PLOS ONE

LEO satellite assisted UAV distribution using combinatorial bandit with fairness and budget constraints

Ehab Mahmoud Mohamed *

Department of Electrical Engineering, College of Engineering in Wadi Addwasir, Prince Sattam Bin Abdulaziz University, Wadi Addwasir, Saudi Arabia

* ehab_mahmoud@aswu.edu.eg

Abstract

In this paper, an integration between a low earth orbit satellite (LEO-Sat) and unmanned aerial vehicle (UAV) is proposed to assist users in post-disaster areas. In this scenario, multiple UAVs will be distributed to fully cover the victims and provide rescue services, while LEO-Sat provides backhaul links for UAVs to the ground base station (GBS). In this regard, we consider the problem of efficient UAVs distribution to maximize the total sum rate of the victims while assuring fairness in their coverage within the limited resources of UAVs batteries and LEO-Sat bandwidth. In this paper, UAV distribution problem is considered as a combinatorial multi-armed bandit (MAB) with arms' fairness and limited UAVs battery budget (CMAB-FB) constraints. Additionally, the utilization of LEO-Sat bandwidth resources is optimized based on the average traffic demands of the LEO-UAV links by means of gradient decent algorithm. The results of numerical analysis indicate that the proposed approach outperforms other naïve ben chmarks.

I. Introduction

In recent years, unmanned aerial vehicles (UAVs) have been increasingly utilized in post-disaster relief efforts, due to their ability to reach remote and hard-to-access areas quickly [1,2]. In post-disaster scenarios, UAVs can provide wireless platforms for user equipments (UEs) belonging to victims and rescue workers. However, effective deployment of UAVs in such scenarios is a complex task that requires balancing multiple conflicting objectives, including coverage, limited UAVs battery life, and communication capabilities [1,2]. The limited transmission range of UAVs may hinder them from communicating directly with the nearest ground base station (GBS) or with each other, impacting their deployment and data gathering capabilities in post-disaster areas. On the other side, low earth orbit satellites (LEO-Sat) have revolutionized wireless communications, offering a myriad of applications that have transformed the way we connect and communicate [3]. LEO-Sats, positioned at an altitude ranging from 500 to 2,000 kilometers above the earth's surface, provide several advantages in wireless communications [3]. Firstly, their proximity to the earth enables lower latency, reducing signal delays and enhancing real-time communication. Additionally, LEO-Sats offer high bandwidth

capabilities, enabling faster data transmission and accommodating the increasing demand for high-speed internet access. These satellites have opened a world of possibilities for wireless communication applications, including global broadband internet coverage, internet-of-things (IoT) connectivity, remote sensing, and disaster management [3]. With their extensive coverage, LEO-Sat provides connectivity to even the most remote areas, bridging the digital divide and facilitating global communication networks. Moreover, they play a crucial role in disaster response, offering immediate connectivity and enabling effective coordination and communication during emergencies [3].

In this paper, an integration between LEO-Sat and UAVs in post-disaster scenario is proposed with the aim of efficiently distributing UAVs among the post-disaster zones. In the proposed scenario, LEO-Sat will be used to provide crucial backhaul links for UAVs, allowing them to communicate with the nearest survival GBS and relay critical data/control information, while considering the limited bandwidth resources of LEO-Sat. This cannot be achieved in traditional scenarios, i.e., without LEO-Sat, due to the destroyed/malfunctioned GBS, and the difficulty for the UAVs to directly connect with the nearest survival GBS due to their limited coverage area. In post-disaster zones, the deployment of UAVs should prioritize maximizing their achievable data rates while considering their limited battery resources and the limited bandwidth resources of LEO-Sat as well as ensuring fairness in coverage among post-disaster zones based on their UEs density. Essentially, subareas with higher UE densities should receive more frequent service than those with lower densities. The challenge of this optimization problem comes from three main folds: 1) How to ensure coverage fairness among UEs in the post-disaster area where the wireless infrastructure is completely destroyed. 2) UAVs flying and hovering battery consumptions should be saved while maximizing the information gathered from the post-disaster area. 3) Finally, LEO-Sat has limited bandwidth resources mandating its optimal utilization based on UAVs traffic needs.

In this paper, we will address this problem using online learning by proposing a combinatorial multi-armed bandit (MAB) algorithm with fairness and UAVs battery budget restrictions for UAVs distribution [4]. Then and based on it, the gradient decent algorithm [5] will be used for optimizing the LEO-Sat bandwidth resources. Generally speaking, MAB [6] is an efficient reinforcement learning approach, where a player aims to maximize his achievable profit while playing over the bandit arms. The only available information to the player is the played arms along with their observed rewards, without any prior information about the game and its statistics. During the game, the player tries to balance between always exploiting the arm with the highest observable reward so far or exploring new ones, which is known as the *exploitation-exploration* tradeoff of the MAB games [7]. The proposed combinatorial MAB with fairness and budget restrictions is a type of MAB games, where the player selects a combination of multiple arms called the "super-arm" in each trial, while assuring fairness in the selected arms limited by the allowable cost budget paid by the player to select this "super-arm" [4]. In the proposed MAB model, the nearest survival GBS will act as the player, and the arms are the post-disaster zones over which UAVs should be distributed limited by their battery budget. Herein, the LEO-Sat will be used to relay data and control information between GBS and UAVs. To this end, the optimization problem will be formulated, and the combinatorial MAB with fairness and budget constraints (CMAB-FB) will be proposed to efficiently address the UAV distribution problem. Then, a gradient decent algorithm will be used to optimize the LEO-Sat bandwidth resources. Thus, the main contributions of this paper can be summarized as follows:

- An integrated LEO-Sat and UAVs network is proposed for rescue services in post-disaster area, where LEO-Sat will be used to relay control/traffic data to/from GBS from/to UAVs. In

this network, the problem of UAVs distribution among users in this post-disaster area will be considered. The distribution of UAVs aims to maximize the total achievable data rate of the victims, while maintaining fairness in their coverage. Also, UAVs' limited battery budget should be considered along with the limited bandwidth resources of the LEO-Sat. To the best of our knowledge, this is the first research effort for proposing LEO-Sat assisted UAV distributions in post-disaster scnearios.

- To efficiently address the efficient UAVs distribution problem within its UAVs energy constraints, it is considered as a MAB game, where "CMAB-FB" algorithm is proposed to implement it. The proposed "CMAB-FB" algorithm will be used to distribute the UAVs among the post-disaster zones to maximize the users' achievable data rate while maintaining coverage fairness among them within the limited battery budget of UAVs. Then, based on the achieved UAVs rates, the bandwidth resources of the LEO-UAV links will be optimized subject to LEO-Sat limited bandwidth resources by means of gradient decent algorithm. By this way, the aforementioned challenges can be addressed as follows: 1) UEs coverage fairness can be assured based on their densities in post-disaster grids, where their densities can be pre-estimated using GPS localization and refined through UAVs exploration during the MAB game. 2) The budget constraint nature of the "CMAB-FB" algorithm results in saving UAVs energy resources. 3) The bandwidth resources of the LEO-Sat will be optimized based on the selected "super-arm", i.e., grids, at each time slot.

- Numerical analyses are conducted to prove the effectiveness of the proposed scheme under different scenario. In this regard, the proposed scheme is compared against other benchmarks such as random and nearest UAV distribution, where the proposed scheme demonstrates superior performance over these benchmark schemes.

The rest of this paper is organized as follows: Section II gives the related works to that presented in this paper. Section III gives the proposed system model including the optimization problem formulation, Section IV gives the proposed "CMAB-FB" and the gradient descent algorithms, Section V gives the conducted numerical simulations, followed by the concluding remarks in Section VI.

