Received 25 March 2023; accepted 1 April 2023. Date of publication 5 April 2023; date of current version 14 April 2023.

Digital Object Identifier 10.1109/OJCOMS.2023.3264465

Machine Learning for Predictive Deployment of UAVs With Multiple Access

FEIYANG GUO[1], LINYAN LU[1], ZELIN ZANG[2], AND MOHAMMAD SHIKH-BAHAEI[1] (Senior Member, IEEE)

[1]Centre for Telecommunications Research, Department of Engineering, King's College London, WC2R 2LS London, U.K.

[2]College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310027, China

CORRESPONDING AUTHOR: F. GUO (e-mail: feiyang.guo@kcl.ac.uk)

ABSTRACT This paper presents a machine learning-based framework for the predictive deployment of unmanned aerial vehicles (UAVs) as flying base stations (BSs) to offload heavy traffic from ground BSs. To account for time-varying traffic distribution, a long short-term memory (LSTM)-based prediction algorithm is introduced to predict future cellular traffic. A joint K-means and expectation maximization (EM) algorithm based on Gaussian mixture models (GMM) is proposed to determine the service area of each UAV based on the predicted user service distribution. Based on the predicted traffic, the optimal positions of UAVs are derived, and four multiple access techniques, namely, rate splitting multiple access (RSMA), frequency domain multiple access (FDMA), time domain multiple access (TDMA), and non-orthogonal multiple access (NOMA), are compared to minimize the total transmit power. Simulation results show that the proposed method can reduce up to 24% of the total power consumption compared to the conventional method without traffic prediction. Furthermore, RSMA is found to require the lowest transmit power among the four multiple access techniques. Therefore, this paper focuses on the comparison of multiple access techniques for UAV deployment, which is essential for the efficient and effective use of UAVs as flying BSs.

INDEX TERMS UAV deployment, LSTM, K-means, EM, GMM, RSMA.

I. INTRODUCTION

SINCE user demands for communication services have grown dramatically, traditional base stations (BSs) cannot meet the required demand of the cellular traffic, which can lead to a bottleneck in cellular communication [1], [2], [3]. Recently, there has been growing interest in the study of unmanned aerial vehicle (UAV) communication due to its excellent attributes of versatility, maneuverability, and flexibility [4]. UAVs gradually play an important role in wireless communications, offering low-cost and efficient wireless connectivity for devices. UAVs acting as flying BSs are one of the most important research subjects in UAV communication. Through location adjustment, obstacle avoidance, and line-of-sight (LoS) link reinforcement, UAVs are able to offload data traffic from loaded BSs and increase the connectivity of wireless networks, thereby improving communication throughput, coverage, and energy efficiency [5].

Therefore, it is a feasible and beneficial option to utilize UAVs to ensure the connectivity of wireless communication networks by meeting the surging data demands. For efficient and rapid dispatch of UAVs, the prediction of potential hotspot areas plays a crucial role in helping network operators acquire the information of occurrence and degrees of congestion in advance to reduce the entire network communication delay [6], [7], [8]. Machine learning (ML) techniques are useful tools that have the ability to efficiently predict the distribution of future traffic data [9], [10], [11], [12], [13], [14], [15]. With such predictions, the target locations of UAVs can be specified beforehand, and the deployment can be more intelligent and on-demand.

There are a number of existing works investigating the applications of UAV deployment in communication. In UAV deployment, UAV service areas and their optimal placements are two critical factors [16], [17], [18], [19]. In [16], UAV altitudes and 2-D locations are optimized based on circle

packing theory. Then, UAV placements and service area boundaries are separately determined for power efficiency in [17]. To efficiently dispatch UAVs, some intelligent and practical methods are proposed. A functional split scheme selection of each UAV is joined with UAV deployment in [18]. In [19], an adaptive deployment scheme is introduced to optimize UAV displacement direction; time-dependent traffic of mobile users can be served in real time. In addition, the recent work by [20] discusses the importance of edge intelligence technologies for achieving critical mission services for sixth-generation (6G) in space-air-ground integrated networks, and proposes novel algorithms and architectures for optimizing resource management and data processing. Moreover, the work by [21] proposes an architecture for building agile and resilient UAV networks based on software-defined networking (SDN) and blockchain, which enables secure and efficient communication, flexible network management, and autonomous decision-making by UAVs. In recent years, several studies [22], [23], [24] have investigated the challenge of managing the heterogeneous and high dynamic resources of UAVs, and proposed novel approaches for optimizing resource allocation and utilization in UAV networks.

In UAV communication, machine learning techniques are also applied to improve system performance [25], [26], [27]. The aerial channel environment is conditional random field position (CRF) in [28]. In [25], in order to provide more efficient downlink service, a learning approach based on the weighted expectation maximization (WEM) algorithm is used to predict downlink user traffic demand; meanwhile a traffic load contract using contract theory is introduced.

Different multiple access methods for cognitive radio networks has been explored in [29], starting from orthogonal multiple access (OMA) and progressing to non-orthogonal multiple access (NOMA) and rate-splitting. In [30], a new anti-interference coding scheme is proposed for achieving reliable transmission in NOMA systems with randomly deployed receivers. The scheme leverages interference among receivers to improve system reliability and throughput, effectively reducing transmission error rates and addressing the challenge of reliable transmission in such systems.

Rate splitting multiple access (RSMA) is a widely used multiple access technique in UAV communication systems, which has been extensively studied in the literature. For example, the downlink performance of RSMA-based UAV communications has been investigated in [31], and the ergodic capacity and placement optimization for RSMA-enabled UAV-assisted communication have been studied in [32]. In addition, RSMA has been combined with other technologies such as Reconfigurable Intelligent Surface (RIS) and Full Duplex (FD) to enhance the communication performance [33]. Moreover, trajectory design and outage probability analysis have also been explored in multi-UAV-assisted RSMA downlink communication [34], [35].

Although there have been numerous studies on RSMA in UAV communication systems, to the best of our knowledge, none of them have focused on optimizing the UAV locations for low power consumption. In this paper, we propose a novel UAV deployment framework that integrates machine learning and optimization techniques to predict the cellular traffic and optimize the UAV locations for low power consumption using RSMA. Our proposed framework aims to improve the energy efficiency and performance of UAV-assisted networking, which is an important research direction in the field of wireless communication systems.

The main contribution of this work is a novel UAV deployment framework for UAV-assisted networking with machine learning approaches. In contrast to existing works, our framework not only predicts the cellular traffic using a backpropagation (BP) neural network model, but also optimizes the UAV locations using a joint K-means and expectation maximization (EM) algorithm of a Gaussian mixture model (KEG), and compares the performance of different multiple access techniques. Specifically, we propose a predictive deployment scheme that divides the entire service area into clusters as UAV service areas in temporal and spatial patterns using the KEG algorithm. The optimized UAV locations are updated iteratively to minimize the total transmit power of the UAV swarm network, taking into account the predicted traffic and the energy consumption of the multiple access techniques.

Our key contributions are summarized as follows:

- A novel predictive UAV deployment framework is proposed, which integrates machine learning and optimization techniques to minimize the total transmit power of the UAV swarm network.
- The joint K-means and EM algorithm of a Gaussian mixture model (KEG) is proposed to divide the entire service area into clusters as UAV service areas in temporal and spatial patterns, which enhances the accuracy and flexibility of the UAV deployment scheme.
- Four different multiple access techniques (RSMA, NOMA, frequency division multiple access (FDMA), time division multiple access (TDMA)) are compared, and the results show that the proposed framework with RSMA can decrease up to 16.9%, 35.5% and 66.4% total power consumption compared to NOMA, FDMA and TDMA, respectively.

We believe that our novel framework and the proposed KEG algorithm provide a significant contribution to the research field, and the simulation results demonstrate the effectiveness and efficiency of our approach. We hope that our work will inspire further research on predictive UAV deployment and optimization for UAV-assisted networking.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Given a time-dependent UAV-assisted wireless network, a group of ground users in the network are distributed in a geographical area C. A set $\boldsymbol{I}$ of I UAVs assist a set $\boldsymbol{J}$ of J

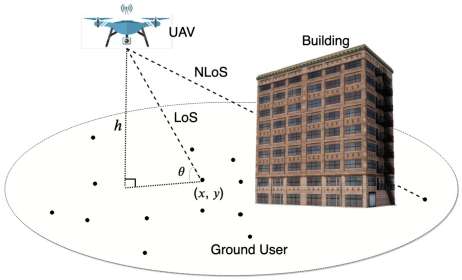

FIGURE 1. The LoS Link and the NLoS Link of UAV Transmission.

ground BSs to offload amounts of cellular traffic for congestion alleviation, so the ground users in a time-variant hotspot area have the air-ground communications with UAVs when ground BSs are overloaded.

It is assumed that the height of all users and BSs is zero compared to the UAVs. Besides, each UAV has directional antennas, so the transmissions of different UAVs will not be interfered with each other. For convenient elaboration, we specify the ground users served by a UAV as *aerial users* and the service area of a UAV i as an *aerial cell* which can be expressed as C_i. In order to consider the communication of all users fairly, the aerial cells are supposed to completely cover the entire area without any overlaps.

Because UAVs have limited energy amounts, UAVs should be efficiently deployed.

