This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2023.3285737

UAV Trajectory and Multi-User Beamforming Optimization for Clustered Users Against Passive Eavesdropping Attacks With Unknown CSI

Aly Sabri Abdalla, Ali Behfarnia, and Vuk Marojevic

***Abstract*—This paper tackles the fundamental passive eavesdropping problem in modern wireless communications in which the location and the channel state information (CSI) of the attackers are unknown. In this regard, we propose deploying an unmanned aerial vehicle (UAV) that serves as a mobile aerial relay (AR) to help ground base station (GBS) support a subset of vulnerable users. More precisely, our solution (1) clusters the single-antenna users in two groups to be either served by the GBS directly or via the AR, (2) employs optimal multi-user beamforming to the directly served users, and (3) optimizes the AR's 3D position, its multi-user beamforming matrix and transmit powers by combining closed-form solutions with machine learning techniques. Specifically, we design a plain beamforming and power optimization combined with a deep reinforcement learning (DRL) algorithm for an AR to optimize its trajectory for the security maximization of the served users. Numerical results show that the multi-user multiple input, single output (MU-MISO) system split between a GBS and an AR with optimized transmission parameters without knowledge of the eavesdropping channels achieves high secrecy capacities that scale well with increasing the number of users.**

Index Terms: UAV-assisted, beamforming, DRL, eavesdropping, MU-MISO, physical layer security, power control, trajectory optimization.

I. Introduction

Unmanned aerial vehicles (UAVs) are envisioned to improve the next generation of wireless communication systems, 6G and beyond, by providing flexible, intelligent, secure, and limitless connectivity [1]–[3]. Steps to identify the challenges and solutions of emerging cellular networks to serve UAVs are being undertaken by the 3rd Generation Partnership Project (3GPP) [4]. A prominent use case for a UAV is the aerial relay (AR) which supports extended coverage or higher system capacity at low-cost. However, the largely line of sight (LoS) air-to-ground (A2G) communications between UAVs and user equipment (UEs) make the system vulnerable to a variety of attacks [5]. Eavesdropping is a major passive attack that can compromise communications channels and gain access to private and sensitive user information.

Physical layer security has been introduced as a powerful tool to secure communication links by using the physical characteristics of wireless communication channels [6]–[8]. UAVs can benefit from physical layer security by applying the latest technologies such as artificial intelligent (AI) methods as well as various communication techniques to mitigate the compromise of malicious behavior in the network. However, applying such techniques is complicated by three factors: (i) requiring to coordinate between a ground base station (GBS) and the AR to determine which users should be served by which stations, (ii) handling resource limitations and trajectory of ARs to serve specific UE(s), and (iii) choosing the well-suited learning and communication techniques for the GBS and the AR to dynamically maximize the security metrics.

A. Related Work

The recent related works can be classified into three groups: i) beamforming-aided secure communications; ii) UAV-aided secure communications; iii) beamforming and UAV-aided secure communications. Table I provides a summery of the prior art and proposed research related to the work presented in this paper.

Beamforming-aided Secure Communications: Transmit beamforming limit the radio frequency (RF) propagation footprint and thus implicitly enable a secure the propagation channel without high computational requirements a the receiver as compared to cryptographic security schemes. Carefully designing the beam patterns of antennas at the transmitter, receiver, or both can enhancing system performance and security parameter, such as signal-to-interference plus noise ratio (SINR) and secrecy rate, respectively. Researchers have studied how to leverage and optimize beamforming for improving the PLS of current and future wireless communication networks. The work presented in [9] investigates the achieved secrecy sum rate for a multi-cell multiple-input multiple-output (MIMO) system which is under a passive eavesdropper attack. The power allocation between artificial noise (AN) and information signal is managed to maximize the sum secrecy rate with imperfect channel state information (CSI), which is derived using regularized channel inversion (RCI) precoding. Reference [10] proposes strategies of combining AN and beamforming to achieve high secrecy performance for massive MIMO systems in spite of single-antenna active eavesdropping attacks that

Manuscript received 8 September 2022; revised 3 November 2022; accepted 4 June 2023. Date of publication - - 2023; date of current version 10 June 2023. This work was supported in part by the NSF PAWR program, under grant number CNS-1939334. The work of A. Behfarnia is supported in part by the Faculty Research Program under the Office of Sponsored Programs and Research at the University of Tennessee at Martin.
Aly Sabri Abdalla and Vuk Marojevic are with the Department of Electrical and Computer Engineering, Mississippi State University, MS, USA e-mail: (asa298@msstate.edu; vm602@msstate.edu).
Ali Behfarnia is with the Department of Engineering, University of Tennessee at Martin, TN, USA e-mail: (a.behfarnia@tennessee.edu).

Table I: Prior Art and Proposed Research.

Category	Ref.	Objective Metric	Attack type	Strategy	Attackers' CSI
Beamforming	[9]	Secrecy sum rate	Passive eavesdroppers	RCI is adopted to drive the power allocation between AN and information signal.	Perfect CSI with channel errors
	[10]	Secrecy rate	Active eavesdropper	Analytical framework to find the best combination of AN and beamforming.	Perfect CSI
	[11]	Secrecy rate	Active and passive eavesdroppers	Analytical framework to design the beamforming.	Statistical CSI
UAV	[12]	Average secrecy rate	Passive eavesdropper	Optimizing the UAV's trajectory and AN allocation via iterative algorithm.	Perfect CSI
	[13]	Secrecy rate	Passive eavesdropper	Jointly optimizing the source/ UAV relay transmit power and the UAV trajectory through an iterative algorithm.	Perfect CSI
	[14]	Secrecy rate	Passive eavesdropper	The UAV's trajectory and transmit power allocation are jointly optimized by applying DQL algorithm.	Unknown CSI
Beamforming and UAV	[15]	Minimum secrecy rate	Passive eavesdropper	Jointly optimizing the UAV beamforming and position to enhance UE's secrecy rate through applying multi-objective dragonfly algorithm.	Perfect CSI
	[16]	Secrecy capacity	Passive eavesdroppers	A DRL is proposed to optimize the UAV trajectory and transmitter and jammer UAVs beamforming.	Perfect CSI
	[17]	Secrecy rate	Passive eavesdropper	Jointly optimize the beamforming of multi-beam satellite and the power allocation of UAV through an iterative alternating optimization approach.	Perfect CSI
This work		Secrecy sum capacity	Passive eavesdroppers	User clustering for association with the GBS and AR, where a DQL is designed to optimize the UAV trajectory, beamforming, and power control without knowledge of the wiretap CSI.	Unknown CSI

attempt to spoil the channel estimation acquisition at the BS. Reference [11] derives the multiple-input, single-output (MISO) beamforming design for random wireless networks with statistical CSI in an environment with eavesdroppers and interferers.

UAV-aided Secure Communications: Reference [12] demonstrates the applicability of maximizing the achievable average secrecy rate by optimizing the AN transmission and UAV trajectory. Reference [13] proposes using the UAV as a relay to improve the secrecy rate by jointly optimizing the source/relay transmit power and the UAV trajectory. In our previous work [14], we have proposed a deep Q-learning (DQL) algorithm to optimize the secrecy rate by optimizing the trajectory of the UAV relay and the transmit power without the availability of the CSI of the wiretap channel.

Beamforming Plus UAV-Aided Secure Communications: Combining both beamforming and UAV has been considered as an enhanced PLS technique in advanced wireless communications. For example, [15] introduces the multi-objective dragonfly algorithm (MODA) to solve the multi-objective optimization problem for enhancing the minimum secrecy rate between the UAV node and a single UE for different clusters. The work presented in that paper assumes perfect CSI conditionsfor the BS and focuses on optimizing the UAV performance. Reference [16] proposes a multi-agent deep reinforcement learning (DRL) algorithm to maximize the secrecy capacity of a multi user system by optimizing the trajectory of the aerial BS and the beamforming matrix of the jammer UAV interfering with the eavesdroppers. The authors of [17] propose an iterative optimization approach that alternately optimizes the beamforming of satellite transmitters and the power allocation of the UAV acting as an aerial relay and friendly jammer supporting multi-beam satellite-enabled vehicle communication in the presence of eavesdropping.

B. Contribution

In this paper, we aim to mitigate passive eavesdropping attacks, where an eavesdropper illegitimately wiretaps the legitimate wireless communications links. To this end, we propose a combination of machine learning, deep reinforcement learning, and multi-antennas techniques at the BS and the AR to maximize the security of UEs in a wireless communication network. The contributions of this paper are:

- We define a practical optimization problem to maximize the channel secrecy capacity without CSI knowledge of the wiretap channel.
- We introduce a framework for effectively solving this problem by means of user clustering, beamforming and power control, and AR trajectory optimization. We design a DRL solution for the trajectory optimization and leverage the closed form solutions for the beamforming and transmit and relay power allocation.
- We provide a comprehensive numerical analysis that demonstrates the effectiveness of the proposed tools.

The rest of paper is organized as follows. Section II presents the system model. Section III formulates the problem and defines the relevant metrics. Section IV derives the solution. Numerical results and analyses are presented in Section V. Section VI provides the concluding remarks.

II. System Model

We consider a ground base station (GBS) serving ground UEs where the communication links are subject to passive eavesdropping attacks. The eavesdroppers have a radio receiver and can wiretap the downlink transmission. A UAV acting as an AR is dispatched to support secure communications. This scenario is illustrated in Fig. 1.

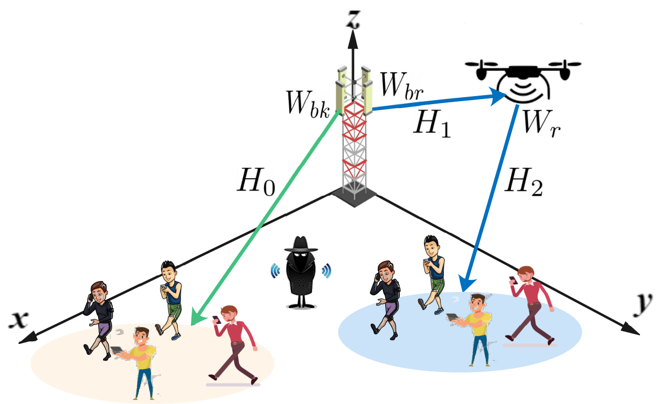


Figure 1: System model.

We use the following notation: lower-case letters represent scalars and bold lower-case letters denote vectors. Bold upper-case letters are used for matrices. Tr($\mathbf{S}$) and $\mathbf{S}^{-1}$ represent the trace and the inverse of a square matrix $\mathbf{S}$, respectively. The operator $(.)^T$ denotes transpose, and the operator $(.)^\dagger$ denotes conjugate transpose. $\mathbf{S}(i,j)$ shows the (i,j)th element of matrix $\mathbf{S}$ and Rank($\mathbf{S}$) shows the rank of the matrix. $||\mathbf{v}||$ represents the Euclidean norm of a complex vector $\mathbf{v}$. Also, $|v|$ denotes the norm of a complex number v. $\mathbb{C}^{a\times b}$ denotes the dimension of $a\times b$ for a complex vector or matrix. Complex normal distribution vector with the mean vector $\mathbf{m}$ and the covariance matrix $\mathbf{\Sigma}$ is denoted by $\mathcal{CN}(\mathbf{m},\mathbf{\Sigma})$, and $\sim$ implies "distributed as".

A. Channel model

1) Air-to-ground: In terms of modelling the A2G communication channel between the UAV and ground receivers, we consider small-scale Rician fading where the line of sight (LoS) component coexist with non-LoS (NLoS) components [18]. The GBS and AR have both a uniform linear array (ULA) of M and N antennas, respectively. The A2G channel model,

$$\boldsymbol{G}_{TR} = \frac{\sqrt{\lambda_0}}{d_{TR}^{\alpha}}\left(\sqrt{\frac{\beta}{1+\beta}}\;\boldsymbol{G_{TR}^{LoS}} + \sqrt{\frac{1}{\beta+1}}\;\boldsymbol{G_{TR}^{NLoS}}\right), \quad (1)$$

is obtained as the superposition of the LoS and NLoS channel components, where λ_0 is the path loss at the reference distance of $1\ m$, d_{TR} is the 3D distance between the GBS and AR, α is the path loss exponent, and β is the Rician factor. Without loss of generality, the entries of $\boldsymbol{G_{TR}^{NLoS}}$ are assumed to be independent and identically distributed (i.i.d.) zero-mean and unit variance circularly symmetric complex Gaussian (CSCG), i.e., $\sim\mathcal{CN}(0,1)$. The LoS component,

$$\boldsymbol{G}_{TR}^{LoS} = \boldsymbol{g_{TR}^{(A)}}\;\boldsymbol{g_{TR}^{(D)}}, \quad (2)$$

where

$$\boldsymbol{g_{TR}^{(A)}} = \left[1, e^{-j\frac{2\pi}{\lambda}\Upsilon\Lambda^{TR}}, \cdots, e^{-j\frac{2\pi}{\lambda}(N-1)\Upsilon\Lambda^{TR}}\right] \quad (3)$$

and

$$\boldsymbol{g_{TR}^{(D)}} = \left[1, e^{-j\frac{2\pi}{\lambda}\Upsilon\Gamma^{TR}}, \cdots, e^{-j\frac{2\pi}{\lambda}(M-1)\Upsilon\Gamma^{TR}}\right] \quad (4)$$

correspond to channel contributions from the angel-of-arrival (AoA) and angel-of-departure (AoD) between the GBS and the AR. Parameter λ is the carrier wavelength, Υ is the antenna separation, $\Lambda^{TR} = cos\ \Theta\ sin\ \varphi$ is the AoA component (Θ–azimuth and φ–elevation AoA), and $\Gamma^{TR} = sin\ \vartheta\ cos\ \psi$ is the AoD component (ϑ–elevation and ψ–azimuth AoD) of the transmitted signal from the GBS to the AR.

