International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys.13, No.1 (May-23)

http://dx.doi.org/10.12785/ijcds/1301109

Optimal Coverage Enhancement for Multiple UAVs Using Multi-agent Learning Technique

Raymond Sabogu-Sumah[1,2], Kwasi Adu-Boahen Opare[1], James Dzisi Gadze[1], Jerry John Kponyo[1], Ahmed Abdul-Rahman[1], Ezer Osei Yeboah-Boateng[2] and Edmund Yirenkyi Fianko[2]

[1]*Department of Telecommunications Engineering,Kwame Nkrumah University of Science and Technology, Kumasi, Ghana*
[2]*National Communications Authority, Accra, Ghana*

Received 31 Aug. 2022, Revised 2 Apr. 2023, Accepted 23 Apr. 2023, Published 30 May. 2023

Abstract: Robustness of terrestrial cellular wireless networks becomes challenging in times of disasters such as earthquakes. This paper studies the deployment of multiple Unmanned Aerial Vehicles (UAVs) above the earth's surface to provide ubiquitous connectivity to under-laid users on earth. We provide an analytical framework using tools from stochastic geometry to model the UAV-user equipment (UE) network. We specifically model the UAVs in a finite three dimensional (3-D) space with their associated UEs as the marks on a two dimensional (2-D) earth surface. Tractable expressions for the UE's received signal strength and signal-to-interference plus noise ratio (SINR) are derived in Nakagami fading environments. A new paradigm in the study of UAV cellular communication is also developed in this work with a multi-agent learning technique. With this technique, the UAVs learn from each other by communicating, as well as interacting with their environment to provide optimal coverage. Our numerical results show that our method drastically reduces the interference from adjacent UAVs leading to improved coverage in terms of SINR values. Also, the results show that, UAV deployed wireless network provides better coverage compared to conventional terrestrial base station (BS) deployment.

Keywords: Unmanned Aerial Vehicle, Stochastic Geometry, Marked Point Process, Multi-agent learning, Q-learning, Reward, Equilibrium, Coverage, Altitude, SINR.

1. Introduction

Ubiquitous wireless connectivity has become imperative in the wake of proliferation of billions of internet-of-things (IoT) devices. One effective way of achieving this global connectivity is by the use of flying base stations installed on unmanned aerial vehicles (UAVs). The use of UAVs for providing wireless communications is also advantageous especially in search and rescue operations or when massive connectivity is required at public places during events.

UAV deployment has been studied in [1]-[10] to provide wireless connectivity for users on the surface of the earth. A trajectory optimization scheme was proposed [1] for the deployment of a single UAV to provide coverage for ground-based users. Mohammed et al [2] presented the use of mobile UAVs for energy efficient IoT communications with minimum transmit power. One advantage pointed out is the existence of dominant line-of-sight as compared to terrestrial base stations due to the high altitude at which the UAVs are positioned [2]. However, continuous movement of the UAVs leads to high power consumption. The authors of [3] applied evolutionary algorithms to find the optimal placement of low altitude platforms (LAPs) where overlapping LAPs coverage areas were allowed by using inter-cell interference coordination (ICIC). ICIC however incurs further communication overhead. The optimization of UAVs altitude to maximize coverage for ground users was proposed in [4]. In [5] and [6] the problem of optimal placement of UAVs with slightly different objectives was studied. The opportunities and challenges associated with the usage of UAVs to provide wireless connectivity were analyzed in [7]. Chetlar in [8] analyzed 3-D finite deployment of UAVs as a Binomial point process (BPP). The authors of [9] and [10] similarly studied the optimization of ground-to-air uplink communication for a single UAV. In [11], the authors studied the performance of WiMax systems deployed through high altitude platforms where the performance was shown to depend on some parameters such as the elevation angle and user traffic type. No mechanism for multi-UAV learning was implemented in any of the aforementioned studies.

An effective mathematical analysis of cellular wireless networks and for that matter BSs and UEs at random locations is by the use of tools from spatial stochastic processes. This approach has been extensively studied in literature [12]-[15]. It has generally been shown that the Poisson point process (PPP) is efficient in modeling various

E-mail address:rsabogusuma@yahoo.com, opare@knust.edu.gh, jdgadze@gmail.com

forms of network configurations such as femto cells [16]-[17] or ad hoc networks [18] or both [19]. In practical networks, the spatial deployment is normally different from being regular [19]. To address this problem, a suitable model was proposed in [20]-[21], in which the locations of BS and UE were modeled as homogeneous PPPs [22]-[23]. All the above-mentioned studies considered terrestrial networks with the exception of [8] which analyzed the UAV BS deployment as a BPP.

Multi-agent learning has been used in various applications for solving problems of a distributed nature [23]-[30]. Its usage in terrestrial based wireless communications has also been vast. For example, in [25], multi-agent reinforcement learning (MARL) was used to model inter-cell interference where inter-cell interference coordination mechanisms were proposed. In the area of cognitive radio, MARL has been widely used for competitive and cooperative spectrum access [26]-[27]. MARL applications in UAV cellular communication has received relatively less attention. A multi-agent path planning for UAV threat avoidance was studied in [29] where a path planner searches for the path; this was however, not for the purpose of wireless communication. In [30], a distributed cooperative UAV network for small drones was proposed and implemented to provide wireless coverage. However in [30] the analysis was done qualitatively as no mathematical model was developed for the UAVs cooperation.

There are national rules and regulations on owning and flying UAVs. The China's Civil Aviation Administration (CCAA) and Federal Aviation Administration (FAA) of the United States recommend that the UAV remains in the visual line-of-sight (VLOS) of the remote pilot [29]. The pilot must hold a certification to fly the drone too. With this backdrop, this work develops a multi-UAV learning technique for cellular communication. The novelty and contributions of this work are:

- By modeling the system as a Marked Point Process, we leverage tools from Stochastic Geometry to analyze the distribution of the UAVs in 3-D space and the users on ground (2-D). We derive new tractable analytical expressions to quantify the UEs received signal power, SINR values and coverage probability.
- Secondly, in order to mitigate the problem of harmful interference in UE's served by the multiple UAVs, we propose a multiagent learning technique with a game theoretic framework where the UAVs are the learning agents. They fully cooperate with each other by communicating their current states, actions and rewards with other agents in their neighborhood. This helps in achieving the common goal of obtaining optimal coverage.

The organization of the rest of this paper is as follows: our system setup as a marked point process is presented in Section II. In Section III, we present our novel learning technique for coverage enhancement. The results of the study are presented in Section IV while Section V concludes the paper.

2. Stochastic Geometry Analysis of UAVs in 3-D space and UEs on Ground (2D)

In Fig.1, the geometry of the UAVs in 3-D space and UEs on the ground is shown. The use of stochastic geometry for analyzing UAVs in 3-D space was reported in [8]. In this section, we model our system as a marked point process (MPP) where the UAV equipped BSs are our transmitters (points) and the ground-based UEs are the marks. By leveraging the Matern hard core model (MHC) we optimally position the UAVs to avert undue overlap for the marks on ground. The received signal strength, user SINR, coverage and capacity expressions are further derived.

A. System Analysis as a Marked Point Process

The UAVs are placed in a d dimensional space $\mathbb{R}^d$, (for our case $d = 3$) as the points x_i (BS transmitters) and the marks m_u (UEs) on an l dimensional space $\mathbb{R}^l$, ($l = 2$). The index i ranges from 1 to N number of points (UAVs) and u ranges from 1 to K number of UEs associated with each point. For simplicity, we assume all the marks are located on the ground. $\mathbb{R}^l$ is a random vector containing the locations of the marks (UEs) associated with each point (UAV). The UAVs in space are analyzed as a simple Binomial Point Process (BPP) [8] with intensity Λ. The UAVs-UEs marked point process is then modeled as a marked point process $\bar{\Phi}$ having points in $\mathbb{R}^d$ and marks in $\mathbb{R}^l$ denoted as

$$\bar{\Phi} = \sum_{i}^{N} \sum_{u}^{K} \epsilon(x_i, m_u), \tag{1}$$

with $\epsilon(.)$ as the delta dirac function, x_i as the collection of the UAV transmitters and m_i are the marks (UEs). Consider bounded spaces S_1 and S_2 in $\mathbb{R}^d$ and $\mathbb{R}^l$ respectively and with the assumption that the MPP is independently