A. AIR-GROUND MODEL

We assume that a ground user $j \in \boldsymbol{J}$ is located at $(x, y) \in C$ and a UAV $i \in \boldsymbol{I}$ is located at (x_i, y_i, h_i) with the aerial cell C_i, so the uplink received power of a UAV i and a ground user j is calculated as:

$$P_{r,ij}[dB] = P_{ij}[dB] + G_{ij}[dB] - L_{ij}[dB] - r_{ij}[dB], \quad (1)$$

where P_{ij} is the transmit power, G_{ij} is the antenna gain, L_{ij} is the free space path loss, r_i is the excessive path loss with a Gaussian distribution which depends on the category of link. For mathematical tractability, we give a hypothesis that all of beam alignment of ground-air links are perfect and UAV antenna gains are same. Thus, G_{ij} can be a constant number G. The free space path loss has a specific formula to be acquired:

$$L_{ij}[dB] = 10n \log\left(\frac{4\pi f_c d_{ij}}{c}\right), \quad (2)$$

where $n \geq 2$ is the path loss exponent, f_c is the system carrier frequency, c is the light speed, $d_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2 + h_i^2]^{1/2}$ is the distance between a UAV i and a user j [36].

In general, the air-to-ground transmission is separated into two main categories: the line-of-sight link (LoS) and the non-line-of-sight link (NLoS). The NLoS link suffers a higher excessive path loss owing to shadowing and blockage. Fig. 1 has shown these two links in the picture. The excessive path loss in these two links can be expressed as $r_i^{LoS} \sim \mathcal{N}(\mu_{LoS}, \sigma_{LoS}^2)$ and $r_i^{NLoS} \sim \mathcal{N}(\mu_{NLoS}, \sigma_{NLoS}^2)$ respectively. The probability of the occurrence of LoS link is similar to a sigmoid function [37]:

$$p_{ij}^{LoS} = \frac{1}{1 + a\exp\left[b\left(a - \theta_{ij}\right)\right]}, \quad (3)$$

where a and b are environment constant coefficients, $\theta_{e,ij} = sin^{-1}(h_i/d_{ij})$ is the elevation angle of a UAV i and a user j, so the NLoS link existence probability is $p_{ij}^{NLoS} = 1 - p_{ij}^{LoS}$ [36]. Therefore, the average excessive path loss in the uplink transmission is:

$$r_{ij} = r_i^{LoS} p_{ij}^{LoS} + r_i^{NLoS} p_{ij}^{NLoS} \quad (4)$$

To this end, the uplink data rate of a UAV i and a user j can be expressed as:

$$R_{ij} = W_{ij} \log_2\left(\frac{P_{r,ij}}{N_0} + 1\right)(bits/s) \quad (5)$$

where W_{ij} is the bandwidth of a UAV i and a user j which can be, $N_0 = n_0 W_{ij}$ is the power of additive white Gaussian noise and n_0 is its average power spectral density. For tractable formulation, each UAV can offer sufficient overall bandwidth and all transmit bandwidths are assumed to be a constant value W.

B. MULTI-ACCESS MODES

In this paper, four different multi-access techniques will be used. They are rate splitting multiple access (RSMA), frequency division multiple access (FDMA), time division multiple access (TDMA) and non-orthogonal multiple access (NOMA) [38]. We set two users as an example to give a series of theoretical formulas and set R_1 and R_2 are the actual data rates of two users, g_1 and g_2 are the channel gains of two users which include the antenna gain, free space path loss and the excessive path loss, P_1 and P_2 are the transmit power of the two users.

RSMA is a multi-access mode with coding and decoding techniques, the equations for maximizing the sum-rate of users for uplink transmission are shown below [38], [39], [40], [41], [42], [43], [44], [45], [46]:

$$\begin{cases} R_1 \leq W \log_2(1 + g_1 P_1/(Wn_0)) \\ R_2 \leq W \log_2(1 + g_2 P_2/(Wn_0)) \\ R_1 + R_2 \leq W \log_2(1 + (g_1 P_1 + g_2 P_2)/(Wn_0)) \end{cases} \quad (6)$$

where R_1 and R_2 represent the rates of users for RSMA, the third line is derived from the optimal rate region.

FDMA is a technique of BS bandwidth allocation for users to avoid interference of transmissions in the same area:

$$\begin{cases} R_1 \leq Wa_1 \log_2(1 + g_1 P_1/(Wn_0 a_1)) \\ R_2 \leq Wa_2 \log_2(1 + g_2 P_2/(Wn_0 a_2)) \\ a_1 + a_2 = 1, a_1 \geq 0, a_2 \geq 0 \end{cases} \quad (7)$$

where R_1 and R_2 represent the rates of users for FDMA, a_1 and a_2 are the frequency weight coefficients.

Similarly, TDMA assigns the partial time period to users who can use the whole bandwidth, the equations have been represented.

$$\begin{cases} R_1 \le Wb_1 \log_2(1 + g_1P_1/(Wn_0)) \\ R_2 \le Wb_2 \log_2(1 + g_2P_2/(Wn_0)) \\ b_1 + b_2 = 1, b_1 \ge 0, b_2 \ge 0 \end{cases} \tag{8}$$

where R_1 and R_2 represent the rates of users for TDMA, b_1 and b_2 are the time weight coefficients.

NOMA is a multiple access technique that allowing the users to share the same time, frequency, and code resources with different power allocation schemes, while introducing interference between users. For the uplink NOMA, BS uses successive interference cancelation (SIC), and its uplink transmission are shown as below [29]:

$$\begin{cases} R_1 \le W \log_2(1 + g_1P_1/(Wn_0)) \\ R_2 \le W \log_2\left(1 + \frac{g_2P_2}{Wn_0} \Big/ \left(\frac{g_1P_1}{Wn_0} + 1\right)\right) \\ R_1 + R_2 \le W \log_2(1 + (g_1P_1 + g_2P_2)/Wn_0) \end{cases} \tag{9}$$

where R_1 is the rate of the first decoded user for NOMA, R_1 is the rate of the second decoded user for NOMA by using SIC.

In the simulation and analysis section, we will compare these four multi-access modes for choosing a best one to minimize the total transmit power.

C. UPLINK TRANSMIT POWER COMPUTATION

Network operators are required to assign UAVs to those traffic congestion areas so as to offload the heavy traffic from busy BSs. For UAVs, continuous movements will consume too much transmit power. Thus, we need to analyze the situation of traffic offloading. A data set is presented as a matrix $\boldsymbol{D}$:

$$\boldsymbol{D} = \left\{D_d^t(x, y) \mid t \in \{T, \ldots, 24T\}, \right. \\ \left. d \in \{1, \ldots, 8\}, (x, y) \in C\right\} \tag{10}$$

where T is the time of an hour, t is the time moment and d is the day. $D_d^t(x, y)$ represents the amount of cellular traffic offloaded from a BS located at (x, y) in a period of T in the day d [36]. In this paper, for the convenience of simulation and analysis, we assume that all cellular traffic of BSs are totally offloaded to UAVs.

Since the positions of mobile users are uncertain and most of mobile users only move around a single BS in a period of an hour, we assume that all ground receivers have the same positions as their nearest BSs.

After obtaining the future predictive traffic information through the implementation of ML methods based on the matrix $\boldsymbol{S}$, the required average data rate within a aerial cell C_i in a period of T will be given. Only when the communication throughput of UAV is not less than the demanded data rate, the communication can be ensured. Therefore, the communication condition is formulated as:

$$\iint_{C_i} R_i(x, y)dxdy \ge \frac{1}{T} \iint_{C_i} D_d^t(x, y)dxdy \tag{11}$$

where R(x, y) is the maximum data rate of the overall transmission between a group of ground users in a ground cell with a BS at (x, y) and a UAV i. We can simplify this equation as

$$R_i(x, y) \ge D_d^t(x, y)/T \tag{12}$$

Let us set $D_d^t(x, y)/T = \alpha(x, y)$, $\alpha(x, y)$ is the minimal requirement of average data rate. Combining the formula (1), (5), (12) and the optimal rate region, the minimum power for transmission should be provided by a UAV will be:

$$P_i^{min}(x, y) = \left(2^{\frac{\alpha(x,y)}{W}} - 1\right) n_0 W L_i(x, y) r_i(x, y)/G \tag{13}$$

where both $L_i(x, y)$ and $r_i(x, y)$ are between the UAV i and the group of ground users in a ground cell of the BS at (x, y), $L_i(x, y)$ is the path loss in free space, $r_i(x, y)$ is the excessive path loss. This equation provides a target basis for UAV location optimization.

This section proposes a novel predictive scheduling scheme of UAV based on ML. According to the real data set in the City Cellular Traffic Map [47] and the characteristics of cellular data traffic, for the sake of the efficiency of UAV deployment, we make some rational assumptions. Because humans have a certain pace of life with periodic activity, the change of cellular data traffic has a repetitive pattern in daily life [48]. Thus, we assume that the cellular traffic amount has specific distribution in the same hour in different days and the data of each hour in the same day is independent. To this end, the real data set can be classified into 24 independent models and we assume each single model follows the Gaussian mixture model which will be explained in Section II-E3.

The logical procedure diagram of UAV predictive deployment is shown in Fig. 2. At first, the acquired real data set is preprocessed to get the cellular traffic amount of every hour in the first 5 days $\{D_d^t\}_{d=1,t=T}^{5,24T}$ and the topology information of every BS $\{\boldsymbol{x}_n\}_{n=1}^N = \{(x_n, y_n)\}_{n=1}^N$, where x_n is the relative longitude of n^{th} BS and y_n is the relative latitude of n^{th} BS. Then, a backpropagation (BP) neural network model for cellular traffic amount prediction is developed to predict the cellular amount in 24 hours in the 6^{th} day. At the same time, a joint K-means and expectation maximization (EM) algorithm relying on a Gaussian mixture models (GMM) for aerial cell classification (ground user clustering) is created, the point cluster label $\{l_n\}_{n=1}^N$ of every single point $\boldsymbol{x}_n$ is obtained, and the point $\boldsymbol{x}_n$ with the same label value consists of an aerial cell. Then, the optimal UAV locations $\{\boldsymbol{x}_i\}_{i=1}^K = \{(x_i, y_i)\}_{i=1}^K$ for minimizing total transmit power P_{min} is derived according to the system model introduced in Section III. The purpose of the whole process is to achieve an UAV aided on-demand, power-effective, latency-free network services.

