INL/CON-24-81158-Revision-0

Distributed and Secure Spectrum Sharing for 5G and 6G Networks

November 2024

Arupjyoti Bhuyan, Xiang Zhang, Mingyue Ji

Changing the World's Energy Future

Idaho National Laboratory

INL is a U.S. Department of Energy National Laboratory operated by Battelle Energy Alliance, LLC

INL/CON-24-81158-Revision-0

Distributed and Secure Spectrum Sharing for 5G and 6G Networks

Arupjyoti Bhuyan, Xiang Zhang, Mingyue Ji

November 2024

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

Distributed and Secure Spectrum Sharing for 5G and 6G Networks

Arupjyoti Bhuyan, Xiang Zhang, Mingyue Ji
Email:arupjyoti.bhuyan@inl.gov, xiang.zhang@tu-berlin.de, mingyueji@ufl.edu

***Abstract*—Secure spectrum sharing or spectrum co-existence of multiple 5G networks and future 6G networks is a powerful enabler technology. The National Spectrum Strategy (NSS) published by the White House in November, 2023, and the subsequent NSS implementation plan led by the National Telecommunication and Information Administration (NTIA) is the driver of a national effort to enable co-existence of government incumbents and commercial networks in selected spectrum bands. Cellular networks such as 5G & 6G and non-cellular Wi-Fi 6E & 7 are the prominent wireless technologies considered for co-existence with incumbent wireless links. Security of the spectrum sharing solutions is a must to make this transformation of spectrum use possible, specially for mission critical communications. However, current spectrum sharing solutions rely on centralized data bases with inherent vulnerabilities. This paper focuses on secure spectrum sharing among multiple 5G networks using unlicensed and shared frequency bands. It presents an innovative AI/ML based distributed spectrum sharing approach that can be autonomously used by multiple networks. Each sharing network uses its own observation of the Radio Frequency (RF) environment, which consists of RF measurements reported from the 5G User Equipment (UE), to adjust the transmission power levels for secure co-existence. Data is presented to illustrate the superior performance of this solution compared to other spectrum sharing solutions where each network can utilize usage data of the other networks. Finally it discusses how this efficient spectrum sharing solution can evolve in the future for the 6G networks.**

***Index Terms*—Spectrum Sharing, Spectrum Co-existence,5G, 6G, Wi-Fi 6E/7, Wireless Security**

I. Introduction

In 2020, the Federal Communications Commission (FCC) repurposed 1200 MHz of spectrum (5.925-7.125 GHz) in the 6 GHz band, previously assigned for exclusive federal use, for the commercial use of Wi-Fi 6E/7 and 5G-NR-U in addition to its existing user. In November 2023, President Biden issued the "National Spectrum Security Memorandum on Modernizing United States Spectrum Policy and Establishing a National Spectrum Strategy" [1]. It "...directs my Administration to build on prior innovation by promoting efficient and effective spectrum use by both agencies and non-Federal users ...goal is to accelerate United States leadership in wireless communications ...and to unlock innovations that benefit the American people, while ensuring necessary access to spectrum for agencies and private-sector users, such as for scientific, public safety, critical infrastructure, and national security uses, now and into the future." The National Spectrum Strategy (NSS) was subsequently published by the White House [2] that states "...identify spectrum bands for potential repurposing .. to meet these growing demands."

