

Title	Lossy Cooperative UAV Relaying Networks: Outage Probability Analysis and Location Optimization
Author(s)	Lian, Ya; Lin, Wensheng; Li, Lixin; Yang, Fucheng; Han, Zhu; Matsumoto, Tad
Citation	2024 IEEE Globecom Workshops (GC Wkshps): 1-6
Issue Date	2024-12-08
Type	Conference Paper
Text version	author
URL	http://hdl.handle.net/10119/19997
Rights	This is the author's version of the work. Copyright (C) 2024 IEEE. 2024 IEEE Globecom Workshops (GC Wkshps), Cape Town, South Africa, pp. 1-6. DOI: https://doi.org/10.1109/GCWkshp64532.2024.11101027. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Description	2024 IEEE Globecom Workshops (GC Wkshps), Cape Town, South Africa, December 8–12, 2024



Lossy Cooperative UAV Relaying Networks: Outage Probability Analysis and Location Optimization

Ya Lian*, Wensheng Lin*, Lixin Li*, Fucheng Yang[†], Zhu Han[‡], and Tad Matsumoto[§]

*School of Electronics and Information, Northwestern Polytechnical University, Xi'an, Shaanxi 710129, China

[†]Research institute of Information Fusion, Naval Aviation University, Yantai, China

[‡]University of Houston, Houston, USA

[§]IMT-Atlantic, France, and University of Oulu, Finland, Emeritus.

Email: lianya@mail.nwpu.edu.cn, {linwest, lilixin}@nwpu.edu.cn, fucheng85@sina.com, zhan2@uh.edu, tadeshi.matsumoto@oulu.fi

Abstract—In this paper, performance of a lossy cooperative unmanned aerial vehicle (UAV) relay communication system is analyzed. In this system, the UAV relay adopts lossy forward (LF) strategy and the receiver has certain distortion requirements for the received information. For the system described above, we first derive the achievable rate distortion region of the system. Then, on the basis of the region analysis, the system outage probability when the channel suffers Nakagami- m fading is analyzed. Finally, we design an optimal relay position identification algorithm based on the Soft Actor-Critic (SAC) algorithm, which determines the optimal UAV position to minimize the outage probability. The simulation results show that the proposed algorithm can optimize the UAV position and reduce the system outage probability effectively.

Index Terms—Outage probability, unmanned aerial vehicle, lossy-forward, relay, reinforcement learning.

I. INTRODUCTION

In recent years, rapid development of communication technologies such as unmanned aerial vehicle (UAV) has been widespread in academia and industry [1]–[3]. Different from traditional ground and fixed communication base stations (BS), UAV communication introduces mobility, which brings new challenges to the assurance of coverage and connectivity.

Cooperative relay [4], as a communication strategy, aims to improve the reliability of data transmission and expand the communication coverage through the cooperative operation among the network nodes including relays. UAV as mobile relay has attracted significant research interest. Chen *et al.* [1] characterize the outage probability of the reconfigurable intelligent surface (RIS)-equipped-UAV system under a novel modified-Fisher-Snedecor $\mathcal{F}$ fading channel model. Tu *et al.* [2] derive a closed-form expression of outage probability for the wireless transmission from a BS to a mobile user via a

UAV relay over Rician fading channels. Dilip Kumar *et al.* [3] derive closed-form expressions for the outage probability and throughput of a UAV-assisted full-duplex wireless system with decode-and-forward (DF) protocol over Rician fading channels.

When UAV participates in the communication, the performance of the communication system is likely affected by the position of UAV, so the location of UAV should be dynamically changed to satisfy different communication's Quality-of-Service (QoS) requirements, such as age-of-information (AoI) [5]–[7]. Presently, a lot of related research results have been published on UAV location deployment and trajectory design. Li *et al.* [8] analyze the system outage probability of the multi-UAV cascaded relay communication network using orthogonal frequency division multiplexing (OFDM) and identify the optimal UAV. In [6], account is taken of the AoI as optimization objectives in UAV-enabled edge Internet-of-Things (IoT) scenarios. To solve the optimization problem, deep reinforcement learning (DRL) is shown to be a powerful tool. Moreover, the impact of the UAV location on the average AoI and the peak AoI is investigated in [9] and [10], respectively.

The sequence may contain intra-link errors; however, the sequence is correlated to the source sequence. Hence, joint decoding at the destination may recover the original source sequence. This strategy is called Lossy-Forward (LF) technique [11], which also inspires the innovation of semantic-forward [12]. In terms of outage probability, LF can usually provide better performance than conventional relay technologies [13]. Due to the emergence of semantic communication [14], [15], lossy communication becomes promising in the future wireless networks. Most of the literatures analyze the lossy communication performance only based on the fading model that match the fixed (on-ground) wireless communication propagation scenario [13], [16]. Thus, it is worth conducting further research with the aim of its application to UAV aided communications.