The A2G channel between the AR and the ground users,

$$\boldsymbol{G}_{RK} = \frac{\sqrt{\lambda_0}}{\boldsymbol{d_{RK}^{\alpha}}}\left(\sqrt{\frac{\beta}{1+\beta}}\;\boldsymbol{g_{RK}^{LoS}} + \sqrt{\frac{1}{\beta+1}}\;\boldsymbol{G_{RK}^{NLoS}}\right), \quad (5)$$

has an LoS and an NLoS term, where $\boldsymbol{d_{RK}}$ is the 3D distance between the AR and the ground user cluster. The $\boldsymbol{G_{RK}^{NLoS}}$ entries follow the same CSCG distribution as $\boldsymbol{G_{TR}^{NLoS}}$. The LoS term,

$$\boldsymbol{g_{RK}^{LoS}} = \left[1, e^{-j\frac{2\pi}{\lambda}\Upsilon\chi^{RK}}, \cdots, e^{-j\frac{2\pi}{\lambda}(N-1)\Upsilon\chi^{RK}}\right], \quad (6)$$

defines the AoD components $\chi^{RK} = cos\ \Phi\ sin\ \Omega$ (Φ–azimuth and Ω–elevation AoD) of the transmitted signal from the ULA of the AR to the single-antenna users.

2) Ground-to-ground: the Alpha-beta-gamma (ABG) [19] channel model is adopted for the ground-to-ground (G2G) communication channels between the GBS and the eavesdropper and between the UEs and the eavesdropper. It is the closest path-loss model approximation to the actual 5G ground communications measurement results and it is employed by standard organizations such as ITU-R, 3GPP, mmMAGIC, and QuaDRiGa [20]. It is defined as

$$\begin{aligned} h_{G2G}(f,d) = {} & 10\ \rho_G \times log\left(\frac{d_{gg}}{1m}\right) + \jmath \\ & + 10\ \gamma_G \times log\left(\frac{f_c}{1GHz}\right) + \chi_{\sigma}^{G2G}, \end{aligned} \quad (7)$$

where d_{gg} is the 2D distance between the transmitter and receiver nodes, $\jmath$ is the intercept, and ρ_G and γ_G correspond to the distance and the frequency-dependent exponents. Shadow fading, χ_{σ}^{G2G}, is modeled as a Gaussian random variable of zero-mean and standard deviation σ_{sh}.

B. Communication Model for Legitimate Users

The M-antenna GBS can communicate with the K single-antenna UEs either directly or using the N-antenna AR at the same frequency, employing space-division multiple access (SDMA) and time-division multiple access (TDMA) [21]. The GBS serves K_b users directly and K_r users via the AR, where $K = K_b + K_r$. In what follows, we provide the corresponding communication models and channel capacities.

1) Direct communication from GBS: For the direct communication, the GBS forms multiple simultaneous beams to spatially separated users employing SDMA. The transmit beamforming assigns one beam vector for each user. However, transmit power leakage can occur between beams causing multi-user interference.

We consider the downlink transmission, where the GBS transfers K_b data streams to K_b users. The transmitted signal model is

$$\boldsymbol{x}_b = \sum_{k=1}^{K_b} \boldsymbol{w}_{b,k}\, s_k, \tag{8}$$

where $\boldsymbol{x}_b \in \mathbb{C}^{M\times 1}$, $\boldsymbol{w}_{b,k} \in \mathbb{C}^{M\times 1}$ is the beamforming vector and s_k the transmitted information symbol for the kth user. The beamforming, or precoding, matrix of the GBS contains K_b beamforming vectors, $\boldsymbol{W}_{bk} \in \mathbb{C}^{M\times K_b}$, where $\boldsymbol{W}_{bk} = [\boldsymbol{w}_{b,1}, \cdots, \boldsymbol{w}_{b,K_b}]$. The allocated transmit power for the kth user can then be calculated by the squared norm of the beamforming vector $\| \boldsymbol{w}_{b,k} \|^2$. The received signal at the K_b users can is expressed as

$$\boldsymbol{y}_0 = \boldsymbol{H}_0\, \boldsymbol{x}_b + \boldsymbol{n}_0, \tag{9}$$

where $\boldsymbol{y}_0 \in \mathbb{C}^{K_b\times 1}$, $\boldsymbol{H}_0 \in \mathbb{C}^{K_b\times M}$ represents the channel between the M antennas of the GBS and the K_b single-antenna users, and $\boldsymbol{n}_0 \in \mathbb{C}^{K_b\times 1}$ represents noise. It is assumed that the distribution of noise at each user is complex normal with zero-mean and unit variance, i.e., $n_k \sim \mathcal{CN}(0,1)$. The received signal at user k,

$$\begin{aligned} y_{0,k} &= \boldsymbol{h}_{0,k}\, \boldsymbol{x}_b + n_k, \\ &= \boldsymbol{h}_{0,k}\Big(\sum_{k=1}^{K_b} \boldsymbol{w}_{b,k}\; s_k\Big) + n_k, \\ &= \boldsymbol{h}_{0,k}\boldsymbol{w}_{b,k}\, s_k + \boldsymbol{h}_{0,k}\Big(\sum_{\substack{i=1\\ i\neq k}}^{K_b} \boldsymbol{w}_{b,i}\, s_i\Big) + n_k \end{aligned} \tag{10}$$

has the signal-to-interference-plus-noise-ratio (SINR)

$$\gamma_{b,k} = \frac{|\, \boldsymbol{h}_{0,k}\boldsymbol{w}_{b,k} \,|^2}{\sum_{i\neq k} |\, \boldsymbol{h}_{0,k}\boldsymbol{w}_{b,i} \,|^2 + 1}, \tag{11}$$

where $\boldsymbol{h}_{0,k} \in \mathbb{C}^{1\times M}$ denotes the MISO channel from the GBS to the kth user. The channel capacity of the direct link is obtained from

$$C_{b,k} = \log_2\big(1 + \gamma_{b,k}\big). \tag{12}$$

2) Indirect communication via AR: We assume a time-slot based synchronization between the GBS transmission and the AR transmission [22], [23]. In odd time-slots (phase), the BS transmits K_r data streams to the AR, each of which is destined to one UE. The transmission between the GBS and AR can be modeled as a standard point-to-to-point MIMO channel. The received signal at the AR can be written as

$$\begin{aligned} \boldsymbol{y}_1 &= \boldsymbol{H}_1\boldsymbol{x}_b + \boldsymbol{n}_1, \\ &= \boldsymbol{H}_1\Big(\sum_{k=1}^{K_r} \boldsymbol{w}_{b,k}\; s_k\Big) + \boldsymbol{n}_1, \end{aligned} \tag{13}$$

where $\boldsymbol{y}_1 \in \mathbb{C}^{N\times 1}$, $\boldsymbol{H}_1 \in \mathbb{C}^{N\times M}$ is the MIMO communication channel between the BS and the AR, and $\boldsymbol{n}_1 \sim \mathcal{CN}(0, \boldsymbol{I}) \in \mathbb{C}^{N\times 1}$ is the noise vector.

In the even time-slots, the AR transmits

$$\boldsymbol{x}_r = \boldsymbol{W}_r\; \boldsymbol{y}_1, \tag{14}$$

where $\boldsymbol{x}_r \in \mathbf{C}^{N\times 1}$ and $\boldsymbol{W}_r \in \mathbb{C}^{N\times N}$ is the beamforming matrix. The received signals at the K_r UEs are modeled as

$$\begin{aligned} \boldsymbol{y}_2 &= \boldsymbol{H}_2\; \boldsymbol{x}_r + \boldsymbol{n}_2, \\ &= \boldsymbol{H}_2\left(\boldsymbol{W}_r\left(\boldsymbol{H}_1\Big(\sum_{k=1}^{K_r} \boldsymbol{w}_{b,k}\; s_k\Big) + \boldsymbol{n}_1\right)\right) + \boldsymbol{n}_2, \end{aligned} \tag{15}$$

where $\boldsymbol{y}_2 \in \mathbb{C}^{K_r\times 1}$, $\boldsymbol{H}_2 \in \mathbb{C}^{K_r\times N}$ is the A2G communication channel between the AR and the K_r UEs, and $\boldsymbol{n}_2 \sim \mathcal{CN}(0,\boldsymbol{I}) \in \mathbb{C}^{K_r\times 1}$ is the noise vector. The kth user receives

$$\begin{aligned} y_{2,k} &= \boldsymbol{h}_{2,k}\boldsymbol{W}_r\; \boldsymbol{H}_1\; \boldsymbol{w}_{b,k}\; s_k \\ &\quad + \boldsymbol{h}_{2,k}\boldsymbol{W}_r\; \boldsymbol{H}_1\Big(\sum_{i\neq k}^{K_r} \boldsymbol{w}_{b,i}\; s_i\Big) \\ &\quad + \boldsymbol{h}_{2,k}\boldsymbol{W}_r\boldsymbol{n}_1 + n_{2,k}, \end{aligned} \tag{16}$$

where $\boldsymbol{h}_{2,k} \in \mathbb{C}^{1\times N}$ denotes the MISO channel from the AR to the kth UE and $n_{2,k} \sim \mathcal{CN}(0,1)$ is the additive noise. The SINR of this relayed communication link from the BS to the kth UE via the AR can then be calculated as

$$\gamma_{r,k} = \frac{|\, \boldsymbol{h}_{2,k}\boldsymbol{W}_r\boldsymbol{H}_1\boldsymbol{w}_{b,k} \,|^2}{\sum_{i\neq k} |\, \boldsymbol{h}_{2,k}\boldsymbol{W}_r\boldsymbol{H}_1\boldsymbol{w}_{b,i} \,|^2 + \|\boldsymbol{h}_{2,k}\boldsymbol{W}_r\|^2 + 1}. \tag{17}$$

The channel capacity $C_{r,k}$ of the indirect link is obtained from (12) using $\gamma_{r,k}$ instead of $\gamma_{b,k}$. Note that $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are directly influenced by UAV mobility due to changes in distance, altitude, and orientation relative to ground receivers.

C. Communication Model for Eavesdroppers

1) Eavesdropping on the direct communication link: The eavesdropper listens on the the direct link between the GBS and the associated UEs and receives

$$\begin{aligned} y_{0,e} &= \boldsymbol{h}_{0,e}\; \boldsymbol{x}_b + n_e, \\ &= \boldsymbol{h}_{0,e}\; \Big(\sum_{k=1}^{K} \boldsymbol{w}_{b,k}\, s_k\Big) + n_e, \\ &= \boldsymbol{h}_{0,e}\boldsymbol{w}_{b,k}\, s_k + \boldsymbol{h}_{0,e}\Big(\sum_{\substack{i=1\\ i\neq k}}^{K} \boldsymbol{w}_{b,i}\, s_i\Big) + n_e, \end{aligned} \tag{18}$$

where $\boldsymbol{h}_{0,e} \in \mathbb{C}^{1\times M}$ is the G2G communication channel between the BS and the eavesdropper and n_e is the noise at the eavesdropper such that $n_e \sim \mathcal{CN}(0,1)$. The SINR associated with the direct link between the GBS and the eavesdropper—for the beam formed to user k—can be calculated as

$$\gamma_{b,e,k} = \frac{|\, \boldsymbol{h}_{0,e}\boldsymbol{w}_{b,k} \,|^2}{\sum_{i\neq k} |\, \boldsymbol{h}_{0,e}\boldsymbol{w}_{b,i} \,|^2 + 1}. \tag{19}$$

Consequently, the capacity of the eavesdropper associated with the direct link from the BS to the k^{th} user can be derived as

$$\begin{aligned} C_{b,e,k} &= \log_2\big(1 + \gamma_{b,e,k}\big), \\ &= \log_2\left(1 + \frac{|\, \boldsymbol{h}_{0,e}\boldsymbol{w}_{b,k} \,|^2}{\sum_{i\neq k} |\, \boldsymbol{h}_{0,e}\boldsymbol{w}_{b,i} \,|^2 + 1}\right). \end{aligned} \tag{20}$$

2) Eavesdropping from relay communication link: The eavesdropper can wiretap the A2G relay communication link between the UAV and the UEs. Similar to the section II-B2, the capacity of the eavesdropper associated with the relay link can be derived as

$$\begin{aligned} &C_{r,e} = \log_2(1 + \gamma_{r,e}) \\ &= \log_2\left(1 + \frac{|\, \boldsymbol{h}_{2,e}\boldsymbol{W}_r\boldsymbol{H}_1\boldsymbol{w}_{b,k} \,|^2}{\sum_{i\neq k} |\, \boldsymbol{h}_{2,e}\boldsymbol{W}_r\boldsymbol{H}_1\boldsymbol{w}_{b,i} \,|^2 + \|\boldsymbol{h}_{2,e}\boldsymbol{W}_r\|^2 + 1}\right), \end{aligned} \tag{21}$$

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2023.3285737

where $\gamma_{r,e}$ is the SINR, and $\boldsymbol{h}_{2,e} \in \mathbb{C}^{1\times N}$ denotes the A2G channel between the UAV and the eavesdropper, and $n_{2,e}$ is the noise at the eavesdropper such that $n_{2,e} \sim \mathcal{CN}(0,1)$.

