On UAV Serving Node Deployment for Temporary Coverage in Forest Environment: A Hierarchical Deep Reinforcement Learning Approach

WANG Li[1], WU Xuewei[2], WANG Yanhui[2], XIAO Zhe[1], LI Liang[1], and FEI Aiguo[1]

(1. *School of Computer Science (National Pilot Software Engineering School), Beijing University of Posts and Telecommunications, Beijing 100867, China*)

(2. *School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100867, China*)

Abstract — **Unmanned aerial vehicles (UAVs) can be effectively used as serving stations in emergency communications because of their free movements, strong flexibility, and dynamic coverage. In this paper, we propose a coordinated multiple points based UAV deployment framework to improve system average ergodic rate, by using the fuzzy C-means algorithm to cluster the ground users and considering exclusive forest channel models for the two cases, i.e., associated with a broken base station or an available base station. In addition, we derive the upper bound of the average ergodic rate to reduce computational complexity. Since deep reinforcement learning (DRL) can deal with the complex forest environment while the large action and state space of UAVs leads to slow convergence, we use a ratio cut method to divide UAVs into groups and propose a hierarchical clustering DRL (HC-DRL) approach with quick convergence to optimize the UAV deployment. Simulation results show that the proposed framework can effectively reduce the complexity, and outperforms the counterparts in accelerating the convergence speed.**

Key words — **UAV deployment, Forest channel model, Deep reinforcement learning.**

I. Introduction

When natural disasters, e.g., earthquakes, debris flow, and fire, occur in forest areas, rescuers need to detect and deliver the on-site disaster situation information back to the command post, so that headquarters can make accurate decisions based on the perceived information. However, forest areas often lack public network coverage, and the infrastructure is usually damaged by disasters, leading to paralyzed service. For instance, the probability of public network coverage in the forest of Muli County, Sichuan Province is only 30%. Therefore, there is an urgent need for a new way to provide temporary network coverage when forest disasters occur. Unmanned aerial vehicle (UAV) is considered as a promising temporary serving base station (BS) to provide communication coverage and improve network capacity [1]. However, with the limited number of UAVs and their battery capacity, it is very critical and challenging to find an efficient deployment strategy with good performance in terms of network coverage, transmission latency, and transmission reliability [2].

With the continuous maturity of UAV deployment schemes, UAVs are widely applied to public safety, traffic control, environmental monitoring, and other industries. In [3], the authors proposed a Line of Sight (LoS) communication link model between the UAV and the users, and derived the altitude at which a single UAV can be deployed for maximum coverage. In [4], the authors optimized the three-dimensional deployment height of a single UAV to maximize coverage. In [5], the authors proposed an optimal placement algorithm for UAV-BSs that maximizes the number of covered users using the minimum transmission power. In [6], the authors proposed a deployment framework

based on user information to achieve optimal deployment of UAVs with maximum user's quality-of-experience and minimum transmission power. In [7], the authors proposed a new wireless network architecture of coordinated multiple points (CoMP) in the air to deploy UAVs, so as to achieve the maximum average throughput of ground mobile users. In [8], a full-duplex UAV relay was employed to improve the transmission distance and rate, where the UAV positioning, beamforming, and power control were jointly optimized to achieve a near-upper-bound performance. Due to the limited hardware loading on UAV platforms, non-orthogonal multiple access (NOMA) is a promising technology to enhance the access capability for UAV base stations [9], [10]. However, these researches only focus on the urban and rural scenarios. When they are applied in forest areas, the mismatch between the realistic signal fading and the used channel model will degrade the transmission performance dramatically.

Unlike the urban and rural environment, the forest area is covered by dense trees, which makes the multipath fading and non-line-of-sight channel very complex. The signal shielding and reflection of tree branches or leaves are different from that of urban buildings [11]. Forests channel models are usually classified into the horizontal channel model and the slant channel model. Particularly, the horizontal channel models can be further divided into the modified exponential decay (MED) models and the maximum attenuation (MA) models. The MED models, such as Weissberger model [12] and COST-235 model [13], are dating back to the 1960s, which indicates that the foliage loss will exponentially increase with the frequency and the propagation distance. However, these exponentially-decaying models are purely based on empirical measurements, and thus fit poorly for the practical foliage loss. For this reason, the MA model in the current ITU recommendation ITU-R P.833-9 for attenuation was proposed in [14], which indicates the slant channel models and can fit the practical attenuation better in the slant propagation scene than the horizontal ones [15]. However, these works only study channel characteristics and do not discuss how to use the channel models in specific scenarios.

Furthermore, the environment in the disaster area is unknown and time-varying. Traditional methods on optimizing UAV deployment are usually formulated as NP-hard problems with high complexity. With the gradual development of computing capability for UAV platform, the researchers introduced many artificial intelligence (AI) methods, which are powerful tools to determine the UAV deployment strategies in unknown and dynamic environments. In [16], the authors proposed a UAV relay scheme based on both reinforcement learning (RL) and deep reinforcement learning (DRL) techniques to minimize the energy consumption. In [17], the authors jointly determined the trajectory and power allocation of UAV to serve static users, aiming to support dynamic user grouping and bring more flexibility to network design. Moreover, in [18], the authors formulated the UAV deployment problem as a continuous control task and proposed a DRL method to maximize the energy efficiency of the UAV network. However, in the above works, owing to the high dimension of the state and action space of UAVs, the convergence speed of the DRL algorithm becomes slow. Therefore, a hierarchical clustering based deep reinforcement learning (HC-DRL) approach is proposed in this paper to accelerate the algorithm convergence and improve the UAVs' deployment efficiency.

To tackle the challenges above, this paper investigates the multiple-UAV deployment problem in forest environment based on the forest exclusive channel model. In the transmission scheme of CoMP, we adopt joint transmission (JT) technology, by using multiple UAVs to cooperatively transmit data for users, so as to improve the quality of received signals for users and enhance the system performance. The main contributions of this paper are summarized as follows:

• First, we propose a CoMP-based clustering and deployment framework for multiple-UAVs to provide services to users affected by forest disasters, taking into account two cases where the base station is broken and available, respectively. Additionally, to accelerate the convergence, we propose the fuzzy C-means (FCM) algorithm to cluster the ground users for a better initial UAVs positions, and derive an analytical upper bound for the corresponding ergodic rate to reduce computational complexity.

• Second, on the basis of DRL dealing with the complex forest environment, we use a ratio cut method to cluster UAVs and further propose a HC-DRL approach to reduce the computational complexity caused by large action and state space of UAVs.

• Third, we compare the simulation results with the actual measurement results, which indicates that our proposed FSM model is close to the actual channel. It is also validated that the proposed framework and algorithm can effectively reduce the complexity compared with the other schemes.

The remainder of this paper is organized as follows: Section II introduces the system model, problem formulation and and transformation. Section III proposes the HC-DRL scheme. In Section IV, the simulation results are illustrated and analyzed to demonstrate the superiority of the proposed scheme. Finally, Section V concludes this paper.

II. System Model and Problem Formulation

In this section, we describe the deployment scenario of UAVs in the forest environment, in terms of two cases with a broken BS and with an available BS, and establish the system model. After that, we derive the upper bound of the average ergodic rate to simplify the model.

