Spectrum Sharing of Mobile Network Operators based on Deep Reinforcement Learning in 5G and beyond

Yi-Huai Hsu, Chen-Fan Chang, and Chao-Hung Lee
Department of Computer Science and Engineering, Yuan Ze University, Taoyuan, Taiwan
Email: yhhsu@saturn.yzu.edu.tw, {s1083303, s1106030}@mail.yzu.edu.tw

GLOBECOM 2022 - 2022 IEEE Global Communications Conference | 978-1-6654-3540-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/GLOBECOM48099.2022.10000876

***Abstract*—Spectrum sharing has been proposed to effectively utilize the limited licensed spectrum resource and provide extra capacity for mobile network operator (MNO) who requires more spectrum in 5G network. In spectrum sharing, the initialized licensed spectrum of each MNO is divided into a reserved and shared spectrum. The reserved spectrum is for the private use of an MNO, while the shared spectrum of all MNOs forms a spectrum pool which will be dynamically used by MNO who requires an additional spectrum. However, the spectrum pool management problem significantly affects the efficiency of spectrum sharing among multi-MNO. In this paper, we formulate this problem into a non-linear programming problem that aims to maximize the average data rate of UEs, which subscribe to an MNO sending the spectrum request event, of all events from all MNOs. We propose an event-driven Deep Reinforcement Learning based spectrum sharing mechanism (EDRL-SSM) that utilizes a spectrum pool manager to effectively manage the spectrum pool to achieve long-term optimization of data rate performance of UEs. The spectrum pool manager intelligently allocates spectrum pool resource for each incoming spectrum request event. The proposed EDRL-SSM fully utilizes a DRL technique, Deep Deterministic Policy Gradient (DDPG), to deal with stochastic spectrum request event arrivals in the spectrum pool manager. The simulation results show that the proposed EDRL-SSM can significantly improve the data rate performance of UEs as compared to greedy spectrum pool allocation and without spectrum sharing under both identical and different initialized allocation of licensed spectrum scenarios.**

***Index Terms*—5G, spectrum sharing, mobile network operator, deep reinforcement learning.**

I. Introduction

With the rapid growth of the number of 5G mobile users expected to grow from 8.8 billion in 2018 to 13.1 billion by 2023 [1] and the emergent development of bandwidth-hungry applications such as 4K video and virtual reality (VR), the demand for additional licensed spectrum to support tremendous mobile services of a mobile network operator (MNO) has significantly increased. However, the available licensed spectrum for an MNO has not been increased accordingly because of the high cost of deployment of additional licensed spectrum. Moreover, the traditional static licensed spectrum allocation, in which the initialized licensed spectrum allocated to each MNO can not be changed and MNOs can not share the allocated licensed spectrum with each other, may cause inefficient utilization of the scarce licensed spectrum for an MNO. Thus, the spectrum sharing [2] [3] [4] [5], in which

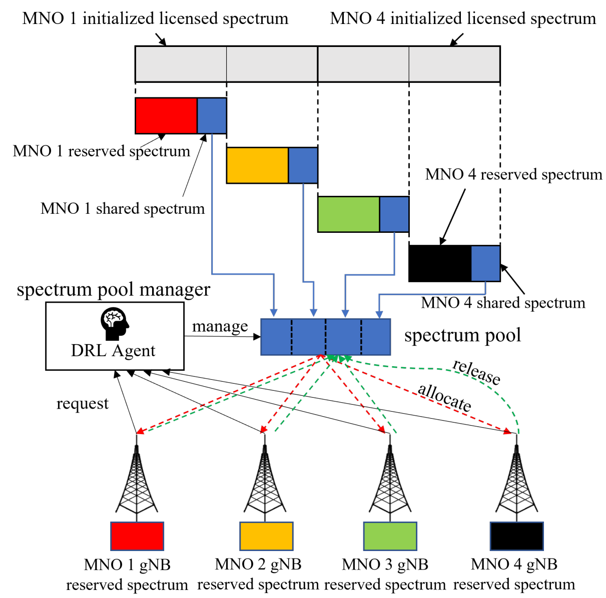

Fig. 1. Spectrum sharing among multi-MNO scenario with identical initialized allocation of licensed spectrum.

the initialized licensed spectrum is partitioned into reserved spectrum for the private use of an MNO and shared spectrum to form the spectrum sharing pool for sharing use of all MNOs, has been proposed to increase the spectrum utilization as each MNO can request for additional spectrum from the spectrum sharing pool and release the spectrum to the spectrum sharing pool according to their loading.