D. Secrecy Capacity

The term *secrecy capacity* is a measure of the information rate that can be transmitted securely without being intercepted. It is obtained as the difference between the achievable data rate of a legitimate receiver and the achievable data rate of an eavesdropper, taking into account the channel conditions and the employed security measures. It corresponds to the rate at which no data will be decoded by the eavesdropper [24].

For the system model of Section II, the average sum-secrecy capacity of the K_b UEs that are directly served by the GBS over T time slots is

$$\begin{aligned} C_{sec,b} &= \frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K_b}\Big(C_{b,k}-C_{b,e}\Big)^{+} \\ &= \frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K_b}\Bigg[\log_2\Bigg(1+\frac{\mid \boldsymbol{h}_{0,k}\boldsymbol{w}_{b,k}\mid^2}{\sum\limits_{i\neq k}\mid \boldsymbol{h}_{0,k}\boldsymbol{w}_{b,i}\mid^2+1}\Bigg)- \\ &\qquad \log_2\Bigg(1+\frac{\mid \boldsymbol{h}_{0,e}\boldsymbol{w}_{b,k}\mid^2}{\sum\limits_{i\neq k}\mid \boldsymbol{h}_{0,e}\boldsymbol{w}_{b,i}\mid^2+1}\Bigg)\Bigg]^{+}. \end{aligned} \tag{22}$$

Likewise, the average sum-secrecy capacity of the K_r UEs served via the AR over T time slots is obtained as

$$\begin{aligned} C_{sec,r} = \frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K_r}\Big(C_{r,k}-C_{r,e}\Big)^{+} = \frac{1}{T}\sum_{t=1}^{T} \\ \sum_{k=1}^{K_r}\Bigg[\log_2\Bigg(1+\frac{\mid \boldsymbol{h}_{2,k}\boldsymbol{W}_r\boldsymbol{H}_1\boldsymbol{w}_{b,k}\mid^2}{\sum\limits_{i\neq k}\mid \boldsymbol{h}_{2,k}\boldsymbol{W}_r\boldsymbol{H}_1\boldsymbol{w}_{b,i}\mid^2+\|\boldsymbol{h}_{2,k}\boldsymbol{W}_r\|^2+1}\Bigg) \\ -\log_2\Bigg(1+\frac{\mid \boldsymbol{h}_{2,e}\boldsymbol{W}_r\boldsymbol{H}_1\boldsymbol{w}_{b,k}\mid^2}{\sum\limits_{i\neq k}\mid \boldsymbol{h}_{2,e}\boldsymbol{W}_r\boldsymbol{H}_1\boldsymbol{w}_{b,i}\mid^2+\|\boldsymbol{h}_{2,e}\boldsymbol{W}_r\|^2+1}\Bigg)\Bigg]^{+}. \end{aligned} \tag{23}$$

where $[\omega]^+ \triangleq max(\omega,0)$. The total secrecy capacity is

$$C_T = C_{sec,b} + C_{sec,r}. \tag{24}$$

Formulas (22)-(24) are derived from information theory and provide quantitative measures of the level of secrecy achieved in the communications channels according to the system model of Fig. 1. The secrecy capacity is maximized when maximizing the SINRs at the legitimate receivers and minimizing the SINRs at the eavesdroppers.

III. Problem Formulation

This paper aims to maximize the total secrecy capacity of the UEs whether they are directly served by the GBS or through the AR. Considering the degrees of freedom for serving UEs directly or via the AR, we formulate two optimization problems.

1) Direct communication: For the directly served UEs, the optimization problem is defined as

$$\begin{aligned} &\max_{\boldsymbol{w}_{b,k}} && C_{sec,b} \\ &\text{subject to (s.t.)} && P_b \le P_{b,max}, \end{aligned} \tag{25}$$

where $C_{sec,b}$ is the secrecy capacity defined in (22), $P_{b,max}$ is the maximum transmit power of the GBS, and P_b is the transmit power of the GBS, that is, $P_b = \text{Tr}\left(\boldsymbol{x}_b\boldsymbol{x}_b^{\dagger}\right)$.

Problem (25) requires the knowledge of the eavesdropping channel. We assume the location of the eavesdropper and thus its CSI to be unknown, which is the scenario of interest in practice where it is difficult to detect or estimate the presence, location, or channel of eavesdroppers because of their passive nature. Therefore, we can only consider the capacity of the legitimate user and reformulate the optimization problem:

$$\begin{aligned} &\max_{\boldsymbol{w}_{b,k}} && \frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K}\Bigg[\log_2\Bigg(1+\frac{\mid \boldsymbol{h}_{0,k}\boldsymbol{w}_{b,k}\mid^2}{\sum\limits_{i\neq k}\mid \boldsymbol{h}_{0,k}\boldsymbol{w}_{b,i}\mid^2+1}\Bigg)\Bigg] \\ &\text{s.t.} && \text{Tr}\big(\boldsymbol{x}_b\boldsymbol{x}_b^{\dagger}\big) \le P_{b,max}. \end{aligned} \tag{26}$$

The eavesdropper location and channel are used only for calculating the resulting secrecy capacity for performance evaluation.

2) Relay communication: For the UEs that are served via the AR, the optimization problem is defined as

$$\begin{aligned} &\max_{\{\boldsymbol{w}_{b,k},x_r,y_r,z_r\}} && C_{sec,r} \\ &\text{s.t.} && P_r \le P_{r,max} \\ &&& (x_r,y_r,z_r) \le (L_x,L_y,L_z), \end{aligned} \tag{27}$$

where $C_{sec,r}$ is the secrecy capacity defined in (23), $P_{r,max}$ is the maximum transmit power of the AR, $P_r = \text{Tr}\left(\boldsymbol{x}_r\boldsymbol{x}_r^{\dagger}\right)$ is the transmit power of the AR and $P_{r,max}$ the maximum transmit power, and (x_r,y_r,z_r) are the 3D coordinates of the UAV bound to (L_x,L_y,L_z). Because of the unknown eavesdropper location and CSI, the optimization problem is rewritten as

$$\begin{aligned} &\max_{\{\boldsymbol{W}_r,x_r,y_r,z_r\}} && \frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K}\Bigg[\log_2\Bigg(1+ \\ &&& \frac{\mid \boldsymbol{h}_{2,k}\boldsymbol{W}_r\boldsymbol{H}_1\boldsymbol{w}_{b,k}\mid^2}{\sum\limits_{i\neq k}\mid \boldsymbol{h}_{2,k}\boldsymbol{W}_r\boldsymbol{H}_1\boldsymbol{w}_{b,i}\mid^2+\|\boldsymbol{h}_{2,k}\boldsymbol{W}_r\|^2+1}\Bigg)\Bigg] \\ &\text{s.t.} && \text{Tr}\big(\boldsymbol{x}_r\boldsymbol{x}_r^{\dagger}\big) \le P_{r,max} \\ &&& (x_r,y_r,z_r) \le (L_x,L_y,L_z). \end{aligned} \tag{28}$$

Although we have incorporated practical system constraints in our model, we acknowledge that there are additional operational aspects, such as UAV energy consumption, flight time, and speed [25], which are not optimized in this paper.

IV. Proposed Solution

Given the available resources, which are one multi-antenna GBS and one multi-antenna AR, the secrecy capacity optimization problem becomes a user association and transmission parameter optimization problem. We perform UE clustering for user association, followed by GBS and AR beamforming and transmit power control, and UAV trajectory optimization. Figure 2 illustrates this. It is important to mention that the beamforming/power control and the UAV trajectory optimization are done through an iterative process. That is, the algorithm obtains the optimal power coefficients for every 3D location of the UAV. Hence, the beamforming and transmit power control of the UAV affects its trajectory adjustment. The details are discussed in Sections IV.B and IV.C

A. User Clustering

The goal of user clustering is to divide K users into two clusters, one cluster is to be served by the GBS and the other cluster is served the UAV.

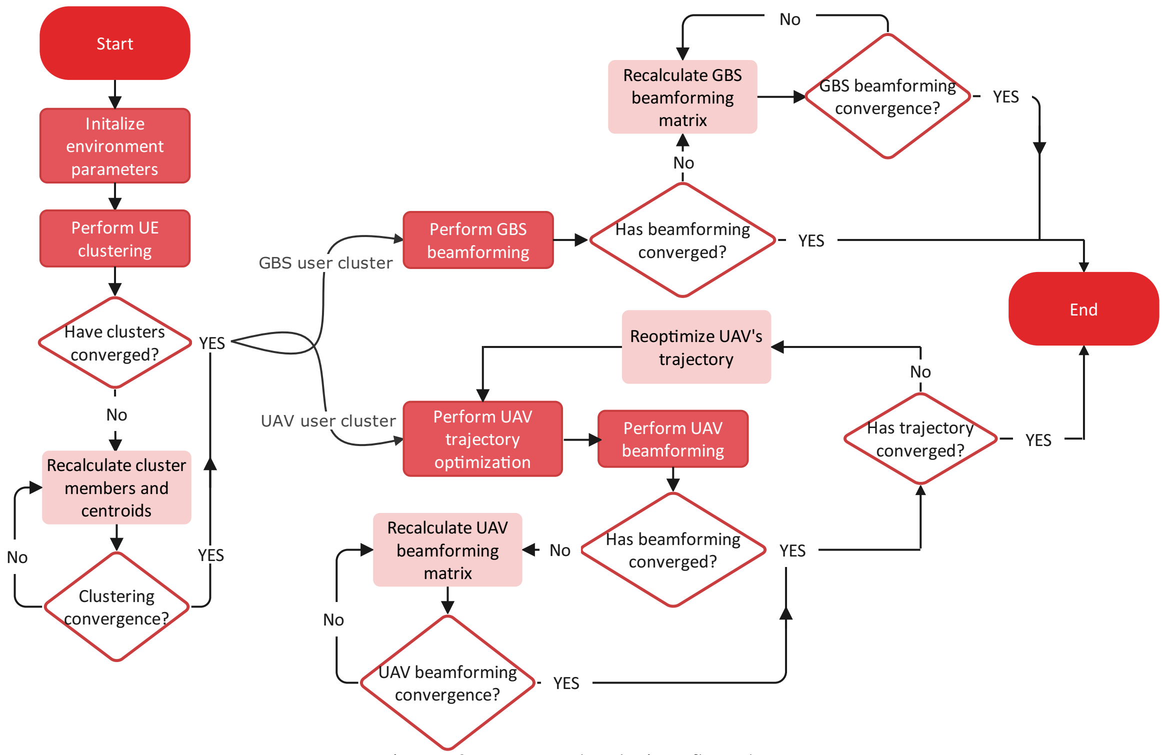


Figure 2: Proposed solution flowchart.

For solving the user clustering problem, we can employ an exhaustive search, but it entails a high computational complexity, which increases exponentially with the number of users. We instead apply K-means clustering, a unsupervised machine learning algorithm that is used for grouping a set of objects so that the similarity criterion of members in a group and the dissimilarity with members of other groups is maximized. K-mean works with any single or multi-dimensional metric that the data captures and user-defined target number of clusters [26]. It is a computationally efficient method compared to other techniques such as graph theory, fuzzy c-means clustering, and hierarchical clustering [27].

We consider the characteristics of wireless communication systems to determine the similarities of the data points. Because the objective it to associate users to base stations, one fixed (GBS) and one mobile (UAV), we take the normalized channel coefficients between UEs and GBS as the data points of the K-means clustering algorithm. This captures the variations of channel gains resulting from different RF propagation effects such as small-scale fading and shadow fading. Hence, we can define

$$h_{b,k}^{n} = \frac{h_{b,k}}{\parallel h_{b,k} \parallel_2}, \tag{29}$$

where $h_{b,k}^{n}$ is the normalized channel gain, $h_{b,k}$ is the channel gain between the GBS and k-th UE, and $\parallel . \parallel_2$ is the L_2 vector norm. Having these channel gains as data points, we apply K-mean clustering algorithm to determine the cluster centers, or centroids, and consequently the UEs associated with each centroid. The goal is to leverage the similarity of channels between the GBS and the UEs to create two UE clusters, where the UEs of one cluster are to be served by the GBS and the UEs of the other cluster by the UAV. This approach is applying the same clustering principle as other studies in the literature [28]–[31].

The K-mean clustering algorithm can be done as follows [32], [33]: (i) the initial centroids $C = \{c_1, c_2, \dots c_n\}$ are randomly selected as the n cluster centers of the K available data points: $U = \{u_1, u_2, \dots, u_k \dots, u_K\}$. Here, we consider two clusters c_1 and c_2 for the GBS and the UAV, and K users where $u_k = h_{b,k}^{n}$ can be considered as a data point of the k-th user. (ii) Distance between each data point, e.g., channel status, and the cluster centers is calculated to assign the data point to the nearest center. Different metrics can be used to measure the distance between data points such as Euclidean distance, Manhattan distance, etc. In this paper, we use L_2^2 norm or Euclidean distance. (iii) the centroids are updated to minimize the sum of squared distances between a user and its centroid,

$$\min_{C} \sum_{k} \min_{r \in R} \ d_{r,k} \tag{30}$$

where $d_{r,k} = \parallel u_k - c_r \parallel_2^2$ is the Euclidean distance between $C \triangleq \{c_r \mid r \in R\}$ and $R = 2$ represents the number of clusters. For example, the distance between the normalized channel gains and the centroids is $d_{r,k} = \parallel h_{b,k}^{n} - c_r \parallel_2^2$. Algorithm 1 represents a pseudocode for the K-mean clustering algorithm. The implementation of algorithm 1 for one scenario is shown in Figure 3 wherein data points are $h_{b,k}^{n}$

In addition to UEs' channel status, other data points can also be considered in the algorithm to study the problem. For example, distance based clustering or rate based clustering, where UEs with the nearest distance to the GBS and the highest downlink rates, respectively, would be grouped. Each scenario can have an effect on the system performance and should be chosen based on the objective and the ability or simplicity to obtain the necessary information to calculate the value for each UE. In Section V.C, we discuss about the results of different scenarios.