1. System model

The UAV-assisted emergency network in the forest environment is shown in Fig.1. There are N users distributed on the ground, which can be denoted by

$\mathcal{N} = \{1, 2, \ldots, N\}$. Considering the impact of hills in the forest environment, all the ground users have 3-dimensional positions. The 3-dimensional position of the n-th user is defined as $s_n = \{x_n, y_n, h_n\}$. Then the 3-dimensional position of all ground users can be denoted by $\mathcal{S} = \{s_1, s_2, \ldots, s_N\}$. There are K UAVs as temporary BSs for communication, which can be denoted by $\mathcal{K} = \{1, 2, \ldots, K\}$. For simplicity, we assume that all the UAVs fly at a fixed altitude. The 3-dimensional position of the k-th UAV is $u_k = \{x_k, y_k, h\}$. Then the deployment strategy of all the UAVs can be denoted by $\mathcal{U} = \{u_1, u_2, \ldots, u_K\}$.
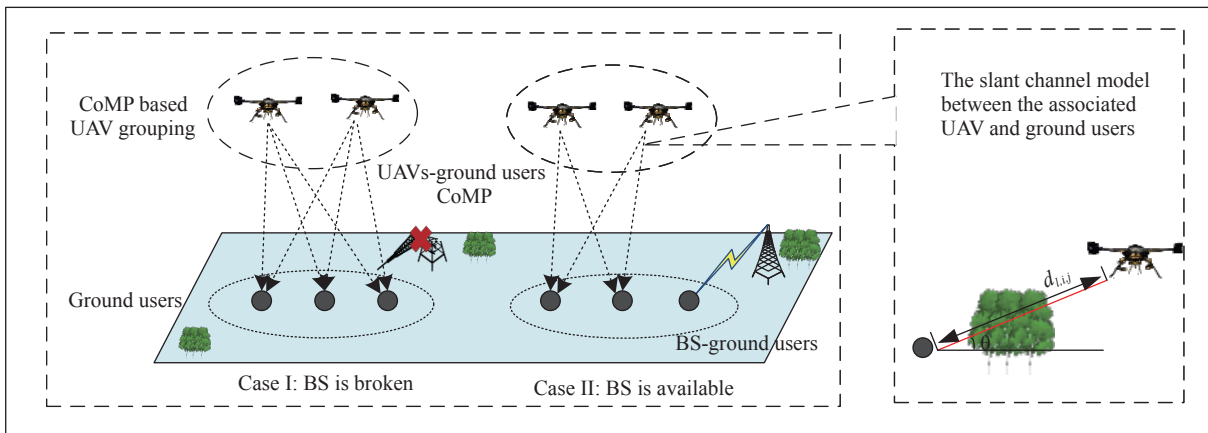

Fig. 1. Scenario description for CoMP based UAV deployment in the forest environment.

Affected by the disasters, the infrastructure is possibly broken. Therefore, the UAV deployment is respectively studied in two cases, i.e., the BS is broken or available. Specifically, there are two feasible communication modes are available for users when the BS is survived. They can download emergency related data from either UAVs or BS. Assume that each user can choose only one communication mode at most and prefers the mode of BS-to-user, since BS is more stable than UAVs that may fly away due to the drained-battery. Furthermore, the CoMP is introduced for the mode of UAV-to-user to further improve the data rate. All the users are assumed to work in the frequency division multiple access (FDMA) manner so that interference between users can be ignored.

There are two aspects that should be taken into account. First, considering the relevance of UAVs, we divide all UAVs into L groups to form CoMP to improve deployment efficiency, which can be denoted by $\mathcal{L} = \{1, 2, \ldots, L\}$. Second, we assign ground users based on UAV CoMP and each UAV CoMP serves its corresponding ground users independently. The number of group users within the covered area of the l-th UAV CoMP is denoted by N_l.

In the following, the two cases are analyzed, respectively.

Case 1　In this case, the ground users can only access to UAVs.

We first introduce the channel model between the UAVs and ground users. As mentioned previously, the forest channel model generally fall into two categories, i.e., the horizontal models and the slant models. In view of the height difference between the UAVs and ground users, the slant models are more practical. Therefore, the path loss between the i-th UAV in the l-th UAV CoMP and the j-th ground user served by this UAV CoMP $L_{l,i,j}^u$ (dB) is given by [14], [15]

$$L_{l,i,j}^u \, (\mathrm{dB}) = L_{l,i,j}^{\mathrm{FSPL}} (d_0) + 10\alpha \lg \left(\frac{d_{l,i,j}}{d_0} \right) + X_\sigma + L_{l,i,j}^{\mathrm{Slant}} \tag{1}$$

where $d_{l,i,j}$ denotes the distance between the i-th UAV in the l-th UAV CoMP and the j-th ground user and X_σ denotes the shadow effect and is a zero-mean Gaussian random variable with a standard deviation of σ dB. Besides, $L_{l,i,j}^{\mathrm{FSPL}} (d_0)$ and $L_{l,i,j}^{\mathrm{Slant}}$ denote the free space path loss (FSPL) and the excess loss in the forest, respectively. They can be expressed as [14]

$$L_{l,i,j}^{\text{FSPL}} = 20\lg\left(4\pi f d_0/c\right) \tag{2}$$

$$L_{l,i,j}^{\text{Slant}} = Af^C d_{l,i,j}^E \left(\theta + G\right)^H \tag{3}$$

where f, d_0, c, α, and θ denote the carrier frequency, reference distance, light speed, path loss exponent, and elevation angle between the UAV and ground user, respectively. A, C, E, G, and H are the multiplication coefficient, the exponential coefficient of the carrier frequency, the exponential coefficient of transmission distance, the correction factor of θ, and the exponential coefficient of θ, respectively, affected by the environments such as the vegetation types and the vegetation density.

Define $\boldsymbol{L}_{l,j}^u = [L_{l,1,j}^u, L_{l,2,j}^u, \ldots, L_{l,Q,j}^u]$ as the channel vector from the l-th UAV CoMP to the j-th user, where $L_{l,i,j}^u = \sqrt{10^{\frac{-L_{l,i,j}^u (\text{dB})}{10}}}$. Then, the signal-to-noise ratio (SNR) between the l-th UAV CoMP and the j-th user $\gamma_{l,j}^u$ can be expressed as [7]

$$\gamma_{l,j}^u = \frac{P_1}{\psi} \boldsymbol{L}_{l,j}^u \left(\boldsymbol{L}_{l,j}^u\right)^{\text{H}} \tag{4}$$

where P_1 and ψ denote the UAV transmission power and the noise power, respectively. Therefore, the ergodic rate of the j-th user served by the l-th UAV CoMP is given by

$$\begin{aligned} R_{l,j}^u &= \mathbb{E}\left[B \log_2\left(1 + \gamma_{l,j}^u\right)\right] \\ &= \mathbb{E}\left[B \log_2\left(1 + \frac{P_1}{\psi} \boldsymbol{L}_{l,j}^u \left(\boldsymbol{L}_{l,j}^u\right)^{\text{H}}\right)\right] \end{aligned} \tag{5}$$

Let a binary variable $x_{l,j} \in \{0,1\}$ indicate the cooperative relationship between UAV CoMP and the corresponding ground users, where $x_{l,j} = 1$ means user j is served by UAV CoMP l; otherwise, $x_{l,j} = 0$. It is given by

$$x_{l,j} = \begin{cases} 1, & P_{l,j}^{\text{CoMP}} < P_{\text{thre}} \\ 0, & P_{l,j}^{\text{CoMP}} \geq P_{\text{thre}} \end{cases} \tag{6}$$

Here, P_{thre} denotes the outage probability threshold, and $P_{l,j}^{\text{CoMP}}$ denotes the outage probability between the j-th ground user and the l-th UAV CoMP, which is calculated by [19]

$$\begin{aligned} P_{l,j}^{\text{CoMP}} &= 1 - \Pr\left(\gamma_{l,j}^u > \gamma_0\right) \\ &= 1 - \Pr\left(\frac{P_1}{\psi} \boldsymbol{L}_{l,j}^u \left(\boldsymbol{L}_{l,j}^u\right)^{\text{H}} > \gamma_0\right) \end{aligned} \tag{7}$$

where γ_0 denotes the SNR threshold.

Therefore, in Case 1, the system throughput, R_1, for all ground users is given by

$$R_1 = \sum_{l=1}^{L} \sum_{j=1}^{N_l} \left(x_{l,j} R_{l,j}^u\right) \tag{8}$$

Case 2 In this case, the ground users can access to either the BS or the UAVs.