The contributions of this paper are summarized as follows:

- In this paper, by using Shannon's lossy sources-channel separation theorem, we derive the achievable rate distortion region of UAV as lossy cooperative relaying network with a specific Binary distortion requirement.
- Then, based on the derived region, we study the outage

Corresponding authors: Wensheng Lin, Lixin Li.

This work was supported in part by National Natural Science Foundation of China under Grants 62001387 and 62101450, in part by the Young Elite Scientists Sponsorship Program by the China Association for Science and Technology under Grant 2022QNRC001, in part by Aeronautical Science Foundation of China under Grants 2022Z021053001 and 2023Z071053007, in part by Shanghai Academy of Spaceflight Technology under Grant SAST2022-052, in part by French National Research Agency Future Program under reference ANR-10-LABX-07-01, in part by NSF CNS-2107216, CNS-2128368, CMMI-2222810, ECCS-2302469, US Department of Transportation, Toyota and Amazon. Tad Matsumoto's contribution to this paper is mainly on outage analysis part.

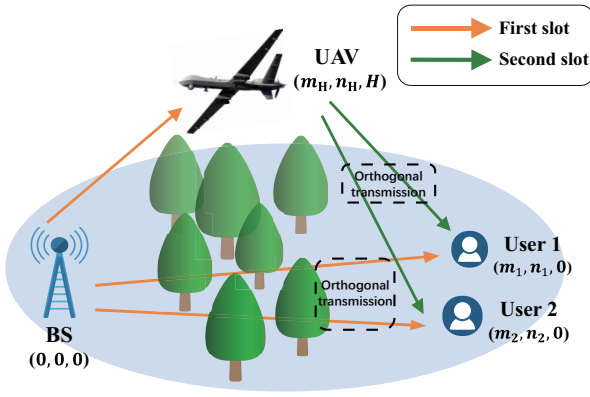


Fig. 1: The lossy cooperative UAV relay communication system.

probability of the network under Nakagami- m fading, and provide the outage derivation and the final result.

- In addition, we design an optimal relay position identification algorithm based on the SAC algorithm to reduce the outage probability of the system while for identify the optimal relay position.

The rest of this paper is organized as follows. Section II introduces the system model of the UAV lossy cooperative ying network. Section III presents the outage probability analysis in detail. Section IV proposes the SAC-based algorithm to minimize the outage probability of the system. Simulation results and numerical results based on the analytical results are presented in Section V. Finally, Section VI draws conclusions.

II. PROBLEM STATEMENT

In Section II-A, we first introduce the communication model of the lossy cooperative UAV relay communication system this paper assumes in detail. Then, Section II-B introduces the channel model adopted in the following sections which provide theoretical results and numerical calculations.

A. System Model

We consider a communication model of the lossy cooperative UAV relay communication system consisting of a BS, a UAV, and K users, as shown in Fig. 1. In this paper we study the case where $K = 2$, i.e., two users are included, however the scenario can be extended to the case of more than two users. The users adopt orthogonal multiple access transmission (OMA) mode, e.g., frequency division multiplexing access (FDMA). Assuming that the poor quality of the communication link between the BS and the user leads to frequent failure. To reduce the outage probability of the communication system, the UAV is located as a relay in the communication network. In lossy communication, when the BS sends the original information sequence, the user, as the receiver, has certain Binary distortion requirements as the QoS on the received information sequence. Communication interruption occurs if the distortion level of the received information does not satisfy the requirement. Since the information on sequences received by the user from the BS and from the UAV relay are correlated, the user can reduce the

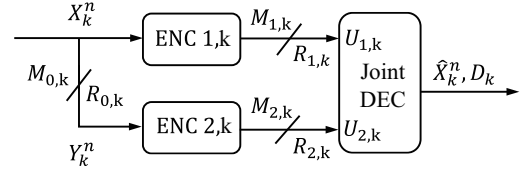


Fig. 2: The multiterminal source coding problem for outage probability analysis of the k^{th} user.

distortion of the received information through joint decoding after receiving these signals to achieve an acceptable degree.

In this paper, we establish the relationship between instantaneous SNR and distortion level based on Shannon's lossy source-channel separation theorem to determine the corresponding achievable link rate, given the distortion requirements.

