A Resource-efficient Placement of Edge Servers for Green Agriculture Consumer Electronics

Yusen Wang, Xiaolong Xu, Ruoshui Wang, Muhammad Bilal, Wei Liu, Guangming Cui*

***Abstract*—Against the backdrop of global carbon neutrality and low-carbon agriculture, the urgency to promote low-carbon agricultural consumer electronics through the integration of sustainable computing is increasingly evident. Edge servers, with their high efficiency and low latency characteristics, have become a crucial component of sustainable computing. Using their local deployment and low-latency advantages, edge servers enable a real-time decision optimization system, optimize energy-efficient resource scheduling, reduce carbon emissions in the agricultural production process, and thereby facilitate low-carbon agriculture. However, for edge computing to deliver efficient, low-latency, and low-energy services, it must rely on the strategic allocation of edge servers. Suboptimal deployment strategies can result in elevated network delays, diminished service reliability, and higher levels of carbon output. The problem of identifying the most effective locations for deploying a limited number of edge servers, while addressing key performance concerns such as latency, reliability, and environmental impact under practical constraints, is commonly known as the *k*ESP problem. Recent research has addressed issues such as high latency, low robustness, and carbon emission reduction in edge computing networks, but has yet to simultaneously reduce latency, improve robustness, and optimize computing resources while lowering carbon emissions. To tackle this challenge, we introduce the *k*ESP-PSO approach, designed to mitigate high latency, enhance service reliability, and reduce carbon emissions by determining an efficient deployment strategy for edge servers. Specifically, *k*ESP-PSO method incorporates a Particle Swarm Optimization (PSO) algorithm, which iteratively refines the location of edge servers based on the spatial distribution of base stations and mobile users across the target region. Through this mechanism, *k*ESP-PSO is capable of theoretically deriving the most effective configuration of edge server placements. Extensive experiments on Melbourne and Shanghai Telecom data sets demonstrate that the proposed method significantly reduces carbon emissions compared to baseline approaches, while also optimizing computing resources and effectively supporting low-carbon agricultural consumer electronics.**

***Index Terms*—Edge server placement, low-carbon agriculture, delocalization, constrained optimization problem, integer programming.**

I. Introduction

THE widespread adoption of agricultural consumer electronics has significantly advanced the intelligence of agricultural production, resulting in higher efficiency and improved quality. In the context of global carbon neutrality and the development of low-carbon sustainable agriculture, integrating sustainable computing into the design of agricultural consumer electronics enables next-generation products to contribute more effectively to low-carbon, green and sustainable agriculture, particularly in terms of energy consumption, efficiency, environmental impact [1], and resource recycling [2]. Edge servers have become a pivotal solution for sustainable computing, attributed to their capability of streamlining computing resources and minimizing carbon footprints [3], [4]. However, achieving these benefits through edge computing [5]–[7] depends on the proper placement of edge servers [8]. Suboptimal placement can result in elevated latency, reduced reliability, and higher carbon emissions within the edge computing network.

The challenge of Edge Server Placement (ESP), which involves deciding where to position edge servers subject to certain constraints, poses a significant and intricate difficulty. Existing studies primarily focus on computation offloading [9], [10], optimizing resource allocation [11], and reducing carbon emissions, such as minimizing mobile user latency [12], [13], maximizing throughput, improving base station load balancing [14], increasing operator profits [15], enhancing data allocation rationality [16], and reducing carbon emissions. Recently, research has begun to explore the k Edge Server Placement (*k*ESP) problem. Given a set of candidate locations for deploying edge servers within a specific region, the main goal of these studies is to minimize both the number of required edge servers and their associated energy consumption [17], while also reducing overall network latency or improving system robustness, all within the bounds of budgetary constraints [18], [19]. However, these methods only consider edge servers serving users based on the signal coverage of their base stations [20], without considering signal transmission between base stations through fiber connections or high carbon emissions.

Existing research on minimizing the total edge server units is deterministic under the scenario where base stations transmit signals to each other. Reducing edge server energy consumption [21], lowering overall network latency [22], and enhancing overall robustness all assume Single Access Edge Server (SAES) scenarios where base stations are independent, information cannot be exchanged, and edge servers rely on the signal coverage of the base station they connect to for service provision. However, in the context of MES (Multi-access Edge

Yusen Wang, Ruoshui Wang and Wei Liu are with the School of Software, Nanjing University of Information Science and Technology, Nanjing 210044, China. E-mail: 202383960014@nuist.edu.cn; WangRsMomo@gmail.com; liuw33997@gmail.com.

Xiaolong Xu and Guangming Cui are with the School of Software and the Jiangsu Province Engineering Research Center of Advanced Computing and Intelligent Services, Nanjing University of Information Science and Technology, Nanjing 210044, China. E-mail: xlxu@nuist.edu.cn; gcui@nuist.edu.cn.

Muhammad Bilal is with the School of Computing and Communications, Lancaster University, United Kingdom. E-mail: m.bilal@ieee.org.

* Guangming Cui is the corresponding author.

Server), the modeling approach changes [23].

Some research has also begun to discuss the challenge of positioning edge servers within the MES context [24], [25]. Considering the quantity of base stations and the connection status between them within a specific range, these studies primarily focus on optimizing server layout [26], reducing information transmission latency [5], improving user experience [27], [28], minimizing server deployment [29], [30], and increasing operator profits [31]. However, most of these studies focus on reducing system latency [32] and do not consider system stability and high carbon emissions [33]. In the MES context, a larger number of users connect to edge servers through base station connections. However, these connections may be disrupted due to weather, connection abnormalities, or other special circumstances. If an edge server is situated at a base station that has only a single connection link to other base stations, and this link fails, the server becomes effectively isolated from the broader edge server network. As a result, it reverts to functioning as a standalone access server. This has a significant impact on the user experience of users connected to this server. The base station housing the server possesses enhanced robustness, the situation of reduced user experience due to base station connection interruptions will be reduced.

To tackle the challenges of high latency, low reliability, and elevated carbon emissions in edge computing networks, this paper proposes an edge server placement model called *k*ESP-PSO, aimed at optimizing the deployment of edge servers. The primary goal of *k*ESP-PSO is to enhance user satisfaction and minimize environmental impact by jointly optimizing transmission delay, handover delay, base station robustness, and the variance between neighboring base stations. The proposed model integrates three core components: the user delay module, the base station robustness module, and the adjacent base station variance module. The user delay module targets the reduction of both transmission and handover delays, with the objective of improving overall user experience through minimized communication latency. The base station robustness module enhances the stability and dependability of the edge network by strengthening the resilience of base stations hosting edge servers. Lastly, the adjacent base station variance module aims to reduce discrepancies between neighboring base stations, promoting a more balanced spatial distribution across the network. This uniformity facilitates effective load balancing among edge servers, ultimately contributing to reduced carbon emissions and optimized resource utilization.

*Summary of Contributions.*To the best of our knowledge, this paper constitutes the initial scholarly examination of the *k*ESP-TSRN issue. The key contributions of this work are outlined below:

- We are the first to consider the balance between transmission delay, switching delay, a variance of nearby base stations, and base station robustness when formulating a *k*ESP strategy.
- We frame the *k*ESP-TSRN issue as a constrained optimization challenge and demonstrate its complexity by relating it to a 0-1 integer linear programming problem.
- We propose a method for solving the *k*ESP-TSRN problem, called *k*ESP-PSO, which utilizes integer programming and Particle Swarm Optimization algorithm.
- An extensive set of experiments is conducted on a widely recognized real-world dataset to evaluate the performance of *k*ESP-PSO in comparison with four established approaches.

II. Related Work

The widespread adoption of agricultural consumer electronics has greatly enhanced the intelligence of agricultural production, leading to increased efficiency and improved quality. However, many of the core technologies in these products lack sufficient emphasis on optimizing computing resources and minimizing carbon emissions. By incorporating sustainable computing into the design of agricultural consumer electronics, the next generation of these products can make a more substantial contribution to low-carbon, green, and sustainable agriculture. This includes advancements in energy efficiency, reduced environmental impact, and improved resource recycling. Edge computing has emerged as a vital approach in sustainable computing, offering the dual benefits of optimizing computing resources and significantly reducing carbon emissions. In recent years, research on the Edge Server Placement (ESP) problem has been scarce. Most studies have focused on reducing the energy consumption of edge servers [31], [32], optimizing server performance [26]–[29], [34]–[36], and improving user experience [18], [37]–[40].As far as we are aware, no studies to date have examined the balance between the performance, robustness, and carbon emissions of edge servers within a MES context.

A. Server Placement Optimization Strategy in Single Access Edge Service (SAES) Environment

In the SAES environment, the optimization of deploying edge servers at base stations can be classified into two primary groups:1) Aiming to position servers to enhance the quality of experience for mobile users by reducing information transmission delay [18]. In particular, Cui et al. [18] set out to identify an edge server placement strategy that maximizes coverage and robustness from the perspective of mobile operators. To tackle this challenge, the authors proposed integer programming algorithms and approximation algorithms to address it.2) Seeking to determine a collection of deployment sites to enhance server functionality [34], [36]. In particular, Zhang et al. [36] focused on tackling the challenge of server placement in the context of online social networking platforms, with consideration for the time expense associated with server-to-server communication. In the context of content distribution networks (CDNs), Huang et al. [34] proposed a methodological approach to pinpoint optimal locations for CDN nodes. This strategy emphasizes enhancing coverage efficiency to reduce end-to-end latency across the network. The above server placement methods are aimed at reducing transmission delay, improving robustness, and optimizing performance. However, these methods are all in the context of ECC, that is, edge servers can only provide services through independent base stations connected to them. However, with the advancement of technology, the concept

of MES has been proposed, that is, edge servers can connect to multiple base stations for data transmission. This way, the server placement method in the SAES environment will not be able to successfully complete the expected goals. Moreover, these methods do not take into account the reduction of carbon emissions in edge service networks.