The National Telecommunications and Information Administration (NTIA) created the 2024 National Spectrum Strategy (NSS) implementation plan [3] to support the NSS. This implementation plan has identified additional bands, including mid bands such as 7/8 GHz band, to be studied for possible co-existence of commercial devices with federal incumbents. Spectrum co-existence requires advanced spectrum sharing technologies which not only function efficiently at large scale, but are also secure and resilient under adversarial operating conditions. Hence security and resiliency in 5G/6G/Wi-Fi spectrum sharing needs further research and development that is critical to the success of the intended spectrum co-existence. Current spectrum sharing solutions for 5G and Wi-Fi 6E/7 rely on centralized data bases that are vulnerable to adversarial attacks. New distributed spectrum sharing solutions are required to remove this primary attack surface and provide a path towards scaling at the national level. Use of unlicensed spectrum, made possible with the 5G New Radio-Unlicensed (NR-U), will increase worldwide to deploy private 5G networks. In the future, 6G will use this powerful capability for continued growth of private cellular networks, and increased focus on secure spectrum sharing research is needed for the future 6G networks as well as advanced Wi-Fi technologies. In U.S., future 6G spectrum allocations are also likely to be intertwined with the outcome of the NSS implementation plan.

A. Additional spectrum repurposing for co-existence

The 2024 NSS implementation plan [3] has identified a total of 2,786 MHz to be studied for co-existence of commercial devices with federal incumbents by repurposing spectrum use in these bands [4]. These bands are listed in Table I.

TABLE I
National Spectrum Strategy Targeted Bands (frequencies in MHz)

Start Freq	End Freq	Width	Center Freq
3100	3450	350	3275
5030	5091	61	5060.5
7125	8400	1275	7762.5
18100	18600	500	18350
37000	37600	600	37300

B. Spectrum Co-existence in CBRS

Citizens Broadband Radio Service (CBRS) is a 150 MHz wide band (3.55-3.7 GHz) allocated by FCC for shared use in the United States. Current spectrum sharing among multiple networks in the CBRS band depends on the enforcement of access rules and arbitration of spectrum access by the Spectrum Access Server (SAS) database [4]. The CBRS band is increasingly used for private LTE and 5G networks across various industries, such as manufacturing, education, and smart cities. By the end of 2023, spending on LTE and 5G NR-based CBRS network infrastructure reached approximately $900 million, with projections to grow at a Compound Annual Growth Rate (CAGR) of 20% through 2026 [5].

C. Spectrum Co-existence for Wi-Fi

The 6 GHz band made available for commercial co-existence by FCC encompasses four U-NII (Unlicensed National Information and Infrastructure) bands: U-NII-5 to UNII-8 [6]. The highest power commercial Wi-Fi operation outdoors, called Standard Power (SP), is limited to U-NII-5 (5.925- 6.425 GHz, 500 MHz wide) and U-NII-7 bands (6.525-6.875 GHz, 350 MHz wide) and requires Automatic Frequency Coordination (AFC) to avoid interference with incumbents. Similar to SAS, AFC is a centralized database with data on incumbent assignments in the shared 6 GHz band. As number of commercial Wi-Fi 6E devices continue to rise at a rapid pace, the actual use of the SP operation is only starting with the recent availability of commercial AFCs. Addressing the incumbents' concern of possible interference from Wi-Fi SP operation outdoors, Department of Energy (DOE) Office of the Cybersecurity, Energy Security, and Emergency Response (CESER) has engaged Idaho National Laboratory (INL) for a science based objective assessment of AFC effectiveness in real world Wi-Fi operation outdoors. This research, already underway, will include a security assessment of the AFC to ensure that the Wi-Fi use continues without any impact to the incumbents under potential threats to the AFC operation. In the worst case, attacks on the AFC can allow commercial Wi-Fi to use the same radio resources used by a nearby utility link and render it nonoperational.

D. Need for Distributed and Autonomous Spectrum Sharing

As discussed, the current spectrum sharing architecture for both CBRS and 6 GHz bands rely on centralized data bases. CBRS also requires additional sensors for the Environment Sensing Capability (ESC) to detect use by incumbents. In addition to these databases being performance bottlenecks, they can become focal points for security attacks. Falsification of reported spectrum sensing data [7] can create sub-optimal and unfair spectrum usage or in the worst case, render incumbent wireless links to be not usable by allowing illegitimate commercial Wi-Fi use.

