This is a peer-reviewed, accepted author manuscript of the following conference paper:Atimati, E, Crawford, D & Stewart, R 2024, Intelligent shared spectrum coordination in heterogeneous networks. in 2023 IEEE Virtual Conference on Communications (VCC). IEEE, pp. 252-257, IEEE Virtual Conference on Communications 2023, 28/11/23. https://doi.org/10.1109/VCC60689.2023.10474686

# Intelligent Shared Spectrum coordination in Heterogeneous Networks

Ehinomen Atimati, David Crawford and Robert Stewart
*Electrical and Electronic Engineering*
*University of Strathclyde,*
Glasgow, United Kingdom
*ehinomen.atimati@strath.ac.uk, david.crawford@strath.ac.uk, r.stewart@strath.ac.uk*

***Abstract*—Global connectivity requires reliable and affordable access to the internet for digital inclusiveness. Shared spectrum technologies are one of many technologies that can help provide affordable connectivity. They require available spectrum (channels) from Primary Users (PUs) and share these among future dynamic heterogeneous secondary user (SU) networks. Coordination of these transient available resources is exacerbated by dynamic SU network scenarios, thus raising the risk of poor SU experience and inefficient spectrum and power usage. Therefore, adopting reinforcement learning (RL), terrain-based propagation models, and IEEE 802.19 coexistence principles, a central intelligent real-time shared spectrum coordination algorithm is proposed to coordinate resource allocation among dynamic SUs operating at low bands. The proposed two-stage algorithm was compared to existing shared spectrum allocation techniques deployed in dynamic spectrum access (DSA) networks to quantify RL's impact on low-band wireless networks.** $66\%$ **to** $100\%$ **of SU nodes/access points (APs) that used the proposed algorithm experienced good quality of service (QoS) in most scenarios examined. A good QoS meant that** $75\%$ **of APs receivers experienced signal-to-noise plus interference ratio (SINR) greater than 5. This was achieved using minimal AP transmission power.**

***Index Terms*—shared spectrum, interference, coexistence, q-learning, reinforcement learning, dynamic spectrum access, TVWS, CBRS.**

## I. Introduction

According to the United Nations (UN), over one-third of the world's population are not connected to the internet [1]. The huge benefits of internet access have driven the need for global wireless connectivity. Smart connections of devices and the Internet of Things (IoT) need affordable connections for global connectivity. One means of providing affordable connectivity to these varied networks is the use of shared spectrum technologies [2]. These are technologies that facilitate the shared use of spectral resources. Affordable connectivity can be facilitated by the shared use of unused incumbent/Primary users' (PUs) spectrum by Secondary users (SUs), as this allows for reduced licensing, capital, and operating costs in the communication system [2]. Television White Space (TVWS) and Citizens Broadband Radio Service (CBRS) in the United States share resources with incumbent/PU TV broadcasters and Fixed Satellite Service providers respectively. These Dynamic Spectrum Access (DSA) networks are examples of shared spectrum networks that make use of a central pool of PU/incumbent information in coordinating PU-SU coexistence and decentralized coordination of SUs-SUs coexistence.

DSA network consists of a central spectrum management system (database) that use the pool of PU information and SU's parameters in estimating available spectral resource for SUs. The central spectrum access system/DB, assures PUs of exclusive protection, while SUs decentralized mechanisms provide for flexible and instantaneous access to unused spectrum [3]. To introduce a level of spectral assurance to SUs a tiered system was adopted in CBRS architecture [4]. SUs are categorized into priority access licensees (PALs) and general authorized access (GAAs) licensees, and its Spectrum Access System (SAS) coordinates the exclusive protection of PAL users from interfering GAAs. Thus, GAAs with equal priority (horizontal spectrum access) do not have the same level of exclusive access to spectrum as PUs and PAL users.

SUs, guided by their Media Access Control (MAC) Protocols, use distributed or decentralized techniques in randomly selecting available channels for safe coexistence. These internal coordination schemes are spectrally flexible and device dependent. However, Heterogeneous networks and devices, with varied network architecture and Media Access Control (MAC) Protocols lack a unified or standard communication mechanism reducing the effectiveness of the proposed intelligent decentralized or distributed resource management coordinating system [5], [6]. Relying on the existing communication structure in DSA networks, and the pool of information in their spectrum management systems, we propose an implementable central intelligent spectrum and power resource matching scheme for improved SU-to-SU coexistence. The proposed algorithm provides exclusive resource allocation to GAAs and SUs with horizontal spectrum access and reduces the degree of resource contention managed by MAC or distributed coordination.

