The following publication E. R. Magsino and I. W. -H. Ho, "An Enhanced Information Sharing Roadside Unit Allocation Scheme for Vehicular Networks," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 9, pp. 15462-15475, Sept. 2022 is available at https://dx.doi.org/10.1109/TITS.2022.3140801.

An Enhanced Information Sharing Roadside Unit Allocation Scheme for Vehicular Networks

Elmer R. Magsino, *Member, IEEE* and Ivan Wang-Hei Ho, *Senior Member, IEEE*

***Abstract*—Sharing up-to-date environment information collected by intelligent connected vehicles is critical in achieving travel comfort, convenience, and safety in vehicular networks. Individually collected information should be made available to other vehicular nodes, adjacent or distant, to achieve an informed and well-managed vehicular traffic. The coverage reach of sharing these road data can be maximized by allocating roadside units in strategic positions. In this work, we propose an Enhanced Information SHAring via Roadside Unit Allocation (EISHA–RSU) scheme that strategically determines where RSUs must be deployed from all spatial candidate locations. The urban area is irregularly partitioned into effective regions of movement (ERM) according to vehicular capacity with priority. For each ERM, EISHA–RSU greedily allocates the initial RSU to an effective position and optimally assigns the remaining RSUs to spatial locations that capture the maximum I2V/V2I information sharing based on the area's average road speed. In effect, the proposed deployment scheme addresses both the issues of coverage and connectivity among vehicles and the infrastructure. We evaluate the proposed RSU allocation scheme by employing three urban empirical mobility datasets and compare its network starvation fairness, effectiveness, and efficiency performance measures with three other deployment benchmarks. Overall, EISHA–RSU reduces the number of required RSUs to cover a certain area, exhibits higher connectivity, and achieves maximum I2V/V2I information sharing among the evaluated schemes.**

***Index Terms*—Roadside Unit Allocation, Information Sharing, Vehicular Networks, Spatiotemporal Coverage, Efficiency and Effectiveness, Starvation and Fairness**

I. INTRODUCTION

An urban transportation network is comprised mainly of public and private vehicles, as well as infrastructure such as buildings, roads, highways, and roadside facilities. With the support of vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications, these vehicles and infrastructure can communicate and exchange important environment data and control messages to provide comfortable, time-efficient, safe, environment- and energy-friendly travel. This interconnection among vehicles and infrastructure has transformed the transport network into an intelligent transportation system (ITS). Years from now, intelligent connected vehicles, as well as autonomous ones, are envisioned to be found more on the urban and rural roads, distributed over a region, and driving with other regular transportation modes [1]. As the number of vehicles plying the road and prevalence of useful road data increases, the need to enhance the information sharing in the vehicular network becomes significant.

Widely distributed in the vehicular network, intelligent vehicles sense their current surroundings, process the raw road data to provide a set of beneficial and dynamic collections of environment information, and then store them in their memory for a limited amount of time. However, these pieces of environment information are only useful when shared among vehicles, pedestrians, road users, and policy makers, which the data dissemination problem in vehicular networks tries to solve. Various concerns and issues have limited the efficient information sharing between/among vehicles and infrastructures, e.g., limited wireless bandwidth and intermittent connectivity due to the high mobility of vehicular nodes.

One solution to address these issues is to strategically deploy static roadside units (RSUs) to provide reliable connectivity and wider coverage for time-critical data download and upload, while considering deployment and maintenance costs. Intelligent vehicles can upload their relevant road information to the RSUs they encounter along the way and can also download readily available data that they can utilize in their trips. For RSUs, aside from receiving, processing, and storing vehicular uploaded information, they can also monitor their surroundings and allow data dissemination and exchanges between nearby passing vehicles. These information upload and download must happen in real-time while the data being exchanged are still valid. To support V2I and V2V communications in achieving enhanced information sharing, the RSU deployment problem must consider the structured road network topology, sensed dynamic environment data locations, and vehicular mobility distributions. Given these considerations, candidate or optimal RSU locations on the streets are based at intersections [2] or other locales which satisfy a set of criteria or constraints [3]. In [4], the infrastructures are deployed on intersections based on the amount of allowable data delivery delay.

One important application of information sharing can be seen in route planning. Given an urban map divided into grids with a deployed RSU, a decentralized route planning based on distributed routing tasks was investigated [5]. Latency was reduced as more grids were covered, but at the

Manuscript received ; revised . The work of I. W.-H. Ho was supported in part by the General Research Fund (Project No. 15201118) established under the University Grant Committee (UGC) of the Hong Kong Special Administrative Region (HKSAR), China. The work of E. R. Magsino was supported in part by the Research Impact Fund (Project No. R5007-18) established under the University Grant Committee (UGC) of the Hong Kong Special Administrative Region (HKSAR), China.

I.W.-H. Ho is with the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong (email:ivanwh.ho@polyu.edu.hk).

E. R. Magsino was with the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong and is currently affiliated with the Department of Electronics and Computer Engineering, Gokongwei College of Engineering, De La Salle University, Manila, Philippines (email:elmer.magsino@dlsu.edu.ph).

expense of route accuracy. Also, as the number of partitions increased, more RSUs were needed to be deployed. In [6], the route planning through information sharing was improved by adding infrastructures such as vehicular traffic servers and base stations. RSU deployment was also done on intersections, but simulations involved only a small-scale setup. Perhaps, the most common use-case setup of information sharing in vehicular networks can be seen in emergency and safety scenarios [7], [8]. Both works approached the allocation scheme from the distance perspective to minimize the delay caused by traveling, thereby, promoting quick transmission of emergency warnings. From these examples, it is imperative that road data must be shared in real-time, as the information sharing in a vehicular network can influence the behavior of traffic flow, emergency response, and driving decisions. This can be addressed by RSUs that have data upload and download capabilities. The majority of upload comes from nearby intelligent vehicles, while the optimally deployed RSUs should be able to enhance information sharing in an optimal manner [9].

This study targets to maximize the spatiotemporal information coverage in vehicular networks. The huge amount of static (e.g., building and road topology) and dynamic (e.g., pedestrians, accidents, and road routing changes) road information must be collected and processed in real-time to aid vehicles and the transport system in performing up-to-date and appropriate decisions. This can be addressed by allowing intelligent vehicles and roadside units to capture road data and stay connected to enable enhanced information exchanges between them. However, a strategic method must be developed to optimally locate the effective positions where stationary roadside units should be deployed.

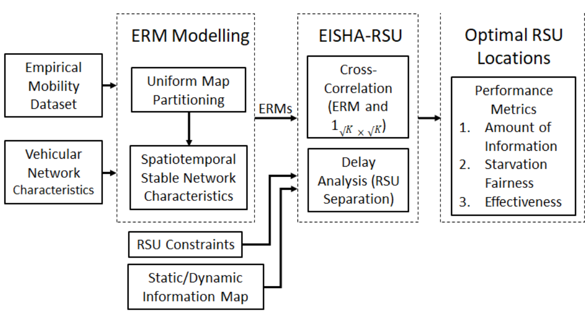

Fig. 1. Proposed block diagram of deriving the optimal RSU locations in an urban map utilizing the EISHA-RSU scheme from empirical mobility dataset, network characteristics, and constraints.

To accomplish this, we propose an Enhanced Information SHAring via RSU (EISHA-RSU) allocation scheme that will tackle the on-time collection and dissemination of road data by strategically positioning RSUs in an urban scenario while at the same time, ensuring the stable connectivity among vehicles and the infrastructure in the network. The framework of our proposed methodology to permit enhanced information exchanges between vehicles and RSUs is depicted in Fig. 1. Since empirical mobility datasets are varying both in space and time, we initially partition the urban map uniformly and compute for its spatiotemporal stable network characteristics. From this step, we can now identify the major contributions of the research work as summarized below.

1) We introduce the concept of Effective Regions of Movement (ERMs) to irregularly partition an urban city based on an area's vehicular capacity threshold. (We note that other vehicular network characteristics such as velocity and density profiles can be used to form ERMs.) ERMs automatically determine the priority of different urban locations and their corresponding coverage areas.
2) Based on ERM, we propose the novel vehicular-mobility-aware EISHA-RSU deployment scheme to capture and share the environment data in the vehicular network. For each formed ERM, effective positions (EPs) are located based on the amount of information of an area and the average road speed between EPs to ensure urban-wide connectivity and wider coverage. Since ERMs are two-dimensional areas, all physical locations are considered as possible EPs in the RSU allocation.
3) Extensive simulation utilizing three urban empirical mobility traces and locations, with different dataset characteristics, is carried out to evaluate EISHA-RSU's efficiency. EISHA-RSU fairly allocates RSUs at EPs to achieve the problem objectives via identifying locations with maximized information-rich sources and data carriers. By comparing EISHA-RSU with three other benchmarks, our proposed allocation scheme saves on average 21%, 21%, and 101% less number of RSUs, respectively, while satisfying the problem objectives. Also, based on the *Effectiveness* and *Network Starvation* metrics, EISHA-RSU performs the best when concerning coverge area and the amount of information shared.

The paper is outlined as follows. Section II provides a concise summary of related works to the RSU deployment problem. Section III discusses the definitions, assumptions, and setup considered to solve the allocation problem. Section IV presents the novel EISHA-RSU algorithm that utilizes the concepts of ERMs and EPs. Section V discusses the results derived from extensive simulation employing three urban empirical mobility traces. The summary, conclusion, and future research directions are given in Section VI.

II. Background and Related Literature

Previous studies dealt with the deployment problem of roadside units (RSUs) according to three general objectives, namely, (i) network connectivity or temporal coverage; (ii) traffic or spatial coverage; and (iii) information dissemination or spatiotemporal coverage [10].

Network connectivity or temporal coverage research focuses on attaining stable wireless connections among vehicles and infrastructure by minimizing V2V and V2I disconnections. On the other hand, traffic or spatial coverage studies center on maximizing V2I contacts in a region where an RSU is deployed. Lastly, information dissemination or spatiotemporal coverage literature engages in improving data exchange in a vehicular network by minimizing delay and packet loss, while maximizing throughput.

