RIS Subset Selection/Assignment Based on Multi-Armed Bandit in Multi-RIS-Aided Communications

Aral Ertug Zorkun
SnT, University of Luxembourg,
Luxembourg, Luxembourg
aral.zorkun@uni.lu

Juan Andres Vasquez Peralvo
SnT, University of Luxembourg,
Luxembourg, Luxembourg
juan.vasquez@uni.lu

Praveen Naidu Vummadisetty
SnT, University of Luxembourg,
Luxembourg, Luxembourg
praveen.vummadisetty@uni.lu

Symeon Chatzinotas
SnT, University of Luxembourg,
Luxembourg, Luxembourg
symeon.chatzinotas@uni.lu

***Abstract*—This paper, proposes a preliminary approach to optimal user-Reconfigurable Intelligent Surfaces (RIS) subset selection/assignment based on Multi-Armed Bandit (MAB) to improve the overall communication quality of wireless systems. Users select optimal RIS subsets based on the observations of previous selections to maximize the Spectral Efficiency (SE) based on instantaneous Channel State Information (CSI). The SE is drawn from Nakagami-*m* distribution and mean. The simulation results show that with the adopted approach faster RIS subset selection/assignment to the users can be achieved.**

Keywords—Multi-armed bandit (MAB), reconfigurable intelligent surfaces (RIS), RIS selection, spectral efficiency.

I. Introduction

The overall Quality of Service (QoS) in Reconfigurable Intellegent Surfaces (RIS)-aided communication systems can be improved by optimum user-RIS selection/assignment in multiple-RIS multiple user scenarios. However, user-RIS selection/assignment is hard to manage since channel estimation, frequent pilot transmission, phase shift optimization of each RIS and the transmission of RIS control signals require high computational time in the system [1].

There are extensive practical studies on reducing the computational time for user-RIS selection/assignment in RIS-aided communication applications. In [1], a power based user selection method is proposed for RIS-aided multi-user massive Multiple-Input Single-Output (MISO). In [2], a user selection strategy is proposed in multi-RIS multiple user communications and lastly, [3] proposes optimum RIS selection method for the wireless networks. In line with these studies, we propose a RIS subset selection method based on Multi-Armed Bandit (MAB) Spectral Efficiency (SE) based on instantenous Channel State Information (CSI) is considered for the selection policy.

II. System Model

This paper considers a downlink multi-RIS-aided multi-user communication systems. There are K_{RIS} RISs in the system serving to K_{UE} User Equipment (UE). Each RIS and UE is equipped with different number of antennas, N_k^{RIS}, and N_k^{UE}, respectively. The system consists of one Base Station (BS) which is equipped with N^{BS} antennas. The BS, RISs and UEs are equipped with Uniform Antenna Arrays (ULA). The RISs are randomly distributed and the locations of RISs are unknown by the central network controller. The direct link, between BS and all UEs are blocked by the obstacles. In each time interval, phase shifts of some RISs may or may not be optimized.

At a given time interval, the transceiver activates a subset of RISs, $k_{RIS} \leq K_{RIS}$. All the possible RIS subsets in a set, S, with cardinality can be expressed as:

$$|S| = \sum_{i=1}^{K_{RIS}} \binom{K_{RIS}}{i} \tag{1}$$

The cascaded channel $\mathbf{H}_n \epsilon \mathbb{C}^{N_k^{UE} \times N^{BS}}$ between BS and the k_{UE}-th UE via k_{RIS}-th RIS can be written as:

$$\mathbf{H}_n = \sum_{l_1=1}^{L_1} \sum_{l_2=1}^{L_2} \beta_{l_1} \alpha_{l_2} \boldsymbol{a}\left(\theta_{l_1}^{UE}\right) \boldsymbol{a}\left(\theta_{l_1}^{RIS}\right)^T . \boldsymbol{\Theta} \boldsymbol{a}\left(\varphi_{l_2}^{RIS}\right) \boldsymbol{a}\left(\varphi_{l_2}^{BS}\right)^T \tag{2}$$

where L_1 and L_2 are the number of scatters between the channels RIS-UE and BS-RIS, respectively, β_{l_1} and α_{l_2} are the path loss coefficients, $\boldsymbol{a}$ is the steering vector, $\theta_{l_1}^{UE}$ and $\varphi_{l_2}^{RIS}$ are the Direction-of-Arrival (DoA) and $\theta_{l_1}^{UE}$ and $\varphi_{l_2}^{RIS}$ Direction-of-Departure (DoD) components, and $\boldsymbol{\Theta} \epsilon \mathbb{C}^{N_k^{RIS} \times N_k^{RIS}}$. Alternating optimization [2] can be adopted for optimizing $\boldsymbol{\Theta}$.

