On Enhancing WiGig Communications With A UAV-Mounted RIS System: A Contextual Multi-Armed Bandit Approach

Sherief Hashima[*1], Ehab Mahmoud Mohamed[†2], Kohei Hatano[*‡3], Eiji Takimoto[‡4], Mostafa M. Fouda[§¶5], and Zubair Md Fadlullah[‖6]

[*]Computational Learning Theory Team, RIKEN-AIP, Fukuoka, 819-0395, Japan.
[†]College of Engineering, Prince Sattam Bin Abdulaziz University, Wadi Aldwaser 11991, Saudi Arabia.
[‡]Department of Informatics, Kyushu University, Fukuoka 819-0395, Japan.
[§]Department of Electrical and Computer Engineering, Idaho State University, Pocatello, ID 83209, USA.
[¶]Center for Advanced Energy Studies (CAES), Idaho Falls, ID, USA.
[‖]Department of Computer Science, Western University, London, ON N6A 3K7, Canada.
Emails: [1]Sherief.hashima@riken.jp, [2]Ehab_mahmoud@aswu.edu.eg,[3]kohei.hatano@riken.jp, [4]eiji@inf.kyushu-u.ac.jp, [5]mfouda@ieee.org, [6]zfadlullah@ieee.org,

***Abstract*—Recently emerging WiGig systems experience limited coverage and signal strength fluctuations due to strict line-of-sight (LoS) connectivity requirements. In this paper, we address these shortcomings of WiGig communication by exploiting two emerging technologies in tandem, namely the reconfigurable intelligent surface (RIS) and unmanned aerial vehicles (UAVs). In ultra-dense traffic sites (referred to as hotspots) where WiGig nodes or User Devices (UDs) experience complex propagation and non-line-of-sight (non-LoS) environment, we envision the deployment of a UAV-mounted RIS system to complement the WiGig base station (WGBS) to deliver services to the UDs. However, commercially available UAVs have limited energy (i.e., constrained flight time). Therefore, the trajectory of our considered UAV needs to be locally estimated to enable it to serve multiple hotspots while minimizing its energy consumption within the WGBS coverage boundaries. Since this tradeoff problem is computationally expensive for the resource-constrained UAV, we argue that sequential learning can be a lightweight yet effective solution to locally solve the problem with a low impact on the available energy on the UAV. We formally formulate this problem as a contextual multi-armed bandit (CMAB) game. Then, we develop the linear randomized upper confidence bound (Lin-RUCB) algorithm to solve the problem effectively. We regard the UAV as the bandit learner, which attempts to maximize its attainable rate (i.e., the reward) by serving distinct hotspots in its trajectory that we treat as the arms of the considered bandit. The context is defined as the hotspots' locations provided using GPS (global positioning system) service and the reward history of each hotspot. Our proposal accounts for the energy expenditure of the UAV in moving from one hotspot to another within its battery charge lifetime. We evaluate the performance of our proposal via extensive simulations that exhibit the superiority of our proposed Lin-RUCB algorithm over benchmarking methods.**

***Index Terms*—WiGig, RIS, UAV, MAB, Lin-RUCB.**

I. Introduction

To cope with the recent explosion of content-rich applications and services demanding large bandwidth, WiGig emerged as a promising technology by allowing Wireless Devices (WDs) to access the non-congested 60 GHz frequency spectrum with significantly wide channels to achieve multi-gigabit per second speeds. Despite this superior capacity, WiGig suffers from constrained coverage area due to stringent line-of-sight (LoS) requirements between the transmitting and receiving nodes [1], [2]. Consequently, WiGig technology does not scale well beyond indoor settings due to intricate propagation and path loss phenomena [2], [3]. This paper aims to enhance the WiGig communication system with unmanned aerial vehicle (UAV)-mounted reconfigurable intelligent surface (RIS). Both UAV and RIS technologies have advanced tremendously in recent years. Their combined exploitation could unlock a unique capability for enhancing WiGig communication coverage and signal strength in complex and challenging environmental settings. Inspired by the recent work in [4]–[6], we anticipate that such a coupling between UAV and RIS technologies may facilitate efficient WiGig communication and data transfer across a range of applications from emergency response and industrial operations to urban connectivity and precision agriculture [4], [5]. Since WiGig communications primarily rely on LoS links that result in poor coverage, beamforming training (BT) extends its communication range. This can be implemented by directing the beams of the WiGig base station (WGBS) towards the RIS board, which adjusts its phase shifts (PSs) to direct the beam to the WiGig User Device (UD) [7], [8].

In order to improve the WiGig communication coverage by providing communication links to non-LoS users via the UAV-mounted RIS, in this paper, we describe the problem of the route planning of the UAV. Our formulated optimization problem aims to maximize the potential data rate across various UD-hotspots while taking into factor the flight cost of the UAV in terms of its energy expenditure. Herein, the UAV is assumed to have knowledge of the UDs location via GPS service. However, the UAV is considered to be unaware of their respective traffic requirements prior to providing WiGig communication links to the hotspots. Additionally, the lack of interactivity of the RIS hinders the investigation of the correlation between WGBS, UAV-mounted RIS, and UDs. To

address this problem, we utilize a contextual bandit, a self-learning system [9], [10]. Contextual bandits are special multi-armed bandit (MAB) types, in which the player maximizes a long-term payoff by attempting the accessible bandit arms within a pre-determined duration by utilizing supplemental information known as the context [11]. We justify the adoption of CMAB for addressing the UAV-mounted RIS problem in this paper due to its advantage in overcoming the complexity of estimating CSI and the inherent passivity of the RIS elements.