Algorithm 1 K-means user clustering in MU-MISO environment.

Input: U and C
Output: K_c and $C, \forall\, c \in C$
Initialize *cluster head set* $C_{\mathcal{CH}} = \emptyset$ *and* $c = 1$;
while $c \leq C$ **do**
  Randomly select a cluster head $\mathcal{CH}_c$ *from* U;
  Update $C_{\mathcal{CH}} = \{\mathcal{CH}_c, \forall\, c \in C\}$;
  $c = c + 1$;
end
repeat
  For each user $m \in U$, *calculate the minimum distance from the* $\mathcal{CH}_c$;
  Fit each user to the closest cluster;
  Update the cluster head $\mathcal{CH}_c$ *by taking the average of all the users* m;
until *The cluster members do not change;*

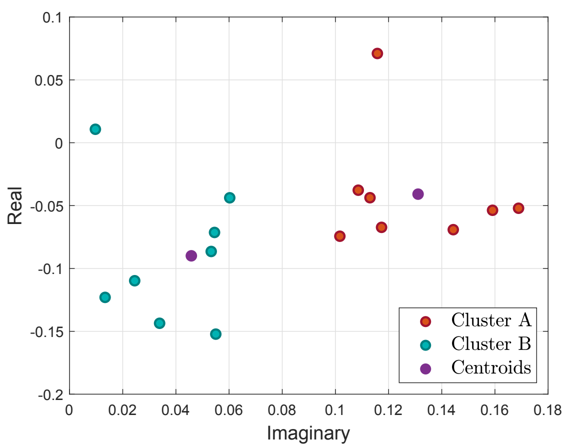


Figure 3: Channel-based clustering.

B. Optimal Beamforming and Power Control

Problem (26) is a traditional beamforming power control problem between a GBS and a user. This problem has been extensively studied in the literature [34], [35]. The beamforming vectors of the GBS are obtained by applying the weighted minimum mean square error (WMMSE) algorithm [36]. The WMMSE is an iterative closed form solution that optimizes the transmitter and receiver precoding vectors to maximize the sum rate of all UEs for a GBS power constraint. The precoding solution for the direct communication links is then [37]

$$\boldsymbol{W_{bk}} = \Big(\boldsymbol{H_0^H Q^H F Q H_0} + \frac{\mathrm{Tr}(\boldsymbol{F Q Q^H})}{P_{b,max}} \boldsymbol{I_M}\Big)^{-1} \boldsymbol{H_0^H Q^H F} \tag{31}$$

where $Q = diag\{q_1, \cdots, q_k\}$ is the receiver precoding, $F = diag\{f_1, \cdots, f_k\}$ is the weight matrix, and $\boldsymbol{I_M}$ is the covariance matrix.

The beamforming and power control for the relay communications problem is solved in the remainder of this section. Inspired by the zero-forcing (ZF) criterion and the channel singular value decomposition (SVD) based structure introduced in [21], we first propose a beamforming matrix structure for the UAV (i.e., $\boldsymbol{W}_r$) to eliminate interference among users. This converts the optimization problem (28) into a simplified convex optimization problem. Then, we solve the modified optimization problem using the Lagrangian function and Karush-Kuhn-Tucker (KKT) conditions to obtain the UAV's optimal beamforming matrix.

1) Beamforming Matrices: Beamforming is done at the GBS and the AR, each serving a distinct set of users. By using the concepts of channel inversion, ZF, and linear algebra, the multi-user interference can be minimized. From (15), the received signal at K UEs transmitted from the UAV can be written as

$$\boldsymbol{y}_2 = \boldsymbol{H}_2 \boldsymbol{W}_r \boldsymbol{H}_1 \boldsymbol{W}_{br}\, \boldsymbol{s}_K + \boldsymbol{H}_2 \boldsymbol{W}_r\, \boldsymbol{n}_1 + \boldsymbol{n}_2, \tag{32}$$

where $\boldsymbol{s}_K \in \mathbb{C}^{K\times 1}$ corresponds to the K transmit signals to the K UEs. The ZF criterion requires that $\boldsymbol{H}_2\boldsymbol{W}_r\boldsymbol{H}_1\boldsymbol{W}_{br}$ is to be a diagonal matrix with rank K, which implies that $\mathrm{Rank}(\boldsymbol{H}_1) \geq K$ and $\mathrm{Rank}(\boldsymbol{H}_2) \geq K$ [21]. By applying the SVD, $\boldsymbol{H}_2$ and $\boldsymbol{H}_1$ can be expressed as

$$\boldsymbol{H}_1 = \boldsymbol{U}_1\, \boldsymbol{\Sigma}_1\, \boldsymbol{V}_1^{\dagger}, \tag{33}$$

$$\boldsymbol{H}_2 = \boldsymbol{U}_2\, \boldsymbol{\Sigma}_2\, \boldsymbol{V}_2^{\dagger}, \tag{34}$$

where $\boldsymbol{U}_i$ and $\boldsymbol{V}_i$, for $i = 1, 2$, are unitary matrices and $\boldsymbol{\Sigma}_i \in \mathbb{C}^{K\times K}$ is a diagonal matrix with positive diagonal elements. Knowing the channel coefficients at the BS and at the UAV, to satisfy the ZF criterion, we propose the following beamforming matrices for the GBS and the UAV

$$\boldsymbol{W}_{br} = \boldsymbol{V}_1\, \boldsymbol{\Lambda}_b\, \boldsymbol{U}_2^{\dagger}, \tag{35}$$

$$\boldsymbol{W}_r = \boldsymbol{V}_2\, \hat{\boldsymbol{\Lambda}}_r\, \boldsymbol{U}_1^{\dagger}\, \boldsymbol{\Lambda}_r, \tag{36}$$

where $\boldsymbol{\Lambda}_b, \hat{\boldsymbol{\Lambda}}_r$, and $\boldsymbol{\Lambda}_r$ are all $K \times K$ diagonal matrices. Without loss of generality, it can be assumed that the elements of these two diagonal matrices are non-negative, representing the allocated beamforming power at the BS and the UAV, respectively.

2) Optimal AR Transmit Power Allocation:

Lemma IV.1. *The objective function defined in (28) can be written as*

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K} log_2\Big(1 + \frac{\lambda_{r,k}^2}{\lambda_{r,k}^2 + 1}\Big), \tag{37}$$

where $\lambda_{r,k}$ *is the* $\boldsymbol{\Lambda}_r(k,k)$.

Lemma IV.2. *The beamforming power constraint defined in (28) can be expressed as*

$$2\sum_{m=1}^{K}\sum_{n=1}^{K} |\boldsymbol{U}_2(m,n)|^2\, \sigma_{2,n}^{-2}\, \lambda_{r,m}^2 \leq P_{r,max}, \tag{38}$$

where $\sigma_{2,n}$ *is the* $\boldsymbol{\Sigma}_2(n,n)$.

The lemmas are proved in the Appendix (Section VI). Leveraging (37) and (38), the beamforming power optimization problem for the UAV at any location can be written as

$$\max_{\{\lambda_{r,m}\}} \quad \frac{1}{T}\sum_{t=1}^{T}\sum_{m=1}^{K} log_2\Big(1 + \frac{\lambda_{r,m}^2}{\lambda_{r,m}^2 + 1}\Big) \tag{39}$$

$$\text{s.t.} \quad 2\sum_{m=1}^{K}\sum_{n=1}^{K} |\boldsymbol{U}_2(m,n)|^2\, \sigma_{2,n}^{-2}\, \lambda_{r,m}^2 \leq P_{r,max}, \tag{40}$$

$$0 \leq \lambda_{r,m} \leq \lambda_{r,max}, \quad m \in \{1, ..., K\}, \tag{41}$$

where $P_{r,max}$ is the maximum available transmit power at the

UAV, and $\lambda_{r,max}$ is the maximum allocated power for each antenna.

The Lagrangian function of the optimization problem can be expressed as

$$\mathcal{L}(\lambda_{r,l}, \alpha_1, \alpha_{2,l}, \alpha_{3,l}) = +\sum_{l=1}^{K} log_2\Big(1 + \frac{\lambda_{r,l}^2}{\lambda_{r,l}^2 + 1}\Big) - \alpha_1\Big(2\sum_{l=1}^{K}\sum_{n=1}^{K} |\boldsymbol{U}_2(l,n)|^2\, \sigma_{2,n}^{-2}\, \lambda_{r,l}^2 - P_{r,max}\Big) - \Big(\alpha_{2,l}\Big(\sum_{l=1}^{K}\lambda_{r,l} - \lambda_{r,max}\Big)\Big) - \Big(\alpha_{3,l}\sum_{l=1}^{K} -\lambda_{r,l}\Big), \quad (42)$$

where α_1, $\alpha_{2,l}$, and $\alpha_{3,l}$ are the non-negative Lagrangian multipliers corresponding to the first and the second constraints, respectively.

Theorem IV.3. *The optimal beamforming power for the lth antenna of the UAV can be obtained as*

$$\lambda_{r,l}^{*} = \begin{cases} 0 & \lambda_{r,l}^{\dagger} \leq 0, (\alpha_1^{*}\, F\, ln\, 2) > 0.25 \\ \lambda_{r,l}^{\dagger} & 0 < \lambda_{r,l}^{\dagger} < \lambda_{r,max} \\ \lambda_{r,max} & \lambda_{r,l}^{\dagger} \geq \lambda_{r,max} \end{cases}$$

in which

$$\lambda_{r,l}^{\dagger} = \sqrt{\frac{1}{4}\Big(\sqrt{1 + \frac{2}{\alpha_1^{*}\, F\, ln\, 2}} - 3\Big)}, \quad (43)$$

where F is a constant and equals to $\sum_{n=1}^{K} |\boldsymbol{U}_2(l,n)|^2\, \sigma_{2,n}^{-2}$, and the Lagrangian multiplier α_1^{*} can be obtained by replacing (43) into the first constraint of (40) when the equality holds.

The optimal beamforming matrix and transmit power formulations for the AR defined above are used for the UAV trajectory optimization.

C. UAV Trajectory Optimization

The objective function of (28) is non-convex with respect to parameters x_r, y_r, z_r, P, and the constraints, and the problem is NP-hard [21], [38]–[40]. We, therefore, propose a machine learning solution where the UAV trajectory is updated through a transition process based on the current system state. Since the next system state is independent from the previous state and action, the process can be modeled as a Markov decision process (MDP). In order to avoid intractably high dimensionality for the high state-action space, we propose a DQN. It is noteworthy that the following proposed DQN is based on basic reinforcement learning algorithms such as Q-learning and deep reinforcement learning. The aim is to use DQN as an alternative tool for solving this NP-hard optimization problem while consuming less power and computational resources [41]–[43]. Depending on an application, one can extend the following framework to more advanced learning models that suit particular use cases.

1) MDP Settings: The MDP for the UAV agent is composed of the state space $\mathcal{S}$, the action space $\mathcal{A}$, the reward space $\mathcal{R}$, and the transition probability space $\mathcal{T}$. At time slot t, the agent observes the state $s_t \in \mathcal{S}$, and takes action $a_t \in \mathcal{A}$ based on its policy. Depending on the distribution of the transition probability $\mathcal{T}(s_{t+1}|s_t, a_t)$, the agent is then transferred to the new state s_{t+1}. Since the transition probability is specific to the operational environment, we choose the Q-learning method as a model-free algorithm to find the best policy for each action in each state. This means that we do not need to know $\mathcal{T}$, but we need to carefully define the states, the actions, and the reward.

State: The set of states is defined as $\mathcal{S} = \{s_1, s_2, ..., s_t, .., s_T\}$, where t is the time slot index. Each state s_t corresponds to the 3D coordinates of the UAV and the users served by the AR.

Action: The states are transitioned according to the defined set of actions defined as $\mathcal{A} = \{a_1, a_2, ..., a_t, .., a_T\}$, where each action consists of three parts related to the UAV movement, $a_t = \{\delta_x, \delta_y, \delta_z\}$, where δ_x , δ_y, and δ_z represent the movement in the x, y and z directions. The movement along each axis is assumed to change positively or negatively, or remain in the original position. Hence, here we consider 3 possible directional movements for the 3 axes of the AR trajectory, resulting in 27 possible actions for the AR.

Reward: After taking action a_t in state s_t, the UAV agent will receive a reward $R_t(s_t, a_t)$. The UAV gets more rewards for actions that lead to higher legitimate user rates. We define the reward function accordingly:

$$R_t(s_t, a_t) = \sum_{k=1}^{K_r} C_{r,k} \quad (44)$$

2) Deep Q-Network Method: The DQN, initially proposed by Google Deep Mind [44], integrates the RL and deep learning methods. This technique uses the power of nonlinear functions, specifically DNNs, in order to approximate the Q-values and handle highly dimensional state-action problems.