B. Server Placement Optimization Strategy in Multi-Access Edge Service (MES) Environment

Optimizing edge server placement at base stations in the MES environment can be divided into two primary categories: 1) Optimizing server placement to reduce information transmission delay and improve service quality [28], [29]. Specifically, Mazloomi et al. [29] proposed an RL-based framework to tackle the challenges of edge server deployment and task distribution in Mobile Edge Computing. The algorithm can simultaneously minimize network latency and the number of edge servers. In their work, Zhao et al. [28] explored optimal edge server deployment approaches within mobile edge computing environments, focusing on mitigating signal interference and improving overall service delivery quality. They introduced an approach that utilizes graph partitioning (GP) and the Upper Confidence Bound (UCB) algorith, effectively improving service quality.2) Finding a set of placement locations to reduce server energy consumption [31], [32]. Specifically, Asghari et al. [32] aimed to reduce access delay, improve load balancing, and increase server energy efficiency. They divided the network's geographic region into multiple small regions and used a Coral Reef Optimization (CRO) algorithm for local resource allocation in each region. Li et al. [31] aimed at minimizing access delay while maximizing the profit of the edge service provider. They developed a method grounded in Particle Swarm Optimization (PSO) to optimize profits by introducing a weighting factor q to balance access latency and base station allocation, ensuring access delay while maximizing the profit of the edge provider. However, these methods primarily focus on reducing latency, improving user service quality, and reducing the carbon emissions of edge servers. Nevertheless, they do not address the robustness of the edge server network.

Although previous methods have demonstrated their superiority, they still face limitations in the context of sustainable agricultural consumer electronics. Specifically, they are unable to minimize the carbon footprint of edge servers while simultaneously ensuring their performance and robustness. The method introduced in this study addresses this issue by optimizing the trade-offs between transmission delay, switching delay, base station robustness, and the variance of nearby base stations through the PSO algorithm. This enables the final edge server placement decision to not only reduce carbon emissions but also ensure the performance and robustness of the edge servers.

III. Problem Formulation

This study seeks to identify the best positioning of edge servers by utilizing the location information of base stations and mobile users within a specified range. This approach helps reduce latency, enhance the robustness of edge computing networks, and simultaneously decrease carbon emissions.

The latency of edge computing networks is affected by transmission delay and handover delay. Transmission delay is the communication lag at the user's access base station and the edge server, which affects the user's experience under normal circumstances. Handover delay determines the latency difference incurred when users switch to a different edge server under special circumstances, reflecting the gap in user experience between special and normal situations. The robustness of base stations significantly influences the overall network reliability. Deploying edge servers on highly robust base stations enhances the resilience of the edge computing network accordingly. Carbon emissions are affected by the distribution characteristics of neighboring base stations. An imbalance in the count of base stations connected to an edge server can lead to uneven server loads, resulting in increased power consumption across the edge service network and consequently higher carbon emissions. Given n base stations $B = \{bs_1, \ldots, bs_n\}$, we choose k of these as the deployment sites for edge servers $ES = \{es_1, \ldots, es_k\}$, which corresponds to the server placement decision $p = \{p_1, \ldots, p_n\}$. Based on the coverage zones of the base stations, we choose m mobile users $U = \{u_1, \ldots, u_m\}$. The best approach to solving the *k*ESP-TSRN problem seeks to minimize the trade-off among transmission delay, handover delay, the variance of neighboring base stations, and base station robustness.

$$\begin{aligned} &TSRN(p) \\ &= \sqrt[\alpha]{w_t T(p)^\alpha + w_s S(p)^\alpha + w_r R(p)^\alpha + w_n N(p)^\alpha}. \end{aligned} \tag{1}$$

The optimization objective of the *k*ESP-TSRN problem is to:

$$minimiz\ TSRN(p). \tag{2}$$

Due to the time constraints associated with traditional optimization methods, which often become impractical when dealing with large-scale data, we adopt a Particle Swarm Optimization (PSO) algorithm to iteratively approach the optimal solution. To ensure that edge servers are deployed at base stations, we assign each edge server to its nearest base station after every iteration. Table 1 summarizes the key symbols and notations used throughout this paper.

IV. Method

A. Overview

In this section, we introduce the *k*ESP-PSO method, designed to improve the rationality of edge server placement. This approach aims to reduce latency in edge computing networks, enhance system reliability, and simultaneously lower the carbon emissions of edge service networks. The *k*ESP-PSO method comprises three key components: the user delay module, the base station robustness module, and the nearby base station variance module. The overall workflow of the algorithm is illustrated in Fig. 1.

TABLE I: Key Notations

Notations	Meanings
bs_i	*i*th base station
b_{es_i}	base station preferred edge server es_i
B	limited collection of base stations
$c(bs_i)$	number of base stations connected to base station bs_i
$cov(bs_i)$	coverage area of bs_i
E	set of edges
$E^1_{bs_i}$	preferred edge server for base station bs_i
$E^2_{bs_i}$	alternate edge server for base station bs_i
$l(bs_i, bs_j)$	length between bs_i and bs_j
L	matrix of length between base stations
m	fixed number of mobile users
n	fixed number of base stations
k	bounded set of edge servers
d_i	deployment strategy for base stations bs_i
d	a *k*ESP-TSRN set
es_k	*t*he edge server at position k
S	limited collection of edge servers
$s(es_i)$	number of base stations selecting edge server es_i as their primary edge server
$sl(bs_i, bs_j)$	shortest length between bs_i and bs_j
SL	matrix of shortest length between base stations
u_i	mobile user u_i

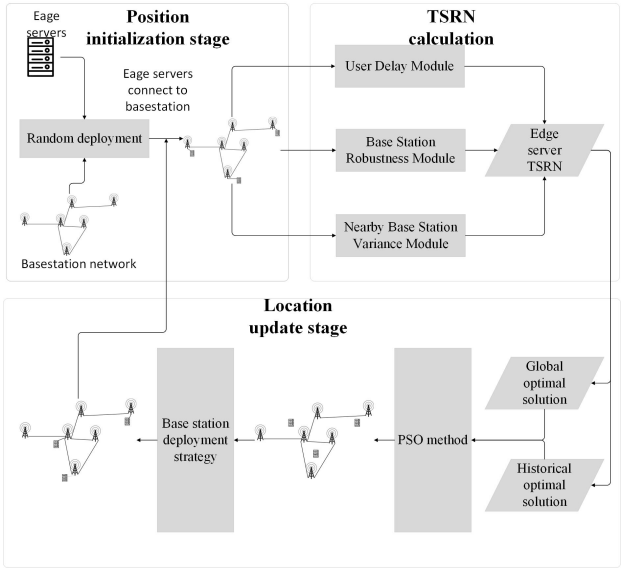

Fig. 1: Flow of *k*ESP-PSO algorithm

B. User Delay Module

In edge computing networks, a mobile user's computation request is initially transmitted to the base station to which they are connected. Subsequently, the base station forwards the data to an edge server via communication links such as fiber optics. A certain amount of delay is introduced during the data transmission between the base station and the edge server. If the distance between them is excessive, the resulting latency can significantly degrade the user experience. The alternative edge server serves as a secondary option, typically the second closest to the base station. When the primary edge server experiences software failures or its computational resources are heavily utilized, preventing it from processing offloading requests, the base station switches to the alternative edge server to ensure service continuity. However, if the alternative edge server is located at a considerable distance from the base station, this may lead to a notable increase in service latency, negatively impacting agricultural consumer electronics that rely on real-time connectivity. Moreover, the increased distance between the base station and the alternative edge server also results in higher energy consumption during signal transmission, thereby contributing to elevated carbon emissions from the edge service network. Therefore, *k*ESP-PSO method uses transmission delay and handover delay to evaluate the user experience. A mobile user u_j can only transmit data to the base station bs_i if u_j is covered by bs_i. The connection between a mobile user and a base station can be represented as:

$$u_i \in cov(bs_i), \forall u_j \in U, \forall bs_i \in B. \tag{3}$$

Base station bs_i can only transmit data if there is a connection with another base station bs_j. The connection between base stations can be represented as:

$$bs_j \in cov(bs_i), \forall bs_i \in U, \forall bs_j \in B, \tag{4}$$

where bs_j is called the connected base station of bs_i.

The set of connected base stations for bs_i is represented as N(bs_i). According to (4), a connection matrix for all base stations can be constructed to represent the connectivity of each base station bs_i to bs_j. Given n base stations $B = \{bs_1, \cdots, bs_n\}$, the connection distance matrix is defined as:

$$L = \begin{bmatrix} l(bs_1, bs_1) & \cdots & l(bs_1, bs_n) \\ \vdots & \ddots & \vdots \\ l(bs_n, bs_1) & \cdots & l(bs_n, bs_n) \end{bmatrix}, \tag{5}$$

where $l(bs_i; bs_j)$ shows the physical separation between the base stations bs_i and bs_j calculated based on their coordinates. $l(bs_i; bs_j) = \infty$ if and only if there is no connection between bs_i and bs_j.

Given n base stations $B = \{bs_1, \cdots, bs_n\}$ and the connection distance matrix L, the shortest distance matrix for base stations is defined as:

$$SL = \begin{bmatrix} sl(bs_1, bs_1) & \cdots & sl(bs_1, bs_n) \\ \vdots & \ddots & \vdots \\ sl(bs_n, bs_1) & \cdots & sl(bs_n, bs_n) \end{bmatrix}, \tag{6}$$

where $sl(bs_i; bs_j)$ represents the shortest path distance between base stations bs_i and bs_j. This distance is calculated by modeling the base stations and their interconnections as an undirected graph structure, and then using the Floyd-Warshall algorithm based on this graph. $sl(bs_i; bs_j) = \infty$ if and only if there is no connecting path between bs_i and bs_j.

The placement decision $d_i \in 0, 1$ for base station bs_i signifies if an edge server is deployed at bs_i. If an edge server is deployed at bs_i, then d_i = 1. Otherwise, d_i = 0.