Based on our considered UAV-mounted RIS system, we regard the UAV as the CMAB player, which attempts to maximize its profit via serving various UD hotspots, i.e., the bandit arms. However, this bandit game is limited by the battery capacity (residual energy) of the UAV. We capture this tradeoff as a CMAB problem and propose a linear randomized upper confidence bound(Lin-RUCB) algorithm to effectively solve the UAV-mounted RIS path planning problem. Based on extensive computer-based simulations, we evaluate the performance of our proposal. Our conducted simulation results clearly demonstrate the superior performance of our proposal over benchmarks and comparable methods.

The remainder of our paper is organized as follows. The relevant research work are surveyed in section II. Next, section III details our considered system model. Section IV formulates the UAV-mounted RIS route planning optimization problem. Our proposed Lin-RUCB algorithm is presented in section V. Section VI presents simulations-based comparisons of Lin-RUCB with comparable methods. Finally, section VII entails concluding remarks and future research directions.

II. Related Work

Recently a number of researchers attempted to employ RIS to improve the coverage area of WiGig communications. In particular, researchers in [12] leveraged stochastic geometry to analyze the possible coverage improvement via RIS-enhanced WiGig communications. The work in [13] investigated WiGig communications with random blockers utilizing multiple RIS panel structures. Subsequently, the authors of [14], [15] addressed the issue of extensive RIS-enhanced WiGig channel estimation by exploiting its cascaded setup. In [16], the RIS's PSs and the hybrid precoding metrics of the WGBS are configured under the assumption of perfect channel state information (CSI).

Although a considerable amount of research has examined the use of RIS-enabled WiGig communications, only a few research works have focused on its potential to enhance WiGig communications through UAV-mounted RIS. However, the combination of UAV and RIS was carried out in the existing research work without considering the need for the path optimization of the resource-constrained UAV to serve multiple groups of UDs, referred to as hotspots, with the minimum possible energy expense. For instance, a deep reinforcement learning (DRL) algorithm was presented in [17] to improve the performance of UAV-mounted RIS systems supporting WiGig communications. However, no consideration on the UAV's path optimization or energy expense was made in that work. Similarly, the work in [18] developed another DRL algorithm to facilitate WiGig-UAV communications. Again, they focused on placing the RIS panels on the WiGig base stations. In contrast, we hypothesize mounting RIS on the UAV will yield a better linkage performance due to its mobility and maneuverability.

Concerning the work mentioned above combining RIS and UAV technologies to complement WiGig communications, neural networks that require offline, supervised training is typically used. This is evident from the recent work in [19] that attempted to combine UAV and RIS operating at even higher frequency bands above 60 GHz by training a deep learning model for predicting beam patterns. Training such models also require acquiring a robust dataset with exhaustive network conditions, such as channel state information (CSI). Researchers in [10], [20] realized this issue and indicated that nested, two-stage bandit algorithms might yield localized solutions for RIS-assisted WiGig systems. Also, in [21], [22], the authors utilized bandits for RIS relay probing in WiGig communications. A reinforcement learning (RL)-aided solutions for UAV-NOMA data offloading in B5G mmWave-enabled communications were investigated in [23]. Furthermore, RIS load balancing and user association using multiplayer bandits were proposed in [24]. However, the above work still did not consider the UAV-mounted RIS scenario.

On the other hand, coauthors of our work investigated the RIS-enabled UAV in an earlier research work [7], [8]. They developed the system preliminaries for this paper that are essential to formulate a multi-armed bandit (MAB) game for the UAV path estimation with energy minimization while designing the relevant side information, i.e., a context, for finding an accurate solution to the formulated problem. To the best of our knowledge, this paper is the first research work to propose a contextual multi-armed bandit (CMAB) game to represent the aforementioned problem and design an algorithm to find a solution that converges fast by making use of the context information, so that the performance and coverage of the WiGiG communication can be improved via our considered UAV-mounted RIS system.

III. System Model

In this section, we describe our considered system model with the aid of Fig. 1, which illustrates a WiGig-enhanced communication network supported by UAV-mounted RIS. The figure shows multiple UD-hotspots are dispersed in the WGBS coverage area. The capacity of UDs per hotspot is adjustable depending on the traffic requirements. It is worth noting that the UAV-mounted RIS significantly improves the coverage of the WGBS, particularly at the remote hotspots (toward the edge of the portrayed area in the figure) that do not enjoy LoS connectivity with the WGBS.

As Fig. 1 illustrates, the UDs are connected to the WGBS directly and indirectly via the UAV-mounted RIS. The WGBS utilizes a control channel to tune the RIS panel's PSs and direct its main beam toward the assisted UDs/hotspot. Therefore, a need arises to optimize the UAV-mounted RIS route planning to maximize the accessible data rate while accounting for the

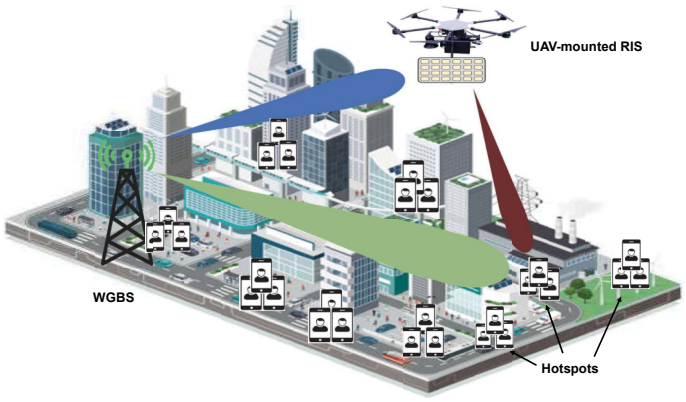


Fig. 1. UAV-mounted RIS-assisted WiGig communication system.