D. CELLULAR DEMAND PREDICTION

In this part, a simple BP neural network model is utilized to predict the future cellular demand in an hour. A brief

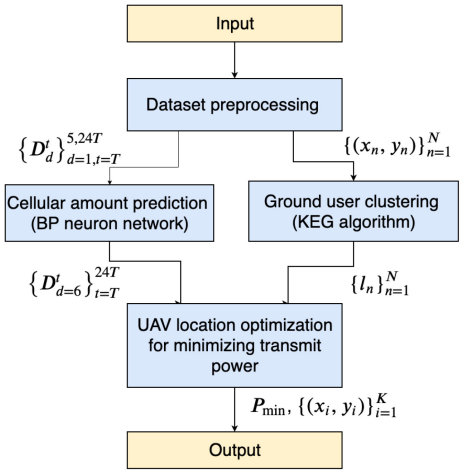

FIGURE 2. The Logical Procedure Diagram of UAV Predictive Deployment.

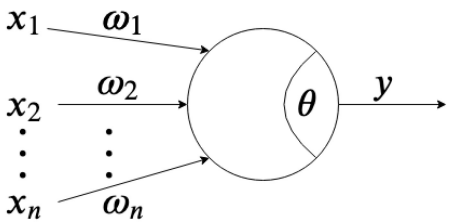

FIGURE 3. M-P Neuron Model.

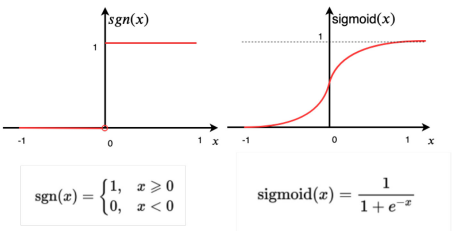

FIGURE 4. The Step Function and the Sigmoid Function.

description of the BP neural network we are using is shown as below:

1) THE NEURON MODEL

For our BP neural network, we model the neuron as McCulloch and Pitts model (M-P) shown in Fig. 3 with the activation function modeled as sigmoid function shown in Fig. 4. The single BP neuron can be formulated as:

$$y = f(\boldsymbol{w}\boldsymbol{x} - \theta) \tag{14}$$

where $\boldsymbol{x} = \{x_i\}_{i=1}^{n}$ is the input of the neuron, $\boldsymbol{w} = \{w_i\}_{i=1}^{n}$ is its corresponding weight, θ is a threshold of activation, and $f(.)$ is an activation function.

With sufficient training data sets, the neural network can learn the appropriate weights and threshold. The threshold can be seen as a dummy node with fixed input 1, so the learning of weights and threshold can be joined together. For a single training data set (x, y), the current output is $\hat{y}$

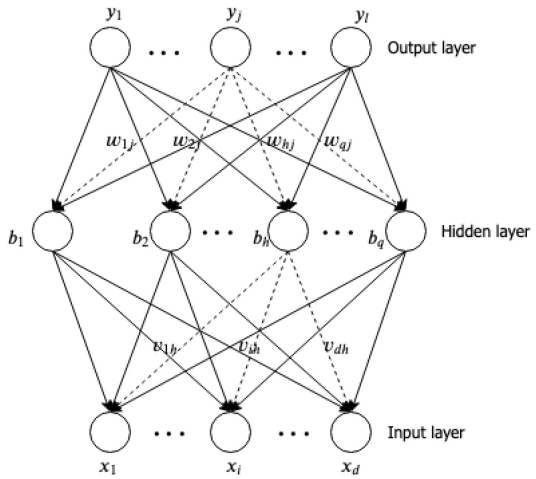

FIGURE 5. Multilayer Network Sketch.

for the input x. Weight updating can be shown as:

$$w_i \leftarrow w_i + \triangle w_i \tag{15}$$

$$\triangle w_i = \eta(y - \hat{y})x_i \tag{16}$$

where $\eta \in (0, 1)$ is a learning rate. In this way, the weight will be adjusted until $y = \hat{y}$. This is a single functional neuron, and its learning ability is very limited. Thus, we always add the hidden layers to compose a multi-layer feedforward neural network.

2) ERROR BACKPROPAGATION

For multi-layer networks, BP(backpropagation) is one of the most successful learning algorithms. The basic structure of the neural networks consists of an input layer and an output layer with multiple hidden layers between them. For a clear and concise description, the number of hidden layers is set to 1.

The input layer is $\boldsymbol{x} = \{x_i\}_{i=1}^{d}$, the hidden layer is $\boldsymbol{b} = \{b_h\}_{h=1}^{q}$, and the output layer is $\hat{\boldsymbol{y}} = \{\hat{y}\}_{j=1}^{l}$. The weight between the input layer node i and the hidden layer node h is v_{ih}, the weight between the hidden layer node h and the output layer node y is w_{ih}. The variable symbols and the BP network model with three layers are shown in Fig. 5, and the weights are also represented.

For h^{th} node of hidden layer and j^{th} node of output layer, the inputs and the outputs are shown below:

$$\begin{cases} \alpha_h = \sum_{i=1}^{d} v_{ih} x_i \\ b_h = f(\alpha_h - \gamma_h) \\ \beta_j = \sum_{h=1}^{q} w_{hj} b_h \\ \hat{y}_j = f(\beta_j - \theta_j) \end{cases} \tag{17}$$

Then, the mean square error can be obtained as follows, where $1/2$ is for the convenience of the subsequent calculation.

$$E = \frac{1}{2} \sum_{j=1}^{l} (\hat{y}_j - y_j)^2 \tag{18}$$

Algorithm 1 BP Algorithm [49]

Input: A training data set $\boldsymbol{S} = \{(\boldsymbol{x}_k, \boldsymbol{y}_k)\}_{k=1}^{m}$, a learning rate η.

Randomly initiate all weights and thresholds w, θ, v, γ

repeat

for all $(\boldsymbol{x}_k, \boldsymbol{y}_k) \in \boldsymbol{S}$ **do**

Calculate the current output $\hat{y}_k$ using Equation (17)

Calculate intermediate variable of gradient g and e using Equation (22) and (23)

Update weights and threshold w, θ, v, γ using Equation (24)

end for

until The stop condition is reached

Output: Weights and thresholds w, θ, v, γ

Because BP algorithm is based on gradient descent algorithm, parameters are adjusted according to the negative gradient. Taking the gradient of output layer as an example, the change of weight can be written as:

$$\triangle w_{hj} = -\eta \frac{\partial E}{\partial w_{hj}} \tag{19}$$

Besides, the chain rule can make the formula as:

$$\frac{\partial E}{\partial w_{hj}} = \frac{\partial E}{\partial \hat{y}_j} \cdot \frac{\partial \hat{y}_j}{\partial \beta_j} \cdot \frac{\partial \beta_j}{\partial w_{hj}} \tag{20}$$

$$\frac{\partial \beta_j}{\partial w_{hj}} = b_h \tag{21}$$

Next, we calculate the negative error gradient of the output layer, which is also the second layer of the network. The intermediate variable of gradient can be obtained as:

$$g_j = -\frac{\partial E}{\partial \hat{y}_j} \cdot \frac{\partial \hat{y}_j}{\partial \beta_j} = \hat{y}_j\big(1 - \hat{y}_j\big)\big(y_j - \hat{y}_j\big) \tag{22}$$

Therefore, the gradient of the hidden layer is similar,

$$e_h = -\frac{\partial E}{\partial b_h} \cdot \frac{\partial b_h}{\partial \alpha_h} = b_h(1 - b_h) \sum_{j=1}^{l} w_{hj} g_j \tag{23}$$

Finally, we can get the update of weights and thresholds for the hidden and output layer,

$$\begin{cases} \triangle w_{hj} = \eta g_j b_h \\ \triangle \theta_j = -\eta g_j \\ \triangle v_{ih} = \eta e_h x_i \\ \triangle \gamma_h = -\eta e_h \end{cases} \tag{24}$$

Note that the aim of BP algorithm is to minimize the accumulated error for all training data sets. The workflow of BP algorithm is represented in Algorithm 1. The stop condition is that the iteration has reached the peak value or the error E is smaller than a minimum threshold.

However, BP algorithm has some inadequacies. Overfitting problem is very obvious because of the powerful representation of BP networks; early stopping method and regularization method are often used to prevent this situation. The limitation of local minimum is another problem; people usually utilize simulated annealing, genetic algorithms and random gradient descent algorithms to solve it [49].

E. GROUND USER CLUSTERING

Ground user clustering is a key step in UAV deployment, which also means the partition of UAV aerial cells. In order to satisfy the fairness and globality of division, we adopt a KEG algorithm to implement service area classification. To demonstrate the algorithm practicability, we use the topology information part of the City Cellular Traffic Map [47] as the real data set.

1) THE K-MEANS ALGORITHM

The K-means algorithm is the most basic non-hierarchical iterative clustering algorithm, belonging to unsupervised learning. The goal of this algorithm is to cluster data points with very low inter-cluster similarity as well as very high intra-cluster similarity [50]. The similarity often denotes the distances between data points.