II. Literature review

Recently, extensive research has been undertaken to explore the utilization of UAVs for the purpose of supporting post-disaster areas. The synergistic employment of UAVs with cellular networks, and wireless sensor networks (WSNs) to facilitate disaster management applications was investigated in [8]. In another research endeavor given in [9], a genetic algorithm was employed to optimize the placement of UAVs, with the objective of enhancing both the overall coverage and data rate of the wireless network. In [10], an effective methodology to aid rescue operations in locating victims affected by a natural disaster was proposed. This approach involved the utilization of UAVs equipped with LiDAR and infrared depth cameras to construct an independent detection system that was not reliant on illumination intensity. Additionally, [11] involved the deployment of UAVs integrated with a video recorder and geolocation module, enabling the search for survivors within a post-disaster area. Furthermore, [12] explored the concept of flying communication services by equipping UAVs with Wi-Fi, video cameras, and web servers. The objective was to empower individuals affected by a disaster to utilize their smartphones for real-time text and video communication. Building upon this research, the authors of [13] proposed a mobility model based on the self-deployment of an aerial ad hoc network, utilizing the Jaccard dissimilarity metric to facilitate post-disaster scenarios. This mobility model incorporated the movement of victims and generated

corresponding UAV mobility patterns to track these individuals. In the realm of disaster management systems, [14] presented a novel approach for energy-efficient task scheduling for the collected data obtained by UAVs from ground IoT networks. As for [15], UAVs were leveraged as on-demand airborne relays to establish connectivity between remote users and GBS, particularly when they were geographically isolated by substantial obstacles. In [16], millimeter wave (mmWave) UAV gateway selection is proposed in post-disaster scenarios. In [17], enhanced dynamic spectrum access was proposed empowered by MAB schemes for UAV networks in post-disaster scenario. This work was then extended in [18] to optimize the UAV 3D trajectory as well. Almost all work existing in literature regarding UAVs applications in post-disaster scenarios assumed that UAVs fly go-and-forth between the post-disaster zones and the ground fusion center in the nearest survival GBS. This will highly consume the UAVs energy and delay the rescue services operations. Moreover, all the above existing research works assumed a full awareness of the network parameters, which cannot be easily obtained in the completely destroyed infrastructure in post-disaster scenarios. In this paper, we leverage LEO-Sat as backhauling for the UAVs networks, which highly relaxes the need of UAVs flying between GBS and post-disaster zones. Moreover, as the proposed scheme is based on online learning, which does not need any prior information about the environment, except users' pre-locations obtained using GPS services.

Regarding the integration between LEO-Sat and UAVs for enhancing/extending aerial coverage. In [19], the authors used a combination of LEO-Sat and UAVs for beyond-5G communication, utilizing millimeter-wave (mmWave) and free-space optical (FSO) links. A multi-agent deep reinforcement learning (MARL) approach is used to optimize communication and energy efficiencies, leading to improved peak and worst-case throughput compared to using only one of the links. In [20], the authors proposed the use of UAVs and LEO-Sat for data collection from internet-of-remote-things (IoRT) sensors. In [21], the authors proposed using LEO-Sat and caching by UAVs for content delivery in terrestrial networks to improve connectivity and capacity. The problem of optimizing cache placement, resource allocation, and trajectory was solved using an alternating algorithm. In [22], LEO-Sats were used for UAV tracking using Gauss Hermite filter based on hybrid TDOA/FDOA geolocation measurements. In [23], an integration between LEO-Sat and UAVs was considered for integrated mobile edge caching IoT system. In this model, LEO-Sat broadcasts data, and UAVs collect it from decentralized ground sensors. UAV deployment and power allocation for secure space-air-ground communications was considered in [24]. The aim of the formulated optimization problem was to maximize the secrecy rate subject to UAV's power and deployment area. Despite the existing research in LEO-UAV integration, none of them considered the problem of LEO-Sat assisted UAV distribution in post-disaster area as presented in this work.

III. System model and optimization problem formulation

The proposed system model for LEO-UAV integration to cover the post-disaster area is shown in Fig 1. The area is divided into a set of M non-overlapped grids collected in $\mathcal{M}$. Each grid $i \in \mathcal{M}$ contains K_i UEs. Dividing the post-disaster area into non-overlapped zones comes from the nature of the victims' distributions which are typically formed in sparse groups due to the destruction happening in the area or grouping them within rescue shelters. This assumption was also considered in [25,26], when using UAVs in rescue services in post-disaster scenarios. In the proposed system model, LEO-Sat relays control and traffic data between GBS and UAVs for both control and traffic data. In this study, it is assumed that a set of N UAVs collected in $\mathcal{N}$, $N = 5$ in Fig 1, where $j \in \mathcal{N}$ and $N<M$, are always within the coverage area of the LEO-Sat. At each time slot t, the GBS allocates the N UAVs to a selected group of the post-
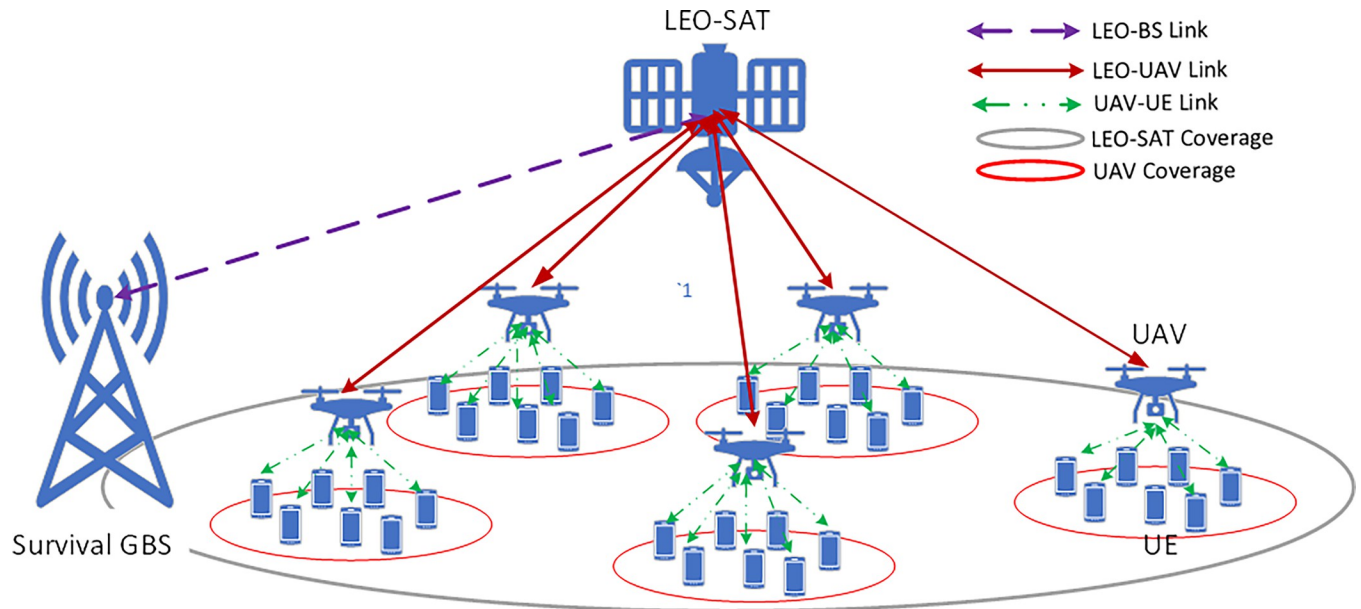
Fig 1. LEO-Sat assisted UAVs in a post-disaster area.

https://doi.org/10.1371/journal.pone.0290432.g001

disaster grids, and it relays the information of the selected grids to the UAVs via the LEO-Sat. The criterion based on which the GBS selects the grids is to maximize the achievable data rates of their users, while assuring fairness in the grids coverage based on their users' densities. Then, UAVs will fly towards the selected grides and hover above them to relay their users' traffic data to the GBS via the LEO-Sat backhaul links. This operation is repeated over the time horizon constrained by UAVs' battery capacities and LEO-Sat bandwidth resources. In the followings, the utilized UAV-UE and LEO-UAV link models will be explained in detail, and then the optimization problem of UAVs distribution will be formulated.

