Collaborative Wideband Spectrum Sensing and Scheduling for Networked UAVs in UTM Systems

Sravan Reddy Chintareddy*, Keenan Roach†, Kenny Cheung†, Morteza Hashemi*

*Department of Electrical Engineering and Computer Science, University of Kansas

†Universities Space Research Association (USRA)

Abstract—In this paper, we propose a data-driven framework for collaborative wideband spectrum sensing and scheduling for networked unmanned aerial vehicles (UAVs), which act as the secondary users to opportunistically utilize detected spectrum holes. To this end, we propose a multi-class classification problem for wideband spectrum sensing to detect vacant spectrum spots based on collected I/Q samples. To enhance the accuracy of the spectrum sensing module, the outputs from the multi-class classification by each individual UAV are fused at a server in the unmanned aircraft system traffic management (UTM) ecosystem. In the spectrum scheduling phase, we leverage reinforcement learning (RL) solutions to dynamically allocate the detected spectrum holes to the secondary users (i.e., UAVs). To evaluate the proposed methods, we establish a comprehensive simulation framework that generates a near-realistic synthetic dataset using MATLAB LTE toolbox by incorporating base-station (BS) locations in a chosen area of interest, performing ray-tracing, and emulating the primary users channel usage in terms of I/Q samples. This evaluation methodology provides a flexible framework to generate *large spectrum datasets* that could be used for developing ML/AI-based spectrum management solutions for aerial devices.

I. INTRODUCTION

Unmanned aerial vehicles (UAVs) have attracted significant interests from communications and networking, robotics, and control societies for exploring novel applications such as on-demand connectivity, search-and-rescue operations, situational awareness, to name a few [1]. In order to truly unleash the potentials of UAVs, real-world and commercial deployments will most likely be in the form of beyond visual line-of-sight (BVLOS) scenarios, which in turn provide easier access to remote or hazardous areas, less human intervention, and reduced cost of operation [2]. For safe operations of multiple UAVs under BVLOS conditions, the NASA and FAA are in the process of defining the Unmanned Aircraft System Traffic Management (UTM) systems [3].

Existing terrestrial mobile networks (e.g., 4G LTE and the upcoming 5G-and-beyond) provide significant wireless coverage with relatively low latency, high throughput, and low cost. This, in turn, will make the cellular network a good candidate for the operation of UAVs in BVLOS scenarios [4]. However, the proliferation of new wireless services and the demand for higher cellular data rates have significantly exacerbated the spectrum crunch that cellular providers are already experiencing. Therefore, developing dynamic spectrum management services to sense, assign, and monitor spectrum usage within the UTM architecture is of utmost importance in order to enable advanced UAV use cases in BVLOS, such as urban air mobility (UAM) and advanced air mobility (AAM) [5].

There exists a multitude of prior works on spectrum management frameworks for ground users [6]–[11]. For instance, the authors in [8]–[10] propose deep learning based wideband spectrum sensing to dynamically detect "spectrum holes". Furthermore, several studies leverage reinforcement learning (RL) techniques for spectrum sharing, assuming that the spectrum sensing results are readily available [11]. While these data-driven spectrum management frameworks for ground users are available, they are not directly applicable for UTM-enabled UAV operations, due to several factors, such as widely different wireless channel models and the overall system architecture [8], [11]. In the context of UAV spectrum systems, the authors in [12], [13] have developed efficient spectrum sharing policies for UAV communications to enhance the spectral efficiency (SE). However, the majority of the existing spatial spectral sensing (SSS) models do not consider the spectrum usage pattern of users under realistic scenarios (e.g., ignoring the I/Q level samples), and/or they consider only a single primary user (PU) or secondary user (SU). Moreover, the problem of joint multi-channel wideband spectrum sensing and scheduling among several SU's has not been fully investigated.

To address these gaps, in this paper, we propose a data-driven model for joint wideband spectrum sensing and scheduling across several UAVs, which act as secondary users to opportunistically utilize detected spectrum holes. Our proposed system model presents a unified framework that is compatible with the UTM deployment models with centralized controlling and monitoring entities (e.g., UAS service suppliers). To make our development more concrete and grounded, the problem of joint spectrum sensing and sharing is formulated as an energy efficiency (EE) maximization in a wideband multi-UAV network scenario. Then, we transform the EE optimization problem into a Markov Decision Process (MDP) to maximize the overall throughput of the SUs. To enable spectrum sensing, we develop a multi-class classification framework to identify spectrum holes based on observed I/Q samples. To enhance the accuracy of the spectrum sensing module, the outputs from the multi-class classification by each individual UAV are fused at the UTM server. In the spectrum scheduling phase, we develop and implement several RL algorithms, including the standard Q-learning methods to dynamically allocate underutilized spectrum sub-channels to multiple UAVs. We further investigate the performance of the

"vanilla" deep Q-Network (DQN) and its variations, including double DQN (DDQN) and DDQN with soft-update.