We assume that a ground user will access to UAVs only if it is not within the coverage of the base station. Let m indicate the BS. The power gain between the BS and the n-th user is given by

$$L_{m,n}^B (dB) = L_{m,n}^{\text{FSPL}} (d_0) + 10\alpha \lg\left(\frac{d_{m,n}}{d_0}\right) + X_\sigma + L_{m,n}^{\text{Slant}} \tag{9}$$

Hence, the SNR of user n with the BS is expressed as

$$\gamma_{m,n}^B = \frac{P_2 L_{m,n}^B}{\psi} \tag{10}$$

where $L_{m,n}^B = 10^{-\frac{L_{m,n}^B(\text{dB})}{10}}$ and P_2 is the transmission power of BS. Further, the ergodic rate of user n is given by

$$\begin{aligned} R_{m,n}^B &= \mathbb{E}\left[B \log_2\left(1 + \gamma_{m,n}^B\right)\right] \\ &= \mathbb{E}\left[B \log_2\left(1 + \frac{P_2 L_{m,n}^B}{\psi}\right)\right] \end{aligned} \tag{11}$$

Let a binary variable $y_{m,n}$ show whether a user is connected to the BS or not, that is,

$$y_{m,n} = \begin{cases} 1, & \text{if } \gamma_{m,n}^B > \gamma_0 \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

Besides, the ground users can also be connected with UAVs. The UAV-user connection can refer to Case 1. Specially, if a UAV does not form CoMP with other UAVs, its service mode to users is the same as that of the BS. Define $z_{l,j} \in \{0,1\}$ to indicate whether the ground user is served by UAV CoMP, i.e.,

$$z_{l,j} = \begin{cases} 1, & P_{l,j}^{\text{CoMP}} < P_{\text{thre}}, \ y_{m,n} = 0 \\ 0, & P_{l,j}^{\text{CoMP}} \geq P_{\text{thre}} \end{cases} \tag{13}$$

In this instance, the ergodic rate of UAV-user can be calculated by (5).

Therefore, in Case 2, the system throughput, R_2, for all the ground users is given by

$$R_2 = \sum_{n=1}^{N} y_{m,n} R_{m,n}^B + \sum_{l=1}^{L} \sum_{j=1}^{N_l} \left(z_{l,j} R_{l,j}^u\right) \tag{14}$$

2. Problem formulation

We consider the cases where BS is broken and

available with probabilistic β and $(1 - \beta)$, respectively. This paper aims at maximizing the system average ergodic rate by optimizing the deployment strategy $\mathcal{U}$ of all the UAVs and the number of UAV CoMP L. Therefore, the problem can be formulated as

$$\text{P1:} \quad \max_{L,\mathcal{U}} R = \frac{1}{N}\left(\beta R_1 + (1 - \beta) R_2\right) \tag{15a}$$

s.t. System throughput (8), (14) for Case 1, 2, resp. (15b)

$$x_{l,j}, y_{m,n}, z_{l,j} \in \{0,1\}, \forall l \in \mathcal{L}, \forall j \in \mathcal{J}, \forall n \in \mathcal{N} \tag{15c}$$

$$\sum_{l=1}^{L} x_{l,j} \leq 1, \forall l \in \mathcal{L}, \forall j \in \mathcal{J} \tag{15d}$$

$$y_{m,n} + \sum_{l=1}^{L} z_{l,j} \leq 1, \forall l \in \mathcal{L}, \forall j \in \mathcal{J}, \forall n \in \mathcal{N} \tag{15e}$$

$$P_{l,j}^{\text{CoMP}} < P_{\text{thre}}, \quad \forall l \in \mathcal{L}, \forall j \in \mathcal{J} \tag{15f}$$

$$\gamma_{m,n}^{B} > \gamma_0, \quad \forall n \in \mathcal{N} \tag{15g}$$

where N denotes the number of ground users. In the above, formulas (15b)–(15e) restrict the connection status of users, i.e., every user chooses at most only one communication model, formula (15f) represents that the required CoMP outage probability P_{thre} should be satisfied, and formula (15g) denotes that the users associated with the BS should meet the SNR threshold γ_0.

3. Bound relaxations and problem transformation

It is computation-intensive to calculate $R_{l,j}^{u}$ for the UAV CoMP communication mode. To further reduce the computational complexity, the upper bound of $R_{l,j}^{u}$ in formula (5) will be derived and used for problem optimization in this article.

It can be shown that given $q > 0$, $f(x) = \log_2(1 + px)$ is concave over $x > 0$. Then, according to Jensen's inequality [17],

$$f(\mathbb{E}[x]) \leq \mathbb{E}[f(x)] \tag{16}$$

On the basis of the above analysis, the upper bound in formula (5) is derived first. Obviously, $\mathbb{E}[B\log_2(1 + \frac{P_1}{\psi}\boldsymbol{L}_{l,j}^{u}(\boldsymbol{L}_{l,j}^{u})^{\text{H}})]$ is concave with respect to $\boldsymbol{L}_{l,j}^{u}(\boldsymbol{L}_{l,j}^{u})^{\text{H}}$, and $\boldsymbol{L}_{l,j}^{u}(\boldsymbol{L}_{l,j}^{u})^{\text{H}} = ||\boldsymbol{L}_{l,j}^{u}||^2$. Therefore,

$$R_{l,j}^{u} = \mathbb{E}\left[B\log_2\left(1 + \frac{P_1}{\psi}||\boldsymbol{L}_{l,j}^{u}||^2\right)\right]$$
$$\leq B\log_2\left(1 + \frac{P_1}{\psi}\mathbb{E}\left[||\boldsymbol{L}_{l,j}^{u}||^2\right]\right) \tag{17}$$

Hence, the upper bound of $R_{l,j}^{u}$ in (5) can be expressed as

$$\bar{R}_{l,j}^{u} = B\log_2\left(1 + \frac{P_1 \cdot \mathbb{E}\left[||\boldsymbol{L}_{l,j}^{u}||^2\right]}{\psi}\right) \tag{18}$$

As mentioned before, the upper bound, $\bar{R}_{l,j}^{u}$, can be used instead to reduce the computational complexity. Since the bound relaxation is exploited, the system throughput in Case 1 and Case 2, i.e., R_1 and R_2, can be rewritten as

$$R_1' = \sum_{l=1}^{L}\sum_{j=1}^{N_l}\left(x_{l,j}\bar{R}_{l,j}^{u}\right) \tag{19}$$

$$R_2' = \sum_{n=1}^{N} y_{m,n}R_{m,n}^{B} + \sum_{l=1}^{L}\sum_{j=1}^{N_l}\left(z_{l,j}\bar{R}_{l,j}^{u}\right) \tag{20}$$

Hence, the original optimization problem P1 related to the average ergodic rate is transformed to P2 with respect to the upper bound of ergodic sum rate, which can be formulated as

$$\text{P2:} \quad \max_{L,\mathcal{U}} R = \frac{1}{N}[\beta(R_1') + (1 - \beta)(R_2')] \tag{21a}$$

$$\text{s.t. (15c)–(15g), (19), (20)} \tag{21b}$$

III. Hierarchical Clustering Based DRL Scheme

In this section, we present the design of Hierarchical Clustering based DRL (HC-DRL) scheme for UAV deployment control and discuss its design rationale.

1. Overall workflow

We develop a hierarchical clustering based DRL scheme to determine the UAVs' deployment locations in low complexity, of which the core idea is to shrink the state space in the DRL process by incorporating two-tier clustering, i.e., ground users clustering and UAVs clustering. As illustrated in Fig.2, in Step 1, we cluster the ground users based on FCM and deploy the UAVs at the center of the user clusters as the initial positions of the UAVs. We calculate the outage probability between UAV and users to obtain the service coverage of each UAV. In Step 2, we use the ratio cut method to cluster the UAVs so that UAVs in a cluster can form CoMP to serve corresponding ground users. In Step 3, we merge the corresponding user clusters of the UAVs which belong to the same UAV cluster, so that each UAV cluster corresponds to the user cluster one by one. The key of the scheme is the design of hierachical clustering and HC-DRL algorithm, which we will discuss next.