A. UAV-UE link model

For the UAV-UE link model, we utilized the simple link model given in [27], which can be given as follows:

$$\Lambda_{j,k_i}(x_{j,k_i}) = \mathbb{P}^{\text{LoS}} \lambda_{j,k_i}^{\text{LoS}}(x_{j,k_i}) + \mathbb{P}^{\text{NLoS}} \lambda_{j,k_i}^{\text{NLoS}}(x_{j,k_i}), \tag{1}$$

$$\begin{cases} \Lambda_{j,k_i}^{\text{LoS}}\left(x_{j,k_i}\right) = 20 \log\left(\dfrac{4\pi x_{j,k_i}}{\lambda_U}\right) + \rho^{\text{LoS}} \\[4mm] \Lambda_{j,k_i}^{\text{NLoS}}\left(x_{j,k_i}\right) = 20 \log\left(\dfrac{4\pi x_{j,k_i}}{\lambda_U}\right) + \rho^{\text{NLoS}} \end{cases} \tag{2}$$

$\Lambda_{j,k_i}(x_{j,k_i})$, $\Lambda(x_{j,k_i})$ and $\Lambda_{j,k_i}^{\text{NLoS}}(x_{j,k_i})$ are the total path loss, its line-of-sight (LoS) component, and non-LoS (NLoS) in dB between UAV j and UE k in grid i as a function of their separation distance x_{j,k_i} Herein, λ_U is the wavelength of the UAV-UE link, and ρ^{LoS} and ρ^{NLoS} indicate the system loss in dB for both LoS and NLoS, respectively. $\mathbb{P}^{\text{LoS}}$ and $\mathbb{P}^{\text{NLoS}}$, where $\mathbb{P}^{\text{NLOS}} = 1 - \mathbb{P}^{\text{LOS}}$, are the probabilities for LoS and NLoS links. $\mathbb{P}^{\text{LoS}}$ is defined as follows [27].

$$\mathbb{P}^{\text{LoS}} = \left[1 + a \exp(-b(\theta_{j,k_i} - a))\right]^{-1}, \tag{3}$$

where a and b are constants based on the environment, while θ_{j,k_i} indicates the elevation angle between UAV j and UE k in grid i. Herein, $\theta_{j,k_i} = \tan^{-1}\left(\frac{h_j}{d_{Hj,k_i}}\right)$, where h_j is the hovering height of UAV j and d_{Hj,k_i} is the horizontal distance between UAV j and UE k in grid i. Without loss of generality, uplink transmission is assumed between UEs in grid i and UAV j. Thus, the average data rate between UAV j and grid i at time slot t, $\Psi_{j,i}^{t}$, can be given as follows:

$$\Psi_{j,i}^{t} = \frac{B}{K_i}\sum_{k=1}^{K_i}\log_2(1 + P_{j,k_i}^{\text{Rx,t}}/\sigma^2), \tag{4}$$

In (4), frequency division multiple access (FDMA) is considered among UEs in grid i, where the total bandwidth B is equally divided among UEs in grid i which is denoted as K_i. $P_{j,k_i}^{\text{Rx,t}}$ is Rx power at UAV j from UE k in grid i at time slot t, and σ^2 is the noise power of the UAV-UE link, respectively. This model assumes uplink transmissions, but the same principle can be applied to the downlink. Moreover, the grids are assumed to be sparse as considered in [25,26], which prevents mutual interferences among grids, which is a reasonable assumption in post-disaster areas as previously explained.

B. LEO-UAV link model

For LEO-UAV communication link, we utilized the link model given in [28], where the received power, $P_{S,j}^{\text{Rx,t}}$, and the achievable uplink data rate in bps at LEO-Sat from UAV j, η_{S,j_i}^{t}, at time slot t are written as follows:

$$P_{S,j}^{\text{Rx,t}} = \frac{P_{j,S}^{\text{Tx}}G_j^{\text{Tx}}G_S^{\text{Rx}}\lambda_S}{4\pi x_{S,j}^{t}}, \tag{5}$$

$$\eta_{S,j}^{t} = B_{S,j}^{t}\log_2\left(1 + \frac{P_{S,j}^{\text{Rx,t}}}{\tau\varepsilon B_{S,j}^{t}}\right), \tag{6}$$

where $P_{j,S}^{\text{Tx}}$ is the Tx power from UAV to LEO-Sat, and G_j^{Tx}, G_S^{Rx} are the Tx and Rx antenna gains from UAV (LEO-Sat) towards LEO-Sat (UAV), respectively. $x_{S,j}^{t}$ is the separation distance between the LEO-Sat and UAV. Also, λ_S is the wavelength of the LEO-UAV link. In (6), $B_{S,j}^{t}$, τ and ε are the allocated bandwidth of the LEO-Sat link at time slot t, the noise temperature, and the Boltzmann constant, respectively.

C. Optimization problem formulation

At each time slot t, the GBS should decide which grids from $\mathcal{M}$ the UAVs should cover, and then assigns UAVs to these selected grids. This information is communicated to the UAVs via the LEO-Sat. The objective is to distribute the UAVs-grids in a manner that maximizes the long term achievable data rates of the UAVs and assures fairness among the grids based on their UEs density, while also considering the limited battery capacities of the UAVs and the limited bandwidth of the LEO-Sat. Mathematically speaking, this optimization problem can be formulated as follows:

$$\max_{\eta_{S,j}^{t}, d_{j,i}^{t}} \frac{1}{T}\sum_{t=1}^{T-1}\sum_{j=1}^{N}\sum_{i=1}^{M} d_{j,i}^{t}\min(\eta_{S,j}^{t}, \Psi_{j,i}^{t}) \tag{7A}$$

s.t.

$$d_{j,i}^{\mathrm{t}} \in \{0, 1\} \tag{7B}$$

$$\sum_{j=1}^{N} \sum_{i=1}^{M} d_{j,i}^{\mathrm{t}} \leq N \tag{7C}$$

$$\sum_{i=1}^{M} d_{j,i}^{\mathrm{t}} = 1 \ \forall j \in \mathcal{N} \tag{7D}$$

$$d_{j,i}^{\mathrm{t}} \Gamma_{j,i}^{\mathrm{t}} \leq \Gamma_{\mathrm{Umax}} \ \forall j \in \mathcal{N} \tag{7E}$$

$$\sum_{j=1}^{N} \sum_{i=1}^{M} d_{j,i}^{\mathrm{t}} \eta_{S,j}^{\mathrm{t}} \leq \eta_{S\mathrm{max}} \tag{7F}$$

$$\frac{1}{T} \sum_{t=0}^{T-1} \sum_{j=1}^{N} d_{j,i}^{\mathrm{t}} \geq \frac{K_i}{\sum_{i=1}^{M} K_i} \ \forall i \in \mathcal{M} \tag{7G}$$

where,

$$\Gamma_{j,i}^{\mathrm{t}} = P_{\mathrm{f}} T_{j,i}^{\mathrm{f,t}} + P_{\mathrm{h}} T_{j,i}^{\mathrm{h,t}}, \tag{8}$$

where T indicates the total time horizon, $d_{j,i}^{\mathrm{t}} \in \{0, 1\}$ is a selection indicator which is equal to one if grid i is selected to be covered by UAV j at time slot t, and zero otherwise. $\eta_{S,j}^{\mathrm{t}}$ is the assigned capacity of the LEO-UAV of UAV j at time slot t. The 2nd constraint in (7) means that the total number of selected grids should be less than or equal to the total number of UAVs N as the battery of some of UAVs may be completely depleted during the coverage process and needs for re-charging. The 3rd constraint means that each UAV j should cover only one grid i at a time slot t. The 4th constraint means that the energy consumption of UAV j, $\Gamma_{j,i}^{\mathrm{t}}$, required to serve grid i at time slot t should not exceed its total battery capacity Γ_{Umax}. $\Gamma_{j,i}^{\mathrm{t}}$ is defined in (8) and considers both the flying and hovering powers (P_{f} and P_{h}) and times ($T_{j,i}^{\mathrm{f,t}}$ and $T_{j,i}^{\mathrm{h,t}}$) of the UAV. $T_{j,i}^{\mathrm{f,t}}$ is the ratio of the distance between the UAV's current location and its target location in grid i at time slot t, divided by the UAV speed, while $T_{j,i}^{\mathrm{h,t}}$ is determined by the traffic demand of grid i relative to $\Psi_{j,i}^{\mathrm{t}}$. Actually, there are eight sources of UAV power consumptions as given in details in [29]. However, the flying and the hovering power consumptions are the most dominant ones as shown in [29], with flying consumes more energy than hovering [29]. Both P_{f} and P_{h} are related to the mass of the UAV, the gravitational force, the radius of the propeller, and the air density. In addition, P_{f} depends on the deviation angle between the UAV vertical axis and the Z axis as shown in [29]. For more details about various sources of UAV power consumption and their mathematical details, interested readers are advised to check [29]. The 5th constraint means that as the link is established between UAV j and grid i at a time slot t, i.e., $d_{j,i}^{\mathrm{t}} = 1$. It should be assigned a bandwidth resource from the LEO-Sat with specific data rate $\eta_{S,j}^{\mathrm{t}}$, where the sum data rates of all UAV-UE links corresponding to $d_{j,i}^{\mathrm{t}} = 1$ must not exceed the total capacity of the LEO-Sat, i.e., $\eta_{S\mathrm{max}}$. The 6th constraint is used to ensure fairness among grides based on their UEs density.