There are two DNNs of the same structure: a training network and a target network. The training network outputs the Q-values associated with the actions of the UAV in each state. The target network supervises the training network by providing the target Q-values obtained from the Bellman equation [45],

$$Q^{*}(s,a) = E_{s'}\Big[R(s,a) + \gamma \times \max_{a \in \mathcal{A}}\ Q(s',a')\Big], \quad (45)$$

which provides the optimal state-action pairs, where s' and a' symbolize the next state and action. Parameter $\gamma \in (0, 1)$ denotes the discount factor that affects the importance of the future reward.

The target values are compared with the outputs of the training network to minimize the loss function,

$$L(\theta) = \mathbb{E}\Bigg[\Big(\Big[r_t + \gamma \times \max_{a \in \mathcal{A}}\ Q(s_{t+1}, a_{t+1}; \theta^{\dagger})\Big] - \Big[Q(s_t, a_t; \theta)\Big]\Big)^2\Bigg], \quad (46)$$

where the Q-value of the first term is obtained from the target network and the Q-value of the second term is obtained from the training network. Parameters $\theta^{\dagger}$ and θ denote the weights of the target network and training network, respectively. The $\theta^{\dagger}$ coefficients are updated every few time slots in order to ensure the stability of

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2023.3285737

the target values and, hence, facilitate stable learning.

Algorithm 2 DQN for UAV trajectory optimization.

Initialize ϵ_{start}, ϵ_{end}, decay
Initialize T time slots, $\mathcal{J}$ episodes
Initialize *replay memory* $\mathcal{M}$ *to capacity N*
Initialize θ, $\theta^{\dagger}$, γ, α, B
for *episode = 1, 2, ...,* $\mathcal{J}$ **do**
  Reset Environment
  for *t = 1, 2, ..., T* **do**
    Obtain the initial observation s_t
    if $\epsilon > random(0, 1)$ **then**
      Select random $a_t \in \mathcal{A}$
    else
      $a_t = \text{argmax}_{a \in \mathcal{A}}\ Q(s_t, a; \theta)$
    end
    UAV executes action a_t *in the environment*
    Observe transition $\mathbf{s}_{t+1}$, r_t
    Store transition in $\mathcal{M}$**:** $\mathcal{M} \leftarrow \mathcal{M} \cup \{\mathbf{s}_t, a_t, r_t, s_{t+1}\}$
    Sample mini-batch from $\mathcal{M}$**:** $\mathbf{e}_i = (s_i, a_i, r_i, s_{i+1})$
    Train the DNN and compute the estimated Q-value
    Calculate the loss between estimated and calculated Q-value
    Derive loss gradient and update θ **in training network**
    Every B steps, copy θ **to** $\theta^{\dagger}$ **to update target network**
  end
  $\epsilon \leftarrow$ *updated via* ϵ*-greedy algorithm*
  Store reward for each episode
end
Result: Optimal UAV trajectory

As the UAV takes an action, the system generates a record of experience. At time step t, the experience contains the current state s_t, the action a_t, the reward r_t, and the next state s_{t+1}, formed as a tuple $e_t = (s_t, a_t, r_t, s_{t+1})$. Each such experience is stored in a replay memory with the capacity of N, such that $\mathcal{M} = \{e_1, ..., e_t, ..., e_N\}$. The memory is a queue-like buffer that stores the latest N experience vectors. We use a mini-batch sample from the replay memory to feed the input of the training network. The main reason for using the mini-batch samples from the reply memory is to break possible correlations between sequential states of the environment, and thereby facilitate generalization.

The UAV applies a gradient descent algorithm,

$$\nabla_{\theta} L(\theta) = -\mathbb{E}\Bigg[2\ \nabla_{\theta} Q(s_t, a_t; \theta)\Bigg(r_t + \gamma \times \max_{a \in \mathcal{A}}\ Q(s_{t+1}, a_{t+1}; \theta^{\dagger}) - Q(s_t, a_t; \theta)\Bigg)\Bigg], \quad (47)$$

to update θ an $\theta^{\dagger}$ as the weights of the DNNs with the aim of minimizing the prediction error.

Finally, we apply the $\epsilon-$greedy algorithm to select an action while balancing the exploration and the exploitation of the UAV in the environment. In this algorithm, the UAV explores the environment with the probability of ϵ by choosing a random action. More precisely, the UAV exploits the environment with the probability of $1-\epsilon$ by choosing the actions that maximize the Q-value function, i.e., $a^* = \text{argmax}_{a \in \mathcal{A}}\ Q(s, a; \theta)$. A high value of ϵ is initially set in the model for the UAV to spend more time for the exploration. As the agent obtains more knowledge about the environment, the ϵ value is gradually decreased to leverage the experience and choose the best actions for the UAV, rather than continuing with the exploration.

Algorithm 2 details the DQN-based algorithm used by the UAV agent for optimizing the sum-rate of the UEs that are served via the AR.

In summary, the proposed techniques accomplish the following: i) User-BS association employing K-means clustering (e.g., channel, rate, or distance based), ii) multi-user beamforming and power management for the GBS, iii) UAV trajectory optimization in conjunction with multi-user beamforming and power management. The objective is to maximize the secrecy rate which can be used to evaluate the effectiveness of the proposed security measurein protecting against eavesdropping attacks [46], [47], especially in the context of wireless communication [48]–[51]. Since the CSI of the eavesdropping channel cannot be obtained for passive, receive-only eavesdroppers, our solution maximizes the user rate for it to be generally applicable without requiring collaboration with eavesdroppers or wasting power for generating artificial noise in random directions, because of the unknown eavesdropper locations, as opposed to using this power to increase the user rate. The secrecy capacity also provides a unified measurement framework for the numerical analyses presented in the following section.

V. Numerical Analysis and Discussion

In this section, we present simulation results to evaluate the secrecy performance of the UAV-assisted communications system, where users are clustered and served by a fixed and a mobile access point. In the presence of an eavesdropping attack, our solution jointly optimizes of the UAV trajectory, GBS beamforming, and AR beamforming coefficients. The numerical analysis quantifies the impact of different user clustering technique, discount factor (Gamma) values, learning rates, and number of users on the achievable secrecy capacity of the system.

The simulation scenario is illustrated in Fig. 1 and consist of multiple single antenna ground UEs, an AR, and a group of malicious nodes that is performing a passive eavesdropping attack on the downlink transmission. The terrestrial users and the eavesdroppers are randomly distributed in a 2D area. The AR is launched at a random location and height and is equipped with an antenna array to enable communications with the GBS and the UEs. Table II captures the simulation parameters. The simulations are performed with Python 3.6 and PyTorch 1.7.

A. Hyper-parameters

The hyper-parameters of the learning algorithm need to be optimized for our specific problem and environment. Therefore, Fig. 4 and Fig. 5 numerically evaluate the secrecy capacity of the UEs served through the AR for different discount factors Gamma and learning rates (LRs). Additionally, Fig. 4 and Fig. 5 verify the convergence of the proposed solution across these different settings. The results presented in both figures are for the case of 16 ground users where there are 8 users in each clusters as shown earlier in Fig. 3. These figures plot the total achieved secrecy capacity of the user cluster served by the AR over the training time for different hyper-parameter values.

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2023.3285737

Table II: Simulation parameters.

Parameter	Value
Area length (L_x)	20 m
Area width (L_y)	20 m
UAV height (z_u)	20-80 m
UAV trajectory step along the x or y axis	0.5 m
UAV trajectory step along the z axis	2 m
GBS height (z_s)	15 m
Path loss at 1 m reference distance (λ_0)	-40 dB [52]
Path loss exponent (α)	2
Rician factor (β)	10 dB
ABG distance-dependent exponents (ρ_G)	2.1
ABG intercept ($\jmath$)	31.7 dB
ABG frequency-dependent exponents (γ_G)	2
Shadow fading (χ_σ^{G2G})	3.9 dB
Central frequency	3.2 GHz
Noise variance	10^{-2}
Number of ground users (K)	4-64
Number of ground eavesdroppers (E)	2-10
Number of K-means clusters (R)	2
Number of episodes	2×10^4
Number time slots per episode	200
Learning rate (LR)	10^{-4}
Discount factor	0.9
Replay memory size	10^5 entries
Mini-batch size	64
Update rate of target network	10

When the discount factor is very high, the agent equally considers the future and current rewards. Fig. 4 shows that this leads to low performance. The best result for our scenario is achieved by slightly discounting the future reward, corresponding to a Gamma of 0.9.

By configuring higher LRs, the agent becomes increasingly biased to take the same action that will enforce the learning policy to be particular to a deterministic environment. On the other hand, for very low LRs the DQL agent keeps exploring the environment in a complete random behavior without learning. A moderate LR provides the equilibrium between a deterministic and stochastic environment. Fig. 5 compares the learning outcome for three LRs, where a LR of 10^{-4} provides the best result.

Note that one reason of DQN failure is related to the choice of the hyperparameters. There are a number of search techniques that can be used to adjust hyperparameters. In this analysis, we have employed the grid search technique that

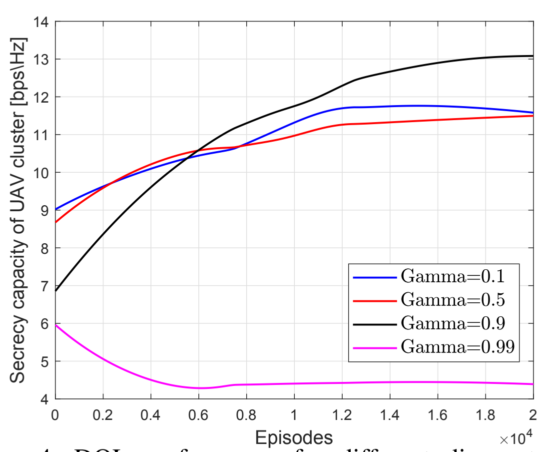


Figure 4: DQL performance for different discount factors Gamma for a learning rate of 10^{-4}.

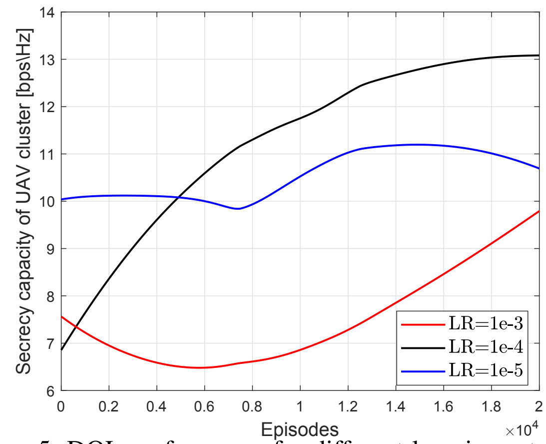


Figure 5: DQL performance for different learning rates (LRs) for Gamma = 0.9.

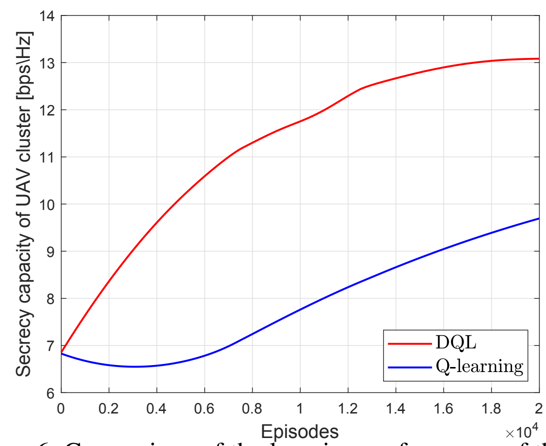
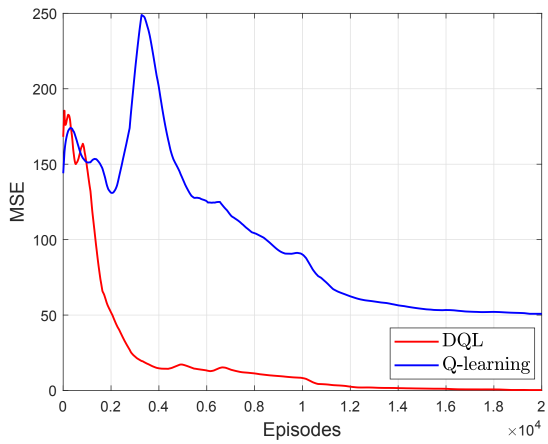


Figure 6: Comparison of the learning performance of the DQL and Q-learning for the UAV trajectory optimization.

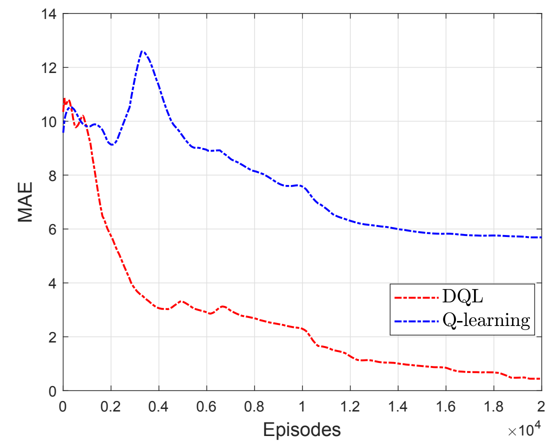
involves specifying a range of values for each hyperparameter and then training the DQN with all possible combinations of these values. The combination of hyperparameters that produces the best performance is then selected.