The research works [11]–[13] tackled the network connectivity objective. A greedy deployment scheme was proposed in [11] for improving the connectivity of the network by formulating it as a mobility clustering problem, where the clusters formed from the union of nearby mobility traces determine the RSU locations. In [12], a two-step solution was developed to determine the road segment's midpoint location where the RSU is installed to enhance connectivity when the vehicular density is low. A RSU scheduling algorithm was presented in [13] where the authors assumed that all RSUs were already deployed for maximum coverage. Their scheduler aimed at minimizing energy consumption through selecting the set of operating RSUs dynamically while maintaining network connectivity. Different from [13], a crowdsensing middleware for vehicular networks was proposed to opportunistically identify and localize RSUs along a route. It enabled vehicles to understand the network topology by allowing them to connect to other RSU candidates that are not congested or have better connectivity [14]. In [15], generic vehicles can localize the nearest available RSU to connect with by using compressive sensing and received signal strength.

The studies in [16]–[18] investigated the traffic coverage issue. GeoCover [16] initially extracted the road's geometry and hotspot locations from the mobility patterns. Given these, GeoCover then employed genetic and greedy algorithms to determine the locations where RSUs should be placed to achieve maximum coverage. In [17], the authors maximized the traffic coverage by considering the number of distinct vehicles having at least one V2I contact. They used the migration ratios between adjacent cells to locate candidate RSU locations. Finally, [18] maximized the number of covered cars while achieving longer connection time for an improved quality of service.

The papers [8], [19]–[21] analyzed the information dissemination problem. [19] considered the delivery delay in the deployment of RSUs while considering bi-directional movements of vehicles and RSU capacity. [8] and [21] considered the effect of road elements on radio signal and the minimal delay constraints for information spreading, respectively. Lastly, [20] presented that downloading up-to-date road data was one of the major responsibilities of having RSUs deployed strategically in the network.

It is also inevitable that some of these deployment objectives have been studied altogether. The work in [22] proposed a maximum coverage scheme based on the number of V2I connections. In [23], a generalized profit function was used to determine the locations of RSUs in a one-dimensional road setup, while [24] maximized coverage, the number of vehicular contacts and contact times for information dissemination in vehicular networks.

We differentiate our work from these previous studies in three distinct ways. Firstly, while most works have already identified intersections as candidate positions for assigning RSUs, we extend the search for candidate locations to the whole considered area that may include landmarks and information-rich places. [16] is a close related work that implements a modified density-based spatial clustering of applications with noise (DBSCAN) clustering algorithm to determine the hotspots of an urban city but heavily relies on the GPS traces of vehicles. If DBSCAN was applied to the JKT and SIN traces, there will be regions that are left out due to insufficient traces from the available datasets. When using DBSCAN, the road structure of an urban city is automatically conceived because these are the places where you can find the vehicular traces. In our proposed EISHA-RSU deployment scheme, determination of a possible deployment location has encompassed the work presented in [16] and has been modified/extended so that we can use other network characteristics to define region homogeneity such as speed, density, volume, wireless performance metrics, etc. In [17] and [18], any spatial location was also considered as candidate RSU position but only focused on one dimensional roads and employed migration patterns, respectively.

Secondly, we irregularly group uniform area grids based on the mobility patterns of vehicles and form homogeneous ERMs based on the vehicular capacity. This ERM-based model automatically discriminates information-rich locations. The works in [11], [12], [18], [25]–[27] utilized uniform partitioning to model and represent road networks. From these equal grids, network graphs were developed to represent intersections, important locations, and landmarks as vertices while road segments characterizing the map were denoted as edges. Others proceeded with defining their optimization problems to minimize the number of RSUs needed while satisfying a set of constraints. Our work belongs to few research studies that utilize non-uniform partitioning, e.g., [16].

Finally, to select the effective positions where RSUs will be deployed, we formulate the optimization problem by maximizing the information shared under the constraints of on-time data delivery, and maintaining the connectivity presented by the unstructured transport network.

III. Effective Regions of Movement

In this section, we discuss the irregular partitioning of an urban map according to its vehicular distribution, thus, automatically removing unnecessary city details and optimally form homogeneous effective regions of movement (ERMs).

A. *Spatiotemporal Stable Network Characteristics*

Consider the uniform partitioning of an urban area under study in Fig. 2 into $N \times N$ map grids, $g_{p,q}$. Fig. 2 is a general assumption of a grid-type urban city that has intersections with four road segments leading to it such as Beijing [28], Singapore [29], and Jakarta [30]. Each map grid, $g_{p,q}$, is characterized by its utility function, $\zeta_{p,q}$, at time t and depends on the map grid's longitude and latitude location. The indexes p and $q \in \{1, \ldots N\}$. We define $\zeta_{p,q}$ as:

$$\zeta_{p,q} = \mathbf{E}\big[\eta_{p,q}\big],$$

where $\mathbf{E}\big[\bullet\big]$ is the expectation of $\big[\bullet\big]$. $\eta_{p,q}$ describes the map grid's current spatial network characteristic, such as the grid's dynamic data, $\delta_{p,q}$, static data, $\gamma_{p,q}$, vehicular capacity, $c_{p,q}$, vehicular density, connectivity, accident rate, etc. Each network characteristic is assumed to be independent from each

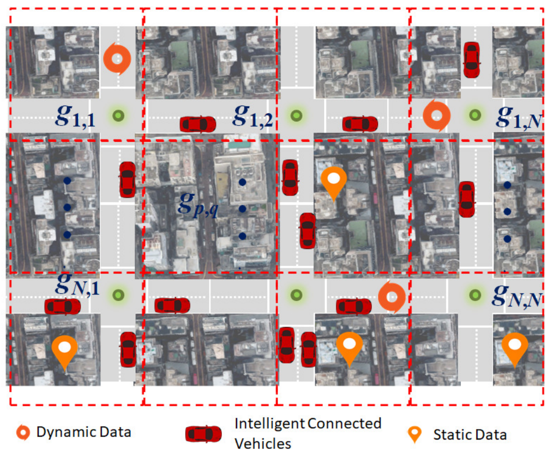

Fig. 2. Uniform partitioning of an urban map revealing various $g_{p,q}$ and its corresponding utility of network parameters (e.g., dynamic and static data, and the number of intelligent connected vehicles) at sampling time $t = iT_S$.

other and the utility function can be a combination of any of these $\eta_{p,q}$'s. In general, there are certain network characteristics that are dependent on each other, e.g., velocity and density. However, a study in [31] showed that the independence of these two network characteristics happens when the traffic load is < 30 cars/min. In this work, $\eta_{p,q} = c_{p,q}$.

If vehicular movements and other dynamic network characteristics are disregarded, the map grid's current spatial network characteristic is constant, i.e., $\zeta_{p,q} =$ constant. However, when dynamic map data, sources, and movement of vehicles are considered, $\zeta_{p,q}$'s vary in both space and time. Therefore, to implement a consistent and reliable grid partitioning, the map grid's spatiotemporal stable $\zeta_{p,q}$, must be determined from available sampling times, iT_S, where $i \in \{1, 2, \ldots, I\}$. $I = \frac{24*60}{T_S}$. T_S denotes the sampling time of each mobility trace measured in minutes. A map grid $g_{p,q}$'s spatiotemporal stable network characteristic, $\zeta_{p,q,STS}$, is established according to (1).

$$\zeta_{p,q,STS} = \sum_{i=0}^{I} \alpha(iT_S)\omega(iT_S), \tag{1}$$

where

$$\alpha(iT_S) = \frac{\zeta(iT_S) - \min\left[\zeta(iT_S)\right]}{\max\left[\zeta(iT_S)\right] - \min\left[\zeta(iT_S)\right]}$$
$$\omega(iT_S) = \frac{\zeta(iT_S)}{\max\left[\zeta(i = 0, \ldots, IT_S)\right]}.$$

$\alpha(iT_S)$ is the feature scaling parameter at time $t = iT_S$, while $\omega(iT_S)$ is the weight correlating all the $\zeta_{p,q}$'s, respectively. Fig. 3 illustrates an example of how the spatiotemporal stable network characteristics $\zeta_{p,q,STS}$'s are generated as defined in (1) when $N = 20$. $\zeta_{p,q}$ is characterized by the map grid's vehicular capacity, $c_{p,q}$.

B. Forming Effective Regions of Movement

To divide a geographical area into various sections and determine the possible candidate locations for deploying road-side units, we introduce the concept of the Effective Regions of Movement (ERMs). An ERM is a grouping of edge-adjacent map grids with its spatiotemporal stable network characteristic, $\zeta_{p,q,STS}$, that possess a unifying characteristic, such as vehicular capacity, density, etc. The merging [32] of spatiotemporal map grids form an ERM, ERM_e, is governed by (2a) below.

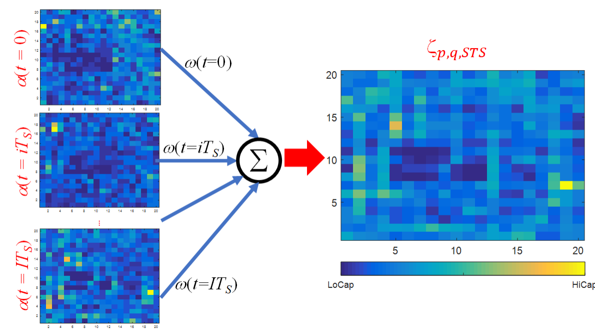

Fig. 3. An example of how $\zeta_{p,q,STS}$ of each map grid $g_{p,q}$ is formed. Each $\zeta(t = iT_S)$ is characterized by its current vehicular capacity, $c_{p,q}$. Darker map grids have lower vehicular capacities (LoCap) over lighter grids (HiCap).

$$ERM_e \equiv g_{p,q} \cup g_{p+\Delta p,q+\Delta q} \tag{2a}$$
$$\text{subject to} \quad \left|\{c_{g_{p,q}}\} \cup \{c_{g_{p+\Delta p,q+\Delta q}}\}\right| \leq \tau_c \tag{2b}$$
$$\min(\rho_{g_{p,q}}, \rho_{g_{p+\Delta p,q+\Delta q}}) \geq \rho_0, \tag{2c}$$

where $\Delta p, \Delta q \in \{-1, 0, 1\}$. $c_{g_{p,q}}$ is the expected vehicular capacity found in $g_{p,q}$, while τ_c is the vehicular capacity threshold of each formed ERM_e. Note that only edge-adjacent grids are considered during the merging. Constraint (2b) allows the merging of grids $g_{p,q}$ and $g_{p+\Delta p,q+\Delta q}$ when the merged grids have vehicular capacity less than the vehicular capacity threshold, τ_c, of each formed ERM_e. $|\bullet|$ defines the cardinality of the union of the vehicular capacities found in grids $g_{p,q}$ and $g_{p+\Delta p,q+\Delta q}$. $\rho_{g_{p,q}}$ and $\rho_{g_{p+\Delta p,q+\Delta q}}$ in Constraint (2c) denote the outbound and inbound vehicular flow from and to $g_{p,q}$, respectively. If the minimum vehicular flow between two map grids is $\geq \rho_0$, then merging proceeds; otherwise, $g_{p+\Delta p,q+\Delta q}$ is dropped.