Furthermore, one of the primary challenges of using machine learning (ML) methods for spectrum management is the need for large amounts of training data. The lack of available spectral data in many cases is a significant obstacle, especially for UAV networks that introduce additional complexity for large-scale experimental data collection. To address this gap, and evaluate the proposed methods, we develop a comprehensive framework for spectrum dataset generation, which accurately models LTE waveform generation and propagation in any environment of interest for UTM-enabled UAV applications. This platform enables modeling cooperative spectrum sensing and sharing for wideband multi-UAV network scenarios, and can be used for scalable generation of large spectrum datasets within an area of interest.

The rest of this paper is organized as follows. In Section II, we review related works. In Section III, we discuss our system model, followed by the problem formulation for joint spectrum sensing and access in Section IV. Section V describes our methodology to generate synthetic spectrum dataset, and Section VI includes our numerical results. Section VII concludes the paper.

II. RELATED WORKS

Works on Spectrum Sensing for UAVs. The authors in [14] address spectrum access and interference management by utilizing SSS for ground based device-to-device (D2D) communications [15], [16]. Furthermore, the authors in [13] extend the usage of SSS to UAVs. The UAVs perform SSS to obtain the received signal strength and compare it with a threshold to identify the spectrum occupancy of a particular uplink channel, similar to energy detection methods. The performance of the SSS is then characterized using the probability of spatial false alarm and probability of spatial miss detection. The SSS methods are not directly applicable to wideband sensing to detect multiple spectrum holes simultaneously by all UAVs.

In addition to the SSS methods, data-driven deep learning (DL) methods for spectrum sensing have been considered in several works [17], [18]. However, multi-channel wideband sensing using DL methods have not been explored by the above studies. To address this, the authors in [8] developed a DL-based fast wideband spectrum sensing. The DL model is based on convolutional neural networks (CNN) and accepts raw I/Q signals and predicts the spectrum holes. This work considers a single PU, a single SU only, and the channel between them is modeled as a Rayleigh fading channel.

In this paper, we consider dataset generation for downlink LTE with multiple transmitters (base stations) and receivers (UAV), thereby incorporating interference caused by neighboring cells. Moreover, we use real deployment LTE BS locations for transmitters and employ ray-tracing to effectively model the dynamic UAV environment. Additionally, we consider collaborative spectrum sensing and aggregate the spectrum hole results. We show that spectrum fusion improves the reliability of overall spectrum hole detection.

Works on Spectrum Sharing methods. Spectrum sharing solutions for ground based communications has been extensively investigated (e.g., [6], [7], [19], [20]). The authors in [10], [21] propose the use of RL for dynamic spectrum access in multi-channel wireless networks. Moreover, [11], [22], [23] proposes the use of DQN where in each time slot a single SU decides to stay idle or transmit using one of the sub-channels in a multi-channel environment. While these studies have provided significant insights, they consider one SU only.

Compared with the existing works in this area, the key contributions of this paper are as follows: **(i)** we develop a *joint* spectrum sensing and access framework using raw LTE I/Q data, **(ii)** the spectrum sensing module identifies multiple spectrum holes in a wideband multi-channel setting, **(iii)** we use RL-based techniques (i.e., DQN and its variations) to allocate the identified spectrum holes to multiple UAVs, and **(iv)** the developed solution is based on realistic channel modeling between several LTE BSs and UAVs.

III. SYSTEM MODEL

Network Model and Communication Protocol. We consider a set of UAVs denoted by $\mathcal{K}$ ($|\mathcal{K}| = K$) where each UAV is capable of performing wideband sensing over M orthogonal primary spectrum resources (sub-channels) independently. Due to the highly dynamic environment in which UAVs operate, it may not be feasible for all the UAVs to observe every vacant sub-channel. Therefore, we leverage a collaborative spectrum sensing by the UAVs, and perform spectrum fusion at the UTM server to increase the reliability of spectrum hole detection. Identified spectrum holes are allocated to the UAVs. Therefore, the overall system model is divided into two major components: (i) collaborative spectrum sensing and fusion policies, (ii) spectrum allocation and access policies.