B. Learning Performance Evaluation

Figure 6 compares of learning and convergence performance of the proposed DQL scheme with the Q-learning as a benchmark learning algorithm. It plots the total secrecy capacity of the user cluster served by the AR over the number of learning episodes for the DQL and Q-learning trajectory and power optimization. The curves show that the total secrecy capacity of user served by the UAV tends to increase over the episodes until convergence. This validates our approach to define the reward so as to maximize the user rates with unknown CSI of the eavesdropping channel. Initially, the total secrecy capacities match for the two algorithms. This is so because of lacking interaction with the environment to provide enough data for training the learning agents. As the learning evolves, favorable actions become more easily discriminated from the unfavorable ones by exploring the environment. It is noticeable that the DQL performance substantially exceeds the Q-learning performance due to its ability to approximate



(a)



(b)

Figure 7: Comparison of the MSE (a) and MAE (b) losses of the DQL and Q-learning over the number of episodes.

the Q-values instead of an inefficient Q-table of the QL, which allows learning gigantic state and action in fewer episodes.

In order to further analyze the effectiveness and the convergence performance of the learning performance of the DQL design for optimizing the UAV trajectory, Fig. 7 plots the mean square error (MSE) and mean absolute error (MAE) of the DQL and Q-learning solutions over the number of episodes. The results show how the loss error is minimized by adjusting the weights of the neural network used to approximate the Q-function. This indicates how well the proposed algorithm is performing and illustrates its convergence toward an optimal policy. The MSE and MAE are calculated by comparing the estimated Q-values using the learned DQL and Q-learning models versus their actually computed values. Those figures reveal that the quality of the UAV actions are rather poor during the early training phase. As the learning continues, more measurements are accumulated that yield to improved actions taken by the UAV agent for both algorithms. The DQL method meets a MSE target of 50 an order of magnitude faster than Q-learning (Fig. 7a). It converges faster and achieves a 35% higher secrecy capacity after 20000 episodes (Fig. 6).

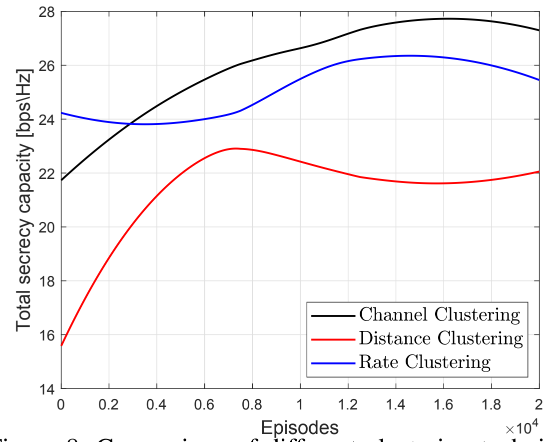


Figure 8: Comparison of different clustering techniques.

C. Clustering Performance Evaluation

Here we evaluate the performance of the proposed user clustering scheme and the employed metric on the total secrecy rate performance of the system. We consider three metrics for the clustering algorithm presented in Algorithm 1: *distance clustering*, where UEs are grouped based on their distances to the GBS, *rate clustering*, where UEs are grouped based on their downlink rates while being served by the GBS, and *channel clustering*, where UEs are grouped based on the normalized channel coefficients.

Figure 8, shows the total secrecy capacity of the system after clustering, beamforming, and UAV trajectory optimization for the three clustering metrics. We observe that the proposed channel clustering metric outperforms the rate and distance clustering metrics in terms of total secrecy capacity.

D. Overall Performance Evaluation

We explain the overall performance evaluation of the proposed method in two parts. In the first part, the impact of the optimal beamforming and power control performance evaluation is studied. In particular, the proposed method is compared with three scenarios where our optimal beamforming and power control are only partially implemented. In the second part, the impact of the UAV trajectory on the secrecy capacity is studied. Specifically, the UAV's 3D movement is shown in a scenario in which two clusters of UEs are simultaneously served, one by the UAV, which relocates to best serve the UEs in the cluster, and the other by the GBS.

The context that we study is unique compared to other studies as captured in Table I. We develop a framework that involves user clustering, multi-user beamforming, power control, and reinforcement learning for solving the problem and there are no existing studies that propose comparable solution. Therefore, we define our own benchmarks to evaluate the proposed framework and the importance of each component comprising it. The baseline techniques are: AR deployment without optimal GBS beamforming (UAV+NoBF), no AR deployment with optimal GBS beamforming (NoUAV+BF), and no AR deployment and without optimal GBS beamforming

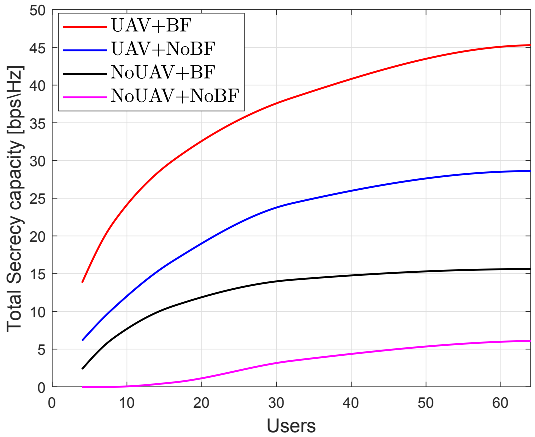


Figure 9: Comparison of different techniques Vs number of users for total secrecy capacity of the system.

(NoUAV+NoBF). In all cases where the AR is deployed the optimal beamforming and power control of the AR is activated.

Figure 9 shows the achieved total secrecy capacity over the number of users. The secrecy capacity improves with the number of users for all schemes. The proposed solution clearly outperforms the other techniques. The UAV+NoBF scheme achieves a better secrecy capacity than the NoUAV+BF scheme. That is, deploying an AR is more useful for improving the secrecy capacity than employing optimal multi-user beamforming at the GBS. Nevertheless, beamforming and power control schemes have a notable contribution to the secrecy performance of the system as can be observed when comparing the performance of the NoUAV+NoBF scheme with the proposed solution and other benchmark techniques. The optimal beamforming and power control increases the SNR and reduces the multi-user interference while minimizing the likelihood of eavesdropping and improving the overall secrecy capacity. The addition of the UAV as an AR allows to serve those users effectively that have a worse channel to the GBS. This is accomplished by the proposed clustering method and the UAV trajectory optimization along with beamforming and power control.

Figure 10 illustrates the dynamic 3D trajectory optimization process for the case of 16 users where there are 8 users in each cluster as a result of Algorithm 1 that employs channel based clustering. We observe that the UAV moves toward the center of the area where the users are located, and its final position is near the minimum height to be as close as possible to the UEs served by the AR for ensuring good channels to be able to lower the transmission power and thus increase the secrecy capacity, while also ensuring that the UAV stays within its operational limits and avoids ground obstacles. Overall, the dynamic 3D trajectory optimization process, combined with optimal multi-user beamforming and power control, helps achieve high secrecy capacities (Fig. 9) in the presence of eavesdroppers.

E. Known vs. Unknown Information of Eavesdroppers

In order to put our contribution in context and provide further justification for our optimization framework, we consider the case where the location and the channel states of eavesdroppers are known to the GBS and the UAV in the proposed communication context. Knowning the CSI of eavesdropping channels allows employing the secrecy capacity (24) as the reward function.

We simulate 16 users that are clustered in two groups to be served by the GBS or AR resulting from the channel based clustering with known eavesdropper locations and CSI. Figure 11 shows the resulting average secrecy capacity per user with and without eavesdropping information available to the network. For the case of unknown eavesdropping channels, we employ the proposed optimization solution and reward function based on the legitimate user rate. The UAV adjusts its power and trajectory according to the available information. As expected, having the information of malicious actors eavesdropping on the wireless links allows the network to adjust its parameters better and increase the secrecy capacity.

Figure 11 also indicates that the secrecy capacity performance gap of not knowing the channel characteristics of eavesdroppers is not significant. In other words, blindly optimizing the secrecy capacity by focusing on the legitimate user rates produces an outcome that is very close to an optimization framework that has and leverages the full information about eavesdroppers. The reason for this is that the proposed practical solution with unknown CSI imperceptibly considers the possible CSI between the base stations and the eavesdroppers.

We conclude that despite the practical assumption of not knowing the CSI of eavesdropping channels and not exploring methods other than optimizing the user rates, the proposition of this paper, the proposed communications and optimization framework can accomplish a performance that is very close to a network that has access to the full information about eavesdroppers. This encourages doing further research on improving the proposed technique, for example, by considering partially known information about eavesdroppers or other reasonable assumptions, or even exploring new physical layer security metrics.

F. User Mobility

In this subsection, we examine the secrecy capacity of mobile users with a static eavesdropper. Without loss of generality, the mobility of the ground users will be over the x-axis with a fixed y-position. We define the distance step parameter (dx), which corresponds to the granularity of movement. Then, $UE_{X_C^{t+1}} = UE_{X_C^t} + dx$, can be defined to model the mobility of the ground users, where $UE_{X_C^{t+1}}$ is the next center x-positions of all the ground users in the next time step and for each new center the users are redistributed randomly around the center. This process enables the users to simulate a realistic movement pattern that reflects their actual movement.

We consider 16 ground users to be served either directly by the GBS or through the AR. Additionally, with each movement of step dx , the proposed solution of Fig. 2 re-clusters the

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2023.3285737

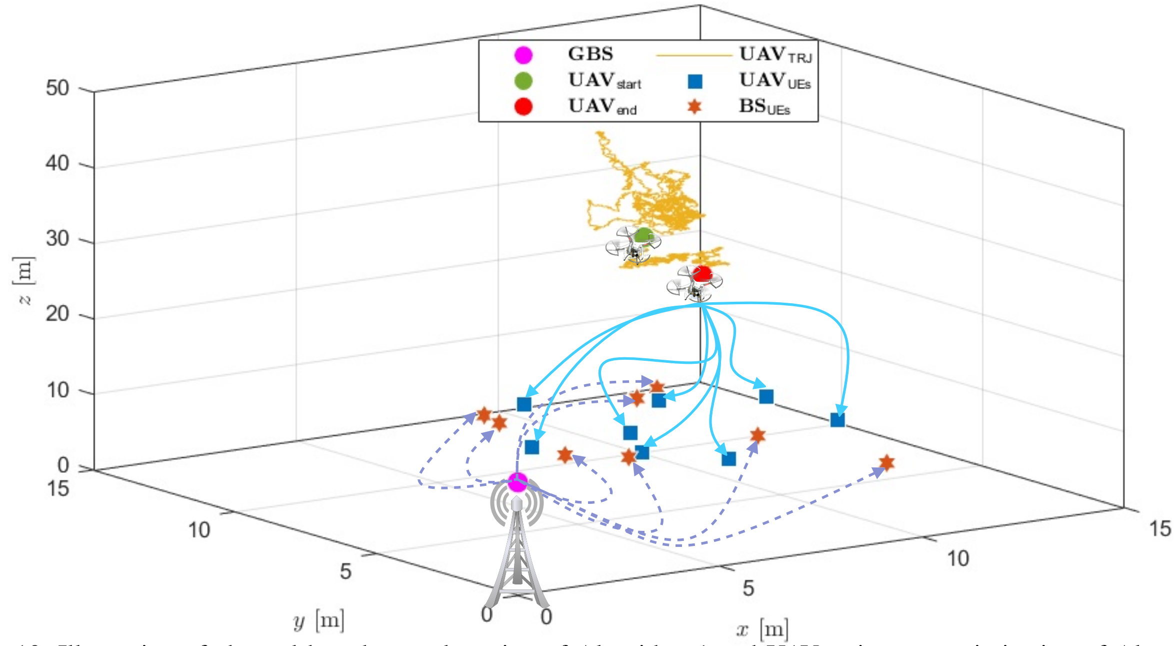


Figure 10: Illustration of channel based user clustering of Algorithm 1 and UAV trajectory optimization of Algorithm 2.

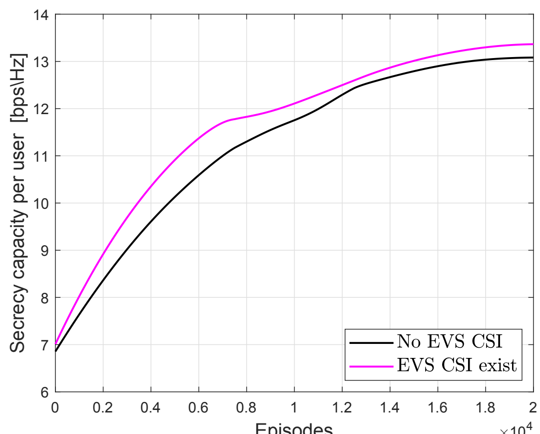


Figure 11: Secrecy capacity of the proposed optimization framework for the cases of known and unknown CSI of eavesdroppers, employing the secrecy capacity and user capacity as the reward function, respectively.

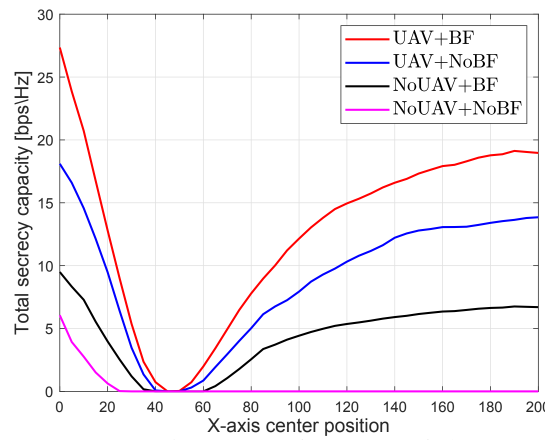


Figure 12: Comparing the performance of the proposed solution under user mobility scenario with other benchmark techniques.

users, re-performs beamforming and power control for the GBS and UAV transmissions, and re-optimizes the trajectory of the UAV given the new positions of users. Figure 12 presents the obtained total secrecy capacity over the center position of the moving user cluster for the proposed solution and the benchmarks introduced in Fig. 9.