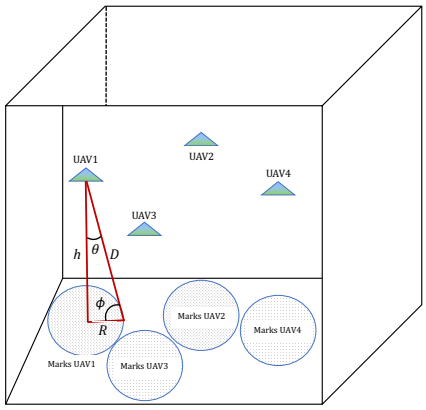


Figure 1. System Model as a Marked Point Process.

marked, the MPP has an intensity and mean measure given respectively as

$$\bar{\Lambda}(S_1 X S_2) = \int_{S_1} \bar{p}(x,k)\Lambda(dx), \tag{2}$$

$$\mathbf{E}[\bar{\Phi}(S_1 X S_2)] = \int_{S_1} F_x(S_2)M(dx), \tag{3}$$

where $\bar{p}(x,k) = \int_{S_2} F_x(d_m)$ is the distribution of the marks m_u in $\mathbb{R}^l$ and $M(S_1) = \mathbf{E}[\Phi(S_1)]$ is the average measure of the points (Φ) of $\bar{\Phi}$.
The Laplace functional of the marked point process with intensity measure $\bar{\Lambda}$ is derived as similarly done in [12] as

$$\mathcal{L}_{\bar{\Phi}}(f) = e^{-\int_{(S_1 X S_2)}(1-e^{-f(x)})\bar{\Lambda}(dx)}, \tag{4}$$

where f runs over a set of non-negative functions on $\mathbb{R}^d$ and $\mathbb{R}^l$.

B. Optimal UAV Positioning

In the usage of multiple low or high altitude platforms (LAPs or HAPs) for providing wireless connectivity to ground users, it is important that they are optimally positioned in space to avoid causing harmful interference to users covered by adjacent cells. Hard core models are generally used to ensure that points are not close to each other less than some given range in spatial geometry. The Matern hard core (MHC) model is applied in this study to arrange the points (UAVs) h_{HCM} distance apart. For equal UAV BS transmit power and similar air-to-ground channel characteristics, the marks (UEs) associated with UAVs are within equal coverage radii. The UAV packing using the MHC model ensures minimal if not zero coverage overlap. Integrating the MHC point process with our discussion of a MPP, the marked MHC point process is defined as

$$\bar{\Phi}_{MHC} = \sum_{i}^{N}\sum_{u}^{K} \epsilon(x_i(U_i, m_u)), \tag{5}$$

where U_i is interpreted as the age of point x_i and its marks m_u. The volume of space occupied by each UAV is

$$v_d = \sqrt{\pi^d}/\Gamma(1 + d/2), \tag{6}$$

with Γ is a gamma distribution function and d is the dimension of the space. The MHC point process parks the points of radius $h_{HCM}/2$ with volume fraction [12]

$$v_f = \frac{1}{v_d h_{HCM}^d} v_d (h_{HCM}/2)^d = \frac{1}{2^d}. \tag{7}$$

(7) represents the lower bound for saturated packing while an upper bound for any hard sphere packing is provided by [12] as:

$$v_{f_u} = (d+1)2^{-d/2}, d \geq 1 \tag{8}$$

For a given volume in space, expressions (7) and (9) can be used to compute the lower and upper bound of the number of UAVs to cover a given area on earth. Each UAV uses the HCM point process to create an optimum exclusion region around itself and therefore prevent interference to its UEs located on the ground. The value of this exclusion region will be used in the sequel when we introduce multi-agent learning technique between the UAVs in Section III.

C. UE SINR analysis

The probability density function (PDF) of the distance between a given UE and its serving UAV is derived as [31]:

$$f_{d_k}(D) = 4\pi\Lambda D^2 \exp(-\frac{4}{3}\pi\Lambda D^3), D \geq 0, \tag{9}$$

where D is the distance between the UE and its serving UAV calculated as $D = \sqrt{h^2 + r^2}$ using Fig.1. Likewise, the PDF of the UE distances r in a single UAV coverage radius (R) is given as:

$$f_r(r) = \frac{2r}{R^2}, 0 \leq r \leq R, \tag{10}$$

We assume the signals are received by omnidirectional antennas at the UEs. The received signal at the i^{th} UE located at a point y_i on Earth is given as:

$$P_r(D) = P_{r\bar{\Phi}}(D) = \sum_{(x_i,p_i)\in\bar{\Phi}} \frac{p_i 10^{A/10}}{l'(|y_i - x_i|)^\alpha X_\sigma}, y_i \in \mathbb{R}^l, x_i \in \mathbb{R}^d. \tag{11}$$

where p_i is the UAV BS transmit power, α is an environment dependent path loss exponent, A denotes the antenna gain, X_σ is the shadowing figure which is Log-normally distributed and $l'(|y_i - x_i| = D)$ is the path loss between the UAV and UE given as [32]:

$$\begin{aligned} l'(D) &= P_D^l (YD)^\alpha \rho_1 + P_D^n (YD)^\alpha \rho_2 \\ &= \frac{1}{1 + \xi\exp(\tau(\theta - \xi))}(YD)^\alpha \rho_1 + \\ &\Big(1 - \frac{1}{1 + \xi\exp(\tau(\theta - \xi))}\Big)(YD)^\alpha \rho_2. \end{aligned} \tag{12}$$

In (12), the values of ξ and τ are environmental and frequency dependent, θ is the angle from the UAV to the UE located r meters away from the origin as shown in Fig.1. Y is a constant defined as $Y = \frac{4\pi f_c}{c}$ with f_c and c as the carrier frequency and the speed of light respectively. Also, α is the path loss exponent, ρ_1 and ρ_2 are extra path loss components for line-of-sight (LOS) and non-LOS (NLOS) respectively. We assume the UAVs use beamforming techniques to transmit signals to UEs. The antenna gain in that case is given as

$$\begin{aligned} A(\phi,\theta) &= G_{max} + A_{BS}(\phi,\theta), \\ A_{BS}(\phi,\theta) &= \Big[- min\Big(12\Big(\frac{\phi}{\phi_{3dB}}\Big)^2, A_\phi\Big) - min\Big(12\Big(\frac{\theta}{\theta_{3dB}}\Big)^2, A_\theta\Big)\Big], \end{aligned} \tag{13}$$

where G_{max} is the maximum BS antenna gain, ϕ_{3dB}, θ_{3dB} are the 3dB beamdwidths, A_ϕ and A_θ are the maximum attenuations in the horizontal and vertical domains respectively. Also ϕ and θ (with abuse of notation) are defined as the angles between the directions of interest and the

boresight of the antenna in the horizontal and vertical domains respectively.

Expression (11) assumes the UE receives signals from all the UAVs which is however, not the case due to their (UAVs) optimal planning and positioning. The signals received from adjacent UAVs farther away from the serving UAV are considered negligible using the concept of extremal noise in [12] (11) then becomes

$$P'_{r,i}(D) = P'_{r,i\bar{\Phi}}(D) = \frac{p_i 10^{A/10}}{l'(|y_i - x_i|)^{\alpha} X_{\sigma}}. \tag{14}$$

There is limited knowledge on air-to-ground channel models for LAPs [8] and this actuated our choice of Nakagami fading in the UE-to-UAV links. Nakagami models give general view of various fading environments depending on the parameters chosen. Incorporating the effect of fading into (14), the received signal strength becomes,

$$P'_{r,i}(D) = P'_{r,i\bar{\Phi}}(D) = F_i \frac{p_i 10^{A/10}}{l'(|y_i - x_i|)^{\alpha} X_{\sigma}}, \tag{15}$$

where F_i is the Nakagami fading gain which follows a Gamma distribution with PDF given as [22]

$$f_F(f) = \frac{m^m f^{m-1}}{\Gamma(m)} \exp(-mf), \tag{16}$$

where m is a parameter which determines the shape of the distribution and Γ is the gamma distribution.

$$m = \begin{cases} 1 & \text{Rayleigh fading,} \\ 2(X+1)/(2X+1) & \text{Rician fading,} \\ \infty & \text{no fading,} \end{cases} \tag{17}$$

where X is the ratio of the LOS to the NLOS components of the signal. Thus, the Nakagami distribution can model Rayleigh and Rician distributions, as well as more general ones. The PDF of the UE received signal is derived as

$$\begin{aligned} f_{P_{r,i}}(D) &= \int_{\pi R^2} P'_{r\bar{\Phi}}(D) f_F(f) \\ &= \int_{\pi R^2} \frac{p_i 10^{A/10}}{l'(|y_i - x_i|)^{\alpha} X_{\sigma}} \frac{m^m f^{m-1}}{\Gamma(m)} \exp(-mf) dD, \end{aligned} \tag{18}$$