2. FCM based ground users clustering

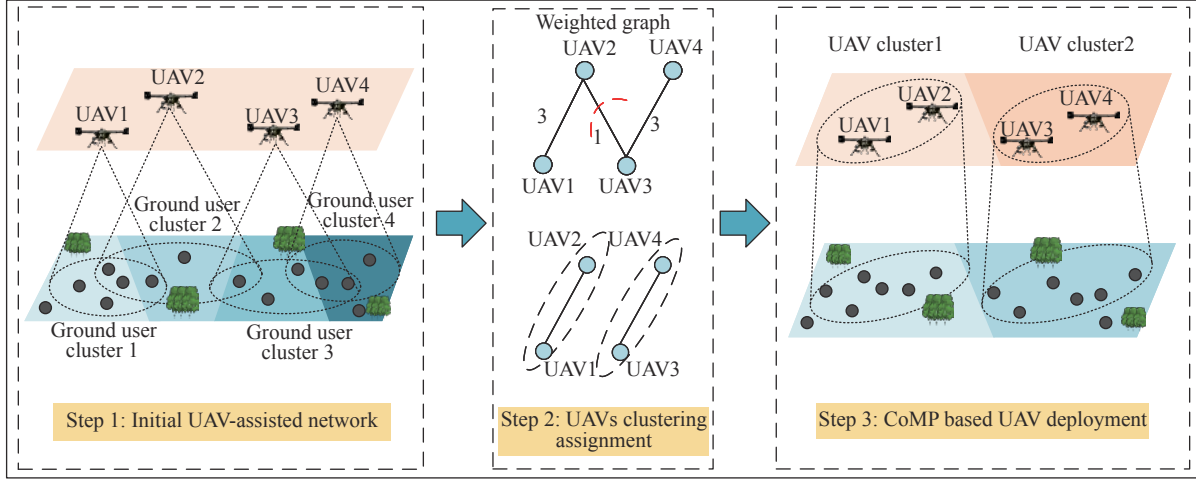A critical issue that is commonly overlooked in ex-

Fig. 2. The process for ratio cut method based UAV clustering and UAV deployment.

isting DRL-based UAV deployment algorithms is to initialize the UAVs' positions, where the initial positions of UAVs are generated randomly. However, such random initialization may result in unsatisfactory performance and consume a relatively long time to converge. To circumvent this drawback, we propose to use the FCM [20] clustering-based initialization to generate the initial positions of UAVs. Compared with the classic clustering algorithm, such as K-means, FCM falls into the soft-clustering category. It utilizes the weight between 0 and 1, rather than the deterministic zero or one, to express the association between the data point and different clusters. More specifically, the FCM clustering-based algorithm is employed to group N users into K clusters by minimizing the following sum of squared error denoted by ϵ:

$$\epsilon = \sum_{k=1}^{K} \sum_{n=1}^{N} \omega_{k,n}^{q} \|s_n - c_k\|^2 \tag{22a}$$

$$\text{s.t.} \sum_{k=1}^{K} \omega_{k,n} = 1, \quad n = 1, 2, \ldots, N \tag{22b}$$

where $\omega_{k,n}$ denotes the weight that the n-th user belongs to the k-th cluster. q is the exponent. c_k is the cluster center of the k-th cluster. The update formulas of $\omega_{k,n}$ and c_k are as follows

$$c_k = \frac{\sum_{n=1}^{N} \omega_{k,n}^{q} \cdot s_n}{\sum_{n=1}^{N} \omega_{k,n}^{q}} \tag{23}$$

$$\omega_{k,n} = \frac{\dfrac{1}{\|s_n - c_k\|^{\frac{2}{q-1}}}}{\sum_{k=1}^{K} \dfrac{1}{\|s_n - c_k\|^{\frac{2}{q-1}}}} \tag{24}$$

The FCM algorithm is shown in Step 1 of Algorithm 1. Therefore, all the initial positions of UAVs can be expressed as $\mathcal{C}_K = \{c_1, c_2, \ldots, c_K\}$.

Algorithm 1 Hierarchical clustering for UAVs and ground users

---

Input: the position of all ground users $\mathcal{S}$, the number of UAVs K, iter$_{\max}$, q, $Q_{\max}$.

Output: the initial position of UAVs $\mathcal{U}$, the number of UAV clusters L.

Step 1: FCM based Ground User Clustering

  Divide all ground users into K clusters according to the number of UAVs;

  Initialize $\omega_{k,n}$ subject to $\sum_{k=1}^{K} \omega_{k,n} = 1, n = 1, 2, \ldots, N$;

  Set $t = 1$.

  while $t \leq$ iter$_{\max}$ **do**

    Calculate the center positions $\mathcal{C}_K = \{c_1, c_2, \ldots, c_K\}$ of all the clusters according to (23), where $c_k = \{x_k, y_k, h_k\}$;

    Update $\omega_{k,n}$ according to (24);

    $t = t + 1$;

  end

Step 2: Ratio Cut based UAV Clustering

  for $k = 1, 2, \cdots, K$ **do**

    Initialize 3-dimensional position of the k-th UAV u_k with $\{x_k, y_k\}$ of c_k and the fixed height h;

  end

  Obtain the number of UAV clusters $L = L_{\text{initial}}$ according to (25);

  repeat

    Construct the weighted graph of UAVs;

    Divide UAVs into L clusters using the ratio cut method;

    $L = L + 1$;

  until the number of UAVs per cluster no more than $Q_{\max}$

---

3. Ratio cut based UAVs clustering

The considered problem in (21) can be solved by using a centralized optimization method. All UAVs are

regarded as an agent, and the state of the agent is defined as the positions of all UAVs. Obviously, the centralized method leads to large-scale action and state spaces, resulting in high computational complexity. To overcome this challenge, this paper proposes a novel solution based on clustering. Specifically, we divide the UAVs into different clusters, where each cluster is regarded as an independent agent to implement a joint optimization of positions. Moreover, the UAV clusters echoe the UAV group to form a CoMP in Section Ⅱ.1.

We take the case where the BS is broken as an example. First, we have obtained the center of ground user clusters $\mathcal{C}_K$ based on FCM in Section Ⅱ.1. Considering the fixed height of UAVs, we initialize the 3-dimensional position of UAVs by utilizing 2-dimensional position $\{x_k, y_k\}$ of c_k. This paper models the UAV network as an undirected graph with edge weights using the spectral clustering method, including the minimum cut method and ratio cut method. Compared to the minimum cut method, the ratio cut is able to avoid the single point being remained, and thereby achieves a more balanced cutting. Therefore, we divide all UAVs into L clusters by means of the ratio cut method. Because the number of UAV clusters L is related to both the system performance and complexity, L should be chosen carefully. Considering the practical computing ability, we limit the maximum number of UAVs in each UAV cluster to $Q_{\max}$ ($Q_{\max}$ can be properly adjusted according to the devices' capability). Then the initial minimum number of UAV clusters L_{initial} is given by

$$L_{\text{initial}} = \lfloor \frac{K}{Q_{\max}} \rfloor \tag{25}$$

As illustrated in Fig.2, we introduce the process of ratio cut method based UAV clustering and UAV deployment. In Step 1, we have determined the initial 3-dimensional positions of the UAVs, the coverage of each UAV, and the results of user clustering. For better distinction, we use different colors to represent different UAV clusters and ground user clusters. Intuitively, the more users that can be served within the overlapping communication coverage of two different UAVs, i.e., users who meet the SNR requirements, the more important to jointly optimize the locations of the two adjacent UAVs.

In Step 2, we first use UAVs and the number of ground users located within the overlapped coverage of two UAVs to denote the vertex and edge weights of the graph, respectively. For example, there are three users located in the overlapped area of UAV1 and UAV2; Therefore, there is an edge between UAV1 and UAV2 with an edge weight 3. The greater the edge weight between the two UAVs, the more users the two adja-

cent UAVs can serve. Then our goal is to cut the graph $G(V, E)$ into L subgraphs that are not connected to each other. The set of subgraphs is expressed as $\mathcal{A} = \{A_1, A_2, \ldots, A_L\}$, for any $l \neq l'$, we have $A_l \cap A_{l'} = \emptyset$, and $A_1 \cup A_2 \cup \cdots \cup A_L = V$. Define the weight sum between the two subgraphs A_l and $A_{l'}$ as

$$W(A_l, A_{l'}) = \sum_{i \in A_l, j \in A_{l'}} \omega_{i,j} \tag{26}$$

The ratio cut method aims to minimize the weighted sum between all the subgraphs and their complements, which can be expressed as

$$\text{cut}(A_1, A_2, \ldots, A_L) = \min \left(\frac{1}{2} \sum_{l=1}^{L} \frac{W(A_l, \bar{A}_l)}{|A_l|} \right) \tag{27}$$

where $\bar{A}_l$ denotes the complement of the subgraph, and $|A_l|$ denotes the number of nodes in the subgraph A_l. At last, the weighted graph is divided into two subgraphs, and each subgraph is regarded as a cluster, i.e., an independent agent in DRL. That is, the deployment scheme of UAVs with more common ground users will be optimized jointly.