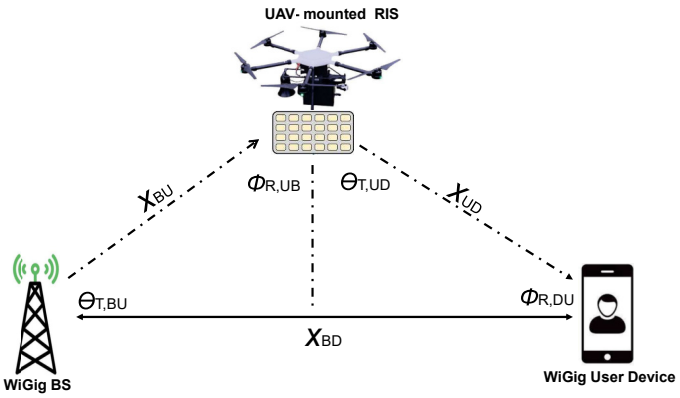


Fig. 2. Simple demonstration of the WGBS, UAV-mounted RIS and UD communication linkages.

UAV's battery power usage. Before this particular problem can be formally treated later in section IV, we provide the system model preliminaries as follows.

First, the received power at a UD can be expressed as:

$$P_R = P_{R,BUD} + P_{R,BD} \tag{1}$$

where $P_{R,BUD}$ and $P_{R,BD}$ denote the indirect power delivered from WGBS via the UAV-RIS and the direct power delivered from the WGBS, respectively.

Next, according to the WiGig-RIS link model presented by [25], $P_{R,BUD}$ can be expressed as:

$$P_{R,BUD} = P_T \left(\frac{\lambda}{4\pi}\right)^4 (Q\gamma)^2 A_{T,BU}\left(\theta_{T,BU}, \theta_{-3dB}\right) \Delta_{R,UB}\left(\phi_{R,UB}\right) \Delta_{T,UD}\left(\theta_{T,UD}\right) A_{R,DU}\left(\phi_{R,DU}, \phi_{-3dB}\right) \left(X_{BU} X_{UD}\right)^{-\alpha}, \tag{2}$$

where, P_T, λ, Q, γ denote the WGBS radiated power, WiGig wavelength, the number of RIS elements number, and the RIS elements' reflection coefficient, respectively [25]. $A_{T,BU}\left(\theta_{T,BU}, \theta_{-3dB}\right)$ indicates the beamforming gain from WGBS to UAV-RIS, while $A_{R,DU}\left(\phi_{R,DU}, \phi_{-3dB}\right)$ refers to the delivered gain from UAV-mounted RIS to UD. θ_{-3dB} and ϕ_{-3dB} are the half power radiating and receiving beamwidths [26], respectively. $\theta_{T,BD}$ is the azimuth angle for the beam between WGBS and UD. $\phi_{R,DU}$ denotes the beam angle between UD and UAV-mounted RIS, as described in Fig. 2. According to this scenario, the UAV-mounted RIS can fly from WGBS and UD once the far-field RIS radiated signal is delivered. Hence, the antenna elements of the RIS panel have an equal gain [25].

By adopting the WiGig antenna layout from WiGig standards [26], a two-dimensional dirigible antenna framework with a Gaussian main loop shape, $A_{T,BU}\left(\theta_{T,BU}, \theta_{-3dB}\right)$, is utilized, which can be expressed as follows,

$$A_{T,BU}\left(\theta_{T,BU}, \theta_{-3dB}\right) = A_0 \; e^{\left(-4ln(2)\left(\frac{\theta-\theta_{T,BU}}{\theta_{-3dB}}\right)^2\right)}, \qquad A_0 = \left(\frac{1.6162}{\sin\left(\frac{\theta_{-3dB}}{2}\right)}\right)^2. \tag{3}$$

On the WiGig UD side, similarly we consider the parameter $A_{R,DU}\left(\phi_{R,DU}, \phi_{-3dB}\right)$, by replacing the parameters $\theta_{T,BU}$ and θ_{-3dB} in eq. (3) by $\phi_{R,DU}$ and ϕ_{-3dB}, respectively. $\Delta_{R,UB}\left(\phi_{R,UB}\right)$ and $\Delta_{T,UD}\left(\theta_{t,UD}\right)$ are the RIS elements' radiation patterns. $\Delta_{R,UB}\left(\phi_{R,UB}\right)$ can be mathematically described as follows [25],

$$\Delta_{R,UB}\left(\phi_{R,UB}\right) = 4\cos\left(\phi_{R,UB}\right). \tag{4}$$

Similar formula can be implemented for $\Delta_{T,UD}\left(\theta_{T,UD}\right)$ with the replacement of $\phi_{R,UB}$ instead of $\theta_{T,UB}$. Regarding $P_{R,BD}$, we leverage the link model provided in [26], [27] for terrestrial WiGig communications, where $P_{R,BU}$ is formulated in eq. (5).

$$P_{R,BD} = P_T A_{T,BD}\left(\theta_{T,BD}, \theta_{-3dB}\right) A_{R,DB}\left(\phi_{R,DB}, \phi_{-3dB}\right) \left(\frac{\eta\left(\mathbb{P}_{LoS}\left(X_{BD}\right)\right)}{L_{LoS}\left(X_{BD}\right)} + \frac{\chi\left(\mathbb{P}_{NLoS}\left(X_{BD}\right)\right)}{L_{NLoS}\left(X_{BD}\right)}\right), \tag{5}$$

where $A_{T,BD}\left(\theta_{T,BD}, \theta_{-3dB}\right)$ and $G_{R,DB}\left(\phi_{R,DB}, \phi_{-3dB}\right)$ are evaluated from eq. (3). $\eta\left(\mathbb{P}_{LoS}\left(X_{BD}\right)\right)$ is the LoS blocking probability and $\chi\left(\mathbb{P}_{NLoS}\left(X_{BD}\right)\right)$ is the NLoS blocking probability of the WiGig link [26]. $L_{LoS}\left(X_{BD}\right)$ and $L_{NLoS}\left(X_{BD}\right)$ reflect the LoS and NLoS path losses, formulated as follows,

$$10\log_{10}\left(L_v(X_{BD})\right) = \beta_v + 10\alpha_v \log_{10}\left(X_{BD}\right) + \varepsilon_v, \tag{6}$$

where $v \in \{LoS, NLoS\}, \beta_v = 82.02 - 10\alpha_v \log_{10}(X_0)$ is the reference five meters path loss. α_v denotes the path loss coefficient, and $\varepsilon_v \backsim \mathcal{N}(0, \delta_v)$ represents the zero mean and δ_v standard deviation log-normal shadowing. Hence, the q^{th} UD's spectral efficiency is formulated as follows.