## II. Literature Review

There is a myriad of research work on resource allocation techniques that have used mathematical, game theory, heuristic, and reinforcement learning methods [7]. Reinforcement Learning (RL) gives dynamic solutions to non-convex wireless communication resource allocation problems [8]. RL approaches to wireless networks have been categorized into centralized and decentralized (distributed) algorithms based on the network architectures in [9] to address several network challenges.

Distributed RL resource management algorithms were designed to assist user equipment (UE) in effectively maximizing subchannels in DSA networks [10], to select spectrum (channels/subchannels) that enable it to coexist with SUs and PUs in DSA networks (vertical spectrum access) [11], to optimize power used by SUs [12] and jointly optimize power and spectrum [13] usage in different networks. These distributed spectrum management algorithms, have SUs (nodes or UEs) as decision-makers informed by limited/local information from neighbouring SUs. Thus nodes need a communication mechanism necessitating a similar MAC/PHY layer or standards. This was evidenced in the practical implementation of these approaches in [14], [15] where network nodes' MAC/PHY layers had to be re-configured to communicate using wireless or wired communication links in CBRS networks. This may not be possible in certain network scenarios.

A unified central and distributed approach to coexistence management was proposed in the 802.19 standard for resource management in heterogeneous DSA networks [16]. They however do not provide a specific implementation. Similarly, resource management suggested in WInnforum standard for CBRS proposed a recursive reuse of spectrum among nodes with future releases to focus on power and intelligent resource management [17], [18]. A heuristic implemented WInnforum's recursive spectrum allocation in central joint resource allocation of a CBRS network was explored in [19]. These approaches were prescriptive, void of the predictive and adaptive benefits of RL methods.

A central and distributed RL resource management among the same priority nodes was suggested in [20], SUs' agents took decisions that were centrally rewarded based on interference prevention. Their robust RL algorithm for large network sizes focused on optimizing power and assumed an efficient channel allocation. A central coordinating RL algorithm achieved optimal resource allocation through an event-triggering training of their offline policy in a dense WLAN was proposed in [21]. Similarly, in [22] a Graphical Convolution Network (GCN) with RL centrally matched multiple APs to the same spectrum in a densely populated WLAN . The network structure (homogeneous networks) and resource management objective (maximizing throughput) differed from our proposed algorithm as some parameters were not applicable in DSA network.

Closely related work used recursive neural networks and Boltzmann optimization equation in resource allocation for coexistence management among SUs in a DSA network [23]. Also, a Q-learning algorithm was developed in [24] to enable the shared use of spectrum between PUs and SUs (underlay coexistence) at optimal data rates. Centralized approaches have been prescriptive, used different learning algorithms for either spectrum or power resource sharing, and focused on vertical spectrum access. Our centralized intelligent resource management algorithm addresses DSA's unique joint resource sharing challenges among horizontal access heterogeneous networks to promote resource reuse and reduce SUs' contention.

## III. Methodology

### A. Network Architecture

PUs are assumed to be protected adequately by our previously designed database in [25]. SUs frequently communicate their device location and transmission parameters to SAS/database, based on local shared spectrum regulation [16]. The heterogeneous network simulated consists of IEEE 802.11 Television Very High Throughput (TVHT) wireless local area network (WLAN) and IEEE 802.22 wireless regional area network (WRAN) access points (AP) described in Figure 1. In the Figure, the colored APs and channels reflect APs with different MAC protocols and available channels.

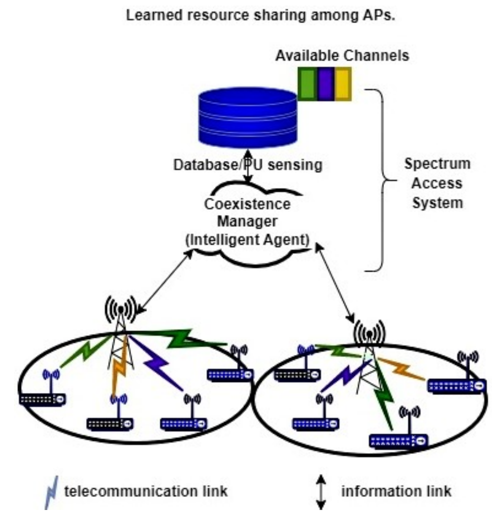


Fig. 1. Intelligent Coordination of Heterogeneous Access Points

### B. Intelligent Coexistence Manager

The objective of SAS, within which the intelligent coexistence manager lies, is to allocate resources such that resources are utilized maximally while SU-to-SU interference is minimized. Future shared spectrum networks may be characterized by dynamic spectra bands and heterogeneous access points that seek to share resources [6]. Identifying the maximum number of APs that can share resources can be challenging when the minimum interference level for SU-to-SU coexistence is not clearly defined. To address this we adopt the 802.19 definition of interference level discovery as the $90^{th}$ percentile experienced by receivers of an AP from another interfering AP. Similarly, finding the balance between reducing the allocated power of an AP to prevent it from interfering with another AP, while maintaining its user equipment's (UEs) QoS, remains a challenge in dynamic and instantaneous networks. We explore Q-learning reinforcement algorithm in solving these optimization problems.