Then, the received signal at the k_{UE}-th UE, $\mathbf{y}_{k_{UE}} \epsilon \mathbb{C}^{N_k^{UE} \times 1}$ can be written as:

$$\mathbf{y}_{k_{UE}} = \mathbf{H}_s \mathbf{x} + \mathbf{n} \tag{3}$$

where $\mathbf{H}_s \epsilon \mathbb{C}^{N_k^{UE} \times N^{BS}}$ is the subset of cascaded channels $s \epsilon S$ and is the sum of the channels formed by the selected RISs, $\mathbf{x} = \left[x_{1,} \ldots, x_{N^{BS}}\right]^T \epsilon \mathbb{C}^{N^{BS} \times 1}$ is the modulated symbols with unitary average transmit power $E[\mathbf{x}\mathbf{x}^H] = 1$ and $\mathbf{n} \epsilon \mathbb{C}^{N_k^{UE} \times 1} \sim \mathcal{CN}\left(0, \sigma^2 \mathbf{I}_{N_k^{UE}}\right)$ is the additive white Gaussian noise (AWGN) with noise power σ^2. The subset of cascaded channel for k_{UE}-th UE where all the RISs are active can be written as:

$$\mathbf{H}_s = \sum_{n=1}^{K_{RIS}} \mathbf{H}_n \tag{4}$$

The achievable spectral efficiency in bps/Hz at the k_{UE}-th UE with an arbitrary subset of RISs, then, can be expressed as:

$$R_s^{k_{UE}} = \log_2 \det\left(\mathbf{I}_{N_k^{UE}} + \frac{\mathbf{H}_s}{\sigma^2}\right) \tag{5}$$

We adopted a MAB RIS subset selection policy. In this sense path loss coefficients are *independent but not identically distributed* (i.n.i.d). In general, RIS-aided communication systems suffer from a heterogenous fading environment, thus, generic Nakagami-m fading channel is more suitable for path loss coefficients rather than Rayleigh or Rician channel models [2]. Then, the variables, p_s, N_k^{RIS} and N_k^{UE} are Nakagami-m RVs. The fading severity and path loss coefficients are distinct. Denoting X ~Nakagami (m, Ω) to represent a random variable (RV) following Nakagami-m distribution, the mean can be computed by:

$$\bar{\theta} = \frac{\Gamma\left(m + \frac{1}{2}\right)}{\Gamma(m)} \left(\frac{\Omega}{m}\right)^{1/2} \quad (6)$$

where $m > 0$ indicates the severity of small-scale fading, $\Omega > 0$ is equivalent to large-scale fading, the Gamma function $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$.

III. RIS SUBSET SELECTION POLICY

In MAB, a set of arms are given as $l = \{1,2, \dots, L\}$ and the player selects an arm at each time t, where $l_t \epsilon l$ and the observe reward is z_{l_t}. At the beginning, the prior information of arms and the rewards are unknown to the player. The goal is to maximize the player's average profit within a certain period of time.

In our approach, UEs are the independent players, channels of the RIS subsets are the arms and achievable spectral efficiencies are the rewards. It should be noted that locations of UEs are randomly distributed, thus, same RIS subset has different rewards for distinct UEs. It is assumed that only the cumulative reward is known where all the RISs are active. Instead of tracking the spectral efficiencies for each subset of RISs, an approximation for the proportional reward gain by the each related RIS channel within a subset is recorded [4].

Due to the limited space, the entire derivation of the adopted MAB approach for RIS subset selection is not given, however, detailed explanation can be found in [4]. Considering $\hat{\theta}_i$ is a counter that tracks the samples mean of the rewards obtained from each RIS, and $\hat{\varphi}_i$ is a counter that tracks the number of times a particular RIS selected. The RIS subset selection is given in **Algorithm 1**.