Saha [6] proposed Flexible License Spectrum Allocation (FLSA) that periodically allocates the initialized licensed spectrum to each MNO according to the ratio of the number of user equipments (UEs) subscribed to each MNO over the total number of UEs subscribed to all MNOs. Saha [7] [8] proposed Licensed Countrywide Full-Spectrum Allocation (LCFSA) to periodically allocate the initialized licensed spectrum to each MNO in a specific area when UEs arrive in this area according to the ratio of the number of UEs subscribed to each MNO over the total number of UEs subscribed to all MNOs. The above works [6] [7] [8] showed that their proposed spectrum allocation techniques can outperform the traditional static licensed spectrum allocation. However, these works only consider the periodic allocation of initialized licensed spectrum

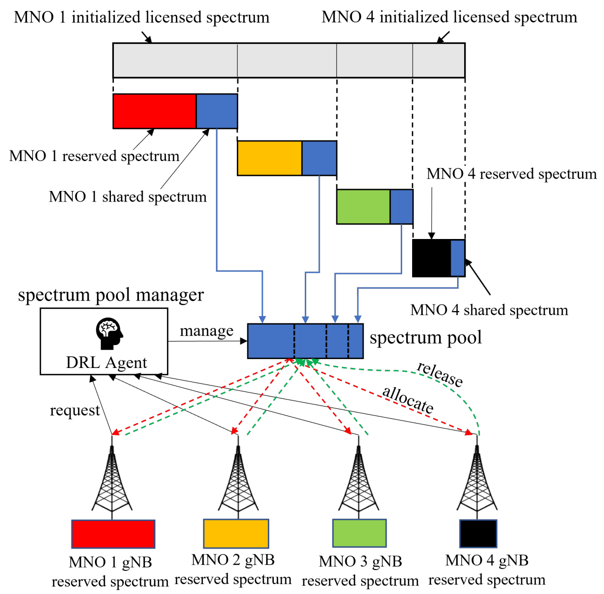

Fig. 2. Spectrum sharing among multi-MNO scenario with different initialized allocation of licensed spectrum.

without spectrum sharing. The frequent change of the number of UEs of each MNO within a specific area in a short time can not be reflected because the change period of the initialized licensed spectrum for all MNOs may take a couple of months with high cost.

Tehrani et al. [4] investigated the concept of spectrum sharing and introduced different spectrum sharing scenarios with their corresponding challenges and benefits. They mentioned that the shared spectrum needs to be regulated by specific rules so as to ensure rational access to the spectrum. Saha [5] proposed FLSA-SS which is a FLSA [6] based spectrum sharing mechanism. In FLSA-SS, each MNO uses the whole spectrum of the sharing pool so that the data rate performance of UEs in each MNO can be significantly improved. However, FLSA-SS is impractical because part of the sharing pool may be occupied by some MNOs when an MNO requires an additional spectrum from the sharing pool. Thus, none of the MNO can use the entire spectrum of the sharing pool. Saha [9] compared the performance of FLSA-SS [5] and LCFSA [7] [8]. They showed that FLSA-SS [5], which utilizes spectrum sharing, outperforms LCFSA [7] [8]. Saha [10] used enhanced inter-cell interference coordination (eICIC) to enable the indoor small cell base stations can share the spectrum of all MNOs' macro cell base stations. Hu et al. [11] introduced the potential, advantages, and challenges of applying blockchain and artificial intelligence (AI) for dynamic resource sharing such as spectrum sharing and computing resource sharing in 6G and beyond. Rony et al. [12] proposed a framework to pool the available spectrum into a spectrum pool of an MNO, use AI to predict the number of UEs who may arrive in each base station of the MNO, and allocate the spectrum resource in the spectrum pool to each base station of the MNO according to the predicted result. Liu and Tian [13] adopted a cooperative game theorem, Bankruptcy game, to allocate the shared spectrum to all MNOs.

The above works [4] [5] [9] [10] [11] [12] adopted the spectrum sharing technique to improve spectrum utilization. However, most of them adopted spectrum sharing, but none of them proposed an efficient spectrum pool management mechanism to effectively manage the spectrum pool. More specifically, when the spectrum pool is not effectively managed, some of the MNOs may occupy most of the spectrum pool so that other MNOs have no chance to use the spectrum pool. Only [13] proposed a shared spectrum management mechanism that tries to optimize network performance in a short time. They assumed that the request event arrivals from all MNOs are synchronized within a short time. However, this assumption is not practical since the request event arrivals from all MNOs should be interleaved. In this paper, we propose an event-driven Deep Reinforcement Learning based spectrum sharing mechanism (EDRL-SSM) that utilizes a spectrum pool manager to effectively manage the spectrum pool so as to achieve long-term optimization of the data rate performance of UEs. Event-driven means that the spectrum pool allocation is immediately triggered by the incoming spectrum request event at any time. To the best of our knowledge, our work is the first to adopt DRL for the spectrum pool management problem of spectrum sharing among multi-MNO in 5G network under event-driven based environment.