In this paper, we summarize our ongoing research with an innovative AI/ML based distributed spectrum sharing approach that can be autonomously used by multiple 5G networks. Each sharing network uses its own observation of the Radio Frequency (RF) environment, which consists of RF measurements reported from the 5G User Equipment (UE), to adjust the UE power levels for secure co-existence. This solution removes the reliance of spectrum sharing decisions on vulnerable databases and the need for an additional sensor network resulting in a secure and robust solution for spectrum sharing and co-existence.

The rest of this paper is organized as follows. Section II provides background on Reinforcement Learning (RL). Section III describes an innovative and autonomous beam scheduler that can be independently utilized by the operators sharing the spectrum. Section IV presents new numerical evaluation of this solution to demonstrate its effectiveness. Finally, section V descrbes how the presented solution will remain important and relevant in 6G and possible areas of enhancements utilizing new capabilities to be introduced in 6G.

II. Background: Reinforcement Learning

Given a Markov Decision Process (MDP) $(\mathcal{S}, \mathcal{A}, R, T)$ where $\mathcal{S}, \mathcal{A}, R$ and T respectively denote the state space, action space, reward function and the environment transition function, Reinforcement Learning (RL) seeks to navigate through the MDP by observing the environment state and choosing proper actions so that the expected reward in the long run is maximal. In particular, given some initial state $s_t \in \mathcal{S}$, the agent takes an action $a_t \in \mathcal{A}$ with probability $\mu(a_t|s_t)$ according to some policy μ which satisfies $\int_a \mu(a|s_t)da = 1$ (summation if the action space is discrete). Impacted by a_t, the environment transitions (governed by T) to a new state s_{t+1} and the agent receives a reward $r_t = R(s_t, a_t, s_{t+1})$ as an indication of how good a_t is. The collection of transition quadruples $\{(s_t, a_t, r_t, s_{t+1}), \forall t\}$ is called experiences of the agent. The return G_t is defined as the cumulative future rewards $G_t \triangleq \sum_{\tau=0}^{\infty} \gamma^\tau r_{t+\tau}$ with $\gamma \in (0, 1]$ being the discount factor. The Q-function Q^μ of the policy μ is defined as the expected return starting from any state-action pair (s, a), i.e., $Q^\mu(s, a) \triangleq \mathbb{E}\left[G_t | s_t = s, a_t = a\right]$.

Due to the recent advancement of the Deep Neural Networks (DNNs), Deep Reinforcement Learning (DRL) uses DNNs to represent the policy (the actor network) and Q-function (the critic network) in order to take advantage of DNN's representation capability. Deep Deterministic Policy Gradient (DDPG) [8] is a DRL algorithm which focuses on deterministic policies that map each state s to a specific action $a = \mu(s)$. It uses an actor-critic architecture in which two separate DNNs with parameters θ^μ and θ^Q are used to represent the policy $\mu(s|\theta^\mu)$ and the Q-function $Q(s, a|\theta^Q)$. Multi-agent DDPG (MA-DDPG) [9] is an adaptation of DDPG to the multiple-agent domain to combat the nonstationarity issue commonly seen in multi-agent RL systems. MA-DDPG uses a centralized-training-distributed-execution framework in which the actors and critics are trained periodically with network-level experiences while the action of each agent is determined based solely on that agent's local observation. In

particular, let $\boldsymbol{a} \triangleq (a_i)_{i=1}^K$ and $\boldsymbol{s} \triangleq (o_i)_{i=1}^K$ denote the joint actions and observations of all agents where o_i denotes the local observation of agent i. The learning process is described as follows. The critic and actor of agent i is represented by two DNNs $Q_i(\boldsymbol{s}, \boldsymbol{a}|\theta_i^Q)$ and $\mu_i(o_i|\theta_i^\mu)$. In order to make the learning more data-efficient, an experience replay buffer $\mathcal{D}$ is used to store the past experiences of all agents in a sliding-window manner. Mini-batches of experiences are then sampled repeatedly from $\mathcal{D}$ to train the actors and critics using Stochastic Gradient Descent (SGD). More specifically, given a mini-batch of samples $\mathcal{B} = \{(\boldsymbol{s}^j, \boldsymbol{a}^j, \boldsymbol{r}^j, \boldsymbol{s}'^j)\}_j$ where $\boldsymbol{r}^j = (r_i^j)_{i=1}^K$ are the rewards of the agents, the critic network θ_i^Q of agent i is trained by minimizing the loss

