Driven by the services mentioned above, a large quantity of researches are arouse to improve the fast-response and high-bandwidth communications, because the delay and bandwidth shortage would affect the driving routes and experiences [5]. Considering the case when a user drives amid dense fog, the neighbor vehicles or local status could not be observed,but all the vehicles couldrely on the surrounding vehicles or the RSUs which could transmit the real-time dynamic information (e.g., the speed of surrounding vehicles, accidents and road status) to help avoid traffic collision.

However, it is a very complex issue due to the fast changing network topology, diverse road types, vehicle's high mobility and extremely huge volume of vehicular data. Moreover, the conventional computing paradigm for VANETs, that is Cloud Computing (CC), is totally centralized. CC manages a data pool to process the data [6]. Although the substantial communication, computation and storage capabilities make CC a common practice, delivering massive-volume vehicular data from users to the remote cloud to process and storage adds the latency and makes the cloud suffer from the bandwidth consumption at the same time. In a word, the CC is unsuitable for VANETs' real-time services and the VANET architecture is thirst for the fast-response and high-bandwidth communications.

Driven by this motivation, the 5G-based Vehicular Edge Computing Network (VECN) architecture which integrates the 5G Micro Base Station (gNB) and Edge Computing (EC) is proposed. Instead of the conventional RSU, VECN architecture places gNBto render VANET a further improvement. The gNB is sweeping around the world recently due to its ultra-high data rate. Its performance requirements include (1) latency: millisecond level, (2) peak data rate: tens of Gbps and (3) connection density: 1 million/km^2 [8]. By offering local computation capabilities and processing the time-sensitive tasks locally, EC could reduce the response time and bandwidth consumption and therefore is recognized as the pillar technology to achieve 5G's soaring demands. Jointly applying gNB and ECP turns out to be an adequate solution for faster processing and is viewed as the key to break through the conventional VANET's bottleneck [9]–[12]. VECN architecture aims at satisfying the following requirements:

- Fast response: The relatively long distance between the vehicles and the remote cloud data center would add the transmission delay. In contrast, EC pushes the cloud closer to where data are generated. The much shorter transmission route of EC could reduce the transmission delay. Moreover, the number of the users that EC support is much smaller than that of CC, which means EC could relieve the computation delay [9]. In a nutshell, ECP could facilitate fast response.
- High bandwidth: CIF recorded 78% of UK companies have already introducing CC. Moving such massive data from users to the remote and centralized cloud data center results in not only the bandwidth consumption, but also an immense load on the cloud. EC could process the data locally, so as to circumvent the bandwidth consumption.

Although VECN architecture could improve the fast-response and high-bandwidth communications, deploying the architecture in the real world is a complex task. The straightforward method is to blindly place more gNBs and ECPs, yet their high cost hinders the placement [13]. In a nutshell, studying the Dimensioning and Layout Planning (DLP) problem under VECN architecture to generate the optimal practicable layout solution is really necessary. The contributions of this paper are summarized as follows.

- This paper investigates an optimization framework for the DLP problem of VECN, which integrates the gNB and ECP to improvefast-response and high-bandwidth communications.
- The DLP problem is formulated as an integer linear program (ILP) with the objective to minimize the total placement cost under the constraints of the full coverage, tree-based structure connectivity, communication range, capacity, etc.
- A Greedy Algorithm (GA) and a Cost-Effective Heuristic Algorithm (CEHA) are proposed to efficiently solve the formulated problem. The proposed GA aims at placing the nodes that could connect with more child nodes to decrease the number of the selected nodes. It could obtain the layout solution efficiently, but it is inclined to be trapped in the local optima. The proposed CEHA initializes a probability matrix P_{ij}. P_{ij} improves the randomness to avoid GA's local optima, while it is directional to improve the computation efficiency, at the same time. This paper conducts a series of simulations to verify the feasibility and scalability of the DLP optimization framework. Simulation results also reveal that the proposed CEHA is fairly effective and efficient to solve the DLP problem.

The remainder of this paper is organized as follows. In Section II, we present related work, which is followed by the network architecture and problem statement in Section III. Section IV shows the problem formulation and Section V presents the proposed Heuristic algorithms. Section VI presents the simulations and numerical results. Finally, Section VII concludes the paper.

II. RELATED WORK

Edge computing could be distributed between the remote cloud and network edges so as to reduce the response time and could provide better services for VANETs [5]–[7], [9], [14]–[17]. Yu *et al.* [6] studied the optimal deployment and dimensioning (ODD) problem of two diverse modes, i.e., the coupling mode (CRF), and the decoupling mode (DRF), respectively. They formulated the ODD problem to minimize the deploymentcost and proposed a heuristic algorithm to solve ODD problem. Zhong *et al.* [7] proposed a message authentication scheme for edge computing based VANET. Certificateless ring signature was introduced to boost the efficiency and security. Zhu *et al.* [14] proposed a fog following me solution, considering two types of fog nodes: the conventional stationary nodes, and the mobile nodes, to literally deal with the time-critical data. Zhang *et al.* [16] investigated the delay-optimal cooperative edge caching in user-centric mobile networks and proposed a greedy content placement algorithm to reduce the transmission delay. Zhou *et al.* [17] proposed a consortium blockchain-based secure trading mechanism and a contract theory-based incentive mechanism for vehicle-to-grid (V2G). An edge computing-based task offloading mechanism was also proposed to improve the probability of block creation.

Numerous researchers have put effort on the study of VANETs [1], [4], [5], [11], [18]–[26]. Zhang *et al.* [1] studied an architecture which integrated the RFID and VANETs. The architecture aimed to monitor traffic flow to avoid traffic accident. The RFID-reader-embedded RSUs placement problem was formulated as an ILP. Wu *et al.* [18] considered two diverse access patterns, and formulated the deployment problem, which aimed tooptimally deploy the RSUs to realize the maximum aggregate throughput in view of vehicle speed, vehicle distribution and wireless interference. Jiang *et al.* [19] proposed a greedy algorithm and an improved greedy algorithm to deal with the RSU and incremental RSU deployment issues for traffic flow coverage requirement. Zhang *et al.* [20] utilized the Geometric Dilution of Precision to assess the proposed RSU deployment method's accuracy. They adopted the asynchronous particle swarm optimization algorithm to achieve a best localization accuracy with minimum RSUs. Mehar *et al.* [21] proposed ODEL architecture to optimally place RSU with the objective to tradeoff between the deployment costand transmission delay. Hafeez *et al.* [22] proposed an analytical model for assess the Dedicated Short-Range Communication (DSRC) reliability and delay. An adaptive algorithm (AMBA) was also proposed to boost VANET's reliability. Hafeez *et al.* [23] proposed a mobility model to obtain the probability of multi-hop connectivity and receive the broadcasted packets successfully. The model could also obtain the number of the vehicles on the road and effectively capture the effects of vehicles' dynamics on the safety services. Wang *et al.* [24] proposed an assistant vehicle localization method, which was composed of three collaborative BSs and based on direction-of-arrival estimation, to improve the vehicle localization accuracy. Bi *et al.* [25] proposed an urban multi-hop broadcast protocol, which included a forwarding node selection scheme to reduce emergency message redundancy and transmission delay. Bi *et al.* [26] studied the urban vehicular communications' user mobility models and addressed the packet loss issue for each handover scenario.

