Edge Server Deployment Approach Based on Uniformity and Centrality

Xinghong Jiang[1], Yong Ma[1](✉), Yunni Xia[2], Qilin Xie[1], and Wenxin Jian[3]

[1] School of Computer and Information Engineering, Jiangxi Normal University, Nanchang 330027, China
{jxh,may,qilinxie}@jxnu.edu.cn
[2] School of Computer Science, Chongqing University, Chongqing, China
[3] School of Digital Industry, Jiangxi Normal University, Shangrao 334000, China
jianwenxin@jxnu.edu.cn

Abstract. In mobile Internet applications that support edge computing, the deployment scheme of edge servers affects the business operation state. Traditional edge servers are deployed on base stations, which do not fully extend the service range of edge servers, resulting in difficult access to edge services. Therefore, this paper proposes the Edge Server Deployment Approach Based on Uniformity and Centrality (ESDA-UC). ESDA-UC considers intersections as candidate deployment locations for edge servers, taking into account traffic density and road network structure. Connection centrality, between centrality, base station centrality, and traffic density are used as the main factors. The intersection centrality of each intersection is calculated as the selection criteria for the deployment location. To avoid concentrating the coverage of edge servers in developed regions of the city, we allocate the number of edge servers according to regional importance. Finally, the improved greedy algorithm is utilized to generate a deployment plan for edge servers. Experiments show that ESDA-UC has higher base station coverage, vehicle coverage, and vehicle coverage time ratios compared to the baseline method.

Keywords: Edge computing · Service scope expansion · Deployment · Intersection centrality

1 Introduction

With the development of artificial intelligence, big data, and other information technologies, business needs based on the mobile Internet continue to emerge. Existing mobile terminals have limited computing and storage capacity. In cases of complex demands, it is necessary to transfer task data to the cloud for processing. However, the cloud computing center is far from the terminal. Network latency can adversely affect real-time tasks. Edge computing introduces storage and computing resources to the user end, greatly reducing the task transmission

H. Gao et al. (Eds.): CollaborateCom 2023, LNICST 561, pp. 93–110, 2024.
https://doi.org/10.1007/978-3-031-54521-4_6

delay. It is an effective means to improve the degree of completion of real-time tasks. The Edge Server (ES) is an important component of edge computing.

There are two ways for mobile devices to get edge services. A multi-hop approach is used [1], where the transmission of tasks is realized through the base station. The other is a single-hop approach, where the mobile device establishes a connection channel directly with the edge server. The former involves issues such as task transmission path selection and transmission delay. Therefore, the method of establishing a transmission path between the mobile device and the ES is the best way to obtain edge services. However, the number of edge servers is sparse, and the coverage is limited. It is worth studying how to expand the scope of edge services with a limited number of ESs.

At present, a large number of scholars have conducted research for the deployment approach of edge service. Moyukh Laha used intersection centrality as intersection gain to find the optimal intersection as the deployment location for ES using the Dynamic Programming (DP) algorithm [2]. But it only considered the vehicle service demand. Chen Xiaoran [3] proposed a road side unit(RSU) deployment mechanism based on intersection prioritization and deployment uniformity. It proposes a uniformity scheme, considering that the deployment scheme based on intersection prioritization leads to overlapping coverage of roadside units. But it still only considers the service demand for vehicles. Meanwhile, its uniform deployment scheme may lead to edge servers being deployed in a few regions. Noting that there is a close relationship between base stations and ES, Zengshi Qin et al. proposed an edge server deployment scheme based on through the improved Top-K algorithm by taking into account the distance between base stations and edge servers, the weight proportion of base stations in the base station cluster, and the upper limit problem of the computational task [4]. The scheme takes the base station as the deployment location of the edge server. It cannot expand the scope of edge services with a limited number of edge servers. Therefore, it is a challenge to expand the area and scenarios of edge services simultaneously in complex urban road networks.

In this paper, we investigate the ES placement problem in the edge service expansion problem and propose the Edge Server Deployment Approach Based on Uniformity and Centrality(ESDA-UC). Firstly, the intersection centrality is calculated by using the intersection information and base station information. Then the intersection traffic density is calculated based on the cab trajectory data. The intersection priority is obtained by combining intersection centrality and traffic density. In order to avoid the service capacity to be clustered only in the developed regions of the city, we use the intersection location information to generate multiple regions by k-means++ clustering [5]. The number of edge servers in each region is calculated by combining the intersection priority. An improved greedy algorithm is utilized to select suitable intersections from the regions as deployment locations for edge servers.

The main contributions of this paper include the following four aspects.

(1) This paper proposes the concept of base station centrality. This mines hot intersections in urban road networks from different perspectives of edge service expansion.
(2) First, the K-means++ algorithm is used to divide the region. Then, the number of edge servers is allocated according to the average value of intersection priority in the region. Avoiding the problem of edge servers being deployed only in developed subregions of the city.
(3) Using a improved greedy algorithm to select intersections for deployment from the subregion to avoid overlapping edge service ranges.

The rest of this paper is organized as follows. We first describe the related work in Sect. 2. Then Sect. 3 presents the basic definitions and the framework underlying the approach of this paper. In Sect. 4, we describe each step of our method in detail. In Sect. 5, we show our dataset and experimental results. Finally, in Sect. 6, we conclude our paper.

2 Related Work

The foundation of edge nodes placement problem starts with comprehensive research available on similar topics such as base station, and road side unit.

