Spectrum Sensing-Enhanced Federated Multi-Armed Bandit for Cognitive Radio

Kanghee Lee, Jungmin So[a]

[a]*Department of Computer Science and Engineering, Sogang University, Seoul, South Korea*

Abstract

Cognitive radio (CR) technology addresses the challenge of wireless spectrum scarcity by intelligently adapting to dynamic spectrum availability. The Federated Multi-Armed Bandit (FMAB) algorithm enables collaborative decision-making among distributed cognitive radios, leveraging federated learning principles for efficient spectrum utilization. This paper enhances the FMAB algorithm by introducing uncertainty in frequency band rewards and employing randomized mean rewards, reflecting real-world CR scenarios. Spectrum sensing using energy detection reduces exploration costs by selectively collecting data from idle frequency bands. The effectiveness of the enhanced FMAB algorithm is evaluated in uncertain CR scenarios. A comparison between Fed1 UCB and Fed2 UCB approaches reveals that client cooperation in Fed2 UCB leads to faster convergence and lower regret, outperforming independent client behavior in Fed1 UCB. These findings emphasize the importance of federated learning in the Multi-Armed Bandit framework for CR systems and provide valuable insights for practical deployments.

1. Introduction

Cognitive radio (CR) has emerged as a promising technology to address the increasing demand for wireless spectrum in the face of spectrum scarcity. By intelligently adapting to the dynamic spectrum availability, cognitive radios can opportunistically access underutilized frequency bands, enabling efficient spectrum utilization and enhancing overall network performance.

In the context of cognitive radio systems, the traditional Multi-Armed Bandit(MAB) algorithms face challenges in scenarios where the available channels and their associated rewards can vary across different geographic locations or user contexts. This variation necessitates the integration of federated learning principles into the MAB framework, leading to the development of the Federated Multi-Armed Bandit (FMAB) algorithm.

The FMAB algorithm leverages the power of federated learning to enable collaborative decision-making among distributed cognitive radios. By allowing individual radios to learn and adapt their strategies based on local observations while exchanging information and experiences with other radios, the FMAB algorithm promotes more efficient and effective decision-making in dynamic and heterogeneous spectrum conditions.

Unlike traditional MAB algorithms that rely solely on individual radio observations, the FMAB algorithm harnesses the collective intelligence of the cognitive radio network. By facilitating information exchange and leveraging shared knowledge, the FMAB algorithm enables radios to make more informed decisions about channel selection, leading to improved spectrum access and utilization.

By building upon previous work of FMAB(Chengshuai Shi, 2021), this paper aims to create a more realistic cognitive radio (CR) scenario by introducing uncertainty into the reward assignment for frequency bands and simulating the process of detecting available frequency bands. To achieve this, a simplified version of energy detection was employed for identifying the availability of frequency bands.

2. Uncertainty in Global Mean Reward

In previous research, the Federated Multi-Armed Bandit (FMAB) algorithm was introduced as a solution to the challenges faced by traditional Multi-Armed Bandit (MAB) algorithms in cognitive radio (CR) systems. However, the original implementation of FMAB relied on a hardcoded global mean reward for each frequency band, which did not accurately reflect the realistic uncertainties present in CR scenarios. In actual situations, the availability of frequency bands as idle or occupied is often unknown, making it challenging to determine the rewards associated with each band.

To address this limitation and achieve a more realistic representation of CR scenarios, we aim to improve the FMAB algorithm. In our enhanced approach, we introduce a random generation of mean rewards for each frequency band, reflecting the uncertainty of their availability. This modification allows FMAB to adapt to the dynamic nature of spectrum conditions and effectively select the most rewarding frequency band based on limited information.

By incorporating this enhancement, we evaluate the effectiveness of FMAB in real-world CR scenarios, where the availability of frequency bands is uncertain. Through collaborative learning and information exchange among distributed cognitive radios, the FMAB algorithm enables efficient decision-making and enhances spectrum access and utilization, even in the face of varying and uncertain channel conditions. By evaluating the performance of FMAB in this more realistic context, we can assess its effectiveness in practical CR deployments and its potential to improve overall network performance.