$$\psi_q = \log_2\left(1 + P_{Rq}/\sigma_0\right), \tag{7}$$

where σ_0 is the noise power.

Based on the system model presented in this section, we now have the preliminaries to formally articulate the actual research problem in the following section.

IV. UAV-mounted RIS Route Planning Problem Formulation

In this section, we present the UAV-mounted RIS route planning problem that targets optimizing the UAV-mounted RIS trajectory that maximizes data rate from serving distinct hotspots while considering the UAV energy boundary over its flight duration. Let S be a set of likelihood spread hotspots, denoted as $\mathcal{S} = \{1, 2, 3, \ldots., s\}$. The UAV trajectory can be expressed mathematically as $j_p = \left\{s_1,\ s_2,\ \ldots..,\ s_{N_{j_p}}\right\}$, where $s_i \in \mathcal{S}$ indicates the hotspot index in the route/path j_p, and $s_{N_{j_p}}$ stands for the last indexed hotspot of the UAV path, i.e., the final served hotspot before the UAV's battery is depleted. Every existing hotspot has random UDs F_{s_i}. Hence, the UAV-mounted RIS route planning optimization can be formulated as follows.

$$\max_{j_p \in \mathcal{J}} \sum_{i=1}^{N_{j_p}} \Psi_{s_i} \tag{8}$$

$$s.t.$$

$$\sum_{i=1}^{N_{j_p}} \Xi_{s_i} \leq \Xi_b$$

$$where \quad \Xi_{s_i} = P_h T_{h_{s_i}} + P_f T_{f_{s_i}},$$

$$\Psi_{s_i} = \sum_{q=1}^{Q_{s_i}} \psi_q,\ T_{h_{s_i}} = \frac{F_{s_i}}{BW\Psi_{s_i}},\ T_{f_{s_i}} = \frac{\left|x_{s_i, s_{i+1}}\right|}{V_f}.$$

Here, $\mathcal{J}$ represents the set of all possible paths, Ψ_{s_i} denotes the spectral efficiency of hotspot $s_i \in \mathcal{S}$, which equals the sum of hotspot's UDs' spectral efficiencies ψ_q, $1 \leq q \leq Q_{s_i}$. Also, let Ξ_{s_i} be the energy depleted/consumed to serve hotspot s_i. The UAV's hovering, and flying powers are denoted as P_h and P_f, respectively. $T_{h_{s_i}}$ determines the hovering duration of the UAV via hotspot k_i, which is the ratio between the k_i's traffic demands, F_{s_i}, and the attained data rate $BW\Psi_{s_i}$. $T_{f_{s_i}}$ is the UAV's duration necessary for its move from hotspot s_i to s_{i+1}, where $\left|x_{s_i, s_{i+1}}\right|$ is the distance between two consecutive hotspots, and V_f is the aerial velocity. The limitation specified in (8) demonstrates that the overall energy depleted by the UAV-mounted RIS throughout its route is constrained by its battery-powered capacity Ξ_b. The UAV should autonomously adapt its route and be adaptable to hotspots' traffic demands. Therefore, the UAV must optimize the data rate of the aided hotspots while extending their battery life through an appropriate selection policy without any CSI calculations.

Given that this problem is computationally hard to solve and the UAV has limited energy and computational resources on board, we seek to find a localized yet efficient solution at the UAV level in the following section.

V. Proposed Lin-RUCB Algorithm

In this section, we consider exploring a sequential algorithm for localized decision-making at the resource-constrained UAV to serve the hotspots to maximize the data rate while minimizing the overall energy expenditure. In this vein, we convert the optimization problem in (8) into a CMAB. In our formulated CMAB game, the UAV acts as the learner/player which progresses through the sequence of trials. It is rewarded with cumulative gains for their choices (choosing arms/ hotspots). At every trial, the learner decides its action based on the context vector (side information) and reward history of the previous rounds. Hence, the learner only collects the reward of the decided arm/hotspot [9]. By using the features to encode the context is derived from supervised machine learning, while exploration is essential for optimizing the learning performance, such as with the reinforcement learning technique. Therefore, the converted problem (i.e., the CMAB game) can be regarded as a commonly accepted compromise between supervised learning and reinforcement learning. The CMAB problem is typically addressed by postulating a linear correlation between the resultant reward and its respective contexts [28], which motivates us to design our algorithmic solution as follows.

As a solution to the aforementioned CMAB game, we design the Lin-RUCB algorithm to efficiently locate the exploration parameter instead of relying on a fixed value like in LinUCB [29]. This approach aims to generate a randomized confidence interval and pick the optimal arm at each process round by adjusting the exploration parameter, $Z_t, Z_1, .. Z_{\mathcal{T}}$, by using a random variable. The value V is determined via the discretization of the interval $[L, U]$ into $N = 100$ points, where Z is calculated as $Z = (L + (U - L)/(N))$ [29].