Given a data set $\boldsymbol{X} = \{\boldsymbol{x}_n\}_{n=1}^{N}$ composed of N instances of M-dimensional variables $\boldsymbol{x}_n$. We divide this data set into K clusters, each cluster has a centroid. At first, we set the centroids of clusters to be K vectors $\{\boldsymbol{\mu}_k\}_{k=1}^{K}$ with M dimensions. These vectors usually randomly take K variables from the given data set $\boldsymbol{X}$ for initialization. To realize the algorithm goal, each data point is supposed to be as close as possible to its cluster centroid, and the distances of each data point and other cluster centroids should be as large as possible. Thus, we set the point cluster label $\boldsymbol{r}_{nk}$ to indicate which cluster the data point $\boldsymbol{x}_n$ belongs to; this variable can be formulated as:

$$\boldsymbol{r}_{nk} = \begin{cases} 1 \text{ if } k = \arg\min_i \|\boldsymbol{x}_n - \boldsymbol{\mu}_i\| \\ 0 \text{ otherwise} \end{cases} \tag{25}$$

Then, we assign every data point to its nearest cluster and update the point cluster label after the distances between each data point and cluster centroids are calculated. The third step is to update the cluster centroids according to those data point labels. The specific K-means algorithm is shown as Algorithm 2.

On the one side, although the K-means algorithm has the capability to cluster data points, it cannot find the latent variables corresponding to the observed data. On the other hand, the principle of K-means is simple, and it is easy to implement the simulation due to its fast convergence and good clustering effect. It can be a practical data initializer for other complex algorithms.

2) THE EM ALGORITHM

The EM algorithm has the ability to recognize the important role of latent variables in a joint distribution. Its goal is to acquire maximum likelihood results for the models with latent variables [50].

In a mixed distribution model, given a sample data set $\boldsymbol{X} = \{\boldsymbol{x}_n\}_{n=1}^{N}$ with an unknown latent variable set $\boldsymbol{Z} = \{z_n\}_{n=1}^{N}$,

Algorithm 2 K-Means Algorithm [50]

Input: The cluster number K, the data point set $\boldsymbol{X} = \{\boldsymbol{x}_n\}_{n=1}^{N}$.

Initialize $\{\boldsymbol{\mu}_k\}_{k=1}^{K}$ as K variables chose from $\boldsymbol{X}$ randomly, initialize $\boldsymbol{r}_{nk}$ as an $n * k$ all-zero matrix
repeat
  for all $\boldsymbol{x}_n \in X$ **do**
    Allocate each data point $\boldsymbol{x}_n$ to cluster $k^* = \arg\min_i \|\boldsymbol{x}_n - \boldsymbol{\mu}_i\|$
    Update point labels $\boldsymbol{r}_{nk}(n, k^*) = 1$
    Calculate cluster centroids $\boldsymbol{\mu}_k = \sum \boldsymbol{r}_{nk}\boldsymbol{x}_n / \sum \boldsymbol{r}_{nk}$, $k = 1, ..., K$
  end for
until $\{\boldsymbol{\mu}_k\}_{k=1}^{K}$ will not be changed.

Output: The point cluster label $\boldsymbol{r}_{nk}$, the cluster centroids $\{\boldsymbol{\mu}_k\}_{k=1}^{K}$

Algorithm 3 EM Algorithm [50]

Input: The observed variable set $\boldsymbol{X}$

Initialize the parameter set $\boldsymbol{\theta}_{old}$,
repeat
  E step Calculate the posterior probability $p(\boldsymbol{Z}|\boldsymbol{X}, \boldsymbol{\theta}_{old})$
  M step Calculate $\boldsymbol{\theta}_{new} = \arg\min_{\boldsymbol{\theta}} \sum_{\boldsymbol{Z}} p(\boldsymbol{Z}|\boldsymbol{X}, \boldsymbol{\theta}_{old}) \ln p(\boldsymbol{X}, \boldsymbol{Z}|\boldsymbol{\theta})$
  Update the parameter set $\boldsymbol{\theta}$, $\boldsymbol{\theta}_{old} \leftarrow \boldsymbol{\theta}_{new}$
until The log likelihood function $\mathcal{L}(\boldsymbol{\theta})$ converges.

Output: The posterior probability $p(\boldsymbol{Z}|\boldsymbol{X}, \boldsymbol{\theta})_{old}$, the parameter set $\boldsymbol{\theta}_{old}$

we want to find the suitable parameter set $\boldsymbol{\theta}$ to well describe the joint distribution $p(\boldsymbol{X}|\boldsymbol{\theta}) = \sum_{\boldsymbol{Z}} p(\boldsymbol{X}, \boldsymbol{Z}|\boldsymbol{\theta})$ [50]. However, the observed variable set $\boldsymbol{X}$ and the latent variable set $\boldsymbol{Z}$ are determined by parameter set $\boldsymbol{\theta}$. Since we are only given the incomplete data set $\boldsymbol{X}$ without $\boldsymbol{Z}$, so we cannot get the optimal parameter set $\boldsymbol{\theta}$. To facilitate analysis, a log likelihood function is defined as:

$$\mathcal{L}(\boldsymbol{\theta}) = \ln\{p(\boldsymbol{X}|\boldsymbol{\theta})\} = \ln\left\{\sum_{\boldsymbol{Z}} p(\boldsymbol{X}, \boldsymbol{Z}|\boldsymbol{\theta})\right\} \tag{26}$$

And a posterior distribution $p(\boldsymbol{Z}|\boldsymbol{X}, \boldsymbol{\theta})$ regarding the latent variable set $\boldsymbol{Z}$ is introduced. Thus, our goal is changed to get the maximum likelihood function $p(\boldsymbol{X}|\boldsymbol{\theta})$ with regard to $\boldsymbol{\theta}$. The iteration of EM mainly consists of two steps: Expectation step (E step) and Maximum step (M step). In the E step, we evaluate the posterior probability $p(\boldsymbol{Z}|\boldsymbol{X}, \boldsymbol{\theta})$; In the M step, we operate the log likelihood maximization to update parameter set $\boldsymbol{\theta}$ [50]. The iteration stops until the convergence of the log likelihood function is checked. The specific algorithm is shown in Algorithm 3.

The EM algorithm can also have good performance in the situation of missing some observed variable values, the observed variable distribution can be acquired by marginalizing those missing value and taking the whole joint variable distribution. In this case, the data returned by the sensor which has some values missing can also be well processed. Therefore, in the scenario of our ground user clustering, the EM algorithm is a useful method to find the latent variables in the data set and classify users in a fair manner.

3) THE KEG ALGORITHM

In this paper, the cellular traffic distribution is complex and time-varying, but a GMM, a linear Gaussian-component superposition model has a remarkable advantage of abundantly representing data distribution. We model the cellular traffic distribution by the GMM as:

$$p(\boldsymbol{X}) = \sum_{k=1}^{K} \pi_k \mathcal{N}\left(\boldsymbol{X}|\boldsymbol{\mu}_k, \boldsymbol{\sigma}_k\right) \tag{27}$$

where $\boldsymbol{X} = \{\boldsymbol{x}_n\}_{n=1}^{N}$ is the topology information of the whole area and $\boldsymbol{x}_n$ is every data point with M dimensions. $p(.)$ represents a probability function, K is the Gaussian component number and $k \in \{1, \ldots, K\}$ denotes a specific Gaussian component, the mixing coefficient π equals one or zero has $\sum_{k=1}^{K} \pi_k = 1$, $\boldsymbol{\mu} = \{\boldsymbol{\mu}_n\}_{n=1}^{N}$ is the mean values corresponding to the cluster centroids with M dimensions, $\boldsymbol{\sigma} = \{\boldsymbol{\sigma}_n\}_{n=1}^{N}$ is the covariance with M dimensions. Besides, in a GMM, the latent variables are discrete.

To this end, we introduce the KEG algorithm based on the K-means algorithm and the EM algorithm. In the KEG algorithm, the EM part aims to find the discrete latent variables and the suitable parameters for analyzing data distribution and clustering data set. Even if a data set is incomplete with some values missing, the EM can also process the data set in a suitable manner. In the EM part, the value of the log likelihood function will increase with the number of iterations rising. When the log likelihood function does not change anymore, the current parameters are the aims we want to obtain. But since this algorithm usually needs many iterations to reach the convergent point, the K-means part is utilized as a data initializer to provide appropriate and rational initialized values. The integrated KEG algorithm is demonstrated in Algorithm 4.

When the parameters are obtained, the predicted cellular traffic amount data is used as the input, only the 4^{th} step and 5^{th} step of Algorithm 4 are operated, we can get the cluster label $\{l_n\}_{n=1}^{N}$of the data points $\{\boldsymbol{x}_n\}_{n=1}^{N}$ that indicates which cluster the data point belongs to.

F. THE REASON OF CHOOSING K-MEANS, EM AND GMM

In this study, we chose the K-means algorithm for clustering due to its simplicity, efficiency, and wide applicability in various fields. K-means is a well-established method for partitioning data into clusters based on similarity, which makes it suitable for our application in dividing the service area into UAV service regions. However, we recognize that K-means alone may not be sufficient to handle data with mixed distribution, and it may not provide the most accurate and flexible clustering results.