The algorithm to determine various ERMs of an area under study is illustrated in Algorithm 1. Fig. 4 shows an illustration of formed ERMs derived from Fig. 3 containing single- and multiple-grid ERMs.

IV. Enhanced Information Sharing RSU Allocation (EISHA-RSU) Scheme

In this section, we discuss the novel Enhanced Information SHAring RSU (EISHA-RSU) allocation technique. EISHA-RSU maximizes information sharing and vehicular connectivity in ERMs by allowing relevant and on-time information exchange among the largest number of vehicles and infrastructure. In essence, by considering the proper spacing between deployed RSUs, EISHA-RSU, can maximize the urban coverage area and achieve its goals by utilizing the effective position (EP) concept for locating ideal RSU deployment position.

Algorithm 1 Determining Effective Regions of Movement (ERMs)

INPUT:
$g_{p,q}$'s – List of spatiotemporal stable geographical grids
$c_{p,q}$ – vehicular capacity of $g_{p,q}$
τ_c – vehicular quantity threshold,
N^2 - number of grids
OUTPUT: ERM, List containing ERMs.

Δp = [-1,0,1,0]; and Δq = [0,1,0,-1];
$ERM = 0_{N \times N}$;
region = 1; ▷ **Initialize first ERM region to 1.**
L_c = sort($c_{p,q}$) in descending order;
for $i = 1$ to length of L_τ **do**
    $[p, q]$ = ind2sub(index($L_\tau(i)$));
    ▷ **Convert linear indices to subscripts**
    $ERM(p, q)$ = region;
    $L_c = L_c \setminus L_c(g_{p,q})$;
    while $L_c \neq \emptyset$ **do**
        for $k = 1$ to 4 **do** ▷ **Get edge-adjacent map grids.**
            if Constraints (2b) AND (2c) == TRUE **then**
                if $g_{p+\Delta p(k),q+\Delta q(k)}$ is not yet visited **then**
                    $ERM(p + \Delta p(k), q + \Delta q(k))$ = region;
                    ▷ **Assign** $g_{p+\Delta p(k),q+\Delta q(k)}$ **to ERM region**
                    $L_c = L_c \setminus L_c(g_{p+\Delta p(k),q+\Delta q(k)})$;
                end if
            end if
        end for
    end while
    region = region + 1; ▷ **Move to next ERM**
end for

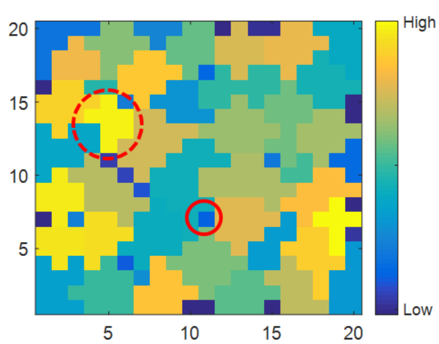

Fig. 4. An example of various ERMs formed by following Algorithm 1. Dark and light colors depict low and high ERM priorities, respectively. In the formation of ERMs, single-grid (bold circle) and multi-grid (dotted circle) ERMs are created.

We define an EP as a physical urban location where we can deploy an RSU for environment information collection, such as an intersection, certain location along a road segment, a combination of both, or any landmark.

A. Problem Formulation

The main goal of EISHA-RSU is to locate RSUs in spots where maximized information is shared between the infrastructures and vehicles. These road and environment data can be either static (landmarks and roads) or dynamic (pedestrian, road accidents, events, etc.) in nature. Given a constraint in the number of RSUs to be deployed, EISHA-RSU prioritizes each effective region of movement to ensure that even the least prioritized map grid still has a chance to obtain relevant and on-time information from high-priority ERMs. EISHA-RSU also assures on-time delivery and storage of dynamic environment data collected from the surrounding vehicles.

In each ERM, the amount of information shared, I_{Sha}, is given in (3).

$$\begin{aligned} I_{Sha} &= U_\gamma + \beta U_\delta \\ &= \sum_{p=1}^{N}\sum_{q=1}^{N} \gamma_{g_{p,q}} + \beta \sum_{p=1}^{N}\sum_{q=1}^{N} \delta_{g_{p,q}}(t), \end{aligned} \tag{3}$$

where $\gamma_{g_{p,q}}$ and $\delta_{g_{p,q}}(t)$ represent the amount of static and dynamic environment road map data, respectively. We note again that each $g_{p,q}$ is dependent on its corresponding longitude (x) and latitude (y) coordinates. β is an importance factor we assign to dynamic road data to signify the repetitive occurrence of instantaneous events, such as accident-prone areas in a $g_{p,q}$, where $1 \leq \beta \leq \frac{1}{\xi}$. ξ denotes the proportionality constant between static and dynamic environment data such that $\delta_{g_{p,q}} = \xi \gamma_{g_{p,q}}$, where $0 < \xi \leq 1$, i.e., dynamic environment data are much less than static environment data. When $\beta = \frac{1}{\xi}$, it implies that the dynamic environment data $\delta_{g_{p,q}}$ have high importance and are treated equally as static environment data.

EISHA-RSU addresses the maximization problem given by (4a) subject to constraints (4b) and (4c).

$$\begin{aligned} &\text{maximize} \quad I_{Sha} && \text{(4a)} \\ &\text{subject to} \quad \sum_{l=1}^{N^2} EP_l \leq \Omega_R, \; EP_l \in \{0, 1\} && \text{(4b)} \\ &\qquad \frac{d(EP_l, EP_m)}{v(EP_l, EP_m)} \leq W \quad \forall EP_l, EP_m = 1. && \text{(4c)} \end{aligned}$$

Constraint (4b) assures that there is only the at most Ω_R RSUs to be deployed in each ERM or the urban map under study, located at effective positions, EP_l. $EP_l = 1$ means an RSU can be deployed there; otherwise, $EP_l = 0$. Constraint (4c) dictates the network's allowable on-time delivery delay of information, W, between deployed RSUs found in effective positions EP_l and EP_m, which is equal to the quotient of $d(EP_l, EP_m)$ and $v(EP_l, EP_m)$. We note that $d(EP_l, EP_m)$ does not automatically equal to the shortest distance between EP_l and EP_m. It is the travel distance with respect to the road network. In Fig. 5, $d_1 = d(EP_l, EP_m) = 4$ and $d_2 = d(EP_l, EP_k) = 3$.

B. Delay Analysis between Two Effective Positions

To achieve on-time and up-to-date delivery of environment data, the distance between two EPs must be minimized, according to the delivery delay W. The calculated separation between EPs will allow vehicles outside an RSU's transmission range to travel and carry valid and relevant road information from one grid to another without any RSU.

The general expression for the total average delivery delay, W, to locate two EPs, EP_l and EP_m found in ERM_e, is given in (5).

$$W = W_g + W_a, \tag{5}$$

where W_g is the average time required for a vehicle to deliver its stored data before becoming invalid to an RSU at an effective position while traversing road distance d on g grids along the path. W_a is the average additional stop time the vehicle encounters during its trip, e.g., passing an intersection or encountering accidents. For simplicity, we assume that W_a is constant, e.g., traffic lights operating at fixed cycles [33]. (5) can be re-written to the expression given in (6) to determine how much time a vehicle takes to traverse grids g along a given path found in ERM_e.

$$W_g = W - W_a. \tag{6}$$

We adopt the delay analysis in [34], [35], and apply it to the 2D scenario in Fig. 5 to determine how far an effective position should be situated from an initial EP, EP_l. The following are assumed to determine the effective positions where RSUs can be allocated.

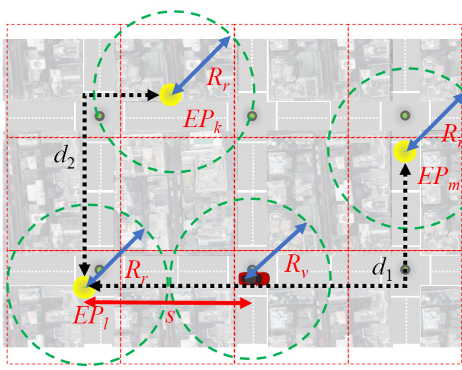

Fig. 5. Determining the effective positions in an ERM given the allowable network data delivery delay and road parameters of the urban city.

1) When outside of an RSU transmission range, intelligent connected vehicles can still sense environment data. These collected vehicular data can only be shared and forwarded to an RSU found at an effective position in the path ahead (V2I operations only). V2V communications are not considered here to model the worst case scenario, i.e., a vehicle must bring its information directly to the RSU.
2) The transmission ranges of a deployed RSU at an effective position and a vehicle are R_r and R_v, respectively.

Upon leaving an RSU at EP_l, the average time for the red vehicle as depicted in Fig. 5 to deliver its newly collected data to a nearby effective position found in another grid, while still being valid, is (7), given that the conditional probability of s (the location of the vehicle with information) in $[0, d]$ is $f(s)$.

$$W_g = \int_0^d WT_g f(s) ds. \tag{7}$$

Given the assumptions above, the time needed for the red vehicle (source vehicle) in Fig. 5 to deliver its valid data along a road of length d, WT_g, is given by (8).

$$\begin{aligned} WT_g &= \phi T \\ \phi &= 1 - (1 - e^{-\lambda R_v})^{\kappa} \\ T &= \frac{d - s - R_r + R_v - E_x}{v}, \end{aligned} \tag{8}$$

where

$$\kappa = \frac{2(d - s - R_r + R_v)}{\mathrm{E}[d_V]}$$

$$E_x = \frac{\mathrm{E}[d_V]\left[1 - (\kappa+1)(1 - e^{-\lambda R_v})^{\kappa} + \kappa(1 - e^{-\lambda R_v})^{\kappa+1}\right]}{2[1 - (1 - e^{-\lambda R_v})^{\kappa}]e^{-\lambda R_v}}.$$

(8) is jointly characterized by the number of departing vehicles at an RSU and the travel delay from that RSU to the next and nearest RSU location. ϕ is the probability that the RSU at EP_m is beyond the range of the source vehicle with transmission range R_v. T is the delivery delay from position s to $d - R_r$, i.e., out of RSU's communication range at EP_m.