It is possible to predict the probable reward of an arm by ascertaining a linear relation between the prior rewards of the arm and its existing context vector. Lin-RUCB utilizes the features vector of the existing round and the coefficients and rewards from the previous rounds to anticipate the reward for the current round by interpreting them into a linear combination of the feature vectors observed in prior rounds. At trial t, let $\mathcal{G}_i$ be an $m \times d$ matrix, in which each of the m rows corresponds to the contexts observed for arm/hotspot i. The application of ridge regression to the training data $(\mathcal{G}_i, b_i)$ provides an estimate of the coefficients as follows,

$$\hat{\theta}_i = (\mathcal{G}_i^T \mathcal{G}_i + I_d)^{-1} b_i. \tag{9}$$

It has been demonstrated that when the components of a_i are independent given the corresponding rows of $\mathcal{G}_i$, then $b_i = \mathcal{G}_i^T a_i$, where a_i is an m-dimensional vector containing the past observed rewards of arm/hotspot i.

$$|C_{i,t}^T \theta_{i,t} - E[\psi_{i,t}|C_{i,t}]| \leq Z_{Lin-RUCB} C_{i,t}^T \mathbf{B}_i^{-1} C_{i,t}, \tag{10}$$

where $\mathbf{B}_i = \mathcal{G}_i^T \mathcal{G}_i + I_d$. Hence, the best hotspot is selected as:

$$i_t^* = \arg\max_{i \in A} \ \left(\ j_{i,t}\right), \tag{11}$$

where $j_{i,t} = {X_{i,t}}^T \hat{\theta}_{i,t} + Z_{Lin-RUCB} \sqrt{{C_{i,t}}^T B_i^{-1} C_{i,t}}$

Algorithm 1 highlights the main steps of Lin-RUCB to optimize the UAV-mounted RIS route via a proper selection of the next hotspot using embedded side information which are the hotspot places and previous reward history of each hotspot. Initially when $t = 0$, the selection counter of each hotspot $s_i \in S, W_{s_{i_t}}$ and their corresponding data rates,$\bar{\psi}_{s_{i_t}}$ are

Algorithm 1 Lin-RUCB for UAV-mounted RIS.

Input: $\mathcal{S}$, T_H, Ξ_{s_i}, Ξ_b
Initialization: $t = 0$, $W_{s_{i_t}} = 0$, $\bar{\psi}_{s_{i_t}} = 0$, $\forall s_i \in \mathcal{S}$, $\mathbf{B}_i = I_d$, $b_i = 0_{d\times 1}$.
For t=1,2,...,T_H
For $i = 1, 2, ...\mathcal{S}$
Attempt each hotspot s_{i_t}.
Collect its corresponding rate $\Psi_{s_{i_t}}$ and traffic request $F_{s_{i_t}}$.
Notice features of $\forall s_i \in A : C_{s_{i_t}} \in C \subseteq \mathbb{R}^d$
While $\Xi_{s_{i_t}} \geq \Xi_b$
$\hat{\theta}_i = \mathbf{B}_i^{-1} b_i$
$j_{i,t} = {C_{i,t}}^T \hat{\theta}_{i,t} + Z_{Lin-RUCB}\sqrt{{C_{i,t}}^T \mathbf{B}_i^{-1} C_{i,t}}$
End While
End For
Choose hotspot $i_t^* = \arg\max_i \left(j_{i,t}\right)$ and observe $\psi_{s_i*,t}$
$\mathbf{B}_{i_t^*} = \mathbf{B}_{i_t^*} + C_{i_t^*,t}\ {C_{i_t^*,t}}^T b_{i_t^*} = b_{i_t^*} + \psi_{s_{i_t}^*} C_{s_{i_t}^*,t}$
$\bar{\Psi}_{s_{i_{t+1}}} = \frac{1}{W_{s_{i_{t+1}}}} \sum_{r=1}^{W_{s_{i_{t+1}}}} \Psi_{s_{i_r}}$
Update $W_{k^*,t} = W_{k^*,t-1} + 1$.
Update $\bar{\Psi}_{s_{i_t^*}} = \frac{1}{W_{s_{i_t^*}}} \sum_{r=1}^{W_{s_{i_t^*}}} \Psi_{s_{i_r}}$
End For

set to zero. The algorithm initializes by attempting each hotspot once and collecting its corresponding reward with the side information (i.e., context) that contains the previous rewards and the distance of the UAV from each hotspot. Afterward, during the remaining time horizon, Lin-RUCB's main selection strategy, as in (10), is applied to select the next hotspot. Finally, the algorithm parameters are updated for the next round.

VI. Performance Evaluation

In this section, simulations are conducted to evaluate the performance of the envisioned Lin-RUCB method against other solutions, such as classical UCB (upper confidence bound), random, and nearby selection methods. In UCB, the hotspot corresponding to the highest spectral efficiency is chosen without taking into consideration the UAV battery. The random selection stochastically draws the following hotspot in the UAV route, whereas the nearby election technique always opts for the closest hotspot. Furthermore, the optimum rate performance is provided as the UAV has the complete environmental information and draws the hotspot that offers the maximum accessible rate. The simulation setup places two hundred hotspots uniformly in a 16 Km^2 area. Each hotspot has UDs that are uniformly distributed between 1 and 50. The simulation parameters are listed in Table I. The performance evaluation metrics are the data rate and energy efficiency (EE). In this context, EE of hotspot s_i is equal to Ψ_{s_i}/Ξ_{s_i}.

The data rate and EE performances of the schemes involved in the comparison versus different numbers of hotspots uniformly located within 16 Km^2 simulation area are presented in Figs. 3(a) and 3(b), respectively.

From Fig. 3(a), we can notice that the data rate is proportionally related to the hotspot numbers except for nearby and random methods since they present a fixed data rate performance. The nearby method delivers the worst rate performance

TABLE I
Considered simulation parameters.

