Enabling a "Use-or-Share" Framework for PAL–GAA Sharing in CBRS Networks via Reinforcement Learning

Chance Tarver, *Student Member, IEEE*, Matthew Tonnemacher, *Member, IEEE*, Vikram Chandrasekhar, Hao Chen, Boon Loong Ng, *Member, IEEE*, Jianzhong Zhang, *Fellow, IEEE*, Joseph R. Cavallaro, *Fellow, IEEE*, and Joseph Camp, *Member, IEEE*

***Abstract*—By implementing reinforcement learning-aided listen-before-talk (LBT) schemes over a citizens broadband radio service (CBRS) network, we increase the spatial reuse at secondary nodes while minimizing the interference footprint on higher-tier nodes. The federal communications commission encourages "use-or-share" policies in the CBRS band across the priority access license (PAL)–general authorized access (GAA) priority tiers by opportunistically allowing the lower-priority GAA nodes to access unused higher-priority PAL spectrum. However, there is currently no mechanism to enable this cross-tier spectrum sharing. In this paper, we propose and evaluate LBT schemes that allow opportunistic access to PAL spectrum. We find that by allowing LBT in a two carrier, two eNB scenario, we see upward of 50% user perceived throughput (UPT) gains for both eNBs. Furthermore, we examine the use of *Q*-learning to adapt the energy-detection threshold (EDT), combating problematic topologies, such as hidden and exposed nodes. With merely a 4% reduction in primary node UPT, we see up to 350% gains in average secondary node UPT when adapting the EDT of opportunistically transmitting nodes.**

***Index Terms*—CBRS, LBT, Q-learning, dynamic spectrum access.**

Manuscript received December 5, 2018; revised April 19, 2019; accepted July 6, 2019. Date of publication July 16, 2019; date of current version September 9, 2019. This work was supported by Samsung Research America, Inc. The work of C. Tarver and J. R. Cavallaro was supported in part by the U.S. NSF under grants ECCS-1408370, CNS-1717218, and CNS-1827940, for the "PAWR Platform POWDER-RENEW: A Platform for Open Wireless Data-driven Experimental Research with Massive MIMO Capabilities." The work of J. Camp was supported in part by the U.S. National Science Foundation under grants CNS-1526269 and CNS-1823304. A preliminary version of this manuscript appeared in the Proceedings of the 2018 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN) [1]. The associate editor coordinating the review of this paper and approving it for publication was H. Kim. *(Corresponding author: Chance Tarver.)*

C. Tarver and J. R. Cavallaro are with the Department of Electrical and Computer Engineering, Rice University, Houston, TX 77005 USA (e-mail: tarver@rice.edu).

M. Tonnemacher is with the Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75205 USA, and also with the Standards and Mobility Innovation Laboratory, Samsung Research America, Plano, TX 75093 USA.

V. Chandrasekhar, H. Chen, B. L. Ng, and J. Zhang are with the Standards and Mobility Innovation Laboratory, Samsung Research America, Plano, TX 75093 USA.

J. Camp is with the Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75205 USA.

Digital Object Identifier 10.1109/TCCN.2019.2929147

I. Introduction

ADDITIONAL spectrum availability and increased efficiency in the use of existing resources are needed to accommodate the rapidly increasing density and subsequent data demands of wireless devices around the world. The United States government recognized this need in 2010 and put into motion what would ultimately become the Citizens Broadband Radio Service (CBRS) [2].

In the standard CBRS architecture, there is a three-tiered system managed by a dynamic database called the Spectrum Access System (SAS), illustrated in Fig. 1. The top tier consists of incumbent users, the second tier consists of Priority Access Licenses (PALs), and the bottom tier is for General Authorized Access (GAA). Many users will likely operate in a pseudo-unlicensed fashion on the GAA tier, and in competitive markets, some carriers may choose to purchase a PAL license to ensure a minimum QoS. For more background on the CBRS band, see the Appendix.

At the same time that the federal government is opening up new bands for shared use, there is increasing congestion on the unlicensed bands. For example, 802.11 Wi-Fi and Bluetooth devices densely occupy the unlicensed 2.4 GHz and 5 GHz industrial, scientific and medical (ISM) radio bands. Cell providers are increasingly interested in using free, unlicensed spectrum to supplement their licensed networks, exacerbating this crowding. The idea of using unlicensed bands to supplement licensed networks has been pushed to multiple standards such as 3GPP's licensed assisted access (LAA) [3], LTE-U [4], and MulteFire [5].

It is clear that spectrum sharing will become more prevalent in future medium access policies. The cornerstone of these policies is the requirement for a way to manage the sharing. Spectrum sharing is traditionally done through contention-based protocols such as a listen-before-talk (LBT) scheme like Carrier-Sense Multiple Access (CSMA) for 802.11 devices. In these schemes, the channel must be determined to be unoccupied for a certain amount of time before it is accessed.

In the current release of the CBRS standard, there is no formal mechanism for GAA users to access PAL allocated channels. Instead, it is framed such that "use-or-share" [6] policies should be adopted to prevent spectrum warehousing by allowing PAL–GAA sharing but without any explanation

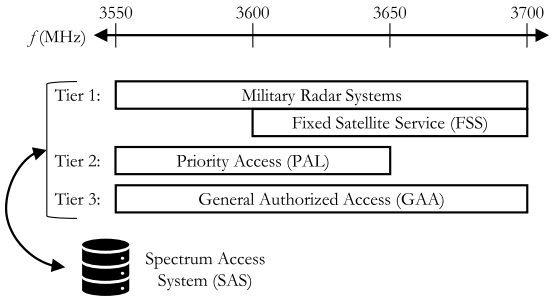


Fig. 1. Structure of the CBRS band. Three tiers of different priorities share the band, and a central database called the SAS dynamically manages the users. PAL users are licensed through auction while GAA users may use any spectrum not reserved for a higher tier [2].

for how it should be done. We propose a method to address this with the following contributions.

A. Contributions

In this paper, we present and compare two LTE-based LBT schemes for use in the CBRS infrastructure to allow the GAA users to access licensed PAL spectrum when available. Specifically, we analyze the trade-off between GAA user gain and PAL user interference through simulations where we aggregate results over many random user topologies and packet arrival rates. Moreover, we test our LBT schemes on a custom testbed with multiple software-defined radios and a real-time signal analyzer to show the feasibility on real hardware. In doing so, we make the following five contributions:

i We design and evaluate two LBT schemes to be used in CBRS networks for PAL–GAA spectrum sharing.
ii We show that while one scheme has higher performance, both schemes significantly improve GAA UPT with a minor decrease in PAL UPT.
iii We find that the decreased PAL UPT is a function of PAL traffic load and problematic network topologies between PAL–GAA users.
iv To reduce the negative consequences of spectrum sharing on the PAL, we formulate a novel Q-learning algorithm that adjusts GAA opportunistic access via learning an improved energy-detection threshold (EDT) for carrier sensing.
v By using average and differential PAL buffer occupancy as the environmental observations, we find the detriment to PAL UPT from spectrum sharing can be greatly reduced.