$$L(\theta_i^Q) = \frac{1}{|\mathcal{B}|} \sum_{j=1}^{|\mathcal{B}|} \left(y_i^j - Q_i\big(\boldsymbol{s}^j, (a_k^j)_{k=1}^K | \theta_i^Q\big) \right)^2. \quad (1)$$

Let $y_i^j \triangleq r_i^j + Q_i'(\boldsymbol{s}'^j, (a_k')_{k=1}^K | \theta_i^{Q'})\big|_{a_k' = \mu_k'(o_k'^j | \theta_k^{\mu'}), \forall k}$ be the regression target generated by two *target networks* $Q_i'(\cdot|\theta_i^{Q'})$ and $\mu_i'(\cdot|\theta_i^{\mu'})$. Note that $\boldsymbol{s}'^j \triangleq (o_k'^j)_{k=1}^K$ denotes the joint observations following the sample $\boldsymbol{s}^j$ in time. Also note that in (1) the Q-function of agent i has an input including the local observations and actions of all agents which is made possible by the use of the replay buffer. The actor network θ_i^μ of agent i is trained by minimizing

$$L(\theta_i^\mu) = -\frac{1}{|\mathcal{B}|} \sum_{j=1}^{|\mathcal{B}|} Q_i\big(\boldsymbol{s}^j, (a_k^j)_{k \neq i}, a_i | \theta_i^Q\big)\big|_{a_i = \mu_i(o_i^j | \theta_i^\mu)} \quad (2)$$

which is equivalent to maximizing the Q-function. Note that in the input of the Q-function in (2), the action of agent i is generated by the actor θ_i^μ while the actions of other agents are sampled from the buffer. Finally, the DNN parameters of the target networks are updated according to

$$\begin{aligned} \theta_i^{Q'} &\leftarrow \tau\theta^Q + (1-\tau)\theta_i^{Q'}, \\ \theta_i^{\mu'} &\leftarrow \tau\theta^\mu + (1-\tau)\theta_i^{\mu'} \end{aligned} \quad (3)$$

for some small number $\tau \in (0, 1)$. In MA-DDPG, exploration is achieved by adding a random noise n_t to the actor output, i.e., $a_i^{(t)} = \mu_i(o_i^{(t)}|\theta_i^\mu) + n_t$ which is clipped to within a proper range to be consistent with the action space definition.

III. DRL-based Spectrum Sharing Approach

We summarize the MA-DDPG-based spectrum sharing scheme in the authors' previously published papers [10], [11] as follows. Each 5G base-station (BS) i is modeled as an RL agent that is equipped with an actor for action selection and a critic for evaluating the Q-function. In particular, the actor μ_i is represented by a DNN θ_i^μ and produces action $a_i = \mu_i(o_i|\theta_i^\mu)$, with o_i being the observation of agent i. The critic Q_i is also represented by a DNN θ_i^Q. It is assumed that each BS does not know the actions of other BSs. The MA-DDPG algorithm with a centralized-training-distributed-execution framework is adopted as shown in Fig. 1. The system is fully synchronized