We formulate the relationship between the distortion and each involved link's supported rate in the framework of multi-terminal source coding. As shown in Fig. 2, taking the k^{th} user as an example, X_k^n is equivalent to the original information sequence sent by the BS to the k^{th} user, where n represents the sequence length. Assuming that the information sequence is compressed into Y_k^n at the bit rate $R_{0,k}$, and then the sequences X_k^n and Y_k^n are compressed into two code words $M_{1,k}$ and $M_{2,k}$ at the bit rates $R_{1,k}$ and $R_{2,k}$. When transmitted to the joint decoder, the lossy versions $U_{1,k}$ and $U_{2,k}$, of X_k^n and Y_k^n , respectively, are obtained. The final reconstructed sequence of information may contain Binary distortion D_k . If $D_k \neq 0$, the reconstructed sequence is a lossy version $\hat{X}_k^n$ of X_k^n .

B. Channel Model

Consider the transmission of the k^{th} user as an example, when the transmit power P_t is fixed, the receive power with the k^{th} user $P_{r,k}$ can be expressed as

$$P_{r,k} = \frac{P_t h_k}{PL_{i,k}}, \quad (1)$$

where h_k represents the instantaneous normalized channel gain, $PL_{i,k}$ represents path loss, where $i \in \{0,1,2\}$ denoting the S-R link, the S-D link and the R-D link, respectively. The channel is expressed by the A2G channel model, with $i \in \{0,2\}$, and the ground channel model with when $i = 1$. According to the literature [17], the formula for the $PL_{i,k}$ calculation is given as below:

$$PL_{i,k} = \frac{\eta_{LoS} - \eta_{NLoS}}{1 + \mathcal{A} \exp(-\mathcal{B}(\theta_k - \mathcal{A}))} + 20 \log_{10} \left(\frac{4\pi f_k d_{i,k}}{c} \right) + \eta_{NLoS}, \quad (2)$$

where $\mathcal{A}$, $\mathcal{B}$, η_{LoS} and η_{NLoS} are the environmental parameters. $\theta_k = \arcsin \left(\frac{H}{d_{i,k}} \right)$ is the elevation of the UAV, with H being the altitude of the UAV, and $d_{i,k}$ the length of the links. f_k is the carrier frequency, and c is the speed of light.

Then the instantaneous signal-to-noise ratio (SNR) of each link can be calculated by

$$\gamma_{i,k} = P_{r,k}/N_0 = P_t h_k / (N_0 \cdot PL_{i,k}), \quad (3)$$

where N_0 is the noise power spectral density of the additive Gaussian noise of the receiver. We normalize the average fading channel gain to the unity, and use Nakagami- m fading channel to describe the fading channel variation at the UAV and the users [18]. It is easy to know that the probability density function of instantaneous SNR is

$$p(\gamma_{i,k}) = \frac{m^m \gamma_{i,k}^{m-1}}{\bar{\gamma}_{i,k}^m \Gamma(m)} \exp\left(-\frac{m\gamma_{i,k}}{\bar{\gamma}_{i,k}}\right), \quad (4)$$

where the exact value of the factor m can be determined through field measurement campaign in real channels. For simplicity, in this paper, we set it as a parameter. $\bar{\gamma}_{i,k}$ is the average SNR and $\Gamma(\cdot)$ is the gamma function.

III. OUTAGE PROBABILITY ANALYSIS

We derive the relationship between instantaneous SNR and final distortion in two steps, which are, first we identify the achievable rate-distortion region, and then calculate analytically the outage probability.

A. Achievable Rate Region

We assume a Binary source. A schematic diagram of the problem in the multi-terminal source coding framework corresponding to the system we investigate is shown in Fig. 2.

Let's continue with the k^{th} user. The information sequences X_k , Y_k , $U_{1,k}$ and $U_{2,k}$ form a Markov Chain $U_{1,k} \rightarrow X_k \rightarrow Y_k \rightarrow U_{2,k}$. For the S-R link, the link rate $R_{0,k}$ should not be less than the information about X_k obtained from the information sequence Y_k in the UAV relay, which is given by the mutual information $I(X_k; Y_k)$. Notice the same rule should apply to the R-D link, the rate $R_{2,k}$ cannot be smaller than the mutual information $I(Y_k; U_{2,k})$. For the S-D link, $U_{2,k}$ provides helper information, by utilizing the compressed side information $U_{2,k}$ in joint decoding, the rate $R_{1,k}$ only has to be greater than or equal to the mutual information $I(X_k; U_{1,k} | U_{2,k})$.

The requirements for the link rates can be simplified as:

$$R_{0,k} \geq 1 - H_b(\rho_{1,k}) \quad (5)$$

$$R_{1,k} \geq H_b(\rho_{1,k} * \rho_{2,k} * D_k) - H_b(D_k) \quad (6)$$

$$R_{2,k} \geq 1 - H_b(\rho_{2,k}) \quad (7)$$

where $\rho_{1,k}$ and $\rho_{2,k}$ represent the crossover probabilities with the S-R link and the R-D link, respectively. D_k denotes the crossover probability with the S-D link, and $H_b(\cdot)$ denotes the Binary entropy function.