Base station placement problem are addressed in [6]. The literature proposes to solve the model by an improved differential evolutionary algorithm. The algorithm has two sets of variation and restart strategies to adapt to different traffic volumes, which can dynamically adjust the base station deployment strategy according to the traffic volume. Dai et al. [7] trained a predictor of received signal strength without propagation model based on machine learning model. The coverage performance of the base station deployment has been optimized by a multi-objective heuristic approach. Literature [8] proposes a strategy for deploying microbase stations in 5G HetNets to obtain high energy efficiency. It optimizes the objective values and makes trade-offs under different user distribution probabilities to improve the adaptability to different user distribution scenarios.

RSU is a communication node installed inside the infrastructure. Ghosh D [9] addresses this trade-off by incorporating a new scheme called Cross-Impact Analysis System for Optimal RSU Deployment (IIA-ORD). The main goal of IIA-ORD is achieved by modeling the transportation network as a connectivity graph and executing a modified K-shell and TOPSIS-based framework. The literature [10] investigates the delay and roadside unit deployment problem for in-vehicle self-organizing networks in cities. It is demonstrated that the roadside unit deployment problem is an uncertain polynomial time hard problem. A binary differential evolution scheme is proposed to maximize the number of roads covered by deployed roadside units. The literature [11] proposes a geometry-based sparse coverage protocol that aims to consider the geometric properties of the road network, the vehicle movement patterns and resource constraints. Sengathir J et al. [12] proposed Honey Badger Optimization Algorithm based

RSU Deployment (HBOA-RSUD) scheme. A multi-objective fitness function is utilized to improve the VANET network coverage.

In recent years, there have been several studies on edge server placement in urban vehicular networks. The literature [13] describes the edge server configuration problem in a smart city mobile edge computing environment as a multi-objective constrained optimization problem. To achieve this, edge servers are placed at some strategic locations. Laha M et al. [2] jointly considered the structural characteristics of the road network, ranked the candidate sites for edge node placement using complex network-based centrality metrics and vehicle traffic distribution of the network. In the literature [14], a more efficient low-energy placement scheme is proposed to formulate the server deployment problem as a multi-objective optimization problem. Luo F et al. [15] proposed a deep reinforcement learning-based edge server placement algorithm. The edge server placement problem is modeled as a Markov decision process, formalized in state space, action space and reward function, and finally solved using a reinforcement learning algorithm to achieve optimal placement. The literature [16] transforms the edge server deployment problem into a minimum dominating set problem in graph theory, and proposes a greedy-based solution algorithm for edge servers whose capacity can be configured on demand, with the key idea of iteratively selecting the nodes that can connect as many nodes as possible under constraints such as delay, degree, and cluster size.

Using base stations directly as the deployment location of edge servers is not able to expand the scope of edge services with a limited number of edge servers. The deployment approach of RSU roadside units only considers the service demand of vehicles. In order to expand the edge service scope and balance the service demand in different regions, this paper designs and optimizes the edge server deployment approach in a two-dimensional city scenario.

3 Preliminaries

We assume that all edge servers have the same communication range. ES can exchange information with vehicles or base stations within the communication range. Dubey B.B. proposed that by using intersections as deployment locations for edge servers, the coverage of edge servers to vehicles will increase by about 15% compared to other locations [3]. The increase in the coverage of edge services results in better coverage continuity for vehicles. Thus, the edge service with better continuity is provided to the vehicles [17]. ES deployed at intersections can also establish connectivity with base stations. In summary, ES deployed at intersections extends the scope and continuity of edge services. Therefore, this paper uses all intersections in the target road network as a set of candidate deployment locations for edge servers. Figure 1 is a schematic diagram of the services provided by ES deployed at the intersection.

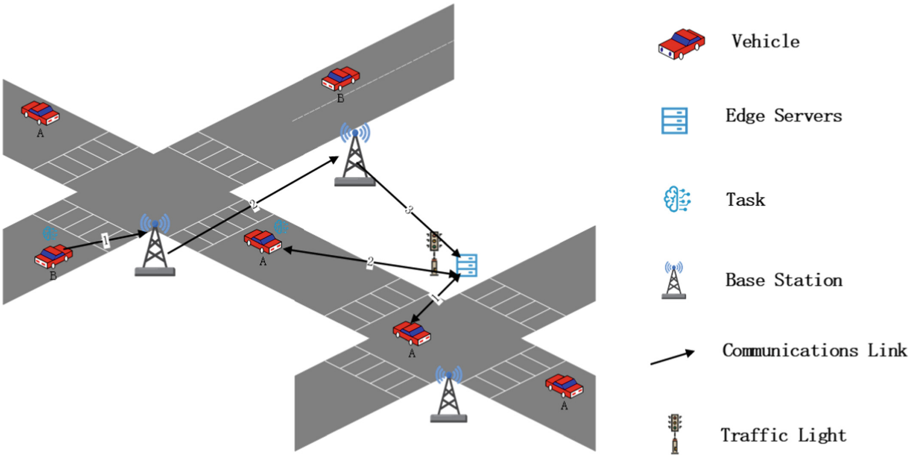


Fig. 1. Schematic diagram of edge services. Car A establishes direct communication links with the ES at both neighboring intersections. Car B establishes an indirect communication link with the ES through the base station.

3.1 Road Model

The urban road network is a complex network system consisting of intersections and the road edges between intersections. $\mathrm{G} = (\mathrm{V}, \mathrm{E}, \mathrm{S})$ represents the road network graph with p intersections, in which V and E denote the intersection set $\mathrm{V} = \{\mathrm{v_1}, \mathrm{v_2}, \ldots, \mathrm{v_p}\}$ and the edge set $\{\mathrm{e_{ij}}\}$ of G. $\mathrm{e_{ij}} = (\mathrm{v_i}, \mathrm{v_j})$ denotes an edge connecting intersection v_i and v_j, and $S \in \{0, 1\}^{p \times p}$ is a adjacency matric storing G's topological structure, in which an entry $S_{ij} = 1$ if $\mathrm{e_{ij}} \in E$; otherwise, $S_{ij} = 0$. Assume that there are s base stations in the city. The set of base stations is denoted as $\mathrm{B} = \{\mathrm{b_1}, \mathrm{b_2}, \ldots, \mathrm{b_s}\}$.