Recent years, the architecture integrating 5G and EC attracts increasingly interest of the researchers [2], [5], [9], [10], [13], [27]. Li *et al.* [27] proposed an orchestrating method, which fused edge-centric computing and content-centric networking together in a 5G mobile environment, to boost 5G network's service capability. The unique features of the architecture integrating 5G and EC, i.e., superb reliability and super-low latency, render the architecture an adequate solution for VANETsand [2], [5], [9] investigate the architecture's advantage. Khan *et al.* [2] proposed a new hierarchical VANET architecture to effectively relieve the delay. Luo *et al.* [5] studied the content prefetching and distribution in the architecture, which integrated the 5G and edge computing together. Luo *et al.* [9] proposed a data sharing architecture in 5G-VANET and proposed a graph theory based algorithm to address the data sharing issue. They also proposed an algorithm to better balance the content distribution.

Though the architecture integrating gNB and EC has become a hot topic in VANETs field (e.g., Khan *et al.* [2] focused on analyzing and comparing the traditional architecture and proposed architecture),due to the high mobility of VANETs, the high cost of the entities, the limited computing resources and so on, it is really necessary to study the dimensioning and layout planning tasks to better place the architecture in the real world. The existing researches rarely focus on the dimensioning and layout planning problem. [9] proposed by Luo *et al.* was the most relevant literature, while it aimed to efficiently solve the data sharing problem. This paper, however, focuses on minimizing the total placement cost under the constraints of the fullcoverage, tree-based structure connectivity, communication range, capacity, etc., and verifies the VECN architecture could provide guidance for real-world deployment.

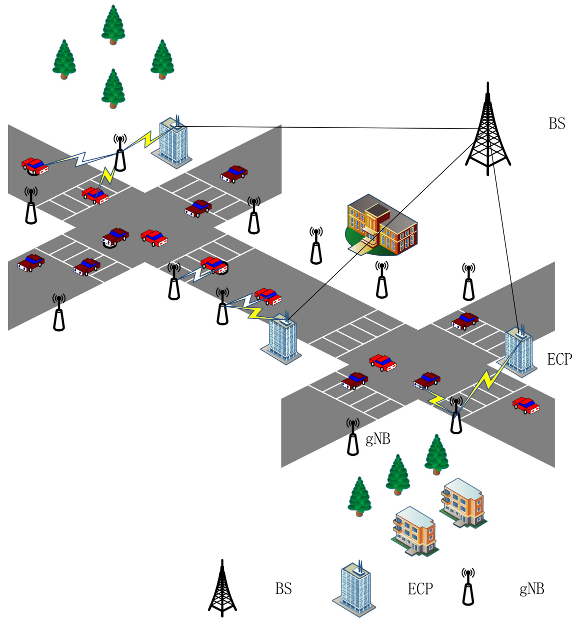

FIGURE 1. **The network architecture of VECN.**

III. NETWORK ARCHITECTURE AND PROBLEM STATEMENT

A. NETWORK ARCHITECTURE

As shown in Fig. 1, the VECN architecture introduces ECPs to extend the cloud to network edges. Associated by gNB, VECN could improve the fast-response and high-bandwidth communications.

VECN architecture is comprised of four entities, namely BS, ECP, gNB and vehicles. The gNB cancollect vehicles' data, and upload them to the surrounding ECPs. ECP is capable of preprocessing the time-sensitive data as the local information, and transmitting the results to BS or back to the vehicles. Then BS could relay the local information to the remote cloud data center. Vehicles can fetch the environmental and surrounding vehicular information from VECN, so as to improve the driving experiences. A concrete explanation is shown as follow.

- BS is regarded as the root node in the tree-based topology. It isintegrated for expanding information's coverage scope and ensuring the network connectivity. BS gathers all the local information from the surrounding ECPs and uploads the data to the remote cloud data center to regulate the traffic with a global view.
- ECP is hosted in the vicinity of end users as a local decider to serve as a complementary device to the cloud in this paper. Its function is to push the computation capabilities closer to where data are generated and preprocess the time-sensitive data immediately, thus reduces the respond time. Moreover, ECP would

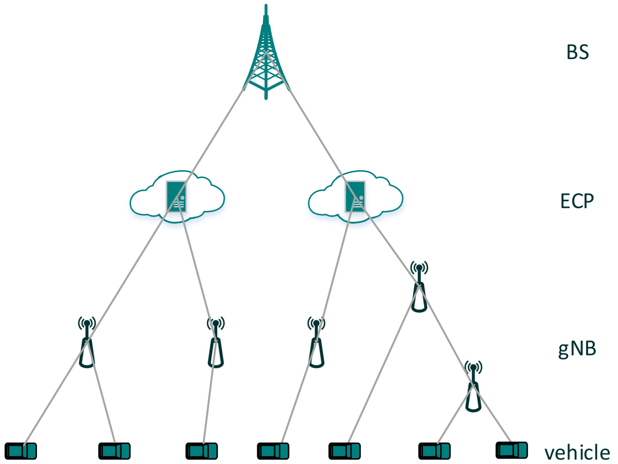

FIGURE 2. The network topology of VECN.

transmit the local information to the BS for boosting the city-level information.