B. Outage Probability Derivation

According to Shannon's lossy source-channel separation theorem, we can establish the relationship between the instantaneous channel SNR $\gamma_{i,k}$ and its corresponding rate constraint $R_{i,k}$ for $i \in \{0, 1, 2\}$, as:

$$R_{i,k} = C(\gamma_{i,k}) / \mathcal{R}_{i,k} = B \log_2(1 + \gamma_{i,k}) / \mathcal{R}_{i,k}, \quad (8)$$

where $C(\gamma_{i,k})$ is the Shannon capacity using the Gaussian codebook, and $\mathcal{R}_{i,k}$ represents the end-to-end rates of joint source channel coding.

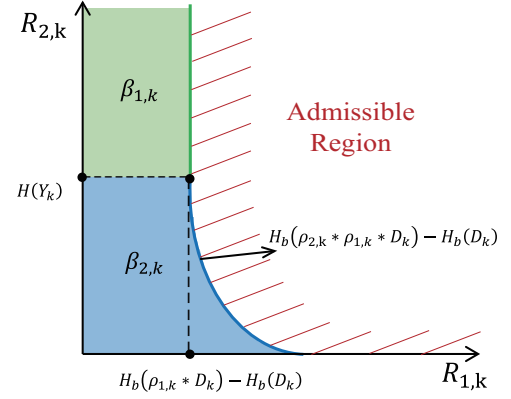


Fig. 3: The admissible rate region of the k^{th} user; the red solid line indicates the achievable rate region with acceptable distortion D_k .

Substituting Eq. (8) into (7), we have

$$\rho_{2,k} = H_b^{-1}\left(1 - \frac{B \log_2(1 + \gamma_{2,k})}{\mathcal{R}_{2,k}}\right), \quad (9)$$

For the achievable rate region of the k^{th} user obtained in Section III-A, we can further derive the outage probability of the k^{th} user. The outage probability of the k^{th} user can be expressed as

$$P_{out,k} = \Pr\{(R_{0,k}, R_{1,k}, R_{2,k}) \in \beta_k\}, \quad (10)$$

where β_k represents the admissible rate distortion region of the k^{th} user, as shown in Fig. 3.

The admissible rate distortion region described in Section III-A needs to be further decomposed for the ease of the calculation. We first focus on the S-D link and the R-D links. For the S-R link, the cross over probability $\rho_{1,k}$ between information X_k and Y_k is regarded as a parameter determined according to the rate-distortion function, given $R_{0,k}$. Through this method, a two-dimensional achievable rate distortion region can be obtained, as shown in Fig. 3. Moreover, the rate region is divided into two parts corresponding to $\beta_{1,k}$ and $\beta_{2,k}$ to facilitate integral calculation. The outage probability of the k^{th} user can then be expressed as

$$P_{out,k} = \Pr\{\rho_{1,k} \in [0, 0.5], (R_{1,k}, R_{2,k}) \in \beta_{1,k} \cup \beta_{2,k}\}, \quad (11)$$

which can further be decomposed into:

$$\begin{aligned} P_{out,k} = & \Pr\{\rho_{1,k} = 0, (R_{1,k}, R_{2,k}) \in \beta_{1,k}\} \\ & + \Pr\{\rho_{1,k} = 0, (R_{1,k}, R_{2,k}) \in \beta_{2,k}\} \\ & + \Pr\{\rho_{1,k} \in (0, 0.5], (R_{1,k}, R_{2,k}) \in \beta_{1,k}\} \\ & + \Pr\{\rho_{1,k} \in (0, 0.5], (R_{1,k}, R_{2,k}) \in \beta_{2,k}\}, \end{aligned} \quad (12)$$

After several steps of mathematical manipulations, we have:

$$\begin{aligned} P_{out,k} = & 1 + \frac{2\Gamma(m, \frac{m}{\bar{\gamma}_{0,k}})\Gamma(m, \frac{m}{\bar{\gamma}_{2,k}})}{[\Gamma(m)]^2} \\ & - \frac{\Gamma(m, \frac{m}{\bar{\gamma}_{0,k}})}{\Gamma(m)} \cdot \int_0^1 p(\gamma_{2,k}) \end{aligned}$$