Algorithm 4 KEG Algorithm

Input: The topology data set $\boldsymbol{X} = \{\boldsymbol{x}_n\}_{n=1}^{N}$, the clustering number K

Set a variable D as the dimension of $\boldsymbol{x}_n$, initialize the means $\boldsymbol{\mu}_k$ as a 1-by-D matrix by using the K-means algorithm, the covariance $\boldsymbol{\sigma}_k$ as a D-by-D identical matrix and mixing coefficients $\pi_k = 1/K$, $\forall k \in \{1, \ldots, K\}$.
repeat
  for all $k \in \{1, \ldots, K\}$ **do**
    E step: Compute the posterior probability of all $\boldsymbol{x_n}$ by
    $\gamma_{nk} = \pi_k \mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_k, \boldsymbol{\sigma}_k) / \sum_{i=1}^{K} \pi_i \mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_i, \boldsymbol{\sigma}_i)$
    For every $\boldsymbol{x}_n$, allocate it to the cluster $l_n = \arg\max_k \gamma_{nk}$
    M step: Calculate the new parameters
    $\boldsymbol{\mu}_k^{new} = \sum_{n=1}^{N} \gamma_{nk} \boldsymbol{x}_n / \sum_{n=1}^{N} \gamma_{nk}$
    $\boldsymbol{\sigma}_k^{new} = \sum_{n=1}^{N} \gamma_{nk} (\boldsymbol{x}_n - \boldsymbol{\mu}_k^{new})(\boldsymbol{x}_n - \boldsymbol{\mu}_k^{new})^T / \sum_{n=1}^{N} \gamma_{nk}$
    $\pi_k^{new} = \sum_{n=1}^{N} \gamma_{nk} / \sum_{k=1}^{K} \sum_{n=1}^{N} \gamma_{nk}$
  end for
  Calculate the log likelihood using $\boldsymbol{\mu}_k^{new}$, $\boldsymbol{\sigma}_k^{new}$ and π_k^{new}, for $k \in \{1, ..., K\}$
  $\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\pi}) = \ln p(\boldsymbol{X}|\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\pi}) = \sum_{n=1}^{N} \ln \{\sum_{k=1}^{K} \pi_k \mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_k, \boldsymbol{\sigma}_k)\}$
  Update the parameters as $\boldsymbol{\mu}_k \leftarrow \boldsymbol{\mu}_k^{new}$, $\boldsymbol{\sigma}_k \leftarrow \boldsymbol{\sigma}_k^{new}$, $\pi_k \leftarrow \pi_k^{new}$
until The log likelihood function $\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\pi})$ is converged.

Output: The parameters $\{\pi_k, \boldsymbol{\mu}_k, \boldsymbol{\sigma}_k\}_{k=1}^{K}$, the cluster labels $\{l_n\}_{n=1}^{N}$.

To address these limitations, we combined K-means with the Expectation Maximization (EM) algorithm and Gaussian Mixture Model (GMM). The combination of these methods enhances the clustering performance by enabling probabilistic interpretation of the data and providing soft clustering results. Specifically, the EM algorithm addresses the limitation of K-means in handling data with mixed distribution, while the GMM generates soft clustering results that can be used as inputs to our optimization model.

We acknowledge that there may be other clustering methods more suitable for our application. However, our approach is not intended to provide the best clustering results per se but rather to generate soft clustering results that can be used as inputs to our optimization model. The effectiveness of our approach should be evaluated in terms of its impact on the performance of the optimization model, rather than the clustering performance itself.

In future work, we plan to conduct further validation studies and sensitivity analysis to investigate the impact of different clustering methods on the performance of our model. This will enable us to better understand the strengths and weaknesses of different methods and make an informed decision about the most suitable method for our application.

G. COMPLEXITY ANALYSIS

The computational complexity of the BP neural network should be $\mathcal{O}(\sum_{i=1}^{L} n_i NT)$, where L is the layer of BP neural network including output layer, n_i is the number of nodes in each layer, N is the size of the input data set, T is the iteration number.

For the KEG algorithm, we derive the computational complexity in two parts. The K-means algorithm has the computational complexity of $\mathcal{O}(NKT_1D)$, where N is the size of the topology data set, K is the cluster number, T_1 is the number of the iterations of K-means algorithm, D is the dimension of each data point. Then, the complexity of EM algorithm should be $\mathcal{O}(NKTD^2)$, where T is the number of iterations of EM algorithm. This is because we need to calculate D-by-D covariance matrix of each cluster centroid in M step. So, the total computational complexity of KEG Algorithm is $\mathcal{O}(NKTD^2)$.

H. UAV LOCATION OPTIMIZATION

After determining the aerial cells using the KEG algorithm, our next aim is to select an optimal location for every UAV so that the minimum transmit power can be obtained. No matter whether UAVs are in a high altitude platform or a low altitude platform, we assume that all UAVs are in the same altitude h, we can formulate this problem as:

$$\min_{x_i, y_i} \quad P_i = Q \iint_{C_i} A_i(x, y) d_i^2(x, y) r_i(x, y) dxdy \tag{28}$$

where $Q = (\frac{4\pi f_c}{c})^2 \frac{Wn_0}{G}$ is not relevant to the locations of BS (x, y) in the service area, $A_i(x, y) = 2^{\frac{TD^t(x,y)}{W}-1}$ is the BS distribution and $D^t(x, y)$ denotes the cellular data amount of BS at (x, y) in a hour in the predictive day, $d_i^2(x, y)$ is the distance between the UAV i and the BS at (x, y), $r_i(x, y)$ is also related to $d_i^2(x, y)$ for the computation of LoS link probability.

According to the Theorem 1 of paper [17], the function to get P_i^{min} is a convex function, UAV optimal locations can be calculated as:

$$x_i^* = \frac{\iint_{C_i} x A_i(x, y) dxdy}{\iint_{C_i} A_i(x, y) dxdy} \tag{29}$$

$$y_i^* = \frac{\iint_{C_i} y A_i(x, y) dxdy}{\iint_{C_i} A_i(x, y) dxdy} \tag{30}$$

where there is a condition $h_i >> (x - x_i)^2 + (y - y_i)^2$ or $h_i << (x - x_i)^2 + (y - y_i)^2$ to be satisfied [17].

At last, we accumulate the transmit power of all operating UAV to get the minimum total power for transmission:

$$P_{min} = \sum_i P_i^{min} \tag{31}$$

III. SIMULATION AND ANALYSIS

A. ILLUSTRATION OF SIMULATION PROCESS

For simulation, we consider two scenarios, one is the entire area given from the raw data, the other is the limited area

TABLE 1. Simulation parameters.

f_c	Carrier frequency	5GHz
n_o	Noise power spectral	-140dBm/Hz
μ_{LoS}	Excessive path loss for LoS link	3dB
μ_{NLoS}	Excessive path loss for NLoS link	23dB
W	Bandwidth	1MHz
h	UAV's altitude	200m
G	Antenna gain	10dB

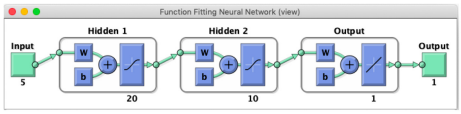

FIGURE 6. Neural Network Structure.

with the relative longitude from 111.055 to 111.07 degrees and the relative altitude from 13.03 to 13.05 degrees. The limited area is the contrast of the entire area. And for a specific contrast, we classify the entire area and the limited area into 8 clusters. Some parameters used in the simulation have been given the specific values which are shown in Table 1 [51]. Moreover, for getting the reasonable simulation results, we assume that all BSs have a basic cellular amount 500 bytes for basic operation, so all the cellular amounts we get from the raw data add 500.

For BP neural network part, we use the *nftool* built-in toolbox of MATLAB to implement the cellular traffic amount prediction. We assume that every BS has the same cellular traffic distribution, we set that the input is the cellular traffic amount of the first six days and the output is the amount of the seventh day. The input in the training data set form is $(\{\boldsymbol{D}_d^t\}_{d=1}^5, \boldsymbol{D}_{d=6}^t)$, where the input layer is the first 5 days cellular amount and the output layer is the 6^{th} day cellular amount.

We choose *trainrp* as the training function and set that the neural network has 2 hidden layers with 20 and 10 neurons respectively, the network structure is shown in Fig. 6. Then we do data training to find the suitable parameters for our neural network model. The user interface of neural network training is shown in Fig. 7.

After training, the input layer $\boldsymbol{x}$ is set as $\{\boldsymbol{D}_d^t\}_{d=2}^6$, then this output layer is what we want to obtain. Furthermore, the ratio of training data is 80%, validation data is 10% and testing data is 10%. The data training is for updating weights, the data validation is to detect the overfitting problem and to avoid it as soon as possible, the data testing is to examine the performance of the neural network. In addition, all the cellular data amounts are normalized using the min-normalization method to make sure every single value is ranged from 0 to 1.

Then, for the application of the KEG algorithm, we give two conditions for finishing the iterations: the first one is the minimum threshold for parameter error of two neighbouring iterations, the second one is the maximum iteration times. If any of these conditions are met, the algorithm ends.

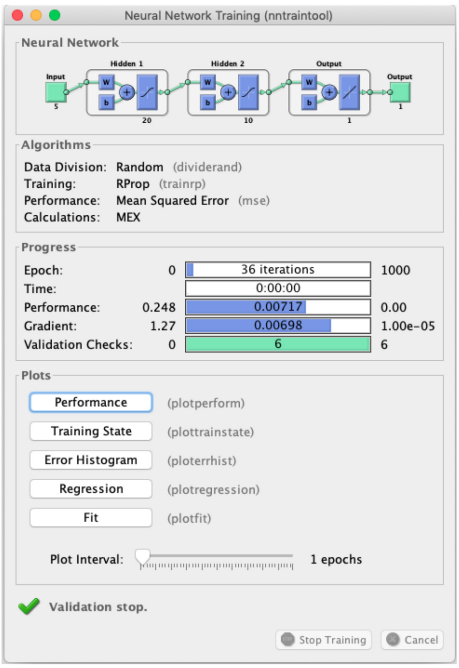

FIGURE 7. Neural Network Training.