- The gNB, a 5G micro base station, has considerable high data-rate advantages, such as sufficient bandwidth resources and millimeter wave-length spectra [5]. These crucial characteristics make it a promising enabler for VANETs. In this paper, instead of conventional RSU, we introduce gNB to connect vehicles with ECP in real time to improve the fast-response and high-bandwidth communications.
- Each vehicle in VANETs has an On Board Unit (OBU), which employs the wireless units inside, to transmit the data they generate to the neighbor vehicles or the road-side implements. A set of Test Points (TPs) specifically for the vehicles' mobile coverage test are defined within the target area. The DLP problem aims to optimally place gNB and ECP to cover all the TPs, in consideration of the tree-based structure connectivity, communication range and capacity.

B. PROBLEM STATEMENT

The DLP problem of VECN architecture is to generate the optimal layout solution with the objective of minimizing the total placement cost. The tree-based network topology of VECN is shown in Fig. 2. It can be observed that BS is regarded as the root node, ECP and gNB as the internal nodes, and vehicles as the leaf nodes, respectively.

Given:

- The fixed locations of BS and TPs.
- The sets of candidate sites for placing ECP and gNB.
- The deployment cost of BS, ECP and gNB.
- The communication range of the ECP and gNB, and the sensing range of gNB.
- The capacity of the BS, ECP and gNB.

Constraints:

- The tree-based network topology should be maintained.
- All the TPs must be covered by gNB.

TABLE 1. The Related Parameters

Parameter	Definition
$\boldsymbol{\Omega}_{BS}$	the set of BS
$\boldsymbol{\Omega}_{ECP}$	the set of ECP
$\boldsymbol{\Omega}_{gNB}$	the set of gNB
$\boldsymbol{\Omega}_{TP}$	the set of TP
d_{ij}	the Manhattan Distance between node i and j
C^B	the deployment cost of the BS
C^E	the deployment cost of the ECP
C^g	the deployment cost of the gNB
E_{com}	the communication range of the ECP
G_{com}	the communication range of the gNB
G_{sen}	the sensing range of the gNB
K_1	the maximum number of ECPs that an BS can accommodate
K_2	the maximum number of gNBs that an ECP can accommodate
K_3	the maximum number of TPs and gNBs that a gNB can accommodate

- The child nodes that BS, ECP and gNB connect with should be within the corresponding communication range or sensing range.
- The number of the child nodes that BS, ECP, and gNB accommodate should not exceed the corresponding capacity.
- The network connectivity should always be maintained.

Output:

- A network optimal layout solution including the selected nodes and the connected routes.
- The whole VECN network's deployment cost.

IV. PROBLEM FORMULATION

The parameters for DLP problem are listed in the Table 1. To formulate the DLP problem of VECN architecture, this paper maps the VECN into a directed graph G = ($\boldsymbol{\Omega}$, E). Note that G consists of a node set $\boldsymbol{\Omega}$ and a directed edge set **E**. Specifically, $\boldsymbol{\Omega} = \boldsymbol{\Omega}_{BS} \cup \boldsymbol{\Omega}_{\mathrm{ECP}} \cup \boldsymbol{\Omega}_{gNB} \cup \boldsymbol{\Omega}_{TP}$ where Ω_{BS}, $\boldsymbol{\Omega}_{ECP}$, $\boldsymbol{\Omega}_{gNB}$ and $\boldsymbol{\Omega}_{TP}$ denote the node sets of BS, ECP, gNB and TP, respectively.

Particularly, we explain the definition of the decision variables below, which are binary to indicate if a node or a connection exists.

$$b_e = \begin{cases} 1, & \text{an ECP exists at the candidate site } e; \\ 0, & \text{Otherwise.} \end{cases}$$

$$c_n = \begin{cases} 1, & \text{a gNB exists at the candidate site } n; \\ 0, & \text{Otherwise.} \end{cases}$$

$$e_{ij} = \begin{cases} 1, & \text{a connection exists between node } i \text{ and node } j; \\ 0, & \text{Otherwise.} \end{cases}$$

This paper formulates the DLP problem as an ILP to minimize the total placement cost.

Objective:

$$\text{minimize } C = C^B + C^E \sum_{e\in\mathbf{\Omega}_{ECP}} b_e + C^g \sum_{n\in\mathbf{\Omega}_{gNB}} c_n \tag{1}$$

The objective function (1) minimizes the total placement cost, where C^B, C^E and C^g are the deployment costs of BS, ECP and gNB, respectively.

Subject to:

$$e_{ij} = b_e,\ \forall i \in \mathbf{\Omega}_{BS},\ \forall j, e \in \mathbf{\Omega}_{ECP} \tag{2}$$

$$\sum_{j\in\mathbf{\Omega}_{gNB}} e_{ij} \geq b_e,\ \forall i, e \in \mathbf{\Omega}_{ECP} \tag{3}$$

$$\sum_{i\in\mathbf{\Omega}_{gNB}\cup\mathbf{\Omega}_{ECP}} e_{ij} = c_n,\ \forall j, n \in \mathbf{\Omega}_{gNB} \tag{4}$$

$$\sum_{j\in\mathbf{\Omega}_{gNB}\cup\mathbf{\Omega}_{TP}} e_{ij} \geq c_n,\ \forall i, n \in \mathbf{\Omega}_{gNB} \tag{5}$$

Constraints (2)–(5) aim to maintain the whole tree-based topology, and make sure that each node only has one root node, while could accommodate multiple child nodes at the same time. Constraints (2)–(3) denote that if an ECP exists, it must be connected with one BS, but it could connect to multiple gNBs. Constraints (4)–(5) denote that if a gNB_j exists, it must be connected with one ECP_i or gNB_i, but it could cover multiple TPs and gNBs.

$$\sum_{i\in\mathbf{\Omega}_{gNB}} e_{ij} \geq 1,\ \forall j \in \mathbf{\Omega}_{TP} \tag{6}$$

Constraint (6) ensures each TP node must be covered once at least.

$$e_{ij}d_{ij} \leq E_{com},\ \forall i \in \mathbf{\Omega}_{BS}, \forall j \in \mathbf{\Omega}_{ECP} \tag{7}$$

$$e_{ij}d_{ij} \leq G_{com},\ \ \forall i \in \mathbf{\Omega}_{ECP} \cup \mathbf{\Omega}_{gNB}, \forall j \in \mathbf{\Omega}_{gNB} \tag{8}$$

$$e_{ij}d_{ij} \leq G_{sen},\ \forall i \in \mathbf{\Omega}_{gNB}, \forall j \in \mathbf{\Omega}_{TP} \tag{9}$$