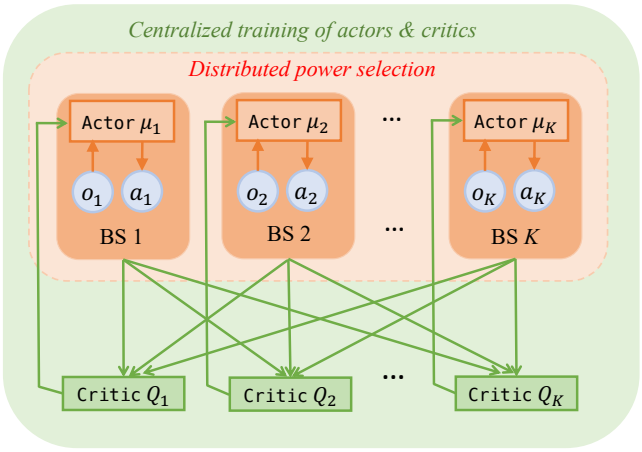

Fig. 1. Proposed MA-DDPG-based power allocation & spectrum sharing scheme. Each BS k is equipped with an actor network μ_k and a critic network Q_k.

where all BSs choose their powers simultaneously at the beginning of each slot based on their local observations. The system then transitions to a new state and the experiences of all agents are pushed to a centralized replay buffer $\mathcal{D}$ in the form of $(o_i, a_i, r_i, o_i'), i = 1, \cdots, K$. The replay buffer operates in a FIFO manner so the oldest experiences are ejected when the buffer is full. In each slot, the actor and critic networks are trained with mini-batches of data sampled from $\mathcal{D}$ according to (1), (2). In addition, the parameters of the target networks are updated softly according to (3). The action, observation and reward are defined as follows.

Action. Each BS needs to determine the transmit power $p_i^{(t)} \in [0, p_i^{\max}]$ to its scheduled UE. Since the tanh() activation is used at the output layer of the actor networks, the actor output $a_i^{(t)} = \mu_i(o_i^{(t)}|\theta_i^\mu)$ falls into $[-1, 1]$. To achieve exploration, a random noise n_t is added to $a_i^{(t)}$, which is then clipped to within the range $[-w, w]$ for some $w \in (0, 1)$. Therefore, the actor output is mapped to the powers by $p_i^{(t)} = \frac{a_i^{(t)} + w}{2w} p_i^{\max}$.

Observation. This step is crucial since the observation must contain useful information of the radio environment that helps the BS make better power selections. In particular, the observation of BS i (in slot t) is defined as $o_i^{(t)} \triangleq o_{i,1}^{(t)} \cup o_{i,2}^{(t)}$ where

$$\begin{aligned} o_{i,1}^{(t)} &\triangleq \left\{ p_i^{(t-1)}, g_{ii}^{(t-1)}, g_{ii}^{(t)}, I_i^{(t-1)}, \widehat{I}_i^{(t)}, C_i^{(t-1)}, \frac{C_i^{(t-1)}}{\sum_{k=1}^K C_k^{(t-1)}} \right\}, \\ o_{i,2}^{(t)} &\triangleq \left\{ g_{ij}^{(t-1)} p_j^{(t-1)}, g_{ij}^{(t)} p_j^{(t-1)}, C_j^{(t-1)}, j = 1, ..., K \right\}. \end{aligned} \quad (4)$$

The first part $o_{i,1}^{(t)}$ contains several general measurements local to BS i. In particular, $p_i^{(t-1)}$ is BS i's power in the previous slot, $g_{ii}^{(t-1)} = \mathrm{PL}(d_{ii}) G_{ii}^{\mathrm{Tx}} G_{ii}^{\mathrm{Rx}} |h_{ii}^{(t-1)}|^2$ is the direct channel gain in slot $t-1$ which can be estimated via pilot training, $g_{ii}^{(t)} = \mathrm{PL}(d_{ii}) G_{ii}^{\mathrm{Tx}} G_{ii}^{\mathrm{Rx}} |h_{ii}^{(t)}|^2$ is the direct channel gain in slot