To coordinate the spectrum sensing, fusion, and access steps, we assume that each time slot is divided into four consecutive sub-slots: UAV resource request (t_{req}), spectrum sensing (t_s), broadcasting to server (t_b), and channel access (t_a). Specifically, at the beginning of each time slot, the UAVs that require PU resources request the server for resource allocation. In the subsequent sub-slot of sensing (t_s), the UAVs perform spectrum sensing and broadcasts the sensed channel information in the following sub-slot (t_b). The server then applies fusion rules and assigns spectrum holes to the requesting UAVs. The UAVs transmit on the allocated spectrum holes in the access sub-slot (t_a).

Collaborative Spectrum Sensing and Fusion Policies. Each individual UAV captures the raw I/Q samples from over the air received signals and predict the availability of spectrum holes across M sub-channels. We assume that there is an associated spectrum sensing cost for each UAV k involved in sensing at time slot t. The spectrum sensing cost is the energy consumed for sensing the spectrum and is proportional to the voltage V_{CC} of the receiver, the system bandwidth B, and the duration allotted for sensing (t_s) [24]. Therefore, it is defined as $SC_{km}(t) = t_s V_{CC}^2 B_m$. Upon the completion of sensing phase, the UAV k has a predicted spectrum occupancy vector
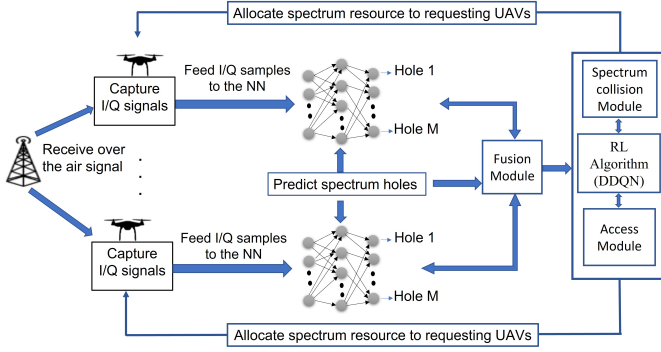
Fig. 1. Proposed system model.

$\mathbf{h}_k(t) = [h_{k,1}(t), ..., h_{k,M}(t)]$ such that $h_{k,m}(t) = 0$ if the m-th sub-channel is detected vacant at time t, and $h_{k,m}(t) = 1$ otherwise. This problem can be considered as a multi-class classification problem, and we leverage deep neural network (DNN) at each UAV to identify the spectrum holes and outputs the prediction vector $\mathbf{h}_k(t)$.

The UTM server receives multiple copies of spectrum holes detected by individual UAVs and applies fusion rules that results in aggregated spectrum holes. In this paper, we use n-out-of-N fusion rule that is defined as follows:

$$f_m(t) = \begin{cases} 0, & \text{if } \sum_{k \in \mathcal{K}} \mathbb{1}\{h_{k,m}(t) = 0\} \geq n; \\ 1, & \text{otherwise,} \end{cases} \quad (1)$$

where $\mathbb{1}\{.\}$ is an indicator function. In this case, $\mathbf{f}(t) = [f_1(t), ..., f_M(t)]$ is the fused prediction of all the M sub-channels at the UTM server. Note that when $n = 1$, the n-out-of-N rule is equivalent to the "OR" rule, and $n = N$ is the same as the "AND" rule.

Spectrum Allocation and Access Policies. Based on the aggregated fusion result, the UTM server then allocates sub-channels to the requesting UAVs. The UAVs then transmit data on the sub-channels allocated to them by the server in the next time step. The transmission energy consumption is denoted by $AC_{km}(t)$. The access cost is the energy consumed for data transmission and is defined as $AC_{km}(t) = t_a P_{tx}$, where, P_{tx} is the transmit power and t_a is the time allotted to transmission. Furthermore, the transmission utility is the amount of data transmitted on the allocated sub-channel and is defined as follows:

$$R_{km}(t) = t_a B_m \log_2 \left(1 + \text{SINR}_{k,m}(t)\right), \quad (2)$$

where B_m is the sub-channel bandwidth allocated for data transmission and SINR_{km} is the signal-to-interference-plus-noise ratio observed on the link between UAV and its receiver over sub-channel m.