Based on the shortest distance matrix for base stations and the placement decisions, considering n base stations $B = \{bs_1, \cdots, bs_n\}$, the connection distance matrix L, and the

placement decisions for k servers $d = (d_1, \cdots, d_k)$, the base station-server connection distance matrix is defined as:

$$BS = \begin{bmatrix} s(bs_1, bs_1) \cdot d_1 & \cdots & s(bs_1, bs_n) \cdot d_n \\ \vdots & \ddots & \vdots \\ s(bs_n, bs_1) \cdot d_1 & \cdots & s(bs_n, bs_n) \cdot d_n \end{bmatrix}. \quad (7)$$

Based on the base station–server connection distance matrix, for a given base station bs_i, its transmission delay is determined by the shortest distance to the base station that is connected to the preferred edge server es_1, i.e., $|c(bs_i)|$. The transmission delay of a base station is calculated as follows:

$$T(bs_i) = s(bs_i, E^1_{bs_i}). \quad (8)$$

According to the base station-server connection distance matrix, given a base station bs_i, its handover delay is evaluated through the minimum distance to the base station linked with the alternative edge server es_2 and the minimum distance to the base station linked with the preferred edge server es_1. When the preferred edge server fails, the base station will offload tasks to the alternative server, i.e., $|\mathrm{s}(bs_i)|$. The handover delay of a base station is calculated as follows:

$$S(bs_i) = s(bs_i, E^1_{bs_i}) - s(bs_i, E^2_{bs_i}). \quad (9)$$

Based on the transmission delay of the base stations, given an edge server es_i, the transmission delay of es_i is evaluated by the transmission delay of its neighboring base stations, i.e., $|\mathrm{c}(es_i)|$. The transmission delay of the edge server is calculated as follows:

$$T(es_i) = \sum T(b_{es_i}). \quad (10)$$

Based on the handover delay of the base stations, given an edge server es_i, the handover delay of es_i is evaluated by the handover delay of its neighboring base stations, i.e., $|\mathrm{s}(es_i)|$. The handover delay of the edge server is calculated as follows:

$$S(es_i) = \sum S(b_{es_i}). \quad (11)$$

For a *k*ESP problem with the goal of user delay, the goal is to reduce both transmission latency and handover latency, which can be defined as follows:

$$minimize \sqrt{T(d)^2 + S(d)^2}, \quad (12)$$

C. Base Station Robustness Module

In edge computing networks, edge servers deliver services to mobile subscribers by exchanging data with other base stations through the base stations where they are deployed. In cases where the base station containing an edge server has few connection paths to other base stations and these paths are interrupted due to failures, the edge server will be unable to connect to other base stations and provide edge computing services, resulting in users who originally offloaded data to this server needing to offload to other servers. This not only leads to a degraded user experience but may also cause other edge servers to fail due to insufficient computing resources, potentially leading to the collapse of the entire edge service network. Therefore, base station robustness is an important component of system robustness. In the *k*ESP-PSO method, the number of connection channels of base stations needs to be considered, where the placement of edge servers serves as a metric for assessing the robustness of the edge computing network.

With respect to a base station bs_i, the network robustness at the base station level is assessed by the number of base stations connected to bs_i, calculated as follows:

$$R(bs_i) = C(bs_i). \quad (13)$$

Given an edge server es_i, the network robustness at the edge server level is assessed by the robustness of the base station bs_i where it is deployed, calculated as follows:

$$R(es_i) = \sum R(b_{es_i}). \quad (14)$$

For a *k*ESP problem with the goal of base station robustness, the objective is to maximize the base station robustness, defined as follows:

$$minimiz\ R(d). \quad (15)$$

D. Nearby Base Station Variance Module

In sustainable computing services, carbon emissions are a critical metric. In edge computing networks, carbon emissions are primarily determined by electricity consumption. The power consumption of these networks consists of two main components: one tasked with signal delivery and the other with facilitating the functioning of edge servers delivering services. The power consumption associated with signal transmission is measured by the user latency module, which calculates the spatial distance from the user to the edge server. This distance directly impacts the power consumption during transmission. The energy usage of edge servers is measured by the adjacent base station variance module. Edge servers, even in an idle state without performing any computing tasks, still consume a significant amount of power, which can reach up to 60% of full load power. The computational workload of an edge server is directly influenced by the number of base stations that connect to it. Through the strategic deployment of edge servers, it becomes feasible to balance the distribution of connected base stations across servers, thereby achieving a more efficient allocation of computational resources. To assess the uniformity of this deployment, the adjacent base station variance module quantifies the number of base stations associated with each edge server. This metric facilitates a more equitable distribution of processing tasks among edge servers. Such an approach not only supports effective load balancing, minimizing energy waste caused by underutilized or idle servers, but also reduces handover latency at base stations that may result from server overloading. As a result, the overall robustness of the edge computing network is enhanced, contributing to lower service latency and improved system performance.

According to the base station-server connection distance matrix, the definition of the count of nearby base stations is as follows:

Given n base stations $B = \{bs_1, \cdots, bs_n\}$, the connection distance matrix L, and the placement decisions for k servers

$d = (d_1, \cdots, d_k)$, the count of nearby base stations can be calculated as follows:

$$nn(es_i) = s(es_i) \cdot d_i. \tag{16}$$

Given the edge server placement strategy $d = (d_1, \cdots, d_k)$, the variance of nearby base stations for the edge server is assessed by setting the number of base stations where es_i is located as the preferred base station. The variance of neighboring base stations can be computed as follows:

$$N(es_i) = c(es_i). \tag{17}$$

The mathematical expression for the variance of nearby base stations, which serves as the primary optimization target in the *k*ESP problem,m is defined as follows:

$$minimiz\ N(d). \tag{18}$$

E. Optimization Strategy

In the *k*ESP-PSO framework, the optimal placement of edge servers aims to minimize $TSRN$, where T and S are influenced by the spatial relationship between mobile users and edge servers, while R and N are determined by the spatial distribution of base stations relative to the edge servers. Consequently, the overall value of $TSRN$ is jointly affected by mobile users, base stations, and edge servers. However, given that the positions of mobile users and existing infrastructure are fixed in practical deployment scenarios, the value of $TSRN$ becomes entirely dependent on the selected locations for the edge servers.

In the *k*ESP-PSO approach, we initially relax the constraint that edge servers must be collocated with base stations and instead treat them as independent points distributed across the deployment area. Under this assumption, the resulting $TSRN$ value from a given edge server placement is solely determined by the positions of the edge servers. To address this optimization problem, we model it as a Particle Swarm Optimization (PSO) problem, in which the locations of edge servers are continuously refined to minimize the $TSRN$ metric associated with the placement decision. After each iteration updates the coordinates of the edge servers, each server is then assigned to the nearest base station. The location of this base station becomes the updated position for the corresponding edge server. The $TSRN$ value of an edge server is influenced by its relative spatial relationship with the associated base station. By leveraging the PSO algorithm, it becomes feasible to identify the optimal geographical location for deploying each edge server. Mapping the edge server to the closest base station ensures that its actual deployment location remains as close as possible to the computed optimal position. This strategy effectively maximizes the performance of the overall edge server network in terms of $TSRN$, thereby enhancing system efficiency and service quality.

The $TSRN$ calculation formula generated by the edge server es_i is as follows:

$$\begin{aligned} &TSRN(es_i) \\ &= \sqrt[\alpha]{w_t T(es_i)^\alpha + w_s S(es_i)^\alpha + w_r R(es_i)^\alpha + w_n N(es_i)^\alpha} \end{aligned} \tag{19}$$

The historical locations and $TSRN$ of edge server es_i are recorded in the sets L_{es_i} and R_{es_i}, respectively. L_{es_i} and R_{es_i} are defined as follows:

$$L_{es_i} = \{l^1_{es_i}, \cdots, l^n_{es_i}\}, \tag{20}$$

$$R_{es_i} = \{r^1_{es_i}, \cdots, r^n_{es_i}\}, \tag{21}$$

where $l^n_{es_i}$ represents the location of edge server es_i after the nth location update, and $r^n_{es_i}$ represents the $TSRN$ of edge server es_i after the nth location update.

The locations and historical $TSRN$ for every individual edge server within the set are recorded in the sets L_m and R_m after the location update. L_m and R_m are defined as follows:

$$L_m = \{l^1_m, \cdots, l^n_m\}, \tag{22}$$

$$R_m = \{r^1_m, \cdots, r^n_m\}, \tag{23}$$

where L^n_m represents the position of the nth edge server within the server ensemble after the mth data update, and R^n_m represents the $TSRN$ of the nth edge server within the server ensemble after the mth data update.

The $TSRN$ generated by edge server es_i can be calculated by its transmission delay, handover delay, base station robustness, and nearby base station variance, as follows:

$$kv^{j+1}_{es_i} = wv^j_{es_i} + c_1 r_1 (L_{im} - x^j_{es_i}) + c_2 r_2 (L_{gm} - x^j_{es_i}), \tag{24}$$

where L_{im} represents the global optimal solution, indicating the server location with the minimum $TSRN$ in the current server layout. L_{jm} represents the historical optimal solution, indicating the location of the server with the minimum $TSRN$ during the update process of the location of the edge server.

With the existing edge server layout as the initial solution, the PSO method computes the next-generation positions of edge servers through an iterative updating mechanism. The updated location of each edge server can be formulated as follows:

$$x^{j+1}_{es_i} = x^j_{es_i} + v^{j+1}_{es_i}. \tag{25}$$

where $v^j_{es_i}$ = $L^j_{es_i}$ - $L^{j-1}_{es_i}$, $x^j_{es_i}$ = $L^j_{es_i}$, R_{im} denotes the historical optimal position of edge server bs_i, where the minimum $TSRN$ value was achieved, and R_{gm} represents the location of the edge server that achieved the minimum $TSRN$ after the nth position update. $v^{j+1}_{es_i}$, $v^j_{es_i}$, $R^m_i - x^j_{es_i}$, R_{gm} - $x^j_{es_i}$ are all vectors.

The following formula is applied to allocate each edge server to the nearest base station:

$$x^{j+1}_{es_i} = L(nbs), \tag{26}$$

Constraint Condition:

$$x^{j+1}_{es_i} \neq x^{j+1}_{es_k}, \tag{27}$$

where $L(nbs)$ searches for the base station closest to the updated location of the edge server and checks if it meets the constraint conditions. If the selected base station violates these constraints, the search continues until the most suitable base station is found.