IV. Proposed CMAB-FB and gradient decent algorithms

The problem given in (5) is a dynamic non-linear programing problem, without a closed form optimal solution. However, this problem can be simplified by splitting it into two stages, where

LEO-Sat capacity resources $\eta_{S,j}^t$ are typically optimized based on required UAVs traffic rates, i.e., $\Psi_{j,i}^t$, which is based on the optimized $d_{j,i}^t$ values. Thus, in the first stage, the UAV-grid association parameters $d_{j,i}^t$ are optimized, while in the second stage, $\eta_{S,j}^t$ values are adjusted based on the optimized $d_{j,i}^t$ and their corresponding $\Psi_{j,i}^t$. To optimize the values of $d_{j,i}^t$, the problem is considered as a combinatorial MAB problem with arms' fairness constrained by UAV battery budget. Thus, in this section, we will explain the MAB model in general, then we introduce the proposed "CMAB-FB" algorithm for adjusting the values of $d_{j,i}^t$. Finally, the values of $\eta_{S,j}^t$ are adjusted using gradient decent algorithm.

A. MAB model

MAB is an efficient online learning methodology, where a player plays over bandit arms and observes their rewards [7]. The player aims to maximize his long-term profit by learning to always play with the arm having the maximum achievable reward. The player has no-prior knowledge about the game except the played arms and their corresponding rewards. At each time slot during the game, the player tries to compromise between always exploiting the arm having the highest reward so far or exploring new unknown ones. This is what is called the *exploitation-exploration* dilemma of the MAB games [8]. There are several algorithms that can efficiently implement the MAB hypothesis like upper confidence bound (UCB), $\epsilon-$greedy, Thompson sampling (TS), etc [30]. In some of the MAB games, selecting an arm comes with paying cost, which is constrained by the player's limited budget. This type of MAB game is called budget constraints games, where budget constraint UCB and budget constraint TS are two well-known MAB algorithm variants that can efficiently implement these types of MAB games [31]. Also, in some cases, the player should select a group of arms from the arm space, which is called a "super-arm"; this type of MAB game is called a combinatorial bandit because a combination of arms should be selected at each time slot [32–34]. Sometimes fairness is required among the selected arms while selecting the super-arms, which are called combinatorial bandit with fairness constraints [4].

B. Optimization of $d_{j,i}^t$

In this stage, we will optimize the values of $d_{j,i}^t$ adhering to constraints 1, 2, 3, 4 and 6. This is a time sequential combinatorial non-linear optimization problem, which can be viewed as a combinatorial MAB game with fairness considerations and UAV battery limitations. In this scenario, the player (i.e., GBS) chooses a "super-arm" by combining several "arms" (i.e., grids) at each time slot t, balancing the selection among the available arms based on their UEs density and minimizing the energy costs for the UAVs covering the chosen grids. Algorithm 1 gives the proposed "CMAB-FB" algorithm, where it is influenced by learning with fairness algorithm given in [4]. This algorithm and the LEO-Sat resource optimization will be run by the GBS platform as it is the player of the MAB game and the most powerful entity in the proposed LEO-UAV network. The inputs to the algorithm are the sets $\mathcal{M}$ and $\mathcal{N}$, and the design parameter Ω. Also, the values of UEs density $\delta_i = K_i / \sum_{i=1}^M K_i$ for grids is input to the algorithm, where K_i can be pre-estimated using GPS localization and then refined by UAVs exploration during the MAB game. For initialization, at t = 0, the selection vector, i.e., w_i^t, the number of times grid i was selected up to time slot t, i.e., h_i^t, the average rate of grid i up to time slot t, i.e., $\hat{\gamma}_i^t$, and the queue of grid i, q_i^t, are all set to zero $\forall i \in \mathcal{M}$. q_i^t is used to assure fairness among the grids as will be explained shortly. For t = 1 to T, the upper confidence bound (UCB) values for $\forall i \in \mathcal{M}$ are set to $\bar{\gamma}_i^t = \hat{\gamma}_i^{t-1} + \sqrt{(3 \log t)/(2h_i^{t-1})}$ if $h_i^{t-1} > 0$ or $\bar{\gamma}_i^t = 1$, if $h_i^{t-1} = 0$. Then q_i^t is

evaluated for $\forall i \in \mathcal{M}$ as given in the algorithm. At every time slot, the value of q_i^t is increased by δ_i and decreased by 1 if grid i was selected in time slot $t-1$. Thus, if grid i is not selected many times, its q_i^t is increased by multiples of its UE density value, and vice versa, which gives it a high priority for being selected in the next time slot. Thus, after evaluating q_i^t and $\bar{\gamma}_i^t$ for $\forall i \in \mathcal{M}$, a super arm $A(t) \subset \mathcal{M}$ is selected based on the following equation:

$$A^t = \underset{A^t \subset \mathcal{M}}{\operatorname{argmax}} \sum_{i \in A^t} (q_i^t + \Omega \bar{\gamma}_i^t), |A^t| \leq N, \tag{9}$$

where Ω is a design parameter used to balance between selecting the grid maximizes the achievable average data rate or that maximizes fairness based on q_i^t values. A^t can be easily evaluated by enumerating the $|A^t|$ grids having the highest values of $(q_i^t + \Omega \bar{\gamma}_i^t)$. After obtaining A^t, the UAVs should be distributed among them, and obtaining $d_{j,i}^{*,t}$ in the way that minimizes their

```
Algorithm 1. Proposed CMAB-FB Algorithm.
Output: d*,t
        j,i
Input: M,N, initial values of δ_i using GPS localization, and Ω
Initialization: At t = 0, w_i^t = 0, h_i^t = 0, γ̂_i^t = 0, and q_i^t = 0 ∀i ∈ M
For t = 1,2,...,T
  For i = 1,2,...,M
    If h_i^{t−1} > 0
        γ̄_i^t = γ̂_i^{t−1} + √((3 log t)/(2h_i^{t−1}))
      else
      γ̄_i^t = 1
      End if
    q_i^t = max{q_i^{t−1} + δ_i − w_i^{t−1}, 0}
  End for
    • Select the super arm A^t as follows:
      A^t = argmax_{A^t⊂M} ∑_{i∈A^t} (q_i^t + Ωγ̄_i^t), |A^t| ≤ N
    • Select d*,t that minimizes UAVs energy consumptions as:
            j,i
      1. Calculate Γ_{j,i}^t matrix, ∀i∈A^t and ∀j ∈ N using (8)
      2. Connect UAV j to grid i as follows:
        For itr = 1:N
        a) {i*,j*} = argmin Γ_{j,i}^t
        b) d_{j,i}^{*,t} = 1
        c) Γ_{j*,i*}^t = Inf
        End for
    • Observe the average rates of the selected super arm A^t and update
its related parameters as follows:
      1. w_i^t = 1 ∀i ∈ A(t).
      2. h_i^t = h_i^{t−1} + 1 ∀i ∈ A^t
      3. γ̂_i^t = (1/h_i^t) ∑_{y=1}^{h_i^t} w_i^y Ψ_{j,i}^y ∀i ∈ A^t
      4. Update δ_i ∀i∈A^t
End for
```

energy consumptions as given in Algorithm 1. As a final step, the average data rate corresponding to the selected A^t are observed, and its related parameters are updated as given in the Algorithm 1 including the actual observed δ_i values corresponding to A^t. As the "super-arm" selection in the proposed "CMAB-FB" algorithm is done in the same way as that given in [4]. The time average accumulative regret of the proposed algorithm is the same, which is defined

as follows [4]:

$$R(T) \leq \frac{M}{2\Omega} + \frac{2\sqrt{6NMT\log T} + \left(1 + \frac{5\pi^2}{12}\right)M}{T}, \qquad (10)$$

For detailed derivation of (10), readers are advised to check [4].