Since it is delayed a transmission, the UAVs transmit on the spectrum holes in the current step that are detected vacant in the previous time step. Spectrum collision occurs when the previously detected spectrum holes are no longer available at the current time step. We assume that the true state of sub-channel m is denoted by $\bar{f}_m(t)$. To capture this, we define the spectrum access collision indicator $r_{km}(t)$ as follows:

$$r_{km}(t) = \begin{cases} 1, & \text{if } \bar{f}_m(t) = 0 \text{ and } f_m(t-1) = 0; \\ -1, & \text{if } \bar{f}_m(t) \neq 0 \text{ and } f_m(t-1) = 0; \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

IV. JOINT SPECTRUM SENSING AND ACCESS

Given the presented model, we cast the problem of joint spectrum sensing and access as an energy efficiency optimization for the UAVs. The overall system model of collaborative spectrum sensing and access is shown in Fig. 1 with the overall algorithm described in Algorithm 1.

Energy Efficiency Optimization. Let $y_{km}(t) = 1$ if UAV k is scheduled to use sub-channel m at time t, and $y_{km}(t) = 0$ otherwise. Given that the spectrum holes are allocated to the requesting SUs based on the sub-channel availability we incorporate sensing and access cost to maximize the overall energy efficiency (EE) of the system. We formulate EE problem as an optimization problem as follows:

$$\begin{cases} \max_{\{y_{km}(t)\}} & \mathbb{E}\left\{ \sum_{t,k,m} \frac{y_{km}(t) \, r_{km}(t) \, R_{km}(t)}{y_{km}(t) AC_{km}(t) + SC_{km}(t)} \right\} \\ \text{subject to:} & \sum_m y_{km}(t) \leq 1, \ \forall \ k = 1, 2, 3, \ldots K, \\ & \sum_k y_{km}(t) \leq 1, \ \forall \ m = 1, 2, 3, \ldots M, \\ & \sum_{k,m} y_{km}(t) \leq M - |\mathbf{f}(t)|, \\ & y_{km}(t) \in \{0, 1\}, \end{cases} \quad (4)$$

where $R_{km}(t)$, $SC_{km}(t)$, and $AC_{km}(t)$ are the amount of data transmitted, the sensing cost, and transmission cost by the SU k on sub-band m simultaneously. The constraints guarantee that each UAV is scheduled to use at most one sub-channel, while the total number of scheduled UAVs is at most equal to the number of detected spectrum holes at time t, which is $M - |\mathbf{f}(t)|$.

This optimization problem is a fractional integer programming problem, which is NP-hard in general. If we consider maximizing the numerator alone, which is the total utility of the UAVs over all sub-channels, the problem will become an integer programming problem. In this case, the utility would depend on the spectrum usage pattern by the PUs, which is captured by $r_{km}(t)$ as well as the channel condition between the BSs and UAVs that determine the amounts of transmitted data $R_{km}(t)$. To tackle this utility optimization problem, we model the channel occupancy $\bar{f}_m(t)$ as a Markov process, which enables us to use an Markov decision process (MDP) formulation to solve this problem [25] and develop a dynamic spectrum allocation policy to the SUs.

Dynamic Spectrum Allocation Using RL. As we assume there exists M sub-channels in the system, each sub-channel can be modeled as an independent two-state Markov chain. The transition probability function $\mathbf{P}$, can then be viewed as a set of transition probability matrices $\{\mathbf{P}_i\}$ for each sub-channel that capture the randomness in the assumed multi-user multi-channel environment. Hence, we can formulate the total utility of the SUs into a traditional MDP which is governed by the tuple $(\mathcal{S}, \mathcal{A}, \{\mathbf{P}_i\}, U, \gamma)$, consisting of the set of states $\mathcal{S}$, set of actions $\mathcal{A}$, a transition probability function $\{\mathbf{P}_i\}$, a reward function U, and a discount factor γ. To solve

an MDP using RL, an agent learns to make decisions in an uncertain environment by maximizing a cumulative reward over a sequence of actions. Specifically, the agent interacts with an environment by taking actions that transition the system from one state to another, and the agent receives a reward that is commensurate with the merit of the action. The discount factor determines the relative importance of immediate and future rewards. One of the most popular RL methods is Q-learning [25]. The classical Q-learning is table-based, i.e. the values of the Q-function are stored in a table of size $|\mathcal{S}| \times |\mathcal{A}|$. However, when the size of the state and action spaces get large, the complexity of tabular Q-learning becomes cumbersome. For example, with $M = 16$ sub-channels, the Q-table will be of size $65,537 \times 17$.