### C. Mathematical Formulation

The objective of the intelligent coexistence manager was split into two: spectrum allocation and power allocation stages.

*1) Spectrum Allocation Phase:* The overall objective of this phase was to allocate spectrum and minimize the interference between APs transmitting at maximum power. Using the idea of graph theory, we attempt to minimize interference in equation (1) by reducing the number of pairwise interference $I_{i,j}^{k}$ between access points (i and j) in order to optimize spectral reuse and reduce interference simultaneously. A pairwise interference existed when the interference level experienced by $m$ (an $i^{th}$ AP's receiver) from a $j^{th}$ AP, exceeded the sensitivity of $m$ [16]. The interference level threshold $I_{level_{i\leftarrow j}}$ in equation (1) and (2) was the $90^{th}$ percentile of received signals strength $(RSS_{i\leftarrow j})$ at 100 $AP_i$ receivers as a result of $AP_j$ transmitting at the same $k$ channel as $AP_i$. At any instance $(t)$, the maximum number of channels assigned to an $i^{th}$ node/AP $(k_{i_{max}})$ is one, however, this channel can be shared by multiple $(M_k)$ nodes defined by $(\beta_i^k(t)$ in equation

(1)). Also, there were more node/APs $(N)$ than available spectrum $(K)$, reflecting spectral contention.

$$min \sum_{j=1}^{N} \sum_{i=1}^{N} I_{i,j}^{k} \quad k = 1, \dots K; \quad i \neq j \tag{1}$$

subject to:

$$k_{i_{max}} = \sum_{k=1}^{K} \beta_i^k(t) = 0 \ or \ 1; \quad i = 1, \dots N;$$

$$M_k = \sum_{i=1}^{N} \beta_i^k(t) \geq 0; \quad k = 1, \dots K;$$

$$K < N$$

$$\beta_i^k(t) = \begin{cases} 1, & \text{if } i^{th} \text{ node is allocated } k^{th} \text{ channel.} \\ 0, & \text{otherwise.} \end{cases}$$

$$I_{i,j}^k = \begin{cases} 1, & \text{if } I_{level_{i \leftarrow j}} \geq sensitivity_{i_{rx}}. \\ 0, & \text{otherwise.} \end{cases}$$

where

$$I_{level_{i \leftarrow j}} = RSS_{90\%} = maxRSS$$
$$subject\ to$$
$$P[RSS_{i \leftarrow j} < RSS] \leq 90\% \tag{2}$$

*2) Power Allocation Phase:* The optimal channel allocation from the previous phase was the input to this phase, aimed at minimizing APs transmitting power while maintaining good QoS for their receivers. Good quality of service was achieved in equation (3) by minimizing the difference between signal-to-noise ratio (SNR) experienced by UEs when an AP was the sole occupier of a channel and its Signal to Interference plus Noise Ratio (SINR) when it shared its resource with other APs.

$$min(SNR_i - SINR_i) \tag{3}$$

subject to

$$P_{min} < P_i < min(P_{max}, P_{db})$$

Where $P_i$ is a set of $AP_i's$ transmitter power bounded between minimum power $(P_{min})$ and maximum power $((P_{max})$ of the AP or the SAS assigned power limit $(P_{db})$.

$$SNR_i = \frac{P_i H_{i,m}}{\sigma^2} \quad i = 1, \dots N, m = 1, \dots M \tag{4}$$

$$SINR_i = \frac{P_i H_{i,m}}{\sigma^2 + \sum_{j=1}^{N} P_j H_{j,m}} \quad i, j = 1, \dots N; \quad i \neq j; \tag{5}$$

SNR and SINR in equations (4) and (5) have $H_{i,m}$ as the channel characteristics defined as path-loss between an $i^{th}$ AP and its $m^{th}$ receiver. The noise power at $m$ is represented as the power of AWGN $(\sigma^2)$. $P_j$ is the Power transmitted by any other $j^{th}$ AP sharing an $i^{th}$ AP's channel and $H_{j,m}$ is channel characteristics between $j$ and $m$ nodes.