3.2 Coverage Model

Mobile devices can establish an indirect connection to the ES through a base station or a direct connection to the ES. Therefore, both the number of base stations within communication range and intersection traffic flow should be considered when selecting ES deployment locations. The ES deployment problem is represented by the coverage model $C\left(C^R, VE, r, m\right)$, where C^R denotes the set of base stations that can be covered by the deployed ES, VE denotes the number of vehicles that can be covered by the deployed ES, r denotes the communication radius of the ES, and m denotes the number of ESs deployed in the road network.

Typically, the coverage scope of an ES is represented by a circle. The location of the center of the circle is the deployment location of the ES. Denoting the set of intersections where ES is deployed by V^R, v_i^R denotes that ES is deployed at intersection v_i. C_i^R is the set of base stations covered at intersection v_i. VE_i is

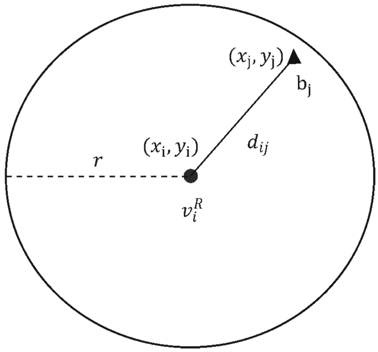


Fig. 2. Coverage model. The black origin in the diagram indicates the ES and the black triangle indicates the base station.

the number of vehicles passing through intersection v_i. As shown in Fig. 2, base station b_j is located within the ES coverage scope at v_i^R. For any base station b_j, the way to determine whether it is within the coverage scope at v_i^R and update C_i^R is as follows:

$$C_i^R = \begin{cases} C_i^R \cup \{b_j\}, & d_{ij} \leqslant r \\ C_i^R, & \text{otherwise} \end{cases}$$

Suppose the coordinates of base station b_j are (x_j, y_j). If base station b_j is in ES coverage scope, then b_j is added to the set C_i^R. where $d_i j$ denotes the distance between v_i and b_j, and is calculated using the Euclidean distance $d_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$.

The set of base station and vehicle coverage for all ESs deployed is $C^R = \mathrm{U}_{i=1}^{m} C_i^R$ and $VE = \mathrm{U}_{i=1}^{m}\mathrm{VE_i}$, respectively. The optimization deployment objective in this paper is to cover as many base stations and vehicles as possible for the determined number of ESs.

3.3 Framework

The simple framework of the methodology of this paper is shown in Fig. 3. The framework consists of four parts: intersection information collection, hotspot intersection mining, region segmentation and ES deployment location selection. Intersection information collection is the first step of the system model. The collected intersection information includes the latitude and longitude of the intersection, vehicle trajectory data, and the latitude and longitude information of the base station. According to the intersection information, the static centrality and traffic density of the intersection are calculated to get the priority of each intersection. Deploying ES at intersections with high priority can obtain a higher coverage scope. In order to avoid ES is only concentrated in a few subregions, we partition the city. When selecting the deployment location within a region, in order to avoid the overlap of ES coverage we use an improved greedy algorithm to generate the final deployment plan.

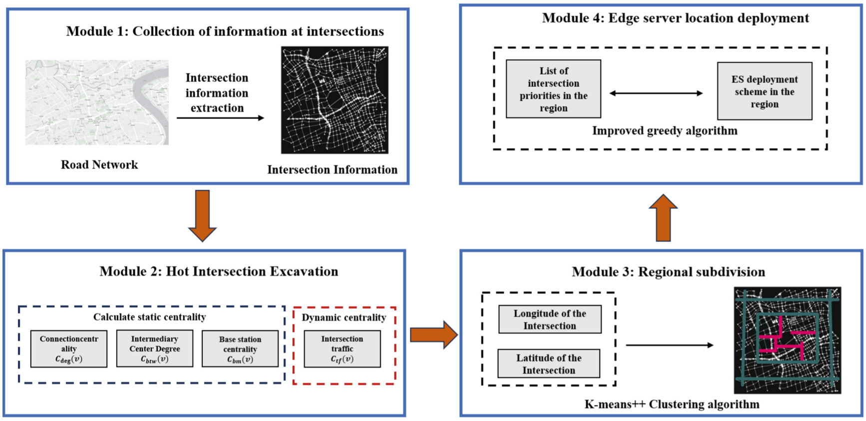


Fig. 3. The framework diagram of the methodology in this paper. The framework consists of four parts: Collection of intersection informations, Hotspot intersection Excavation, Regional subdivision and Edge server location deployment.

4 Edge Server Deployment Approach Based on Uniformity and Centrality

In this section our approach is discussed in detail. Firstly, the hotspot intersection mining algorithm is introduced. Then the edge server deployment scheme is discussed, which consists of two parts: region partitioning and intersection deployment location selection.

4.1 Hot Intersection Excavation

We use centrality [18] to mine urban hotspot intersections. In this section, we discuss four influencing factors for hotspot intersection mining: traffic density, connection centrality, base station centrality, and intermediary centrality.