B. Related Work

Many works in the cognitive radio and dynamic spectrum access literature deal with sharing between a higher priority primary node (PN) and a lower priority secondary node (SN). Some recent investigations also consider this while adopting machine learning with promising results. In [7], the authors use Q-learning for dynamically choosing the channel for cells as opposed to static assignments. In [8], a decentralized Q-learning scheme is used for reducing the interference seen by 802.22, PN users. Although Q-learning has been considered for power allocations and channel assignments, it has not been used, to the best of the authors' knowledge, for adapting a dynamic EDT for SNs in a shared spectrum environment.

Overall, CBRS is still an emerging standard, but we can learn from similar experiences on other bands. When considering the coexistence of GAA users, there are many similarities to unlicensed bands which have been studied extensively. Notably, there is substantial work that has been done for the coexistence of LTE and Wi-Fi nodes in unlicensed bands for LAA [9]–[12] with [13] using reinforcement learning to alter the duty cycle of the LTE nodes and [14] using Q-learning to adjust the channel occupancy time of the LAA nodes.

Given that CBRS is a new band, the case of LTE nodes coexisting with another LTE node that has a higher priority is a new area. Much of the previously mentioned LTE–LAA works are not sufficient as the higher-tiered node does not necessarily engage in LBT. There have not yet been extensive studies for CBRS-specific performance improvements; so far, there are only initial proof of concept demonstrations reported. In [15], a field trial of CBRS devices (CBSDs) working with the SAS was shown where they suggest improvements to the SAS protocol based on their results, and recently Verizon has deployed a CBRS network in Florida [16].

The rest of the paper is organized as follows. We present our new listen-before-talk schemes for sharing the spectrum amongst LTE devices on the CBRS band in Section II. In Section III, we give an overview of the reinforcement learning scheme we apply. We additionally provide Q-learning simulation results in Section III and then conclude the paper in Section IV. Following the conclusion is an Appendix devoted to the state of the CBRS band.

II. Listen Before Talk

Although the CBRS band allows for spectrum sharing between multiple tiers and encourages such "use-or-share" frameworks, there is currently no mechanism to facilitate this sharing, a deficiency that we address here.

Throughout the analysis in the remainder of the paper, we adopt the conventional spectrum sharing terminology of primary and secondary nodes. In most scenarios, the PAL acts as a PN which will transmit whenever it has traffic for a connected user in what is called "On/Off" mode. The GAA acts as an SN which will need to contend for access to the PAL's channel via LBT. However, we also present a mutual sharing scenario in Section II-C where two PAL operators may use each other's spectrum as GAA users. In this case, an operator would be a PN on its licensed PAL spectrum and an SN on other operators' spectrum.

When it comes to opportunistic random access, LBT schemes are a proven method that can be used to allow GAA users to access PAL spectrum opportunistically. Wi-Fi, one of the most popular random access schemes available, has used LBT in the form of CSMA for sharing the spectrum between multiple users to great success. A version of LBT is essential in any shared-spectrum environment and is legally required in the European Union and Japan for operation on an unlicensed

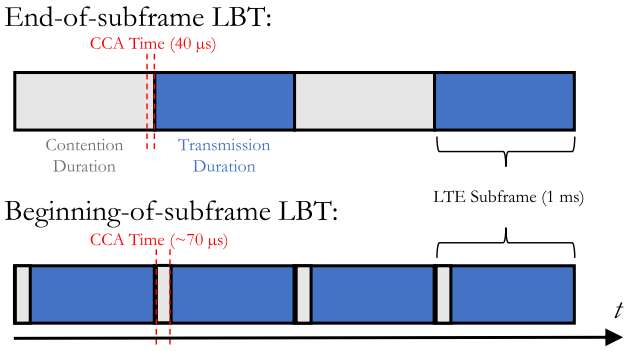


Fig. 2. Comparison of two proposed LBT schemes for CBRS. In End-of-subframe LBT, the SN uses the last 40 μs of a subframe for a CCA, leading to a 50% duty cycle but relatively easier implementation. In Beginning-of-subframe LBT, the 1st symbol of a subframe acts as the CCA leading to a higher, 93% duty cycle at the cost of a more complex implementation.

band. Ideally, under such a scheme, the secondary GAA nodes (SNs) gain additional spectrum resources, increasing throughput, while the primary PAL nodes (PNs) are unaffected. While it is impossible to achieve such perfect coordination due to the inability for secondary nodes to predict future PN packet arrival, it is possible for LBT to adequately allow LTE nodes to coexist with marginal harm to the PN.

LBT has already been adopted in some LTE standards, for example, 3GPP's LAA specification. The LBT scheme used in LAA is as follows. Whenever a device needs to transmit, it needs an initial clear-channel assessment (CCA). It must sense that the channel is idle for at least 34 μs. If so, it can transmit for the length of one transmit opportunity (TXOP). If there is additional traffic to send, an exponential random backoff mechanism is used, similar to Wi-Fi [11].

However, in the context of CBRS Alliance LTE devices, it is possible to further tailor the LBT scheme specifically for LTE devices. Although there could be many ways for performing LBT, we develop and compare the performance of two schemes that seem to be the natural choices: sensing at the end or sensing at the beginning of a subframe. These are shown in Fig. 2, and each is evaluated below. Similar schemes have also been considered for LTE/Wi-Fi coexistence in [12]. For the developed LBT schemes, we consider only the downlink, assuming that devices are operating similarly to LAA in Release 13 of LTE where the CBRS carrier is considered as a supplemental downlink carrier [11]. For more information on the typical structure of an LTE frame, see [17] and the references therein.

A. End-of-Subframe LBT

In end-of-subframe LBT, an entire subframe functions as a contention window for the SN with the last 40 μs in a subframe used for the CCA. If the channel is determined to be idle, the SN uses the next subframe for transmission. Using this scheme has an advantage of not altering the structure of a subframe. However, it results in, at most, a 50% transmission duty cycle. Moreover, the scheme could measure the channel to be idle during the contention window only for the primary node to start transmitting on the next subframe, leading to a collision. As this is a tiered access system, collisions with the PN need to be avoided at all costs.

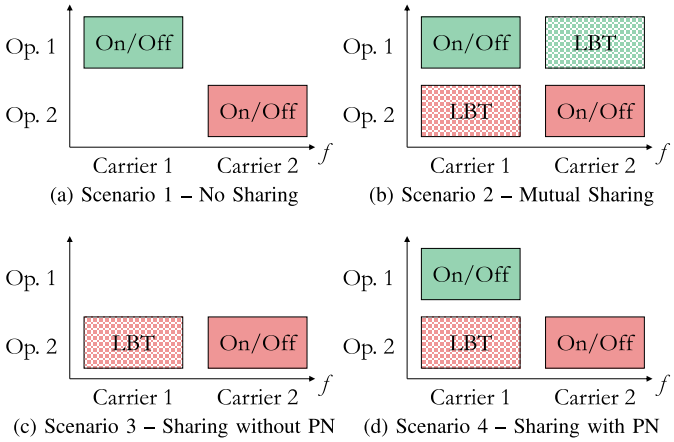
B. Beginning-of-Subframe LBT

In beginning-of-subframe LBT, the potential transmitter always senses in the first symbol of every subframe. If successful, the SN transmits in the remaining 13 symbols of the subframe. This LBT scheme has the advantage of sensing at the time that a PN would start a transmission, reducing the likelihood of a collision, given synchronization of subframe boundaries between PN and SN. Moreover, this scheme sacrifices a single symbol out of each subframe, a 93% duty cycle. However, LTE uses the first symbol for a control channel, so this scheme may require altering the subframe structure, though this omission of the control channel may be inconsequential in cases where SN's use cross-carrier scheduling.