$$\begin{aligned}
& \cdot \frac{\Gamma(m, \frac{m \cdot [2^{H_b(\varphi(\gamma_{2,k}) * D_k) - H_b(D_k)} - 1]}{\bar{\gamma}_{1,k}})}{\Gamma(m)} d\gamma_{2,k} \\
& - \frac{\Gamma(m, \frac{m}{\bar{\gamma}_{2,k}})}{\Gamma(m)} \cdot \int_0^1 p(\gamma_{0,k}) \\
& \cdot \frac{\Gamma(m, \frac{m \cdot [2^{H_b(\varphi(\gamma_{0,k}) * \varphi(\gamma_{2,k}) * D_k) - H_b(D_k)} - 1]}{\bar{\gamma}_{1,k}})}{\Gamma(m)} d\gamma_{0,k} \\
& - \int_0^1 d\gamma_{0,k} \int_0^1 p(\gamma_{0,k}) p(\gamma_{2,k}) \\
& \cdot \frac{\Gamma(m, \frac{m \cdot [2^{H_b(\varphi(\gamma_{0,k}) * \varphi(\gamma_{2,k}) * D_k) - H_b(D_k)} - 1]}{\bar{\gamma}_{1,k}})}{\Gamma(m)} d\gamma_{2,k}, \quad (13)
\end{aligned}$$

where $\varphi(\gamma_{i,k})$ represents the inverse function of the Binary entropy of $H_b^{-1}(1 - \log_2(1 + \gamma_{i,k}))$, $\Gamma(\cdot, \cdot)$ stands for the upper incomplete gamma function. It should be noted that the boundaries of the integrals have been properly replaced corresponding to Eq.(9) and no further integral calculation can be performed analytically.

IV. SAC-BASED OUTAGE PROBABILITY MINIMIZATION

In this section, we design the algorithm based on SAC method. Compared with deterministic strategies, SAC algorithm uses random strategies, which has advantages yielding good performance in continuous actions and state space.

A. Problem Framework

We model the considered outage probability optimization problem as a Markov Decision Process (MDP). At the time slot t , the agent obtains the state s_t by observing the environment and performs an action a_t . If the environmental state happens, of which the occurrence probability is $p(s_{t+1}|s_t, a_t)$, agent will receive a reward r_{t+1} according to reward function. The state, action, and reward functions of the MDP are defined as follows.

1) *States*: The state of the UAV at each time slot t contains the information about the horizontal location of the UAV, which is describe as $s_t = (n_{1,t}, n_{2,t})$.

We consider a constant UAV altitude and only solve the problem of UAV deployment in the two dimensions. Because in the actual scenario, the UAV will usually use the lowest altitude to obtain the best relay performance, optimizing the drone height becomes irrelevant.

2) *Actions*: The UAV can fly in any direction, and hence its space of action is continuous. Since the action of each time slot t is denoted by $a(t) = (\Delta n_{1,t}, \Delta n_{2,t})$. The action space range is limited, such scenario is equivalent to limiting the flight speed of the UAV.

3) *Rewards*: Since we want to minimize the outage probability of the communication system, we consider incorporating the outage probability into the reward function. At each time epoch t , the reward is calculated based on the reward function given by

$$r_{t+1} = \frac{K}{\mu \cdot \sum_{k=1}^K P_{out,k}(t)}, \quad (14)$$

where the constant μ needs to be tuned empirically.

B. Algorithm Description

SAC algorithm is an off-policy algorithm developed for maximum entropy RL [19]. It not only maximizes the long-term return, but also maximizes the entropy of each output action of the strategy, so as to achieve larger exploration, faster convergence, and higher stability. Its optimization objective can be expressed as

$$J(\pi) = \mathbb{E}_{(s_t, a_t) \sim \rho^\pi} \sum_t r(s_t, a_t) + \alpha \mathcal{H}(\pi(\cdot|s_t)), \quad (15)$$

where π represents the action function of the agent, the policy. ρ^π denotes the state-action marginal distribution induced by the policy π . $\mathcal{H}(\pi(\cdot|s_t))$ is the information entropy of $\pi(\cdot|s_t)$, and α is the temperature coefficient which determines the importance of entropy with respect to the reward, thus controlling the degree of randomness of the policy.

The temperature coefficient needs to be adjusted according to specific factors such as the optimization objective, which imposes some level of difficulties to the training. Thus, we use the method of automatically adjusting the temperature coefficient. At this time, the objective function becomes as follow:

$$\begin{aligned}
& \max_{\pi} \mathbb{E}_{\rho^\pi} \left[\sum_t r(s_t, a_t) \right] \\
& \text{s.t. } \mathbb{E}_{(s_t, a_t) \sim \rho^\pi} [-\log(\pi_t(a_t|s_t))] \geq \mathcal{H}_{\min}, \forall t, \quad (16)
\end{aligned}$$

Here, $\mathcal{H}_{\min}$ is entropy target, which is set as $\mathcal{H}_{\min} = -\dim(A)$, the opposite of the dimension of the action space.