πR^2 is the coverage area of a UAV with R as the radius. We use the UE SINR value as an appropriate metric of the system performance. It is denoted by ς and computed as

$$\varsigma = F_i \frac{p_i 10^{A/10}}{l'(|y_i - x_i|)^{\alpha} X_{\sigma}} \Big(\sum_{j \in N \setminus i} F_j p_j D_j^{-\alpha} 10^{A_j/10} X_{\sigma_j} + \sigma^2 \Big)^{-1}. \tag{19}$$

In (19), $j \in N \setminus i$ is a set operation which indicates the summation over all N UAVs except the i^{th} one which in this case is the serving UAV. F_j, p_j, X_{σ_j}, A_j, σ^2, and D_j are the channel fading parameter, interfering UAV transmit power, shadowing figure, interferer antenna gain, noise power and the distance between the interfering UAV and the UE respectively. The rest of the parameters are defined previously in (15). Note also that $\sum_{j \in N \setminus i} F_j p_j D_j^{-\alpha} 10^{A_j/10} X_{\sigma_j}$ is the aggregate interference (I_{agg}) on a given UE from multiple UAVs with mean and variance given as $\mu_{I_{agg}}$ and $\sigma_{I_{agg}}$ respectively. The lognormal shadowing figure X_{σ} introduces some randomness in to the interference which makes the cumulative interference random. This gives rise to theorem 1.

Theorem 1: If I_1 and I_2 are two independent random interference from two UAVs, the sum $I_{agg} = I_1 + I_2$ is also random.

The CDF of I_{agg} is similarly approximated as [33]

$$F_{I_{agg}}(z) = \frac{1}{2}\Big(1 + \text{erf}\Big(\frac{\ln(z) - \mu_{I_{agg}}}{\sqrt{2\sigma_{I_{agg}}}}\Big)\Big), \tag{20}$$

D. Downlink Coverage Probability

We start this subsection by outlining two fundamental necessary and sufficient conditions for the analysis of coverage probability in our study.

Proposition 1: $P'_{r,i}(D) \geq P'_{r,i}(D)_{thr}$ is a necessary condition for the UE to effectively operate.
$P_r(D)_{thr}$ is technically known as the receiver sensitivity which specifies the signal level below which the UE cannot intelligibly decode.

Proposition 2: The necessary and sufficient conditions for optimal coverage and efficient UE operation are:
1) $P'_{r,i}(D) \geq P'_{r,i}(D)_{thr}$.
2) $\varsigma \geq \varsigma_{thr}$.
Note that ς_{thr} is the SINR threshold for excellent UE operation which takes into account the interference from other transmitters and noise in the channel as well. The second proposition defines a complete condition to characterize coverage probability. This study therefore resorts to proposition 2 as a more meaningful metric in characterizing coverage probability. The interference received at each UE from the interfering UAVs are i.i.d. and appealing on the Central Limit Theorem [8] the coverage probability is given in (22); Where $\bar{\varsigma}$ and σ_{ς} are the mean and standard deviation of the UE SINR values. The mean UE SINR value is $\bar{\varsigma} = \mathbb{E}(\varsigma)$. See proof in Appendix B.

σ_{ς} is similarly computed by first deriving the variance and taking the square root of it. An interested reader is referred to [34].

Due to the non existence of a close form solution for the integral in (22), its solution can be evaluated using the Q-function

$$Q(z) = \frac{1}{2\pi} \int_z^{\infty} \exp\left(\frac{-x^2}{2}\right) dx \tag{21}$$

In our case the Q-function takes the form $Q\left(\frac{\varsigma_{thr} - \bar{\varsigma}}{\sigma_{\varsigma}}\right)$.

$$P_c = P\Big(\Big(F_i \frac{p_i 10^{A/10}}{l'(|y_i - x_i|)^{\alpha} X_{\sigma_i}} \Big(\sum_{j \in N \setminus i} F_j p_j D_j^{-\alpha} 10^{A_j/10} X_{\sigma_j} + \sigma^2 \Big)^{-1} \Big) > \varsigma_{thr} \Big)$$

$$= \int_{\varsigma_{thr}}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_{\varsigma}} \exp\left[-\frac{1}{2} \left(\frac{\Big(F_i \frac{p_i 10^{A/10}}{l'(|y_i - x_i|)^{\alpha} X_{\sigma_i}} \Big(\sum_{j \in N \setminus i} F_j p_j D_j^{-\alpha} 10^{A_j/10} X_{\sigma_j} + \sigma^2 \Big)^{-1} \Big) - \bar{\varsigma}}{\sigma_{\varsigma}} \right)^2 \right] d\varsigma, \tag{22}$$

E. Capacity and System Sum Rate Computation

The performance of the UAV empowered cellular network is analyzed here. Having derived expressions to quantify SINR of the UEs, we can as well derive with ease, the capacity and system sum rate expressions based on the Shannon Hartley formula. Assuming our operational bandwidth is denoted as W, the capacity is

$$C = W \log_2 \big(1 + \varsigma\big), \tag{23}$$

where all the parameters are the same as in (19). The system sum rate is overall rate of all users covered by a given UAV. By this definition, we formulate expressions (24) to compute sum rate of the i^{th} UAV covering UEs of intensity λ, in a radius R.

$$R = 2\pi\lambda W \log_2 \big(1 + \varsigma\big) P_c. \tag{24}$$

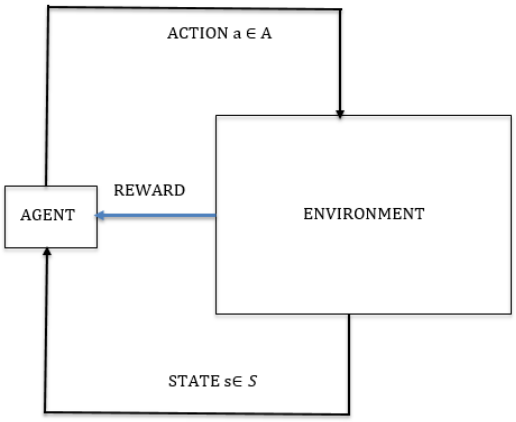
3. Coverage Enhancement with multi-agent learning Technique

In this section we consider a multi UAV cooperative system where each UAV (hereafter, an agent) is a learning entity. We first present a brief description of what a multiagent system is. According to [30] an agent is an entity which observes, interacts as well as learn from/with its environment in a stochastic fashion mostly described as a Markov decision process (MDP) to earn some reward. For a multiagent system, the agents interact with their environment and also among themselves, take actions and gain some rewards in the process. Agents can also teach others in a multiagent system [30] . In both scenarios, the main object is to get optimal reward also known as payoff.

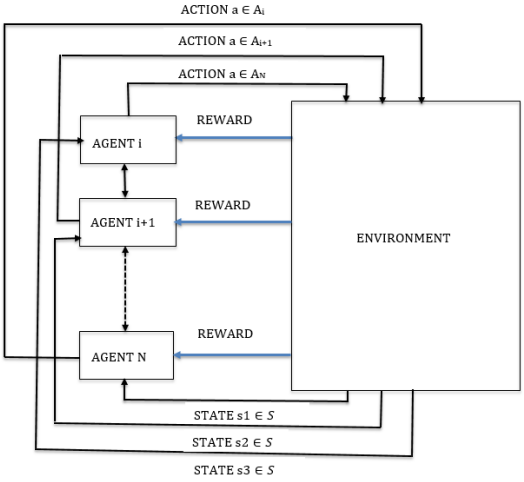
We show in Fig.2a, a pictorial illustration of a single and multiagent system. The single agent interacts with its environment and make necessary adjustments in terms of action selection to increase its rewards. In the multiagent system shown in Fig.2b, the agents not only interact with each other or their environment, but also share their states as well as rewards for effective coordination.

A. Communication in a Cooperative Multiagent Learning Environment

Communication between agents in a cooperative multiagent system is vital. In [35] the authors proposed a motion gesture based communication between the interacting agents and this could be applied to a multi-UAV embedded with cameras. This technique is however not suitable for



(a) Single Agent System



(b) Multiagent System

Figure 2. Description of Single and Multiagent Systems

our case due to the high separations between the UAVs. This will lead to ineffective image or video surveillance of each others' actions.