Traffic Density: The traffic density of an intersection refers to the number of vehicles passing through the intersection per unit time. Edge servers are deployed at an intersection with high traffic density, which improves the number and duration of vehicle coverage for the ES.The traffic density of intersection v_i is denoted as $C_{Tf}(v_i)$. The calculation formula is as follows:

$$C_{Tf}(v_i) = \frac{t(v_i)}{Tl} \tag{1}$$

$t(v_i)$ represents the number of vehicles passing through intersection v_i during a fixed period. Tl represents the length of the unit time.

Degree Centrality: The degree centrality of an intersection refers to the number of intersections directly connected to it by roads. An intersection with a large degree of connectivity centrality has a higher structural importance. Intersections with higher structural importance are more likely to cover more base stations. It is defined as follows:

$$\mathrm{c_{deg}}\left(\mathrm{v_i}\right) = \sum_{j=1}^{p} S_{ij} \tag{2}$$

It is assumed that there are p intersections in the road network.$S \in \{0,1\}^{p\times p}$ is a adjacency matric storing road network topological structure, in which an entry $S_{ij} = 1$ if $\mathrm{e_{ij}} \in E$; otherwise, $S_{ij} = 0$.

Between Centrality: In an urban road network, there are always multiple paths between two destinations. People usually choose the shortest path as the best route. When an intersection appears multiple times in multiple shortest paths, it is considered a central intersection. Using ES at intersections with a high the between centrality improves ES vehicle coverage. It is defined as follows:

$$\mathrm{c_{btw}}\left(\mathrm{v_i}\right) = \sum_{\mathrm{v_s} \neq \mathrm{v_i} \neq \mathrm{v_t}} \frac{\sigma_{\mathrm{s,t}}\left(\mathrm{v_i}\right)}{\sigma_{\mathrm{s,t}}} \tag{3}$$

$\sigma_{\mathrm{s,t}}$ represents the number of shortest paths between any two intersections v_s and v_t. $\sigma_{\mathrm{s,t}}\left(\mathrm{v_i}\right)$ represents the number of shortest paths between intersections v_s and v_t that pass through intersection v_i.

Base Station Centrality: In urban environments, telecom departments tend to deploy more base stations for prosperous subregions, so the number of base stations around an urban intersection reflects the prosperity of that intersection [19]. The higher the prosperity of the intersection, the more suitable as an edge server deployment location. The base station centrality of intersection v_i is denoted as $C_{bm}\left(v_i\right)$ and is calculated as follows:

$$C_{bm}\left(v_i\right) = \frac{\mathrm{m}\left(v_i\right)}{M} \tag{4}$$

$\mathrm{m}\left(v_i\right)$ denotes the number of base stations that can be covered by deploying ES at intersection v_i, while M denotes the total number of base stations.

Traffic density, connection centrality, between centrality, and base station centrality reflect the importance of the intersection in different dimensions. In order to conveniently measure the importance of the four influencing factors, the four influencing factors are normalized in terms of values. The calculation formula is as follows:

$$X = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{5}$$

where X is the data after normalization, x is the source data, $\min(x)$is the minimum value in the source data, and $\max(x)$ is the maximum value in the source data.

After normalization of the influence factors, the four influence factors are considered together to calculate the intersection priority.

$$\begin{aligned} p_{v_i} &= wf \times C_{de.g.}(v_i) + wc \times C_{btw}(v_i) \\ &+ wl \times C_{bm}(v_i) + wb \times C_{Tf}(v_i) \end{aligned} \tag{6}$$

where p_{v_i} is the intersection priority of urban intersection v_i. wf, wc, wl and wb are the weights of connection centrality, between centrality, base station centrality, and traffic density, respectively, all of which take positive values and $wf + wc + wl + wb = 1$. Parameters are set with reference to the deployment realities and are set by the edge server deployers.

4.2 Edge Servers Deployment Rules

There are developed subregions in every city and multiple intersections in this subregion will have higher priority. Selecting intersections for edge servers deployment only from the dimension of intersection centrality will result in edge servers being concentrated in a few subregions. We perform the following steps to improve service coverage and server continuity for edge services. Firstly, the city region is divided into multiple subregion based on the intersection latitude and longitude coordinates. Second, the number of edge servers in the region is obtained based on the mean value of intersection priority in the subregion. Finally, a greedy algorithm is used to select a suitable intersection from the subregion as the deployment location.

Region Division. In this paper, K-means++ clustering is performed using the latitude and longitude location information of the intersections to generate multiple clusters of intersections, and the range of the clusters of intersections is treated as an subregion. As shown in Fig. 4, the k-means++ algorithm adds the initial centroid generation scheme to the k-means algorithm. Assume that s city subregions are generated. We finally output the set of intersections under all city subregions $\mathrm{K} = \{K_1, K_2, \ldots, K_s\}$.

ES Allocation Strategy Under Subregion. In this paper, using the intersection priority, we can calculate the average value of intersection priority between subregions. It is found that due to the large differences in the level of development between urban subregions, the average values of priority among subregions differed significantly. If the number of edge servers is allocated entirely based on the regional intersection priority average value. It will lead to an unbalanced distribution of the number of edge servers. In extreme cases, there will be excess edge service capacity in developed subregions and insufficient or even non-existent edge services capacity in less developed subregions.