C. LBT Scheme Comparison

There are apparent differences between the LBT schemes by construction. Given that the best-case duty cycle for end-of-subframe LBT and beginning-of-subframe LBT are 50% and 93%, respectively, beginning-of-subframe LBT is preferable. However, there is a tradeoff between performance and implementation complexity between the two schemes. With end-of-subframe LBT, we do not require any modification to the subframe structure, but beginning-of-subframe LBT would require proposing a change to the relevant standards. This change to the standards could be done through the CBRS Alliance, but since it may modify the subframe this change may also require 3GPP efforts, resulting in higher implementation complexity. Such a proposal would be a lengthy effort. Ultimately, after further study, there would need to be a mechanism to allow the control channel information to be sent on any symbol, or there would need to be a requirement of cross-carrier scheduling for LBT Scheme 2. However, standardization efforts are beyond the scope of the paper. Instead, we merely present our interpretation of the most likely scenario for the benefits that these LBT schemes could have for coexistence purposes.

To evaluate the performance of these LBT schemes in a CBRS-like framework, we first simulate their operation in various scenarios. In Fig. 3, we show each scenario considered to help understand the effect of the LBT scheme on both the PN and SN. For a baseline, we consider in Fig. 3a, Scenario 1, where two operators are operating on their carriers without any sharing. They operate in an "On/Off" mode without contention. In Fig. 3b, Scenario 2, we test the result when both operators engage in mutual sharing onto each other's carrier. In Fig. 3c, Scenario 3, we consider the case that a single operator is on its carrier and performs LBT on another carrier that is entirely unoccupied. This case represents an upper bound on gains for an SN. In Fig. 3d, Scenario 4, sharing is performed on a single component carrier to see the realistic gains for operator 2 (Op. 2) when sharing and the effect it has on the PN, operator 1 (Op. 1). Due to the nature of the band, PAL–GAA sharing will likely require an agreement between



(a) Scenario 1 – No Sharing (b) Scenario 2 – Mutual Sharing

(c) Scenario 3 – Sharing without PN (d) Scenario 4 – Sharing with PN

Fig. 3. Simulation scenarios considered throughout the paper. Simulations using these scenarios give us a performance baseline, a possible gain when two carriers mutually share their spectrum, an approximate best case SN gain when sharing, and an expected PN loss.

two operators on a secondary market. Considering this case, more complicated topologies with many operators and carriers are not likely to be realistic, and hence, the simulation scenarios in Fig. 3 give a realistic insight to potential sharing scenarios.

For the sake of comparing the performance of each LBT scheme, we considered the scenarios with no sharing (Scenario 1) and sharing with a PN (Scenario 4) from Figs. 3a and 3d as before and after cases. We report the change in the user-perceived throughput (UPT) for Op. 1 and Op. 2. In this figure, UPT is given by

$$UPT = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{P_{total}}\left[\sum_{j=1}^{P_{served}}\frac{M\cdot r_{ij}}{t_{ij}} + \frac{b_i}{t_{serving,i}}\right] \quad (1)$$

where N is the number of UEs served by the eNB, and i indexes the UEs. P_{total} is the total number of packets, elaborated by $P_{total} = P_{served} + P_{serving}$, where P_{served} and $P_{serving}$ are the number of packets served and being served, respectively. M is the number of bits per packet, r_{ij} is the ratio of successfully transmitted bits over all bits in the packet to UE i for packet j, and t_{ij} is the time taken to send the same packet. b_i is the number of bits sent to UE i as a partial packet still in flight, and $t_{serving,i}$ is the time spent by the packet.

We simulated our spectrum-sharing scheme using MATLAB by reusing the 3GPP LAA evaluation assumptions for an indoor scenario [18]. Fig. 4 shows this topology, and the rest of the simulation settings are as follows:

- Two operators with four small cells each in a single floor building (Fig. 4)
- 18 dBm TX power
- 10 randomly distributed UEs per operator
- −72 dBm EDT
- 20 MHz system bandwidth
- 10 independent simulations with random positioning of the UEs
- 20,000 subframes per simulation.

Fig. 5 shows the simulation results. For each test, we show the mean, 5th percentile, median, and 95th percentile to illustrate the variance across all simulations. In Fig. 5a, we see a

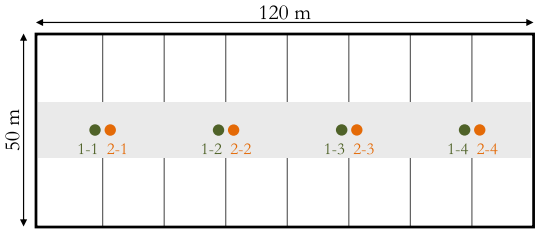
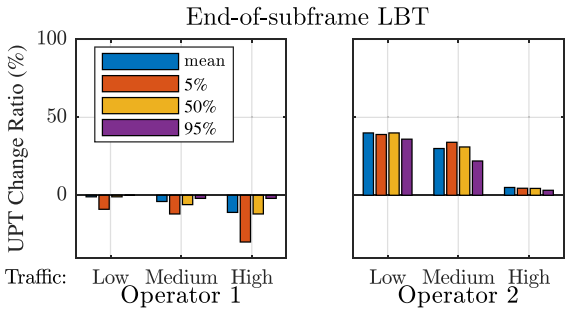
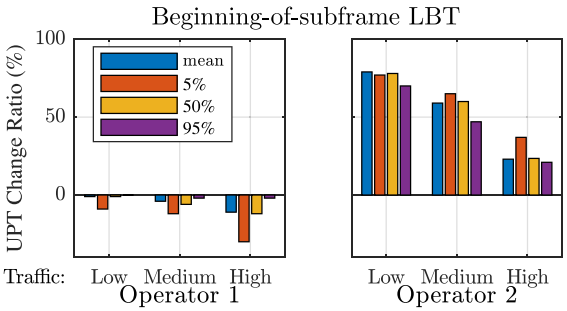


Fig. 4. 3GPP indoor scenario for LAA coexistence evaluations with two operators and four nodes per operator [18]. This standard scenario provides an industry agreed-upon simulation scenario for our LBT evaluations.



(a) UPT change when secondary node uses end-of-subframe LBT.