In our proposed learning scheme explicated in the sequel, agents need to communicate between themselves in order to achieve an optimal global payoff. We develop an optimal communication mechanism in this subsection for efficient information sharing. Agents communicate by broadcasting messages containing their states, actions and rewards to each other in space. Due to the general power constraints in UAV empowered cellular communication, we assume transmit power of these broadcast messages to be very low (0.1% of the agent transmit power to UEs). This assumption stems from the fact that the primary goal of an agent is to cover users on ground and substantial percentage of its limited transmit power needs to be used for that purpose.

UAV multicast communication: Each agent broadcasts signals at successive intervals to the entire network in space. The broadcast signal contains information about their positions (x_i', y_i, h_i), where h_i represents the heights of the UAVs from ground. The received signal from an i^{th} agent to the rest is $10^{-3} p_i l' |x|^{-\alpha}$, with $|x|$ as the separation between the agents. For simplicity, we assume free space pathloss between agent broadcast message transmission.

The successive time intervals between agents broadcast message transmission may however lead to a delay in communication. For example, if each agent takes 2 second to transmit and there are 50 UAVs in the network, an agent will take as long as 100 seconds to update the others about its location and other important information for learning. For optimal coverage and avoidance of coverage overlap, the agents separation in space must be equal to or greater than the sum of the coverage radius of adjacent agents [6]. The latter statement implies that the separation between the agents in space will be large for high coverage radius, and the probability distribution of guaranteed reception of each agent's broadcast message by the others is derived as

$$f_X(x) = \begin{cases} 1 & x < \kappa x_{max}, \\ \frac{1}{(2x/x_{max})^2} & x \geq \kappa x_{max}, \end{cases} \tag{25}$$

where κ is an arbitrary value representing the percentage of maximum distance that will result in effective communication and x_{max} is the maximum radius of the entire agents in space.

Fig.3 depicts the probability of guaranteed communication between agents with an increase in their distances apart. It is clear from the plot that, the probability of reception of the broadcast messages decreases with an increase in x which implies that effective communication would be compromised. Besides the ineffective communication associated with large separation between agents, developing a very effective mechanism to coordinate agents behavior in a dynamic environment is hard as the number of agents grow [36].

Nearest neighbor communication: Due to the challenge of multicast inter-agent communication, this study resorts to nearest neighbor communication in which the agents are divided in clusters consisting of few agents adjacent to each other.

Definition 1 : A point $x_{i\pm1}$ is in the neighborhood of x_i if and only if

$$|x_i - x_{i\pm1}| \leq h_{HCM} \tag{26}$$

The above equation uses the hardcore point packing radius devised in Section II as the upper bound in defining the neighborhood of points (agents).

B. Stochastic Game Theoretic Multiagent Learning

We model the multi-agent learning process as a stochastic game with $< S, A...A_N, p, r_1...r_N >$ tuple. where N is the number of agents in the system, S is the set of environment states, A_i' is the set of agent i available actions, p is a conditional probability function determining the probability of transiting to the next state and r_i is the local immediate reward function of agent i.

Definition 2: In this study, we define the reward for an agent as the percentage of its covered UEs having an SINR greater than the predefined threshold ($\%\varsigma > \varsigma_{thr}$).

$$r(t) = \%\varsigma > \varsigma_{thr} \tag{27}$$

This implies that the reward is a function of the UEs SINR values. While each agent tries to maximize its payoff also known as the reward, its effect on the others is kept in mind because the prime aim is to solve a common universal problem. The immediate global reward is therefore the culmination of the individual rewards of the agents. It is given as

$$r = \sum_{i=1}^{N} r_i \tag{28}$$

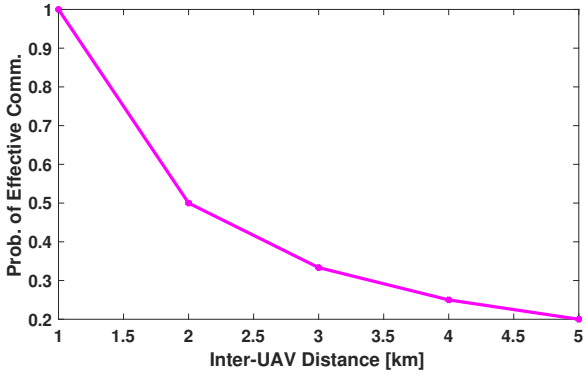


Figure 3. Probability of Guaranteed Inter-agent Communication

The rewards of an agent at times $(t+1), (t+2), (t+3)...(t+V)$ are denoted as $r(t+1), \gamma r(t+2), \gamma^2 r(t+3) + ...\gamma^{V-1} r(t+V)$ respectively. The time rewards constitute a time series and applying techniques from series and sequences, the discounted global reward is calculated as

$$R(t) = r(t+1) + \gamma r(t+2) + \gamma^2 r(t+3) + ...\gamma^{V-1} r(t+V) = \sum_{v=0}^{V} \gamma^v r(t+v+1), \tag{29}$$

γ is the discounted factor (0,1), V is the ending time of an episode and $r(t+1)$ is the global immediate reward received at time $t+1$. The discounted factor determines the importance given to future rewards, thus a 0 discounted factor will make the agent consider only current rewards and on the other hand a factor approaching 1 will make it strive for a long-term high reward. For a stochastic game with $N > 1$ agents as considered in this study, each agent chooses an action or strategy Δ_i from its action set A_i'. The set of possible actions available for an agent are to change its position by either moving up or down (Δh), move right or left $(\Delta r, \Delta l)$ or move forward or backward $(\Delta b, \Delta f)$. Thus,

$$A_i' = \left\{\Delta h, \Delta r, \Delta l, \Delta b, \Delta f\right\} \tag{30}$$

In other applications of multiagent learning specifically for cognitive radio applications, agents can take an action by altering their transmit power [26], we could model our stochastic game by including that as one of the actions. However, due to the power constraints of UAV empowered communication, we leave out power alteration but rather focus on optimal agent orientation to provide excellent coverage. When agent i takes an action Δ_i from A_i', its state s_t is transformed to the next state say s_{t+1} and the agent receives a payoff $r(t)$. The agent broadcasts its selected action and reward to other agents in its neighborhood using definition 3. We assume the UEs communicate in the uplink by sending signals containing their SINR values and their channel state information (CSI) to their serving agents. The agents are therefore able to determine the quality of the actions they take.

This study applies the Q-learning which is a type of reinforcement learning used in both single and multiagent learning systems. It is applied as a Markov decision process (MDP) in the single agent learning paradigm where the agent learns by transiting from state to state. For a multiagent learning scheme as considered here, it is applied as a muti MDP, where all actions transit from one state to another while learning from each other. The Q-value denotes how good or bad an action is. The action selection procedure for an i^{th} agent follows the Boltzmann [26] exploration scheme as:

$$p_i(t) = \frac{e^{Q_i(t)}/T}{\sum_{k \in A_i'} e^{Q_k(t)/T}}, \tag{31}$$

with $p_i(t)$ as the probability of selecting action Δ_i, $Q_i(t)$ is the Q-value of action Δ_i at time t, A_i' is the action set for agent i, and T is temperature parameter controlling the balance between exploration and exploitation. A higher T ensures equal opportunity for all actions to be selected whereas actions with high Q-values are favored with a small T. The agents environments are dynamic and thus, an optimal action at a particular time may change due to a change in the environmental conditions. To mitigate this challenge of remaining in the most rewarding state, the stochastic action selection strategy given in (48) provides an optimal solution. The Q-learning values of an agent are computed and updated according to [28]

$$V_i(s_{t+1}) := Q_i(s_{t+1}, \Delta^*), \Delta^* \in \text{argmax} \sum_{i=1}^{n} Q_i(s_{t+1}, \Delta^*)$$
$$Q_i(s, \Delta_i) := (1-\lambda) Q_i(s, \Delta_i) + \lambda[r_i + \gamma V_i(s_{t+1})], \tag{32}$$

where $0 < \lambda \leq 1$ is a parameter that determines the learning rate and γ is the discounted future reward. If λ is close to zero, the UAV tends to consider only instant rewards and the vice versa and Δ^* is the action with optimal payoff. The goal of each UAV is to get to a state with the highest reward. If it does, it may try to remain there forever which leads to the over exploitation of that state-action pair.