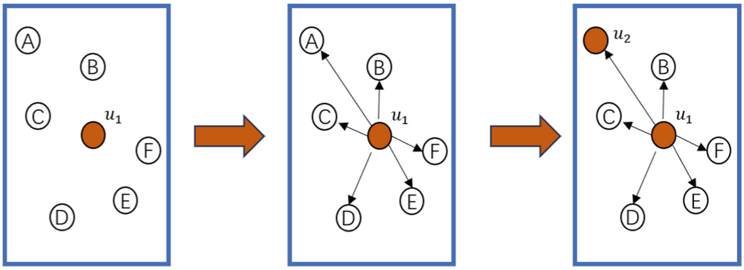


Fig. 4. Schematic diagram of center point selection in the diagram, point u_1 is the center point. u_1 is calculated with the distance from all the remaining points. The point A, which is furthest away, is used as the center point.

In order to improve the utilization rate of edge services. We call a certain percentage of ES, as the number of edge servers that are evenly distributed among all subregions. This part of the ES quantity is called the base number. The ratio of the base number to the total number of ESs is denoted as λ. The remaining ESs are allocated proportionally. Subregions with higher mean intersection priority are allocated more. The specific steps are as follows.

(1) Calculate the base number of subregions to obtain the number of remaining edge servers that denoted as n.
(2) Calculate the average value of intersection priority of each subregion. a_i indicates the average value of intersection priority of the ith subregion. The range of i is $[1, s]$.s denotes the number of urban subregions. p_j indicates the intersection priority of intersection j, and M indicates the number of intersections existing in subregion i. The calculation formula is as follows:

$$a_i = \frac{\sum_{j=1}^{M} p_j}{\text{M}} \tag{7}$$

(3) Calculation of the importance of intersection subregion. A_i denotes the importance of intersection subregion i, and the formula is as follows:

$$A_i = \frac{a_i}{\sum_{j=1}^{s} a_j} \tag{8}$$

(4) The number of edge servers added to the base number in the subregion is calculated based on the intersection subregion importance degree. The calculation formula is as follows:

$$n_i = [\text{n} \times A_i + \epsilon] \tag{9}$$

The number of ES added to subregion i, which are denoted as n_i. Where $[\cdot]$ indicates rounding down. To ensure that ES is not left over, we set ϵ to 0.5.

Allocation Strategy of Edge Servers in the Subregion. The intersections within the subregion are spatially close to each other. When edge server deployment intersections are selected based solely on intersection priority, the resulting service scope may overlap. To avoid this, a greedy algorithm is used in this study to select deployment intersections within the subregion. The specific steps are as follows.

(1) Select an intersection with the highest priority from the intersections in the subregion as the deployment intersection of the edge server.
(2) Remove the deployment intersection and the intersections under its coverage from the region.
(3) Repeat steps (1) (2) until the required number of intersections are identified.

Algorithm 1. ESDA-UC

Require: Urban road network: G=(V,E,S), Intersection latitude and longitude information:(lat, lon).
Ensure: Abnormal sub-regions sets: AR_t
1: Initialize the intersection priority set:$U \leftarrow \varnothing$;
2: Initialize the set of intersections under the city subregion:$K \leftarrow \varnothing$;
3: Initialize the set of intersection means :$A \leftarrow \varnothing$;
4: Initialize a set of deployment intersections for edge servers:$AR \leftarrow \varnothing$;
5: **for** $i = 1$ to $len(V)$ **do**
6: Calculate $C_{Tf}(v_i), C_{btw}(v_i), C_{Tf}(v_i), C_{bm}(v_i)$;
7: $p_{v_i} = wf \times C_{deg}(v_i) + wc \times C_{btw}(v_i) + wl \times C_{bm}(v_i) + wb \times C_{Tf}(v_i)$;
8: $U.\text{append}(p_{v_i})$;
9: **end for**
10: $\text{K} = \{K_1, K_2, \ldots, K_s\} \leftarrow \text{K-means}++(lat, lon)$
11: **for** $i = 1$ to $len(K)$ **do**
12: $a_i = \frac{\sum_{j=1}^{M} p_{v_j}}{\text{M}}$
13: $A \leftarrow A.\text{add}(a_i)$;
14: **end for**
15: **for** $i = 1$ to $len(A)$ **do**
16: $A = \frac{a_i}{\sum_{j=1}^{s} a_j}$
17: $n = [\text{n} \times A + \epsilon]$
18: **for** $t = 1$ to n **do**
19: $k_{\max} = K_i.\text{selectMax}()$
20: **if** $distance(k_{\max}) > coverage$ **then**
21: $AR \leftarrow AR.\text{add}(k_{\max})$;
22: $K_i \leftarrow K_i.\text{remove}(k_{\max})$;
23: **end if**
24: **end for**
25: **end for**
26: **return** AR

Algorithm 1 shows how the three main functional modules are organized in the ESDA-UC framework. Lines 5–9 show the steps of hotspot intersection mining. We first calculate degree centrality, between centrality, base station centrality and traffic density for each intersection. Then, the intersection priority is calculated by fusing the centrality metrics. In line 10, the intersection area is divided according to the intersection information. In lines 11–14, we compute the priority metrics to obtain the region based on the intersection priorities in the region. Lines 15–17, we compute the number of ES deployments in the region based on the region prioritization. Lines 18–24, we discover the intersections deployed with ES from the region using the improved greedy algorithm. Line 26, returns the set of deployed intersections for the edge server.

5 Experiment

To evaluate the effectiveness of our method, we conduct extensive experiments on real datasets. The experiments are run on a workstation equipped with an Intel i7-12700H CPU and an NVIDIA GeForce RTX 3060 Laptop GPU. The experimental code is all implemented in python. We first perform quantitative analysis on the real dataset. In order to enhance the effectiveness of the scheme, optimization experiments were conducted for the relevant parameters.

5.1 Baselines

Uniform-Based Deployment Approach (UDA): UAD [20] uses the k-means++ algorithm to cluster the intersections and generate small regions equal to the number of edge servers. Equal number of edge servers are assigned to each small region.