C. Optimization of $\eta_{S,j}^t$

In the second step after obtaining $d_{j,i}^{*,t}$, the term $\Psi_{j,i}^t$ becomes a constant in (7), then we can optimize the values of $\eta_{S,j}^t$ under constraint 5 as follows:

$$\max_{\eta_{S,j}^t} \frac{1}{T} \sum_{t=1}^{T-1} \sum_{j=1}^{N} \min(\eta_{S,j}^t, \Psi_{j,i}^t) \qquad (11)$$

s.t.

$$\sum_{j=1}^{N} \eta_{S,j}^t \leq \eta_{S\max}$$

Typically, the goal of optimizing LEO-Sat resources $\eta_{S,j}^t$ is to accommodate the traffic demands of UAVs $\Psi_{j,i}^t$ within LEO-Sat maximum capacity constraint $\eta_{S\max}$. This allows us to simplify the optimization problem given in (11) as to minimize the absolute difference between LEO-Sat offering rate $\eta_{S,j}^t$ and $\Psi_{j,i}^t$ as follows:

$$\min_{\eta_{S,j}^t} \frac{1}{T} \sum_{t=1}^{T-1} \sum_{j=1}^{N} \text{abs}|\eta_{S,j}^t - \Psi_{j,i}^t| \qquad (12)$$

This problem is a non-linear programing problem due to the absolute value. However, it can be linearized by simplifying the absolute function, then solved by any famous iterative method like gradient decent method as given in Algorithm 2. In this Algorithm, we omit t for notation simplification. The inputs to the algorithm are $\Psi_{j,i}$ obtained after UAVs distribution done by Algorithm 1, $\eta_{S\max}$, the value of the learning rate α, the maximum number of iterations *MaxIter*, the tolerance value *Tolernc*, and the number of UAVs N. The output is the optimized values of $\eta_{S,j}^{*t}$. For initialization, the values of $\eta_{S,j}$ are randomly selected. Then for *iter* = 1 to *MaxIter*, the gradient vector, *Gradient*, of size1×N is set to 0. Then for j = 1 to N the difference between $\eta_{S,j}$ and $\Psi_{j,i}$ is calculated, if it is greater than or equal to 0, the *Gradient* {j} is set to +1, otherwise it is set to -1. Then, the values of η_{Sj} are updated as follows:

$$\eta_{S,j} = \eta_{S,j} - \alpha \times Gradient, \quad \forall j \in \mathcal{N}, \qquad (13)$$

After updating $\eta_{S,j}$, the conditions of maximum capacity constraint, the gradient tolerance and the maximum number of iterations are tested successively and if any one of them is satisfied, the value of $\eta_{S,j}^*$ is returned and the Algorithm will be terminated. After adjusting the values of $\eta_{S,j}^{*t}$, the values of $B_{S,j}^{*t}$ can be easily obtained by solving (6).

From Algorithm 1 and 2, we can conclude that the fairness, UAVs energy conserving and LEO-Sat resource optimization provided by the proposed scheme comes at the expense of slight increase in computational complexity compared to other naïve benchmarks.

```
Algorithm 2. Gradient Decent Algorithm for η_{S,j}^t Optimization.
Output: η*_{S,j}
Input: Ψ_{j,i}, η_{Smax}, α, MaxIter, Tolernc, N
Initialization: η_{S,j} is randomly initialized
```

```
For iter = 1,2,...,MaxIter
     Gradient{1:N} = 0
   For j = 1,...,N
     If η_{S,j}-Ψ_{j,i}≥0
         Gradient{j} = +1
     else
         Gradient{j} = -1
     End if
   End for
```
$$\eta_{S,j} = \eta_{S,j} - \alpha \times Gradient \ \forall j \in \mathcal{N}$$
$$\sum_{j=1}^{N} \eta_{S,j} \leq \eta_{Smax}$$
$$\eta_{S,j}^{*} = \eta_{S,j}$$
```
   Break For
   nd If
   If ‖Gradient‖_2≤Tolernc
```
$$\eta_{S,j}^{*} = \eta_{S,j}$$
```
   Break For
   End If
   If iter = MaxIter
```
$$\eta_{S,j}^{*} = \eta_{S,j}$$
```
   End If
End for
```

V. Numerical analysis

In this section, the effectiveness of the proposed "CMAB- FB" algorithm is evaluated through comprehensive Monto-Carlo

simulations. A post-disaster area of 1 Km^2 is considered, which is divided into 36 grids, each with a varying number of users. The altitude of the LEO-Sat is set at 550 kilometers, and the UAV are flying at 100 meters in height. It is assumed that the LEO-Sat has full coverage of the post-disaster area. The Tx power of the LEO-Sat is set to 10 watts, while the Tx power of the UAV and UE are set to 1 watt. The total bandwidth available to the LEO-Sat is 100 MHz, and the bandwidth assigned to the UAV is 40 MHz. Additional simulation parameters are listed in Table 1.

As there are no comparable schemes exist in literature that address the same problem, the performance of the proposed "CMAB-FB" algorithm is compared with two naïve benchmarks: random "Rand" selection and nearest "Nearest" selection. In the random selection method,

Table 1. Simulation parameters.

Parameter	Value
η_{Smax}, Γ_{Umax}	100 MHz [28], 10 KJ
K_i	Uniformly random in the range [1,100]
f_U, f_S	2 GHz [27], 2.4 GHz [28]
σ^2(dBm)	-174 + 10log$_{10}(W)$ + 10 [35]
P_f, P_h	4, 2 Watt [16] [25]
τ and ε	1000 and 1.38×10^{-23} [28]
ρ^{LOS} and ρ^{NLOS}	0.1 dB and 21 dB [27]
UE data load	5 Gbit
T, Ω	1000, 0.01
$G_{j_i}^{Tx} G_S^{Rx}$	15 dB [28]
a, b	4.88, 0.429 [27]

https://doi.org/10.1371/journal.pone.0290432.t001

grids are randomly chosen by GBS with the constraint that each UAV can only serve one grid at a time. The LEO-Sat bandwidth $B_{S,j}^{*t}$ is also randomly distributed across [1, total Sat bandwidth/N]. In the nearest selection method, the GBS chooses the closest grid for each UAV to serve, with the same constraint that each UAV can only serve one grid at a time. Moreover, the LEO-Sat bandwidth $B_{S,j}^{t}$ is equally divided among UAVs.

Fig 2 shows the average total system rate in Gbps against the number of UAVs. The total system rates of all compared schemes are increased as we increase the number of UAVs due to covering more grids at a time. The proposed "CMAB-FB" scheme has the best performance among the compared schemes due to its objective of maximizing the sum of average UAVs rates at each time step. Additionally, it is noteworthy that the "Nearest" scheme has better performance compared to the "Rand" scheme, because in the nearest selection, the LEO-Sat bandwidth is shared equally among the UAVs, whereas in the random selection, the LEO-Sat bandwidth is randomly assigned to each LEO-UAV link. At $N = 2$, the proposed "CMAB-FB" scheme outperforms "Nearest" and "Rand" schemes by 1.327 and 1.5 times, respectively. These values become 1.32 and 2 times at $N = 14$, respectively.