In Step 3, after clustering the UAVs, we merge the corresponding ground user clusters. Then UAVs in the same cluster can form CoMP to enhance the communication service to the ground users. Specially, the method of CoMP based UAVs clustering can improve the overall system performance, but it cannot guarantee the service of each user. For example, UAV2 and UAV3 can form CoMP to provide service for a user. According to the ratio cut method, the UAV cluster corresponding to the user is composed of UAV3 and UAV4, and UAV2 will no longer provide service. Since the user is not within the coverage of UAV4, only UAV3 provides service for the user. Although this user's ergodic rate decreased, the system's average ergodic rate increased.

Notice that the proposed UAV clustering method will be repeated with an additional UAV cluster, if any UAV cluster contains UAVs more than $Q_{\max}$. Otherwise, the process of UAV clustering will be terminated. The details of hierarchical clustering method for UAVs and ground users are shown in Algorithm 1.

4. DRL-based UAVs deployment adjustment

Taking the UAV cluster $\mathcal{V}$ as an example, we denote the set of UAVs contained in the cluster $\mathcal{V}$ as $\mathcal{V} = \{1, 2, \ldots, V\}$. The key components of DRL, i.e., state, action, and reward, are defined as follows, and their interactions are shown in Fig.3 .

State   The state at time slot t is expressed as

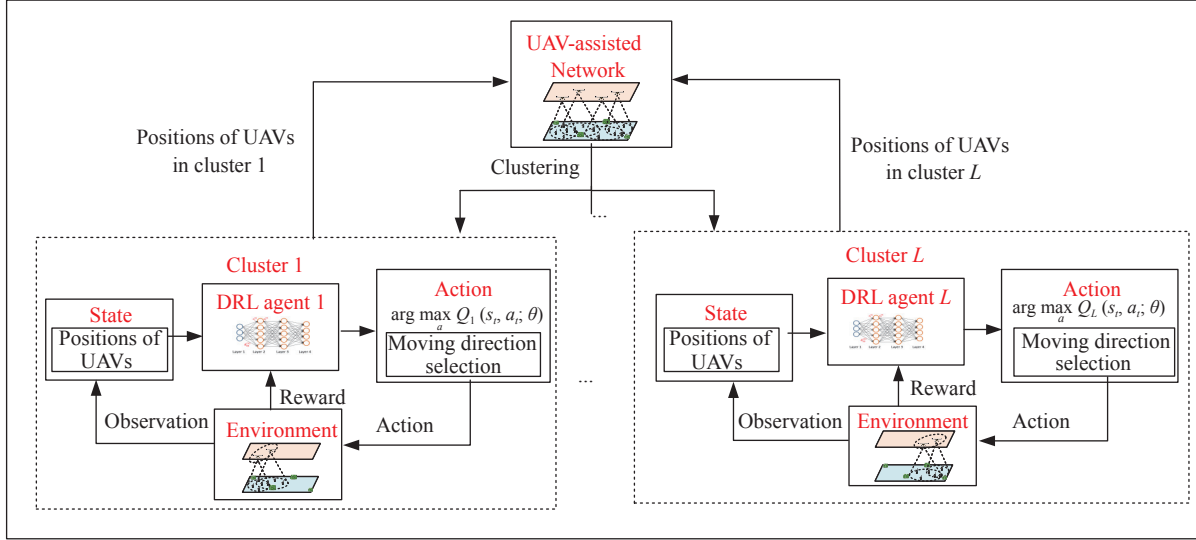$$s_t = [u_1(t), u_2(t), \ldots, u_V(t)] \tag{28}$$

Fig. 3. The detailed framework of HC-DRL.

where $u_v(t)$ denotes the position of UAV v clustered in $\mathcal{V}$. For simplicity, we divide the whole area into fine-grained grids and suppose that UAVs will be deployed at the grid center.

Action The action at time slot t is expressed as

$$a_t = \left\{ a_t^1, a_t^2, \ldots, a_t^V \right\} \tag{29}$$

where a_t denotes all the UAVs movements in the cluster. Here, we discretize the movement of the UAV into 5 optional directions $a_t^v = \{\text{front, behind, left, right, stay}\}$.

Reward The reward at time slot t is related to the throughput at the last time slot t, which can be defined as follows:

$$r_t = \begin{cases} 1, & \text{if } \overline{R(t)} > \overline{R(t-1)} \\ -0.2, & \text{if } \overline{R(t)} = \overline{R(t-1)} \\ -2, & \text{if } \overline{R(t)} < \overline{R(t-1)} \end{cases} \tag{30}$$

It is widely known that Q-learning works well if the state and action spaces of the problem are small, and a Q table can be used to accomplish the update rule. However, this becomes impossible when the state-action space becomes very large. In this situation, many states may be rarely visited, and thus the corresponding Q-values are seldom updated, leading to a much longer time to converge. Deep Q-network combines Q-learning with deep learning. In a deep Q-network, we use a replay memory D to store the training samples (s_t, a_t, r_t, s_{t+1}), and the Q function is approximated by a deep neural network. The basic idea behind the deep Q network is a deep neural network (DNN) function approximator with weights θ as a Q-network. Once θ is given, Q-values, $Q(s_t; a_t)$ will be determined.

The Q-network updates its weights θ at each iteration to minimize the following loss function derived from the same Q-network with old weights on a data set D:

$$\text{Loss}(\theta) = E\left[(\text{Target}Q - Q(s, a; \theta))^2 \right] \tag{31}$$

$$\text{Target}Q = r + \varrho \max_{a'} Q(s', a'; \theta') \tag{32}$$

where ϱ denotes the discount factor. The proposed HC-DRL algorithm is shown in Algorithm 2.

Algorithm 2: Hierarchical Clustering Based Deep Reinforcement Learning (HC-DRL) Algorithm

---

Input: the initial position of UAVs $\mathcal{U}$ and the number of UAV clusters L according to algorithm 1.

Output: the optimal deployment position of UAVs $\mathcal{U}^*$, the system average ergodic rate $\bar{R}$.

  for $l = 1, 2, \ldots, L$ **do**

  Randomly initialize Q-network and target Q-network of each cluster with weights $\{\theta\}$ and $\{\theta'\}$;

  For each cluster, initialize replay memory D with capacity S;

  for epoch $= 1$ to I **do**

  Initialize the beginning state s_0;

  for $t = 1$ to T **do**

  With probability ε select a random action, otherwise select $a_t = \arg\max_a Q(s_t, a_t; \theta)$;

  Execute action a_t in emulator and observe reward r_t and new state s_{t+1};

  Store the experience (s_t, a_t, r_t, s_{t+1}) into the replay memory D;

  Get a random minibatch of M samples $\{(s_i, a_i, r_i, s_{i+1})\}$ from replay memory D;

  Set $\text{Target}Q = r + \varrho \max_{a'} Q(s', a'; \theta')$;

  Perform a gradient descent step on the $\text{Loss}(\theta) =$

$E[(\text{Target}Q - Q(s, a; \theta))^2]$ w.r.t. the network parameters θ;

    Every D' steps reset $\theta' = \theta$;

  end

 end

end

Obtain the optimal deployment position of UAVs $\mathcal{U}^*$ and calculate the system average ergodic rate $\bar{R}$.

IV. Simulation Results

In this section, we present simulation results to demonstrate the performance of the proposed scheme HC-DRL. We list the default simulation parameter settings in Table 1. The simulation environment setting is presented as follows except otherwise stated. We take Muli county of Sichuan Province as an example, where the probability of public network coverage of forest is only 30%. Therefore, we set β as 0.7.