Edge Server Deployment Approach Based on Based on Degree Centrality (ESDA-DC): ESDA-DC uses the connectivity centrality of intersections as an evaluation metric. Suitable intersections in urban region are selected as deployment locations for edge servers.

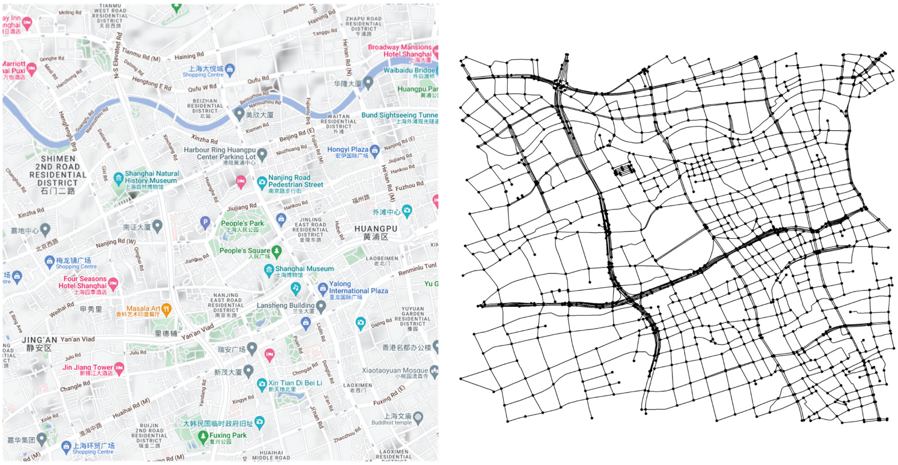
Edge Server Deployment Approach Based on Based on Between Centrality (ESDA-BC): ESDA-BC uses the between centrality of intersections as an evaluation metric. Suitable intersections in urban region are selected as deployment locations for edge servers.

Edge Nodes Placement in 5G Enabled Urban Vehicular Networks: A Centrality-Based Approach(CBA): The Closeness centrality, Betweenness centrality, and traffic flow of the intersection are taken as the deployment benefits, and the deployment budget is taken as the backpack capacity. The dynamic programming method is used to obtain the deployment scheme [2].

Edge Server Placement Method Based on Delay and Energy Awareness (ESPE-DE): An edge server deployment scheme based on improved Top-K algorithm [4].

5.2 Real Cities Dataset

In this paper, a rectangular region of 12 km^2 between 121°45′27″ ∼ 121°49′ east longitude and 31°21′49″ ∼ 31°24′84″ north latitude in Shanghai is selected as the target region. The region contains 1236 vehicle intersections. Figure 5 shows the electronic map and road network diagram of the target region.



(a) Electronic map (b) Road network diagram

Fig. 5. (a) Electronic maps of the target region. (b) A road map of the target region. The black nodes in the map are intersections. Black lines between nodes indicate roads.

The urban data we collected includes location information for each intersection, centrality data, intersection traffic flow data, and base station location information. To obtain the intersection location information and centrality metrics, we first download the geospatial data through OpenStreetMap and then use the OSMnx package [21] to extract the intersection information and calculate the Between Centrality and Degree Centrality for each intersection. We calculate the base station centrality of each intersection based on the intersection location information and the base station location information. The base station location information in the target region is obtained from the real-edge environment dataset Telecom [22–24], which records the latitude and longitude data of 3,233 base stations within the urban region of Shanghai. We utilized the Shanghai cab trajectory data to calculate the traffic density of each intersection in the month of July 2007. The detailed experimental parameter are shown in Table 1.

We use base station coverage, vehicle coverage, and coverage time ratio as metrics to evaluate the effectiveness of the deployment scheme. Vehicle coverage is a direct reflection of the service scope. It refers to the ratio of the number of vehicles covered within the communication range of the edge server to the total number of vehicles. Base station coverage is an indirect reflection of the

Table 1. Experimental parameter setting

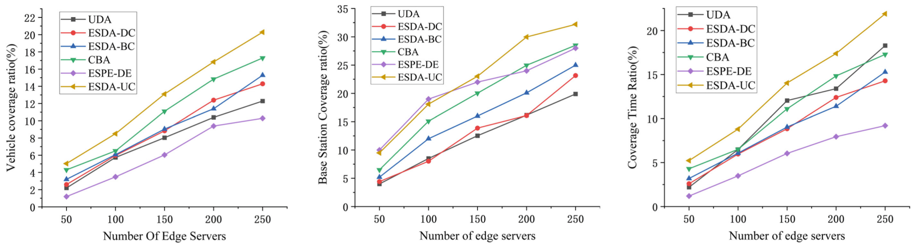
Parameter	Value
Scenario	12 km^2
Number of intersections	1236
$wf/wc/wl/wb$	0.2/0.2/0.3/0.3
Number of junction subregions	5
Number of edge servers	50,100,150,200,250
λ	0.4

edge service range. It refers to the ratio of the number of base stations covered within the communication range of the edge server to the total number of base stations. The coverage time ratio is a reflection of the coverage continuity during vehicle movement. It refers to the percentage of time that a vehicle is located in the ES coverage during its movement. We use the comprehensive coverage ratio and coverage time ratio to evaluate the impact of the number of subregions on the effectiveness of the scheme. We obtain the comprehensive coverage ratio by weighting the sum of base station coverage and vehicle coverage. The weights of both base station coverage and vehicle coverage are 0.5.