Definition 3 : An agent's action Δ_i is strictly dominant over other agent's if $Q(\Delta_i) > Q(\Delta_{-i})$. The payoffs associated with strictly dominated actions are optimal. Since all agents strive to maximize their utilities, this translates the learning game into a multi objective optimization problem. Solutions to this kind of problems are hard to find because while maximizing the utility for one agent, the utilities on the others deteriorate. Thanks to the coordination and learning between the agents proposed in this work, each agent has the best response to the strategies of the others. No two or more adjacent UAVs move towards each other, while one moves right, its adjacent neighbors takes any of the available actions apart from moving left to avoid undue coverage overlap. Also, no agent will have a strictly dominated strategy over the others.

In competitive multiagent games, Nash equilibrium is usually the optimal solution for all agents [28], however, the solution for our collaborative multiagent learning game is a Pareto optimal equilibrium, hence, no agent can improve its reward by deviating from the rules of the game. The payoff of agent j in response to agent i taking action Δ_i is thus given as

$$Q_j(\Delta_{-i}) \leq Q_i(\Delta_i). \tag{33}$$

In other words, when agent i takes action Δ_i, agent j respond with a joint action Δ_{-i} which has a payoff less than or equal to the payoff of agent i. At the equilibrium solution, the payoff of an agent in terms of its UE SINR values in (22) becomes

$$\varsigma = F_i \frac{p_i A}{l'(|y_i - x_i|)^{\alpha} X_{\sigma}} \left(\psi I_{agg}\right)^{-1}, \tag{34}$$

with ψ as the interference reduction factor from the multiagent learning and coordination scheme. Our multi-UAV learning algorithm is presented below.

Multi-UAV Learning ALGORITHM

1: for each agent action $\Delta_i \in A_i'$
2: Compute the reward from (45)
3: Broadcast the reward in step 2 to all agents in the neighborhood of agent i
4: **if** $\%R(\Delta_i, s) < R(\Delta_i, s)_{thr}$ is high,(i.e. most UEs SINR values are less than the threshold),
i) Agent i changes strategy by choosing actions from (47) using (48)
ii) Neighboring agents respond using the best response strategy **then**
5: Repeat step 2 to 3
6: **else**
7: Agents remain in optimal positions to provide optimal connectivity.
8: **end if**

4. Results and Analysis

We present our results in this section. For our Monte Carlo simulation, we randomly generated UEs within a coverage radius R for each UAV and computed their corresponding distances from their serving UAVs. The efficacy of our multiagent learning scheme is also shown.

A. Effect of Varying UAV Heights on Coverage

We analyzed the effect of varying the UAVs heights on the UEs SINR value with fixed coverage radius of R=1000m, m=1, N=10 UAVs and h ranging between 200m to 500m. As shown in the plots of Fig.4, with decreasing heights of the UAVs, we obtained improved UEs SINR values deployed within the coverage radius R of the serving UAVs. This correspondingly translates to better coverage probability for a given SINR threshold. Increasing the UAVs heights results in an increase in the interference from the others as similarly reported in [8]. However in [8], the SIR values are quite lower than our values because of the dynamic UAV deployment with multiagent learning technique implemented here. Also, a decrease in h results in a decrease in the the propagation distance which gives low path loss values. There is also a good agreement between our Monte Carlo simulation and analytical results.

TABLE I. General Simulation Parameters.

Parameter	Description	Value
p_i	UAV Transmit Power	30 dBm
A_i	UAV Maximum Antenna Gain	8 dB
f_c	Carrier Frequency	2 GHz
$\xi, \tau, \rho_1, \rho_2, \alpha$	Pathloss Parameters [2]	11.95,0.14,3,23 &2
X_σ	Mean Shadowing Figure	8 dB
W	Bandwidth	1 MHz
σ^2	Noise Power	-120 dBm

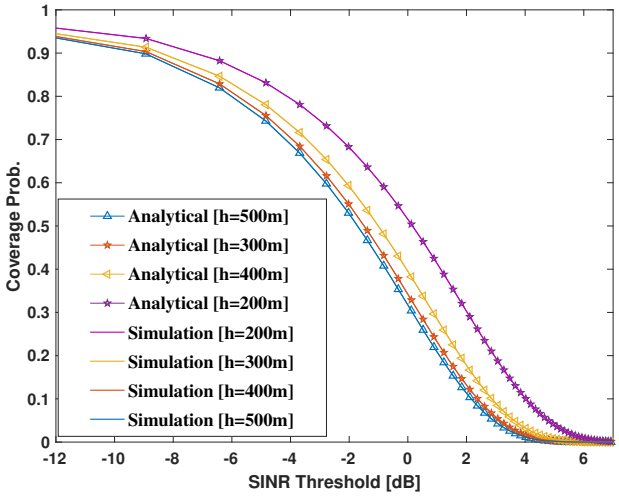


Figure 4. Coverage probability for Varying UAV Heights.

Cellular communication systems are uplink limited, which makes these results quite desirable as a very high UAV altitude could result in very high distance for UEs uplink communication. To ensure an optimal bidirectional communication, h must therefore be carefully selected.

B. Results of Multiagent Learning with Varying Coverage Radius

We used the percentage of UEs SINR values greater than a given threshold as our performance metric here. As observed in Fig.5, the percentage of UEs SINR values less than a given threshold decreases with an increase in the value of R which is quite expected because an increase in R results in a corresponding increase in the UAVs-to-UEs propagation distances. For a given coverage radius, our proposed multiagent learning approach gives better SINR performance compared to the static UAV deployment scheme especially at the tail end of the SINR curves (ie for cell edge users) which proves the efficacy of our approach. This better performance with regards to our proposed approach stems from the full cooperation and learning among the agents, that is the interference is well managed through the avoidance of undue coverage overlap.

C. SINR Versus Coverage Radius

We positioned the UAVs at an altitude of 400m above the ground. By varying the coverage radius from 0 to 1000m, we compute the UE SINR values. As depicted in Fig.6, the UE SINR values increases with an increase in R until a point is reached where it begins to decrease. The decrease arises as a result of the close proximity between

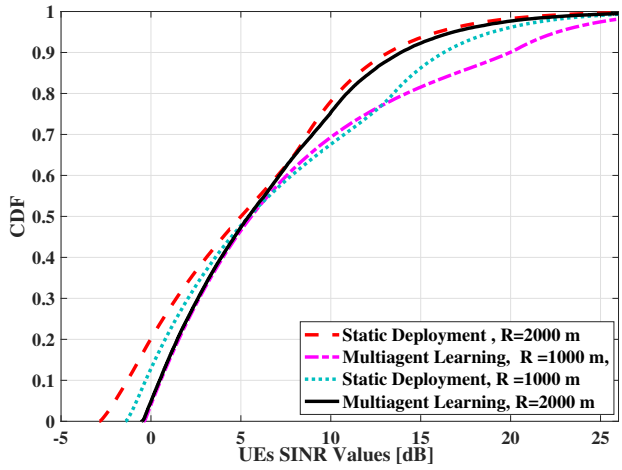


Figure 5. CDF plot for UEs SINR Values, with fixed Coverage Radii

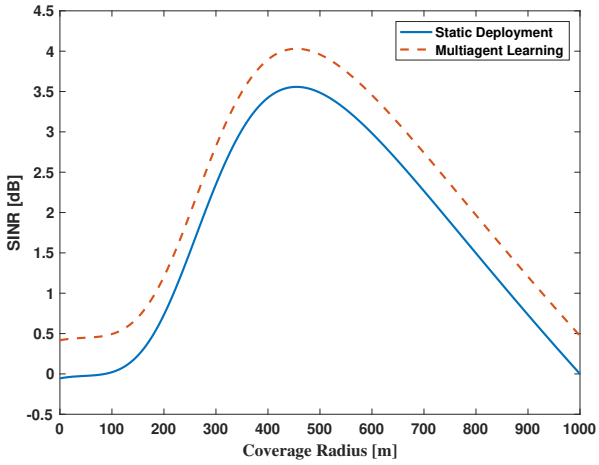


Figure 6. UE SINR versus variable coverage Radius.

the interferers and the UE. Also increasing R to some point results in a corresponding increase in d (with high path loss) for cell edge users. Finally, the impact of our multiagent learning is clearly shown with improved SINR values for any given coverage radius as compared to the static deployment studied in [8].

D. Coverage Probability for Different Fading Environments

Here, the effect of different fading environments on the coverage probability is shown. We fixed h=400m and R=1000m. It can be seen from Fig.7 that the coverage probability for a given SINR threshold improves with increasing m. An increase in m implies a high probability of LOS transmission which results in less signal fading. As mentioned in (17), a high m value leads to Rician or no fading which has high gain. This is clearly manifested in this results.