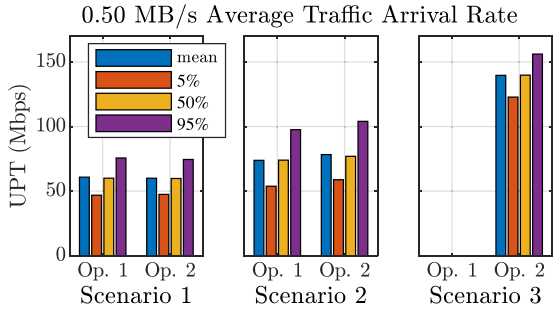
(b) UPT change when secondary node uses beginning-of-subframe LBT

Fig. 5. Performance of each LBT scheme for different traffic loads. In the results, a low-traffic load corresponds to a traffic arrival rate per UE of 0.5 MB/s, the medium is 0.75 MB/s, and the high is 1.05 MB/s.

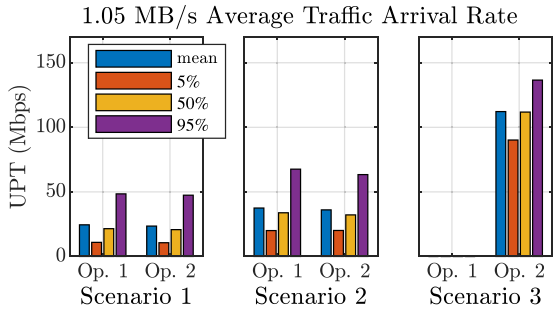
maximum increase in UPT for the SN of 40%. However, for the same case, there can be a 10% reduction in UPT for the PN. In Fig. 5b, we see a nearly 80% increase in performance for the SN with a similar drop in performance for the PN. Overall, beginning-of-subframe LBT performed significantly better for both the PN and SN. So, in the next subsection, we select this scheme for use in additional simulations to determine the possible spectrum-sharing gain. Although throughout the paper we show

D. Simulations With Static EDT

Fig. 6 shows the results for simulating beginning-of-subframe LBT across different spectrum sharing scenarios for two different traffic arrival rates to see the effect on UPT. In Fig. 6a, the average traffic arrival rate was 0.5 MB/s for PNs and SNs. The first cluster of results is the Scenario 1 baseline from Fig. 3a, where both operators are on their carriers without sharing the spectrum. The second cluster of results shows Scenario 2 from Fig. 3b, where each operator mutually shares

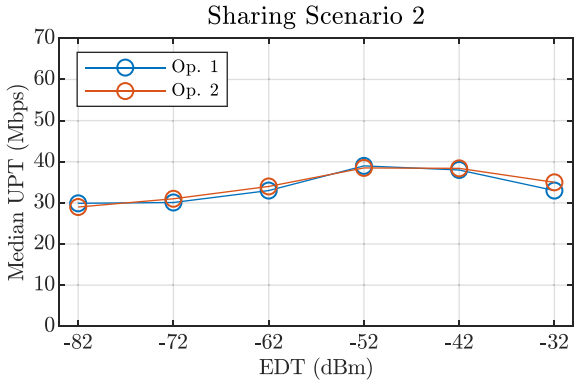


(a) UPT for spectrum-sharing scenarios when both operators use a relatively slow average traffic arrival rate of 0.50 MB/s.

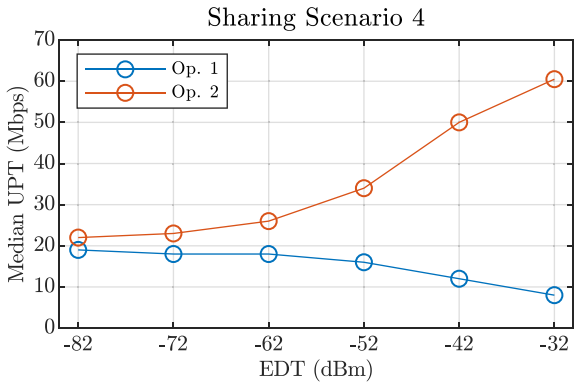


(b) UPT for spectrum-sharing scenarios when both operators use a relatively fast average traffic arrival rate of 1.05 MB/s, similar to each user streaming a 1080p video.

Fig. 6. Performance of LBT for two different average traffic arrival rates.



(a) Mutual sharing on two carriers, Scenario 2.



(b) Sharing on one carrier, Scenario 4.

Fig. 7. Median UPT vs. EDT for tests with a 20 MHz system bandwidth and a traffic arrival rate of 1.05 MB/s.

its primary spectrum with the other operator. Here, we can see that each operator experiences an increase in the mean UPT by about 25%.

The third cluster shows Scenario 3 from Fig. 3c, which provides a (coarse) upper bound on the maximum achievable spectrum sharing gain. Here, we can see that an operator can achieve a maximum of 133% gain when adopting a spectrum sharing scheme.

The simulation is repeated for the case of a higher-traffic arrival rate of 1.05 MB/s in Fig. 6b, which is similar to each user streaming a 1080p video using the H.264 codec. Similar to the results for the slower traffic rate, when each operator engages in a mutual sharing as in Scenario 2, each operator sees an improvement in UPT. For the higher traffic case, the gains are doubled with approximately 50% increase in UPT for both operators.

We then use beginning-of-subframe LBT and consider the performance for various EDTs. In Fig. 7a, we show the results when we consider different EDTs for Scenario 2 from Fig. 3b. Here, we see that there is an "optimal" EDT around -52 dBm. These results highlight that for different scenarios, there may be different "ideal" EDTs.

In Fig. 7b, we plot the UPT vs. EDT for Scenario 4 from Fig. 3d. In this figure, we can see the effect of a higher EDT. Here, a higher EDT at Op. 2 implies more frequency of channel access at the expense of increased downlink interference at Op. 1. As the EDT increases, the UPT of Op. 1 decreases, and the UPT of Op. 2 increases.

It is worth noting that the UPT decreases for the PN are significantly smaller in the low-traffic load case and more significant in the high-traffic load case for both schemes. Ideally, in situations where the PN has a high-traffic load, the SN would behave more passively when on the PN's carrier. In Section III, we will explore the use of machine learning in adjusting SN EDT to improve LBT performance in hidden and traffic-heavy node scenarios. Using our algorithm, we show that scenario-specific, poor-LBT performance can be significantly reduced.

E. Multiple Secondary Nodes

In many cases where PAL–GAA sharing may occur, it is likely that a sharing arrangement between the two will need to be formed on a secondary market [6]. In this case, there would be only two operators, and the operator on the GAA tier would perform its cell planning so that it would not experience significant self interference.

In the case where multiple GAA tier operators attempt to use the same PAL spectrum, the LBT scheme could be modified to accommodate this. One simple way would be to apply random back-off in the unit of subframes whenever an SN experiences a collision. Another possibility is to enforce a Maximum Channel Occupancy Time (MCOT) to limit the number of subframes an SN can transmit on consecutively.

F. Shared-Spectrum Testbed

To further evaluate the LBT schemes outlined so far, we developed a shared-spectrum testbed shown in Fig. 8. The testbed consists of the following.