The results of Fig. 12 show that the proposed solution achieves a higher total secrecy capacity compared to the other schemes. By optimizing the trajectory of the UAV, the system ensures that it flies as close as possible to the users being served by the AR enabling good channels and lower transmit powers. We observe that the secrecy capacity of the proposed solution drops to zero only for the case where the center of the user cluster matches the eavesdropper position. After the users pass the eavesdropper, the secrecy capacity rapidly recovers.

Notice the secrecy capacity after passing the eavesdroppers is lower than before reaching it. This is because of the lower data rates achieved by the direct GBS links experiencing a higher path loss with increasing distance. On the other hand, the NoUAV+NoBF scheme has reached the zero secrecy capacity much earlier and remains at this state even after leaving the eavesdropper behind. When comparing the performance of the UAV+NoBF and the NoUAV+BF schemes, we again realize the effectiveness of deploying the AR for achieving a higher secrecy capacity.

VI. Conclusions

This paper addressed the major eavesdropping problem in present-day wireless communications. We developed a practical framework against passive eavesdroppers in multi-user cellular networks without knowledge of the eavesdroppers' locations and CSI channels. Considering the unknowns, we optimized the user rates employing advanced wireless techniques at the physical layer to improve the sum-secrecy capacity among all users in a cell. Specifically, we suggested employing multi-user beamforming and deploying a UAV that serves as an AR. We clustered the users into two groups wherein users are either served by the GBS or by the AR, whose 3D position, multiuser beamforming matrix, and transmit powers are optimized combining closed-form expressions with machine learning techniques. Specifically, we designed and analyzed a DQN for the UAV trajectory optimization subproblem. Numerical results showed that the proposed system achieves highest secrecy capacities and scales well over the number of users to be served.

Lessons learned from this work can lead to a number of research directions for solving open research challenges. We will examine additional UAV specific operational constraints, including energy consumption, flight time, and speed, in future work. These are especially important for implementation and deployments of ARs with today's small UAVs. One can prototype and validate the presented techniques on the Aerial Experimentation and Research Platform for Advanced Wireless (AERPAW) [53], which facilitates implementing the proposed communications system with software radios and conducting different types of mobility experiments by leveraging AERPAW's unmanned ground vehicles.

Appendix

A. Proof of Lemma IV.1

Proof. The substitution of $\boldsymbol{H}_1$, $\boldsymbol{H}_2$, $\boldsymbol{W}_{br}$, and $\boldsymbol{W}_r$, which are defined in (33), (34), (35), and (36), respectively, in $\boldsymbol{y}_2$ defined in (32) yields

$$\begin{aligned}\boldsymbol{y}_2 = \boldsymbol{U}_2\,\boldsymbol{\Sigma}_2\,\underbrace{\boldsymbol{V}_2^\dagger\,\boldsymbol{V}_2}_{\boldsymbol{I}}\,\hat{\boldsymbol{\Lambda}}_r\,\boldsymbol{U}_1^\dagger\,\boldsymbol{\Lambda}_r\,\boldsymbol{U}_1\,\boldsymbol{\Sigma}_1\,\underbrace{\boldsymbol{V}_1^\dagger\,\boldsymbol{V}_1}_{\boldsymbol{I}}\,\boldsymbol{\Lambda}_b\,\boldsymbol{U}_2^\dagger\,\boldsymbol{s}_K \\ +\,\boldsymbol{U}_2\,\boldsymbol{\Sigma}_2\,\underbrace{\boldsymbol{V}_2^\dagger\,\boldsymbol{V}_2}_{\boldsymbol{I}}\,\hat{\boldsymbol{\Lambda}}_r\,\boldsymbol{U}_1^\dagger\,\boldsymbol{\Lambda}_r\,\boldsymbol{n}_1\,+\,\boldsymbol{n}_2. \qquad (48)\end{aligned}$$

Therefore we have

$$\begin{aligned}\boldsymbol{y}_2 = \underbrace{\boldsymbol{U}_2\,\boldsymbol{\Sigma}_2\,\hat{\boldsymbol{\Lambda}}_r\,\boldsymbol{U}_1^\dagger}_{\boldsymbol{I}}\,\boldsymbol{\Lambda}_r\,\underbrace{\boldsymbol{U}_1\,\boldsymbol{\Sigma}_1\,\boldsymbol{\Lambda}_b\,\boldsymbol{U}_2^\dagger}_{\boldsymbol{I}}\,\boldsymbol{s}_K \\ +\,\boldsymbol{U}_2\,\boldsymbol{\Sigma}_2\,\hat{\boldsymbol{\Lambda}}_r\,\boldsymbol{U}_1^\dagger\,\boldsymbol{\Lambda}_r\,\boldsymbol{n}_1\,+\,\boldsymbol{n}_2, \qquad (49)\end{aligned}$$

where the matrices $\boldsymbol{I}$ are obtained using the ZF criterion, i.e., $\boldsymbol{U}_2\,\boldsymbol{\Sigma}_2\,\hat{\boldsymbol{\Lambda}}_r\,\boldsymbol{U}_1^\dagger = \boldsymbol{I}$ and $\boldsymbol{U}_1\,\boldsymbol{\Sigma}_1\,\boldsymbol{\Lambda}_b\,\boldsymbol{U}_2^\dagger = \boldsymbol{I}$. As a result, the simplified equation is

$$\boldsymbol{y}_2 = \boldsymbol{\Lambda}_r\,\boldsymbol{s}_K\,+\,\boldsymbol{\Lambda}_r\,\boldsymbol{n}_1\,+\,\boldsymbol{n}_2 \qquad (50)$$

in which

$$\boldsymbol{\Lambda}_r = \begin{pmatrix}\lambda_{r,1} & 0 & \dots & 0 & 0\\ \vdots & \vdots & \dots & \vdots & 0\\ 0 & \dots & \lambda_{r,k} & 0 & 0\\ 0 & \vdots & \dots & \vdots & 0\\ 0 & 0 & \dots & 0 & \lambda_{r,K}\end{pmatrix}_{K\times K},\ \boldsymbol{s}_K = \begin{pmatrix}s_{1,1}\\ \vdots\\ s_{k,1}\\ \vdots\\ s_{K,1}\end{pmatrix}_{K\times 1}.$$

$\boldsymbol{\Lambda}_r$ is a diagonal matrix and $\boldsymbol{n}_1, \boldsymbol{n}_2 \in \mathbb{C}^{K\times 1}$ are the noise vectors. The simplified SINR can then be written as

$$SINR = \frac{\lambda_{r,k}^2}{\lambda_{r,k}^2+1}, \qquad (51)$$

which proves Lemma IV.1.

□

B. Proof of Lemma IV.2

Proof. The beamforming power at the relay can be simplified as follows

$$\begin{aligned} P_r &= \mathrm{Tr}\big(\boldsymbol{x}_r\boldsymbol{x}_r^\dagger\big) && (52)\\ &= \mathrm{Tr}\bigg(\boldsymbol{W}_r\Big(\underbrace{\boldsymbol{H}_1\boldsymbol{W}_b\boldsymbol{W}_b^\dagger\boldsymbol{H}_1^\dagger}_{\textit{term i}}+\boldsymbol{I}\Big)\boldsymbol{W}_r^\dagger\bigg), && (53)\\ &= \mathrm{Tr}\bigg(\boldsymbol{W}_r\Big(\underbrace{\boldsymbol{U}_1\boldsymbol{\Sigma}_1\boldsymbol{\Lambda}_b\boldsymbol{\Lambda}_b^\dagger\boldsymbol{\Sigma}_1^\dagger\boldsymbol{U}_1^\dagger}_{\boldsymbol{I}}+\boldsymbol{I}\Big)\boldsymbol{W}_r^\dagger\bigg), && (54)\\ &= 2\times\mathrm{Tr}\Big(\boldsymbol{W}_r\,\boldsymbol{W}_r^\dagger\Big) && (55)\end{aligned}$$

where (53) is obtained by replacing (13) and (14) in (52), (54) is derived by substituting (33) and (35) in *term i* of (53), and (54) is obtained from $\boldsymbol{U}_1\,\boldsymbol{\Sigma}_1\,\boldsymbol{\Lambda}_b\,\boldsymbol{U}_2^\dagger = \boldsymbol{I}$. Subsequently, by replacing (36) into (55), we have

$$\begin{aligned} P_r &= 2\times\mathrm{Tr}\Big(\boldsymbol{W}_r\,\boldsymbol{W}_r^\dagger\Big)\\ &= 2\times\mathrm{Tr}\Big(\boldsymbol{V}_2\,\underbrace{\hat{\boldsymbol{\Lambda}}_r\,\boldsymbol{U}_1^\dagger}_{\textit{term ii}}\,\boldsymbol{\Lambda}_r^2\,\underbrace{\boldsymbol{U}_1\,\hat{\boldsymbol{\Lambda}}_r}_{\textit{term iii}}\,\boldsymbol{V}_2^\dagger\Big) && (56)\\ &= 2\times\mathrm{Tr}\Big(\boldsymbol{V}_2\boldsymbol{\Sigma}_2^{-1}\boldsymbol{U}_2^\dagger\boldsymbol{\Lambda}_r^2\boldsymbol{U}_2\boldsymbol{\Sigma}_2^{-1}\boldsymbol{V}_2^\dagger\Big) && (57)\\ &= 2\sum_{m=1}^{K}\sum_{n=1}^{K}|\boldsymbol{U}_2(m,n)|^2\,\sigma_{2,n}^{-2}\,\lambda_{r,m}^2, && (58)\end{aligned}$$

where considering $\boldsymbol{U}_2\,\boldsymbol{\Sigma}_2\,\hat{\boldsymbol{\Lambda}}_r\,\boldsymbol{U}_1^\dagger = \boldsymbol{I}$ for*term ii* and *term iii* of (56) yields (57). □

C. Proof of Theorem IV.3

Proof. We need to obtain the optimum beamforming power elements and Lagrange multipliers, i.e., $\lambda_{r,l}^*$, α_1^*, $\alpha_{2,l}^*$, and $\alpha_{3,l}^*$, where $l = \{1,\cdots,K\}$. To this end, we apply the Karush-Kuhn-Tucker (KKT) conditions to this problem as has been applied to similar ones [54] [55] [56]. From the gradient condition and the complementary slackness condition, we have

$$\begin{aligned} &\nabla_{\lambda_{r,l}}\mathcal{L}(\lambda_{r,l}^*,\alpha_1^*,\alpha_{2,l}^*,\alpha_{3,l}^*) = 0, && (59)\\ &-\alpha_1^*\bigg(2\sum_{l=1}^{K}\sum_{n=1}^{K}|\boldsymbol{U}_2(l,n)|^2\,\sigma_{2,n}^{-2}\,{\lambda_{r,l}^*}^2 - P_{r,max}\bigg) = 0, && (60)\\ &-\alpha_{2,l}^*\Big(\lambda_{r,l}^* - \lambda_{r,max}\Big) = 0, && (61)\\ &-\alpha_{3,l}^*\Big(-\lambda_{r,l}^*\Big) = 0. && (62)\end{aligned}$$

By simplifying (59), we obtain

$$\begin{aligned}\frac{2\,\lambda_{r,l}^*}{\big(2{\lambda_{r,l}^*}^2+1\big)\,\big({\lambda_{r,l}^*}^2+1\big)\ln 2} - 4\,\alpha_1^*\,\lambda_{r,l}^* \\ \times\sum_{n=1}^{K}|\boldsymbol{U}_2(l,n)|^2\,\sigma_{2,n}^{-2} - \alpha_{2,l}^*\,+\,\alpha_{3,l}^* = 0. \qquad (63)\end{aligned}$$

Applying the KKT conditions yields the optimal beamforming power as follow

$$\lambda_{r,l}^* = \begin{cases} 0 & \lambda_{r,l}^\dagger \le 0, (\alpha_1^*\,F\ln 2) > 0.25\\ \lambda_{r,l}^\dagger & 0 < \lambda_{r,l}^\dagger < \lambda_{r,max}\\ \lambda_{r,max} & \lambda_{r,l}^\dagger \ge \lambda_{r,max}\end{cases}$$

in which

$$\lambda_{r,l}^\dagger = \sqrt{\frac{1}{4}\Big(\sqrt{1+\frac{2}{\alpha_1^*\,F\ln 2}}-3\Big)}, \qquad (64)$$

where F is a constant that equals to $\sum_{n=1}^{K} |\boldsymbol{U}_2(l,n)|^2 \sigma_{2,n}^{-2}$, and the Lagrangian multiplier α_1^* can be obtained by replacing (43) into the first constraint of (40) when the equality holds: $\alpha_1^* = f_1(P_{r,max}, \boldsymbol{U}_2, \Sigma_2)$. Note that if the last constraint defined in the condition (62) is binding, i.e., if $\lambda_{r,l}^* = 0$, then $\alpha_1^* = \alpha_{2,l}^* = 0$ due to the complementary slackness conditions. Substituting these multipliers in (59) results in $\alpha_{3,l}^* = 0$. Also, replacing $\lambda_{r,l}^* = 0$ in the objective function of (39) results in a zero capacity rate, which is not desired. In the same way as in [55], [56], it can be considered that $\lambda_{r,l}^* = \lambda_{r,max}$ for the values of beamforming powers above than the maximum. In addition, the value of $\lambda_{r,l}^{\dagger}$ in (64) can be numerically obtained for different values of channel coefficients and the UAV's power limitation, i.e., $\boldsymbol{U}_2$, $\sigma_{2,n}^{-2}$, and $P_{r,max}$. However, using the Taylor series in (64) at $x = \frac{2}{\alpha_1^* F \ln 2}$ with negligible $O(x^2)$, one can further simplify the obtained $\lambda_{r,l}^{\dagger}$ as

$$\lambda_{r,l}^{\dagger 2} = \frac{1}{4 f_1 \sum_{n=1}^{K} |\boldsymbol{U}_2(l,n)|^2 \sigma_{2,n}^{-2} \ln 2} - 0.5, \qquad (65)$$

where $\lambda_{r,l}^{\dagger}$ is a positive number and f_1 is the above defined function of the UAV power and channel coefficients. Finally, it is worth pointing out that if the UAV uses only one antenna for communicating with its users, then the beamforming power for the antenna is set to be $\lambda_{r,max}$. □

REFERENCES

[1] M. Mozaffari, X. Lin, and S. Hayes, "Toward 6g with connected sky: Uavs and beyond," *IEEE Communications Magazine*, vol. 59, no. 12, pp. 74–80, 2021.