λ is the departing rate of vehicles from an effective position that can overtake the source vehicle and can become a forwarding node. v is the space mean speed of the road segment d. $\mathrm{E}[d_V]$ is the average distance of vehicles found between the two effective positions, d is the maximum separation between two effective positions and is not necessarily the shortest distance but the distance defined by the road topology. s is the location (distance traveled from EP_l) of a source vehicle having new environment data to be shared.

The worst-case scenario happens when a vehicle has real-time environment data and has no immediate RSU to offload its contents. This scenario also occurs when it has no leading vehicle(s) within its transmission range, R_v, to which it can forward its information. As such, the worst case scenario happens where $s = R_r + R_v$, $\lambda = 0$, and $\mathrm{E}[d_V] = d - s - R_r$.

Given the values of W and W_a, the maximum allowable separation, d, between two effective positions is given in (9). Note here that the conditional probability of s in $[0, d]$ follows a uniform distribution since there is no prior knowledge of the locations in $[0, d]$. We allow all s's in $[0, d]$ to be equiprobable points where a car can sense its environment [36].

$$d = 2W_g v + R_r. \tag{9}$$

C. EISHA-RSU *Algorithm*

The EISHA-RSU scheme allocates RSUs to effective positions found in each ERM by satisfying the maximization problem in (4). Its detailed operation is illustrated in the pseudo-code in Algorithm 2.

ERMs can be categorized into two configurations, namely: (1) single-grid and (2) multiple-grid. We discuss for each configuration how the EPs are identified.

1) Single-Grid ERMs: For single-grid ERMs, an example is shown in Fig. 6, where its static environment data, $\gamma_{g_{p,q}}$, is represented by the blue shade. The grid is further sub-divided into $k = 1, 2, \ldots, K$ sub-grids to introduce additional dynamic data, $\delta_{g_{p,q}} = \sum_{k=1}^{K} \mu_k \delta_{k,g_{p,q}}$, where $\mu_k \in \{0, 1\}$ and is used to

Algorithm 2 EISHA-RSU Algorithm

INPUT:
ERM's – Formed ERMs with its priority
Ω_R – maximum number of RSU's to be deployed
W_g – Waiting time for data to be considered valid
R_r and R_v – transmission ranges of RSU's and vehicles, respectively
K – number of sub-grids

OUTPUT: EP_{List}, List containing EPs of each ERM.

Determine ERM ascending priority list, ERM_{List};
Divide each grid of ERM_e into K sub-grids.
$\Omega_{R_{EP}}$ = ceil(Ω_R / ERM_{List}); ▷ # of EPs for high-priority ERMs
$EP_{List} = \emptyset$.
$\Omega_{R_{ctr}} = \Omega_R$;
for $e = 1 : \max(ERM_{List})$ **do**
  if ERM_e is single-grid ERM **then**
    if $\Omega_{R_{ctr}} > 0$ **then**
      $EP_{temp} = [x^*, y^*]$;
      $EP_{List} = EP_{List} \cup EP_{temp}$;
      $\Omega_{R_{ctr}} = \Omega_{R_{ctr}} - 1$;
    end if
  else
    Determine how many $\Omega_{R_{EP}}$ for ERM_e.
    if $\Omega_{R_{EP}} > 0$ **then**
      M = Convolve ERM_e with $\mathbf{1}_{\sqrt{K} \times \sqrt{K}}$
      $\mathbf{M_{List}}$ = nchoosek$(\mathbf{M}, \Omega_{R_{EP}})$;
      $ctr = 1$;
      while size$(EP_{List}) \neq \Omega_{R_{EP}}$ **do**
        Compute distances in $\mathbf{M_{List}}(ctr)$.
        EP_{temp} = center locations of $\mathbf{M_{List}}(ctr)$.
        if (9) is satisfied **then**
          $EP_{List} = EP_{List} \cup EP_{temp}$;
        end if
        $ctr = ctr + 1$;
      end while
    end if
    $\Omega_{R_{ctr}} = \Omega_{R_{ctr}} - \Omega_{R_{EP}}$;
  end if
end for
Output EP_{List}.

reduce the computation cost for dynamic data. A sub-grid k contributes dynamic data when $\mu_k = 1$, else zero.

(4a) reduces to an optimization problem requiring only one RSU to be deployed, since this is the least deployable number of RSUs. The effective position where the RSU is located is at the point where maximum static and dynamic data are shared, as defined in (10). We set $\beta = 1$.

$$\begin{aligned} I_{Sha} = & \int_{x_1}^{x_1+\Delta x+R_r} \int_{y_1}^{y_1+\Delta y+R_r} \gamma_{g_{p,q}}(x,y)dxdy \\ & + \sum_{k=1}^{K} \int_{\Delta x_{1,k}}^{\Delta x_{2,k}} \int_{\Delta y_{1,k}}^{\Delta y_{2,k}} \mu_k \delta_{k,g_{p,q}}(x,y)dxdy. \end{aligned} \tag{10}$$

The center of an RSU with transmission range R_r is moved from the corner point x_1, y_1 by $\Delta x, \Delta y$ until maximum static data are covered. By doing this, the maximum static data are shared when $\Delta x = \frac{x_2 - x_1}{2}$ and $\Delta y = \frac{y_2 - y_1}{2}$, i.e., the center of the grid. Thus, the location of the effective position EP to cover static data in a single grid is:

$$x_\gamma = x_1 + \Delta x \text{ and } y_\gamma = y_1 + \Delta y. \tag{11}$$

Likewise, for covering all K sub-grids of an ERM containing dynamic data, the location is at:

$$x_\delta = \sum_{k=1}^{K} \frac{\mu_k \delta_{k,g_{p,q}}(x,y)}{\sum_{k=1}^{K} \mu_k \delta_{k,g_{p,q}}(x,y)} \varepsilon_k. \tag{12}$$

$$y_\delta = \sum_{k=1}^{K} \frac{\mu_k \delta_{k,g_{p,q}}(x,y)}{\sum_{k=1}^{K} \mu_k \delta_{k,g_{p,q}}(x,y)} \zeta_k. \tag{13}$$

ε_k and ζ_k are the centroid coordinates of sub-grid k with available dynamic data, where $\varepsilon_k = \frac{\Delta x_{2,k} - \Delta x_{1,k}}{2}$ and $\zeta_k = \frac{\Delta y_{2,k} - \Delta y_{1,k}}{2}$.

$x_1 \leq x_\gamma, x_\delta \leq x_2$ and $y_1 \leq y_\gamma, y_\delta \leq y_2$. Given these two possible EP locations, the maximum static and dynamic information shared is achieved when the RSU is situated at an EP having coordinates in (14), i.e., at the center of the single-grid ERM.

$$x^* = x_\gamma \text{ and } y^* = y_\gamma. \tag{14}$$

under the constraint that $x_2 - x_1 = y_2 - y_1 \leq \sqrt{2} R_r$.

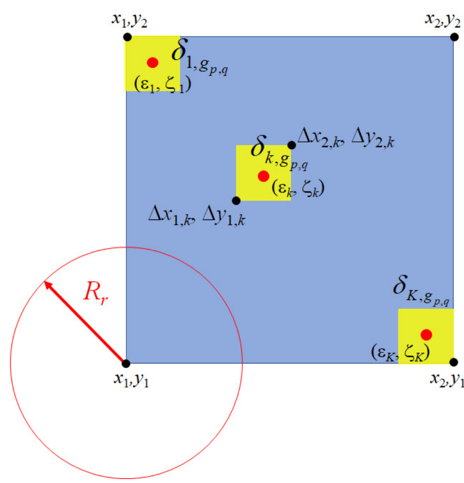

Fig. 6. Determining the effective position in a single-grid ERM, given the grid's static and dynamic data.

2) Multiple-Grid ERMs: There are two cases for multiple-grid ERMs, 1) the number of map grids is equal to Ω_R, and 2) the number of map grids is greater than Ω_R. For case 1), locating EPs follows the single-grid ERM deployment, where each map grid of the ERM has an effective position at the center. However, if the number of grids is higher than the desired number of deployable RSUs, then the EPs are heuristically searched.

Given Ω_R RSUs to be deployed in an urban setup, EISHA-RSU allocates an RSU to all ERMs by following a round-robin procedure. If Ω_R is higher than the lowest priority ERM, then all ERMs are guaranteed to have at least one effective position where an RSU can be deployed. Another round robin deployment ensues and ends until the target number Ω_R has been reached.

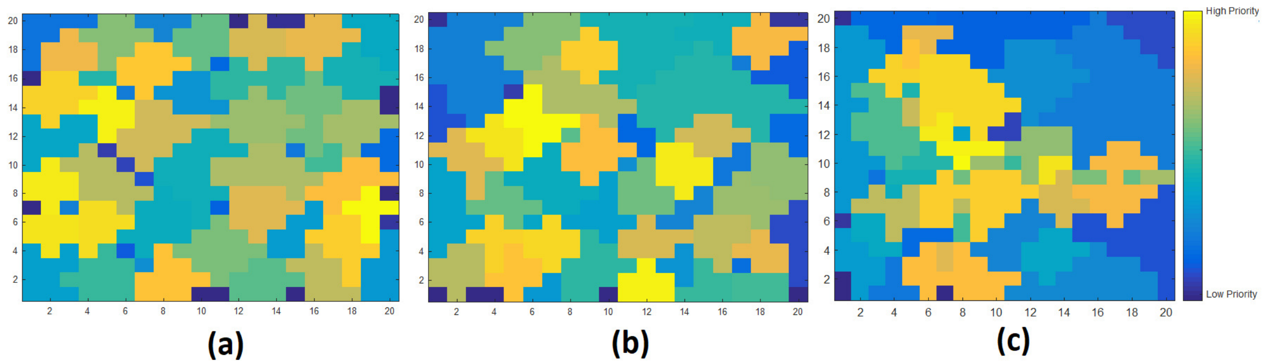

Fig. 7. Spatiotemporal stable ERMs with $N = 20$ sampled at $T_S = 10$ min for (a) Beijing, (b) Jakarta, and (c) Singapore. Dark color ERMs have lower priorities over lighter ERMs.