DDQN-Based Spectrum Allocation. To address the complexity issue, we adapt the deep Q-learning approach in [26] to approximate the Q-function by a neural network $Q_{\boldsymbol{\theta}}$ called Double Deep Q-Network (DDQN) and train its weights $\boldsymbol{\theta}$ using experience replay. As the name suggest, we have two networks when using DDQN where, $Q_{\boldsymbol{\theta}}$ is called the primary network and $Q'_{\boldsymbol{\theta}}$ is called the target network and the weights of the target network are updated periodically. In the original DDQN, the weights of target network are directly copied from the primary network every few episodes. In DDQN-soft, the target networks are updated using polyak averaging to smoothly update the weights ("soft-update") [26].

The input to the DDQN agent is a state **s** of size $1 \times M$. The output of the network is a vector of size $1 \times (M+1)$ that contains the values of the Q-function with respect to state **s** and each of the $M + 1$ actions. In all the hidden layers, we use the rectified linear unit (ReLU) as an activation function.

Algorithm 1 Collaborative Spectrum Sensing and Access

Phase 1 – Spectrum Sensing and Broadcasting

1: **for** each UAV in $\mathcal{K}$ **do**
2:    Capture N I/Q samples from over the air signal, where, $\mathbf{X} \in \mathbb{C}^{N \times 2}$.
3:    Feed I/Q samples to the pre trained ML model that predicts the spectrum holes s.t $f : \mathbf{X} \to \mathbf{h}$.
4:    Broadcast the individual spectrum hole observations $\mathbf{h} \in \{0,1\}^{1 \times m}$ to the fusion center.
5: **end for**

Phase 2 – Spectrum Fusion and Access

6: Apply fusion rule in Eq. (1) to predict spectrum holes $\mathbf{f}(t)$.
7: Allocate a single spectrum hole to each requesting UAV using pre trained RL algorithm, $y_{km}(t)$, s.t. the constraints in Eq. (4) are satisfied.
8: UAVs transmit on the sub-channel allocated in the previous allocated time slot
9: Given the spectrum allocation $y_{km}(t)$ and spectrum access collision indicator $r_{km}(t)$, the total utility U(t) can be computed using Eq. (4).
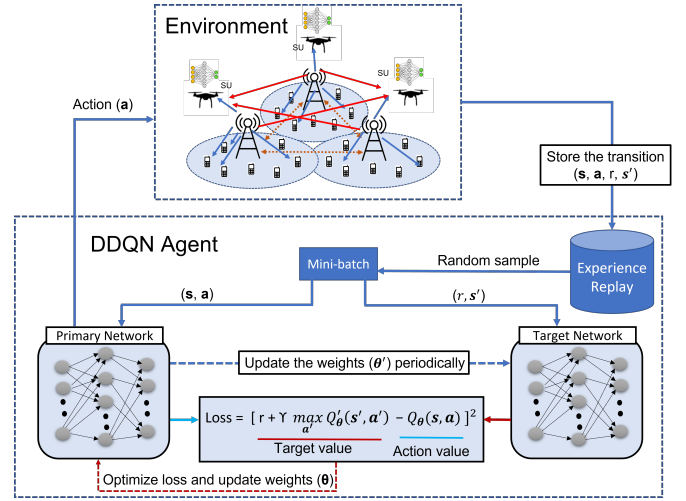

Fig. 2. DDQN for spectrum access.

Given the neural networks input-output dimensions, the overall DDQN architecture and its interaction with the environment is shown in Fig. 2. As shown in Fig. 2, the major components are primary network, target network, experience replay and the interaction with the environment to pick an action.

To train the DDQN agent, the experiences are initially stored in the memory using $\epsilon-$greedy policy i.e., for a state s_t, an action a_t is taken randomly with probability ϵ_t or taken greedily with probability $1 - \epsilon_t$ from the current state of the DDQN network. Then, when we have sufficient samples in the memory a mini-batch of B experiences $\{(\mathbf{s}_i, \mathbf{a}_i, r_i, \mathbf{s}'_i)\}_i \in \mathcal{B}_t$ is randomly sampled from the memory for every time step t to train the neural networks. Here, $\mathcal{B}_t$ is the set of experiences currently available in the memory. Based on the mini-batch selected, we compute and update the weights $\boldsymbol{\theta}$ of the primary network $Q_{\boldsymbol{\theta}}$ that minimize the loss function $L_t(\boldsymbol{\theta})$. Fig. 2 captures the overall DDQN architecture and the agents interaction with the environment [25], [26].