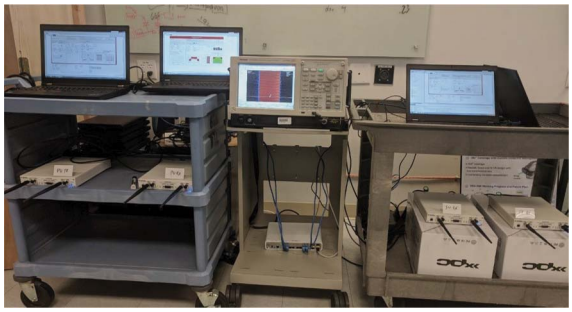


Fig. 8. Photograph of the shared-spectrum testbed. Four USRPs connected to host PCs running LabVIEW Communications with a real-time signal analyzer.

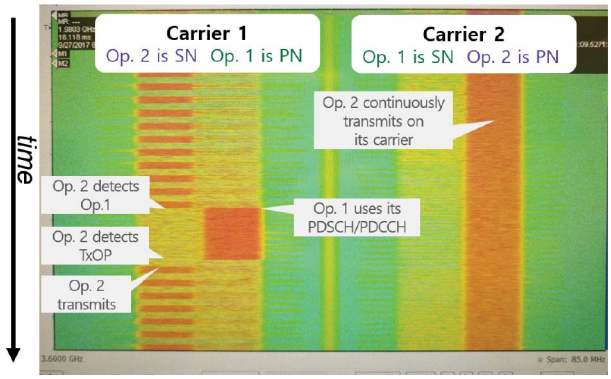


Fig. 9. Real-time signal analyzer spectrogram for end-of-subframe LBT, Scenario 2 from the testbed. Each operator is restricted to half of the full carrier to allow for visual distinction on the spectrum analyzer — Op. 2 transmits on the outer resource blocks, and Op. 1 transmits on the inner resource blocks. Here, Op. 2 has high traffic while Op. 1 has low traffic. On Carrier 2 where Op. 2 is the PN, Op. 2 continuously transmits as indicated by the orange color representing higher measured power. It can aggregate onto Carrier 1 while Op. 1 has no traffic. Whenever Op. 1 does transmit, Op. 2 can sense and avoid until the channel is free again.

- 4 USRP SDRs with the possibility of including more for more extensive tests with many nodes and UEs.
- Ideal ethernet backhaul via python/UDP for statistics.
- Real-time spectrum analyzer.

The nodes of the testbed can be arranged to emulate various topologies such as hidden/exposed nodes with additional possibilities of including mobility, allowing us to see how well the LBT schemes behave under real channel conditions where there may be a rich, multipath environment that changes quickly.

We modified the National Instruments LTE Application Framework to implement our LBT schemes. Fig. 9 shows an example result from the real-time signal analyzer spectrogram. Here, we use end-of-subframe LBT where the SN uses a subframe for contention and a subframe for transmission. We show Scenario 2, where each operator can engage in spectrum sharing. In this case, Op. 2 is under heavy load while Op. 1 is not. Op. 2 augments its services by aggregating onto Op. 1's spectrum while Op. 1 has no traffic. For this demo, we restrict each operator to occupy only half of each 20-MHz carrier so that we can easily distinguish between the operators on the spectrogram. In this figure, red represents high measured power for the corresponding time/frequency unit, and green represents low power. We can see that Op. 2 is using

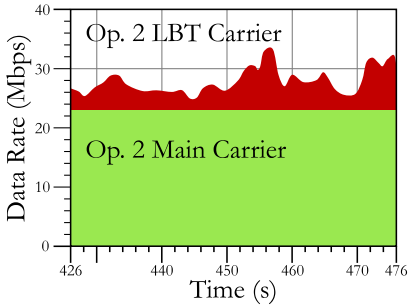


Fig. 10. Downlink throughput result from the testbed. Here, the green represents the throughput achieved on Op. 2's main carrier where it does not need to perform LBT. The red represents the throughput on its secondary carrier where it performs end-of-subframe LBT to avoid collisions with Op. 1. Periodically, Op. 2 is able to get an additional 50% throughput by utilizing the other carrier opportunistically.

Carrier 1 with a 50% duty cycle as it is the SN on this carrier and must perform end-of-subframe LBT. We can also see that when Op. 1 begins using its primary spectrum, Op. 2 detects the presence of the PN and waits for the carrier to become available again before transmitting.

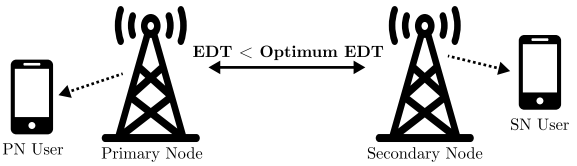
Fig. 10, which shows an example output from the GUI corresponding to the custom receiver design, shows the downlink performance for this example. Here, Op. 2 is under high traffic. It fully utilizes its main carrier to achieve approximately 22 Mbps UPT. To enhance performance, it performs end-of-subframe LBT on Op. 1's carrier. Throughout the experiment, Op. 2 is able to aggregate opportunistically onto Op. 1's primary spectrum. At times where Op. 1 has nearly no traffic, Op. 2 can get an additional 50% throughput.

In this testbed, the SN synchronizes with the timing of the PN of a channel. Using the existing LTE synchronization signals, the SN detects the subframe boundaries, measures the energy in the channel at the appropriate time, and then if the SN determines the PN not to be transmitting, the SN transmits to its users during the available TXOP. This result highlights the feasibility for the SN to sync to a PN for performing LBT in real-time. Moreover, Fig. 10 highlights what the throughput gain for the SN may realistically be as the PN transmits with a low traffic-arrival rate.

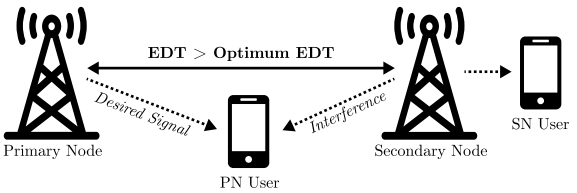
In this testbed, the SN synchronizes with the timing of the PN of a channel. The SN detects the primary synchronization signal (PSS) of the PN every 5 ms. It then uses this to estimate the subframe boundaries so that it can measure the energy in the channel at the appropriate time. If the SN determines the PN not to be transmitting, the SN transmits to its users during the available TXOP. This result highlights the feasibility for the SN to sync to a PN for performing LBT in real-time. Moreover, Fig. 10 highlights what the throughput gain for the SN may realistically be as the PN transmits with a low traffic-arrival rate.

III. Reinforcement Learning

While LBT schemes have been successfully implemented in 802.11 standards to great effect, certain network situations can result in poor performance. While other applications of



(a) Exposed node case with a Primary Node (PN) inside the sensing range of a Secondary Node (SN). Here, the energy detection threshold (EDT) is too sensitive, so whenever the PN transmits, the SN defers to avoid possible collisions with the PN's users. However, given the positions of the users, the SN could transmit without interference. This wasted opportunity could be remedied with a higher EDT.



(b) Hidden terminal case where a PN user is impacted by SN interference, despite PN being outside of SN sensing range. This collision could be avoided with a lower EDT.

Fig. 11. Example primary node (PN)–secondary node (SN) topologies where an adaptive EDT could benefit the network.