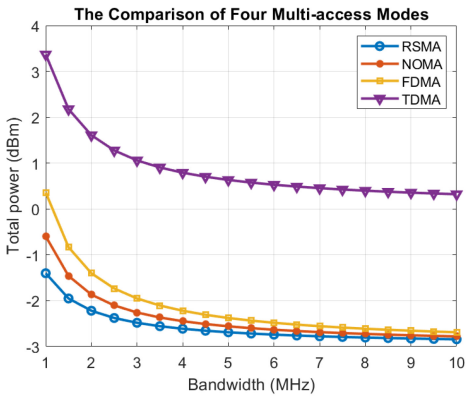

FIGURE 8. Comparison of Four Multi-access Modes.

B. THE COMPARISON OF FOUR MULTI-ACCESS TECHNIQUES

In the proposed UAV deployment framework, we adopt RSMA as the preferred option for uplink transmission due to its excellent robustness and energy efficiency. To provide a comprehensive evaluation, we have compared RSMA with three other widely used multi-access techniques, including FDMA, TDMA, and NOMA, as shown in Fig. 8.

Fig. 8 presents the performance comparison of RSMA, FDMA, TDMA, and NOMA for varying bandwidths. The results indicate that with an increase in bandwidth, the total power consumption of the system decreases for all four multi-access techniques. However, RSMA outperforms the other three techniques and achieves the lowest power consumption for all bandwidths. NOMA is the second-best option, followed by FDMA, and TDMA is the least energy-efficient technique. RSMA achieves a power reduction of 16.9% compared to

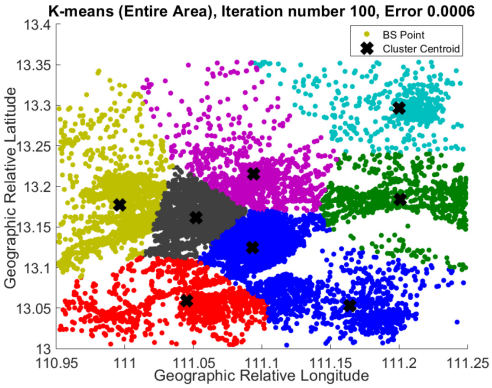

FIGURE 9. K-means for the Entire Area.

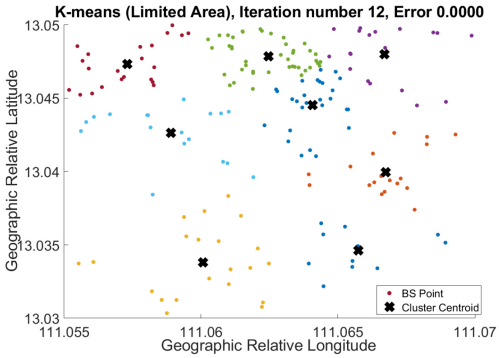

FIGURE 10. K-means for the Limited Area.

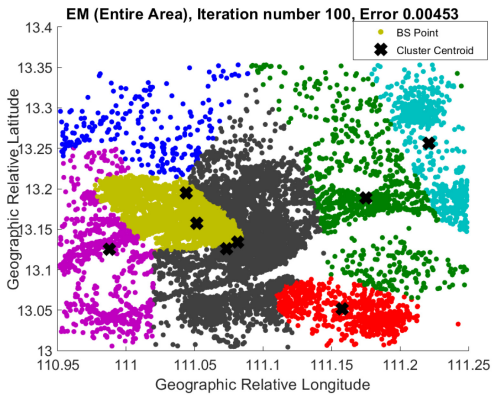

FIGURE 11. EM for the Entire Area.

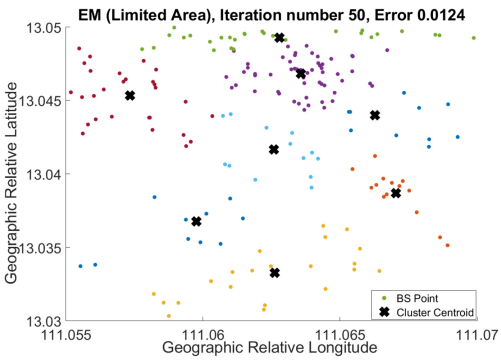

FIGURE 12. EM for the Limited Area.

NOMA, 35.3% compared to FDMA, and 66.4% compared to TDMA. These results demonstrate that RSMA is the most energy-efficient and reliable option for uplink transmission in our proposed framework. By comparing four different multi-access techniques, we show that RSMA is the best choice for our proposed framework, offering the lowest power consumption and highest reliability.

In our study, we have chosen to apply RSMA due to its advantages over other multiple access methods, such as NOMA. Although RSMA might require a more complicated receiver architecture, its benefits in terms of performance and adaptability to the UAV communication systems outweigh the increase in complexity. Specifically, RSMA offers lower complexity in decoding at the base station, as it does not rely on successive interference cancellation (SIC) like NOMA. Furthermore, RSMA requires less stringent channel conditions compared to NOMA, making it more suitable for UAV communication systems.

Moreover, RSMA provides more stable signal quality, as users are spatially isolated, unlike the overlapping of users in the physical layer of NOMA systems. Overall, RSMA enables users to share the same time and frequency resources without complex coordination among them, while capitalizing on interference diversity achieved through the randomness of user locations. This helps to mitigate interference between users and improve the overall system performance.

We acknowledge the concern regarding the low power rule at the UAV and the potential impact of RSMA's receiver complexity. However, we believe that the advantages of RSMA in terms of performance, adaptability, and interference management make it a viable choice for UAV communication systems, despite the increased receiver complexity.

C. THE SIMULATION RESULTS OF THE LIMITED AREA AND THE ENTIRE AREA

In order to verify the usefulness and robustness of the KEG algorithm, we use the data of the limited area and the entire area to do the simulation. The comparison results of the K-means part are shown in Fig. 9 and Fig. 10 and the EM part are shown in Fig. 11 and Fig. 12.

In the Fig. 9, 10, 11 and 12, the scattering points are the locations of BS. The group of the points with the same color represents a cluster which is an aerial cell of a UAV. The black crosses denotes the centroids of every cluster. As we can see, for the K-means part, both the entire area and the limited area have distinct boundaries of service cell and the centroids locate at the clusters with the corresponding colors. But for the EM part, the clusters have the containing and contained states, the centroids may not appear in the clusters with their own colors. The corresponding situation of the limited area is much better than that the entire area.

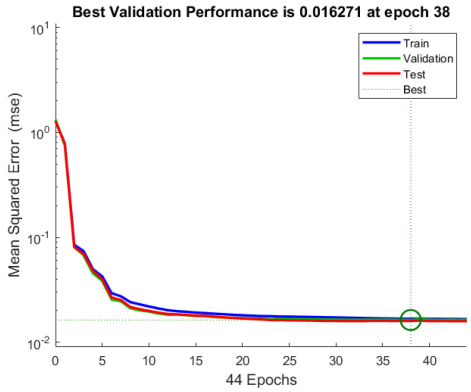

FIGURE 13. Convergence Performance of BP Neural Network for the Entire Area.

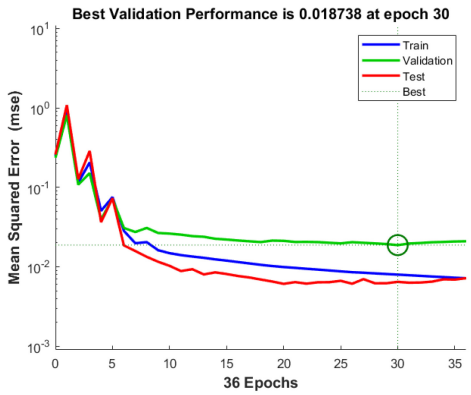

FIGURE 14. Convergence Performance of BP Neural Network for the Limited Area.

D. THE CONVERGENCE PERFORMANCE OF BP NETWORK AND KEG ALGORITHM

The convergence performance of the BP neural network for the entire and limited areas has been evaluated by measuring the mean squared error (MSE), as shown in Fig. 13 and Fig. 14. The results indicate that the network's convergence is stable after approximately 37 and 30 epochs for the entire and limited areas, respectively.

In addition, the convergence performance of the KEG algorithm has been evaluated by measuring the log-likelihood error, as shown in Fig. 15 and Fig. 16. The results indicate that the convergence of KEG is stable after approximately 30 and 20 iterations for the entire and limited areas, respectively. Since KEG is a soft clustering method, we only report the convergence performance of the training set.

These results demonstrate that our proposed framework achieves stable convergence performance for both the BP neural network and KEG algorithm. This contributes to the overall effectiveness of our proposed approach for optimizing UAV positioning and reducing power consumption.

E. THE EFFECT OF PROPOSED UAV SCHEDULING FRAMEWORK

The main goal of UAV deployment in this paper is to minimize the total transmit power in the uplink transmission. Thus, based on this goal, we divide the area into the ground

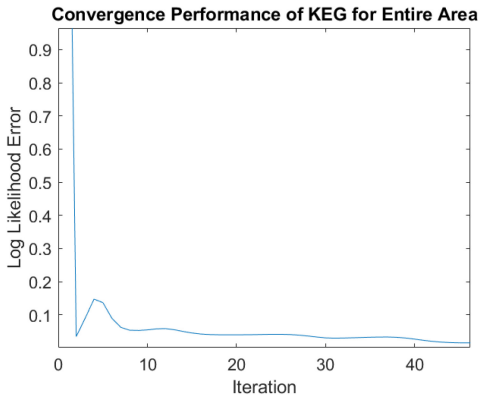

FIGURE 15. Convergence Performance of KEG for Entire Area.