The sets $L^{j+1}_{es_i}$, $R^{j+1}_{es_i}$, L^i_{j+1}, R^i_{j+1} will be updated by the following formulas:

$$L_{es_i}(j+1) = x^{j+1}_{es_i}. \tag{28}$$

Algorithm 1. kESP-PSO

Input: Iteration number R,
base station dataset, user dataset, number of edge servers.
OutputEdge server placement strategy.
1.Initialize space;
2.Randomly assign edge servers to base stations;
3.**while** r<R **do**
4. **for** n edge servers **do**
5. **Calculate** the TSRN of edge servers.
6. **Calculate** the historical optimal solution and the global optimal solution of the edge server.
7. **Update** of each edge server by incorporating its the historical optimal solution, the global optimal solution, the and an inertia-based velocity component.
8. **Update** the position of edge servers according to (25);
9. **Place** edge servers on base stations according to (26);
10. **Update** $L_{es_i}^{j+1}$;
11. **Update** $R_{es_i}^{j+1}$;
12. **Update** L_{j+1}^{i};
13. **Update** R_{j+1}^{i};
14. **end for**
15. r ← r+1;
16.**end while**
17.return L_{j+1}^{i};

$$R_{es_i}(j+1) = TSRN(es_i). \tag{29}$$

$$L_{j+1}(i) = x_{es_i}^{j+1}. \tag{30}$$

$$R_{j+1}(i) = TSRN(es_i). \tag{31}$$

F. Feasibility Analysis

The PSO algorithm converges to the optimal solution on the premise that the expected value of the variance of particle positions reaches a certain location, meaning it converges to the Dirac delta function. Moreover, the second moment of the average velocity tends to zero.

$$\begin{aligned}\mathcal{H}(n) :=& \left(\frac{-w}{2k}\right)^2 \left|\bar{p}_n - \mathbb{E}\left[\bar{p}_n\right]\right|^2 \\ &+ \frac{3}{2}\left|\bar{V}_n\right|^2 + \frac{1}{2}\left(\frac{3c_1}{k} + \frac{(-w)^2}{k^2}\right)\left|\bar{p}_n - \bar{R_{im}}\right|^2 \\ &+ \frac{-w}{2k}\left|\bar{p}_n - \mathbb{E}\left[\bar{p}_n\right], \bar{V}_n\right| + \frac{-w}{k}\left|\bar{p}_n - \bar{R}_{im}, \bar{V}_n\right|,\end{aligned} \tag{32}$$

The last two terms are necessary for the technique. According to the equivalence of Lemma 1, it can be proven that $\mathbb{E}[\mathcal{H}(n)]$ and $\mathbb{E}[|\bar{p}_n - \mathbb{E}[\bar{p}_n]|^2 + |\bar{V}_n|^2 + |\bar{p}_n - \bar{R_{im}}|^2]$ decay at the same rate.

Lemma 1 The decay rate of $\mathcal{H}(n)$ is the same as that of $|\bar{p}_n - \mathbb{E}[\bar{p}_n]|^2 + |\bar{V}_n|^2 + |\bar{p}_n - \bar{R_{im}}|^2$, which indicates

$$\begin{aligned}&\frac{1}{2}\left(\frac{-w}{2k}\right)^2 \left|\bar{p}_n - \mathbb{E}\left[\bar{p}_n\right]\right|^2 + \frac{1}{2}\left|\bar{V}_n\right|^2 + \frac{3c_1}{2k}\left|\bar{p}_n - \bar{R}_{im}\right|^2 \\ &\le \mathcal{H}(n) \\ &\le \frac{5}{2}\left(\left(\frac{-w}{2k}\right)^2 + 1 + \frac{3c_1}{k} + \frac{-2w^2}{k^2}\right) \\ &\quad \cdot \left(\left|\bar{p}_n - \mathbb{E}\left[\bar{p}_n\right]\right|^2 + \left|\bar{V}_n\right|^2 + \left|\bar{p}_n - \bar{R}_{im}\right|^2\right).\end{aligned} \tag{33}$$

Lemma 2 The position that minimizes $TSRN$ is $\bar{p}_n, \bar{R}_{im}, \bar{V}_n$. Then $\mathcal{H}$, as defined in (32), satisfies

$$\begin{aligned}&\frac{d}{dn}\mathbb{E}[\mathcal{H}(n)] \le \frac{w}{2k}\mathbb{E}\left[\left|\bar{V}_n\right|^2\right] \\ &- \left(\frac{-\left(c_1+2c_2\right)w}{(2k)^2} - \left(\frac{9c_2^2}{-wk} + \frac{-3c_1w}{(2k)^2}\right)\frac{6e^{-\alpha\mathcal{E}}}{\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{Y}_n\right)\right)\right]}\right) \\ &\cdot \mathbb{E}\left[\left|\bar{p}_n - \mathbb{E}\left[\bar{p}_n\right]\right|^2\right] \\ &- \left(\frac{-\left(c_1+c_2\right)w}{k^2} + \kappa\theta\left(\frac{3c_1}{k} + \frac{(-w)^2}{k^2}\right) + \frac{8\kappa^2 w}{k}\right. \\ &\left. + \frac{c_2^2 w}{2k^2c_1} - \frac{9c_2^2}{-wk} - \left(\frac{9c_2^2}{-wk} - \frac{3c_1w}{(2k)^2}\right)\frac{12e^{-\alpha\mathcal{E}}}{\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{Y}_n\right)\right)\right]}\right) \\ &\cdot \mathbb{E}\left[\left|\bar{p}_n - \bar{Y}_n\right|^2\right].\end{aligned} \tag{34}$$

Proof: Let $\delta\bar{p}_n := \bar{p}_n - \mathbb{E}\left[\bar{p}_n\right]$. We can obtain

$$\frac{d}{dn}\mathbb{E}\left[\left|\delta\bar{p}_n\right|^2\right] = 2\mathbb{E}\left[\left\langle\delta\bar{p}_n, \bar{V}_n\right\rangle\right]. \tag{35}$$

By applying Itô's formula and Young's inequality, we can deduce

$$\begin{aligned}&\frac{d}{dn}\mathbb{E}\left[\left|\bar{V}_n\right|^2\right] \\ &\le -\left(\frac{-2w}{k} - \frac{c_2}{\varepsilon k}\right)\mathbb{E}\left[\left|\bar{V}_n\right|^2\right] \\ &\quad + \frac{\varepsilon c_2}{k}\mathbb{E}\left[\left|y_\alpha\left(R_{gm}\right) - \bar{p}_n\right|^2\right] \\ &\quad - \frac{2c_1}{k}\mathbb{E}\left[\left[\bar{V}_n, \bar{p}_n - \bar{Y}_n\right|\right], \quad \forall\varepsilon > 0.\end{aligned} \tag{36}$$

Using Itô's formula, we can conclude

$$\begin{aligned}&\frac{d}{dn}\mathbb{E}\left[\left\langle\delta\bar{p}_n, \bar{V}_n\right\rangle\right] \\ &\le \mathbb{E}\left[\left|\bar{V}_n\right|^2\right] \\ &\quad + \frac{w}{2k}\frac{d}{dn}\mathbb{E}\left[\left|\delta\bar{p}_n\right|^2\right] - \frac{c_1+2c_2}{2k}\mathbb{E}\left[\left|\delta\bar{p}_n\right|^2\right] \\ &\quad + \frac{c_1}{2k}\mathbb{E}\left[\left|\bar{R}_{im} - y_\alpha\left(R_{gm}\right)\right|^2\right],\end{aligned} \tag{37}$$

In the second line, when C is a constant, we can use $\mathbb{E}\left[\langle\delta\bar{p}_n, C\rangle\right] = 0$ as well as Eq. (34) to expand the expression. After rearranging the inequality, we get the following result

$$\begin{aligned}&\frac{-w}{2k}\frac{d}{dn}\mathbb{E}\left[\left|\delta\bar{p}_n\right|^2\right] + \frac{d}{dn}\mathbb{E}\left[\left\langle\delta\bar{p}_n, \bar{V}_n\right\rangle\right] \\ &\le \mathbb{E}\left[\left|\bar{V}_n\right|^2\right] - \frac{c_1+2c_2}{2k}\mathbb{E}\left[\left|\delta\bar{p}_n\right|^2\right] \\ &\quad + \frac{c_1}{2k}\mathbb{E}\left[\left|\bar{R}_{im} - y_\alpha\left(R_{gm}\right)\right|^2\right].\end{aligned} \tag{38}$$

Applying Itô's formula again, we get

$$\begin{aligned}&\frac{d}{dn}\mathbb{E}\left[\left|\bar{p}_n - \bar{R}_{im}\right|^2\right] \\ &\le 2\mathbb{E}\left[\left(\bar{p}_n - \bar{R}_{im}, \bar{V}_n\right)\right] - 2\kappa\theta\mathbb{E}\left[\left|\bar{p}_n - \bar{R}_{im}\right|^2\right].\end{aligned} \tag{39}$$

Finally, based on Itô's formula and Young's inequality, we can conclude the following constraint

$$\begin{aligned}&\frac{d}{dn}\mathbb{E}\left[\left\langle \bar{p}_n-\bar{R}_{im},\bar{V}_n\right\rangle\right]\\&\leq-\left(\frac{c_1+c_2}{k}-8\kappa^2-\frac{c_2^2}{2k\lambda_1}\right)\mathbb{E}\left[\left|\bar{p}_n-\bar{R}_{im}\right|^2\right]\\&+\frac{3}{2}\mathbb{E}\left[\left|\bar{V}_n\right|^2\right]+\frac{c_1}{2k}\mathbb{E}\left[\left|y_\alpha\left(R_{gm}\right)-\bar{R}_{im}\right|^2\right]\\&+\frac{w}{k}\mathbb{E}\left[\left\langle \bar{p}_n-\bar{R}_{im},\bar{V}_n\right\rangle\right].\end{aligned}\tag{40}$$

Therefore, we have

$$\begin{aligned}&\frac{d}{dn}\mathbb{E}[\mathcal{H}(n)]\\&\leq-\left(\frac{-w}{k}-\frac{3c_2}{2\varepsilon k}\right)\mathbb{E}\left[\left|\bar{V}_n\right|^2\right]+\frac{\left(c_1+2c_2\right)w}{(2k)^2}\mathbb{E}\left[\left|\delta\bar{p}_n\right|^2\right]\\&-\left(\frac{-\left(c_1+c_2\right)w}{k^2}+\frac{8\kappa^2w}{k}+\frac{c_2^2w}{2k^2c_1}\right.\\&+\kappa\theta\left(\frac{3c_1}{k}+\frac{w^2}{k^2}\right)-3\frac{\varepsilon c_2}{k}\mathbb{E}\left[\left|\bar{p}_n-\bar{R}_{im}\right|^2\right]i\\&\left.+\left(3\frac{\varepsilon c_2}{k}+\frac{3c_1\gamma}{(2k)^2}\right)\right)\mathbb{E}\left[\left|y_\alpha\left(R_{gm}\right)-\bar{R}_{im}\right|^2\right].\end{aligned}$$

Lemma 3 By choosing PSO parameters, we ensure that the coefficients of $\cdot\mathbb{E}\left[\left|\bar{p}_n-\mathbb{E}\left[\bar{p}_n\right]\right|^2\right]$ and $\cdot\mathbb{E}\left[\left|\bar{p}_n-\bar{Y}_n\right|^2\right]$ in inequality (34) are negative, thus achieving exponential decay of E[H(t)].

$$\begin{aligned}&c_1>0,c_2>6\max\left\{\frac{D_t^Yc_1}{4},0\right\},\\&\kappa>-\frac{3c_2^2\left(1+D_t^Y\right)}{w\theta c_1},\\&\text{and }m<\min\left\{-\frac{w\theta}{16\kappa},\frac{c_1w^2}{18D_t^Yc_2^2}\right\},\end{aligned}\tag{41}$$

Where we reduce $D_t^Y=12e^{-\alpha\underline{\mathcal{E}}}/\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{R}_{im}\right)\right)\right]$.