Constraints (7)–(9) represent the distance constraints, where G_{sen} represents the sensing range of gNB, and E_{com} and G_{com} represent the communication range of ECP, and gNB respectively. Constraint (7) denotes if an ECP node connects to a BS, the BS must be within ECP's communication range. Constraint (8) denotes if a gNB_jconnects to an ECP_i or a gNB_i, the ECP_i or gNB_i must be within gNB's communication range. Constraint (9) denotes if a TP connects to a gNB, the TP node must be within gNB's sensing range.

$$\sum_{j\in\mathbf{\Omega}_{ECP}} e_{ij} \leq K_1,\ \forall i \in \mathbf{\Omega}_{BS} \tag{10}$$

$$\sum_{j\in\mathbf{\Omega}_{gNB}} e_{ij} \leq K_2,\ \forall i \in \mathbf{\Omega}_{ECP} \tag{11}$$

$$\sum_{j\in\mathbf{\Omega}_{gNB}\cup\mathbf{\Omega}_{TP}} e_{ij} \leq K_3,\ \forall i \in \mathbf{\Omega}_{gNB} \tag{12}$$

Constraints (10)–(12) limit the maximum number of the child nodes that BS, ECP, and gNB could accommodate, respectively.

$$e_{ij} \leq b_e,\ \forall i \in \mathbf{\Omega}_{BS},\ \forall j, e \in \mathbf{\Omega}_{ECP} \tag{13}$$

$$e_{ij} \leq b_e,\ \forall i, e \in \mathbf{\Omega}_{ECP},\ \forall j \in \mathbf{\Omega}_{gNB} \tag{14}$$

$$e_{ij} \leq c_n,\ \forall i \in \mathbf{\Omega}_{gNB} \cup \mathbf{\Omega}_{ECP}, \forall j, n \in \mathbf{\Omega}_{gNB} \tag{15}$$

$$e_{ij} \leq c_n, \forall i, n \in \mathbf{\Omega}_{gNB}, \forall j \in \mathbf{\Omega}_{gNB} \cup \mathbf{\Omega}_{TP} \tag{16}$$

Constraints(13)–(16) limit the network connectivity to ensure the integrity of the network.

$$e_{ij}, b_e, c_n \in \{0, 1\}, \forall i, j \in \mathbf{\Omega}, \forall e \in \mathbf{\Omega}_{ECP}, \forall n \in \Omega_{gNB} \tag{17}$$

Constraint (17) represents that the value of each entry in e_{ij}, b_e and c_n is 0 or 1.

This section formulates the DLP problem as an ILP problem. Gurobi solver could solve ILP problem and the related simulation results are shown in Section VI. While Gurobi is adequate for the small-size scenarios, it is unsuitable to solve the large-size DLP problem. As the network size increases, the computation time of Gurobi grows dramatically, or it cannot even solve the problem. To this end, two heuristic algorithms, namely GA and CEHA, are proposed in Section V.

V. THE PROPOSED HEURISTIC ALGORITHMS

To improve the computation efficiency problem, GA is proposed. The proposed GA could reduce the computation time efficiently, but it is inclined to be trapped in the local optima. Then, the CEHA is proposed to solve the local optima issue. For simplicity, this section denotes child node and father node by CN and FN, respectively.

A. GA

A greedy algorithm is a paradigm that follows the thought of choosing the locally optimal solution. In this paper, the proposed GA aims at placing the FNs that could connect with more CNs to reduce the selected FNs to the least, under the constraints of the full coverage, the FNs' capacity limit, etc. The detailed procedure is listed in Algorithm 1.

GA aims to form the corresponding FN-**CN** pairs for each FN. FN-**CN** records the set of CNs that could be covered by FN and consists of gNB-**TP**, gNB-**gNB** and ECP-**gNB**. In lines 1–3, eachFN would form the corresponding FN-**CN** pair. In lines 4–8 GA would compare the FN-**CN** pairs of every two FNs, i.e., FN_i and FN_j. If the $\mathbf{CN}_{same}$ exists, where $\mathbf{CN}_{same}$ denotes the set of CNs that exist in both FN_i-**CN** and FN_j-**CN**, the node pair which has more CNs would retain $\mathbf{CN}_{same}$ while the other remove the $\mathbf{CN}_{same}$. In this way, we choose the FNs that could cover the most child nodes and reduce the selected FNs to the least, thus minimizing the total placement cost. In lines 9–11, the algorithm would judge that if $\mathbf{CN}_{left}$, which denotes the set of CNs thathave not been covered yet, exists or not. If $\mathbf{CN}_{left}$ exists, other FNs would be added and the selected FNs would be redistributed. Then in lines 12–14, wesearch for the FNs whose FN-**CN** is not

---

Algorithm 1: GA.

Input: The given topology and problem parameters: C^B, C^E, C^g, E_{com}, G_{com}, G_{sen}, K_1, K_2, K_3;
Output: The optimal placement solution (**S**) and the optimal deployment cost of VECN ($C_{\min}$).

1: **for** each FN **do**
2: form the corresponding FN-**CN** pairs;
3: **end for**
4: **for** every two FN-**CN** pairs **do**
5: **if** $\mathbf{CN}_{same}$ exists **then**
6: the node pair FN-**CN** which has more CNs retains $\mathbf{CN}_{same}$;
7: **end if**
8: **end for**
9: **while** $\mathbf{CN}_{left}$ exists **do**
10: redistribute the selected FNs;
11: **end while**
12: **for** the FN whose FN-**CN** is not null **do**
13: add the FN into **S**;
14: end for
15: Calculate $C_{\min}$ based on **S**;
16: **Output: S** and $C_{\min}$;

---

null and add them to placement solution set **S**. Finally, in lines 15–16, **S** is generated and the minimum cost $C_{\min}$ could be calculated based on **S**.

B. CEHA

The proposed GA, which follows the thought of reducing the FNs to the least, could obtain the locally optimal solution efficiently. However, it is inclined to be trapped in the local optima. To avoid GA's local optima, a GEHA is proposed in this section to tradeoff between the local optimal and global optimal.