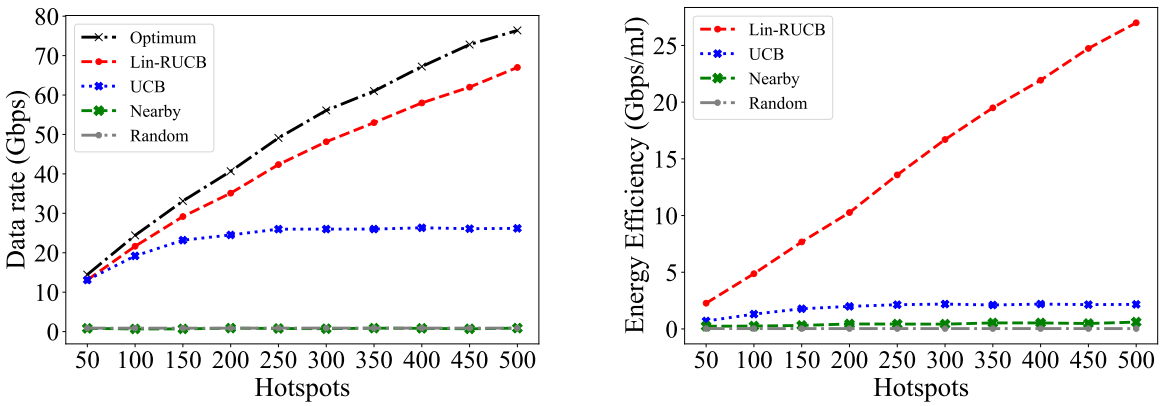
Parameters	Value
P_T, BW	1 Watt, 2.16 GHz
λ, M, Γ	0.005, 64, 0.9
h, T_h	10m,60 sec
θ_{-3dB}, ϕ_{-3dB}	30°
T_H	1000
$\sigma_0(dBm)$	$-174 + 10log10(W) + 10$
R_{k_i}	Uniform [10, 70] Gbit

because of its policy that achieves a low accessible data rate. Our proposed Lin-RUCB algorithm delivers the maximal rate performance, followed by UCB. At 300 hotspots, our proposal (Lin-RUCB), UCB, nearby, and random methods achieve around 86%, 45%, 1.3%, and 1.5% of the ideal selection scheme, accordingly.

Next, Fig. 3(b) exhibits the EE performance evaluation versus varied hotspot numbers. The proposed Lin-RUCB scheme delivers a superior EE performance due to valuable side information (context). Furthermore, note that the nearby election technique outperforms the random one because of its policy of drawing the closest hotspot, hence minimizing the energy cost. At 300 hotspots, the envisioned Lin-RUCB, UCB, nearby and random techniques achieve EE of around 16.7, 2.19, 0.032, and 0.42 Gbps/mJ, respectively.

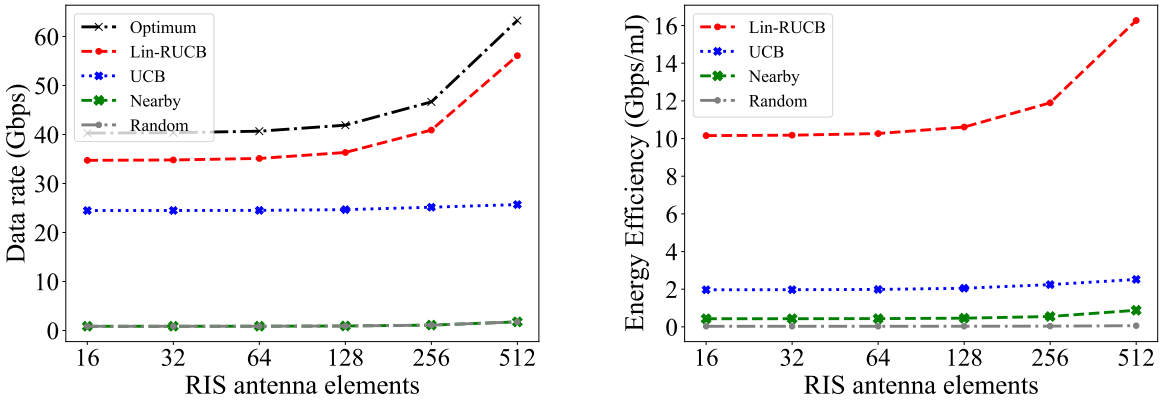
Furthermore, the data rate and EE performances of various methods versus the utilized number of RIS antenna elements Q via 200 randomly spread hotspots within $16Km^2$ are presented in Figs. 4(a) and 4(b), respectively. As Q increases, both data rate and EE performances slightly improve because of the major effect of the LoS link between the WGBS and UD. However, at large numbers of antenna elements 256, 512, the performance is improved due to the increased effect of the UAV-UD path. The proposed Lin-RUCB delivers near-optimal rate performance and the best EE followed by the UCB scheme. At 512 elements, the data rate performance of Lin-RUCB, UCB, nearby, and random techniques attained 88.6%, 40.5%, 2.7%, and 2.6% of the optimum rate performance. Also, they attained EE performance of 16.3, 2.5,0.88, and 0.06 Gbps/mJ, respectively.