In this paper, we investigate the spectrum pool management problem of spectrum sharing among multi-MNO in 5G network. We formulate this problem into a non-linear programming problem that aims to maximize the average data rate of UEs, which subscribe to an MNO sending the spectrum request event, of all events from all MNOs. For this purpose, we propose an event-driven Deep Reinforcement Learning based spectrum sharing mechanism (EDRL-SSM) that utilizes a spectrum pool manager to effectively manage the spectrum pool so as to achieve long-term optimization of the data rate performance of UEs. We consider two spectrum sharing among multi-MNO scenarios in 5G network which are identical and different initialized allocation of licensed spectrum as shown in Fig. 1 and Fig. 2, respectively. For simplicity, in this paper, we assume that each MNO has only one gNB and all gNBs of all MNOs are located in the same area. In both scenarios, the initialized allocation of licensed spectrum will be partitioned into reserved spectrum and shared spectrum according to the sharing rate. The reserved spectrum is for the private use of an MNO, and the shared spectrum forms a spectrum pool which will be dynamically used by MNO who requires an additional spectrum. If an MNO requires an additional spectrum due to the increase of UEs, it will send a request event to the spectrum pool manager. The spectrum pool manager then intelligently allocates the spectrum pool resource for the incoming spectrum request event. Finally, the MNO releases the spectrum resource to the spectrum pool after the number of UEs is decreased. The proposed EDRL-SSM makes full use of a DRL technique, Deep Deterministic Policy Gradient (DDPG), to deal with stochastic spectrum request event arrivals in the spectrum pool manager. The simulation results show that the proposed EDRL-SSM can significantly improve the data rate performance of UEs as compared to greedy spectrum pool allocation and without spectrum sharing under both identical and different initialized allocation of licensed spectrum scenarios.

The rest of the paper is organized as follows. In Sec. II, the system model is described. In Sec. III, the spectrum pool management problem of spectrum sharing among multi-MNO in 5G network is formulated. In Sec. IV, the design details of the proposed EDRL-SSM are described. In Sec. V, the simulation results are shown. Finally, the paper is concluded in Sec. VI.

II. System Model

We consider two spectrum sharing among multi-MNO scenarios with a set $\mathcal{N}$ of N MNO in 5G network which are identical and different initialized allocation of licensed spectrum as shown in Fig. 1 and Fig. 2, respectively. For simplicity, we assume that each MNO has only one gNB and all gNBs of all MNOs are located in the same area. In both scenarios, $W_n^{m,ini}$ is the bandwidth of initialized licensed spectrum allocated to MNO n. The initialized licensed spectrum of MNO n will be partitioned into reserved spectrum and shared spectrum according to the sharing rate r. Thus, the bandwidth of reserved spectrum and shared spectrum are calculated by (1) and (2), respectively.

$$W_n^{m,res} = W_n^{m,ini} \times (1-r), \forall n \in \mathcal{N} \tag{1}$$

$$W_n^{m,sha} = W_n^{m,ini} \times r, \forall n \in \mathcal{N} \tag{2}$$

The shared spectrum of each MNO forms a spectrum pool which will be dynamically used by MNO who requires an additional spectrum, so the bandwidth of spectrum pool is calculated by (3).

$$W^{pool} = \sum_{n=1}^{N} W_n^{m,sha} \tag{3}$$

If the number of UE subscribes to an MNO is greater than a certain threshold, the MNO sends a spectrum request event to the spectrum pool manager. We assume the bandwidth of the reserved spectrum of MNO n can serve C_n^{m,res_max} UEs at most as given by the left hand side of (4), so the number of UE subscribes to MNO n when MNO n sends the spectrum request event i can be served by the bandwidth of reserved spectrum and additional request spectrum from event i of MNO n as given by the right hand side of (4).

$$\frac{C_n^{m,res\_max}}{W_n^{m,res}} = \frac{C_n^m(i)}{W_n^{m,res} + W_n^{m,req}(i)} \tag{4}$$

where $C_n^m(i)$ is the number of UE subscribes to MNO n when MNO n sends the spectrum request event i and $W_n^{m,req}(i)$ is the additional required bandwidth of event i of MNO n. We can calculate $W_n^{m,req}(i)$ as given by (5) according to (4).

$$W_n^{m,req}(i) = \frac{W_n^{m,res}(C_n^m(i) - C_n^{m,res\_max})}{C_n^{m,res\_max}} \tag{5}$$

We assume that the MNO n has a set $\mathcal{I}_n$ of I_n events, where each event means one spectrum request from an MNO, and I_n is a random variable. We also assume that each UE which subscribes to an MNO shares the same amount of spectrum. Thus, the data rate of UE e which subscribes to MNO n in event i is calculated based on the Shannon capacity model as given by (6).