The core idea is by introducing the probability matrix P_{ij} to guide the CN_i to connect to FN_j whose P_{ij} is larger, while improving the randomness (that is, CN_i does not have to connect FN_{lo}, which has the largest value of P_{ij}) . The probability coefficient matrix R_{ij} is calculated by :

$$R_{ij} = \frac{\lambda_1}{D_{ij}} + \lambda_2 * N_j \tag{18}$$

where D_{ij} denotes the distance between CN_i and FN_j, and N_j denotes the number of the CNs that could be covered by FN_j. $\frac{\lambda_1}{D_{ij}}$ favors the FN that is closer to CN_i, and $\lambda_2 * N_j$ prefers the FN that could connect with more CNs, where λ_1 and λ_2 are the scaling factors to balance R_{ij}. R_{ij} would subsequently determine P_{ij} by:

$$P_{ij} = \frac{R_{ij}}{\Sigma_{j=1}^{N} R_{ij}} \tag{19}$$

Algorithm 2: **CEHA**.

Input: The given topology, problem parameters: C^B, C^E, C^g, E_{com}, G_{com}, G_{sen}, K_1, K_2, K_3 and algorithm parameters: λ_1, λ_2, A, h, GENERATION;
Output: The optimal placement solution (**S**), the optimal deployment cost ($C_{\min}$).

// **Phase I: Initialization**
1: Initialize R_{ij} and P_{ij};
2: **while** gen < GENERATION **do**
3: Reset **C**, R_{ij} and P_{ij};
//**Phase II: FNs allocation**
4: **for** each CN **do**
5: **if** rand(0,1) < A **then**
6: Add FN_{lo} into **C**;
7: **else**
8: **for** each FN **do**
9: **if** $d < P_{ij}$ **then**
10: Add FN_j into **C** and break the loop;
11: **end if**
12: **end for**
13: **end if**
14: Update R_{ij} and P_{ij};
15: **end for**
16: Calculate C_{temp} based on **C**;
17: **if** $C_{temp} < C_{\min}$ **then**
18: $\mathbf{S} \leftarrow \mathbf{C}$, $C_{\min} \leftarrow C_{temp}$ and gen $\leftarrow 0$;
19: **end if**
20: gen++;
21: **end while**
22: **Output: S** and $C_{\min}$;

---

where P_{ij} denotes the probability that CN_i chooses to connect to FN_j and N is the number of the candidate FNs. Each CN would randomly generate a decimal d. If $d < P_{ij}$, CN_i would connect to FN_j, otherwise CN_i would connect to other FNs. Let FN_{lo} denote the candidate FN that best fits the thought of local optimal. Though FN_{lo} has the biggest probability to be selected, there is a certain probability for CNs to connect to other FNs so as to improve the randomness to avoid the local optima. The proposed CEHA mainly involves two phases: (1) initialization, (2) FNs allocation. The detailed procedure is listed in Algorithm 2.

Phase I is to initialize R_{ij} and P_{ij}. The loop terminates only when gen reaches the iteration times GENERATION, where gen records the iterations. In line 3, each iteration would reset **C**, R_{ij} and P_{ij}, where **C** denotes the set of the chosen FNs.

Phase II is to search FN for each CN. If rand(0, 1) < A, where A is a decimal, we would directly add FN_{lo} into **C** (that is, directly connect CN_i to the local optimal solution FN_{lo}) in line 6. The larger A means the easier to be trapped in the local optima. However, the iterations would be reduced, thus improving the efficiency, and vice versa. If rand(0, 1) >= A, in lines 8–12, CN_i would randomly generate a decimal d for each FN (i.e., FN_j) and the algorithm would compare the value of d and P_{ij}. If $d < P_{ij}$, then FN_j would be added into **C**, otherwise another d would be generated, and the algorithm would judge if CN could connect to FN_{j+1} (that is, to judge if $d < P_{i(j+1)}$).

Once CN_i chooses to connect to FN_j, we would update the R_{kj} in line 14 by:

$$R_{kj} = R_{kj} + h \tag{20}$$

where h denotes CN_k that could communicate with FN_j (i.e., $R_{kj}! = 0$) and h is a constant. The increment of R_{kj} would subsequently guide the CNs to connect the selected FN_j. The iterations in lines 4–15 would be conducted until all the CNs complete their selection.

C_{temp} would be calculated based on **C** in line 16. If C_{temp} is smaller than $C_{\min}$, that is, **C** is superior to **S**, then the algorithm would update the **S** as **C**, $C_{\min}$ as C_{temp} and set gen = 0, in line 18.

VI. SIMULATIONS AND NUMERICAL RESULTS

In this section, the Gurobi solver is applied to obtain the optimal layout solutions. Simulations are conducted to verify the feasibility of the DLP optimization formulation. Gurobi is adequate for the small-size scenarios, however, as the network size increases, the computation time of Gurobi grows dramatically, or the problem cannot be even solved. Then the simulations are conducted to demonstrate the scalability of the DLP optimization formulation and verifythat the proposed heuristic algorithms could solve the DLP problem efficiently as the network size increases dramatically. Many scenarios are investigated, and several of them are adopted to illustrate the simulations. Specifically, the scenarios take diverse road types and the unbalanced traffic flow into consideration to verify the practicality of the VECN.

A. SIMULATION SETTING

The experimental parameters are listed as the Table 2. This paper estimates the whole placement cost with a generic cost unit (gcu) [28], [29]. C^B, C^E and C^g indicate the price of the BS, ECP and gNB, and they are 300 gcu, 150 gcu and 100 gcu, respectively. The Manhattan distance is used to indicate the corresponding edge distance between each pair of nodes. We take the communication range of the ECP and gNB (i.e., E_{com} and G_{com}) as 900 m and 600 m, respectively. G_{sen} indicates the sensing range of the gNB and is set to 300 m. The maximum number of the child nodes that the BS, ECP and gNB can accommodate (i.e., K_1, K_2 and K_3) is 4, 5 and 7 respectively. The CEGA parameters λ_1, λ_2, A, h and GENERATION are 0.25, 0.02, 0.7, 1 and 1000, respectively.

B. SIMULATION RESULTS OF GUROBI

Gurobi solver is applied to derive the optimal layout solutions. A series of simulations in the different scenarios are conducted, and several of them are shown in Fig. 3. For simplicity, the VECN is formulated into a 2-dimensional grid. The unit distance is set as 10 m, the grid is digitized into 10 ∗ 10, and the horizontal and vertical coordinates denote the normalized distance. The scenarios take diverse road types into consideration, and Scenario 1, 2, 3 and 4 represent the one-way, forked, crossroad and complex scenarios, respectively.

The selected devices are marked as solid and the hollow ones are unselected. We can observe that the two nodes, which could communicate with each other, are connected. The red, pink and blue lines denote the connections between BS and ECP, ECP and gNB, gNB and gNB, respectively. The TPs within the circles are covered by the corresponding gNB.