Theorem 1 Assume the initial data meets the requirements. P1 $\mu_1>0$ with

$$\begin{aligned}\mu_1:=&-\frac{\left(c_1+2c_2\right)w}{(2k)^2}\\&+\left(\frac{9c_2^2}{wk}+\frac{3c_1w}{4k^2}\right)\frac{12e^{-\alpha\underline{\mathcal{E}}}}{\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{R}_{im}\right)\right)\right]}.\end{aligned}\tag{42}$$

P2 $\mu_2>0$ with

$$\begin{aligned}\mu_2:=&-\frac{\left(c_1+c_2\right)w}{k^2}+\kappa\theta\left(\frac{3c_1}{k}+\frac{w^2}{k^2}\right)+\frac{8\kappa^2w}{k}\\&+\frac{c_2^2w}{2k^2c_1}+\frac{9c_2^2}{wk}\\&+\left(\frac{9c_2^2}{wk}+\frac{3c_1w}{(2k)^2}\right)\frac{24e^{-\alpha\underline{\mathcal{E}}}}{\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{R}_{im}\right)\right)\right]}.\end{aligned}\tag{43}$$

P3 it holds

$$\begin{aligned}&\left(\frac{\alpha\kappa k}{c_1\chi}\left(C_{\mathcal{E}}+2\alpha^2\right)+\frac{24C_{\mathcal{E}}^2\kappa}{\alpha\chi^3}\right)\\&\frac{\mathbb{E}[\mathcal{H}(0)]}{\mathbb{E}\left[\exp\left(-\alpha\left(\mathcal{E}\left(\bar{R}_{im}\right)-\underline{\mathcal{E}}\right)\right)\right]}\\&+\frac{6\kappa}{\alpha\chi}\frac{\mathbb{E}\left[\left|\nabla\mathcal{E}\left(\bar{p}_0\right)\right|^2\right]}{\mathbb{E}\left[\exp\left(-\alpha\left(\mathcal{E}\left(\bar{R}_{im}\right)-\underline{\mathcal{E}}\right)\right)\right]}<\frac{3}{32},\end{aligned}\tag{44}$$

where

$$\chi:=\frac{2}{5}\frac{\min\left\{-w/(2k),\mu_1,\mu_2\right\}}{\left((-w/(2m))^2+1+3c_1/m+2(-w/m)^2\right)}.\tag{45}$$

$E[H(t)]$ converges to 0 exponentially fast when t tends to positive infinity.

Proof: First, define the time range

$$\begin{aligned}T:=&\inf\{t\geq0:\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{R}_{im}\right)\right)\right]\\&<\frac{1}{2}\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{R}_{im}\right)\right)\right]\}\quad\text{with }\inf\emptyset=\infty.\end{aligned}\tag{46}$$

Since continuity implies $N>0$, we assert $N=\infty$, and then we will prove this by contradiction.

$$\begin{aligned}&\frac{d}{dn}\mathbb{E}[\mathcal{H}(n)]\\&\leq-\frac{2}{5}\frac{\min\left\{-w/(2k),\mu_1,\mu_2\right\}}{\left((-w/(2k))^2+1+3c_1/k+2w^2/k^2\right)}\mathbb{E}[\mathcal{H}(n)]\\&=:-\chi\mathbb{E}[\mathcal{H}(n)],\end{aligned}\tag{47}$$

In the inequality, we used the upper bound from Lemma 4. The rate χ is implicitly defined, and $0<\chi<-\frac{w}{m}$, where χ is a positive number derived from preparatory conditions p_1, p_2.

$$\mathbb{E}[\mathcal{H}(n)]\leq\mathbb{E}[\mathcal{H}(0)]\exp(-\chi t).\tag{48}$$

Begin to study the evolution of the function $\mathcal{Y}(n)=\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{R}_{im}\right)\right)\right]$. According to Itô's formula

$$\begin{aligned}&\frac{d}{dt}\mathcal{Y}(t)\geq-4\alpha\kappa e^{-\alpha\underline{\mathcal{E}}}C_{\mathcal{E}}\\&\cdot\mathbb{E}\left[\left|\bar{p}_n-\bar{R}_{im}\right|^2\right]-4\alpha\kappa e^{-\alpha\underline{\mathcal{E}}^E\left[\left|\nabla\mathcal{E}(\bar{p}_n)\right|\left|\bar{p}_n-\bar{R}_{im}\right|\right]},\end{aligned}\tag{49}$$

By Young's inequality, we note that $\mathbb{E}\left[\left|\nabla\mathcal{E}\left(\bar{p}_n\right)\right|\left|\bar{p}_n-\bar{R}_{im}\right|\right]\leq e^{(\chi/2)t}\alpha^2\mathbb{E}\left[\left|\bar{p}_n-\bar{R}_{im}\right|^2\right]+e^{-(\chi/2)t}/\alpha^2\mathbb{E}\left[\left|\nabla\mathcal{E}\left(\bar{p}_n\right)\right|^2\right]$. Therefore, we have

$$\begin{aligned}&\mathbb{E}\left[\left|\nabla\mathcal{E}\left(\bar{p}_n\right)\right|^2\right]\\&=\mathbb{E}\left[\left|\nabla\mathcal{E}\left(\bar{p}_0\right)+\int_0^n\nabla^2\mathcal{E}\left(\bar{p}_s\right)\bar{V}_sds\right|^2\right]\\&\leq2\mathbb{E}\left[\left|\nabla\mathcal{E}\left(\bar{p}_0\right)\right|^2\right]+4C_{\mathcal{E}}^2t\mathbb{E}[\mathcal{H}(0)]\int_0^n\exp(-\chi s)ds\\&=2\mathbb{E}\left[\left|\nabla\mathcal{E}\left(\bar{p}_0\right)\right|^2\right]+4C_{\mathcal{E}}^2t\mathbb{E}[\mathcal{H}(0)]\frac{1}{\gamma}(1-\exp(-\chi t)).\end{aligned}\tag{50}$$

The penultimate step uses the explicit solution in Eq. (50). Combining the above observations, we can obtain

$$\begin{aligned}&\frac{d}{dn}\mathcal{Y}(n)\\&\geq -4\alpha\kappa e^{-\alpha\underline{\mathcal{E}}}\left(C_{\mathcal{E}}\exp(-\chi n)+\exp\left(-\frac{\chi}{2}n\right)\alpha^2\right)\frac{2k}{3c_1}\mathbb{E}[\mathcal{H}(0)]\\&-\frac{4}{\alpha}\kappa e^{-\alpha\underline{\mathcal{E}}}\exp\left(-\frac{\chi}{2}n\right)\left(2\mathbb{E}\left[\left|\nabla\mathcal{E}\left(\bar{p}_0\right)\right|^2\right]+\frac{4C_{\mathcal{E}}^2 n}{\chi}\mathbb{E}[\mathcal{H}(0)]\right).\end{aligned} \tag{51}$$

By integrating Eq. (51), we know that t belongs to $[0, T]$

$$\begin{aligned}&\mathcal{Y}(n)\\&\geq \mathcal{Y}(0)-4\alpha\kappa e^{-\alpha\underline{\mathcal{E}}}\left(\frac{C_{\mathcal{E}}}{\chi}+\frac{2\alpha^2}{\chi}\right)\frac{2k}{3c_1}\mathbb{E}[\mathcal{H}(0)]\\&-\frac{4}{\alpha}\kappa e^{-\alpha\underline{\mathcal{E}}}\left(2\mathbb{E}\left[\left|\nabla\mathcal{E}\left(\bar{p}_0\right)\right|^2\right]\frac{2}{\chi}+\frac{16C_{\mathcal{E}}^2}{\chi^3}\mathbb{E}[\mathcal{H}(0)]\right).\end{aligned} \tag{52}$$

Reviewing the definition of Y and applying condition p_3, we infer that $\forall n \in [0, N]$, the following holds

$$\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{R}_{im}\right)\right)\right] > \frac{3}{4}\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{R}_{i0}\right)\right)\right], \tag{53}$$

This means there exists $\delta > 0$ such that in $[N, N+\delta]$ we have $\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{R}_{im}\right)\right)\right] \geq \mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{R}_{i0}\right)\right)\right]/2$, which contradicts the definition of N, hence $N = \infty$.

According to Eq. (48), we can deduce

$$\begin{aligned}&\mathbb{E}[\mathcal{H}(n)] \leq \mathbb{E}[\mathcal{H}(0)]\exp(-\chi n)\\&\text{and}\quad \mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{R}_{im}\right)\right)\right] \geq \frac{1}{2}\mathbb{E}\left[\exp\left(-\alpha\mathcal{E}\left(\bar{Y}_0\right)\right)\right].\end{aligned} \tag{54}$$

For some constant $C > 0$, we get

$$\begin{aligned}\mathbb{E}\left[\left|\bar{p}_n-\mathbb{E}\left[\bar{p}_n\right]\right|^2\right] &\leq C\exp(-\chi n),\\ \mathbb{E}\left[\left|\bar{V}_n\right|^2\right] &\leq C\exp(-\chi n),\\ \text{and } \mathbb{E}\left[\left|\bar{p}_n-\bar{R}_{im}\right|^2\right] &\leq C\exp(-\chi n).\end{aligned} \tag{55}$$

By Jensen's inequality

$$\begin{aligned}\left|\frac{d}{dt}\mathbb{E}\left[\bar{p}_n\right]\right| &\leq \mathbb{E}\left[\left|\bar{V}_n\right|\right]\\ &\leq C\exp(-\chi n/2)\to 0\\ &\text{as } n\to\infty.\end{aligned} \tag{56}$$

Therefore, we conclude that when n is sufficiently large, meaning when the number of iterations is large enough, p can reach the optimal solution. Thus, the PSO algorithm is capable of finding the optimal placement of edge servers.