t. Since the channel changes from $h_{ii}^{(t-1)}$ to $h_{ii}^{(t)}$ at the very beginning of slot t before the new power $p_i^{(t)}$ is determined, $g_{ii}^{(t)}$ can also be obtained by BS i. $I_i^{(t-1)} = \sum_{j\neq i} g_{ij}^{(t-1)} p_j^{(t-1)} + \sigma^2$ is the total received interference at BS i in slot $t-1$ and $\widehat{I}_i^{(t)} = \sum_{j\neq i} g_{ij}^{(t)} p_j^{(t-1)} + \sigma^2$ is the interference measured at the beginning of slot t where the channels have changed but the powers have not changed. $C_i^{(t-1)}$ is the throughput of BS i in slot $t-1$, and $C_i^{(t-1)} / \sum_j C_j^{(t-1)}$ represents the contribution of BS i to the total throughput. The second part $o_{i,2}^{(t)}$ contains measurements of received power and throughput of other BSs. In particular, $g_{ij}^{(t-1)} p_j^{(t-1)}$ and $g_{ij}^{(t)} p_j^{(t-1)}$ are the received power of BS $j(\neq i)$ measured at BS i in slot $t-1$ and the beginning of slot t respectively. $C_j^{(t-1)}$ is the throughput of BS j in the previous slot. Note that $C_j^{(t-1)}$ has to be delivered to BS i from BS j despite all other interference measurements can be directly obtained by BS i. We include one previous slot in order for the agents to better keep track of the temporal correlated channels.

State transition. The measured interference in each BS's observation (4) is determined by the random small-scale fading and the BS transmit powers. Given the current state (i.e., joint observations of all BSs) in slot t, once the powers are chosen, the system will transition to a random new state.

Reward. The reward of BS i is defined as the total throughput of the network, i.e., $r_i^{(t)} = \sum_j C_j^{(t)}$. This definition is intuitive and avoids complicated reward design at the cost of a slight communication overhead.

IV. Experimental Evaluation

The proposed spectrum sharing scheme is evaluated on two network configurations where three and seven BSs belonging to different operators share a spectrum band with possible interference if they transmit simultaneously (See Fig. 2 and Fig. 4). The detailed experiment setup is described as follows. The transmit beams (with 60° coverage) are aligned with the UEs. The maximum power is chosen as $p_i^{\max} = 39$ dBm for all BSs. The BS antenna gain, main to side-lobe ratio (MSR), is chosen as 20 dB. The total noise is calculated according to $\sigma^2\,(\text{dBm}) = 10\lg(\kappa_B T_0 \times 10^3) + \text{NR}\,(\text{dB}) + 10\lg W$ where W is the shared bandwidth. κ_B, NR and T_0 denote Boltzmann's constant, receiver noise figure and temperature respectively. Given the $W = 500$ MHz bandwidth at 6 GHz and the typical values of NR $= 1.5$ dB, $T_0 = 290$ K, we have $\sigma^2 = -85.49$ dBm. The small-scale Nakagami fading parameters are chosen as $m = 50, \Omega = 1$ with temporal correlation of $\rho = 0.1$. Each actor and critic is represented by a fully-connected feedforward DNN with five layers including the input and output layers. The three hidden layers contain $64, 128$ and 64 neurons respectively with ReLU activation. Each actor network has one output port with Tanh activation clipped to within the range $[-0.96, +0.96] (w = 0.96)$. The proposed scheme is implemented with PyTorch. The Adam optimizer is used with learning rates 10^{-4} and 10^{-3} for the actor and critic networks respectively. The action noise $\{n_t, \forall t\}$ is chosen as i.i.d. Gaussian noise with a decreasing variance, i.e., $n_t \sim \mathcal{N}(0, \sigma_t^2)$ where $\sigma_{t+1} = \max\{(1-10^{-4})\sigma_t, 0.001\}$, $\sigma_0 = 1$. This guarantees adequate exploration in the early stage of learning. The replay buffer is implemented as a FIFO queue with size $|\mathcal{D}| = 2 \times 10^5, \forall i$. Batch size is chosen as $|\mathcal{B}| = 64$. Other parameters are chosen as $\gamma = 0.9, \tau = 0.005$. The proposed scheme is compared with the state-of-the-art power allocation scheme WMMSE [12] which is a centralized algorithm that requires the knowledge of all cross channels.