Algorithm 1 RIS Subset Selection Policy

1: **Inputs:** $\hat{\theta}_i = 0, \hat{\varphi}_i = 0, i = 1, \dots, K_{RIS}$
2: **Initialize:** Play each arm once.
    Update $\hat{\theta}_i$, $\hat{\varphi}_i$
3: **Repeat**
4: $$\arg\max_{s \epsilon S} \sum_{i \epsilon s} \hat{\theta}_i + \sqrt{\frac{(K_{RIS} + 1) \ln t}{\sum_{i \epsilon s} \hat{\varphi}_i}}$$
5: Update $\hat{\theta}_i$, $\hat{\varphi}_i$
6: $t = t + 1$
7: **Until** $t = T$

IV. SIMULATION RESULTS

In this section, we evaluate the proposed RIS subset selection policy. We consider there are $K_{RIS} = 3$ RISs and $K_{UE} = 4$ UEs and the number of total RIS element is $\sum_k^{K_{RIS}} N_k^{RIS} = 600$, The reward distribution (spectral efficiencies) and means are unknown to the UEs at the beginning; however, distributions and means are modeled using (4) and (5). In order to calculate the respective means each RIS subset has different m and Ω for distinct UEs. The respective means for the k_{UE}-th UE can be expressed as:

$$\bar{\boldsymbol{\theta}} = \{0.13, 0.26, 0.44, 0.73, 0.58, 0.86, 0.67\} \quad (7)$$

where $\bar{\boldsymbol{\theta}}_1$ denotes the subset which only contains the first RIS as $s = \{1\}$ and $\bar{\boldsymbol{\theta}}_7$ denotes the subset of all RISs as $s = \{1,2,3\}$. The 6th RIS subset is the optimal choice for the k_{UE}-th UE.

Fig. 1 compares the regrets achieved from the naïve and the proposed MAB solution in [4] for 4000 trails averaged over 100 runs. The optimal RIS subset can be selected faster with the propose approach in [4], thus, the overall QoS can be established and users are assigned to best RIS subsets in a short time.

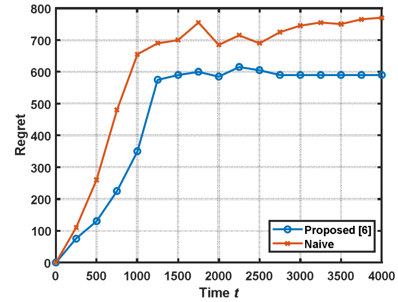

Fig. 1. Regret with respect to time.

V. CONCLUSION

This paper proposed a preliminary approach for optimal user-RIS selection/assignment in multiple-RIS multiple user scenarios. Nakagami-m fading model was adopted to determine the means of SE. An MAB strategy was adopted for the optimal RIS subset selection. It was shown that the fast RIS subset selection can be achieved, thus, user-RIS subset assignment can be performed in a short period of time. Future scopes are to propose a faster and more robust RIS subset selection method and extend the analysis.

ACKNOWLEDGMENT

This work has been supported by the SNS JU TERRAMETA project under EU's Horizon Europe research and innovation program under Grant Agreement number 101097101.

REFERENCES

[1] K. Min and M. Jung, "A power method based user selection for reconfigurable intelligent surface (RIS) assisted multiuser massive MISO systems," ICT Express, p. S2405959522001114, Aug. 2022, doi: 10.1016/j.icte.2022.08.002.

[2] W. Jiang and H. D. Schotten, "User Selection for Simple Passive Beamforming in Multi-RIS-Aided Multi-User Communications," in 2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring), Florence, Italy: IEEE, Jun. 2023, pp. 1–5. doi: 10.1109/VTC2023-Spring57618.2023.10201204.

[3] Y. Fang, S. Atapattu, H. Inaltekin, and J. Evans, "Optimum Reconfigurable Intelligent Surface Selection for Wireless Networks," IEEE Trans. Commun., vol. 70, no. 9, pp. 6241–6258, Sep. 2022, doi: 10.1109/TCOMM.2022.3193152.

[4] A. Mukherjee and A. Hottinen, "Learning Algorithms for Energy-Efficient MIMO Antenna Subset Selection: Multi-Armed Bandit Framework".