V. Performance Evaluation

This section presents a comprehensive evaluation of the performance of *k*ESP-PSO through experiments carried out using popular real-world data sets.

TABLE II: Experimental Settings

Data Sets		K	N	M
EUA Datasets	Set#1.1	4	5,10,∼,30,35	80
	Set#1.2	2,3,∼,8,9	20	80
	Set#1.3	4	20	20,40,∼,140,160
Telecom Shanghai Dataset	Set#1.1	4	5,10,∼,30,35	80
	Set#1.2	2,3,∼,8,9	20	80
	Set#1.3	4	20	20,40,∼,140,160

A. Data Introduction

This study utilizes two datasets: EUA datasets [41], [42] containing data on 1,465 actual base stations located in the urban area of Melbourne, and Telecom Shanghai Dataset [31], [43], [44] with data on 1,417 base stations in urban Shanghai. Base station coverage areas are categorized as large, medium, and small based on mobile user density within their range, with corresponding radii of 8,000 to 7,000 meters, 3,000 to 2,000 meters, and 430 to 750 meters. Each base station's coverage area is randomly selected within its category. Base station connections are established based on coverage areas, with a connection established between two base stations if one is within the other's coverage area. Base stations need to be connected to other base stations to ensure normal data transmission. If a base station has a small number of external connections, some base stations may interconnect, forming a closed network, which contradicts the background of our experiment. Therefore, we set a rule that if a base station has fewer connections than one-seventh of the total number of selected base stations, it will connect to the nearest base station to meet the minimum requirement. In reality, base stations located in city centers have a higher deployment density, serve a larger number of users, and pose a greater challenge for edge server placement algorithms. The experiment selected data from 125 base stations in Melbourne's central business district and 164 base stations in Shanghai's urban area based on their coordinate information.

B. Parameter Settings

To evaluate the effectiveness of these nine methods, three parameters were varied: 1) the count of base stations used for positioning edge servers, as changes in the number of base stations directly affect the choice of edge server placement locations, which in turn influences decisions on the process of deploying edge servers; 2) the total number of edge servers, which affects the distribution of edge server coverage areas; In scenarios characterized by a constrained quantity of edge servers, how to allocate them reasonably so that each user can access sufficient computing resources poses a challenge for server placement algorithms; 3) The quantity of mobile device users, which determines the load on each edge server; The allocation strategy for edge servers under low user density differs significantly from that required in high-demand scenarios, necessitating an adaptive placement algorithm capable of making intelligent and context-aware resource distribution decisions. The weights for transmission

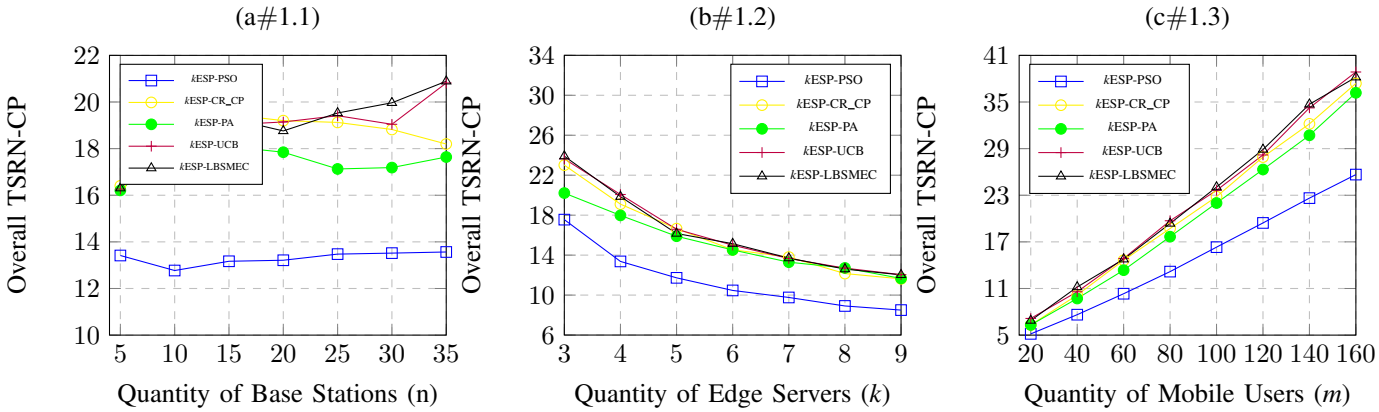

Fig. 2: Assessment of Effectiveness under EUAdatasets experiments (Set #1).

delay, handover delay, base station robustness, and variance of nearby base stations were set to be equal, i.e., $w_1 = w_2 = w_3 = w_4 = 1/4$. Based on the summary of previous experiments, when the particle swarm iteration steps reach 800, a good balance between computation time and result quality can be achieved. Therefore, each experiment runs the particle swarm optimization for 800 iterations. To ensure statistical reliability, every experiment is executed repeatedly for 100 runs, and the reported result represents the mean value across all trials.

C. Comparison Experiments

The experiments compare *k*ESP-PSO with four other methods to evaluate their effectiveness: 1) a method compromising between base station coverage and robustness (*k*ESP-CR_CP), 2) a method maximizing the overall profit of edge servers (*k*ESP-PA), 3) a method compromising between transmission delay, throughput, and user density (*k*ESP-UCB), and 4) a method compromising between edge server coverage area and edge server overlap area (*k*ESP-LBSMEC).

· *k*ESP-CR_CP [18]: Considers a compromise between edge server coverage and robustness.

· *k*ESP-PA [31]: Considers maximizing the profit of edge servers.

· *k*ESP-UCB [28]: Considers a compromise between transmission delay, throughput, and user density.

· *k*ESP-LBSMEC [45]: Considers a compromise between edge server coverage area and edge server overlap area.

Figure 2 compares the effectiveness of *k*ESP-PSO with *k*ESP-CR_CP, *k*ESP-PA, *k*ESP-UCB, and *k*ESP-LBSMEC in the experiment set #1. Overall, *k*ESP-PSO achieves the minimum TSRN-CP in all cases. *k*ESP-PSO has an average advantage of 39.8081%, 32.39304%, 44.16238%, and 44.64895% over *k*ESP-CR_CP, *k*ESP-PA, *k*ESP-UCB, and *k*ESP-LBSMEC, respectively, over the course of experiment sets #1.1, #1.2, and #1.3. In detail, as illustrated in Figure 2(a), within experiment set #1.1, *k*ESP-PSO has advantages of 39.84406%, 31.3656%, 42.41398%, and 42.99296% over *k*ESP-CR_CP, *k*ESP-PA, *k*ESP-UCB, and *k*ESP-LBSMEC, respectively. As shown in Figure 2(b), within experiment set #1.2, *k*ESP-PSO has advantages of 37.90927%, 30.57407%, 42.30149%, and 42.4766% over *k*ESP-CR_CP, *k*ESP-PA, *k*ESP-UCB, and *k*ESP-LBSMEC, respectively. As shown in Figure 2(c), within experiment set #1.3, *k*ESP-PSO, ESP-SR has advantages of 41.38031%, 34.83816%, 47.28061%, and 47.94681% over *k*ESP-CR_CP, *k*ESP-PA, *k*ESP-UCB, and *k*ESP-LBSMEC, respectively.

Figure 3 compares the effectiveness of *k*ESP-PSO with *k*ESP-CR_CP, *k*ESP-PA, *k*ESP-UCB, and *k*ESP-LBSMEC in experiment set #2. Overall, *k*ESP-PSO achieves the minimum TSRN-CP in all cases. *k*ESP-PSO has an average advantage of 75.45565%, 62.20537%, 82.53591%, and 75.51733% over *k*ESP-CR_CP, *k*ESP-PA, *k*ESP-UCB, and *k*ESP-LBSMEC, respectively, over the course of experiment sets #2.1, #2.2, and #2.3. In detail, as illustrated in Figure 3(a), within experiment set #2.1, *k*ESP-PSO has advantages of 72.84820%, 59.96922%, 80.23471%, and 73.94777% over *k*ESP-CR_CP, *k*ESP-PA, *k*ESP-UCB, and *k*ESP-LBSMEC, respectively. As shown in Figure 3(b), within experiment set #2.2, *k*ESP-PSO has advantages of 64.90660%, 52.88316%, 75.89349%, and 67.68684% over *k*ESP-CR_CP, *k*ESP-PA, *k*ESP-UCB, and *k*ESP-LBSMEC, respectively. As shown in Figure 3(c), within experiment set #2.3, *k*ESP-PSO has advantages of 87.000265%, 72.348301%, 90.377505%, and 83.769215% over *k*ESP-CR_CP, *k*ESP-PA, *k*ESP-UCB, and *k*ESP-LBSMEC, respectively.

D. Ablation Experiments

The experiments compare *k*ESP-PSO with four other methods to evaluate their effectiveness: 1) a transmission delay-oriented method (*k*ESP-T), 2) a switching delay-oriented method (*k*ESP-S), 3) a base station robustness-oriented method (*k*ESP-R), and 4) a variance of nearby base stations-oriented method (*k*ESP-N).

· *k*ESP-T: Considers only transmission delay, ignoring switching delay, base station robustness, and variance of nearby base stations.

· *k*ESP-S: Considers only switching delay, ignoring transmission delay, base station robustness, and variance of nearby base stations.

· *k*ESP-R: Considers only base station robustness, ignoring transmission delay, switching delay, and variance of nearby base stations.