The data rate and EE performances of various methods, when simulated across a range of areas from 1 to 100 Km^2 with 128 RIS antenna elements and 200 randomly distributed hotspots, are evaluated in Figs. VI and VI, respectively. In small areas, the high density of UDs leads to increased data rates and, consequently, elevated EEs. The opposite is true for large areas. The proposed Lin-RUCB experiences a superior data rate and EE performances, followed by the classical UCB. Also, the random scheme's data rate exceeds that of the nearby method, and the opposite observation is made regarding the EE performance. This also can be attributed to the previously presented rationale. In a simulated area of 1 km^2, 78.6%, 14%, 4.8%, and 4.7% of the optimum rate performance were achieved using the proposed Lin-RUCB, UCB, nearby, and



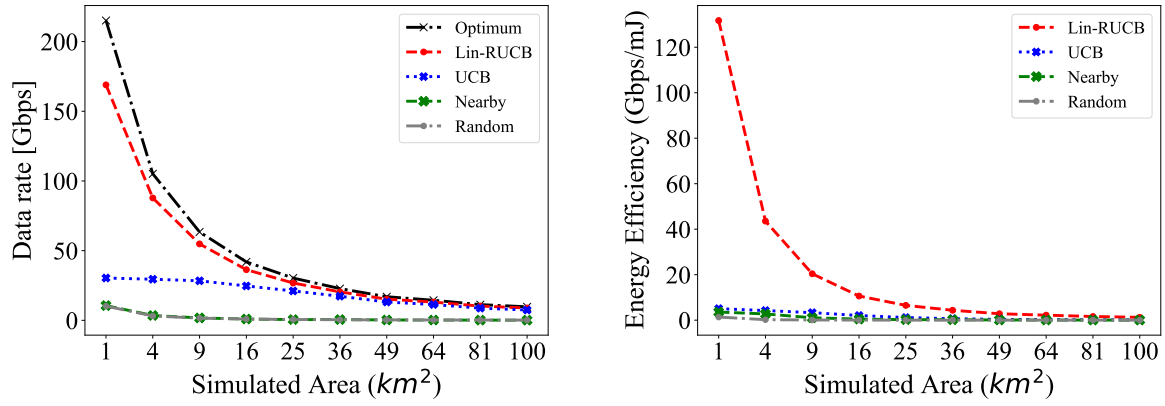
(a) Data rate. (b) Energy efficiency.

Fig. 3. Data rate and EE Comparison of Lin-RUCB against UCB, nearby, and Random schemes at distinct hotspots and 64 RIS elements within 16 Km^2 coverage area.



(a) Data rate. (b) Energy efficiency.

Fig. 4. Data rate and EE Comparison of Lin-RUCB against UCB, nearby, and Random schemes at distinct RIS antenna elements and 200 hotspots within 16 Km^2 coverage area.



(a) Data rate. (b) Energy efficiency.

Fig. 5. Data rate and EE Comparison of Lin-RUCB against UCB, nearby, and Random methods at distinct simulated areas using 200 hotspots and 128 antenna elements.

random methods, respectively. When the simulated area was increased to 81 km^2, these rates became 90%, 77%, 1.3%, and 1.7%, respectively.

In an area of 1 Km^2, the EE of the proposed Lin-RUCB algorithm was superior to that achieved by UCB, nearby, and random selection methods by a factor of 26, 44, and 97, respectively. Furthermore, these values changed to 10.3,41.5, and 520 when the simulation area was set to 81 Km^2.

The computational complexity of the proposed in-RUCB scheme is analogous to the traditional UCB since its major complexity arises from the size of the context vector and the number of probed hotspots/arms as it is in the order of $O\left(d^2K\right)$ [29]. Since d is fixed to 2, the complexity depends mainly on the number of hotspots.

VII. Conclusion

This work investigated the utilization of UAV-mounted RIS to enhance the coverage of WGBS in high-traffic areas. It also presented the rationale behind optimizing the UAV route to serve a large number of hotspots, hence maximizing the data rate and concurrently prolonging the UAV flight time by minimizing the UAV's travel cost in terms of energy expenditure. The optimization problem was reformulated into a contextual multi-armed bandit (CMAB) game, and the Lin-RUCB algorithm was proposed to find a real-time solution to this problem by choosing and serving the appropriate UD-hotspots along the UAV route. Performance evaluations verified the superior performance of the proposed algorithm over benchmark solutions. Future work may include embedding CSI as a context and multiplayer bandit formulation.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Numbers JP21K14162 and JP22H03649, JP19H04067, and JP20H05967 Japan and also supported by the National Science Foundation (NSF) under Award No. 2210252.

References

[1] E. M. Mohamed, S. Hashima, K. Hatano, H. Kasban, and M. Rihan, "Millimeter-wave concurrent beamforming: A multi-player multi-armed bandit approach," *Computers, Materials & Continua*, vol. 65, no. 3, pp. 1987–2007, 2020.

[2] T. K. Rodrigues, S. Verma, Y. Kawamoto, N. Kato, M. M. Fouda, and M. Ismail, "Smart handover with predicted user behavior using convolutional neural networks for WiGig systems," *arXiv preprint arXiv:2303.15731*, 2023.

[3] S. Hashima, K. Hatano, H. Kasban, M. Rihan, and E. M. Mohamed, "Multiagent multi-armed bandit techniques for millimeter wave concurrent beamforming," in *2020 8th IEEE JAC-ECC*, 2020, pp. 56–59.

[4] D. Tyrovolas, P.-V. Mekikis, S. A. Tegos, P. D. Diamantoulakis, C. K. Liaskos, and G. K. Karagiannidis, "Energy-aware design of UAV-mounted RIS networks for IoT data collection," *IEEE Transactions on Communications*, vol. 71, no. 2, pp. 1168–1178, 2023.

[5] J. Luo, T. Liang, C. Chen, and T. Zhang, "A UAV mounted RIS aided communication and localization integration system for ground vehicles," in *2022 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2022.

[6] M.-S. Van Nguyen, D.-T. Do, V.-D. Phan, W. Ullah Khan, A. L. Imoize, and M. M. Fouda, "Ergodic performance analysis of double intelligent reflecting surfaces-aided NOMA–UAV systems with hardware impairment," *Drones*, vol. 6, no. 12, article no. 408, 2022.