### D. Reinforcement Learning Algorithm Design

We proposed the inclusion of an intelligent coexistence manager (an RL algorithm or agent) in the database/SAS such that it uses the APs's ID, location, sensitivity, and antenna height, together with PU's available channels and their permitted power limits to create an RL environment. The RL agent/algorithm is split into two stages:

*1) Spectrum Allocation:* In algorithm 1, the RL environment provides a starting AP's index $(i)$ defined in equation (6) as $(s(t))$, on which an action (equation (7)) of allocating $(a(t) = k)$ or not allocating $(a(t) = 0)$ a channel to the AP was taken. The reward function in equation (8), captures the total number of APs $(M_k)$ sharing a channel $k$, $(M_k \subseteq S)$, and all interfering pair edges (when $I_{ij}^k(t)$ or $I_{ji}^k(t)$ = 1).

$$s(t) = i \in S; \quad S = \{1, ...N\}; \quad and \quad N = n\{S\} \tag{6}$$

$$a(t) = a_i \in \{A\}; \quad A = \{0, k, ..., K\}; \quad K \quad channels \tag{7}$$

$$r_i(t+1) = (2 + e^{-n(M_k)}) \sum_{j=1}^{M_k} \sum_{i=1}^{M_k} I_{i,j}^k \quad k = 1, \dots K; \quad i \neq j; \tag{8}$$

**Algorithm 1:** Spectrum Allocation

**Input:** Initialize $Q(s, a)$ values, learning rate $(\alpha)$, discount factor $(\gamma)$ and $\epsilon$-greedy. Initialize RL environment initial state $(s(t))$ from equation 6
**Output:** $\pi^*(s, a) = Q^*(s, a)$
**for** *epi = 1 to #episodes* **do**
  **for** *t = 1 to #steps* **do**
    from the current state $s(t)$;
    An $\epsilon$-greedy rule:;
    **if** *a random number > $\epsilon$-greedy* **then**
      an action $(a(t))$= $argmax_A(Q(s, a))$ from algorithm's Q-Table is taken
    **end**
    **else**
      A random action (a(t)) from equation 7
    **end**
    Obtain a reward $r_i(t + 1)$ equation 8 and a random next state $(s(t + 1))$ equation 6 from RL environment;
    To minimize number of edges in equation (1) $r(t + 1) = -r_i(t + 1)$;
    From Q-table, obtain all possible next state actions' value $(Q(s', a')$;
    Update the Q-table using equation (12)
  **end**
**end**

*2) Power Allocation:* The agent in algorithm 2 minimized the objective function in equation (3) by observing as its state, an AP's id and its previous transmission power, (equation (9)), and taking an action (equation (10)) of a single-digit increase (2), decreased (1) or no action (0), on the observed power. Controlled by the reward function in equation (11).

$$s(t) = \{i, P_i(t-1)\} \quad P_i(t-1) \in P_i, \ in \ equation \ (3) \quad (9)$$

$$a(t) = \{0, 1, 2\} \quad (10)$$

$$r_p(t+1) = \|SNR(t)–SINR(t)\| \quad (11)$$

**Algorithm 2:** Power Allocation Q-Learning

**Input:** Initialize $Q(s,a)$ values, learning rate $(\alpha)$, discount factor $(\gamma)$,$\epsilon$-greedy, and RL environment initial state $(s(t))$ from equation 9
**Output:** $\pi^*(s,a) = Q^*(s,a)$
**for** *epi = 1 to #episodes* **do**
  **for** *t = 1 to #steps* **do**
    from the current state s(t);
    Follow an $\epsilon$-greedy rule to generate an action: $(a(t))$ same as algorithm 1;
    Obtain a reward $r_p(t+1)$ from equation 11 and a random next state $(s(t+1))$ equation 9 from RL environment;
    Minimize change in SINR in equation (3) $r(t+1) = -r_p(t+1)$;
    Obtain all possible next state $Q$ values from Q-Table $(Q(s', a')$;
    Update the Q-table using equation (12)
  **end**
**end**

*3) Policy search:* The algorithms search through $2^{a(t)}$ policies while taking actions from different starting states, to determine the sequence of actions that assures it of the best reward [26]. The algorithms greedily searched through these options in each iteration and updated a current q-table's $Q(s,a)$ values with $Q'(s,a)$ using the equation (12). A shuffle between a selection of best policy (exploitation) and a thorough search (exploration) was done in each iteration to prevent the algorithms from being stuck at local minimum.

$$Q'(s,a) \leftarrow Q(s,a) + \alpha[R(s,a) + \gamma \arg max_A Q(s', a') - Q(s,a)] \quad (12)$$

where $s = s(t)$, $a = a(t)$ and $Q(s', a')$ = are Q-values from all possible action at the next state $s(t+1)$. $R(s,a)$ is the reward for being in state $s(t)$ and taking action $a(t)$. Learning rate $\alpha$ and discount factor $\gamma$ are training hyperparameters.