V. RAW I/Q DATASET GENERATION

In the previous section, we described the role of DDQN in allocating spectrum holes to the SUs such that the overall utility is maximized. However, each SU has to first send in their spectrum hole detection results based on observed I/Q samples. Implementing a data-driven ML model for such wideband sensing demands large amounts of raw I/Q data. While it is desirable to capture over-the-air raw I/Q signals using actual hardware, it is challenging to accomplish this goal due to the intricate nature of flying multiple UAVs within a specific environment for collaborative sensing implementation. Hence, we resort to generating synthetic dataset that accurately resembles collecting dataset via experimentation. To this end, we use MATLAB's LTE toolbox as outlined in [8] and extend the dataset generation to incorporate UAV specifics. In particular, we generate the dataset by incorporating the LTE base-stations and UAVs locations, as well as 3D environment including buildings and vegetation for performing ray-tracing.

Dataset Generation Methodology. As shown in Fig. 3, we assume three neighboring cells. In order to perform ray-

3067

Fig. 3. Simulation setup for dataset generation

Fig. 4. **Precision** performance at location 3.

Fig. 5. **Recall** performance at location 3.

tracing experiments, the maps required are downloaded from OpenStreetMap [27].

The simulation area we considered is of 3 km × 3 km wide with buildings and vegetation. The location of the base stations are obtained from cellmapper [28], an open crowd sourced cellular tower and coverage mapping service. We define base-stations as the transmitter sites and UAV locations as the receiver sites. Furthermore, we consider three UAVs and three base stations in the region of interest, and use MATLAB's ray-tracer to find the channel between UAV and base-station locations. Note that this scenario can be easily extended to any numbers of LTE cells and UAVs. It is important to emphasize that we consider UAVs hovering in one location. Nevertheless, by executing the ray-tracing engine multiple times for different locations, we can effectively simulate a flight trajectory.

In the next step, we use MATLAB's LTE Toolbox to generate LTE-M waveform. The entire cell bandwidth of 10 MHz (50 resource blocks) is assumed to be split into 16 sub-channels each of size three resource blocks. In general, the base station can allocate a single sub-channel or multiple sub-channels to a PU to transfer user specific data on the downlink shared channel and also can use multiple access techniques for transmitting data to different PUs. However, when we generate the dataset we assume that base station partitions the bandwidth into separate sub-channels. While generating the downlink waveform of the base station we do not generate any UE specific reference signals. Also, we do not corrupt the broadcast channels with user-specific data. We find the appropriate indices and embed the data samples into the downlink shared channel. As we assumed we partition the cell bandwidth into 16 sub-channels, considering each combination as a label, the base station can generate 2^{16} labels, ranging from no sub-channel allocation to a fully busy cell site. Since we assumed all SUs are capable of wideband sensing, each SU samples the RF signal and stores I/Q samples. We vary the noise variance such that the effective SINR varies from -10 dB to 20 dB in steps of 10 dB. Here, the interference caters for the additional received signals from the neighboring base stations which are modeled using the ray-tracing setup as shown in Fig. 3.

VI. NUMERICAL RESULTS

Collaborative Spectrum Sensing Results. As mentioned earlier, identifying spectrum holes falls into the realm of classical multi-class classification problem. We consider Precision, Recall, and F1-score as the metrics to assess the performance of such classifier. The general definitions of Precision and Recall are as follows:
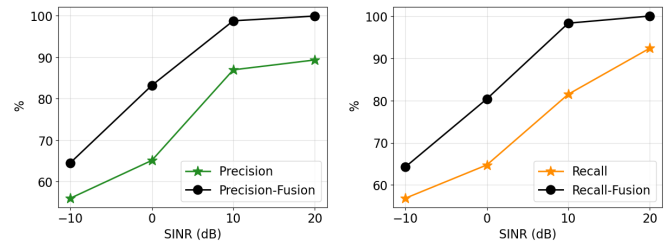
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (5)$$

where TP, FN, FP accounts for the number of true positives, false negatives, and false positives, respectively. To concretely capture the performance of spectrum sensing across 16 sub-channels, we compute the micro-averages for Precision and Recall [29].