[2] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6g wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proceedings of the IEEE*, vol. 109, no. 7, pp. 1166–1199, 2021.

[3] A. S. Abdalla, K. Powell, V. Marojevic, and G. Geraci, "UAV-Assisted Attack Prevention, Detection, and Recovery of 5G Networks," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 40–47, 2020.

[4] A. S. Abdalla and V. Marojevic, "Communications Standards for Unmanned Aircraft Systems: The 3GPP Perspective and Research Drivers," *IEEE Commun. Standards Mag.*, vol. 5, no. 1, pp. 70–77, 2021.

[5] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan, "Survey on uav cellular communications: Practical aspects, standardization advancements, regulation, and security challenges," *IEEE Communications surveys & tutorials*, vol. 21, no. 4, pp. 3417–3442, 2019.

[6] H. V. Poor and R. F. Schaefer, "Wireless physical layer security," *Proceedings of the National Academy of Sciences*, vol. 114, no. 1, pp. 19–26, 2017.

[7] M. Yang, D. Guo, Y. Huang, T. Q. Duong, and B. Zhang, "Physical layer security with threshold-based multiuser scheduling in multi-antenna wireless networks," *IEEE Transactions on Communications*, vol. 64, no. 12, pp. 5189–5202, 2016.

[8] X. Jiang, P. Li, Y. Zou, B. Li, and R. Wang, "Physical layer security for cognitive multiuser networks with hardware impairments and channel estimation errors," *IEEE Transactions on Communications*, 2022.

[9] J. Bai, T. Dong, Q. Zhang, S. Wang, and N. Li, "Coordinated Beamforming and Artificial Noise in the Downlink Secure Multi-Cell MIMO Systems Under Imperfect CSI," *IEEE Wireless Communications Letters*, vol. 9, no. 7, pp. 1023–1026, 2020.

[10] M. A. Teeti, "Downlink Secrecy Rate of One-Bit Massive MIMO System With Active Eavesdropping," *IEEE Access*, vol. 8, pp. 37 821–37 842, 2020.

[11] C. Zhu, J. Xu, and J. Xue, "Secure MISO Beamforming Optimization with Randomly Distributed Eavesdroppers and Interferers," in *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2019, pp. 1–5.

[12] M. T. Mamaghani and Y. Hong, "Joint Trajectory and Power Allocation Design for Secure Artificial Noise Aided UAV Communications," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2850–2855, 2021.

[13] Y. Ning and R. Chen, "Secure UAV Relay Communication via Power Allocation and Trajectory Planning," *IEEE Systems Journal*, pp. 1–10, 2022.

[14] A. S. Abdalla, A. Behfarnia, and V. Marojevic, "Aerial Base Station Positioning and Power Control for Securing Communications: A Deep Q-Network Approach," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, 2022, pp. 2470–2475.

[15] G. Sun, J. Li, A. Wang, Q. Wu, Z. Sun, and Y. Liu, "Secure and Energy-Efficient UAV Relay Communications Exploiting Collaborative Beamforming," *IEEE Transactions on Communications*, pp. 1–1, 2022.

[16] Y. Zhang, Z. Mou, F. Gao, J. Jiang, R. Ding, and Z. Han, "UAV-Enabled Secure Communications by Multi-Agent Deep Reinforcement Learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 599–11 611, 2020.

[17] Z. Yin, M. Jia, N. Cheng, W. Wang, F. Lyu, Q. Guo, and X. Shen, "UAV-Assisted Physical Layer Security in Multi-Beam Satellite-Enabled Vehicle Communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 3, pp. 2739–2751, 2022.

[18] H. Wu, Y. Wen, J. Zhang, Z. Wei, N. Zhang, and X. Tao, "Energy-Efficient and Secure Air-to-Ground Communication With Jittering UAV," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 3954–3967, 2020.

[19] S. Sun, T. A. Thomas, T. S. Rappaport, H. Nguyen, I. Z. Kovacs, and I. Rodriguez, "Path Loss, Shadow Fading, and Line-of-Sight Probability Models for 5G Urban Macro-Cellular Scenarios," in *2015 IEEE Globecom Workshops (GC Wkshps)*, 2015, pp. 1–7.

[20] T. Jiang, J. Zhang, P. Tang, L. Tian, Y. Zheng, J. Dou, H. Asplund, L. Raschkowski, R. D'Errico, and T. Jämsä, "3GPP Standardized 5G Channel Model for IIoT Scenarios: A Survey," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 8799–8815, 2021.

[21] R. Zhang, C. C. Chai, and Y.-C. Liang, "Joint beamforming and power control for multiantenna relay broadcast channel with qos constraints," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 726–737, 2009.

[22] B. Li, S. Zhao, R. Zhang, and L. Yang, "Full-duplex UAV relaying for multiple user pairs," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4657–4667, 2020.

[23] B. Li, R. Zhang, and L. Yang, "Joint user scheduling and uav trajectory optimization for full-duplex uav relaying," in *ICC 2021-IEEE International Conference on Communications*. IEEE, 2021, pp. 1–6.

[24] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. D. Renzo, "Safeguarding 5G wireless communication networks using physical layer security," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, 2015.

[25] J. Wang, C. Jiang, Z. Wei, C. Pan, H. Zhang, and Y. Ren, "Joint UAV Hovering Altitude and Power Control for Space-Air-Ground IoT Networks," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1741–1753, 2019.

[26] V. Marojevic, R. M. Rao, S. Ha, and J. H. Reed, "Performance analysis of a mission-critical portable lte system in targeted rf interference," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1–5.

[27] R. Xu and D. Wunsch, "Survey of clustering algorithms," *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 645–678, 2005.

[28] J. Li, B. Ai, R. He, M. Yang, Z. Zhong, and Y. Hao, "A Cluster-Based Channel Model for Massive MIMO Communications in Indoor Hotspot Scenarios," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 3856–3870, 2019.

[29] A. Naveed and S. S. Kanhere, "Cluster-based channel assignment in multi-radio multi-channel wireless mesh networks," in *2009 IEEE 34th Conference on Local Computer Networks*, 2009, pp. 53–60.

[30] A. Ali *et al.*, "Channel Clustering and QoS Level Identification Scheme for Multi-Channel Cognitive Radio Networks," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 164–171, 2018.

[31] N. Czink *et al.*, "A Framework for Automatic Clustering of Parametric MIMO Channel Data Including Path Powers," in *IEEE Vehicular Technology Conference*, 2006, pp. 1–5.

[32] E. Alpaydin, *Introduction to machine learning*. MIT press, 2020.

[33] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. " O'Reilly Media, Inc.", 2022.

[34] L.-N. Tran *et al.*, "Fast Converging Algorithm for Weighted Sum Rate Maximization in Multicell MISO Downlink," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 872–875, 2012.

[35] Y. Jeon, C. Song, S. Maeng, M. Park, and I. Lee, "MMSE based two-stage beamforming for large-scale multi-user MISO systems," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2016, pp. 1–6.

[36] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, 2011.

[37] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. M. Cioffi, "Weighted Sum-Rate Maximization Using Weighted MMSE for MIMO-BC Beamforming Design," in *2009 IEEE International Conference on Communications*, 2009, pp. 1–6.

[38] A. S. Abdalla and V. Marojevic, "Securing Mobile Multiuser Transmissions with UAVs in the Presence of Multiple Eavesdroppers," *IEEE Transactions on Vehicular Technology*, pp. 1–1, 2021.

[39] Y. Zhang *et al.*, "UAV-Enabled Secure Communications by Multi-Agent Deep Reinforcement Learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 599–11 611, 2020.

[40] J. Li *et al.*, "Deep reinforcement learning based computation offloading and resource allocation for mec," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, 2018, pp. 1–6.

[41] C. E. Mariano and E. F. Morales, "Dql: A new updating strategy for reinforcement learning based on q-learning," in *Machine Learning: ECML 2001*, L. De Raedt and P. Flach, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 324–335.

[42] M. Andrews, C. Dibek, and K. Palyutina, "Evolution of Q Values for Deep Q Learning in Stable Baselines," 2020.

[43] B. Jang, M. Kim, G. Harerimana, and J. W. Kim, "Q-Learning Algorithms: A Comprehensive Classification and Applications," *IEEE Access*, vol. 7, pp. 133 653–133 667, 2019.

[44] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[45] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.

[46] K. Cumanan, Z. Ding, B. Sharif, G. Y. Tian, and K. K. Leung, "Secrecy rate optimizations for a MIMO secrecy channel with a multiple-antenna eavesdropper," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 4, pp. 1678–1690, 2013.

[47] P.-H. Lin, S.-H. Lai, S.-C. Lin, and H.-J. Su, "On secrecy rate of the generalized artificial-noise assisted secure beamforming for wiretap channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1728–1740, 2013.

[48] M. T. Mamaghani and Y. Hong, "Intelligent trajectory design for secure full-duplex mimo-uav relaying against active eavesdroppers: A model-free reinforcement learning approach," *IEEE Access*, vol. 9, pp. 4447–4465, 2020.

[49] H. Kang, J. Joung, J. Ahn, and J. Kang, "Secrecy-aware altitude optimization for quasi-static uav base station without eavesdropper location information," *IEEE Communications Letters*, vol. 23, no. 5, pp. 851–854, 2019.

[50] D. Wang, B. Bai, G. Zhang, and Z. Han, "Optimal placement of low-altitude aerial base station for securing communications," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 869–872, 2019.

[51] Y. Zhang, Z. Mou, F. Gao, J. Jiang, R. Ding, and Z. Han, "Uav-enabled secure communications by multi-agent deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 599–11 611, 2020.

[52] X. Lin *et al.*, "The sky is not the limit: LTE for unmanned aerial vehicles," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 204–210, 2018.

[53] A. S. Abdalla, A. Yingst, K. Powell, A. Gelonch-Bosch, and V. Marojevic, "Open Source Software Radio Platform for Research on Cellular Networked UAVs: It Works!" *IEEE Communications Magazine*, vol. 60, no. 2, pp. 60–66, 2022.

[54] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.

[55] Y. Su, X. Lu, L. Huang, X. Du, and M. Guizani, "Tac-u: A traffic balancing scheme over licensed and unlicensed bands for tactile internet," *Future Generation Computer Systems*, vol. 97, pp. 41–49, 2019.

[56] Y. Su *et al.*, "Cooperative relaying and power control for uav-assisted vehicular networks with deep q-network," in *2021 IEEE/CIC International Conference on Communications in China (ICCC)*, 2021, pp. 318–323.

Aly Sabri Abdalla received the Ph.D. degree in Electrical and Computer Engineering at Mississippi State University, MS, USA in 2023. He received the B.S. and M.S. degrees in Electronics and Communications Engineering from the Arab Academy for Science Technology and Maritime Transport, Egypt, in 2014 and 2019, respectively.

Since 2019 he has been a Research Assistant in the Department of Electrical and Computer Engineering at Mississippi State University. His research interests include wireless communication and networking, software radio, spectrum sharing, wireless testbeds and testing, and wireless security with application to mission-critical communications, open radio access network (O-RAN), unmanned aerial vehicles (UAVs), and reconfigurable intelligent surfaces (RISs). He is serving as a student member of the IEEE Vehicular Technology Society's Ad Hoc Committee on Drones and IEEE 1920.1: Aerial Communications and Networking.

Ali Behfarnia received the Ph.D. degree in electrical engineering and computer science from Wichita State University, KS, USA in 2020. He is currently serving as an Assistant Professor in the Department of Engineering with the University of Tennessee at Martin. His research interests include wireless communication systems, game theory, error correction coding, cyber-physical systems, and the applications of machine and deep learning in communications over wireless networks.

Vuk Marojevic (Senior Member, IEEE) received the M.S. degree from the University of Hanover, Germany, in 2003, and the Ph.D. degree from Barcelona Tech-UPC, Spain, in 2009, all in electrical engineering.

He is currently an Associate Professor with the Department of Electrical and Computer Engineering, Mississippi State University. His research interests include 4G/5G security, spectrum sharing, software radios, testbeds, resource management, and vehicular and aerial communications technologies and systems.

Prof. Marojevic is an Editor of the IEEE Transactions on Vehicular Technology, an Associate Editor of IEEE Vehicular Technology Magazine, and an Officer of the IEEE ComSoc Aerial Communications Emerging Technology Initiative.