Fig 3 illustrates the relationship between the number of UAVs and the average total energy consumption of UAVs. According to this figure, the "Rand" method has the highest energy consumption performance due to its random grid selection. The "Nearest" method, on the other hand, consistently chooses the closest grids to the UAVs and thus has the lowest energy consumption performance. The proposed "CMAB-FB" method, with its focus on UAVs energy consumption minimization, performs similarly to the "Nearest" method. It is worth mentioning that all the methods have comparable energy efficiency performance due to the constraint that only one UAV can cover one grid at a
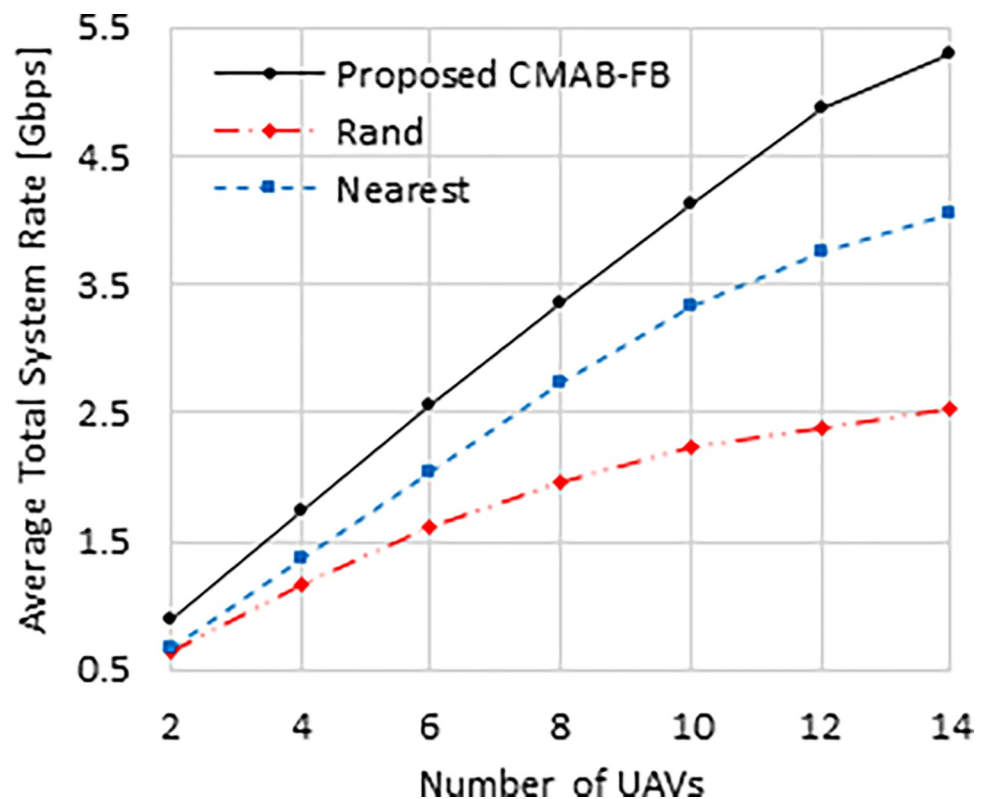

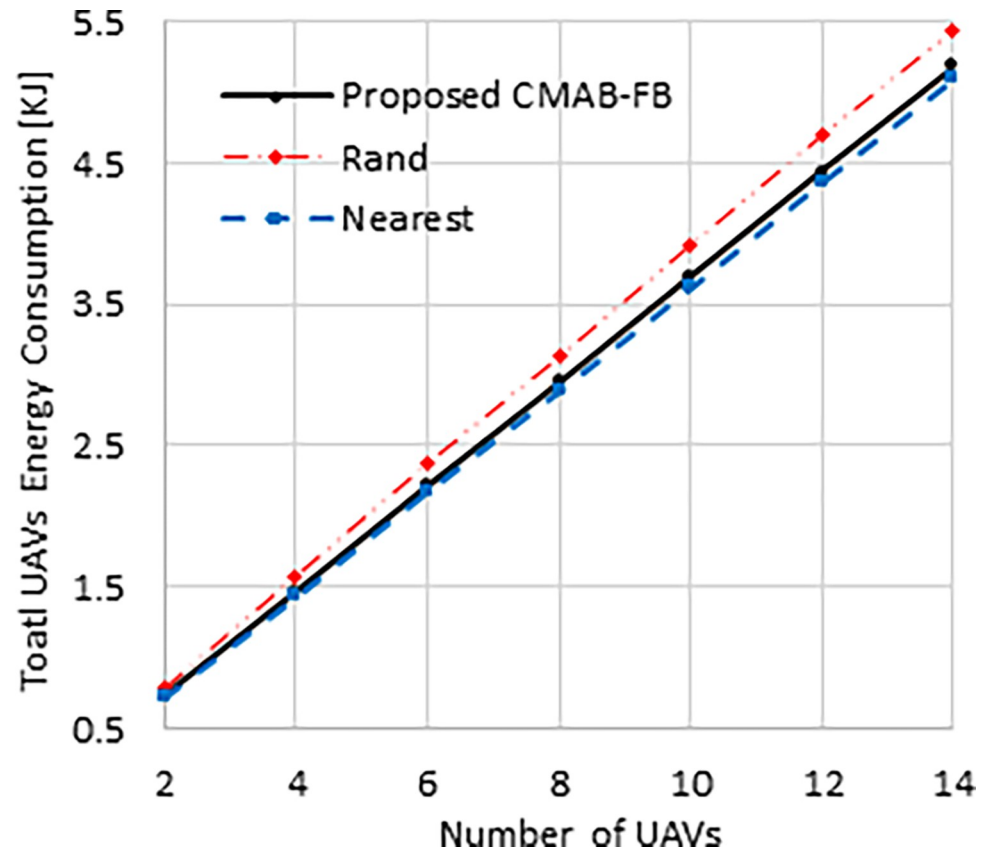**Fig 2. Average total system rate against number of UAVs.**

Fig 3. Energy consumption against number of UAVs.

time. This means that a UAV may have to choose its second or another nearest grid if its closest grid is already being covered by another UAV. Despite this constraint, the proposed "CMAB-FB" method still has performance comparable to the "Nearest" method. At N= 14, both the "CMAB-FB" and the "Nearest" schemes show better energy efficiency performance than the "Rand" method by 5% and 7%, respectively.

Fig 4 displays the fairness index in grids selection, which is calculated as:

$$\chi = \sum_{i=1}^{M} \text{abs}| \frac{h_i^{\text{T}}}{\sum_{i=1}^{M} h_i^{\text{T}}} - \delta_i|, \tag{14}$$

where the term $\left(\frac{h_i^{\text{T}} h_i(T)}{\sum_{i=1}^{M} h_i^{\text{T}} h_i(T)} \right)$ indicates the number of times grid i is selected relative to the total number of grids selections. A value of χ close to 0 indicates that grids are selected based on their UE densities, as determined by the value of δ_i. From Fig 4.

The proposed "CMAB-FB" scheme has the lowest values of χ, which demonstrates its ability to distribute UAVs over grids fairly based on their UEs density. The "Nearest" scheme has the worst fairness performance as it always selects the nearest grids, while "Rand" scheme has better fairness performance than "Nearest" scheme as it selects grids uniformly. This is the reason why χ remains constant, regardless of the number of UAVs tested, in the "Rand" scheme. However, using "Nearest" scheme, χ tends to decrease as we increase the number of UAVs because many UAVs will be better distributed among their nearest available grids. Also, as the
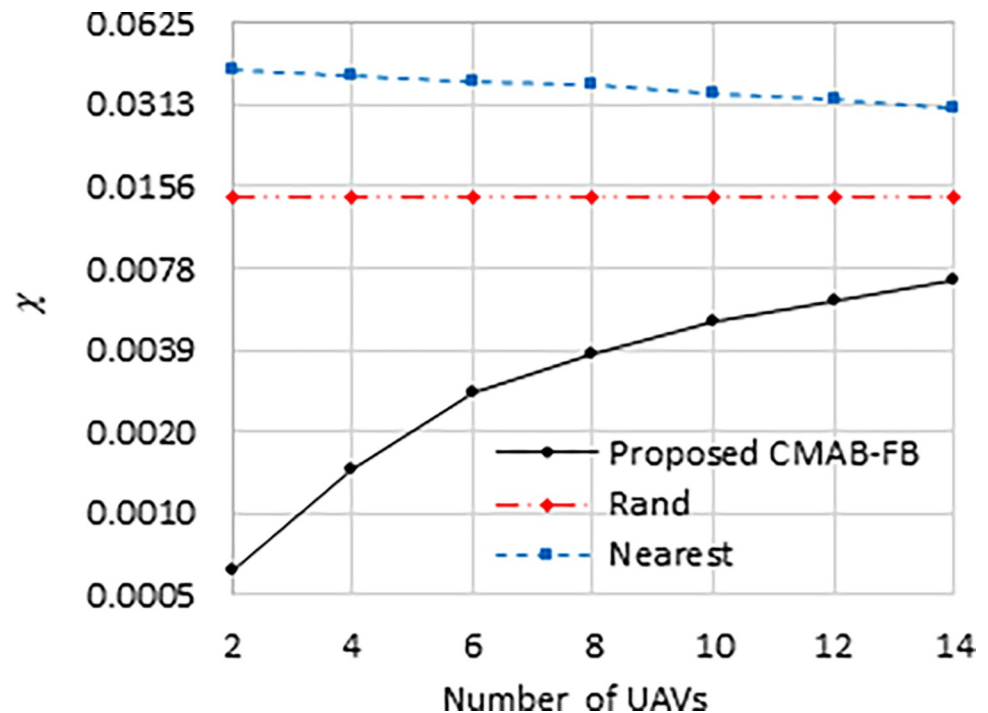
Fig 4. Grids selection fairness index against number of UAVs.