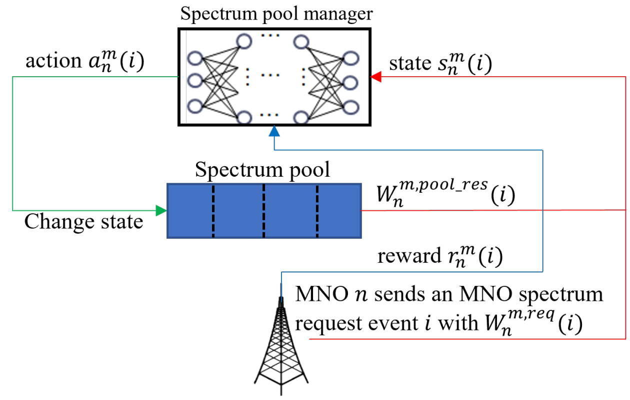

Fig. 3. Workflow of the proposed EDRL-SSM.

$$r_{n,e}^{m,u}(i) = \frac{(W_n^{m,res} + W_n^m(i) W^{pool} W_n^{m,pool\_res}(i))}{C_n^m(i)} \log_2\left[1 + \frac{ph_{n,e}^{m,u}(i)}{\frac{(W_n^{m,res} + W_n^m(i) W^{pool} W_n^{m,pool\_res}(i))}{C_n^m(i)} N_0}\right] \tag{6}$$

where $W_n^m(i)$ is the spectrum allocation ratio of spectrum pool to event i of MNO n, $W_n^{m,pool\_res}(i)$ is the residual spectrum ratio of spectrum pool when spectrum request event i of MNO n arrives in spectrum pool manager, p is the transmission power of 5G gNB, $h_{n,e}^{m,u}(i)$ is the channel condition of UE e of MNO n in event i, and N_0 is the power spectral density (PSD) of the complex white Gaussian channel noise. The average data rate of UEs which subscribe to MNO n in event i as given by (7).

$$R_n^{m,u}(i) = \frac{\sum_{e=1}^{C_n^m(i)} r_{n,e}^{m,u}(i)}{C_n^m(i)} \tag{7}$$

III. Problem Formulation

We formulate the spectrum pool management problem of spectrum sharing among multi-MNO in 5G network into a non-linear programming problem. The objective of this problem is to maximize the average data rate of UEs, which subscribe to an MNO sending the spectrum request event, of all events from all MNOs. This problem is formulated as follows:

$$\mathbf{P}: \max_{\mathbf{W}} \sum_{n=1}^{N} \sum_{i=1}^{I_n} R_n^{m,u}(i)$$

$$s.t. \quad 0 < W_n^m(i) \leq 1, \forall n \in \mathcal{N}, \forall i \in \mathcal{I}_n. \tag{8}$$

(8) shows the range of the spectrum allocation ratio of spectrum pool to event i of MNO n allocated by the spectrum pool manager.

In problem $\mathbf{P}$, $\mathbf{W}$ denotes the set of spectrum allocation ratios of spectrum pool to MNOs. $\mathbf{P}$ can be solved by finding the optimal values of $\mathbf{W}$. Nevertheless, $\mathbf{P}$ is a non-linear programming problem and can be proved as NP-hard problem by the reduction from the knapsack problem. Moreover, it is also difficult to know the spectrum request event arrival distribution and the UE arrival distribution in advance. Therefore, in our proposed EDRL-SSM, we utilize the DRL technique in spectrum pool manager to efficiently solve problem $\mathbf{P}$.

IV. Event-Driven DRL Based Spectrum Sharing Mechanism (EDRL-SSM)

The proposed EDRL-SSM utilizes a spectrum pool manager to effectively manage the spectrum pool so as to achieve long-term optimization of the data rate performance of UEs. The state, action, and reward of the spectrum pool manager when the spectrum request event i of MNO n arrives are defined as follows.

1) *state*. The wireless network state is $s_n^m(i) = (W_n^{m,pool\_res}(i), W_n^{m,req}(i))$.
2) *action*. The action to take for state $s_n^m(i)$ is $a_n^m(i) = W_n^m(i)$.
3) *reward*. Since the goal of the spectrum pool manager is to maximize the average data rate of UEs which subscribe to an MNO, the reward is $r_n^m(i) = R_n^{m,u}(i)$.

The workflow of the proposed EDRL-SSM is shown in Fig. 3. If the number of UEs subscribe to MNO n is greater than a certain threshold, the MNO n sends a spectrum request event i to the spectrum pool manager. When the event i of MNO n arrives, the spectrum pool manager will take action $a_n^m(i)$, which is the spectrum allocation ratio of spectrum pool to event i of MNO n, according to the current state $s_n^m(i)$, which consists of residual spectrum ratio of spectrum pool when event i of MNO n arrives and the additional required bandwidth of event i of MNO n. After the spectrum pool manager takes this action, the spectrum pool manager receives the reward $r_n^m(i)$, which is the average data rate of UEs which subscribe to MNO n in event i, to update its neural network's gradients and the state of spectrum pool will be changed. Note that the state of spectrum pool will also be changed when the MNO releases the spectrum to the spectrum sharing pool.