LBT may have ways of reducing network problems in topologies that include hidden and exposed terminals, shown in Fig. 11 [19], such as the collision avoidance in CSMA, similar schemes are not applicable when applying LBT to the CBRS tiered architecture as the primary node does not necessarily engage in LBT behavior. In this section, we mitigate this issue by proposing a novel reinforcement Q-learning technique to adapt an energy-detection threshold (EDT) for secondary nodes in a shared spectrum environment. We show that by using machine learning, we can increase both SN and PN gains over LBT schemes that use a static EDT.

A. Reinforcement Learning Primer

We first review the general reinforcement-learning strategy. Roughly speaking, reinforcement learning addresses the general problem of learning from interactions to achieve a goal [20]. The learner and decision maker is called the agent. Everything outside the agent that the agent interacts with is called the environment.

Agents interact with the environment via actions. For each action, α, that the agent executes, it influences the state of the environment and receives an evaluative feedback, or reward, *r*. This reward is used to learn/adapt its subsequent actions, should it encounter the same state in a subsequent time slot. We define periodic time intervals, $t = 0, T, 2T, \ldots$, in which each agent represents its observation, *o*, of the surrounding environment at time *t* as a state $s \in S$, where *S* designates a finite set of environmental states. In summary, at each step *t*:

The agent:

- Executes action α_t
- Receives observation α_t of s_t
- Receives reward r_t

The environment:

- Receives action α_t
- Emits observation o_{t+1} of s_{t+1}
- Emits scalar reward r_{t+1}

At each time step, the agent implements a mapping from states to probabilities of selecting each possible action. This mapping is called the agent's policy π_t, where $\pi_t(s, \alpha)$ is the probability that $\alpha_t = \alpha$ if $s_t = s$. Reinforcement learning methods specify how the agent changes its policy as a result of its experience. The agent's goal is to maximize the total amount of reward it receives over the long term.

One popular reinforcement-learning algorithm is Q-learning [20]. This model-free learning strategy can be used to learn an optimal decision policy for any Markov decision process. We adopt Q-learning with the objective to minimize interference at an incumbent or PN due to spectrum sharing with an opportunistic SN. In this scenario, each SN acts as an agent adapting its action in response to the reward obtained for its previous action.

B. Target Improvement Areas

Our objective is to use reinforcement learning to assist the SN in harvesting unused bandwidth from the PN in an efficient fashion. Specifically, we leverage Q-learning to dynamically adjust the SN's EDT to maximize network UPT while subsequently minimizing the impact on PN UPT. We identify two scenarios in which EDT adjustment can mitigate poor LBT performance.

1) Hidden/Exposed Terminals: In hidden or exposed terminal topologies illustrated in Fig. 11 an adaptive EDT can benefit network performance. For example, in the hidden node case, a UE served by the primary node potentially sees a significant interference if a secondary node transmits at the same time. If the queue size at the PN increases, a possible reason is because of interference from a (hidden) SN. In this scenario, the network would benefit if the SN had a more conservative EDT.

2) Load Adaptation: The SN node opportunistically adapts to fluctuations in offered traffic at PNs. If the PN traffic load is low, it may be able to use lower modulation and coding schemes (MCS) while maintaining a similar quality of service. By using more robust coding, higher interference can be tolerated without an increase in packet loss. Thus, the SN EDT can be reduced, allowing for more aggressive SN behavior, depending on the distance of the PN. Alternatively, if the PN traffic load is high, the SN EDT should be increased to prevent interference, even if the PN is further away, allowing higher PN MCS schemes to be used.

With these scenarios in mind, we now present our Q-learning algorithm followed by scenario-specific results.

C. Q-Learning Algorithm Description

In designing the reinforcement-learning algorithm, our objective is to determine a policy (sequence of state/action pairs) by which the agent (SN eNB) adapts its EDT based on observations taken during the latest epoch to maximize long-term rewards. In our setup, we assume that the PN shares transmit buffer occupancy/queue length information

with the overarching CBRS network architecture, making this information available to the SN. In turn, the SN uses this information as the environmental observation for the reinforcement learning.

Our assumption of the PAL sharing its buffer occupancy and queue lengths is not currently part of the CBRS Alliance standard. However, the topic of PAL–GAA coexistence is also not addressed in the current standard. PAL–GAA coexistence is expected to be addressed in a future version. The standard is actively evolving, and such information could be added to the measurement reports that currently exist between CBSDs, SASs, and the CBRS Alliance Coexistence Manager (CxM) [21]. In fact, the governing bodies are actively considering new measurements to include in these reports. We believe that buffer occupancy and queue length should be considered as they could be, as we demonstrate in this paper, useful tools to improve coexistence in future standards.

The buffer occupancy and queue length are helpful in facilitating coexistence because they, in part, provide a snapshot of the interference observed by the PN. The buffer occupancy is a percentage of the time in the previous epoch that the eNB had data in the buffer waiting to be sent. The queue length is the instantaneous amount of data in bytes to be sent at the end of each training epoch. The choice of buffer occupancy is a result of it being able to capture partial observations of three key factors in the environment: physical interference topology of the entire network, traffic load at the PN, and the virtual topology (sensing topology based on the EDT). Effectively, the buffer occupancy can be seen as a function of both the PN traffic arrivals and the inconsistencies between the interference topology and virtual topology (collisions). Since the goal of the EDT adjustment is to alter the transmission aggressiveness of the SN based on the PN traffic load and to make the virtual topology match the interference topology, this makes buffer occupancy the most appropriate metric to use.

The basic Q-learning implementation is as follows. Let epoch m, with duration T, refer to time interval $mT \leq t < (m+1)T$. The epoch duration, T, needs to be long enough (e.g., 10s - 100s of sub-frames) to avoid adapting to short-lived flows. At time $t = mT$, the agent chooses an action which maximizes its Q-table. At time $t = (m+1)T$, the agent receives the observation of the environment state from its last action, receives the associated reward, updates the Q-table, and then chooses an action α_{m+1} for epoch $m+1$. Fig. 12 depicts this iterative process.

Given state space S, the environment lies in one of two states $s \in \{1, 2\}$ defined in Table I. Here, L_m is the instantaneous PN transmit queue size at the end of epoch m, and γ_1 is a threshold used to differentiate high and low traffic loads. γ_1 selection can be used to adjust the relative weight between PN and SN users. We define these states such that there are binary light/heavy traffic conditions to reduce variables in our performance evaluation. However, this definition can easily be extended to multiple states if it is necessary to define more nuanced packet load conditions by defining multiple thresholds.

The agent is rewarded or punished according to the intuitive guidelines listed in Table II. Table III elaborates on this by showing specifically provide rewards in our Q-learning algorithm. In general, a positive reward is given if the PN's state improves or if a higher EDT threshold is chosen at the SN without a negative impact on the PN. Consequently, a negative reward is given in cases where the PN transitions to a worse state or if the SN chooses a low EDT value without any benefits. More specifically, we define numerical rewards according to the state transition and the average buffer occupancy over the previous epoch, B_m. Each γ in Table III is a tunable parameter outlined in Table IV. Each can help control the Q-learning to better tailor it for various goals and constraints. Z_m can be considered as a soft reward when outcomes are between actionable thresholds.