TABLE 2. Experimental Parameters

Parameter	Value
C^B	300
C^E	150
C^g	100
E_{com}	1000m
G_{com}	600m
G_{sen}	300m
K_1	4
K_2	5
K_3	7
λ_1	0.25
λ_2	0.02
A	0.7
h	1
GENERATION	1000

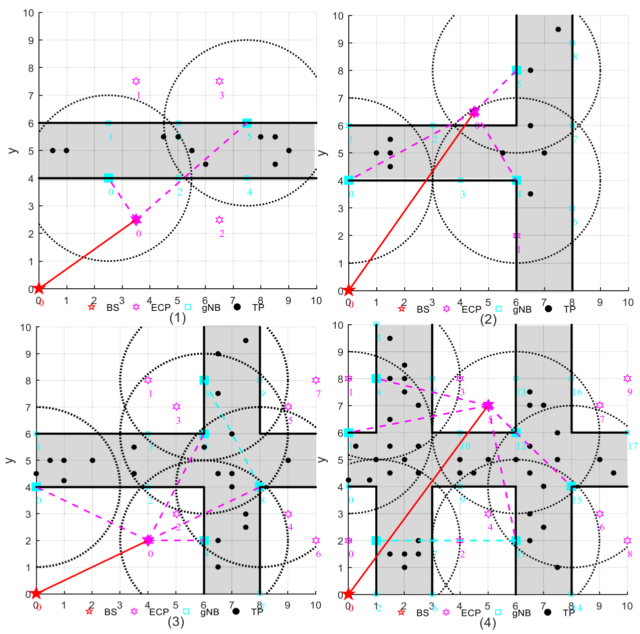

FIGURE 3. The layout solution of DLP problem with Gurobi.

TABLE 3. The Experimental Results of Gurobi

Scenario	The Number of BS, ECP, gNB, TP	The number of selected BS, ECP, gNB	Placement cost (gcu)
1	1,4,6,10	1,1,2	650
2	1,2,9,10	1,1,3	750
3	1,3,8,20	1,1,5	950
4	1,10,18,40	1,1,6	1050
5	1,15,30,70	--	--
6	1,20,45,100	--	--

The layout solutions describedin Fig. 5 show that the DLP problem could be solved and the tree-based topology is maintained with the BS serving as the root node, ECP and gNB as the internal nodes, and TPs as the leaf nodes. The results have verified that the DLP optimization is feasible. The detailed experimental results are shown in the Table 3. The number of BS, ECP and gNB that Gurobi selects is respectively 1-1-2 in Scenario 1, 1-1-3 in Scenario 2, 1-1-5 in Scenario 3, and 1-1-6 in Scenario 4. The placement cost for each scenario is 650 gcu, 750 gcu, 950 gcu, 1050 gcu.

Obviously, we could observe the feasibility of DLP optimization formulation. However,it is worth mentioning that Gurobi cannot solve Scenario 5 and Scenario 6. Gurobi is unsuitable when the network size increases dramatically.

C. SIMULATION RESULTS OF THE PROPOSED ALGORITHMS

Numerical simulations are conducted to demonstrate the scalability of the proposed DLP optimization formulation and the practicability of the proposed algorithms. In this section, several scenarios are illustrated as example. The results of proposed GA are shown in Fig. 4. The results of proposed CEHA are shown in Fig. 5.

The detailed experimental results are shown in the Table 4 and Table 5. The number of BS, ECP and gNB the GA selects is respectively 1-1-2 in Scenario 1, 1-1-3 in Scenario 2, 1-1-5 in Scenario 3, 1-2-8 in Scenario 4, 1-2-14 in Scenario 5 and 1-4-18 in Scenario 6. The placement cost for each scenario is 650 gcu, 750 gcu, 950 gcu, 1400 gcu, 2000 gcu and 2700 gcu, respectively. The number of BS, ECP, gNB the CEHA selects is respectively 1-1-2 in Scenario 1, 1-1-3 in Scenario 2, 1-1-5 in Scenario 3, 1-1-7 in Scenario 4, 1-2-11 in Scenario 5, 1-2-16 in Scenario 6. The placement cost for each scenario is 650 gcu, 750 gcu, 950 gcu, 1150 gcu, 1700 gcu and 2200 gcu, respectively.

Obviously, we could observe that the DLP problem, for each scenario, could besolved. The simulation results reveal the feasibility and scalability of DLP optimization. DLP optimization formulation is practical and could provide guidance for real-world deployment. The results also verify that, for each Scenario, both the proposed GA and CEHA could solve the DLP problem.

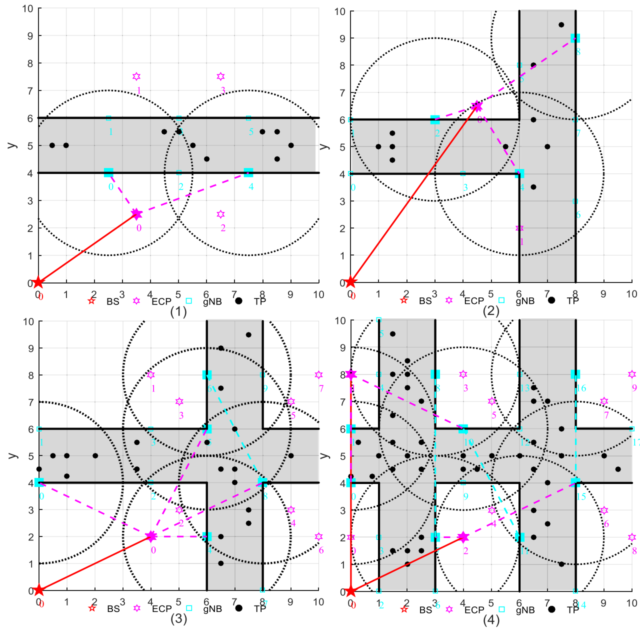

FIGURE 4. The layout solution of DLP problem with GA.