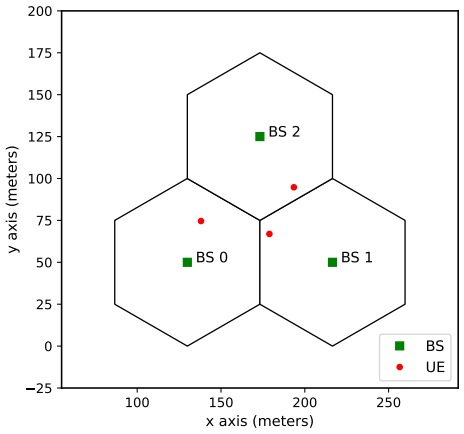

Fig. 2. Network with three BSs.

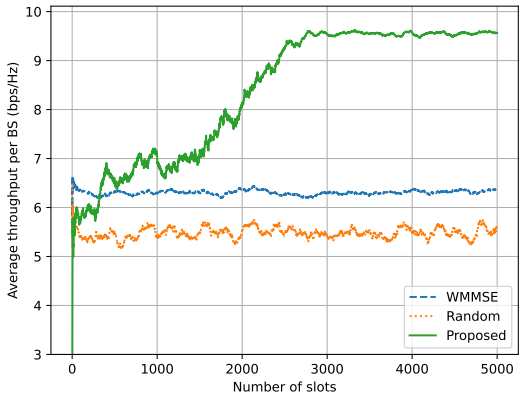

Fig. 3. Throughput of the three-BS network.

Fig. 3 shows the average per-BS throughput achieved along the learning process. It can be seen that the proposed scheme converges in around 3000 iterations and outperforms the baselines by a significant margin, demonstrating its superiority. Note that although the actor and critic networks are trained in a centralized manner in the proposed scheme, the power allocation phase is fully distributed, i.e., each BS chooses its power based solely on it local measurements. Fig. 5 shows the throughput performance in the larger seven-BS network. It can be seen that even with more complex network topologies, the proposed scheme is able to achieve very close performance as

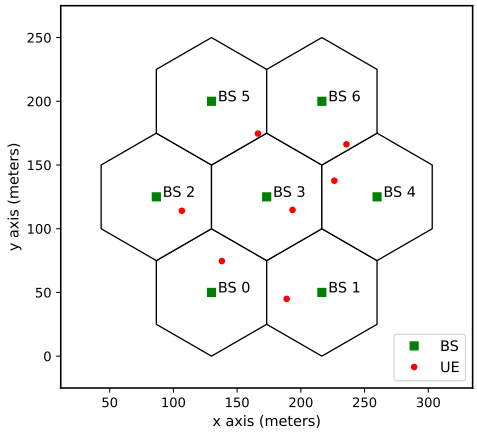

Fig. 4. Network with seven BSs.

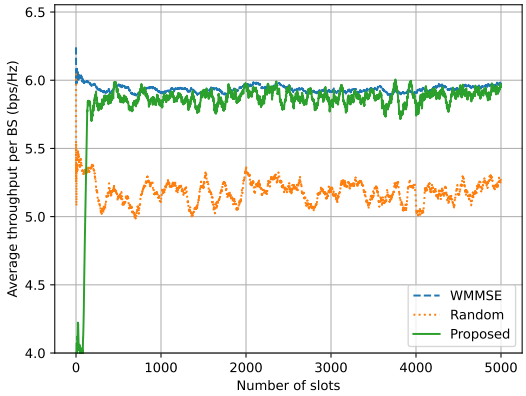

Fig. 5. Throughput of the seven-BS network.

WMMSE.