TABLE I: System Parameters

Parameters	Value	Parameters	Value
N_0	3.9811×10^{-14} W	η_{LoS}	0.1
f_1	2 GHz	η_{NLoS}	21
f_2	1.985 GHz	P_t	0.5 W
$\mathcal{A}$	4.88	m	2
$\mathcal{B}$	0.43	μ	5000

TABLE II: DRL Hyperparameters

Hyperparameters	Value	Hyperparameters	Value
Number of episodes	300	Replay memory size	10000
Mini-batch size	128	Learning rate λ_α	0.0003
Learning rate λ_Q	0.003	Learning rate λ_π	0.001
Discount factor ξ	0.9	Activation function	ReLU

In the SAC-based algorithm, we use the actor network to generate action policy π , and the critic network to approximate the soft Q-function to evaluate the performance of the actor network, where the soft Q-function is the expected cumulative reward of starting from state s_t , taking action a_t , and following policy π .

To avoid duplication of letters, the discount factor in this part is denoted by ξ . The soft Q-function parameters θ and the policy parameter ϕ can be trained by minimizing the loss functions (17) and (18) as follow, respectively,

$$J_Q(\theta) = \mathbb{E}_{(s_t, a_t) \sim \mathcal{D}} \left[\frac{1}{2} (\mathbf{Q}_\theta(s_t, a_t) - r(s_t, a_t)) \right]$$

Algorithm 1 Optimal Relay Position Identification Algorithm

```

1: Initialization the position of UAV;
2: Initialization the experience replay buffer  $\mathcal{D}$ ;
3: for episode 1, 2, 3..., do
4:   Obtain the initial state  $s$  for UAV;
5:   Reset episode rewards;
6:   for step 1, 2, 3..., do
7:     choose the action  $a_t$  according to the current state  $s_t$ 
       the current policy  $\pi_t$ ;
8:     Get the next state  $s_{t+1}$  after performing action  $a_t$ ;
9:     Get reward  $r_{t+1}$ ;
10:    Store the quadruple  $(s_t, a_t, r_{t+1}, s_{t+1})$  to  $\mathcal{D}$ ;
11:    if Memory counter > Memory Capacity then
12:      Sample a batch of data in  $\mathcal{D}$  randomly;
13:      Update the critic network parameter  $\theta$  and the action
        network parameter  $\phi$  by minimizing Equation (17)
        and (18);
14:      Update the temperature coefficient  $\alpha$  by minimizing
        Equation (19);
15:      Soft update the parameter  $\bar{\theta}$  of the target Critic
        network;
16:    end if
17:  end for
18: end for

```

$$-\xi \mathbf{Q}_{\bar{\theta}}(s_{t+1}, a_{t+1}) + \alpha \log(\pi(a_{t+1}|s_{t+1}))^2 \Big], \quad (17)$$

$$J_{\pi}(\phi) = \mathbb{E}_{s_t \sim \mathcal{D}} [\mathbb{E}_{a_t \sim \pi} [\alpha \log(\pi(a_t|s_t)) - \mathbf{Q}_{\theta}(s_t, a_t)]], \quad (18)$$

The temperature parameter α is adjusted by minimizing the loss function, as follows

$$J(\alpha) = \mathbb{E}_{a_t \sim \pi_t} [-\alpha_t \log \pi_t(a_t|s_t) - \alpha \mathcal{H}_{min}]. \quad (19)$$

The detailed algorithm is described in Algorithm 1.

V. NUMERICAL RESULTS

We consider a lossy cooperative UAV relaying network consisting of one UAV and two users with their distortion requirements, and the whole scene is distributed in a 20×20 km² area. The flying altitude of the UAV is 500 m, and the initial position is set at being equal to the location of the BS, which is the origin. The locations of the two users are fixed and known. Other parameter Settings in the environment are shown in Table I unless otherwise specified. All neural networks in the algorithm are implemented based on the Pytorch framework, and the optimizers adopt Adam. In addition, we implement DDPG for comparison. The hyperparameters are shown in Table II.

Fig. 4 depicts the value of the reward function in 300 episodes when $D_1 = 0.1, D_2 = 0.3$. The slight instability observed in the SAC based algorithm during the initial phase can be attributed to the algorithm's necessity for extensive exploration to adequately learn the environment and optimize the strategy, which is a typical characteristic of RL algorithms. As the

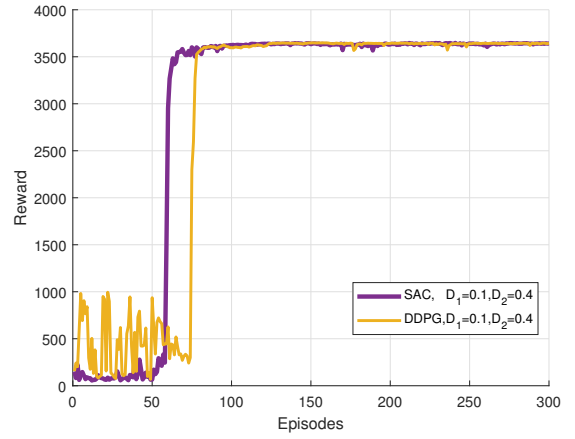


Fig. 4: The value of reward function over 300 iterations.