number of UAVs grows, the χ values of all schemes tend to become more like one another due to the reduced number of grid groups available for selection at each iteration. At $N = 2$, the proposed "CMAB-FB" scheme has lower χ values than "Nearest" and "Rand" schemes by 70 and 24 times, respectively. These values become 4.2 and 2.14 at $N = 14$, respectively.

The utilization ratio of the LEO-SAT bandwidth in the compared schemes is displayed in Fig 5. This figure demonstrates that the proposed "CMAB-FB" scheme optimizes the LEO-Sat bandwidth utilization. With a low number of UAVs, only a small portion of the LEO-Sat bandwidth is used relative to their traffic needs, but with a high number of UAVs, a larger portion of the LEO-Sat bandwidth is utilized. The "Nearest" scheme has a fixed utilization ratio of one, as the LEO-Sat bandwidth is equally divided among UAVs, regardless of their number or traffic needs. The "Rand" scheme has a fixed utilization of 0.5, as the bandwidth is uniformly distributed in the range [1, total Sat bandwidth/N]. With a high increase in the number of UAVs, the proposed "CMAB-FB" scheme approaches a utilization ratio of one, as the UAVs' traffic needs require the full utilization of the LEO-Sat bandwidth.

The computational complexity of the proposed "CMAB-FB" scheme composed of three components. The first one comes from minimizing energy consumption with computational complexity of $\mathcal{O}(N^2)$, the second one comes from the sorting operation with computational complexity of $\mathcal{O}(N)$, and the third one comes from the gradient decent algorithm used to optimize LEO-Sat bandwidth resources with computational complexity of $\mathcal{O}(N)$. Therefore, the computational complexity of the proposed algorithm mainly depends on the square of the number of UAVs. Compared to the other benchmarks, the computational complexity of the rand selection is of order $\mathcal{O}(N)$ as it will generates N random numbers in the range of [1, M]. Also, the computational complexity of the nearest selection is of order $\mathcal{O}(N)$ as each UAV will select its nearest grid.
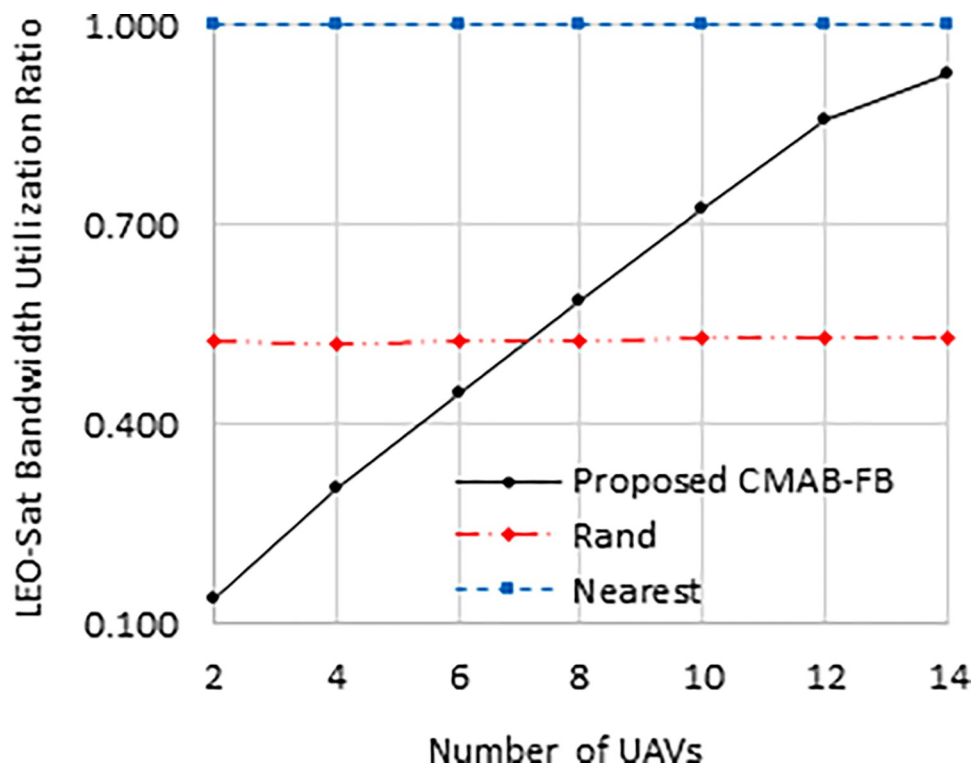
Fig 5. LEO-Sat bandwidth utilization ratio against number of UAVs.

VI. Conclusion

In conclusion, this study has focused on the distribution of UAVs in a post-disaster area with the help of LEO-Sat. The goal was to achieve a balance between maximizing the total sum rate of UAVs and ensuring fairness in the coverage of post-disaster grids based on their UEs density. This is done subject to the limited energy and bandwidth resources of UAVs and LEO-Sat. To address this challenge, we have deemed it as a combinatorial MAB with fairness and budget constraints, and we have proposed the "CMAB-FB" algorithm to solve it efficiently. The study also has proposed a way to optimize the LEO-Sat bandwidth based on UAVs' traffic needs. The results of the numerical analysis have demonstrated that the proposed "CMAB-FB" scheme has outperformed other naïve benchmark approaches.

Author Contributions

Conceptualization: Ehab Mahmoud Mohamed.

Data curation: Ehab Mahmoud Mohamed.

Formal analysis: Ehab Mahmoud Mohamed.

Funding acquisition: Ehab Mahmoud Mohamed.

Investigation: Ehab Mahmoud Mohamed.

Methodology: Ehab Mahmoud Mohamed.

Project administration: Ehab Mahmoud Mohamed.

Resources: Ehab Mahmoud Mohamed.

Software: Ehab Mahmoud Mohamed.

Supervision: Ehab Mahmoud Mohamed.

Validation: Ehab Mahmoud Mohamed.

Visualization: Ehab Mahmoud Mohamed.

Writing – original draft: Ehab Mahmoud Mohamed.

Writing – review & editing: Ehab Mahmoud Mohamed.

References

1. Mozaffari M., Saad W., Bennis M., Nam Y., and Debbah M., "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Communications Surveys and Tutorials*, vol. 21, no.3, pp. 2334–2360, 2019.

2. Mohamed E. M., Hashima S., and Hatano K., "Energy Aware Multi-Armed Bandit for Millimeter Wave Based UAV Mounted RIS Networks," *IEEE Wireless Communications Letters*, vol. 11, no. 6, pp. 1293–1297, 2022.

3. Leyva-Mayorga I., and et al, "LEO small-satellite constellations for 5G and beyond-5G communications," *IEEE Access*, vol. 8, pp. 184 955–184 964, 2020.

4. Li F., Liu J. and Ji B., "Combinatorial sleeping bandits with fairness constraints," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 3, pp. 1799–1813, 2020.

5. Ruder S., "An overview of gradient descent optimization algorithms" in Arxiv, 2016, Online: https://arxiv.org/abs/1609.04747.