We implement a DRL technique, DDPG, in spectrum pool manager. For simplicity, we use s_t, a_t, and r_t to denote the state, action, and reward at event t, and s_{t+1}, a_{t+1}, and r_{t+1} to represent the state, action, and reward at next event $t+1$. The concept and training of DDPG are explained as follows. DDPG [14], an actor-critic based DRL, can operate over continuous action spaces. DDPG comprises an evaluating network and a target network. Both networks have an actor and a critic. The evaluating network actor outputs $a_t = \pi(s_t; \theta_a)$ the best action a_t from all possible actions based on current state s_t that can maximize the output of evaluating network critic $Q(s_t, a_t; \theta_c)$. The meaning of $Q(s_t, a_t; \theta_c)$ is the expected accumulated reward from state s_t by taking action a_t until the end of this episode. The current state s_t, current action a_t, current reward r_t, and next state s_{t+1} will be stored in the replay buffer. The gradients of evaluating network actor and critic are updated by randomly selecting K samples from the replay buffer. We assume that one sample selected from the replay buffer is (s_j, a_j, r_j, s_{j+1}). The gradient of evaluating network actor is updated by the average of the results of calculating gradient descent of the output of evaluating network critic $Q(s_j, a_j; \theta_c)$ and then calculating gradient descent of the output of evaluating network actor $\pi(s_j; \theta_a)$ from K samples as given by (9), where η_a is the learning rate of evaluating network actor and ∇ denotes the gradient descent which takes

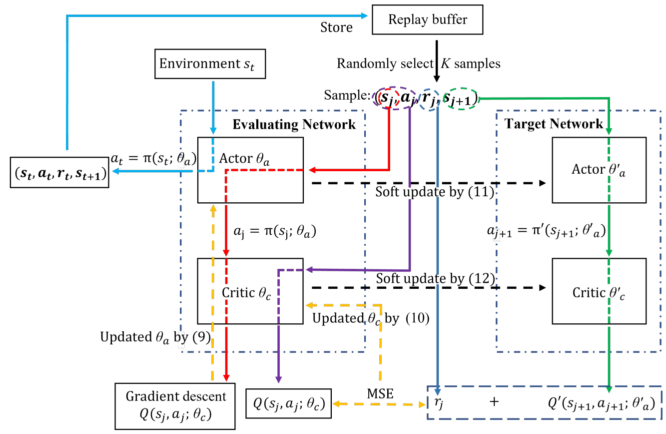

Fig. 4. The structure of DDPG.

partial difference with respect to weight and bias (neural network's gradient).

$$\theta_a \leftarrow \theta_a + \frac{1}{K}\sum_{k=1}^{K} \eta_a \nabla Q(s_j, a_j; \theta_c) | \nabla \pi(s_j; \theta_a) \tag{9}$$

The gradient of evaluating network critic is updated by the average of the results of calculating gradient descent of the mean square error (MSE) of $Q(s_j, a_j; \theta_c)$ and $r_j + Q^{'}(s_{j+1}, a_{j+1}; \theta_c^{'})$ from K samples as given by (10), where $Q^{'}$ is the output of the target network critic and η_c is the learning rate of evaluating network critic.

$$\theta_c \leftarrow \theta_c + \frac{1}{K}\sum_{k=1}^{K} \eta_c \nabla [(r_j + Q^{'}(s_{j+1}, a_{j+1}; \theta_c^{'})) - Q(s_j, a_j; \theta_c)]^2 \tag{10}$$

The gradients of both evaluating network actor and critic are updated after the DDPG agent takes action in each event. After updating the gradients of both evaluating network actor and critic by K samples, the gradients of both target network actor and critic are updated by the "soft update" method as given by (11) and (12), respectively, where τ is the soft update factor. The structure of DDPG is shown in Fig. 4.

$$\theta_a^{'} \leftarrow \tau\theta_a + (1-\tau)\theta_a^{'}. \tag{11}$$

$$\theta_c^{'} \leftarrow \tau\theta_c + (1-\tau)\theta_c^{'}. \tag{12}$$

V. Performance Evaluation

We conducted a system level simulation using Python [15] and TensorFlow [16]. We consider two spectrum sharing among multi-MNO scenarios in 5G network which are identical and different initialized allocation of licensed spectrum as shown in Fig. 1 and Fig. 2, respectively. In our simulation setup, the number of MNOs [18], the range of 5G spectrum [18], and the bandwidth of initialized licensed spectrum of each MNO in different initialized spectrum scenario which is calculated by the population of market analysis of mobile communication [19] are based on the report of National Communications Commission (NCC) of Taiwan. For simplicity, we assume that each MNO has only one gNB and all gNBs of all MNOs are located in the same area. The channel gain of UE e which subscribes to MNO n in event i, $h_{n,e}^{m,u}(i)$, follows the