We use 70% of the samples generated to train the DNN while the rest of samples are used for testing and validation purposes. We observed that the performance metrics improve as the SINR improves. Specifically, for UAV locations 1 and 2, the spectrum sensing performance metrics are greater than 90% for the SINR values above 10 dB. However, for UAV location 3, the overall performance metrics are worse than locations 1 and 2. As shown in Fig. 4 and 5, the performance metrics are about 90% at 20 dB SINR. This is due to weaker received signal strength in location 3, leading to more noisy I/Q samples. Hence, the ML model was not able to predict the spectrum holes accurately, thereby making the rationale for exploring collaborative sensing more apparent. From the spectrum fusion results (described in Eq. (1)) shown in Fig. 4 and 5, we note that incorporating predictions from UAV locations 1 and 2 significantly improved the spectrum prediction performance.

Resource Allocation and Spectrum Access Results. As mentioned in Section IV, we use deep Q-learning methods for allocating spectrum resources to the UAVs. In Fig. 6, we compare the training performance of three variants of Q-learning methods for allocating a sub-channel to a single UAV whenever the fusion rule detects at least a single spectrum hole. It is observed that DDQN with soft update performs slightly better and converges earlier than DDQN and vanilla-DQN. Next, we extend the model to allocate spectrum holes to two UAVs. In this case, we have augmented the DDQN algorithm with soft update to generate two best actions. From the results in Fig. 7, we observe that the utility performance with two SUs is slightly less than two times of the performance with a single SU. We further note that this paper tries to explore the possibility of integrating spectrum sensing and sharing by making use of existing RL algorithms. Though we explored Q-learning techniques, different RL algorithms can be integrated into the proposed framework.

VII. CONCLUSION

In this paper, we proposed a collaborative wideband spectrum sensing and sharing method for networked UAVs. We
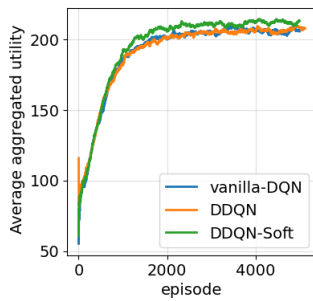
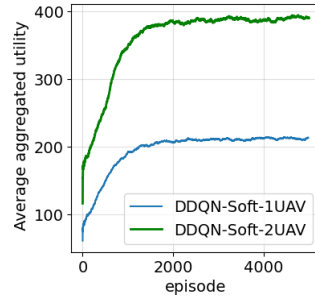Fig. 6. Training results for allocating spectrum holes to **one UAV**.

Fig. 7. Training results for allocating spectrum holes to **two UAVs**.

implemented a multi-class classifier as the spectrum sensing module to detect spectrum holes based on observed I/Q samples. To enhance the accuracy, we considered the spectrum fusion rule at the UTM server. By leveraging deep Q-learning methods, detected spectrum holes are dynamically allocated to the secondary users (i.e., UAVs). To evaluate the proposed methods, we generated a near-realistic synthetic dataset using MATLAB LTE toolbox by incorporating base-station locations in a chosen area of interest, performing ray-tracing, and emulating the primary users channel usage in terms of I/Q samples. Overall, our numerical results suggest that collaborative spectrum sensing to allocate spectrum holes to multiple SUs provide promising results. As a future work, we are planning to conduct aerial tests to collect real-world dataset and further refine our framework.

VIII. ACKNOWLEDGEMENT

REFERENCES

[1] H. Menouar, I. Guvenc, K. Akkaya, A. S. Uluagac, A. Kadri, and A. Tuncer, "UAV-enabled intelligent transportation systems for the smart city: Applications and challenges," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 22–28, 2017.

[2] S. Li, Y. Gu, B. Subedi, C. He, Y. Wan, A. Miyaji, and T. Higashino, "Beyond visual line of sight UAV control for remote monitoring using directional antennas," in *2019 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2019, pp. 1–6.

[3] P. Kopardekar, J. Rios, T. Prevot, M. Johnson, J. Jung, and J. E. Robinson, "Unmanned aircraft system traffic management (UTM) concept of operations," 2016.

[4] A. S. Abdalla and V. Marojevic, "Communications standards for unmanned aircraft systems: The 3GPP perspective and research drivers," *IEEE Communications Standards Magazine*, vol. 5, no. 1, pp. 70–77, 2021.

[5] M. Rimjha and A. Trani, "Urban Air Mobility: Factors Affecting Vertiport Capacity," in *2021 Integrated Communications Navigation and Surveillance Conference (ICNS)*. IEEE, 2021, pp. 1–14.

[6] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 28–41, 2019.