After calculating the number of RSUs needed to be deployed in each ERM, $\Omega_{R_{EP}}$, then EISHA-RSU follows a greedy heuristic method for finding these $\Omega_{R_{EP}}$ EPs. When Ω_R is not exactly divisible by the number of ERMs, higher priority ERMs will have $\Omega_{R_{EP}}$ EPs, while lower priority ERMs will have $(\Omega_{R_{EP}} - 1)$ EPs. A map grid is further divided into K sub-grids to accommodate the presence of dynamic data, if any. $K = 9$ sub-grids is considered for the results in this paper. The ERM is then convolved with a $\mathbf{1}_{\sqrt{K}\times\sqrt{K}}$ filter to determine how much information can be shared when one vehicle travels from one sub-grid to another in the ERM. With respect to Algorithm 2, the convolution results are stored in $\mathbf{M}$. From $\mathbf{M}_{\mathbf{List}}$, EISHA-RSU selects the first $\Omega_{R_{EP}}$ locations with maximum shared information. Their distances should be equal to (9) to avoid overlapping between RSU coverage and prevent invalid data delivery. If their separations do not satisfy (9), then the next maximum information location combination is considered until the optimization problem in (4a) is satisfied. The discussion of this is seen in lines 13–25 of Algorithm 2.

Let us consider the partitioning and ERM formations depicted in Fig. 4. For simplicity, let us consider only the two formed ERMs highlighted by the circles and let $\Omega = 3$. The ERM with the dotted circle has the higher priority over the ERM with the bold circle, therefore, two effective positions must be determined in this ERM, $\Omega_{REP} = 2$. Since there is a single-grid ERM, the effective position is determined by line 9 in Algorithm 2. On the other hand, the multiple-grid ERM has seven grids that contain vehicular environment information. Lines 16 to 17 determine all the possible combinations of locating the two optimal effective positions. For example, if the grid combination of (5, 12) and (5, 14) is compared to the grid combination of (4, 14) and (5, 14), and both combinations have the same information content (denoted in Line 17), EISHA-RSU picks grids (5, 12) and (5, 14) as the optimal effective positions because it satisfies the required distance between the chosen effective positions over the other combination. RSU placements at (5, 12) and (5, 14) also provide a wider coverage between the two.

V. Simulation Results and Discussion

In this section, we present extensive simulation results employing empirical mobility traces and city locations to evaluate the performance of the proposed EISHA-RSU allocation scheme.

A. Simulation Setup

We utilize three public transport mobility trace datasets, namely, (1) Beijing (BJS) [37], (2) Singapore (SIN) [38], and (3) Jakarta (JKT) [38]. The statistics of these three mobility traces, as well as the simulation parameters, are summarized in Table I. Note that the number of vehicles in the JKT and SIN mobility traces also correspond to the number of trips, i.e., each of the 28,000 and 16,174 vehicles has only one trajectory happening at a certain time interval of the day. On the other hand, the 24,845 taxis found in the BJS dataset has more than one trip per vehicle occurring throughout the day.

TABLE I
Empirical Mobility Traces Attributes and Simulation Parameters

Urban City Parameter	**BJS**	**JKT**	**SIN**
Total Area (in $\approx$ km^2)	51	51	51
Grid Area (in $\approx$ m^2)	125,000	125,000	125,000
Total number of vehicles	24,845	28,000	16,174
$R_v = R_r$ (in m)	250	250	250
τ_c (#of vehicles)	993	560	485
Sampling Time T_S (in min)	10	10	10
ρ_0	0.25	0.25	0.25
W_g (in min)	2.8070	2.8070	2.8070
v (in m/s)	5.5556	5.5556	5.5556

The generation of the static environment data (in MB) for each $g_{p,q}$ is governed by:

$$\gamma_{g_{p,q}} = 5000\big(\sin(150y + 15)\cos(150x + 20) + 1\big), \quad (15)$$

where x and y are the corresponding longitude and latitude coordinates of each $g_{p,q}$, respectively. On the other hand, we generate dynamic data by further subdividing a map grid into nine smaller grids. We then perform uniform selection across all sub-grids to randomly select locations where 'accidents' happen, since we do not have enough accident data of BJS, JKT, or SIN. Thus, the generation of additional dynamic environment data is governed by:

$$\delta_{g_{p,q}} = \xi \gamma_{g_{p,q}}. \tag{16}$$

To represent the dynamic environment data available in each $g_{p,q}$, we let $\xi = 0.01$, thus, the importance factor $1 \leq \beta \leq 100$.

B. Effect of Vehicular Capacity in the Formation of ERMs

By implementing Algorithm 2 in the three empirical mobility traces, the ERMs formed per city is shown in Fig. 7. By appropriately selecting the values of τ_c and ρ_0, there are approximately 50 spatiotemporal ERMs found in each urban map, both having single- and multiple-grid ERMs. As the value of τ_c becomes smaller and the value of ρ_0 increases, the number of ERMs employed by EISHA-RSU in determining where to optimally deploy the RSUs approaches the number of grids used in the three baseline schemes, i.e., more single-grid ERMs are present. As more single-grid ERMs are formed, priority takes place and the determining of EP locations will be hastened, i.e., found at the center of the single-grid ERMs. The worst case scenario for EISHA-RSU deployment happens when the ERM priority matches those of the MaxInfo grid selection, i.e., single-grids with the highest available information. On the other hand, when the value of τ_c becomes larger and the value of ρ_0 decreases, larger ERMs are formed implying less distinctions in the given urban map. At the limit, EISHA-RSU will perform like the CityWide baseline scheme if and only if the chosen effective positions correspond to lines 19 to 26 in Algorithm 2 and if the linear interval between effective positions is defined as in (9). In Fig. 8, the τ_c values are equal to the percentage with respect to the number of vehicles in the urban city under study. A system designer can refer to this figure when the number of formed ERMs is critical.

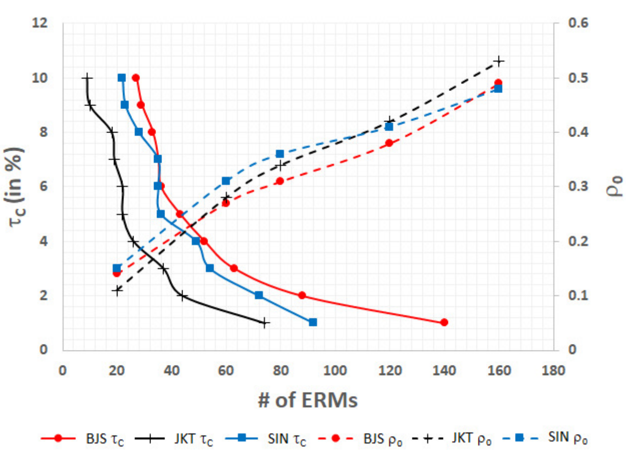

Fig. 8. The effect of varying τ_c and ρ_0 in the formation of ERMs.

C. Stationarity of ERMs

Determining the appropriate sampling time in establishing the ERMs is the first step. Faster sampling time yields a higher number of mobility traces and can depict the movement within the area under study, at the expense of high computational cost. On the other hand, slower sampling time reduces the mobility traces under study and may reduce vehicular network information necessary to provide valid results. To measure ERM stationarity, the root-mean-squared-error (RMSE) (17) of a formed ERM at sampling time T_S is compared to the ERM formed at a reference sampling time, $T_{S_{ref}} = 60$ min.

$$RMSE = \sqrt{\frac{1}{N^2}\sum_{p=1}^{N}\sum_{q=1}^{N}\left[ERM_{p,q}(T_{S_{ref}}) - ERM_{p,q}(T_S)\right]^2}. \tag{17}$$

Fig. 9 illustrates the effect of varying the sampling time from 2 to 30 min. Notice that as we decrease the value of the sampling time, there is an approximate flat response. If the ERMs are dynamic, the RMSE value should approach 400 (i.e., $N \times N$, $N = 20$), signifying that the vehicular trajectories change hastily. However, from Fig. 9, the average RMSE value is only six. The small average RMSE value implies that formed ERMs with sampling times $T_S = 5$, 10, 15, 20, and 30 min have minimal differences with $T_{S_{ref}} = 60$ min and can be regarded as stationary. With this finding, we can use any of these sampling times, and still be able to form approximately the same ERMs. We note that $T_S = 5$, 10, 15, 20, and 30 min are practical sampling times to sense the environment data.

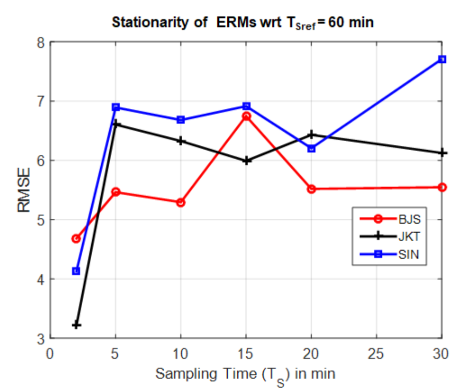

Fig. 9. Determining stationarity of ERMs by varying the sampling time where each $g_{p,q,STS}$ is determined. $T_{S_{ref}} = 60$ min is the reference to test stationarity.

D. Deployment Performance Evaluation

We compare the performance of the EISHA-RSU allocation scheme with the following benchmarks, as described below.

1) Uniform Deployment (UnifDep) [39] allocates the effective positions Ω_R by following the uniform distribution to allow all places to be equally chosen since there are no prior information of each of the spatial location. However, this model can be changed to another distribution function once the selection probability of places is defined, e.g., putting more weight on intersections, bus stops, popular landmarks, etc. For this deployment strategy, simulations are run for 1000 times to capture uniformity.
2) Citywide Deployment (CityWide) [40] chooses the effective positions Ω_R based on maximum urban area coverage. As the coverage is maximized, the greater

amount of information can be sensed and disseminated in the network. This represents the deployment based on the road topology of a certain urban locations. In this work, we simply represent the maximum coverage by following linearly-spaced locations.

3) Maximum Information Deployment (MaxInfo) [27] pinpoints the effective positions Ω_R as the locations that have maximum information. The strategy of [27] to place static RSU in locations with the highest weights is comparable to our baseline method MAXINFO which deploys RSUs to locations with the highest amount of information that can be sensed and disseminated.

References [12], [22], and [25] fall under the UnifDep benchmark. The RSU deployment in [12] depicts the characterization (distribution) of urban locations based on connectivity probability. [22] models the urban setup according to the location's importance factor or deployment priority determined by connectivity robustness and vehicular capacity. [25] also works the same way as [22] but added connectivity duration in its optimization problem. On the other hand, previous studies [4], [8], and [21] are closely related to the CityWide benchmark. The work in [4] explored a 2D road network scenario with intersections equally spaced apart but discriminated by road densities. Simulations done in [8] placed candidate locations in linearly-spaced and prioritized intersections. [21] also follows this setup but considered RSU placement based on intersection radio signal spreading. Lastly, literatures [26], [27], and [41] are categorized under the MaxInfo benchmark that focuses on deterministic properties. The RSUs in [26] were placed on locations that minimizes both the capital and operating expenses. A comprehensive RSU deployment based on fixed locations, structured, and unstructured transportation modes have been studied in [27]. Finally, the study presented here is also an extension of [41] that does not only focus on V2I information exchanges. This is summarized in Fig. 10.