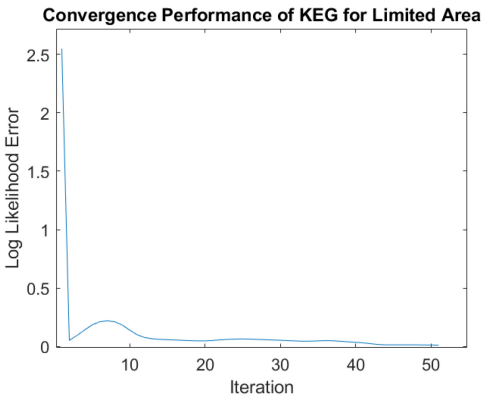

FIGURE 16. Convergence Performance of KEG for Limited Area.

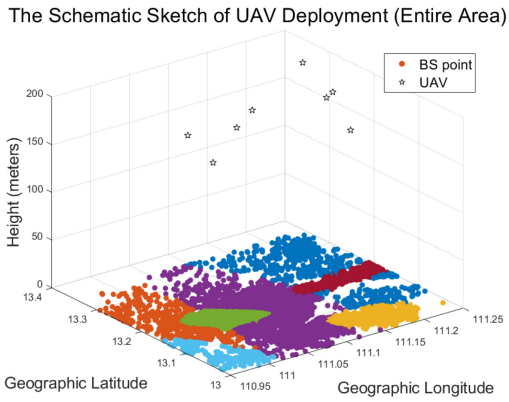

FIGURE 17. Deployment Result for Entire Area.

user clusters for aerial cell classification and determine the optimal locations of UAVs. The Fig. 17 and Fig. 18 shows the schematic sketch of UAV deployment for the entire area and the limited area. As shown in the Fig. 17 and Fig. 18, the number of UAV of the limited area is 8, but the one of the entire area changes into 7 because the system with our proposed framework judges that 7 UAVs are enough to supply the offloaded cellular data amount, this decision is made during the clustering of the K-means part. Because the initial centroid of the merged category

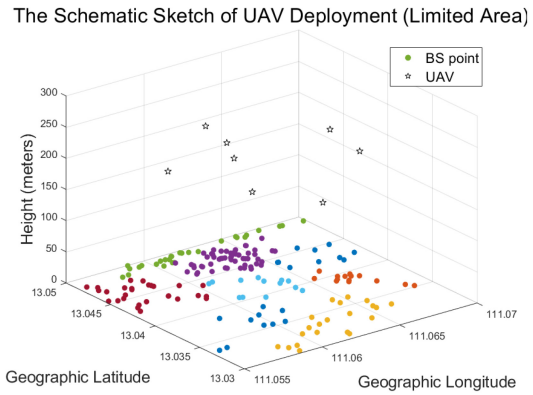

FIGURE 18. Deployment Result for Limited Area.

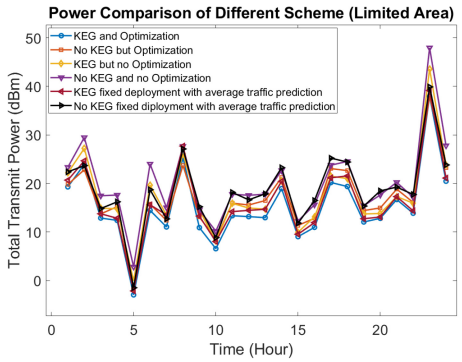

FIGURE 19. Power Comparison for the Limited Area.

chooses an unfavorable position leading to a low similarity in the cluster.

We evaluate the performance of our proposed framework by comparing the total transmit power of six different schemes, including a scheme with KEG and location optimization, a scheme without KEG but with location optimization, a scheme with KEG but without location optimization, a scheme without neither KEG or location optimization, a scheme with KEG but fixed deployment with average traffic prediction, and a scheme without KEG but fixed deployment with average traffic prediction. Specifically, we compare our scheme with KEG and location optimization with the scheme without KEG and with or without location optimization, as well as the scheme with KEG but without location optimization. Moreover, as none of the previous works has been done with dynamic UAV deployment after clustering and traffic deployment, we choose the scheme with fixed UAV deployment as a comparison.

For the limited area, the power comparison of six schemes is represented in Fig. 19. In general, the scheme with KEG and location optimization consume the least total transmit power. The performance of the scheme with no KEG clustering and no location optimization is the worst among the six types of schemes. The remaining four schemes have similar performance in general. Our proposed scheme with KEG clustering and dynamically optimized UAV location

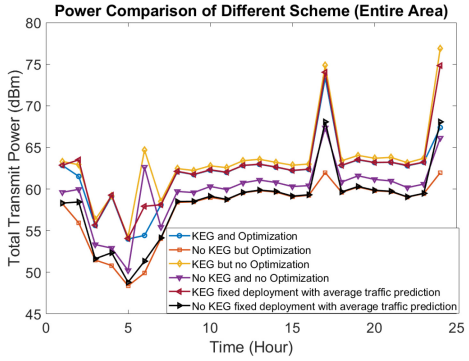

FIGURE 20. Power Comparison for the Entire Area.

reduces power consumption by 24% compared to the worst-performing scheme. Furthermore, when compared with the scheme without KEG clustering and location optimization, and with the scheme with fixed UAV deployment, our scheme with KEG clustering and dynamically optimized location performs the best. For the entire area, six schemes have been contrasted in Fig. 20. Obviously, the scheme without KEG but with location optimization has the best effect for system performance improvement. Next, the scheme without KEG but with fixed deployment with average traffic prediction is better than the one without location optimization but no KEG. The scheme with KEG and location optimization is slightly better than the scheme with KEG but fixed deployment with average traffic prediction. The scheme with KEG but with no location optimization is the worst scheme. The scheme with KEG and location optimization reduce 0.47% power consumption than the worst one.

Based on the above simulation results, our proposed framework is not suitable to apply only several UAVs in the entire area, as the number of UAVs is too small to carry the full range of traffic of an entire city. However, in such situations, our scheme with KEG clustering and location optimization still makes significant contributions to reducing the total transmit power consumption. Moreover, when compared with the scheme without KEG clustering and location optimization, and with the scheme with fixed UAV deployment, our scheme with KEG clustering and dynamically optimized location performs the best. Our UAV deployment framework shows good performance for relatively small areas below dozens of square kilometers, especially for those areas with the approximate GMM cellular data amount distribution.

IV. CONCLUSION

In this paper, we have explored the optimization of UAV locations in an uplink system. Our primary goal was to minimize total power consumption while maintaining efficient communication. To achieve this, we first predicted user traffic distribution by employing a joint K-means and EM algorithm based on the Gaussian Mixture Model (GMM).

This prediction enabled us to accurately determine the optimal locations for UAV deployment.

Our proposed framework not only optimizes UAV locations but also considers the dynamic nature of cellular traffic, ensuring a more adaptive and efficient system. Simulation results demonstrate that our approach significantly reduces total power consumption in comparison to traditional methods such as NOMA, FDMA, and TDMA. This highlights the effectiveness of our method in creating a more energy-efficient and sustainable uplink system for UAV communication.

In conclusion, this paper presents a comprehensive and robust solution for optimizing UAV locations in uplink systems, considering both the prediction of user traffic distribution and the minimization of total power consumption. The results showcase the potential of our approach in improving the performance and sustainability of future UAV communication networks.

REFERENCES

[1] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, May/Jun. 2020, doi: 10.1109/MNET.001.1900287.

[2] J. Lyu, Y. Zeng, and R. Zhang, "UAV-aided offloading for cellular hotspot," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3988–4001, Jun. 2018, doi: 10.1109/TWC.2018.2818734.

[3] M. Chen, H. V. Poor, W. Saad, and S. Cui, "Wireless communications for collaborative federated learning," *IEEE Commun. Mag.*, vol. 58, no. 12, pp. 48–54, Dec. 2020, doi: 10.1109/MCOM.001.2000397.

[4] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2334–2360, 3rd Quart., 2019, doi: 10.1109/COMST.2019.2902862.

[5] M. Mamdouh, M. A. I. Elrukhsi, and A. Khattab, "Securing the Internet of Things and wireless sensor networks via machine learning: A survey," in *Proc. Int. Conf. Comput. Appl. (ICCA)*, 2018, pp. 215–218.

[6] J. Zhang, X. Zhu, and Z. Zhou, "Design of time delayed control systems in UAV using model based predictive algorithm," in *Proc. 2nd Int. Asia Conf. Inf. Control, Autom. Robot. (CAR)*, vol. 1, 2010, pp. 269–272.

[7] Z. Li, M. Chen, C. Pan, N. Huang, Z. Yang, and A. Nallanathan, "Joint trajectory and communication design for secure UAV networks," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 636–639, Apr. 2019, doi: 10.1109/LCOMM.2019.2898404.

[8] Z. Yang et al., "Joint altitude, beamwidth, location, and bandwidth optimization for UAV-enabled communications," *IEEE Commun. Lett.*, vol. 22, no. 8, pp. 1716–1719, Aug. 2018, doi: 10.1109/LCOMM.2018.2846241.

[9] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3039–3071, 4th Quart., 2019, doi: 10.1109/COMST.2019.2926625.