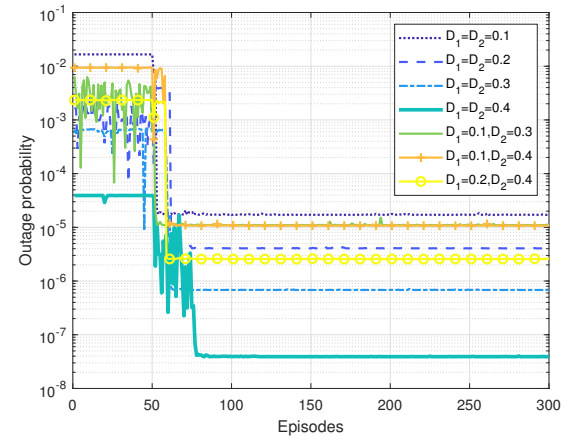
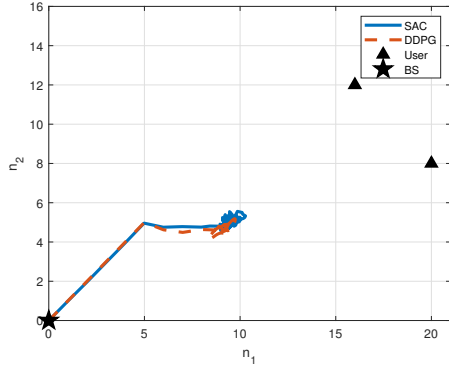


Fig. 5: The value of reward function over 300 iterations.

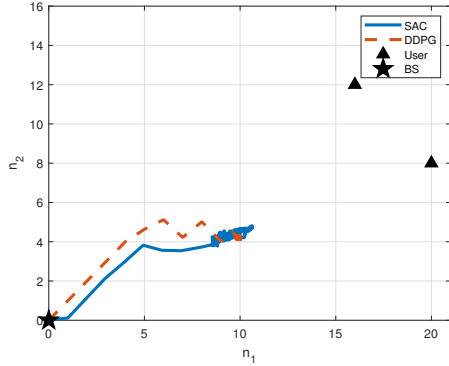
algorithm iterates and learns, the SAC based algorithm gradually converges, exhibiting a high degree of stability. Conversely, the DDPG experiences greater fluctuations and slower convergence.

As shown in Fig. 5, with the increase of the number of episodes, the outage probability gradually converges into the minimum, which also means that after continuous iterations, the UAV gradually converges to the position yielding the minimum system outage probability. At the same time, it is obvious that the lower the distortion requirement, the larger D value, and hence, the smaller the outage probability of the system. When the distortion requirements are different among users, the outage probability value is mainly limited by the largest distortion requirement.

The final trajectory of the UAV is described in Fig. 6. It can be seen that compared with DDPG, after the SAC based algorithm converges, the UAV can find a more reasonable trajectory to reach the position with the minimum outage probability. There will be oscillation in the final trajectory of the UAV. This is because when the UAV flies near the optimal position, the received signal power is already strong, and the influence of horizontal movement on the received power near the optimal position is



(a) $D_1 = D_2 = 0.2$.



(b) $D_1 = 0.1, D_2 = 0.3$.

Fig. 6: The obtained UAV paths of the two algorithms under different distortion conditions.

small, so the UAV oscillates near the optimal position. In future work, the trajectory end point can be stabilized with Kalman filtering, or by a weighted average of historical positions.

VI. CONCLUSION

In this paper, considering the distortion requirements of the users, we have analyzed the performance of the lossy cooperative UAV relay system, and optimize the UAV trajectory and position with the goal of minimizing the probability of interruption. It has been shown through simulations that the designed optimal relay position identification algorithm can satisfy the requirements of a variety of distortions and it has been shown that the optimal UAV position which minimizes the outage probability can be identified quickly and steadily.

ACKNOWLEDGEMENT

This paper in part includes intellectual results which the second and the last authors achieved when they were with Japan Advanced Institute of Science and Technology (JAIST) but not published yet, and their extensions. This paper is closely related to the results presented in [13] which has been published, given the permission by JAIST Research Management Section. Even though the last author's contributions are on behalf of his current affiliation, they are strictly consistent to the permission for [13]

by JAIST Research Management Section, as well as editorial matters for the readability improvement.