[7] W. S. H. M. W. Ahmad, N. A. M. Radzi, F. Samidi, A. Ismail, F. Abdullah, M. Z. Jamaludin, and M. Zakaria, "5G technology: Towards dynamic spectrum sharing using cognitive radio networks," *IEEE access*, vol. 8, pp. 14 460–14 488, 2020.

[8] D. Uvaydov, S. D'Oro, F. Restuccia, and T. Melodia, "Deepsense: Fast wideband spectrum sensing through real-time in-the-loop deep learning," in *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*. IEEE, 2021, pp. 1–10.

[9] J. Cui, Y. Liu, and A. Nallanathan, "Multi-agent reinforcement learning-based resource allocation for UAV networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 729–743, 2019.

[10] Y. Li, W. Zhang, C.-X. Wang, J. Sun, and Y. Liu, "Deep reinforcement learning for dynamic spectrum sensing and aggregation in multi-channel wireless networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 2, pp. 464–475, 2020.

[11] H. Q. Nguyen, B. T. Nguyen, T. Q. Dong, D. T. Ngo, and T. A. Nguyen, "Deep Q-Learning with Multiband Sensing for Dynamic Spectrum Access," in *2018 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2018, pp. 1–5.

[12] J. Kakar and V. Marojevic, "Waveform and spectrum management for unmanned aerial systems beyond 2025," in *2017 IEEE 28th Annual international symposium on personal, indoor, and mobile radio communications (PIMRC)*. IEEE, 2017, pp. 1–5.

[13] B. Shang, V. Marojevic, Y. Yi, A. S. Abdalla, and L. Liu, "Spectrum sharing for UAV communications: Spatial spectrum sensing and open issues," *IEEE Vehicular Technology Magazine*, vol. 15, no. 2, pp. 104–112, 2020.

[14] B. Shang, L. Liu, H. Chen, J. Zhang, S. Pudlewski, E. S. Bentley, and J. Ashdown, "Spatial spectrum sensing-based D2D communications in user-centric deployed HetNets," in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.

[15] H. Chen, L. Liu, T. Novlan, J. D. Matyjas, B. L. Ng, and J. Zhang, "Spatial spectrum sensing-based device-to-device cellular networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7299–7313, 2016.

[16] H. Chen, L. Liu, H. S. Dhillon, and Y. Yi, "QoS-aware D2D cellular networks with spatial spectrum sensing: A stochastic geometry view," *IEEE Transactions on Communications*, vol. 67, no. 5, pp. 3651–3664, 2018.

[17] C. Liu, J. Wang, X. Liu, and Y.-C. Liang, "Deep CM-CNN for spectrum sensing in cognitive radio," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 10, pp. 2306–2321, 2019.

[18] D. Chew and A. B. Cooper, "Spectrum sensing in interference and noise using deep learning," in *2020 54th Annual conference on information sciences and systems (CISS)*. IEEE, 2020, pp. 1–6.

[19] A. Sharmila and P. Dananjayan, "Spectrum sharing techniques in cognitive radio networks–A survey," in *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*. IEEE, 2019, pp. 1–4.

[20] N. Slamnik-Kriještorac, H. Kremo, M. Ruffini, and J. M. Marquez-Barja, "Sharing distributed and heterogeneous resources toward end-to-end 5G networks: A comprehensive survey and a taxonomy," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1592–1628, 2020.

[21] O. Naparstek and K. Cohen, "Deep multi-user reinforcement learning for dynamic spectrum access in multichannel wireless networks," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–7.

[22] Y. Bokobza, R. Dabora, and K. Cohen, "Deep reinforcement learning for simultaneous sensing and channel access in cognitive networks," *IEEE Transactions on Wireless Communications*, 2023.

[23] S. Wang, H. Liu, P. H. Gomes, and B. Krishnamachari, "Deep reinforcement learning for dynamic multichannel access in wireless networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 2, pp. 257–265, 2018.

[24] X. Zhang and K. G. Shin, "E-MiLi: Energy-minimizing idle listening in wireless networks," in *Proceedings of the 17th annual international conference on Mobile computing and networking*, 2011, pp. 205–216.

[25] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[26] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 30, no. 1, 2016.

[27] OpenStreetMap contributors, "Planet dump retrieved from https://planet.osm.org ," https://www.openstreetmap.org , 2017.

[28] "CellMapper," https://www.cellmapper.net/map.

[29] M. Grandini, E. Bagli, and G. Visani, "Metrics for multi-class classification: an overview," *arXiv preprint arXiv:2008.05756*, 2020.