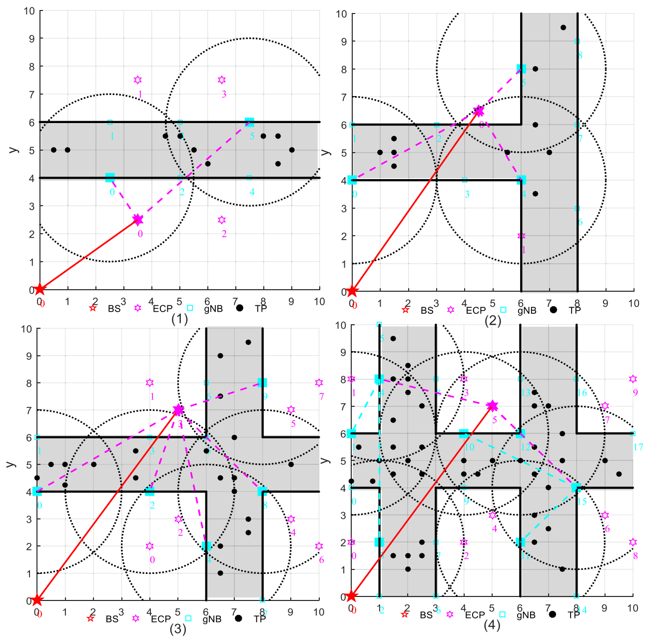

FIGURE 5. The layout solution of DLP problem with CEHA.

D. ANALYSIS

It is observed in Fig. 6 that the total cost of all the three has an upward trend as the network size increases, and they cost the same in Scenario 1, Scenario 2, Scenario 3. However, the expansion of TPs gradually reveals GA's defeats. GA skyrockets and is the highest of the three in Scenario 4, Scenario 5 and Scenario 6. CEHA costs slightly higher than Gurobi in Scenario 4, however, from Scenario 5 on, CEHA continues

TABLE 4. The Experimental Results of GA

Scenario	The Number of BS, ECP, gNB, TP	The number of selected BS, ECP, gNB	Placement cost (gcu)
1	1,4,6,10	1,1,2	650
2	1,2,9,10	1,1,3	750
3	1,3,8,20	1,1,5	950
4	1,10,18,40	1,2,8	1400
5	1,15,30,70	1,2,14	2000
6	1,20,45,100	1,4,18	2700

TABLE 5. The Experimental Results of CEHA

Scenario	The Number of BS, ECP, gNB, TP	The number of selected BS, ECP, gNB	Placement cost (gcu)
1	1,4,6,10	1,1,2	650
2	1,2,9,10	1,1,3	750
3	1,3,8,20	1,1,5	950
4	1,10,18,40	1,1,7	1150
5	1,15,30,70	1,2,11	1700
6	1,20,45,100	1,2,16	2200

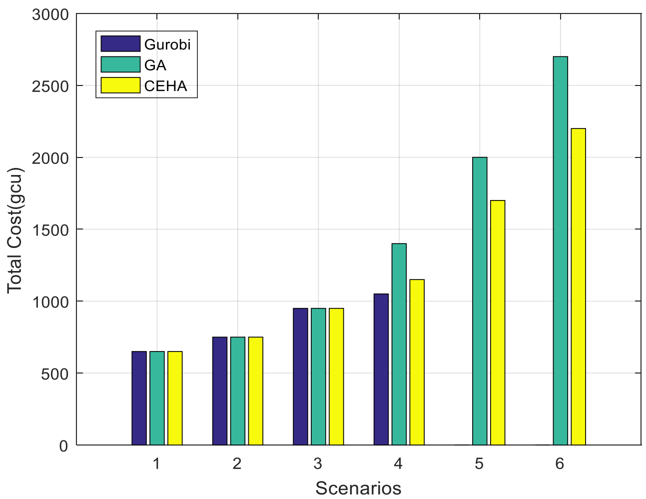

FIGURE 6. Total cost comparison.

to perform the best whereas Gurobi cannot solve the problem anymore.

Table 6 reveals that GA has an absolute superiority in computation time. In the small-size network, Gurobi performs slightly better than CEHA and there is not a great deal of difference between Gurobi and GA. However, in Scenario 4, Gurobi's computation time increases to 1132.58 s due to the network size's expansion, while GA and CEHA consume 0.188 s and 5.569 s, respectively. Though CEHA indeed takes longer time than GA, ithas already improved the computation time to 200 times shorter than Gurobi in Scenario 4.

TABLE 6. The Computation Time Results Comparison

Scenario	Gurobi	GA	CEHA
1	0.101 s	0.138 s	0.468 s
2	0.135 s	0.14 s	1.007 s
3	2.38 s	0.141 s	1.098 s
4	1132.58 s	0.188 s	5.569 s
5	--	0.21 s	27.605 s
6	--	0.257 s	34.863 s

Based on the simulation results, we conclude that Gurobi outperforms the proposed heuristic algorithms, and the proposed CEHA is slightly higher than Gurobi while the proposed GA performs the worst, in terms of placement cost. However, the computation time Gurobi takes grows rapidly or Gurobi cannot solve the problem, as the network size increases dramatically. Note that the GA outperforms the others in terms of computation time, but the CEHA is just slightly higher than CEHA andhas already improved the computation time to 200 times shorter than Gurobi in Scenario 4. The CEHA could not only consumeless computation time than Gurobi, but also help avoid GA's local optima. In conclusion, the proposed CEHA is more efficiently than Gurobi and more effectively than GA, and it is suitable to solve the DLP problem.

VII. CONCLUSION

This paper formulates the DLP problem of the VECN, which integrates gNB and ECP, to minimize the placement cost. Placing the ECP could improve fast-response and high-bandwidth communications, and ease the burden of the cloud. The gNB could support a stable connectivity in a high-speed scenario and facilitate the fast response. The results of Gurobi demonstrate the feasibility of DLP optimization framework. The GA is proposed to improve Gurobi's computation efficiency problem, but it is inclined to be trapped in the local optima. Then CEHA is proposed to tradeoff between the local optimal and global optimal. The results have verified the feasibility and scalability of the DLP formulation and shows that the proposed CEHA outperforms Gurobi and the proposed GA to solve the DLP problem.

REFERENCES

[1] W. Zhang, B. Lin, C. Gao, Q. Yan, S. Li, and W. Li, "Optimal placement in RFID-integrated VANETs for intelligent transportation system," in *Proc. IEEE Int. Conf. RFID Technol. Appl.*, 2018, pp. 1–6.
[2] A. A. Khan, M. Abolhasan, and W. Ni, "5G next generation VANETs using SDN and fog computing framework," in *Proc. 15th IEEE Annu. Consum. Commun. Netw. Conf.*, 2018, pp. 1–6.
[3] Y. Xie, I. W. Ho, and E. R. Magsino, "The modeling and cross-layer optimization of 802.11p VANET unicast," *IEEE Access*, vol. 6, pp. 171–186, 2018.
[4] F. Qu, Z. Wu, F. Wang, and W. Cho, "A security and privacy review of VANETs," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 2985–2996, Dec. 2015.