Table 1. Summary of parameters

Parameters	Symbol	Value
The area size	area	1 km × 1 km
The number of users	N	100
The number of UAVs	K	4
Carrier frequency	f	1.4 GHz
The maximum transmission power of UAV	P_1	20 dBm
The maximum transmission power of BS	P_2	22 dBm
The minimum height of users	$h_{\min}$	0 m
The maximum height of users	$h_{\max}$	10 m
Reference distance of FSPL	d_0	1 m
The height of UAVs	h	100 m
The SNR threshold	γ_0	0 dB
The outage probability threshold	P_{thre}	0.1
The environment parameters	A, C, E, G, H	0.25, 0.39, 0.25, 0, 0.05
The pathloss exponent	α	3.5
The noise power	ψ	−140 dBm/Hz
The number of Episode	$\text{iter}_{\max}$	1000
The number of hidden layer	num_{cell}	2
The learning rate	ϱ	0.01
The maximum number of UAV per cluster	$Q_{\max}$	3

The simulation results consist of three parts: foliage channel model verification, comparison with other schemes, and influence of parameter setting. Next, the three parts are introduced respectively.

1. Foliage channel model verification

In this subsection, we describe several foliage channel models. The SPM model is originated in the cost and Hata prediction models [21], but it adopts the variable coefficients rather than the fixed coefficients, which provides more flexibility and applicability. The standard propagation model (SPM) channel model was well corrected based on measured data in Shuangdao National Forest Park, Weihai City, Shandong Province. We will compare SPM with the following two channels model: 1) our proposed foliage slant model (FSM); 2) foliage horizontal maximum attenuation model (FHM).

- FSM: As stated before, this model can be referred to (1).

- FHM: Compared with our proposed model, this model use the horizontal foliage model rather than slant one, which is proposed by ITU-R P.833-9 and given by

$$L_{l,i,j}^u (\text{dB}) = L_{l,i,j}^{\text{FSPL}} (d_0) + 10\alpha \lg \left(\frac{d_{l,i,j}}{d_0} \right) + X_\sigma + L_{l,i,j}^{\text{Horizontal}} \tag{33}$$

$L_{l,i,j}^{\text{Horizontal}}$ is provided as follows

$$L_{l,i,j}^{\text{Horizontal}} = A_m \left[1 - \exp\left(-d_{l,i,j} \mu / A_m \right) \right] \tag{34}$$

where μ denotes the specific attenuation for very short vegetative paths (unit: dB/m) and A_m denotes the maximum attenuation for one terminal within a specific type and depth of vegetation (unit: dB).

The comparison of different models for total loss is illustrated in Fig.4. As is shown, the gaps between SPM with correction and FSM, FHM are 0.6 dB, 8.5 dB, respectively. Obviously, our proposed FSM is closer to the SPM with correction and fits the measured data points better than FHM. The reason is that there is a certain height difference between the UAV and users, making the slant model more suitable than the horizontal one. Therefore, our proposed FSM has more adaptability to the forest environment.
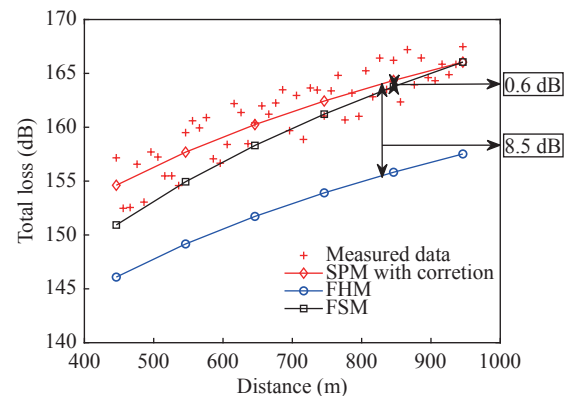

Fig. 4. Comparison of different models.

2. Comparison with other schemes

First, we compare the system average ergodic rate solved by our proposed UB-ergodic rate with the system average ergodic rate solved by the ergodicl rate. Then, the other two schemes as following are presented respectively for the comparison with the proposed HC-
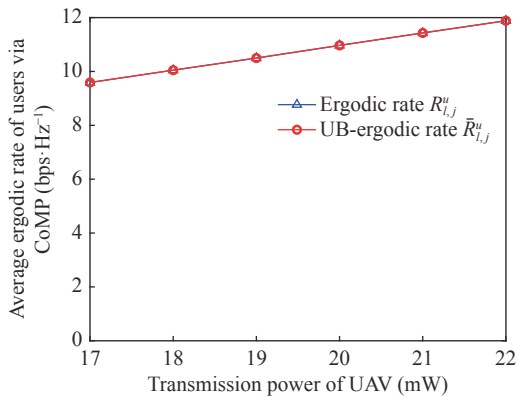
DRL:

- Centralized DRL (C-DRL): All the associated UAVs are regarded as an agent and trained jointly. The action of the agent is set as the directions of the movement of all UAVs.
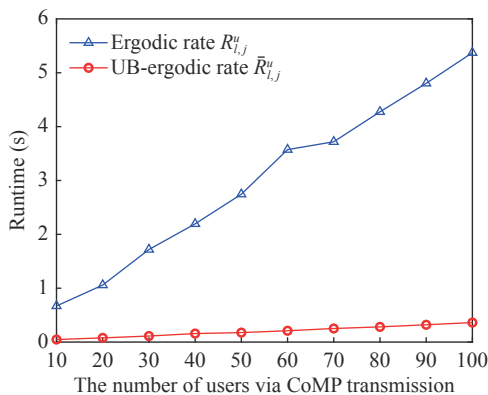
- Parallel single DRL (PS-DRL): Every UAV is regarded as a single agent and trains independently. The action of each UAV will not be affected by others.

In Fig.5(a), we compare the results solved by the proposed UB-ergodic rate with the results solved by the ergodic rate. In this scenario, every user is served by CoMP of 2 UAVs. As can be seen from Fig.5(a), the system average ergodic rate solved by UB-ergodic rate $\bar{R}_{l,j}^u$ is the same as the system average ergodic rate solved by the ergodic rate $R_{l,j}^u$. Since the number of UAVs in the scenario is relatively small, the gap between UB-ergodic rate $\bar{R}_{l,j}^u$ and the ergodic rate $R_{l,j}^u$ does not affect the deployment positions of UAVs. As shown in Fig.5(b), we compare the computation time between $\bar{R}_{l,j}^u$ and $R_{l,j}^u$. Compared with $R_{l,j}^u$, $\bar{R}_{l,j}^u$ has a shorter computation time and performs better. Therefore, we solve the system average ergodic rate by relaxing $R_{l,j}^u$ in (5) and considering $\bar{R}_{l,j}^u$ in (18) instead.

Fig.6 illustrates the user distribution. 100 users are distributed with 3-D Poisson point process (PPP) in the range of a 1000 m × 1000 m square area. Their height was generated randomly between 0 m and 10 m given the mountains' terrain relief. Considering the hills and valleys in the forest environment, it is more practical for the three-dimensional positions of users.
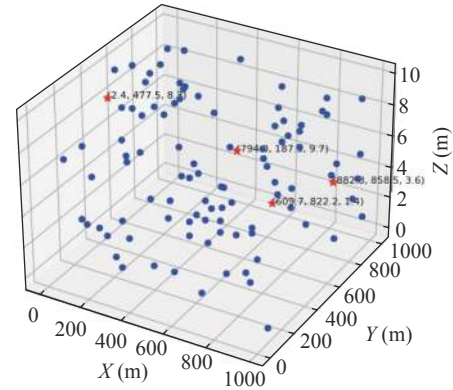

Fig. 6. User distribution with 3-D PPP.

Fig.7 compares the performance of our proposed HC-DRL with C-DRL and PS-DRL in terms of system average ergodic rate. The result demonstrates that our proposed HC-DRL fairly balances system performance and convergence time. Compared with C-DRL converging after about 600 episodes, our proposed HC-DRL greatly accelerates the convergence speed with only 200 episodes roughly. However, the performance of the proposed HC-DRL is just slightly inferior to the C-DRL. Additionally, our proposed HC-DRL gains much higher throughput than the PS-DRL scheme, which converges after about 100 episodes. This is because that the actions of UAVs in the same cluster are jointly optimized in HC-DRL. It indicates that our proposed scheme is both time-saving and effective by jointly considering the system performance and the convergence.
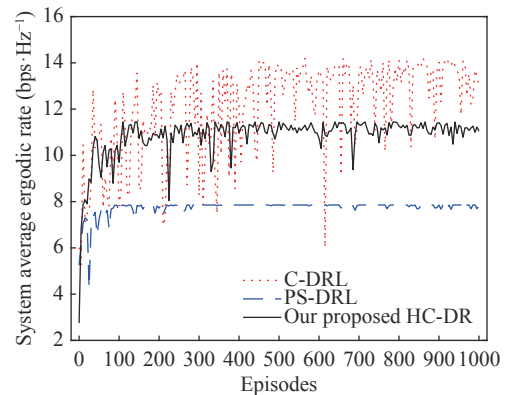

Fig. 7. System average ergodic rate comparison of different schemes.