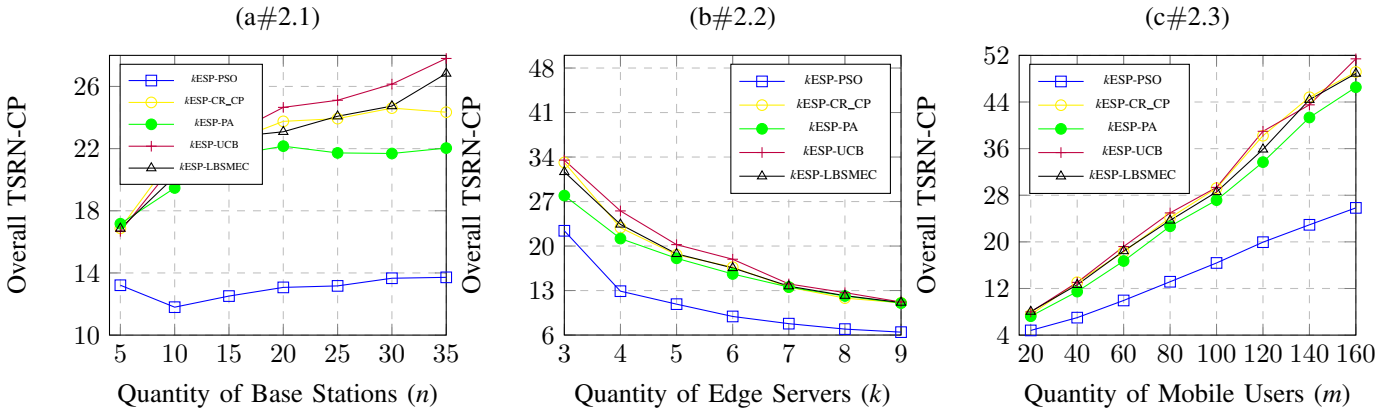

Fig. 3: Assessment of Effectiveness under Telecom Shanghai Dataset experiments (Set #2).

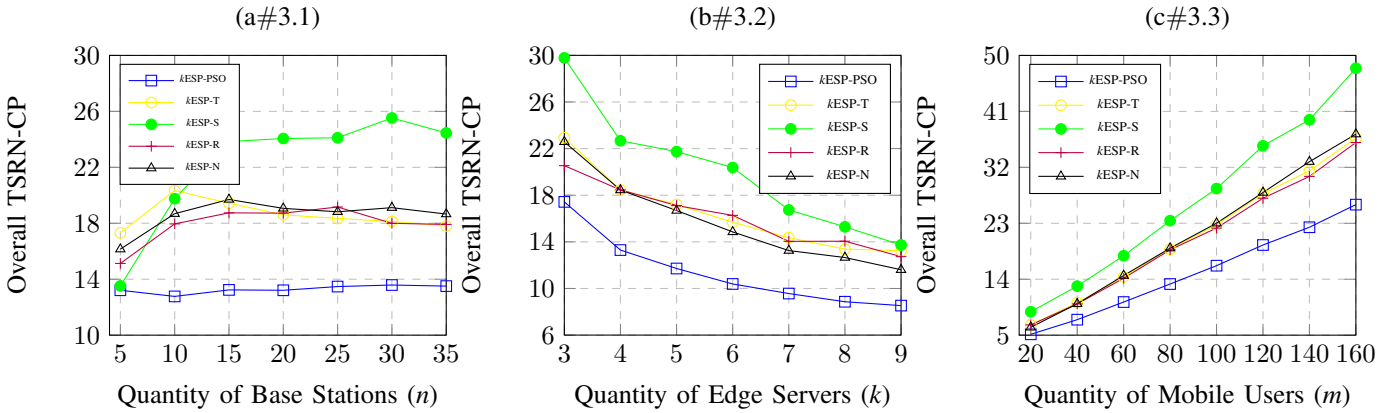

Fig. 4: Assessment of Effectiveness under EUAdatasets experiments (Set #3).

· *k*ESP-N: Considers only variance of nearby base stations, ignoring transmission delay, switching delay, and base station robustness.

Figure 4 compares *k*ESP-PSO with *k*ESP-T, *k*ESP-S, *k*ESP-R, and *k*ESP-N in terms of TSRN-CP in the experiment set #1. Overall, *k*ESP-PSO achieves the minimum TSRN-CP, the theoretical minimum trade-off between transmission delay, switching delay, base station robustness, and variance of nearby base stations. Its performance significantly surpasses *k*ESP-T, *k*ESP-S, *k*ESP-R, and *k*ESP-N by 41.65133%, 71.30973%, 36.05192%, and 40.27124%, respectively. This means *k*ESP-PSO can find a *k*ESP-TSRN strategy with 41.65133%, 71.30973%, 36.05192%, and 40.27124% more TSRN-CP than *k*ESP-T, *k*ESP-S, *k*ESP-R, and *k*ESP-N, respectively. Figure 4(a) illustrates that as the number of base stations increases, the $TSRN$ value gradually stabilizes. This trend can be attributed to the fact that more base stations result in smaller user groups, thereby reducing the number of users connected to each individual base station. However, for edge servers, although a higher number of base stations may lead to an increased connection count, the total number of users they serve remains relatively constant. The addition of more base stations brings the placement of edge servers closer to their ideal locations. Nevertheless, since the user population does not change significantly, its influence on transmission delay, handover delay, and variance among neighboring base stations is minimal. In terms of base station robustness, increasing the number of base stations can indeed improve system resilience; however, the overall impact on $TSRN$ remains limited. Therefore, the growth in the number of base stations has only a marginal effect on the $TSRN$ metric. Figure 4(b) demonstrates that increasing the number of edge servers results in a reduction in $TSRN$. This occurs because a larger number of edge servers leads to a decrease in the number of base stations assigned to each one, which in turn reduces the number of users served by each edge server. As edge servers are positioned closer to end users, both transmission and handover delays are reduced. Moreover, the distribution of edge servers becomes more balanced, lowering the variance among nearby base stations. Consequently, expanding the number of edge servers contributes to a decline in $TSRN$. Figure 4(c) reveals that an increase in the number of users leads to a noticeable rise in $TSRN$. This is primarily due to the increased load on edge servers, as more users connect to them, resulting in higher overall transmission and handover delays. As a result, the $TSRN$ value increases accordingly.

VI. Conduction

In this study, we introduce an algorithm for edge server deployment strategy based on 0-1 programming to address

the challenges of high latency, low reliability, and high carbon emissions in multi-user edge computing (MEC) scenarios. The algorithm aims to enhance user experience while reducing carbon emissions. By utilizing the Particle Swarm Optimization Algorithm, the algorithm can find an optimal edge server placement method that minimizes user latency for edge services, fortifies the resilience of the edge server ecosystem, and reduces server power consumption. Empirical findings show that the *k*ESP-PSO algorithm surpasses alternative approaches. In the future, we will investigate server placement choices under server capacity constraints and optimize the algorithm to save computation time.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China under Grant 62372242 and Grant 92267104, Jiangsu Provincial Major Project on Basic Research of Cutting-edge and Leading Technologies under Grant BK20232032, the Natural Science Foundation of Jiangsu Province under Grant BK20240692, the Natural Science Foundation of the Jiangsu Higher Education Institutions of China under Grant 24KJB520021, the Startup Foundation for Introducing Talent of NUIST, the National Students' Platform for Innovation Training Program, and the NUIST Students' Platform for Innovation Training Program XJDC202510300236.

REFERENCES

[1] J. Xu, H. Xiang, S. Zang, M. Bilal, M. Khan, and G. Cui, "A dqn-based edge offloading method for smart city pollution control," *Tsinghua Science and Technology*, 2024.

[2] O. Rawlley, S. Gupta, K. Mahajan, and S. Rathore, "Green-emulto: A next generation edge-assisted multi-level traffic orchestrator for green computing in consumer autonomous vehicles," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 4, pp. 7291–7301, 2024.

[3] X. Xu, F. Wu, M. Bilal, X. Xia, W. Dou, L. Yao, and W. Zhong, "Xrl-shap-cache: an explainable reinforcement learning approach for intelligent edge service caching in content delivery networks," *Science China Information Sciences*, vol. 67, no. 7, p. 170303, 2024.

[4] Y. Liang, M. Yin, W. Wang, Q. Liu, L. Wang, X. Zheng, and T. Wang, "Collaborative edge server placement for maximizing qos with distributed data cleaning," *IEEE Transactions on Services Computing*, pp. 1–15, 2025.

[5] X. Xu, X. Zhou, X. Zhou, M. Bilal, L. Qi, X. Xia, and W. Dou, "Distributed edge caching for zero trust-enabled connected and automated vehicles: A multi-agent reinforcement learning approach," *IEEE Wireless Communications*, vol. 31, no. 2, pp. 36–41, 2024.

[6] X. Xu, H. Dong, L. Qi, X. Zhang, H. Xiang, X. Xia, Y. Xu, and W. Dou, "Cmclrec: Cross-modal contrastive learning for user cold-start sequential recommendation," in *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2024, pp. 1589–1598.

[7] X. Xu, H. Dong, H. Xiang, X. Hu, X. Li, X. Xia, X. Zhang, L. Qi, and W. Dou, "C2lrec: Causal contrastive learning for user cold-start recommendation with social variable," 2025. [Online]. Available: https://doi.org/10.1145/3711858

[8] H. Xiang, X. Zhang, X. Xu, A. Beheshti, L. Qi, Y. Hong, and W. Dou, "Federated learning-based anomaly detection with isolation forest in the iot-edge continuum," 2024. [Online]. Available: https://doi.org/10.1145/3702995

[9] S. Goudarzi, S. A. Soleymani, M. H. Anisi, A. Jindal, F. Dinmohammadi, and P. Xiao, "Sustainable edge node computing deployments in distributed manufacturing systems," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 1471–1481, 2024.

[10] J. Jing, Y. Yang, X. Zhou, J. Huang, L. Qi, and Y. Chen, "Multi-uav cooperative task offloading in blockchain-enabled mec for consumer electronics," *IEEE Transactions on Consumer Electronics*, vol. 71, no. 1, pp. 2271–2284, 2025.

[11] M. K. Hasan, N. Jahan, M. Z. A. Nazri, S. Islam, M. Attique Khan, A. I. Alzahrani, N. Alalwan, and Y. Nam, "Federated learning for computational offloading and resource management of vehicular edge computing in 6g-v2x network," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 3827–3847, 2024.

[12] B. Liu, H. Tian, Z. Shen, Y. Xu, and W. Dou, "A consortium blockchain-based edge task offloading method for connected autonomous vehicles," *ACM Trans. Auton. Adapt. Syst.*, 2024. [Online]. Available: https://doi.org/10.1145/3696004

[13] Q. Qi, T. Shi, K. Qin, and G. Luo, "Completion time optimization in uav-relaying-assisted mec networks with moving users," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 1246–1258, 2024.