Algorithm	Regret	Communication cost	delta
$Fed1UCB_{\text{old}}$	13381.613	992.1	0.02
Fed1 UCB	9327.629	1016.15	0.02
$Fed2UCB_{\text{old}}$	14876.184	1540.58	0.02
Fed2 UCB	1993.906	187.12	0.057

Table 1: Results of implementing each algorithm based on spectrum sensing and a more realistic randomized global mean reward. A comparison between the old hardcoded approach and the improved method incorporating labeling. The improved method demonstrates lower regret values, indicating more efficient attainment of optimal rewards. Communication cost is reduced for Fed2 UCB, whereas Fed1 UCB's cost is increased. Lastly, the observed delta values for the improved approach are equal to or larger, indicating increased differentiation between reward intervals compared to the previous hardcoded implementation.

2.1. *Randomized Mean Reward*

The probability density function of a normal distribution is given by:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{1}$$

where x is the random variable, μ is the mean, and σ is the standard deviation.

3. Spectrum Sensing

In this study, we employed energy detection as a representative method for spectrum sensing, which is widely used in practice. Energy detection traditionally determines the presence of primary users by comparing the received energy with a predetermined threshold. However, within the framework of multi-armed bandit (MAB) algorithms, where the objective is to maximize the rewards of different arms, we adapted the energy detection mechanism to align with the optimization goal of the MAB framework. Specifically, we assumed a positive correlation between higher rewards and the availability of idle frequencies. As a result, we implemented spectrum sensing by selectively collecting data from frequency indices associated with idle bands. This approach effectively reduced the exploration cost, leading to decreased communication overhead and expedited convergence with diminished regret, thus facilitating the attainment of the optimal solution.

3.1. *Code Implementation*

Considering both the randomized generation of global mean rewards and spectrum sensing, we made the assumption that only frequency bands with a reward value of 0.5 indicate idle status. We captured and stored the indices associated with these specific frequency bands. By focusing our exploration on arms with rewards surpassing a certain threshold, 0.5 in this study, we expect to effectively reduce exploration costs and achieve faster convergence with lower communication expenses. This approach allowed us to identify and exploit optimal frequencies more efficiently.

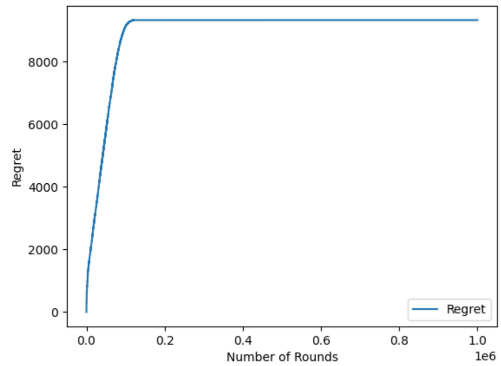

Figure 1: Fed1 UCB

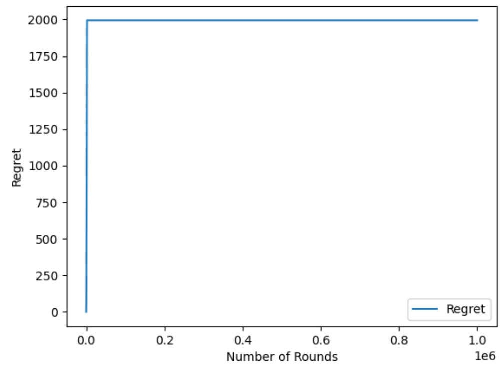

Figure 2: Fed2 UCB

4. Discussion

This study presents a comparative analysis between two approaches: Fed1 UCB and Fed2 UCB. By examining the findings presented in Table 1 and Figures 1 and 2, we can gain valuable insights into the results.