3. Influence of parameter setting

Fig.8 shows the coverage ratio of users for UAV de-

(a) Ergodic rate comparison

(b) Runtime comparison

Fig. 5. Performance comparison between ergodic rate and proposed UB-ergodic rate in CoMP mode.

ployment schemes with different transmission power. It is obvious that the coverage capacity is limited with only a single BS. It is feasible to increase the transmission power of UAVs and the number of UAVs to serve more users. Hence, based on the practical situation of forest rescue, the number of UAVs can be chosen accordingly to enhance the communication capacity.
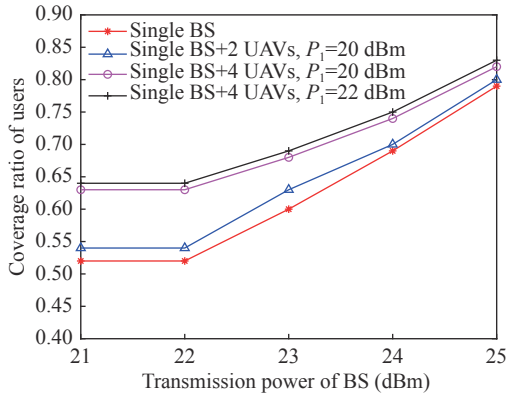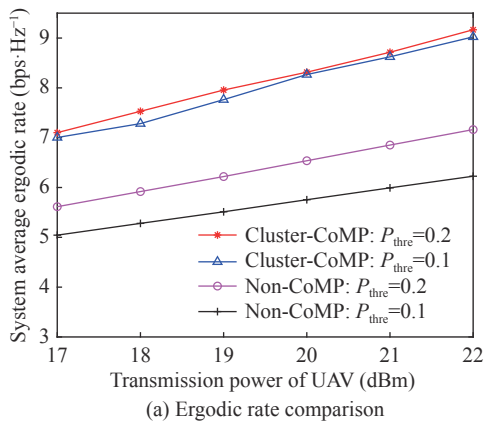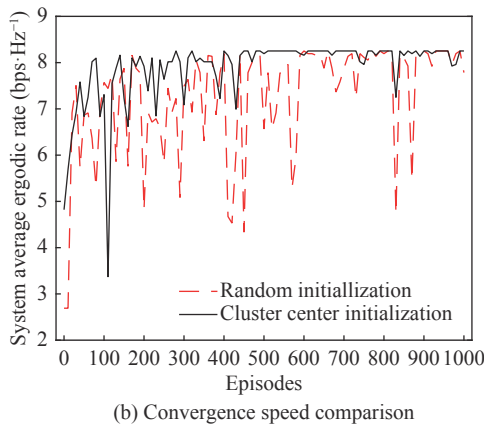

Fig. 8. Performance comparison for UAV deployment schemes with different transmission power.

Fig.9(a) shows the system average ergodic rate with different UAV transmission power and different
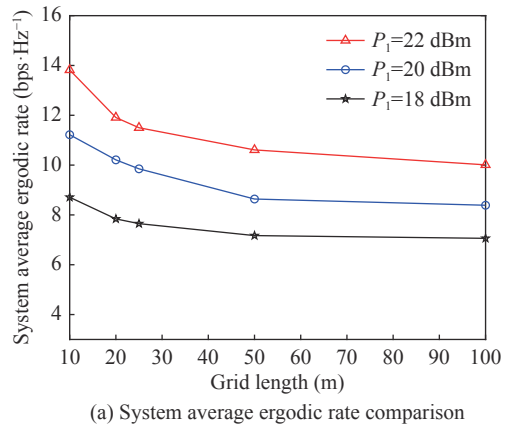

(a) Ergodic rate comparison
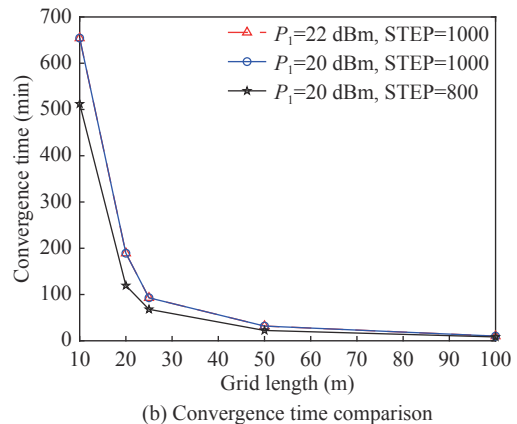

(b) Convergence speed comparison

Fig. 9. Performance comparison of ergodic rate and convergence speed for UAVs with different initial positions and different transmission power.

outage probability threshold. When the SNR threshold parameter is set to 10 dB, it shows that higher outage probability threshold can obtain better performance. The higher outage probability threshold leads to the lower probability of user transmission outage, therefore more users can transmit and improve system performance. The performance of CoMP is better than the method without CoMP (abbreviated as Non-CoMP). This is because CoMP transmission combines multiple UAVs jointly, which increases the user data rate and reduces the outage probability. Fig.9 shows the scheme for UAV position initialized with the center of the user cluster converges faster than random initialization. Compared with random initialization converging after about 500 episodes, cluster center initialization greatly accelerates the convergence speed with only 200 episodes roughly. When the number of iterations is small, the learning process may fall into the local optimum and cannot come out, so a better initial position is more likely to get an accurate result.

Fig.10 shows the effect of the grid size on the system throughput and the convergence time. The configuration of the computer used in the simulation is as follows: the CPU uses Intel (R) Core (TM) i5-9300H, and the graphics card uses NVIDIA GeForce GTX 1066ti.

(a) System average ergodic rate comparison

(b) Convergence time comparison

Fig. 10. Performance comparison with different grid length.

The STEP in Fig.10(b) indicates the number of times each episode of the program runs. As shown, the agent will find a better deployment scheme and need a longer convergence time with fine-grained grid division. On the contrary, the agent converges faster, but the effect is worse with coarse-grained grid division. This is due to that the grid division is directly related to the state space. The UAVs have more deployable positions and it is more likely to find the best solution. However, the algorithm complexity will increase as enlarging the state space. Therefore, there exists a trade-off between complexity and system performance. In practical scenarios, the grid division scheme requires careful design.

V. Conclusions

In this paper, we investigated UAV deployment as serving stations for emergency communication coverage of mountainous forest areas. Based on forest coverage fading, we proposed a CoMP-based clustering and deployment framework by considering both the broken BS and the available BS cases. Then, we proposed a hierarchical clustering algorithm to improve the network performance. To reduce complexity controlled by the size of action and state spaces, we developed an HC-DRL scheme and derived a bound relaxation on ergodic rate. Our simulation results demonstrated that the proposed scheme can effectively achieve a better performance while reducing computation complexity.

References

[1] K. Gomez, S. Kandeepan, M. M. Vidal, *et al.*, "Aerial base stations with opportunistic links for next generation emergency communications," *IEEE Communications Magazine*, vol.54, no.4, pp.31–39, 2016.

[2] C. C. Lai, C. T. Chen, and L. C. Wang, "On-demand density-aware UAV base station 3D placement for arbitrarily distributed users with guaranteed data rates," *IEEE Wireless Communications Letters*, vol.8, no.3, pp.913–916, 2019.

[3] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Communications Letters*, vol.3, no.6, pp.569–572, 2014.

[4] R. I. Bor-Yaliniz, A. El-Keyi, and H. Yanikomeroglu, "Efficient 3-D placement of an aerial base station in next generation cellular networks," in *Proceedings of the 2016 IEEE International Conference on Communications (ICC)*, Kuala Lumpur, Malaysia, pp.1–5, 2016.