## IV. Simulation Results

### A. Simulation Environment

The dynamic heterogeneous network in Figure 1 was simulated in Python, it consisted of WRAN (diamond) and WLAN (circular) APs randomly located in a wide 10 km by 25 km area, to capture low band's wide coverage. The path loss was estimated with the terrain-based Longley Rice propagation model for distances greater than 1 km, as low-frequency bands are susceptible to environmental constraints and with the free-space model at below 1 km. SUs regularly updated their status with the central database, providing real-time training of the RL algorithm. In each iteration, of the algorithm in all scenarios explored the algorithm converged at optimal accumulated rewards as illustrated in Fig. 2.

### B. Resource Allocation

Allocation techniques, random (rand), and recursive(recur) were compared with the first stage of our algorithm (Qfirst), and it behaved similar to other algorithms in the Figure 3 bar chart. The random APs locations made it difficult to share resources maximally, it however did as well as other techniques in terms of the quality of service shown in the figure's box plot. We define a good QoS experience as when only the first quantile of an AP's receivers suffer degradation, by having SINR less than 5.

As the number of APs was increased in Figure 4, our algorithm allocated the same number of channels to APs as other methods in the bar chart. It, however, did better in QoS as it all its APs, transmitting less power, achieved good QoS. It therefore provides less contention among sharing SUs, when compared to other methods.

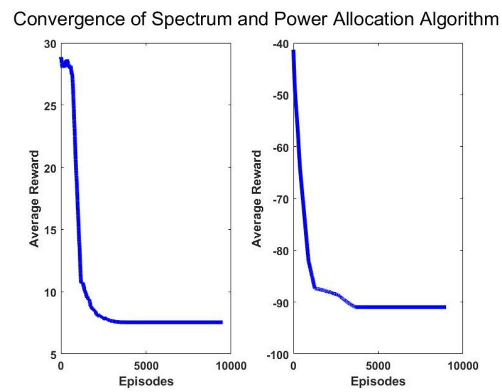


Fig. 2. Convergence of channel and power allocation algorithm

### C. Scalability

Further increase in the number of APs to 6, when only 2 channels were available in Figure 5 bar chart, showed only 4 APs were allocated. Our algorithm learned not to reuse channels when it resulted in more APs having poor QoS. Thus 3 out of its 4 APs enjoyed good QoS in the figure's box plot. In these instances, other methods reallocated these channels resulting poor QoS.

Similarly, when the number of channels were increased to 3 and APs were few (4), the algorithm learned to optimize channels by reusing only two channels as shown in the barchart of Figure 6 while ensuring good QoS of all or most APs. This was repeated if Figure 7 where 4 out of 5 AP had good QoS when it optimally matched 3 channels to 5 APs, even when it had 4 available channels. It therefore chose reuse of a channel based on permitted network interference and QoS degradation. Its overall performance in this figure was similar to random allocation's performance and better than recursive allocation.

### D. QoS Evaluation

In the same vein our algorithm automatically matched 8 APs to 4 channels in Figure 8 and one out of the 8 APs experiencing poor QoS. On average the algorithm's allocation ensured that 66% to 100% of APs allocated enjoyed great QoS in each network scenario investigated. This was slightly better than the 50% to 100% performance of other prescriptive allocation algorithm which could not adapt to varied QoS conditions/scenarios. The algorithm slightly reduced contention as compared to its counterparts, but its performance was significantly affected when the number of APs and Channels went beyond 8 and 4 respectively. This Q-learning algorithms perform poorly in large state spaces, as its states were a function of the number of APs.