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Number of MNO (N) [18]	4
Range of 5G spectrum [18]	3300 MHz $-$ 3570 MHz
Total number of 5G spectrum	270MHz
The bandwidth of initialized licensed spectrum of each MNO ($W_n^{m,ini}$) in identical initialized spectrum scenario	67.5MHz
The maximum number of UE can be served in initialized licensed spectrum of each MNO ($C_n^{m,ini}$) in identical initialized spectrum scenario	200
The bandwidth of initialized licensed spectrum of each MNO ($W_n^{m,ini}$) in different initialized spectrum scenario [19]	97.47, 65.88, 65.61, 41.04 (MHz)
The maximum number of UE can be served in initialized licensed spectrum of each MNO ($C_n^{m,ini}$) in different initialized spectrum scenario	288, 195, 194, 121
The maximum number of UE can be served in reserved spectrum of each MNO (C_n^{m,res_max})	$C_n^{m,ini} \times (1-r)$
Threshold of sending spectrum request event	$C_n^{m,res\_max} \times 0.8$
Power of gNB (p)	46 dBm
Spectrum request event inter arrival time (Poisson distribution)	3.9 s $-$ 4.1 s
Antenna channel gain (A_d)	4.11
Carrier frequency (f_c)	3.5 GHz
Power spectral density (N_0)	2×10^{-13}
Number of events in an episode	500
Number of episode	20000
Learning rate of DDPG actor (η_a)	5×10^{-8}
Learning rate of DDPG critic (η_c)	10^{-7}
Number of hidden layers of DDPG actor and DDPG critic	5
Number of neurons in each layer of DDPG actor and DDPG critic	600, 300, 200, 100, 50
DDPG activation function of neuron	Rectified Linear Unit (ReLU)
Number of samples from replay buffer (K)	32

free-space path loss (FSPL) model [17] and is formulated as $A_d(\frac{3\times10^8}{4\pi f_c d_{n,c}^{m,u}(i)})^2$, where A_d is the antenna channel gain, f_c is the carrier frequency, and $d_{n,c}^{m,u}(i)$ is the distance between UE e and MNO n's gNB in event i. The complete set of simulation setting is shown in Table I.

We compare the proposed EDRL-SSM with two other schemes such as greedy and without sharing. The concepts of these schemes are explained as follows.

1) *Greedy*. The spectrum pool manager allocates all the remaining spectrum in the spectrum pool to an MNO if the number of spectrum the MNO request is larger than the remaining spectrum in the spectrum pool. The spectrum pool manager allocates the number of spectrum the MNO request if the number of spectrum the MNO request is less than the remaining spectrum in the spectrum pool.
2) *Without Sharing*. It only considers the initialized allocation of licensed spectrum for the private use of each MNO.

We denote R as the average data rate of UEs, which subscribe to an MNO, averaged over all events from all MNOs

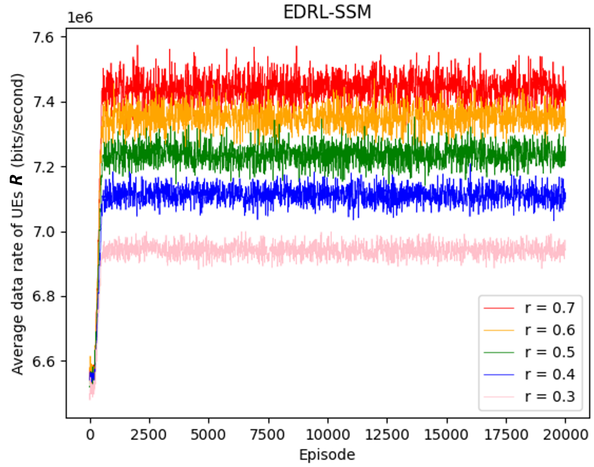

Fig. 5. Learning curve of the average data rate of UEs, which subscribe to an MNO, averaged over all events from all MNOs in the proposed EDRL-SSM with different sharing rate r in identical initialized spectrum scenario.