[5] M. Alzenad, A. El-Keyi, F. Lagum, *et al.*, "3-D placement of an unmanned aerial vehicle base station (UAV-BS) for energy-efficient maximal coverage," *IEEE Wireless Communications Letters*, vol.6, no.4, pp.434–437, 2017.

[6] M. Z. Chen, M. Mozaffari, W. Saad, *et al.*, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE Journal on Selected Areas in Communications*, vol.35, no.5, pp.1046–1061, 2017.

[7] L. Liu, S. W. Zhang, and R. Zhang, "CoMP in the sky: UAV placement and movement optimization for multi-user communications," *IEEE Transactions on Communications*, vol.67, no.8, pp.5645–5658, 2019.

[8] L. P. Zhu, J. Zhang, Z. Y. Xiao, *et al.*, "Millimeter-wave full-duplex UAV relay: Joint positioning, beamforming, and power control," *IEEE Journal on Selected Areas in Communications*, vol.38, no.9, pp.2057–2073, 2020.

[9] Z. Y. Xiao, L. P. Zhu, and X. G. Xia, "UAV communications with millimeter-wave beamforming: Potentials, scenarios, and challenges," *China Communications*, vol.17, no.9, pp.147–166, 2020.

[10] Q. Wei, L. Wang, L. M. Xu, *et al.*, "Hierarchical coded caching for multiscale content sharing in heterogeneous vehicular networks," *IEEE Transactions on Vehicular Technology*, vol.71, no.6, pp.5770–5786, 2022.

[11] J. Hejselbæk, J. Ø. Nielsen, W. Fan, *et al.*, "Empirical study of near ground propagation in forest terrain for internet-of-things type device-to-device communication," *IEEE Access*, vol.6, pp.54052–54063, 2018.

[12] M. A. Weissberger, "An initial critical summary of models for predicting the attenuation of radio waves by trees," *Technical Report*, ESD-TR-81-101, Electromagnetic Compatibility Analysis Center, Annapolis, MD, USA, 1982.

[13] G. A. J. Van Dooren, H. J. F. G. Govaerts, and M. H. A. J. Herben, *COST 235: Radiowave Propagation Effects on Next-Generation Fixed-Services Terrestrial Telecommunications Systems*, Technische Universiteit Eindhoven, 1997.

[14] ITU-R P.833-9 (09/2016): 2016, Attenuation in Vegetation, P Series: Radio propagation, Available at: *https://www.itu.int/rec/R-REC-P.833-9-201609-S/en*.

[15] F. Teschl, F. P. Fontan, M. Schonhuber, *et al.*, "Attenuation of spruce, pine, and deciduous woodland at C-band," *IEEE Antennas and Wireless Propagation Letters*, vol.11, pp.109–112, 2012.

[16] X. Z. Lu, L. Xiao, and C. H. Dai, UAV-aided 5G communications with deep reinforcement learning against jamming, *arXiv preprint*, arXiv: 1805.06628, 2018.

[17] Y. W. Liu, Z. J. Qin, Y. L. Cai, *et al.*, "UAV communications based on non-orthogonal multiple access," *IEEE Wireless Communications*, vol.26, no.1, pp.52–57, 2019.

[18] C. H. Liu, Z. Y. Chen, J. Tang, *et al.*, "Energy-efficient UAV control for effective and fair communication coverage: A deep reinforcement learning approach," *IEEE Journal on Selected Areas in Communications*, vol.36, no.9, pp.2059–2070, 2018.

[19] J. B. Seo, S. Y. Kim, and V. C. M. Leung, "Outage probability characterization of CoMP-joint transmission with pathloss and Rayleigh fading," *IEEE Communications Letters*, vol.19, no.1, pp.78–81, 2015.

[20] Y. Z. Jiang, F. L. Chung, S. T. Wang, *et al.*, "Collaborative fuzzy clustering from multiple weighted views," *IEEE Transactions on Cybernetics*, vol.45, no.4, pp.688–701, 2015.

[21] P. Bayan, S. Ghazi-Maghrebi, and A. Rahmanian, "A new radio wave propagation model optimization approach for Tehran city based on standard model tuning," in *Proceedings of the 2018 International Symposium on Networks, Computers and Communications (ISNCC)*, Rome, Italy, pp.1–5, 2018.

WANG Li received the Ph.D. degree from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2009. She is currently a Full Professor with the School of Computer Science (National Pilot Software Engineering School), BUPT, where she is also an Associate Dean and heads the High Performance Computing and Networking Laboratory. She is also a member of the Key Laboratory of the Universal Wireless Communications, Ministry of Education, China. She also held Visiting Positions with the School of Electrical and Computer Engineering, Georgia Tech, Atlanta, GA, USA, from December 2013 to January 2015, and with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden, from August to November 2015 and July to August 2018. She has authored/coauthored almost 50 journal papers and two books. Her current research interests include wireless communications, distributed networking and storage, vehicular communications, social networks, and edge AI. She currently serves on the Editorial Boards for *IEEE Transactions on Vehicular Technology*, *IEEE Transactions on Green Communications and Networking*, *Computer Networks*, *IEEE Access*, and *China Communications*. She was the Symposium Chair of IEEE ICC 2019 on Cognitive Radio and Networks Symposium and a Tutorial Chair of IEEE VTC 2019-fall. She was the vice chair of Meetings and Conference Committee (MCC) for IEEE Communication Society (ComSoc) Asia Pacific Board (APB) for the term of 2021, and chairs the special interest group (SIG) on Sensing, Communications, Caching, and Computing (C3) in Cognitive Networks for IEEE Technical Committee on Cognitive Networks. She was the recipient of the 2013 Beijing Young Elite Faculty for Higher Education Award, Best Paper Awards from several IEEE conferences, e.g., IEEE ICCC 2017, IEEE GLOBECOM 2018, IEEE WCSP 2019, and so forth. She was also the recipient of the Beijing Technology Rising Star Award in 2018. She has served on TPC of multiple IEEE conferences, including IEEE Infocom, Globecom, International Conference on Communications, IEEE Wireless Communications and Networking Conference, and IEEE Vehicular Technology Conference in recent years.
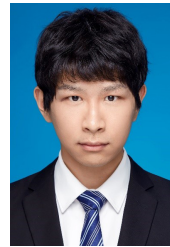(Email: liwang@bupt.edu.cn)

WU Xuewei received the bachelor's degree from Beijing University of Posts and Telecommunications, Beijing, China, in 2018, where she is currently pursuing the M.S. degree in electronics and communication engineering. Her research interests include UAV communications, machine learning-based wireless communications, and cooperative cache.
(Email: xuewei.wu@bupt.edu.cn)

WANG Yanhui received the B.S. degree from Beijing University of Posts and Telecommunications Beijing, China, in 2019, where he is currently pursuing the M.S. degree in electronics and communication engineering. His research interests include distributed caching, deep reinforcement learning, and incentive mechanism design in wireless communications. (Email: 1911128152@qq.com)

XIAO Zhe received the M.S. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2019, where he is currently pursuing the Ph.D. degree in computer science and technology. His research interests include UAV-deployment, deep reinforcement learning, and wireless communications.
(Email: xiaozhe@bupt.edu.cn)

LI Liang received the Ph.D. degree in the School of Telecommunications Engineering at Xidian University, China, in 2021. She is currently a postdoctoral faculty member with the School of Computer Science (National Pilot Software Engineering School), Beijing University of Posts and Telecommunications. She was also a visiting Ph.D. student with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA, from 2018 to 2020. Her research interests include edge computing, federated learning, data-driven robust optimization, and differential privacy.
(Email: liliang1127@bupt.edu.cn)

FEI Aiguo received the M.S. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 1981, and the Ph.D. degree from the University of Science and Technology Beijing, Beijing, in 2004. He is a Professor with the School of Computer Science (National Pilot Software Engineering School), BUPT. He is also a Member of the State Key Laboratory of Networking and Switching Technology and the Academician of the Chinese Academy of Engineering, Beijing. His current research interests include Internet of things, intelligent emergency communication systems, intelligent information systems, big data, cloud computing, and intelligent software development and testing.
(Email: aiguofei@bupt.edu.cn)