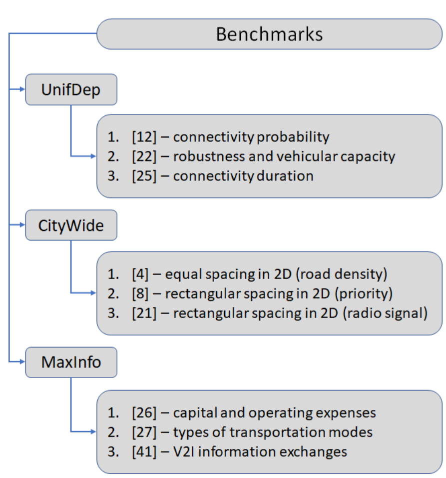

Fig. 10. Summary of state-of-the-art methods used in the benchmarks.

The amount of information shared by the four allocation schemes is illustrated in Fig. 11. This covers information within the RSU coverage area and those carried by vehicles beyond the RSU transmission range but within the allowable distance dictated by (9). For all deployment schemes, it is noticeable that as more effective positions are selected, more information is gathered and exchanged in the vehicular network. EISHA-RSU also performs the best by correctly placing effective positions in the urban area, therefore, enhancing the amount of information shared.

One may argue that the MaxInfo strategy should collect the most data, as its name implies. However, we note that once we placed the RSUs at locations where there is maximum information shared, we observed that the RSUs' transmission ranges are overlapping or at least adjacent to one another, thus, there is redundancy in shared information among RSUs in MaxInfo. This type of deployment allows EPs to be very close to each other, e.g., several meters, leading to the reduced coverage area and vehicular connectivity. Therefore, there will be less or no information-carrying vehicles found within a distance $\leq d$ that will arrive at an EP to deliver additional contents. On the other hand, the other three allocation schemes, especially EISHA-RSU, are able to accurately discriminate grid locations as possible effective positions to sense more environment data.

For all allocation methods, there is an assumption that RSUs are not connected. However, if all RSUs have a wireless or wired connection, then, EISHA-RSU will still be the scheme with the highest amount of shared information in the network. This is attributed to fact that the RSU allocation done by EISHA-RSU is well-positioned in the urban area to collect more base information, when compared to the other three benchmarks.

In addition, the fairness of the network's starvation is measured. To measure how fairly the installed RSUs collect information over the urban map, we relate the proportional fairness [42] to the average throughput, S_i, for each RSU assuming equal data upload and download rate. According to [42], proportional fairness happens when the deployed RSUs receive an equal amount of environment information. We then calculate the Jain's Network Starvation Fairness Index (18) to evaluate the performance of the deployment scheme as the number of deployed RSUs increases [43]. Ω_R is equal to the total number of deployed RSUs.

$$J(S_i) = \frac{\left(\sum_{i=1}^{\Omega_R} S_i\right)^2}{\Omega_R \sum_{i=1}^{\Omega_R} S_i^2}. \tag{18}$$

A higher value of the Jain's starvation fairness index implies that there is almost an equal average amount of information collected by the RSUs allocated on EPs across the urban map under study. Interpreting this index reveals that EISHA-RSU fairly allocates effective positions to allow RSUs to capture approximately equal amount of environment information, as shown in Fig. 12. As the number of deployed RSUs ($>$ 40 deployed RSUs) increases, it is noticeable that the other three deployment schemes have a fast rate of decreasing Jain's starvation fairness index when compared to EISHA-RSU. This decline in index value highlights that there is a huge discrepancy in the collected data among deployed RSUs.

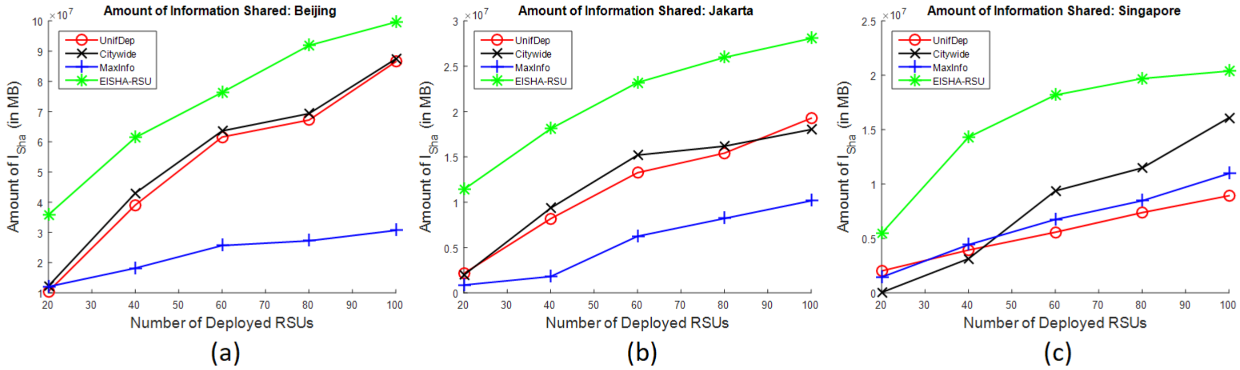

Fig. 11. The amount of information shared in (a) Beijing, (b) Jakarta, and (c) Singapore by utilizing UnifDep, CityWide, MaxInfo, and EISHA-RSU deployment schemes.

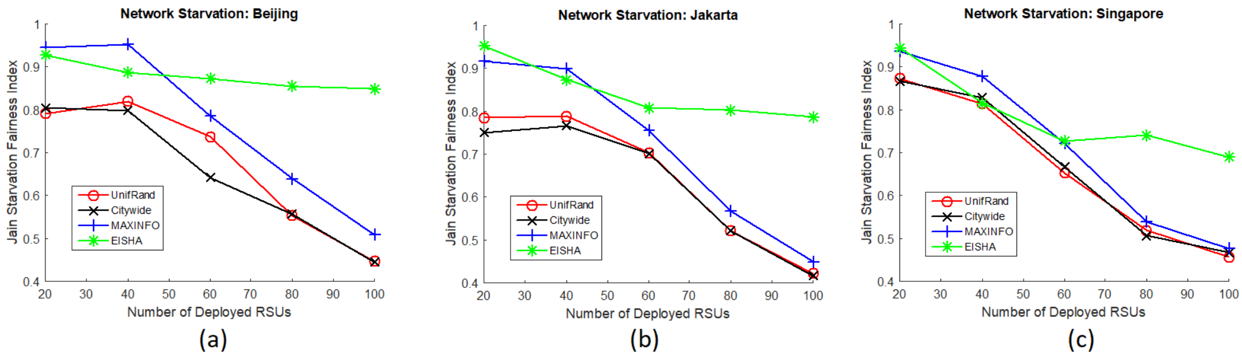

Fig. 12. The Jain's network starvation fairness index indicates how much equal the average throughput there is per deployed RSU by each deployment scheme for (a) Beijing, (b) Jakarta, and (c) Singapore.

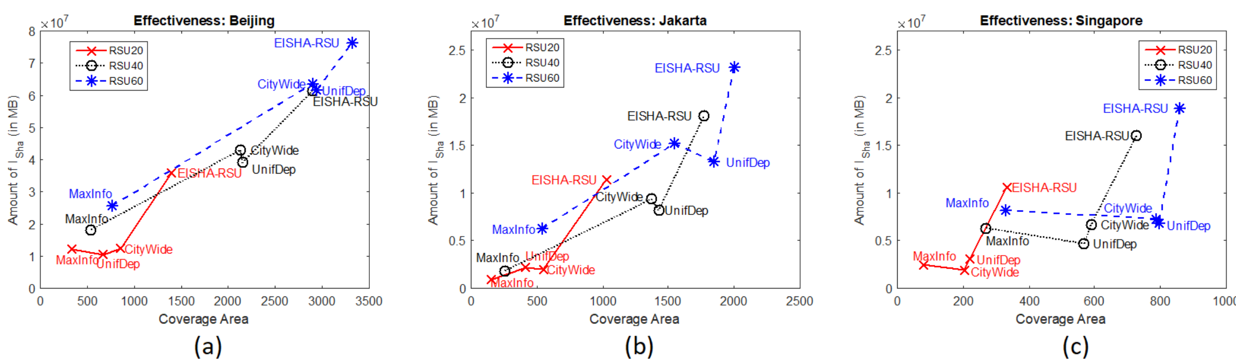

Fig. 13. Performance comparison of each deployment scheme on the cities of (a) Beijing, (b) Jakarta, and (3) Singapore based on Effectiveness.

Though EISHA-RSU also experiences the index decline, the deployment method still achieves a higher value compared to the three, signifying a more balanced data collection.

We compare the performance of the deployment schemes according to its *Effectiveness* [44]. Let $I(D_i)$ and $C(D_i)$ denote the amount of information shared and coverage area of a deployment scheme i for a given number of deployed RSUs, respectively. The effectiveness of a deployment scheme i is defined by $E(D_i) = \big[I(D_i), C(D_i)\big]$. If $E(D_1) \succ E(D_2)$, then we say that any of these three conditions is true: 1) $\big[I(D_1) > I(D_2)$ and $C(D_1) > C(D_2)\big]$, 2) $\big[I(D_1) > I(D_2)$ and $C(D_1) = C(D_2)\big]$, or 3) $\big[I(D_1) = I(D_2)$ and $C(D_1) > C(D_2)\big]$.

Fig. 13 shows the *Effectiveness* plot against the amount of information shared and coverage area of all the deployment strategies for a given RSU deployment density. From Fig. 13, MaxInfo is the least effective in terms of the amount of information shared and coverage area. Given any RSU deployment density, i.e., 20, 40, or 60 deployed RSUs, it is evident that EISHA-RSU is the most effective since it captures the most amount of information with the widest coverage urban area. Hence, I(EISHA-RSU)$>$ I(UnifDep) and C(EISHA-RSU)$>$ C(UnifDep), I(EISHA-RSU)$>$ I(CityWide) and C(EISHA-RSU)$>$ C(CityWide), and I(EISHA-RSU)$>$ I(MaxInfo) and C(EISHA-RSU)> C(MaxInfo). This result implies that EISHA-RSU properly allocates RSUs in effective positions that will both capture directly-sensed (either by RSU or vehicle) and single-hop environment data. In some cases, a lower deployment density while employing EISHA-RSU has more shared information than other deployment schemes at a higher deployment density, e.g., RSU20 vs RSU40 in the Jakarta and Singapore datasets.