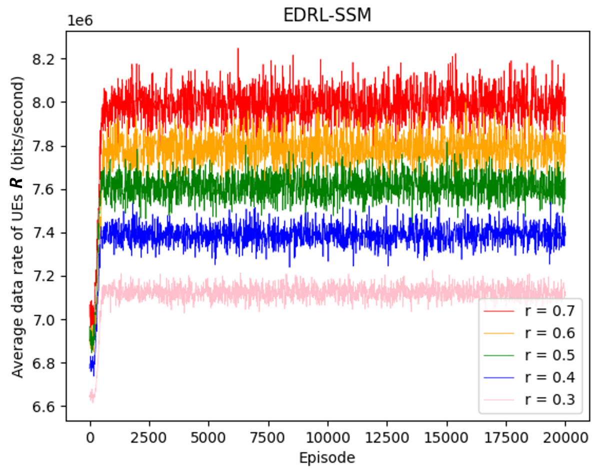

Fig. 6. Learning curve of the average data rate of UEs, which subscribe to an MNO, averaged over all events from all MNOs in the proposed EDRL-SSM with different sharing rate r in different initialized spectrum scenario.

calculated by (13). Fig. 5 and Fig. 6 shows the learning curves of R in our proposed EDRL-SSM with different sharing rate r in identical and different initialized spectrum scenario, respectively. We can see that in both scenarios the data rate performance of UEs is increased as the value of sharing rate r increased. This is because the more the shared spectrum is the better the flexibility of the proposed EDRL-SSM can efficiently allocate the sharing spectrum to the MNOs that require an additional spectrum. We can also observe that the variance of Fig. 6 is bigger than that of Fig. 5 which reflects the initialized allocation of licensed spectrum.

$$R = \frac{\sum_{n=1}^{N}\sum_{i=1}^{I_n} R_n^{m,u}(i)}{\sum_{n=1}^{N} I_n} \tag{13}$$

Fig. 7 and Fig. 8 shows the average convergence of R in identical and different initialized spectrum scenarios, respectively. In both scenarios, the average convergence of R of the proposed EDRL-SSM is increased with the increase of sharing rate r, while that of greedy is decreased when the sharing rate is increased. This is because the more the shared

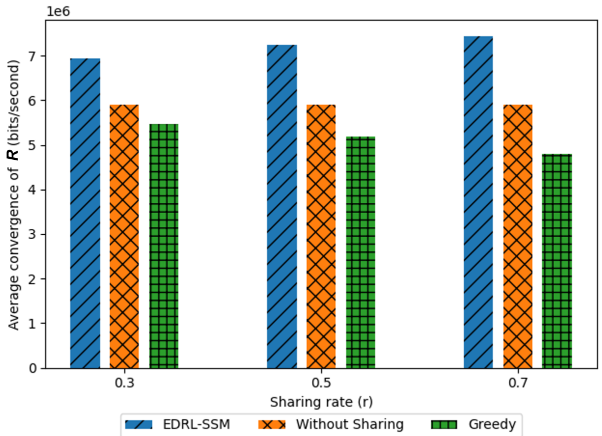

Fig. 7. Average convergence of R in identical initialized spectrum scenario.

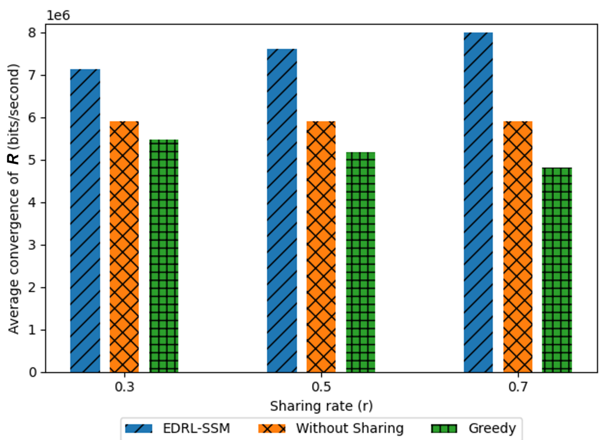

Fig. 8. Average convergence of R in different initialized spectrum scenario.

spectrum is, the less the reserved spectrum can be used by each MNO. Moreover, the shared spectrum can not be fully utilized without efficient management. That is, in greedy, most of the shared spectrum in spectrum pool may be allocated to certain MNOs in advance so that other MNOs may not have chance to use the spectrum pool when they send the request. We can also observe that the performance of without sharing is better than that of greedy. Thus, if we can not efficiently manage the spectrum pool, spectrum sharing may reduce the data rate performance. In both scenarios, the data rate performance of the proposed EDRL-SSM can outperforms that of greedy and without sharing.