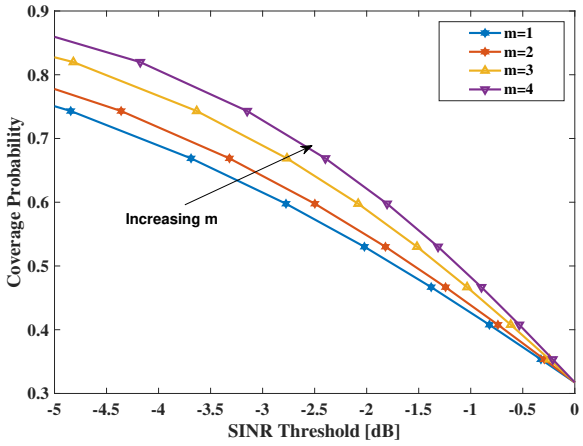


Figure 7. Coverage probability for Different Fading Parameter (m).

E. UE Capacity for Different Coverage Radius

In Fig.8 we show the UE down link data rate for static UAV and multiagent learning scenarios. We assume the UE is located at the boundary of the serving UAV coverage area and all other parameters are the same as stated above. As depicted in the figure, there is a performance improvement (about 20%) in the cell edge downlink rate with the multiagent learning technique proposed in this work. Another insight revealed in this scenario is the slight decrease in the downlink rate as we increase the coverage radius. This decrement arises from the increment in the UE-to-UAV propagation distance as the coverage radius increases. A higher coverage radius will ensure that more users are served but with lesser cell edge performance. An appropriate trade-off is therefore needed to ensure optimal performance.

F. System Sum Rate

The overall system sum rate is evaluated in this section. We used expression (24) in our computation as similarly done in [34]. 1000 UEs were randomly generated within the cell coverage radius R and the corresponding SINR values appropriately computed. The result of our static UAV deployment is quite comparable to the sum rate as reported in [34]. As expected, our proposed approach gives better results (150 bps/Hz as compared to 137 bps/HZ) for the given cell radius and number of UEs. This is shown in Fig.9. It is worth noting that, standard sum rate analysis typically comapares the sum rate against transmit power and or SINR, however, the above results seeks to show the performance of our proposed UAV MARL scheme and static deployment. Thus, the transmit power is fixed (30dBm).

G. Comparison Between Terrestrial and UAV Deployment with the Same Parameters

We simulate the coverage probability for terrestrial and UAV based cellular communication with the same transmit power and antenna gain. The simplified path loss model [37] is employed for the terrestrial BS-to-UE link with the

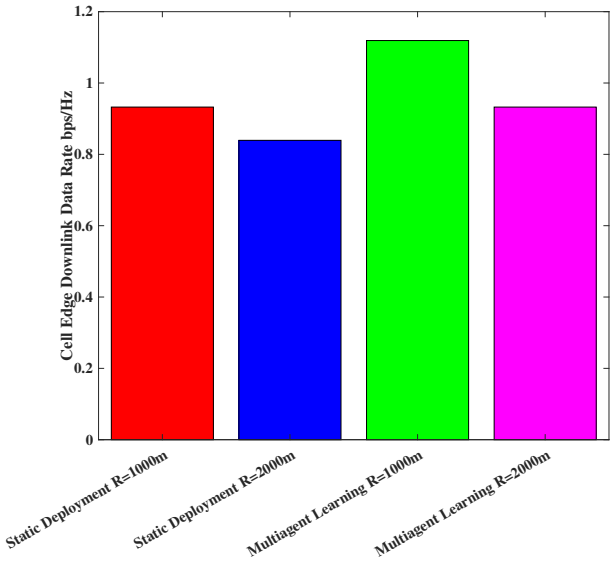


Figure 8. UE Capacity.

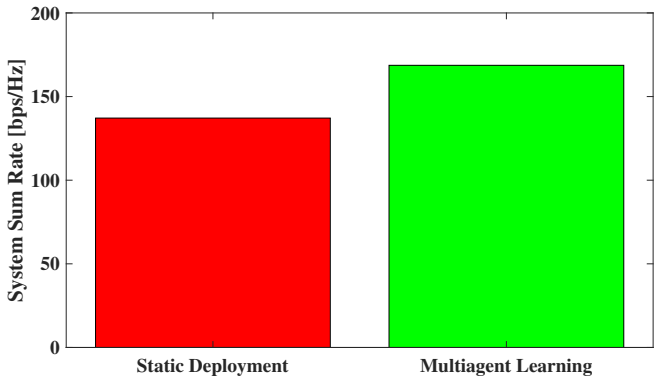


Figure 9. Expected System Sum Rate for R=1000m, 1000 UES and N=10.

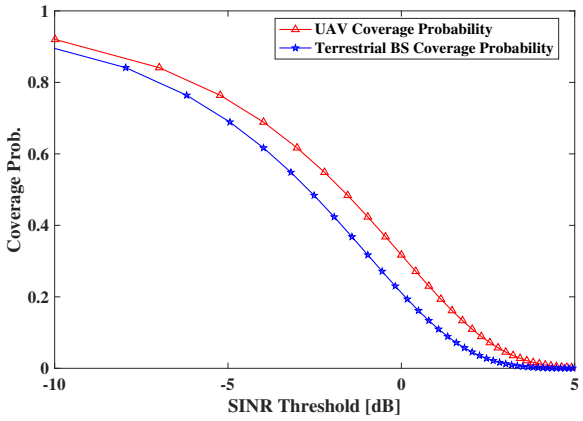


Figure 10. Coverage Comparison for Terrestrial and Aerial Deployment.

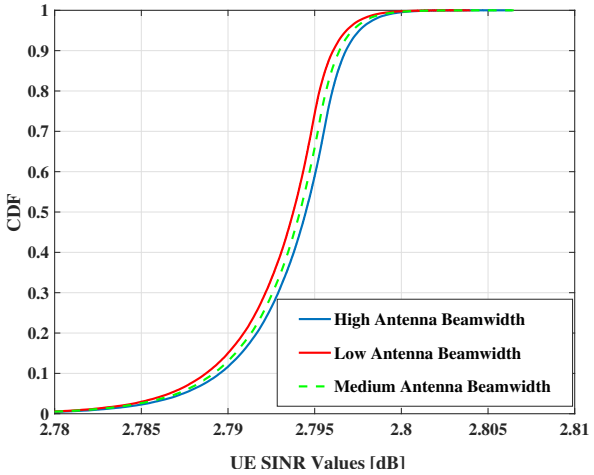


Figure 11. CDF Plot for Varying BS Antenna Gain.

following parameters, pathloss exponent of 3.17, d_0 = 10m, f_c = 2GHz and 8 dB shadowing value. The same parameters for the UAV system as we have in previous subsections are used. As clearly depicted in Fig.10, the UAV deployed BS outperforms the terrestrial system in terms of coverage probability. Besides the improvement in the coverage probability, it is easy to deploy UAV based system to compensate for cell outage in times of disasters [38], and it could also be used for load balancing for overloaded cells.

H. UE SINR Values for Varying Antenna Gain

By altering the values of $\phi_{3\text{dB}}$ and $\theta_{3\text{dB}}$ in (14), we obtained the plots in Fig. 11 with $A_\phi = A_\theta = 20$dB. For a low $\phi_{3\text{dB}}$ and $\theta_{3\text{dB}}$, the antenna gain approaches the maximum value which results in improved UEs received signal strength and SINR values. On the other hand, by increasing the values of the above-mentioned parameters, the antenna gain decreases. The underlying revelation of this results is that, antenna beamforming can be used to directly transmit signals with high gain to UEs as similarly mentioned in [8].

5. CONCLUSION

The deployment of multiple UAVs to provide ground-based users with wireless connectivity is studied in this paper. Tools from Stochastic Geometry are used to derive tractable analytical expressions for modeling the UAV-UE system. We also developed a multi-agent learning technique which enables the UAVs to cooperatively learn from each other by sharing information about their states-actions pairs with their neighbors.

In order to evaluate the impact of interference from other UAVs, different scenarios are analyzed considering varying UAV heights, environments and antenna gains. The feasibility of our proposed learning scheme is evaluated with typical parameters of a base station and the results

show a substantial performance improvement in terms of coverage probability, UEs capacity and system sum rate. Our analytical and simulation results are in agreement with each other.