However, the CityWide and UnifDep allocation methods provide contrasting results between each other. As noticed, CityWide offers higher amount of information shared but at a less coverage area than UnifDep. Though, it can be said that these small changes may denote that these two allocation techniques are interchangeable.

In Table II, we investigate the three urban cities using the same number of sub-grids, i.e., 2500. Such results can be attributed to the nature of the dataset. Table I shows the number of taxis per mobility dataset, i.e., row 4. Since the Beijing dataset (BJS) has more taxis and number of trips per taxi, more GPS traces are provided that allows the user to visualize trips/trajectories undertaken by the taxi throughout the day. However, in the Singapore (SIN) and Jakarta (JKT) traces, each taxi ID has one trip/trajectory and happens only in one time interval of the day. This means that some roads are not covered by the datasets, thus, more RSUs are needed to cover the urban map to sense the environment and share the information with other vehicles. For each urban map, EISHA-RSU, on the average, saves up to 21%, 21%, and 101% of RSUs when compared to UnifDep, CityWide, and MaxInfo, respectively.

TABLE II
REQUIRED NUMBER OF RSUS TO BE INSTALLED BY EACH ALLOCATION SCHEME GIVEN A DESIRED COVERAGE AREA AND AMOUNT OF SHARED INFORMATION.

RSU Allocation Scheme	BJS	% Diff	JKT	% Diff	SIN	% Diff
UnifDep	49	45.00	105	11.06	171	7.90
CityWide	50	46.91	104	10.10	170	7.32
MaxInfo	265	158.11	253	91.64	275	54.04
EISHA-RSU	31		94		158	

To evaluate network connectivity, we only consider how many vehicles are within the single-hop transmission range of its nearest EP, without the consideration of V2V communication. Fig. 14 displays the number of one-hop vehicles for each Ω_R deployment constraint. Because EISHA-RSU considers the appropriate spacing between EPs, the results show that EISHA-RSU captures more single-hop vehicles in the network, when compared to the other deployment schemes for all cities. We emphasize that this spacing between EPs covers the worst-case scenario when there are no leading vehicles. Hence, if V2V is allowed in areas without RSU coverage, data delivery time will be further reduced and would provide fresher and more up-to-date environment data.

In summary and supported by Figs. 11–14, our extensive simulation involving three empirical mobility traces have demonstrated that the proposed EISHA-RSU has enhanced information sharing by accurately selecting effective positions that can provide broader and fairer coverage, stable network connectivity, and maximum shared information. It is also evident that EISHA-RSU has outperformed the three presented benchmarks.

E. Effect of β in EISHA-RSU

Previous results have used $\beta = 1$ as the importance level of all dynamic data, implying equal importance and no prior information on each of the sensed location. However, if dynamic data such as the frequency of accidents and roadblocks/constructions are known in an urban city through intensive monitoring, then, β can be changed accordingly. Since EISHA-RSU locates the maximum information location for initial RSU deployment, changing β will affect how the algorithm will deploy RSUs to achieve maximum coverage and high vehicular connectivity.

We provide an example in Fig. 15(b) where we highlight the sub-grids with dynamic data importance values $\beta = 100$ (lighter color) and $\beta = 1$ (dark color). When these importance factors are multiplied with the dynamic data in Fig. 15(a), the normalized results are shown in Fig. 15(c). For example, if only 20 RSUs are to be deployed in an urban map with 20 ERMs each encompassing one of these sub-grids, then these 20 RSUs will be deployed on those sub-grids with high dynamic data importance. However, if ERMs cover more than one sub-grid, the lighter sub-grid, i.e., with a higher importance value, will automatically be chosen by EISHA-RSU. Therefore, β greatly affects the deployment method and must be used when accurate and correct data distribution is available.

F. Time Complexity Analysis of RSU Deployment Schemes

We evaluate the time complexity, $\mathcal{O}(\bullet)$, of each deployment scheme. To evaluate the schemes, four major processes have been identified, namely, 1) Partitioning of the urban map into $N \times N$ uniform grids, 2) Getting the spatiotemporal stable network characteristics ζ_{STS} of each grid, 3) Forming ERMs, and 4) Deploying RSUs. The time complexities of each process are shown in Table III.

TABLE III
TIME COMPLEXITY COMPARISON OF THE RSU DEPLOYMENT SCHEMES

	Partitioning	Getting ζ_{STS}	Forming ERMs	Deploying RSUs
UnifDep	$\mathcal{O}(N^2)$	–	–	$\mathcal{O}(\Omega_R)$
CityWide	$\mathcal{O}(N^2)$	–	–	$\mathcal{O}(\Omega_R)$
MaxInfo	$\mathcal{O}(N^2)$	$\mathcal{O}(N^2)$	–	$\mathcal{O}(\Omega_R)$
EISHA-RSU	$\mathcal{O}(N^2)$	$\mathcal{O}(N^2)$	$\mathcal{O}(N^2)$	$\mathcal{O}(\Omega_R)$

For optimal deployment, EISHA-RSU takes additional time, i.e., $2\mathcal{O}(N^2)$, in characterizing each grid with respect to its spatiotemporal stable network characteristics, ζ_{STS}, and then forming the effective regions of movement (ERM). As the value of N increases, the resolution of the urban location becomes high, allowing more distinctions and discriminations between grids, e.g., road topology, environment information content, vehicular capacity, density, etc. The grid's spatiotemporal stable network characteristics also become more pronounced and will create some empty grids, which allows the

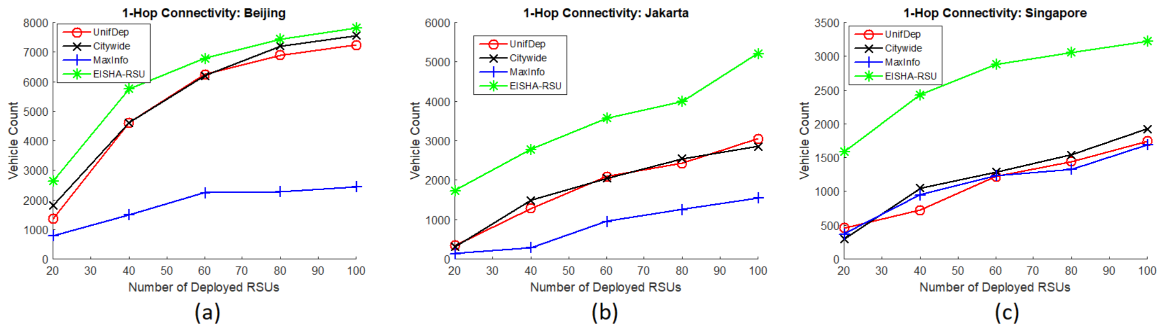

Fig. 14. The amount of vehicles within single hop from an effective position in (a) Beijing, (b) Jakarta, and (c) Singapore that are still capable of delivering valid and up-to-date environment information.

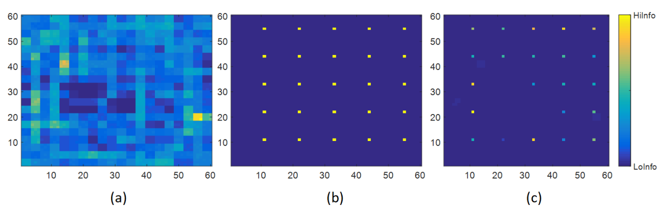

Fig. 15. (a) The amount of static and dynamic information ($\beta = 1$) in each grid, (b) Assigning $\beta = 100$ (lighter color) and $\beta = 1$ (darker color) for dynamic data importance at specific sub-grids, and (c) the resulting possible initial RSU location in an ERM where these sub-grids are found.

formation of higher detailed ERMs. On the other hand, the effect of increasing the value of Ω_R does not really affect the computational complexity, since Ω_R is always much smaller than N^2, especially when the magnitude of Ω_R is close to the number of ERMs.

VI. Conclusion and Future Work

In this work, we have presented an Enhanced Information SHAring RSU (EISHA-RSU) allocation scheme that targets maximal area coverage and vehicular connectivity, resulting to an enhanced amount of information shared between RSUs and vehicles in a vehicular network. To achieve these objectives, an urban map is partitioned according to its Effective Regions of Movement (ERMs) based on its vehicular capacity. EISHA-RSU then locates the effective positions (EPs) that are separated by an optimal distance where RSUs should be deployed to allow maximum information sharing and delivery of vehicular data. The performance of the proposed RSU allocation scheme has been validated by employing three urban empirical mobility traces. Simulation results have verified the fairness, effectiveness, and efficiency of EISHA-RSU when compared to three other benchmarks. In summary, EISHA-RSU allocates fewer RSUs to maximize information sharing, provides wider coverage, and improves connectivity in urban vehicular networks.

References

[1] K. F. Chu, E. R. Magsino, I. W.-H. Ho, and C.-K. Chau, “Index coding of point cloud-based road map data for autonomous driving,” *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pp. 1–7, 2017.

[2] J. Chi, Y. Jo, H. Park, T. Hwang, and S. Park, “An effective rsu allocation strategy for maximizing vehicular network connectivity,” *International Journal of Control and Automation*, vol. 6, no. 4, pp. 259–270, 2013.

[3] C. Lochert, B. Scheuermann, M. Caliskan, and M. Mauve, “The feasibility of information dissemination in vehicular ad-hoc networks,” *Wireless on Demand Network Systems and Services, 2007. WONS’07. Fourth Annual Conference on*, pp. 92–99, 2007.

[4] J. He, Y. Ni, L. Cai, J. Pan, and C. Chen, “Optimal dropbox deployment algorithm for data dissemination in vehicular networks,” *IEEE Transactions on Mobile Computing*, vol. 17, no. 3, pp. 632–645, 2017.

[5] J. P. Talusan, M. Wilbur, A. Dubey, and K. Yasumoto, “On decentralized route planning using the road side units as computing resources,” in *2020 IEEE International Conference on Fog Computing (ICFC)*. IEEE, 2020, pp. 1–8.

[6] A. Ahmad, S. Din, A. Paul, G. Jeon, M. Aloqaily, and M. Ahmad, “Real-time route planning and data dissemination for urban scenarios using the internet of things,” *IEEE Wireless Communications*, vol. 26, no. 6, pp. 50–55, 2019.