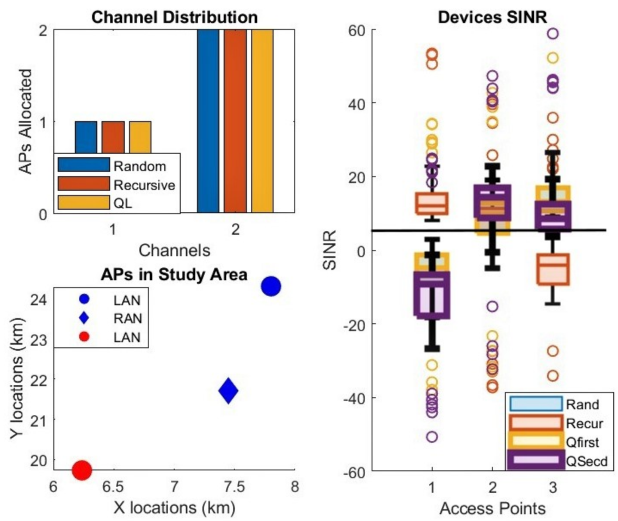


Fig. 3. Comparing allocation techniques $N = 3$ and $K = 2$

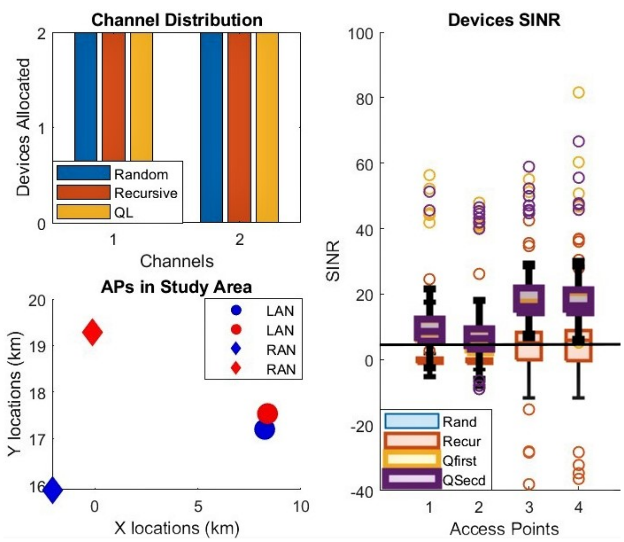


Fig. 4. Comparing allocation techniques $N = 4$ and $K = 2$

## V. Conclusion

Learned algorithms generate predictive responses/allocations to dynamic network demands, unlike prescriptive algorithms, making them responsive to future spontaneous shared spectrum networks. We proposed an

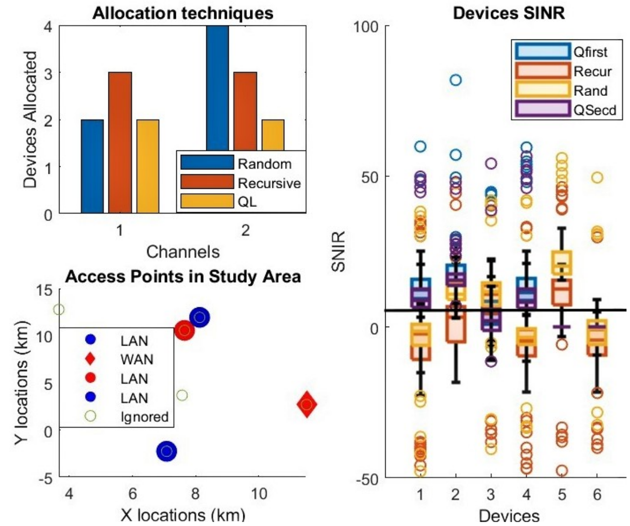


Fig. 5. Comparing allocation techniques $N = 6$ and $K = 2$

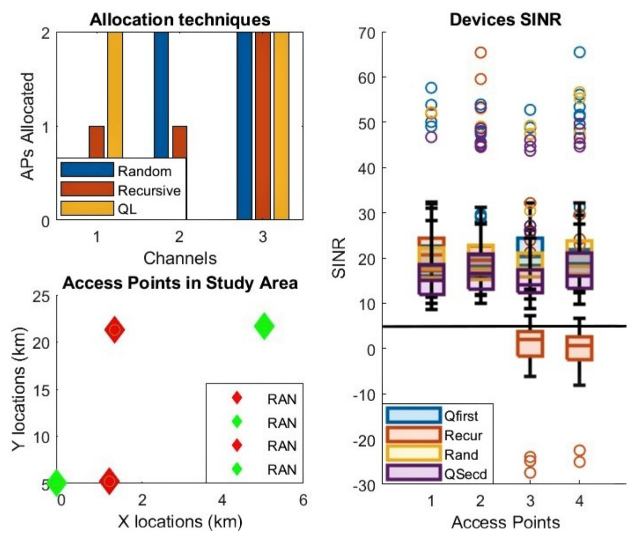


Fig. 6. Comparing allocation techniques $N = 4$ and $K = 3$

intelligent real-time central coordinating resource allocating algorithm for a heterogeneous DSA network. In all scenarios of channel to AP allocation examined, our algorithm ensured that 66% to 100% APs matched with channels experienced good QoS. It struggled to maintain this performance as the size of network increased. Compared to other algorithms, the learning algorithm slightly improved resource contention among SUs as more of its SUs enjoyed good QoS prior to any contention management measures. The algorithm's reuse of channels and reduced power consumption decreased resource contention in small sized networks and can foster affordable connectivity.

## Acknowledgment

Work by this author was (partially) supported by the Schlumberger Foundation Faculty for the Future program.