[5] G. Luo *et al.*, "Cooperative vehicular content distribution in edge computing assisted 5G-VANET," *China Commun.*, vol. 15, no. 7, pp. 1–17, Jul. 2018.

[6] C. Yu, B. Lin, P. Guo, W. Zhang, S. Li, and R. He, "Deployment and dimensioning of fog computing-based Internet of Vehicle infrastructure for autonomous driving," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 149–160, Feb. 2019.

[7] H. Zhong, L. Pan, Q. Zhang, and J. Cui, "A new message authentication scheme for multiple devices in intelligent connected vehicles based on edge computing," *IEEE Access*, vol. 7, pp. 108211–108222, 2019.

[8] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5G network edge cloud architecture and orchestration," *IEEE Commun. Surveys Tut.*, vol. 19, no. 3, pp. 1657–1681, Jul.–Sep. 2017.

[9] G. Luo *et al.*, "Software defined cooperative data sharing in edge computing assisted 5G-VANET," *IEEE Trans. Mobile Comput*, to be published, doi: 10.1109/TMC.2019.2953163.

[10] D. Kombate and Wanglina, "The Internet of Vehicles based on 5G communications," in *Proc. IEEE Int. Conf. Internet Things (iThings) IEEE Green Comput. Commun. (GreenCom) IEEE Cyber, Physical Social Comput. (CPSCom) IEEE Smart Data (SmartData)*, 2016, pp. 445–448.

[11] Y. Sun, L. Xu, Y. Tang, and W. Zhuang, "Traffic offloading for online video service in vehicular networks: A cooperative approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7630–7642, Aug. 2018.

[12] B. Liu, D. Jia, J. Wang, K. Lu, and L. Wu, "Cloud-assisted safety message dissemination in VANET–cellular heterogeneous wireless network," *IEEE Syst. J.*, vol. 11, no. 1, pp. 128–139, Mar. 2017.

[13] X. Cheng, C. Chen, W. Zhang, and Y. Yang, "5G-enabled cooperative intelligent vehicular (5GenCIV) framework: When benz meets marconi," *IEEE Intell. Syst.*, vol. 32, no. 3, pp. 53–59, May/Jun. 2017.

[14] C. Zhu, G. Pastor, Y. Xiao, Y. Li, and A. Ylae-Jaeaeski, "Fog following me: Latency and quality balanced task allocation in vehicular fog computing," in *Proc. 15th Annu. IEEE Int. Conf. Sens., Commun., Netw.*, 2018, pp. 1–9.

[15] X. Xu, Y. Xue, X. Li, L. Qi, and S. Wan, "A computation offloading method for edge computing with vehicle-to-everything," *IEEE Access*, vol. 7, pp. 131068–131077, 2019.

[16] S. Zhang, P. He, K. Suto, P. Yang, L. Zhao, and X. Shen, "Cooperative edge caching in user-centric clustered mobile networks," *IEEE Trans. Mobile Comput.*, vol. 17, no. 8, pp. 1791–1805, 1 Aug. 2018.

[17] Z. Zhou, B. Wang, M. Dong, and K. Ota, "Secure and efficient vehicle-to-grid energy trading in cyber physical systems: Integration of blockchain and edge computing," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 50, no. 1, pp. 43–57, Jan. 2020.

[18] T. Wu, W. Liao, and C. Chang, "A cost-effective strategy for road-side unit placement in vehicular networks," *IEEE Trans. Commun.*, vol. 60, no. 8, pp. 2295–2303, Aug. 2012.

[19] P. Jiang *et al.*, "Roadside units placement for traffic flows coverage requirement in vehicular networks," in *Proc. IEEE Int. Conf. Smart Cloud (SmartCloud)*, 2018, pp. 145–152.

[20] R. Zhang, F. Yan, W. Xia, S. Xing, Y. Wu, and L. Shen, "An optimal roadside unit placement method for VANET localization," in *Proc. GLOBECOM 2017 - IEEE Global Commun. Conf.*, 2017, pp. 1–6.

[21] S. Mehar, S. M. Senouci, A. Kies, and M. M. Zoulikha, "An optimized roadside units (RSU) placement for delay-sensitive applications in vehicular networks," in *Proc. 12th Annu. IEEE Consum. Commun. Netw. Conf.*, 2015, pp. 121–127.

[22] K. A. Hafeez, L. Zhao, B. Ma, and J. W. Mark, "Performance analysis and enhancement of the DSRC for VANET's safety applications," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3069–3083, Sep. 2013.

[23] K. A. Hafeez, L. Zhao, Z. Liao, and B. N. Ma, "Impact of mobility on VANETs' safety applications," in *Proc. IEEE Global Telecommun. Conf. GLOBECOM 2010*, 2010, pp. 1–5.

[24] H. Wang, L. Wan, M. Dong, K. Ota, and X. Wang, "Assistant vehicle localization based on three collaborative base stations via SBL-based robust DOA estimation," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5766–5777, Jun. 2019.

[25] Y. Bi, H. Shan, X. S. Shen, N. Wang, and H. Zhao, "A multi-hop broadcast protocol for emergency message dissemination in urban vehicular Ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 3, pp. 736–750, Mar. 2016.

[26] Y. Bi, G. Han, C. Lin, X. Wang, Q. Zhang, and Z. Liu, "Effective packet loss elimination in IP mobility support for vehicular networks," *IEEE Netw.*, vol. 34, no. 1, pp. 152–158, Jan./Feb. 2020.

[27] H. Li, K. Ota, and M. Dong, "ECCN: Orchestration of edge-centric computing and content-centric networking in the 5G radio access network," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 88–93, Jun. 2018.

[28] B. Lin, X. Pan, R. He, and S. Li, "Joint wireless-optical infrastructure deployment and layout planning for cloud-radio access networks," in *Proc. Int. Wireless Commun. Mobile Comput. Conf.*, 2014, pp. 1027–1032

[29] P. Guo, B. Lin, X. Li, R. He, and S. Li, "Optimal deployment and dimensioning of fog computing supported vehicular network," in *Proc. IEEE Trustcom/Big Data SE/ISPA*, 2016, pp. 2058–2062.