REFERENCES

- [1] Y. Chen, W. Cheng, and W. Zhang, "Reconfigurable intelligent surface equipped UAV in emergency wireless communications: A new fading-shadowing model and performance analysis," *IEEE Transactions on Communications*, vol. 72, no. 3, pp. 1821–1834, 2024.
- [2] H. Tu, J. Zhu, and Y. Zou, "Optimal power allocation for minimizing outage probability of UAV relay communications," in *International Conference on Wireless Communications and Signal Processing*, Xi'an, China, 2019, pp. 1–6.
- [3] K. Dilip Kumar, A. Jeganathan, M. Gupta, and P. Muthuchidambaramanathan, "Outage probability analysis of full-duplex UAV-assisted wireless system over rician fading channel," in *IEEE International Conference on Communication, Networks and Satellite*, Purwokerto, Indonesia, 2021, pp. 263–268.
- [4] J. He, V. Tervo, X. Zhou, X. He, S. Qian, M. Cheng, M. Juntti, and T. Matsumoto, "A tutorial on lossy forwarding cooperative relaying," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 66–87, Aug. 2019.
- [5] Y. Xiao and Q. Du, "Statistical age-of-information optimization for status update over multi-state fading channels," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 4, pp. 6018–6023, Apr. 2024.
- [6] Y. Liu and J. Gong, "DRL-based UAV trajectory planning for AoS and energy consumption minimization assisted by AoI," in *IEEE Global Communications Conference*, Kuala Lumpur, Malaysia, 2023, pp. 1866–1871.
- [7] Y. Xiao, Q. Du, W. Cheng, and W. Zhang, "Adaptive sampling and transmission for minimizing age of information in metaverse," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 3, pp. 588–602, Mar. 2024.
- [8] D. Li and Y. Meng, "Trajectory optimization methods of the OFDM-UAV cascade relay communication system," in *IEEE International Conference on Information, Communication and Networks*, Xi'an, China, 2023, pp. 294–300.
- [9] W. Lin, L. Li, J. Yuan, Z. Han, M. Juntti, and T. Matsumoto, "Cooperative lossy communications in unmanned aerial vehicle networks: Age-of-information with outage probability," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10 105–10 120, Oct. 2021.
- [10] W. Lin, L. Li, Y. Liu, Y. He, and Y. Liu, "Timeliness optimization of unmanned aerial vehicle lossy communications for Internet-of-Things," *Chinese Journal of Aeronautics*, vol. 36, no. 6, pp. 249–255, Jun. 2023.
- [11] X. Zhou, M. Cheng, X. He, and T. Matsumoto, "Exact and approximated outage probability analyses for decode-and-forward relaying system allowing intra-link errors," *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 7062–7071, 2014.
- [12] W. Lin, Y. Yan, L. Li, Z. Han, and T. Matsumoto, "Semantic-forward relaying: A novel framework toward 6G cooperative communications," *IEEE Communications Letters*, vol. 28, no. 3, pp. 518–522, Mar. 2024.
- [13] W. Lin, S. Qian, and T. Matsumoto, "Lossy-forward relaying for lossy communications: Rate-distortion and outage probability analyses," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 3974–3986, 2019.
- [14] Y. Fu, W. Cheng, W. Zhang, and J. Wang, "Scalable extraction based semantic communication for 6G wireless networks," *IEEE Communications Magazine*, vol. 62, no. 7, pp. 96–102, Jul. 2024.
- [15] W. Lin, Y. Yan, L. Li, Z. Han, and T. Matsumoto, "SemantIC: Semantic interference cancellation toward 6G wireless communications," *IEEE Communications Letters*, vol. 28, no. 8, pp. 1810–1814, Aug. 2024.
- [16] T. Matsumoto, A. Zribi, R. Asvadi, E. Dupraz, and W. Lin, "Two-stage successive Wyner-Ziv lossy forward relaying for lossy communications: Rate-distortion and outage probability analyses," *IEEE Transactions on Vehicular Technology*, 2024, Early Access.
- [17] 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.
- [18] A. A. Khuwaja, Y. Chen, N. Zhao, M.-S. Alouini, and P. Dobbins, "A survey of channel modeling for UAV communications," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2804–2821, 2018.
- [19] T. Haarnoja, A. Zhou, K. Hartikainen, G. Tucker, S. Ha, J. Tan, V. Kumar, H. Zhu, A. Gupta, P. Abbeel, and S. Levine, "Soft actor-critic algorithms and applications," *arXiv*, vol. abs/1812.05905, 2018.