[7] Y. Shi, L. Lv, H. Yu, L. Yu, and Z. Zhang, “A center-rule-based neighborhood search algorithm for roadside units deployment in emergency scenarios,” *Mathematics*, vol. 8, no. 10, p. 1734, 2020.

[8] A. Jalooli, M. Song, and X. Xu, “Delay efficient disconnected rsu placement algorithm for vanet safety applications,” *Wireless Communications and Networking Conference (WCNC), 2017 IEEE*, pp. 1–6, 2017.

[9] I. W. Ho, S. C. Chau, E. R. Magsino, and K. Jia, “Efficient 3d road map data exchange for intelligent vehicles in vehicular fog networks,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3151–3165, 2020.

[10] T. Yeferny and S. Allani, “Mpc: A rsus deployment strategy for vanet,” *International Journal of Communication Systems*, vol. 31, no. 12, p. e3712, 2018.

[11] C. Wang, X. Li, F. Li, and H. Lu, “A mobility clustering-based roadside units deployment for vanet,” *Network Operations and Management Symposium (APNOMS), 2014 16th Asia-Pacific*, pp. 1–6, 2014.

[12] S. Mehar, S. M. Senouci, A. Kies, and M. M. Zoulikha, “An optimized roadside units (rsu) placement for delay-sensitive applications in vehicular networks,” *Consumer Communications and Networking Conference (CCNC), 2015 12th Annual IEEE*, pp. 121–127, 2015.

[13] F. Zou, J. Zhong, W. Wu, D.-Z. Du, and J. Lee, “Energy-efficient roadside unit scheduling for maintaining connectivity in vehicle ad-hoc network,” *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication*, p. 64, 2011.

[14] D. Wu, Q. Liu, Y. Zhang, J. McCann, A. Regan, and N. Venkatasubramanian, “Crowdwifi: efficient crowdsensing of roadside wifi networks,” in *Proceedings of the 15th International Middleware Conference*, 2014, pp. 229–240.

[15] D. Wu, D. I. Arkhipov, Y. Zhang, C. H. Liu, and A. C. Regan, “Online war-driving by compressive sensing,” *IEEE Transactions on Mobile Computing*, vol. 14, no. 11, pp. 2349–2362, 2015.

[16] H. Cheng, X. Fei, A. Boukerche, and M. Almulla, “Geocover: An efficient sparse coverage protocol for rsu deployment over urban vanets,” *Ad Hoc Networks*, vol. 24, pp. 85–102, 2015.

[17] C. M. Silva, A. L. Aquino, and W. Meira Jr, “Deployment of roadside units based on partial mobility information,” *Computer Communications*, vol. 60, pp. 28–39, 2015.

[18] F. V. C. Martins, J. F. Sarubbi, and E. F. Wanner, “A multiobjective strategy to allocate roadside units in a vehicular network with guaranteed

levels of service," *International Conference on Evolutionary Multi-Criterion Optimization*, pp. 120–134, 2017.

[19] Y. Ni, J. He, L. Cai, J. Pan, and Y. Bo, "Joint roadside unit deployment and service task assignment for internet of vehicles (iov)," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3271–3283, 2018.

[20] Y. Liu, J. Ma, J. Niu, Y. Zhang, and W. Wang, "Roadside units deployment for content downloading in vehicular networks," *Communications (ICC), 2013 IEEE International Conference on*, pp. 6365–6370, 2013.

[21] K. Gomi, Y. Okabe, and H. Shigeno, "Rsu placement method considering road elements for information dissemination," *VEHICULAR 2017 : The Sixth International Conference on Advances in Vehicular Systems, Technologies and Applications*, pp. 1–6, 2017.

[22] L. Xue, Y. Yang, and D. Dong, "Roadside infrastructure planning scheme for the urban vehicular networks," *Transportation Research Procedia*, vol. 25, pp. 1380–1396, 2017.

[23] Z. Gao, D. Chen, S. Cai, and H.-C. Wu, "Optimal and greedy algorithms for the one-dimensional rsu deployment problem with new model," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7643–7657, 2018.

[24] O. Trullols, M. Fiore, C. Casetti, C.-F. Chiasserini, and J. B. Ordinas, "Planning roadside infrastructure for information dissemination in intelligent transportation systems," *Computer Communications*, vol. 33, no. 4, pp. 432–442, 2010.

[25] J. F. Sarubbi, T. R. Silva, F. V. Martins, E. F. Wanner, and C. M. Silva, "A grasp based heuristic for deployment roadside units in vanets," *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, pp. 369–376, 2017.

[26] N. Nikookaran, G. Karakostas, and T. D. Todd, "Combining capital and operating expenditure costs in vehicular roadside unit placement," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 7317–7331, 2017.

[27] D. Kim, Y. Velasco, W. Wang, R. Uma, R. Hussain, and S. Lee, "A new comprehensive rsu installation strategy for cost-efficient vanet deployment," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 4200–4211, 2016.

[28] H. Xu, "The time after the wei and jin dynasties—city wall and grid pattern," in *Dynamic Interpretation of Early Cities in Ancient China*. Springer, 2021, pp. 19–30.

[29] Z. Shi, S. Hsieh, J. A. Fonseca, and A. Schlueter, "Interdependencies between the design of street grids and the cost-effectiveness of district cooling systems," 2020.

[30] S. Trikariastoto, "Canals to streets: Transformation of jakarta in colonial era."

[31] H. J. Qiu, I. W.-H. Ho, K. T. Chi, and Y. Xie, "A methodology for studying 802.11 p vanet broadcasting performance with practical vehicle distribution," *IEEE transactions on vehicular technology*, vol. 64, no. 10, pp. 4756–4769, 2014.

[32] Q. Zhao, Y. Shi, Q. Liu, and P. Fränti, "A grid-growing clustering algorithm for geo-spatial data," *Pattern Recognition Letters*, vol. 53, pp. 77–84, 2015.

[33] S. Mishra, D. Bhattacharya, and A. Gupta, "Congestion adaptive traffic light control and notification architecture using google maps apis," *Data*, vol. 3, no. 4, p. 67, 2018.

[34] N. Wisitpongphan, F. Bai, P. Mudalige, V. Sadekar, and O. Tonguz, "Routing in sparse vehicular ad hoc wireless networks," *IEEE journal on Selected Areas in Communications*, vol. 25, no. 8, pp. 1538–1556, 2007.

[35] Y. Wang, J. Zheng, and N. Mitton, "Delivery delay analysis for roadside unit deployment in vehicular ad hoc networks with intermittent connectivity," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 8591–8602, 2016.

[36] M.-Y. Yu, R. Vasudevan, and M. Johnson-Roberson, "Occlusion-aware risk assessment for autonomous driving in urban environments," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 2235–2241, 2019.

[37] Y. Li, D. Jin, Z. Wang, L. Zeng, and S. Chen, "Exponential and power law distribution of contact duration in urban vehicular ad hoc networks," *IEEE Signal Processing Letters*, vol. 20, no. 1, pp. 110–113, 2013.

[38] X. Huang, Y. Yin, S. Lim, G. Wang, B. Hu, J. Varadarajan, S. Zheng, A. Bulusu, and R. Zimmermann, "Grab-posisi: An extensive real-life gps trajectory dataset in southeast asia," *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Prediction of Human Mobility*, p. 1–10, 2019.

[39] J. Barrachina, P. Garrido, M. Fogue, F. J. Martinez, J.-C. Cano, C. T. Calafate, and P. Manzoni, "Road side unit deployment: A density-based approach," *IEEE Intelligent Transportation Systems Magazine*, vol. 5, no. 3, pp. 30–39, 2013.

[40] P.-C. Lin, "Optimal roadside unit deployment in vehicle-to-infrastructure communications," *2012 12th International Conference on ITS Telecommunications*, pp. 796–800, 2012.

[41] E. R. Magsino and I. W.-H. Ho, "Roadside unit allocation for fog-based information sharing in vehicular networks," *Proceedings of the 1st ACM International Workshop on Smart Cities and Fog Computing*, pp. 7–12, 2018.

[42] L. B. Jiang and S. C. Liew, "Proportional fairness in wireless lans and ad hoc networks," *IEEE Wireless Communications and Networking Conference, 2005*, vol. 3, pp. 1551–1556, 2005.

[43] R. K. Jain, D.-M. W. Chiu, W. R. Hawe *et al.*, "A quantitative measure of fairness and discrimination," *Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA*, 1984.

[44] I. W.-H. Ho and S. C. Liew, "Impact of power control on performance of ieee 802.11 wireless networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 11, pp. 1245–1258, 2007.

Elmer R. Magsino received the B.Sc. degree in Electronics and Communications Engineering and M.Sc. degree in Electrical Engineering from The University of the Philippines-Diliman in 2002 and 2006, respectively. He completed the Ph.D. degree in Electronic and Information Engineering from The Hong Kong Polytechnic University, Kowloon, Hong Kong, in 2021. Before undertaking the academic career, he served as a design engineer in a power electronics company. He is with the Department of Electronics and Computer Engineering, Gokongwei College of Engineering, De La Salle University-Manila, Philippines where he has also served as the Program Coordinator for the Computer Engineering program from 2013-2015.

Ivan Wang-Hei Ho (M'10–SM'18) received the B.Eng. and M.Phil. degrees in information engineering from The Chinese University of Hong Kong, Hong Kong, in 2004 and 2006, respectively, and the Ph.D. degree in electrical and electronic engineering from the Imperial College London, London, U.K., in 2010. He was a Research Intern with the IBM Thomas J. Watson Research Center, Hawthorne, NY, USA, and a Postdoctoral Research Associate with the System Engineering Initiative, Imperial College London. In 2010, he cofounded P2 Mobile Technologies Ltd., where he was the Chief Research and Development Engineer. He is currently an Associate Professor with the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong. His research interests include wireless communications and networking, specifically in vehicular networks, intelligent transportation systems (ITS), and Internet of things (IoT). He primarily invented the MeshRanger series wireless mesh embedded system, which received the Silver Award in Best Ubiquitous Networking at the Hong Kong ICT Awards 2012. His work on indoor positioning and IoT also received a number of awards, including the Gold Medal at iENA 2019, the Gold Medal with the Organizer's Choice Award at iCAN 2020, and the Gold Medal at the International Exhibition of Inventions Geneva in 2021. He is currently an Associate Editor for the IEEE Access and IEEE Transactions on Circuit and Systems II, and was the TPC Co-Chair for the PERSIST-IoT Workshop in conjunction with ACM MobiHoc 2019 and IEEE INFOCOM 2020.