Appendix A; Proof of Theorem 1: In discrete domain

$$\begin{aligned} p_{I_{agg}}(i_{agg}) &= \mathbb{P}(I_{agg} = i_{agg}) \\ &= \sum_{i_1+i_2=i_{agg}} p_{I_1,I_2}(i_1, i_2) \\ &= \sum_{i_1} \mathbb{P}(I_1 = i_1, I_2 = i_{agg} - i_1) \\ &= \sum_{i_1} p_{I_1,I_2}(i_1, i_{agg} - i_1) \\ &= \sum_{i_1} p_{I_1}(i_1) p_{I_2}(i_{agg} - i_1). \end{aligned} \tag{35}$$

Since I_1 and I_2 are independent, (27) is the simplified probability mass function (PMF) which is the convolution of the PMFs of I_1 and I_2. For the case of continuous random variables (RVs) the proof of theorem 1 is

$$\begin{aligned} F_{I_{agg}}(i_{agg}) &= \mathbb{P}(I_{agg} \le i_{agg}) \\ &= \mathbb{P}(I_1 + I_2 \le i_{agg}) \\ &= \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{i_{agg}-i_1} f_{I_1,I_2}(i_1, i_2) di_2 \Big) di_1 \\ &= \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{i_{agg}} f_{I_1,I_2}(i_1, t - i_1) dt \Big) di_1 \\ &= \int_{-\infty}^{i_{agg}} \Big(\int_{-\infty}^{\infty} f_{I_1,I_2}(i_1, t - i_1) di_1 \Big) dt. \end{aligned} \tag{36}$$

Since I_1 and I_2 are identical and independently distributed (i.i.d.),

$$f_{I_{agg}}(i_{agg}) = \int_{-\infty}^{\infty} f_{I_1,I_2}(i_1, i_{agg} - i_1) di_1 = f_{I_1} * f_{I_1}, \tag{37}$$

which is the convolution of the pdfs of I_1 and I_2.
End of proof.

Appendix B; Proof of Mean SINR

$$\begin{aligned} \bar{\varsigma} &= \mathbb{E}(\varsigma) \\ &= \int_{-\infty}^{\infty} \varsigma \frac{1}{\sqrt{2\pi}\sigma_\varsigma} \exp\Big(\frac{-(\varsigma - \bar{\varsigma})^2}{2\sigma_\varsigma^2}\Big) d\varsigma \\ &= \int_{-\infty}^{\infty} (\varsigma + \bar{\varsigma}) \frac{1}{\sqrt{2\pi}\sigma_\varsigma} \exp\Big(\frac{-(\varsigma)^2}{2\sigma_\varsigma^2}\Big) d\varsigma \\ &= \underbrace{\int_{-\infty}^{\infty} \varsigma \frac{1}{\sqrt{2\pi}\sigma_\varsigma} \exp\Big(\frac{-\varsigma^2}{2\sigma_\varsigma^2}\Big) d\varsigma}_{\text{solving this integral}} + \int_{-\infty}^{\infty} \bar{\varsigma} \frac{1}{\sqrt{2\pi}\sigma_\varsigma} \exp\Big(\frac{-\varsigma^2}{2\sigma_\varsigma^2}\Big) d\varsigma \\ &= -\int_{0}^{-\infty} \varsigma \frac{1}{\sqrt{2\pi}\sigma_\varsigma} \exp\Big(\frac{-\varsigma^2}{2\sigma_\varsigma^2}\Big) d\varsigma + \int_{0}^{\infty} \varsigma \frac{1}{\sqrt{2\pi}\sigma_\varsigma} \exp\Big(\frac{-\varsigma^2}{2\sigma_\varsigma^2}\Big) d\varsigma \\ &= \int_{0}^{\infty} -\varsigma \frac{1}{\sqrt{2\pi}\sigma_\varsigma} \exp\Big(-\frac{-\varsigma^2}{2\sigma_\varsigma^2}\Big) d\varsigma + \int_{0}^{\infty} \varsigma \frac{1}{\sqrt{2\pi}\sigma_\varsigma} \exp\Big(\frac{-\varsigma^2}{2\sigma_\varsigma^2}\Big) d\varsigma \\ &= \underbrace{-\int_{0}^{\infty} \varsigma \frac{1}{\sqrt{2\pi}\sigma_\varsigma} \exp\Big(\frac{-\varsigma^2}{2\sigma_\varsigma^2}\Big) d\varsigma + \int_{0}^{\infty} \varsigma \frac{1}{\sqrt{2\pi}\sigma_\varsigma} \exp\Big(\frac{-\varsigma^2}{2\sigma_\varsigma^2}\Big) d\varsigma}_{0} \\ &= \int_{-\infty}^{\infty} \bar{\varsigma} \frac{1}{\sqrt{2\pi}\sigma_\varsigma} \exp\Big(\frac{-\varsigma^2}{2\sigma_\varsigma^2}\Big) d\varsigma \\ &= \int_{-\infty}^{\infty} \bar{\varsigma} \frac{1}{\sqrt{\pi}} \exp\Big(-\varsigma^2\Big) d\varsigma \ \text{ multiplying by } \sigma_\varsigma \sqrt{2} \\ &= \bar{\varsigma} \frac{2}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\Big(-\varsigma^2\Big) d\varsigma \\ &= \bar{\varsigma} \underbrace{\lim_{t\to\infty} \frac{2}{\sqrt{\pi}} \int_{0}^{t} \exp\Big(-\varsigma^2\Big) d\varsigma}_{1} \\ &= \bar{\varsigma}. \end{aligned} \tag{38}$$

End of proof

References

[1] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3747–3760, June 2017.

[2] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile unmanned aerial vehicles (UAVs) for energy-efficient internet of things communications," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7574–7589, Nov 2017.

[3] J. Kosmerl and A. Vilhar, "Base stations placement optimization in wireless networks for emergency communications," in *IEEE International Conference on Communications Workshops (ICC)*, 2014, pp. 200–205.

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

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

[6] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage," *IEEE Communications Letters*, vol. 20, no. 8, pp. 1647–1650, 2016.

[7] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, 2016.

[8] V. V. Chetlur and H. S. Dhillon, "Downlink coverage analysis for a finite 3-D wireless network of unmanned aerial vehicles," *IEEE Transactions on Communications*, vol. 65, no. 10, pp. 4543–4558, Oct 2017.

[9] F. Jiang and A. L. Swindlehurst, "Optimization of UAV heading for the ground-to-air uplink," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 5, pp. 993–1005, June 2012.

[10] D. W. Matolak and R. Sun, "Unmanned aircraft systems: Air-ground channel characterization for future applications," *IEEE Vehicular Technology Magazine*, vol. 10, no. 2, pp. 79–85, June 2015.

[11] A. Imran and R. Tafazolli, "Performance & capacity of mobile broadband WiMAX (802.16 e) deployed via high altitude platform," in *IEEE European Wireless Conference*, 2009, pp. 319–323.

[12] F. Baccelli, B. Błaszczyszyn *et al.*, "Stochastic geometry and wireless networks: Volume ii applications," *Foundations and Trends® in Networking*, vol. 4, no. 1–2, pp. 1–312, 2010.

[13] F. Baccelli, M. Klein, M. Lebourges, and S. Zuyev, "Stochastic geometry and architecture of communication networks," *Telecommunication Systems*, vol. 7, no. 1, pp. 209–227, 1997.

[14] P. J. Fleming, A. L. Stolyar, and B. Simon, *Closed-form expressions for other-cell interference in cellular CDMA*. University of Colorado at Denver, Center for Computational Mathematics, 1997.

[15] T. X. Brown, "Cellular performance bounds via shotgun cellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 11, pp. 2443–2455, Nov 2000.

[16] V. Chandrasekhar, M. Kountouris, and J. G. Andrews, "Coverage in multi-antenna two-tier networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 10, pp. 5314–5327, October 2009.

[17] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, September 2009.

[18] M. Z. Win, P. C. Pinto, and L. A. Shepp, "A mathematical theory of network interference and its applications," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 205–230, Feb 2009.

[19] J. Peng, H. Tang, P. Hong, and K. Xue, "Stochastic geometry analysis of energy efficiency in heterogeneous network with sleep control," *IEEE Wireless Communications Letters*, vol. 2, no. 6, pp. 615–618, December 2013.

[20] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, November 2011.

[21] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of k-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550–560, April 2012.

[22] D. Wackerly, W. Mendenhall, and R. Scheaffer, *Mathematical statistics with applications*. Nelson Education, 2002.