5.3 Result for Real Urban Dataset

We first perform quantitative analysis on the real dataset. The communication radius of the edge server is 100 m. The number of edge servers are set as 50, 100, 150, 200 and 250 respectively. The weights of the four influencing factors in the ESDA-UC scheme are set as $wf/wc/wl/wb = 0.2/0.2/0.3/0.3$. The target region is partitioned into 5 sub-regions. The weight of both base station coverage and vehicle coverage in the comprehensive coverage is 0.5. Setting $\lambda = 0.4$ during the experiment.

The experimental results of the methods in this paper and the baseline method are presented in Fig. 6. As the number of ESs increases, all methods improve on every metric. On all the metrics, our method outperforms ESDA-DC and ESDA-BC. This shows that our method is computationally effective for centrality improvement. From the experimental results, our method is not worse than ESPE-DE in terms of base station coverage. Our method outperforms ESPE-DE in terms of vehicle coverage and vehicle coverage time. This proves that the method in this paper is stronger than the base-station-based deployment method in terms of service scenarios and scope. CBA outperforms ESDA-DC and ESDA-BC in each metric. This is due to the fact that CBA integrates the structural characteristics of the road network and the traffic flow characteristics. But the CBA does not overcome the problem of overlapping service scope. Therefore it is weak with our method in all evaluation metrics. In ESDA-DC, ESDA-BC, and CBA, ES is deployed at intersections with high centrality. The probability of obtaining edge service is low when the vehicle leaves the central subregion. This leads to weak edge service coverage continuity. UDA uses a uniform deployment scheme, and vehicle movement characteristics have relatively little impact on coverage continuity. Therefore, UDA outperforms ESDA-DC, ESDA-BC, and CBA in terms of coverage continuity. Our approach considers both region characteristics and intersection centrality. Therefore, it outperforms the baseline approach in coverage continuity.



(a) Performance of 6 deployment schemes on vehicle coverage

(b) Performance of 6 deployment schemes on vehicle coverage

(c) Performance of 6 deployment schemes on coverage time ration

Fig. 6. The performance of this paper's methodology on each of the evaluation indicators.

5.4 Parameter Analysis

This section also verifies the effect of the number of subregions on the comprehensive coverage in the ES number 250 dataset. The experimental results are shown in Fig. 7. From Fig. 7(a), it can be seen that the comprehensive coverage rate shows a significant decrease with the increase in the number of subregions. However, there is a brief rebound in the interval from 15 to 25. As the number of subregions increases, ESs that were originally assigned to hot intersections are assigned to other intersections. This leads to a decrease in the comprehensive coverage rate. The brief recovery of the combined coverage is due to the increase in the number of regions, which elevates the number of ESs at the hotspot intersections. Then the number of regions continues to increase, making the method in this paper similar to the UDA effect. As a result, the comprehensive coverage rate remains stable. From Fig. 7(b), it can be observed that as the number of regions increases, the coverage time first appears to be elevated and then decreases. This is due to the fact that ES is concentrated in developed regions when the number of regions is small. As the number of regions increases, ES is deployed at most of the hotspot subregions. This leads to a decrease in coverage time.

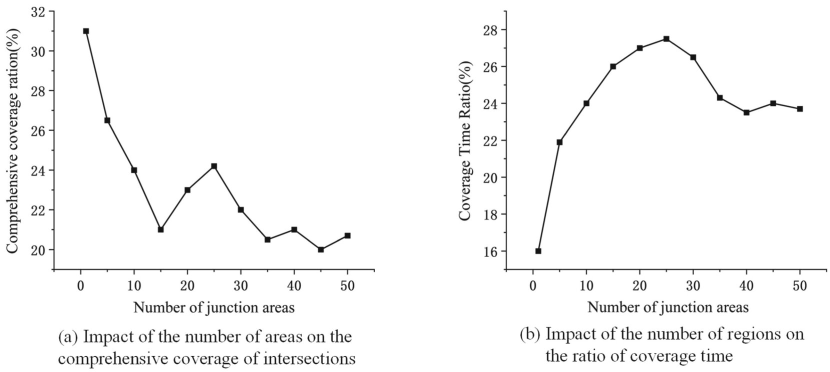


Fig. 7. The effect of the number of regions on the effectiveness of our method

6 Conclusion and Future Work

In this paper, we propose an edge service expansion method based on uniformity and centrality. We find that most intersections in urban environments can realize D2D connectivity with surrounding base stations by collecting base station location information and intersection location information. Based on this finding, in this paper, we take the intersection location as the deployment location of the edge server, and expand the edge service range through the connection between the base station and the edge server. In the specific intersection selection process, we introduce the concept of intersection centrality for discussing the importance degree of intersections. Deploying edge servers only in terms of the importance degree of intersections tends to lead to the concentration of edge coverage in a few regions. Therefore, we first divide the city into sub-regions and allocate the number of edge servers based on the regional importance degree. Finally, the greedy algorithm is utilized to determine the deployment intersections in the entire set of regional intersections.

There exists some extended work in this paper that can be addressed in future work. In this paper, the city is divided into sub-regions in order to avoid concentrating the edge service capacity in the busy subregions of the city. This paper empirically assigns a base number of edge servers to each sub-region, a practice that does not adequately expand the edge service scope. Therefore, in future work, the impact of the value of the base number on the edge service scope needs to be investigated. It should be noted that when mining intersection dynamic centrality in this paper, the intersection dynamic centrality is calculated only from traffic flow due to limited data. However, the dynamic centrality of urban intersections can be mined more deeply by combining the data of intersection pedestrian flow, vehicle flow and urban POI (Point of Interest). Based on the existing data, city managers can fully explore the centrality of intersections to achieve a higher degree of edge service expansion.