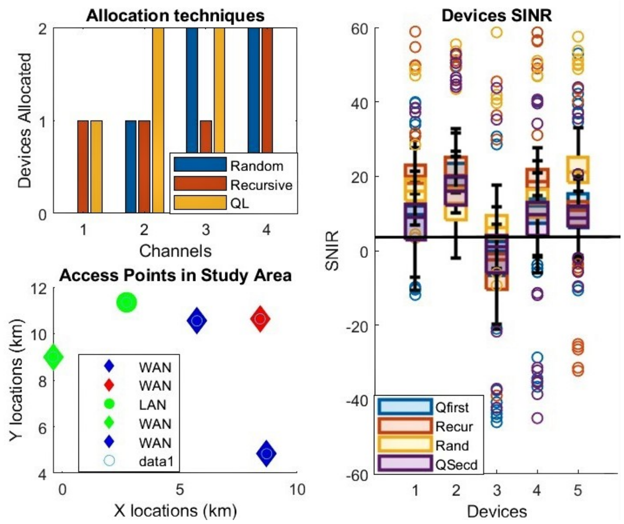


Fig. 7. Comparing allocation techniques $n = 5$ and $k = 4$

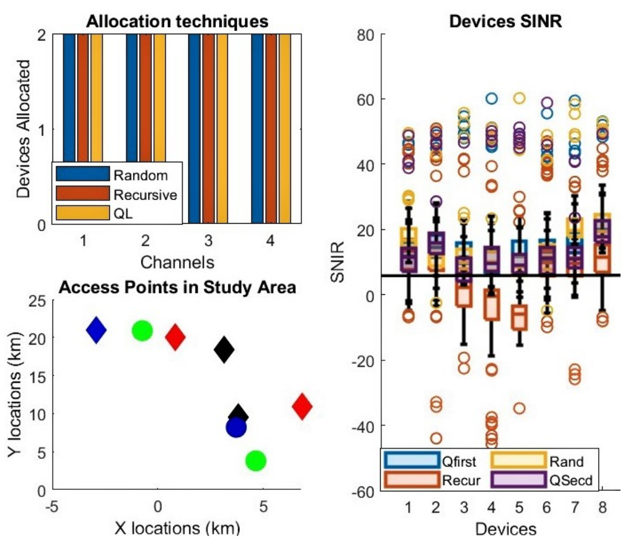


Fig. 8. Comparing allocation techniques $n = 8$ and $k = 4$

## References


[1] United Nations, "Achieving universal and meaningful digital connectivity Setting a baseline and targets for 2030," 2021. [Online].

[2] Dynamic Spectrum Alliance, "Automated Frequency Coordination," 2019.

[3] Y. Han, E. Ekici, H. Kremo, and O. Altintas, "Spectrum sharing methods for the coexistence of multiple RF systems: A survey," Ad Hoc Networks, vol. 53. Elsevier B.V., pp. 53–78, Dec. 15, 2016. doi: 10.1016/j.adhoc.2016.09.009.

[4] Marshall, P. (2017). Three-Tier Shared Spectrum, Shared Infrastructure, and a Path to 5G. Cambridge University Press. https://doi.org/10.1017/9781108165020

[5] O. Naparstek and K. Cohen, "Deep Multi-User Reinforcement Learning for Distributed Dynamic Spectrum Access," IEEE Trans Wirel Commun, vol. 18, no. 1, pp. 310–323, 2019, doi: 10.1109/TWC.2018.2879433.

[6] Mosleh, S., Ma, Y., Rezac, J. D., and Coder, J. B. (2020). Dynamic Spectrum Access with Reinforcement Learning for Unlicensed Access in 5G and beyond. IEEE Vehicular Technology Conference, 2020-May(Ml). https://doi.org/10.1109/VTC2020-Spring48590.2020.9129381

[7] Y. Wang, Z. Ye, P. Wan, and J. Zhao, "A survey of dynamic spectrum allocation based on reinforcement learning algorithms in cognitive radio networks," Artif Intell Rev, vol. 51, no. 3, pp. 493–506, 2019, doi: 10.1007/s10462-018-9639-x.

[8] N. C. Luong et al., "Applications of Deep Reinforcement Learning in Communications and Networking: A Survey," IEEE Communications Surveys and Tutorials, vol. 21, no. 4, pp. 3133–3174, 2019, doi: 10.1109/COMST.2019.2916583.

[9] H. Song, J. Bai, Y. Yi, J. Wu, and L. Liu, "Artificial Intelligence Enabled Internet of Things: Network Architecture and Spectrum Access," IEEE Comput Intell Mag, vol. 15, no. 1, pp. 44–51, 2020, doi: 10.1109/MCI.2019.2954643.