6. Auer P., Cesa-Bianchi N., and Fischer P., "Finite-time Analysis of the multiarmed bandit problem," *Machine Learning* vol. 47, pp. 235–256, 2002. https://doi.org/10.1023/A:1013689704352.

7. Audibert J-Y., Munos R., and Szepesvári C., "Exploration–exploitation tradeoff using variance estimates in multi-armed bandits," *Theoretical Computer Science*, vol. 410, no. 19, 2009.

8. M. Erdelj, E. Natalizio, "UAV-assisted disaster management: Applications and open issues,". In Proceedings of the 2016 International Conference on Computing, Networking and Communications (ICNC), Kauai, HI, USA, 15–18 February 2016; pp. 1–5.

9. Merwaday A., Tuncer A., Kumbhar A. and Guvenc I., "Improved Throughput Coverage in Natural Disasters: Unmanned Aerial Base Stations for Public-Safety Communications," in *IEEE Vehicular Technology Magazine*, vol. 11, no. 4, pp. 53–60, Dec. 2016, https://doi.org/10.1109/MVT.2016.2589970

10. S. Lee, D. Har, and D. Kum, "Drone-assisted disaster management: finding victims via infrared camera and lidar sensor fusion," In Proceedings of the 2016 3rd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), Nadi, Fiji, 4–6 December 2016; pp. 84–89.

11. A. Rivera, A.Villalobos, J. Monje, J. Mariñas, and C. Oppus, "Post-disaster rescue facility: Human detection and geolocation using aerial drones," In *Proceedings of the 2016 IEEE Region 10 Conference (TENCON)*, Singapore, 22–26 November 2016; pp. 384–386.

12. T. Kobayashi, H. Matsuoka, and S. Betsumiya, "Flying communication server in case of a largescale disaster," In *Proceedings of the 2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, Atlanta, GA, USA, 10–14 June 2016; vol. 2, pp. 571–576.

13. J. Sánchez-García, J. García-Campos, S. L. Toral, D. G. Reina, and F. Barrero, "A Self organizing aerial ad hoc network mobility model for disaster scenarios," In *Proceedings of the 2015 International Conference on Developments of E-Systems Engineering (DeSE),* Dubai, United Arab Emirates, 13–14 December 2015; pp. 35–40.

14. Ejaz W., Ahmed A., Mushtaq A., and Ibnkahla M., "Energy-efficient task scheduling and physiological assessment in disaster management using UAV-assisted networks,". In *Computer Communications*, vol. 155, pp.150–157, 2020.

15. Zhan P., Yu K. and Swindlehurst A. L., "Wireless relay communications with unmanned aerial vehicles: performance and optimization," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 3, pp. 2068–2085, July 2011, https://doi.org/10.1109/TAES.2011.5937283

16. Mohamed E. M., Hashima S., Aldosary A., Hatano K., and Abdelghany M. A., "Gateway selection in millimeter wave UAV wireless networks using multi-player multi-armed bandit. *Sensors*, vol. 20, no. 3947, 2020 https://doi.org/10.3390/s20143947 PMID: 32708559

17. Amrallah A., Mohamed E.M., Tran G. K., and Sakaguchi K., "Enhanced dynamic spectrum access in UAV wireless networks for post-disaster area surveillance system: A multi-player multi-armed bandit approach," *Sensors*, vol. 21, no. 7855, 2021 https://doi.org/10.3390/s21237855 PMID: 34883856

18. Amrallah A., Mohamed E. M., Tran G. K., and Sakaguchi K., "Optimization of UAV 3D trajectory in a post-disaster area using dual energy-aware bandits," In *IEICE Communications Express*, 2023 https://doi.org/10.1587/comex.2023TCL0015.

19. Singya P. K. and lim Alouini M. S., "Performance of UAV assisted multiuser terrestrial-satellite Communication system over mixed FSO/RF channels," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 2, pp. 781–796, 2021.

20. Ma T. and et al., "UAV-LEO integrated backbone: A ubiquitous data collection approach for B5G internet of remote things networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 11, pp. 3491–3505, 2021.

21. Tran D. -H., Chatzinotas S. and Ottersten B., "Satellite- and cache-assisted UAV: A joint cache placement, resource allocation, and trajectory optimization for 6G aerial networks," *IEEE Open Journal of Vehicular Technology*, vol. 3, pp. 40–54, 2022.

22. Elgamoudi, Benzerrouk H., Elango G. A. and Landry R., "Gauss Hermite $H\infty$ Filter for UAV Tracking Using LEO Satellites TDOA/FDOA Measurement—Part I," in *IEEE Access*, vol. 8, pp. 201428–201440, 2020, https://doi.org/10.1109/ACCESS.2020.3032825

23. Gu S., Wang Y., Wang N. and Wu W., "Intelligent optimization of availability and communication cost in satellite-UAV mobile edge caching system with fault-tolerant codes," in *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 4, pp. 1230–1241, Dec. 2020, https://doi.org/10.1109/TCCN.2020.3005921

24. Han C., Bai L., Bai T. and Choi J., "Joint UAV deployment and power allocation for secure space-air-ground communications," in *IEEE Transactions on Communications*, vol. 70, no. 10, pp. 6804–6818, Oct. 2022, https://doi.org/10.1109/TCOMM.2022.3203471

25. Lin Y., Wang T. and Wang S., "UAV-assisted emergency communications: an extended multi-armed bandit perspective," in *IEEE Communications Letters*, vol. 23, no. 5, pp. 938–941, May 2019, https://doi.org/10.1109/LCOMM.2019.2906194

26. Zhao C., Liu J., Sheng M., Teng W., Zheng Y. and Li J., "Multi-UAV trajectory planning for energy-efficient content coverage: A decentralized learning-based approach," in *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 10, pp. 3193–3207, Oct. 2021, https://doi.org/10.1109/JSAC.2021.3088669

27. Sabzehali J., Shah V. K., Fan Q., Choudhury B., Liu L. and Reed J. H., "Optimizing number, placement, and backhaul connectivity of multi-UAV networks," *IEEE Internet of Things Journal*, vol. 9, no. 21, pp. 21548–21560, 2022.

28. Jia Z., Sheng M., Li J., Niyato D. and Han Z., "LEO-satellite-assisted UAV: Joint trajectory and data collection for internet of remote things in 6G aerial access networks," in *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9814–9826, 2021.

29. Abeywickrama H. V., Jayawickrama B. A., He Y. and Dutkiewicz E., "Comprehensive energy consumption model for unmanned aerial vehicles, based on empirical studies of battery performance," in *IEEE Access*, vol. 6, pp. 58383–58394, 2018, https://doi.org/10.1109/ACCESS.2018.2875040

30. Hashima S., Hatano K., Takimoto E. and Mahmoud Mohamed E., "Neighbor discovery and selection in millimeter wave D2D networks using stochastic MAB," in *IEEE Communications Letters*, vol. 24, no. 8, pp. 1840–1844, Aug. 2020, https://doi.org/10.1109/LCOMM.2020.2991535

31. Mohamed E. M., Alnakhli M., Hashima S., and Abdel-Nasser M., "Distribution of multi mmWave UAV mounted RIS using budget constraint multi-player MAB,". *Electronics* vol. 12, no.12, 2023, https://doi.org/10.3390/electronics12010012

32. Chen W., Wang Y., Yuan Y., and Wang Q., "Combinatorial multi-armed bandit and its extension to probabilistically triggered arms," In *The Journal of Machine Learning Research*, vol. 17, no. 1, pp 1746–1778, 2017.

33. B. Wu, T. Chen and X. Wang, "A Combinatorial bandit approach to UAV-aided edge computing," *2020 54th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, 2020, pp. 304–308, doi: 10.1109/IEEECONF51394.2020.9443306.

34. Wu B., Chen T., Ni W. and Wang X., "Multi-agent multi-armed bandit learning for online management of edge-assisted computing," in *IEEE Transactions on Communications*, vol. 69, no. 12, pp. 8188–8199, Dec. 2021, https://doi.org/10.1109/TCOMM.2021.3113386

35. Mohamed E. M., Sakaguchi K., and Sampei S., "Wi-Fi coordinated WiGig concurrent transmissions in random access scenarios," in *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10357–10371, 2017.