[10] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. NaderiAlizadeh, "Deep CNN-based channel estimation for mmWave massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 989–1000, Sep. 2019, doi: 10.1109/JSTSP.2019.2925975.

[11] Y. Shi, K. Yang, T. Jiang, J. Zhang, and K. B. Letaief, "Communication-efficient edge AI: Algorithms and systems," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2167–2191, 4th Quart., 2020, doi: 10.1109/COMST.2020.3007787.

[12] G. Jia, Z. Yang, H.-K. Lam, J. Shi, and M. Shikh-Bahaei, "Channel assignment in uplink wireless communication using machine learning approach," *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 787–791, Apr. 2020, doi: 10.1109/LCOMM.2020.2968902.

[13] M. Chen, Z. Yang, W. Saad, C. Yin, H. V. Poor, and S. Cui, "A joint learning and communications framework for federated learning over wireless networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 269–283, Jan. 2021, doi: 10.1109/TWC.2020.3024629.

[14] Z. Yang, M. Chen, W. Saad, C. S. Hong, and M. Shikh-Bahaei, "Energy efficient federated learning over wireless communication networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1935–1949, Mar. 2021, doi: 10.1109/TWC.2020.3037554.

[15] Y. Wang and V. Friderikos, "Caching as an image characterization problem using deep convolutional neural networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2020, pp. 1–6.

[16] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1647–1650, Aug. 2016, doi: 10.1109/LCOMM.2016.2578312.

[17] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Optimal transport theory for power-efficient deployment of unmanned aerial vehicles," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2016, pp. 1–6.

[18] L. Wang and S. Zhou, "Energy-efficient UAV deployment with flexible functional split selection," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, 2018, pp. 1–5.

[19] Z. Wang, L. Duan, and R. Zhang, "Adaptive deployment for UAV-aided communication networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4531–4543, Sep. 2019, doi: 10.1109/TWC.2019.2926279.

[20] X. Hou, J. Wang, Z. Fang, Y. Ren, K.-C. Chen, and L. Hanzo, "Edge intelligence for mission-critical 6G services in space-air-ground integrated networks," *IEEE Netw.*, vol. 36, no. 2, pp. 181–189, Mar./Apr. 2022, doi: 10.1109/MNET.121.2100324.

[21] N. Hu et al., "Building agile and resilient UAV networks based on SDN and blockchain," *IEEE Netw.*, vol. 35, no. 1, pp. 57–63, Jan./Feb. 2021, doi: 10.1109/MNET.011.2000176.

[22] Q. V. Do, Q.-V. Pham, and W.-J. Hwang, "Deep reinforcement learning for energy-efficient federated learning in UAV-enabled wireless powered networks," *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 99–103, 2022, doi: 10.1109/LCOMM.2021.3122129.

[23] T. Fang, D. Wu, M. Wang, and J. Chen, "Multi-stage hierarchical channel allocation in UAV-assisted D2D networks: A stackelberg game approach," *China Commun.*, vol. 18, no. 2, pp. 13–26, 2021, doi: 10.23919/JCC.2021.02.002.

[24] Y. Wu et al., "On optimal resource allocation of VAV customized cloud platform based on bi-level programming," in *Proc. IEEE 11th Data Driven Control Learn. Syst. Conf. (DDCLS)*, 2022, pp. 1142–1147, doi: 10.1109/DDCLS55054.2022.9858426.

[25] Q. Zhang, W. Saad, M. Bennis, X. Lu, M. Debbah, and W. Zuo, "Predictive deployment of UAV base stations in wireless networks: Machine learning meets contract theory," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 637–652, Jan. 2021, doi: 10.1109/TWC.2020.3027624.

[26] Y. Wang, M. Chen, Z. Yang, T. Luo, and W. Saad, "Deep learning for optimal deployment of UAVs with visible light communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7049–7063, Nov. 2020, doi: 10.1109/TWC.2020.3007804.

[27] M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1046–1061, May 2017, doi: 10.1109/JSAC.2017.2680898.

[28] S. Q. Zhang, F. Xue, N. A. Himayat, S. Talwar, and H. Kung, "A machine learning assisted cell selection method for drones in cellular networks," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, 2018, pp. 1–5.

[29] S. Gamal, M. Rihan, S. Hussin, A. Zaghloul, and A. A. Salem, "Multiple access in cognitive radio networks: From orthogonal and non-orthogonal to rate-splitting," *IEEE Access*, vol. 9, pp. 95569–95584, 2021, doi: 10.1109/ACCESS.2021.3095142.

[30] Y. Zhang, J. Wang, L. Zhang, Y. Zhang, Q. Li, and K.-C. Chen, "Reliable transmission for NOMA systems with randomly deployed receivers," *IEEE Trans. Commun.*, vol. 71, no. 2, pp. 1179–1192, Feb. 2023, doi: 10.1109/TCOMM.2022.3230847.

[31] W. Jaafar, S. Naser, S. Muhaidat, P. C. Sofotasios, and H. Yanikomeroglu, "On the downlink performance of RSMA-based UAV communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 16258–16263, Dec. 2020, doi: 10.1109/TVT.2020.3037657.

[32] S. K. Singh, K. Agrawal, K. Singh, and C.-P. Li, "Ergodic capacity and placement optimization for RSMA-enabled UAV-assisted communication," *IEEE Syst. J.*, early access, Nov. 17, 2022, doi: 10.1109/JSYST.2022.3220249.

[33] S. K. Singh, K. Agrawal, K. Singh, B. Clerckx, and C.-P. Li, "RSMA enhanced RIS-FD-UAV-aided short packet communications under imperfect SIC," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2022, pp. 1549–1554.

[34] D. T. Hua, Q. T. Do, T. V. Nguyen, C. M. Ho, and S. Cho, "Trajectory design in multi-UAV-assisted RSMA downlink communication," in *Proc. 13th Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, 2022, pp. 1048–1050.

[35] S. K. Singh, K. Agrawal, K. Singh, and C.-P. Li, "Outage probability and throughput analysis of UAV-assisted rate-splitting multiple access," *IEEE Wireless Commun. Lett.*, vol. 10, no. 11, pp. 2528–2532, Nov. 2021, doi: 10.1109/LWC.2021.3106456.

[36] Q. Zhang, M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Machine learning for predictive on-demand deployment of UAVs for wireless communications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–6.

[37] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569–572, Dec. 2014, doi: 10.1109/LWC.2014.2342736.

[38] Z. Yang, M. Chen, W. Saad, W. Xu, and M. Shikh-Bahaei, "Sum-rate maximization of uplink rate splitting multiple access (RSMA) communication," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2019, pp. 1–6.

[39] Z. Yang, M. Chen, W. Saad, and M. Shikh-Bahaei, "Optimization of rate allocation and power control for rate splitting multiple access (RSMA)," *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 5988–6002, Sep. 2021, doi: 10.1109/TCOMM.2021.3091133.

[40] B. Clerckx, H. Joudeh, C. Hao, M. Dai, and B. Rassouli, "Rate splitting for MIMO wireless networks: A promising PHY-layer strategy for LTE evolution," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 98–105, May 2016, doi: 10.1109/MCOM.2016.7470942.

[41] Y. Mao, B. Clerckx, and V. O. Li, "Rate-splitting multiple access for downlink communication systems: Bridging, generalizing, and outperforming SDMA and NOMA," *EURASIP J. Wireless Commun. Netw.*, vol. 2018, no. 1, p. 133, 2018.

[42] Y. Mao, B. Clerckx, and V. O. Li, "Rate-splitting for multi-user multi-antenna wireless information and power transfer," in *Proc. IEEE 20th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, 2019, pp. 1–5.

[43] Y. Mao, B. Clerckx, and V. O. K. Li, "Rate-splitting for multi-antenna non-orthogonal unicast and multicast transmission: Spectral and energy efficiency analysis," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8754–8770, Dec. 2019, doi: 10.1109/TCOMM.2019.2943168.

[44] Y. Mao, B. Clerckx, and V. O. Li, "Energy efficiency of rate-splitting multiple access, and performance benefits over SDMA and NOMA," in *Proc. 15th Int. Symp. Wireless Commun. Syst. (ISWCS)*, 2018, pp. 1–5.

[45] M. Dai and B. Clerckx, "Multiuser millimeter wave beamforming strategies with quantized and statistical CSIT," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7025–7038, Nov. 2017.

[46] B. Clerckx, Y. Mao, R. Schober, and H. V. Poor, "Rate-splitting unifying SDMA, OMA, NOMA, and multicasting in MISO broadcast channel: A simple two-user rate analysis," *IEEE Wireless Commun. Lett.*, vol. 9, no. 3, pp. 349–353, Mar. 2020, doi: 10.1109/LWC.2019.2954518.

[47] X. Chen, Y. Jin, S. Qiang, W. Hu, and K. Jiang, "Analyzing and modeling spatio-temporal dependence of cellular traffic at city scale," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2015, pp. 3585–3591.

[48] U. Paul, A. P. Subramanian, M. M. Buddhikot, and S. R. Das, "Understanding traffic dynamics in cellular data networks," in *Proc. IEEE INFOCOM*, 2011, pp. 882–890.

[49] Z. Zhou, *Machine Learning*. Beijing, China: Tsinghua Univ. Press, 2016.

[50] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[51] Z. Yang et al., "Energy-efficient wireless communications with distributed reconfigurable intelligent surfaces," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 665–679, Jan. 2022, doi: 10.1109/TWC.2021.3098632.