[7] E. M. Mohamed, S. Hashima, and K. Hatano, "Energy aware multiarmed bandit for millimeter wave-based UAV mounted RIS networks," *IEEE Wireless Communications Letters*, vol. 11, pp. 1293–1297, 2022.

[8] S. Hashima, K. Hatano, and E. M. Mohamed, "Advanced MAB schemes for WiGig-aided aerial mounted RIS wireless networks," in *2023 IEEE Consumer Communications & Networking Conference (CCNC)*, 2023.

[9] L. Li, W. Chu, J. Langford, and R. E. Schapire, "A contextual-bandit approach to personalized news article recommendation," *ArXiv preprint ArXiv:1003.0146*, 2010.

[10] E. M. Mohamed, S. Hashima, K. Hatano, and S. A. Aldossari, "Two-stage multiarmed bandit for reconfigurable intelligent surface aided millimeter wave communications," *Sensors*, vol. 22, no. 6, article no. 2179, 2022.

[11] A. O. Hashesh, S. Hashima *et al.*, "AI-enabled UAV communications: Challenges and future directions," *IEEE Access*, vol. 10, pp. 92 048–92 066, 2022.

[12] M. Nemati, J. Park, and J. Choi, "RIS-assisted coverage enhancement in millimeter-wave cellular networks," *IEEE Access*, vol. 8, pp. 188 171–188 185, 2020.

[13] Y. Chen, Y. Wang, and L. Jiao, "Robust transmission for reconfigurable intelligent surface aided millimeter wave vehicular communications with statistical CSI," *IEEE Transactions on Wireless Communications*, vol. 21, no. 2, pp. 928–944, 2022.

[14] Y. Liu, S. Zhang, F. Gao, J. Tang, and O. A. Dobre, "Cascaded channel estimation for RIS assisted mWave MIMO transmissions," *IEEE Wireless Communications Letters*, vol. 10, no. 9, pp. 2065–2069, 2021.

[15] J. He, H. Wymeersch, and M. J. Juntti, "Channel estimation for RIS-aided mmWave MIMO systems via atomic norm minimization," *IEEE Transactions on Wireless Communications*, vol. 20, pp. 5786–5797, 2021.

[16] C. Pradhan, A. Li, L. Song, B. Vucetic, and Y. Li, "Hybrid precoding design for reconfigurable intelligent surface aided mmWave communication systems," *IEEE Wireless Communications Letters*, vol. 9, no. 7, pp. 1041–1045, 2020.

[17] Q. Zhang, W. Saad, and M. Bennis, "Reflections in the sky: Millimeter wave communication with UAV-carried intelligent reflectors," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019.

[18] X. Guo, Y. Chen, and Y. Wang, "Learning-based robust and secure transmission for reconfigurable intelligent surface aided millimeter wave UAV communications," *IEEE Wireless Communications Letters*, vol. 10, no. 8, pp. 1795–1799, 2021.

[19] N. Abuzainab, M. Alrabeiah, A. Alkhateeb, and Y. E. Sagduyu, "Deep learning for THz drones with flying intelligent surfaces: Beam and handoff prediction," *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2021.

[20] E. M. Mohamed, S. Hashima, and orhers, "Reconfigurable intelligent surface-aided millimeter wave communications utilizing two-phase minimax optimal stochastic strategy bandit," *IET Communications*, vol. 16, pp. 2200–2207, 2022.

[21] E. M. Mohamed, S. Hashima, K. Hatano, and M. M. Fouda, "Cost-effective MAB approaches for reconfigurable intelligent surface aided millimeter wave relaying," *IEEE Access*, vol. 10, pp. 81 642–81 653, 2022.

[22] S. Hashima, K. Hatano, E. Takimoto, and E. M. Mohamed, "Budgeted Thompson sampling for IRS enabled WiGig relaying," *Electronics*, vol. 12, no. 5, 2023.

[23] A. Gendia, O. Muta, S. Hashima, and K. Hatano, "UAV positioning with joint NOMA power allocation and receiver node activation," in *33rd IEEE PIMRC*, 2022, pp. 240–245.

[24] E. M. Mohamed *et al.*, "Load balancing multi-player MAB approaches for RIS-aided mmWave user association," *IEEE Access*, vol. 11, pp. 15 816–15 830, 2023.

[25] K. Ntontin, D. Selimis, A.-A. A. Boulogeorgos, A. Alexandridis, A. Tsolis, V. Vlachodimitropoulos, and F. Lazarakis, "Optimal reconfigurable intelligent surface placement in millimeter-wave communications," in *2021 European Conference on Antennas and Propagation (EuCAP)*, 2021.

[26] E. M. Mohamed, S. Hashima, K. Hatano, S. A. Aldossari, M. Zareei, and M. Rihan, "Two-hop relay probing in WiGig device-to-device networks using sleeping contextual bandits," *IEEE Wireless Communications Letters*, vol. 10, no. 7, pp. 1581–1585, 2021.

[27] S. Hashima, K. Hatano, and E. M. Mohamed, "Multiagent multi-armed bandit schemes for gateway selection in UAV networks," in *2020 IEEE Globecom Workshops (GC Wkshps)*, 2020.

[28] S. Hashima, K. Hatano, H. Kasban, and E. Mahmoud Mohamed, "Wi-Fi assisted contextual multi-armed bandit for neighbor discovery and selection in millimeter wave device to device communications," *Sensors*, vol. 21, no. 8, article no. 2835, 2021.

[29] S. Vaswani, A. Mehrabian, A. Durand, and B. Kveton, "Old dog learns new tricks: Randomized UCB for bandit problems," in *International Conference on Artificial Intelligence and Statistics*, 2019.