[23] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3, pp. 279–292, May 1992. [Online]. Available: https://doi.org/10.1007/BF00992698

[24] L. Panait and S. Luke, "Cooperative multi-agent learning: The state of the art," *Autonomous agents and multi-agent systems*, vol. 11, no. 3, pp. 387–434, 2005.

[25] M. Simsek, M. Bennis, and A. Czylwik, "Dynamic inter-cell interference coordination in hetnets: A reinforcement learning approach," in *2012 IEEE Global Communications Conference (GLOBECOM)*, Dec 2012, pp. 5446–5450.

[26] H. Li, "Multi-agent q-learning for competitive spectrum access in cognitive radio systems," in *2010 Fifth IEEE Workshop on Networking Technologies for Software Defined Radio Networks (SDR)*, June 2010, pp. 1–6.

[27] K. Zhu, E. Hossain, and D. Niyato, "Pricing, spectrum sharing, and service selection in two-tier small cell networks: A hierarchical dynamic game approach," *IEEE Transactions on Mobile Computing*, vol. 13, no. 8, pp. 1843–1856, Aug 2014.

[28] N. Vlassis, "A concise introduction to multiagent systems and distributed artificial intelligence," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 1, no. 1, pp. 1–71, 2007.

[29] G. Lei, M. z. Dong, T. Xu, and L. Wang, "Multi-agent path planning for unmanned aerial vehicle based on threats analysis," in *2011 3rd International Workshop on Intelligent Systems and Applications*, May 2011, pp. 1–4.

[30] J. Wang, C. Jiang, Z. Han, Y. Ren, R. G. Maunder, and L. Hanzo, "Taking drones to the next level: Cooperative distributed unmanned-aerial-vehicular networks for small and mini drones," *IEEE Vehicular Technology Magazine*, vol. 12, no. 3, pp. 73–82, Sept 2017.

[31] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications*. John Wiley & Sons, 2013.

[32] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 3949–3963, 2016.

[33] K. W. Sung, M. Tercero, and J. Zander, "Aggregate interference in secondary access with interference protection," *IEEE Communications Letters*, vol. 15, no. 6, pp. 629–631, June 2011.

[34] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

[35] E. Peshkova, M. Hitz, and B. Kaufmann, "Natural interaction techniques for an unmanned aerial vehicle system," *IEEE Pervasive Computing*, vol. 16, no. 1, pp. 34–42, Jan 2017.

[36] C. Yu, M. Zhang, F. Ren, and G. Tan, "Multiagent learning of coordination in loosely coupled multiagent systems," *IEEE Transactions on Cybernetics*, vol. 45, no. 12, pp. 2853–2867, Dec 2015.

[37] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[38] S. Rohde and C. Wietfeld, "Interference aware positioning of aerial relays for cell overload and outage compensation," in *2012 IEEE Vehicular Technology Conference (VTC Fall)*, Sept 2012, pp. 1–5.

Raymond Sabogu-Sumah is a Senior Spectrum Engineer at the National Communications Authority, Ghana. He received the B.Sc. degree (Hons.) in Telecommunications Engineering from the Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, in 2013, and the master's degree in Electronic Engineering from Hanbat National University, Daejeon, South Korea, in 2017. He is currently studying for the PHD at the Kwame Nkrumah University of Science and Technology, Kumasi, Ghana with research interest in Reconfigurable Intelligent Surfaces, 5G and Beyond Networks, Unmanned Aerial Vehicles and Intelligent Spectrum Management Schemes.

Dr. Kwasi Adu-Boahen Opare received the BSc. Electrical/Electronic Engineering and MSc. Telecommunication Engineering degrees in 2004 and 2010, respectively, from the Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana. He obtained a PhD in Information and Communication Engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, Sichuan, China. His research interests include Wireless Communications and Data Communication Networks.

Dr. James Dzisi Gadze received his BSc, MSc and PhD degrees in Electrical Engineering from Kwame Nkrumah Uninversity of Science and Technology, Tuskegee Univesity, AL USA, and Florida Intenational Univesity, Miami USA respectively. He is currently a Senior lecturer at the department of Telecommuniction Enginering in Kwame Nkrumah University of Science and Technology. His research interests include Software Defined Networking, Machine-and Deep-Learning application in Communication Systems, RoF-Based Fronthaul in 5G and Beyond Networks, IoT and Blockchain in Smart Grids.

Prof. Jerry John Kponyo is the Dean of the Quality Assurance and Planning Office of the Kwame Nkrumah University of Science and Technology (KNUST) under the Vice-Chancellor's Office. He is currently the Deputy Project Lead of the KNUST Engineering Education Project (KEEP), a 5.5 Million Dollar Africa Center of Excellence (ACE) Impact project sponsored by the World Bank with a focus on Digital Development and Energy. He is Co-Founder of the Responsible AI Network (RAIN) Africa, which is a collaborative effort between KNUST and TUM Germany. He has published over 50 articles in refereed Journals and Conference proceedings. He is a member of the Ghana Institution of Engineers. Prof. Jerry John Kponyo is currently the coordinator of the West Africa Sustainable Engineering Network for Development (WASEND). Prof Kponyo is the PI and Scientific Director of the Responsible Artificial Intelligence Lab which is a 1 Million Canadian Dollar grant sponsored by IDRC and GIZ. He is also PI for the Partner-Afrika project which is sponsored by BMZ.

Prof. Abdul-Rahman Ahmed has been a lecturer in the Department of Electrical and Electronic Engineering since December 2005. He received his BSc. Degree in Electrical and Electronic Engineering from the Kwame Nkrumah University of Science and Technology in 2002 and was retained by the department to serve as a teaching assistant until 2003 when he received a British Government Scholarship to undertake a postgraduate program in Radio Systems Engineering at the University of Hull, U.K. He returned in 2005 to join the Department of Electrical and Electronic Engineering and subsequently went further to obtain a PhD in Radio Sciences and Engineering from Chungnam National University in South Korea in 2014. He has been teaching mainly telecommunication courses to undergraduate and postgraduate students in the Telecommunication Engineering program of the faculty.

His teaching include courses in Electromagnetic Fields Theory, Electromagnetic Compatibility, Radio Receivers, Telecommunication Circuits, Microwave Engineering and Information Theory. He has served as a course facilitator for the Institute of Distance Learning and has written the course books in Electromagnetic Field Theory and Information Theory. Positions he has occupied in the faculty includes Examination Officer, Postgraduate Programs Coordinator and Vacation Training Officer. He is also an External Examiner in Telecommunications Engineering for a number of Polytechnics in Ghana.

Prof. Ezer Osei Yeboah-Boateng is currently the Deputy Director General in charge of Technical Operations at the National Communications Authority, Ghana. He is a professional Telecoms Engineer and ICT Specialist with a strong hands-on expertise in a wide range of telecommunications switching systems, revenue assurance, cyber-security, digital forensics, business development, digital transformation, project management, change management, knowledge management, strategic IT-enabled business value creation and capabilities to develop market-oriented strategies aimed at promoting growth and market share.
He has had professional training in UK, Japan, Malaysia, South Africa and USA. In 1993, he received the prestigious Japan International Cooperation Agency (JICA) Kenshu-in scholarship; and in 1995, he was adjudged a top-ranked scholar of the Confederation of British Industry (CBI) scholarship for Graduate Engineers program.

Ezer has a Ph.D. in Cyber-Security from Aalborg University in Copenhagen, Denmark; an M.S. in Telecommunications (magna-cum-Laude) from Stratford University, Falls Church, VA in USA; and a B.Sc. (honors) in Electrical and Electronic Engineering from University of Science and Technology (UST) in Kumasi, Ghana.

Edmund Yirenkyi Fianko is an electronic communications engineer with depth of experience in radio frequency spectrum management, telecom and broadcasting regulation, policy formulation, change management, ICT industry research, writing, publishing, teaching and public speaking. He has been actively involved in the electronic communications regulatory environment in Ghana and within the Africa region over the last fifteen (15) years.

At the international front, he is; a member of the Radio Regulations Board (RRB) of the International telecommunications Union representing Region D (Africa) for the 2023-2026 term, he is also the current Chairman of the African Telecommunications Union (ATU) Task Group on Development of Spectrum Recommendations for Rural Connectivity and a member of the Technical Committee of the ITU Policy and Regulation Initiative for Digital Africa (PRIDA). He is currently the Acting Head of Engineering at the National Communications Authority, Ghana.

Edmund obtained his first degree in Electrical/Electronic Engineering from the Kwame Nkrumah University of Science and Technology (KNUST) and a Masters in Communications Management degree from the Buckinghamshire (Bucks) New University, England.