As mentioned earlier, the selection of γ_1 plays a critical role in how the learning algorithm behaves between the thresholded values, with higher values allowing the SN to be more responsive to changes in the PN queue length. These threshold values need to be tuned experimentally for different deployments, as there is no absolute rule for how they should be set. In general, each threshold contributes in one way or another

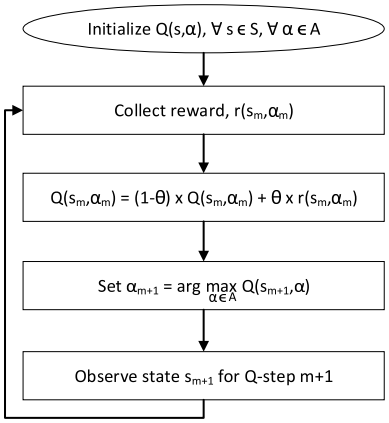


Fig. 12. Overview of Q-learning algorithm. In each epoch, an action is taken based on the state and learned "quality" of each state/action pair. Based on the change in environment from our action, the quality of the previous action is updated.

TABLE I
STATES FOR THE Q-LEARNING

State	Average Primary Node Queue Size	Comment
1	$0 \leq L_m < \gamma_1$	Primary node traffic load is light
2	$L_m \geq \gamma_1$	Primary node traffic load is heavy

TABLE II
REWARD INTUITION

Reward	Conditions
Positive	• The average primary node queue length is low (state 1), and the last action was to choose a high EDT value (e.g. -62 dBm). • The average primary node queue length goes from high to low (state 2 to 1) following the last epoch.
Negative	• The average primary node queue length is low (state 1), but the last action was to choose a low EDT value (e.g. -77 dBm). • The average primary node queue length goes from low to high (state 1 to 2) following the last epoch, and the buffer occupancy is large.

TABLE III
REWARDS FOR THE Q-LEARNING

(s_m, s_{m+1})	Reward r_m
(1, 1)	• γ_4, if $B_m \leq \gamma_2$ and $\alpha_m \geq \gamma_3$ • $-\gamma_4$, if $B_m \leq \gamma_2$ and $\alpha_m < \gamma_3$ • Z_m, otherwise
(1, 2)	• γ_4, if $B_m \leq \gamma_2$ and $\alpha_m \geq \gamma_3$ • $-\gamma_4$, if $B_m \leq \gamma_2$ and $\alpha_m < \gamma_3$ • $-\gamma_4$, if $B_m > \gamma_2$
(2, 1)	• 0, if $B_m \leq \gamma_2$ • γ_4, otherwise
(2, 2)	• 0, if $B_m \leq \gamma_2$ • Z_m, otherwise

TABLE IV
TUNABLE PARAMETERS FOR THE Q-LEARNING

Parameter	Description	
γ_1	Boundary between low-traffic and high-traffic states for the primary node.	Number of bytes.
γ_2	Buffer occupancy threshold for determining rewards.	$\in (0, 1)$
γ_3	EDT setting threshold.	dBm
γ_4	Standard reward value	Scalar
ϵ	Exploration parameter	$\in(0,1)$
θ	Discount factor	$\in(0,1)$
β	Learning Rate	$\in(0,1)$

to how quickly the algorithm adapts to changes. Depending on the specific situation, more or less rapid responses could be advantageous.

The agent updates the Q-table after each action according to (2). Here, $\theta \in (0, 1)$ is a discount factor that is used to control the importance of the current reward, r_m, in terms of updating the Q-function. Larger θ values will prioritize longer-term reward, while a lower θ applies more weight to the next iteration reward. The learning rate is controlled by $\beta \in (0, 1)$ which is a weighted average between the previous value of the Q-table for a state/action pair. This affects the tradeoff between convergence time and the relative stability of the entries.

$$Q(s_m, \alpha_m) \leftarrow (1-\beta) Q(s_m, \alpha_m) + \beta\Big(r_m + \theta \max_{\alpha} Q(s_{m+1}, \alpha)\Big) \quad (2)$$

The next action is chosen at each epoch according to the probability distribution in (3).

$$P(\alpha_{m+1}) = \begin{cases} 1-\epsilon, & \arg\max_{\alpha \in A} Q(s_{m+1}, \alpha) \\ \epsilon, & \text{rand}(\alpha \in A) \end{cases} \quad (3)$$

Here, $\epsilon \in [0,\ 1]$ is an exploration parameter, allowing for occasional random actions to be taken. In general, allowing for exploration prevents the learning algorithm from getting locked into suboptimal operation by filling in more of the Q-table than would occur otherwise. Furthermore, the exploration probability can be reduced over time as more iterations of the algorithm have occurred.

Although it is possible that PN queue length could increase due to factors besides SN interference such as increased MAC-layer contention, the Q-learning can be resilient to these factors as the Q-table is updated after an epoch consisting of 100s of subframes. Moreover, to further minimize the effects of non-interference caused changes to the PN queue length, the learning rate of the Q-learning algorithm, β, can be chosen to be small. The combination of large epochs and a low learning rate has the effect of "averaging out" most non-interference caused changes to the queue.

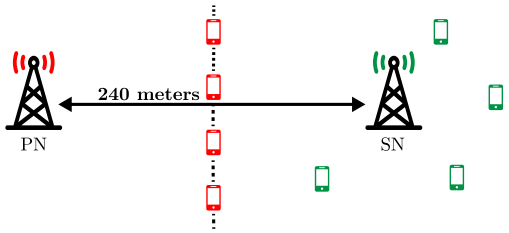


Fig. 13. Hidden node test topology where the PN UEs are equidistant from the PN and SN. Here, the PN users experience 0 dB SINR and are susceptible to collisions from the SN. The SN needs to adjust its EDT to be more sensitive to PN transmissions.

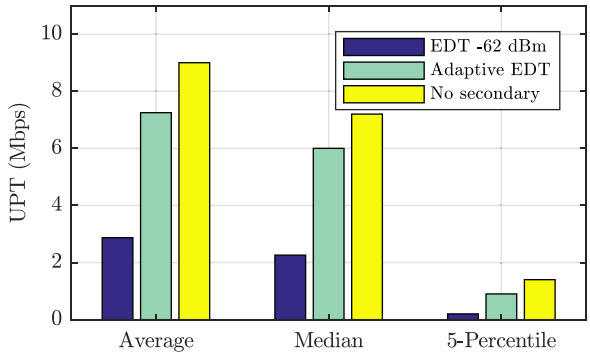


Fig. 14. PN performance when its UEs are hidden terminals to the SN. An adaptive EDT in the SN allows the SN to reduce its interference to the PN.

D. Simulations With Adaptive EDT

To evaluate the performance of the adaptive EDT, we perform system simulation in MATLAB. We examine several distinct scenarios to examine how the performance of LBT compares with and without the Q-learning based adaptive EDT.