VI. Conclusion

In this paper, we have formulated the spectrum pool management problem into a non-linear programming problem that aims to maximize the average data rate of UEs, which subscribe to an MNO sending the spectrum request event, of all events from all MNOs. For this purpose, we have proposed an event-driven DRL-based spectrum sharing mechanism called EDRL-SSM that utilizes a spectrum pool manager to effectively manage the spectrum pool so as to achieve long-term optimization of the data rate performance of UEs. The spectrum pool manager intelligently allocates spectrum pool resource for each incoming spectrum request event. The proposed EDRL-SSM makes full use of a DRL technique, DDPG, to deal with stochastic spectrum request event arrivals in spectrum pool manager. The simulation results have shown that the proposed EDRL-SSM can significantly improve the data rate performance of UEs as compared to greedy spectrum pool allocation and without spectrum sharing under both identical and different initialized allocation of licensed spectrum scenarios.

Acknowledgment

This work was supported by the Ministry of Science and Technology of Taiwan under Grant Numbers MOST 110-2222-E-155-002-MY2.

References

[1] Cisco, “Cisco Annual Internet Report (2018–2023) White Paper,” Update: Mar. 2020.

[2] 3GPP, “Service Aspects and Requirements for Network Sharing,” 3GPP TR 22.951 V17.0.0, 3GPP, Sophia Antipolis, France, Apr. 2022.

[3] J. Opadere, Q. Liu, T. Han, and N. Ansari, “Energy-Efficient Virtual Radio Access Networks for Multi-Operators Cooperative Cellular Networks,” in *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 3, pp. 603-614, Sep. 2019.

[4] R. H. Tehrani, S. Vahid, D. Triantafyllopoulou, H. Lee, and K. Moessner, “Licensed Spectrum Sharing Schemes for Mobile Operators: A Survey and Outlook,” in *IEEE Communications Surveys and Tutorials*, vol. 18, no. 4, pp. 2591–2623, Jun. 2016.

[5] R. K. Saha, “Approaches to Improve Millimeter-Wave Spectrum Utilization Using Indoor Small Cells in Multi-Operator Environments Toward 6G,” in *IEEE Access*, vol. 8, pp. 207643-207658, Nov. 2020.

[6] R. K. Saha, “A Flexible Licensed Spectrum Allocation Technique for Millimeter Wave Mobile Systems Toward 6G,” in *IEEE International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Oct. 2020, pp. 1-6.

[7] R. K. Saha, “A Countrywide Licensed Full Spectrum Allocation Method for Millimeter-Wave Mobile Systems for 6G,” in *IEEE Vehicular Technology Conference (VTC2020-Fall)*, Nov. 2020, pp. 1-7.

[8] R. K. Saha, “Licensed Countrywide Full-Spectrum Allocation: A New Paradigm for Millimeter-wave Mobile Systems in 5G/6G Era,” in *IEEE Access*, vol. 8, pp. 166612–166629, Sep. 2020.

[9] R. K. Saha, “Millimeter-Wave Spectrum Utilization Improvement in Multi-Operator Networks: A Framework Using the Equal Likelihood Criterion,” in *IEEE Access*, vol. 9, pp. 72980–72999, May. 2021.

[10] R. K. Saha, “Nationwide Spectrum Sharing of Mobile Network Operators with Indoor Small Cells,” in *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, Nov. 2019, pp. 1-2.

[11] S. Hu, Y. Liang, Z. Xiong, and D. Niyato, “Blockchain and Artificial Intelligence for Dynamic Resource Sharing in 6G and Beyond,” in *IEEE Wireless Communications*, vol. 28, no. 4, pp. 145-151, Aug. 2021.

[12] R. I. Rony, E. Lopez-Aguilera, and E. Garcia-Villegas, “Dynamic Spectrum Allocation Following Machine Learning-Based Traffic Predictions in 5G,” in *IEEE Access*, vol. 9, pp. 143458-143472, Oct. 2021.

[13] B. Liu and H. Tian, “A Bankruptcy Game-Based Resource Allocation Approach among Virtual Mobile Operators,” in *IEEE Communications Letters*, vol. 17, no. 7, pp. 1420–1423, Jul. 2013.

[14] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, “Continuous Control with Deep Reinforcement Learning,” in *International Conference on Learning Representations (ICLR)*, May 2016.

[15] “Python,” [Online]. Available: https://www.python.org/

[16] “TensorFlow,” [Online]. Available: https://www.tensorflow.org/

[17] L. Huang, S. Bi, and Y. J. Zhang, “Deep Reinforcement Learning for Online Computation Offloading in Wireless Powered Mobile-Edge Computing Networks,” in *IEEE Transactions on Mobile Computing*, vol. 19, no. 11, pp. 2581-2593, Jul. 2020.

[18] “National Communications Commission of Taiwan: Application for 5G Mobile Broadband,” [Online]. Available: https://www.ncc.gov.tw/chinese/files/20022/8_42757_200221_1.pdf

[19] “National Communications Commission of Taiwan: Market analysis of Mobile Communication,” [Online]. Available: https://www.ncc.gov.tw/chinese/files/22032/3773_47334_220328_1.pdf