[14] X. Xia, Z. Wang, R. Sun, B. Liu, I. Khalil, and M. Xue, "Edge unlearning is not "on edge"! an adaptive exact unlearning system on resource-constrained devices," in *2025 IEEE Symposium on Security and Privacy (SP)*, 2025, pp. 2546–2563.

[15] Q. Li, X. Huang, B. Liu, P. Li, J. Zhang, and K. Chen, "Cache-aware i/o rate control for rdma," in *Proceedings of the 9th Asia-Pacific Workshop on Networking*, 2025, p. 9–16. [Online]. Available: https://doi.org/10.1145/3735358.3735376

[16] F. Wang, J. Xu, X. Wang, and S. Cui, "Joint offloading and computing optimization in wireless powered mobile-edge computing systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1784–1797, 2018.

[17] F. Zeng, Y. Ren, X. Deng, and W. Li, "Cost-effective edge server placement in wireless metropolitan area networks," *Sensors*, vol. 19, no. 1, 2019.

[18] G. Cui, Q. He, F. Chen, H. Jin, and Y. Yang, "Trading off between user coverage and network robustness for edge server placement," *IEEE Transactions on Cloud Computing*, vol. 10, no. 3, pp. 2178–2189, 2022.

[19] X. Xu, K. Meng, H. Xiang, G. Cui, X. Xia, and W. Dou, "Blockchain-enabled secure, fair, and scalable data sharing in zero-trust edge-end environment," *IEEE Journal on Selected Areas in Communications*, vol. 43, no. 6, pp. 2056–2069, 2025.

[20] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile edge computing—a key technology towards 5g," *ETSI white paper*, vol. 11, no. 11, pp. 1–16, 2015.

[21] Y. Li and S. Wang, "An energy-aware edge server placement algorithm in mobile edge computing," in *2018 IEEE International Conference on Edge Computing (EDGE)*, 2018, pp. 66–73.

[22] S. Wang, Y. Zhao, J. Xu, J. Yuan, and C.-H. Hsu, "Edge server placement in mobile edge computing," *Journal of Parallel and Distributed Computing*, vol. 127, pp. 160–168, 2019.

[23] A. A. Vali, S. Azizi, and M. Shojafar, "Resp: A recursive clustering approach for edge server placement in mobile edge computing," *ACM Trans. Internet Technol.*, vol. 24, no. 3, Jul. 2024. [Online]. Available: https://doi.org/10.1145/3666091

[24] A. Asghari and M. K. Sohrabi, "Server placement in mobile cloud computing: A comprehensive survey for edge computing, fog computing and cloudlet," *Computer Science Review*, vol. 51, p. 100616, 2024.

[25] T. Wang, Y. Liang, X. Shen, X. Zheng, A. Mahmood, and Q. Z. Sheng, "Edge computing and sensor-cloud: Overview, solutions, and directions," vol. 55, no. 13s, 2023.

[26] X. Zhang, Z. Li, C. Lai, and J. Zhang, "Joint edge server placement and service placement in mobile-edge computing," *IEEE Internet of Things Journal*, vol. 9, no. 13, pp. 11 261–11 274, 2022.

[27] R. Li, Z. Zhou, X. Zhang, and X. Chen, "Joint application placement and request routing optimization for dynamic edge computing service management," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 12, pp. 4581–4596, 2022.

[28] Z. Zhao, H. Cheng, X. Xu, and Y. Pan, "Graph partition and multiple choice-ucb based algorithms for edge server placement in mec environment," *IEEE Transactions on Mobile Computing*, vol. 23, no. 5, pp. 4050–4061, 2024.

[29] A. Mazloomi, H. Sami, J. Bentahar, H. Otrok, and A. Mourad, "Reinforcement learning framework for server placement and workload allocation in multiaccess edge computing," *IEEE Internet of Things Journal*, vol. 10, no. 2, pp. 1376–1390, 2023.

[30] V. Tiwari, C. Pandey, A. Dahal, D. S. Roy, and U. Fiore, "A knapsack-based metaheuristic for edge server placement in 5g networks with heterogeneous edge capacities," *Future Generation Computer Systems*, vol. 153, pp. 222–233, 2024.

[31] Y. Li, A. Zhou, X. Ma, and S. Wang, "Profit-aware edge server placement," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 55–67, 2022.

[32] A. Asghari and M. K. Sohrabi, "Multiobjective edge server placement in mobile-edge computing using a combination of multiagent deep q-network and coral reefs optimization," *IEEE Internet of Things Journal*, vol. 9, no. 18, pp. 17 503–17 512, 2022.
[33] S. Li, G. Liu, L. Li, Z. Zhang, W. Fei, and H. Xiang, "A review on air-ground coordination in mobile edge computing: Key technologies, applications and future directions," *Tsinghua Science and Technology*, 2024.
[34] C. Huang, A. Wang, J. Li, and K. W. Ross, "Measuring and evaluating large-scale cdns," in *ACM IMC*, vol. 8, 2008, pp. 15–29.
[35] B. Krishnamurthy, C. Wills, and Y. Zhang, "On the use and performance of content distribution networks," in *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement*, ser. IMW '01. New York, NY, USA: Association for Computing Machinery, 2001, p. 169–182.
[36] Y. Zhang, D. Li, and M. Tatipamula, "The freshman handbook: A hint for server placement in online social network services," in *2012 IEEE 18th International Conference on Parallel and Distributed Systems*, 2012, pp. 588–595.
[37] B. Li, M. Golin, G. Italiano, X. Deng, and K. Sohraby, "On the optimal placement of web proxies in the internet," in *IEEE INFOCOM '99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now (Cat. No.99CH36320)*, vol. 3, 1999, pp. 1282–1290 vol.3.
[38] L. Qiu, V. Padmanabhan, and G. Voelker, "On the placement of web server replicas," in *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No.01CH37213)*, vol. 3, 2001, pp. 1587–1596 vol.3.
[39] E. Cronin, S. Jamin, C. Jin, A. Kurc, D. Raz, and Y. Shavitt, "Constrained mirror placement on the internet," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 7, pp. 1369–1382, 2002.
[40] S. Jamin, C. Jin, A. Kurc, D. Raz, and Y. Shavitt, "Constrained mirror placement on the internet," in *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No.01CH37213)*, vol. 1, 2001, pp. 31–40 vol.1.
[41] P. Lai, Q. He, M. Abdelrazek, F. Chen, J. Hosking, J. Grundy, and Y. Yang, "Optimal edge user allocation in edge computing with variable sized vector bin packing," in *Service-Oriented Computing: 16th International Conference, ICSOC 2018, Hangzhou, China, November 12-15, 2018, Proceedings 16*. Springer, 2018, pp. 230–245.
[42] Q. He, G. Cui, X. Zhang, F. Chen, S. Deng, H. Jin, Y. Li, and Y. Yang, "A game-theoretical approach for user allocation in edge computing environment," *IEEE Transactions on Parallel and Distributed Systems*, vol. 31, no. 3, pp. 515–529, 2020.
[43] Y. Guo, S. Wang, A. Zhou, J. Xu, J. Yuan, and C.-H. Hsu, "User allocation-aware edge cloud placement in mobile edge computing," *Software: Practice and Experience*, vol. 50, no. 5, pp. 489–502, 2020.
[44] S. Wang, Y. Guo, N. Zhang, P. Yang, A. Zhou, and X. Shen, "Delay-aware microservice coordination in mobile edge computing: A reinforcement learning approach," *IEEE Transactions on Mobile Computing*, vol. 20, no. 3, pp. 939–951, 2021.
[45] G. Cui, Q. He, F. Chen, H. Jin, Y. Xiang, and Y. Yang, "Location privacy protection via delocalization in 5g mobile edge computing environment," *IEEE Transactions on Services Computing*, vol. 16, no. 1, pp. 412–423, 2023.

BIOGRAPHY SECTION

Yusen Wang is currently pursuing the B.S. degree with the School of Software, Nanjing University of Information Science and Technology, China. His research interests include edge computing, cloud computing, and Deep Learning.

Xiaolong Xu (Senior Member, IEEE) received the Ph.D. degree in computer science and technology from Nanjing University, China, in 2016. He is currently a Full Professor with the School of Soft-ware, Nanjing University of Information Science and Technology. His research interest includes Cloud Computing, Big Data, Edge Computing, Deep Learning, and Federated Edge Learning . He has published over 100 peer-review papers in international journals and conferences, including IEEE TPDS, IEEE TKDE, IEEE TSC, ACM TOSN, IEEE TITS, IEEE TII, ACM TOIT, ACM TOMM, ACM TIST, IEEE TVT, IEEE IOT, IEEE TCC, IEEE TBD, IEEE TCSS, IEEE TETCI, Software: Practice and Experience, World Wide Web journal, Information Sciences, Journal of Network and Computer Applications, etc. He was selected as the Highly Cited Researcher of Clarivate 2021, 2022 and 2023.

Ruoshui Wang WangRsMomo@gmail.com, School of Software, Nanjing University of Information Science and Technology.

Muhammad Bilal (Senior Member, IEEE) received the Ph.D. degree in information and communication network engineering from the School of Electronics and Telecommunications Research Institute (ETRI), Korea University of Science and Technology, in 2017. In 2023, he joined Lancaster University as a Senior Lecturer (Associate Professor) with the School of Computing and Communications. He is a prolific author, known for his wide-ranging contributions to numerous articles published in internationally renowned, top-tier journals. His pioneering work has also led to the successful acquisition of multiple U.S. and Korean patents. His research interests include network optimization, cyber security, the Internet of Things, vehicular networks, information-centric networking, digital twins, artificial intelligence, and cloud/fog computing.

Wei Liu received the B.S. degree in computer science and technology engineering from the Nanjing University of Information Science and Technology in 2022, where he is currently pursuing the master's degree in software engineering. His research interests include mobile edge computing, big data, Internet of Things, and machine learning.

Guangming Cui received his Master's degree from Anhui University, China, in 2018 and his PhD degree from Swinburne University of Technology, Australia, in 2022, in computer science. Currently, he is an associate professor at Nanjing University of Information Science & Technology, China. He has published more than 30 peer-reviewed articles in international journals and conferences, including the IEEE TMC, IEEE TPDS, IEEE TSC, IEEE TDSC, JSAC, ICWS, ICSOC, etc. His research interests include edge computing, service computing, mobile computing, and software engineering.