Acknowledgment. This work was funded by the Jiangxi Normal University Postgraduate Study Abroad Programme Fund. This work was supported by the Jiangxi Provincial Education Department Postgraduate Innovation Fund Project Grant (YC2022-s351).

References

1. Deng, Y., Chen, Z., Chen, X., et al.: Task offloading in multi-hop relay-aided multi-access edge computing. IEEE Trans. Veh. Technol. **72**(1), 1372–1376 (2022)
2. Laha, M., Kamble, S., Datta, R.: Edge nodes placement in 5G enabled urban vehicular networks: a centrality-based approach. In: 2020 National Conference on Communications (NCC), Kharagpur, India, pp. 1–6 (2020). https://doi.org/10.1109/NCC48643.2020.9056059
3. Chen, X., Tang, X., Chen, W., Chai, M.: Roadside unit deployment mechanism for urban vehicular networks. J. Chin. Comput. Syst. **42**(3), 601–608 (2021)
4. Qin, Z., Xu, F., Xie, Y., et al.: An improved top-K algorithm for edge servers deployment in smart city. Trans. Emerg. Telecommun. Technol. **32**(8), e4249 (2021)
5. Ren, Y.Y., Wang, H., Wang, J.X., et al.: The sub-block demarcation with K-Means++ in each province's interior and establishment analysis of the relative horizontal velocity field model in Mainland China. Chin. J. Geophys. **63**(7), 2516–2533 (2020)
6. Sun, X., Zhang, T., Xu, J., et al.: Energy efficiency-driven mobile base station deployment strategy for shopping malls using modified improved differential evolution algorithm. Appl. Intell. **53**, 1–21 (2022)
7. Dai, L., Zhang, H.: Propagation-model-free base station deployment for mobile net works: integrating machine learning and heuristic methods. IEEE Access **8**, 83375–83386 (2020)
8. Guo, W., Koo, J., Siddiqui, I.F., et al.: QoS-aware energy-efficient MicroBase station deployment for 5G-enabled HetNets. J. King Saud Univ.-Comput. Inf. Sci. **34**(10), 10487–10495 (2022)
9. Ghosh, D., Katehara, H., Rawlley, O., et al.: Artificial intelligence-empowered optimal roadside unit (RSU) deployment mechanism for internet of vehicles (IOV). In: 2022 IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), pp. 495–500. IEEE (2022)
10. Cheng, H., Fei, X., Boukerche, A., et al.: GeoCover: an efficient sparse coverage protocol for RSU deployment over urban VANETs. Ad Hoc Netw. **24**, 85–102 (2015)
11. Sengathir, J., Deva Priya, M.: Christy Jeba Malar A, et al. Honey Badger Optimization Algorithm-Based RSU Deployment for Improving Network Coverage in VANETs. In: Sharma, D.K., Peng, S.L., Sharma, R., Jeon, G. (eds.) ICMETE 2022, pp. 179–193. Springer, Singapore (2023). https://doi.org/10.1007/978-981-19-9512-5_16
12. Wang, S., Zhao, Y., Xu, J., et al.: Edge server placement in mobile edge computing. J. Parallel Distrib. Comput. **127**, 160–168 (2019)
13. Li, Y., Wang, S.: An energy-aware edge server placement algorithm in mobile edge computing. In: 2018 IEEE International Conference on Edge Computing (EDGE), pp. 66–73. IEEE (2018)
14. Luo, F., Zheng, S., Ding, W., et al.: An edge server placement method based on reinforcement learning. Entropy **24**(3), 317 (2022)

15. Zeng, F., Ren, Y., Deng, X., et al.: Cost-effective edge server placement in wireless metropolitan area networks. Sensors **19**(1), 32 (2018)
16. Dubey, B.B., Chauhan, N., Pant, S.: Effect of position of fixed infrastructure on data dissemination in vanets. Int. J. Res. Rev. Comput. Sci. **2**(2), 482 (2011)
17. Landherr, A., Friedl, B., Heidemann, J.: A critical review of centrality measures in social networks. Wirtschaftsinformatik **52**, 367–382 (2010)
18. Kibiłda, J., Galkin, B., DaSilva, L.A.: Modelling multi-operator base station deployment patterns in cellular networks. IEEE Trans. Mob. Comput. **15**(12), 3087–3099 (2015)
19. Kui, X., Du, H., Xiao, X., Li, Y.: Realistic vehicular mobility trace driven RSU deployment scheme. J. Beijing Univ. Posts Telecom **38**(1), 114–118 (2015)
20. Boeing, G.: OSMnx: new methods for acquiring, constructing, analyzing, and visualizing complex street networks. Comput. Environ. Urban Syst. **65**, 126–139 (2017)
21. Li, Y., Zhou, A., Ma, X., et al.: Profit-aware edge server placement. IEEE Internet Things J. **9**(1), 55–67 (2021)
22. Guo, Y., Wang, S., Zhou, A., et al.: User allocation-aware edge cloud placement in mobile edge computing. Softw. Pract. Exp. **50**(5), 489–502 (2020)
23. Wang, S., Guo, Y., Zhang, N., et al.: Delay-aware microservice coordination in MO bile edge computing: a reinforcement learning approach. IEEE Trans. Mob. Comput. **20**(3), 939–951 (2019)
24. Yang, B., Ma, Y., Ma, Z., et al.: The study on key technology of secure access to the resource pool management. J. Jiangxi Normal Univ. Nat. Sci. Ed. (06), 639–643 (2020). https://doi.org/10.16357/j.cnki.issn1000-5862.2020.06.16