[10] S. Wang, H. Liu, P. H. Gomes, and B. Krishnamachari, "Deep Reinforcement Learning for Dynamic Multichannel Access in Wireless Networks," IEEE Trans Cogn Commun Netw, vol. 4, no. 2, pp. 257–265, Feb. 2018, doi: 10.1109/tccn.2018.2809722.

[11] O. Naparstek and K. Cohen, "Deep Multi-User Reinforcement Learning for Distributed Dynamic Spectrum Access," IEEE Trans Wirel Commun, vol. 18, no. 1, pp. 310–323, 2019, doi: 10.1109/TWC.2018.2879433.

[12] Y. S. Nasir and D. Guo, "Multi-Agent Deep Reinforcement Learning for Dynamic Power Allocation in Wireless Networks," IEEE Journal on Selected Areas in Communications, vol. 37, no. 10, pp. 2239–2250, 2019, doi: 10.1109/JSAC.2019.2933973.

[13] N. Zhao, Y. C. Liang, D. Niyato, Y. Pei, M. Wu, and Y. Jiang, "Deep Reinforcement Learning for User Association and Resource Allocation in Heterogeneous Cellular Networks," IEEE Trans Wirel Commun, vol. 18, no. 11, pp. 5141–5152, 2019, doi: 10.1109/TWC.2019.2933417.

[14] M. Camelo et al., "An AI-Based Incumbent Protection System for Collaborative Intelligent Radio Networks," IEEE Wirel Commun, vol. 27, no. 5, pp. 16–23, 2020, doi: 10.1109/MWC.001.2000032.

[15] C. Tarver et al., "Enabling a 'Use-or-Share' Framework for PAL-GAA Sharing in CBRS Networks via Reinforcement Learning," IEEE Trans Cogn Commun Netw, vol. 5, no. 3, pp. 716–729, 2019, doi: 10.1109/TCCN.2019.2929147.

[16] IEEE, IEEE Standard for Information technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements, no. June. 2018, pp. 39–44.

[17] CBRS WInnForum Standards, Operations for Citizens Broadband Radio Service (CBRS); GAA Spectrum Coordination-Approach 2 Technical Report, no. May. 2019.

[18] CBRS WInnForum Standards, Operations for Citizens Broadband Radio Service (CBRS); GAA Spectrum Coordination-Approach 3 Technical Report, no. May. 2019.

[19] L. Kulacz et al., "Coordinated Spectrum Allocation and Coexistence Management in CBRS-SAS Wireless Networks," IEEE Access, vol. 7, pp. 139294–139316, 2019, doi: 10.1109/ACCESS.2019.2940448.

[20] N. Naderializadeh, J. J. Sydir, M. Simsek, and H. Nikopour, "Resource Management in Wireless Networks via Multi-Agent Deep Reinforcement Learning," IEEE Trans Wirel Commun, vol. 20, no. 6, pp. 3507–3523, 2021, doi: 10.1109/TWC.2021.3051163.

[21] G. Zhao, Y. Li, C. Xu, Z. Han, Y. Xing, and S. Yu, "Joint Power Control and Channel Allocation for Interference Mitigation Based on Reinforcement Learning," IEEE Access, vol. 7, pp. 177254–177265, 2019, doi: 10.1109/ACCESS.2019.2937438.

[22] K. Nakashima, S. Kamiya, K. Ohtsu, K. Yamamoto, T. Nishio, and M. Morikura, "Deep Reinforcement Learning-Based Channel Allocation for Wireless LANs with Graph Convolutional Networks," IEEE Access, vol. 8, pp. 31823–31834, 2020, doi: 10.1109/ACCESS.2020.2973140.

[23] K. Khalil, G. Farhadi, and A. Ito, "Coexistence management for heterogeneous networks in white spaces," 2014 International Conference on Computing, Networking and Communications, ICNC 2014, pp. 691–697, 2014, doi: 10.1109/ICCNC.2014.6785420.

[24] A. and Y. J. Zhang Jingjing and Dong, "Intelligent Dynamic Spectrum Access for Uplink Underlay Cognitive Radio Networks Based on Q-Learning," in Wireless Algorithms, Systems, and Applications, F. and Y. J. Yu Dongxiao and Dressler, Ed., Cham: Springer International Publishing, 2020, pp. 691–703.

[25] E. Atimati, D. Crawford, R. Stewart, I. Achumba, L. Ezema, and U. Diala, "An Interference Management System for a Shared Spectrum Access Network," Proc 2022 IEEE Niger 4th Int Conf Disruptive Technol Sustain Dev NIGERCON 2022, 2022, doi: 10.1109/NIGERCON54645.2022.9803165.

[26] R. S. Sutton and A. G. Barto, Reinforcement Learning An Introduction, 2nd ed., vol. 258, no. 6685. The MIT Press, Cambridge, Massachusetts, 2018.