The evaluation of these approaches reveals that Fed1 UCB, which allows independent client behavior, achieved reduced regret by employing efficient data collection through spectrum sensing even in realistic cognitive radio scenarios. However, it is noteworthy that this approach led to an increase in communication costs. While it cannot be assumed that communication costs always increase due to the random generation of global mean rewards, it can be interpreted that the independent operation of each client in Fed1 UCB leads to inefficiencies in communication. Since each client relies solely on its own information and communicates it to the global server, we can conclude that there is no improvement in communication costs.

In contrast, Fed2 UCB, which emphasizes client cooperation, exhibits notable performance enhancement. By leveraging selectively identified idle frequencies, it demonstrates significantly lower regret and communication costs. The collaborative nature of Fed2 UCB enables faster convergence towards optimal rewards, outperforming the Fed1 UCB and even original Fed2 UCB, which also exhibits rapid convergence.

5. Summary and conclusions

This study compares two approaches, namely Fed1 UCB and Fed2 UCB, in the context of cognitive radio systems. The Fed1 UCB approach allows independent client behavior, while Fed2 UCB emphasizes client cooperation. By integrating federated learning principles into the Multi-Armed Bandit (MAB) framework, the Federated Multi-Armed Bandit (FMAB) algorithm is introduced to enable collaborative decision-making among distributed cognitive radios.**?**

The FMAB algorithm leverages the power of federated learning to enhance decision-making in dynamic and heterogeneous spectrum conditions. Unlike traditional MAB algorithms that rely solely on individual radio observations, FMAB harnesses the collective intelligence of the cognitive radio network. It facilitates information exchange and shared knowledge among radios, leading to improved spectrum access and utilization.

To address the limitations of the original FMAB implementation, the study introduces uncertainty in the reward of frequency bands and incorporates a randomized mean reward. By introducing realistic uncertainties and spectrum sensing using energy detection, the FMAB algorithm adapts to dynamic spectrum conditions and selects rewarding frequency bands based on limited information.

The results show that Fed1 UCB achieves efficient data collection through spectrum sensing in realistic scenarios but increases communication costs. On the other hand, Fed2 UCB, with client cooperation and selective identification of idle frequencies, demonstrates significantly lower regret and communication costs, outperforming both Fed1 UCB and the original Fed2 UCB.

In conclusion, the study highlights the importance of incorporating federated learning principles into the Multi-Armed Bandit framework for cognitive radio systems. The FMAB algorithm, with its collaborative decision-making and utilization of shared knowledge, proves effective in improving spectrum access and utilization. By introducing uncertainty in rewards and implementing spectrum sensing, the algorithm adapts to dynamic spectrum conditions and selects optimal frequency bands.

The comparison between Fed1 UCB and Fed2 UCB reveals the advantages of client cooperation in Fed2 UCB, leading to faster convergence and lower regret. While Fed1 UCB achieves efficient data collection, it incurs increased communication costs due to independent client behavior. These findings provide valuable insights into the performance of different approaches in realistic cognitive radio scenarios

References

Chengshuai Shi, C.S., 2021. Federated multi-armed bandits. CoRR abs/2101.12204.

Liang, Y.C., Chen, K.C., Li, G.Y., Mahonen, P., 2011. Cognitive radio networking and communications: An overview. IEEE transactions on vehicular technology 60, 3386–3407.

Oksanen, J., Koivunen, V., 2015. An order optimal policy for exploiting idle spectrum in cognitive radio networks. IEEE Transactions on Signal Processing 63, 1214–1227.

Wang, B., Liu, K.R., 2010. Advances in cognitive radio networks: A survey. IEEE Journal of selected topics in signal processing 5, 5–23.

Xia, W., Quek, T.Q., Guo, K., Wen, W., Yang, H.H., Zhu, H., 2020. Multi-armed bandit-based client scheduling for federated learning. IEEE Transactions on Wireless Communications 19, 7108–7123.