V. 6G EVOLUTION

6G, the next generation of worldwide cellular standards, will continue the communication transformation started by 5G. 3GPP Rel 19, planned for late 2025, will be the last release focused on 5G. However, it will also include some longer-term technologies that will become the foundation of 6G, thus setting the direction for Rel 20, the first 6G specification in 2026. Meanwhile, research on needed capabilities and key changes in 6G over 5G to realize these new capabilities has been well underway.

6G is expected to increase data rate and reduce latency 10 times over 5G along with similar increases in connected devices and network capacity. It will also introduce 1) Joint Communications and Sensing (JCAS), a transformational change for the 6G BS to sense and geolocate its surrounding space, thereby enhancing its capabilities significantly; 2) use of full duplex radio; 3) use of new sub-THz frequency bands (100-300 GHz) with Giga MIMO (gMIMO) antenna arrays composed of significantly higher number of antennas; and 4) pervasive use of AI/ML to become AI Centric.

Efficient use of spectrum will remain critical for the projected 6G improvements and new capabilities. Secure coexistence of government and commercial networks to share spectrum will be needed at a larger scale to harvest the power of 6G for economic growth worldwide. Distributed and autonomous spectrum sharing solutions such as the one presented in this paper can be further enhanced to realize secure and scalable spectrum sharing solutions for the future with new capabilities of 6G that include 1) JCAS to assess current spectrum use by 6G BSs , and 2) use of narrower beams in the sub-THz bands to reduce overlap of RF transmissions from multiple networks sharing spectrum, resulting in increased use of the shared spectrum.

ACKNOWLEDGEMENT

Part of the work has been supported by the Department of Energy Office of Chief Information Officer (DOE-OCIO) on the Study of Spectrum Security for advanced wireless systems.

REFERENCES

[1] JOSEPH R. BIDEN JR., "Memorandum on modernizing united states spectrum policy and establishing a national spectrum strategy," 2023. [Online]. Available: http://tinyurl.com/3hcm49ms

[2] The White House, "National spectrum strategy," 2023. [Online]. Available: https://www.ntia.gov/sites/default/files/publications/national_spectrum_strategy_final.pdf

[3] NTIA, "National strategy implementation plan," 2024. [Online]. Available: https://www.ntia.gov/sites/default/files/publications/national_spectrum_strategy_final.pdf

[4] N. Kaminski, J. Beck, R. Smith, and A. Bhuyan, "Toward practical federal spectrum sharing for advanced wireless technologies," in *2024 IEEE DySPAN Conference*. IEEE, May 2024.

[5] RCR Wireless News, "Spending on US LTE/5G CBRS networks," https://tinyurl.com/RCRonCBRSgrowth, Accessed Aug 31, 2024.

[6] S. Dogan Tusha *et al.*, "Evaluating the interference potential in 6 ghz: An extensive measurement campaign of a dense indoor wi-fi 6e network," in *ACM Workshop on Wireless Network Testbeds, Experimental evaluation & Characterization 2023*, October 2023.

[7] A. Hyils, S. Magdalene, and L. Thu;asimani, "Analysis of spectrum sensing data falsification (ssdf) attack in cognitive radio networks: A survey," *Journal of Science & Engineering Education volume 2*, pp. 89–100, 2017.

[8] 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," *arXiv preprint arXiv:1509.02971*, 2015.

[9] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, O. Pieter Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," *Advances in neural information processing systems*, vol. 30, 2017.

[10] X. Zhang, A. Bhuyan, S. K. Kasera, and M. Ji, "A novel multi-agent deep reinforcement learning-enabled distributed power allocation scheme for mmwave cellular networks," in *2023 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2023, pp. 73–79.

[11] ——, "Distributed power allocation for 6-Ghz unlicensed spectrum sharing via multi-agent deep reinforcement learning," in *2023 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2023, pp. 1–6.

[12] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted mmse approach to distributed sum-utility maximization for a mimo interfering broadcast channel," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, 2011.