1) Hidden Node – Mitigating Interference: For the first simulation, we consider the topology shown in Fig. 13. All of the PN UEs are placed equidistant from the PN and SN so that if the SN is transmitting, the SINR that they would receive would be approximately 0 dB. This simulation emulates a hidden-node case, where the distance between the PN and SN is far greater than the distance between the SN and the PN UEs. In Fig. 14, we compare PN UPT with a fixed, -62 dBm EDT to the adaptive EDT using Q-learning. In this figure, the upper bound on PN transmission is the situation in which there is no secondary user; thus, the PN can transmit interference free. We can see that when using a fixed EDT, the PN UPT drops drastically as expected in a hidden node scenario. However, by allowing the EDT to increase in response to the detection of increasing buffer occupancy at the PN, the penalty received by the PN is greatly reduced.

2) Exposed Node: For this simulation, we consider the topology shown in Fig. 15. All of the PN UEs are placed

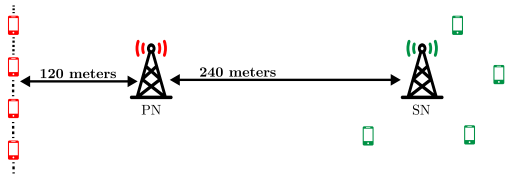


Fig. 15. Exposed node test topology where the PN UEs are at the cell edge opposite of the SN. Here, the PN users experience a low SINR but are not susceptible to collisions from the SN. The SN needs to adjust its EDT to be less sensitive to PN transmissions.

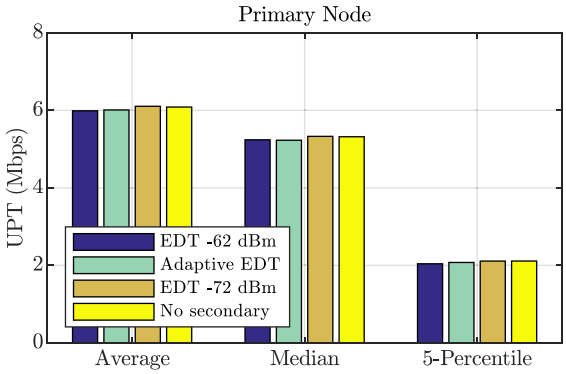


Fig. 16. PN performance when the SN is an exposed node. An adaptive EDT in the SN allows the SN to reduce its interference to the PN.

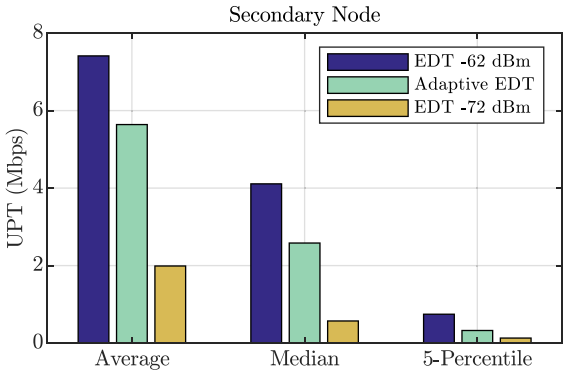


Fig. 17. SN performance for the exposed node topology. An adaptive EDT allows the node to learn that a higher EDT is acceptable with negligible effect on the PN.

at the cell edge opposite of the SN so that they are not susceptible to interference from the SN. This simulation emulates an exposed-node case, where the distance between the PN and SN is far less than the distance between the SN and the PN UEs. In Fig. 16, we compare PN UPT with two fixed EDTs to the adaptive EDT using Q-learning. In this figure, the upper bound on PN transmission is the situation in which there is no secondary user; thus, the PN can transmit interference free. We can see that no matter what the EDT of the SN is, the PN UPT remains nearly constant.

For the SN, we see in Fig. 17 that by allowing an adaptive EDT, the SN UPT is 2.8 times greater than the case in which the SN uses a static -72 dBm EDT. However, when the SN uses a static EDT of -62 dBm, the average UPT is 7.4 Mbps. Although the SN with an adaptive EDT achieves 76% of the performance of the static -62 dBm case, we see an impressive gain in the UPT over the -72 dBm EDT. It is important to

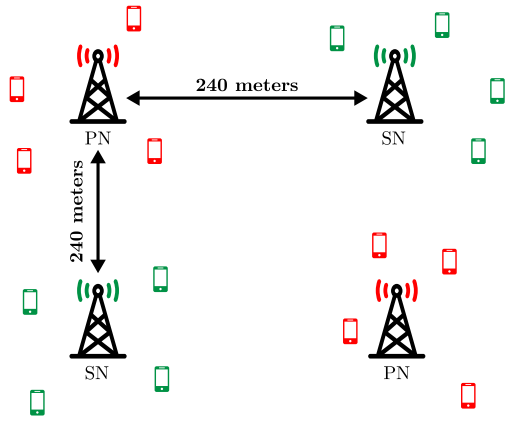
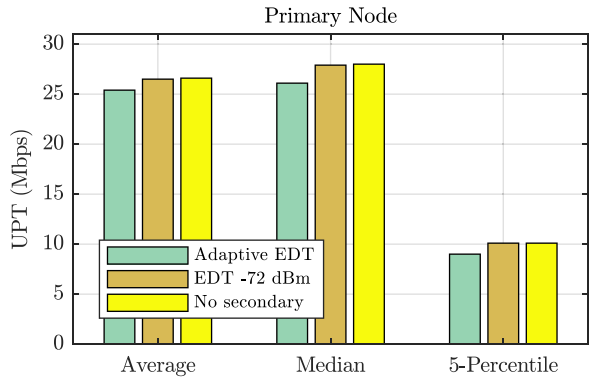
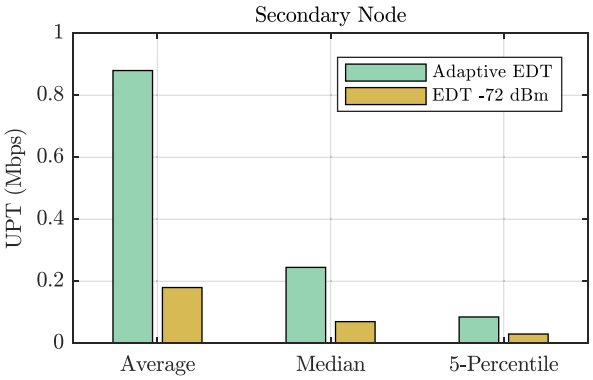


Fig. 18. Four-node test topology with two SNs and two PNs.



(a) PN UPT for the case where the SN adapts to PN load.

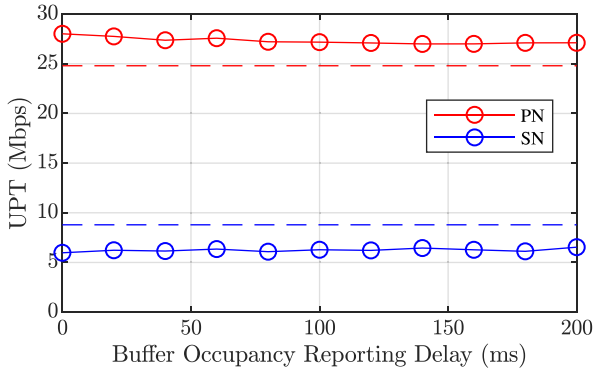


(b) SN UPT for the case where the SN adapts to PN load.

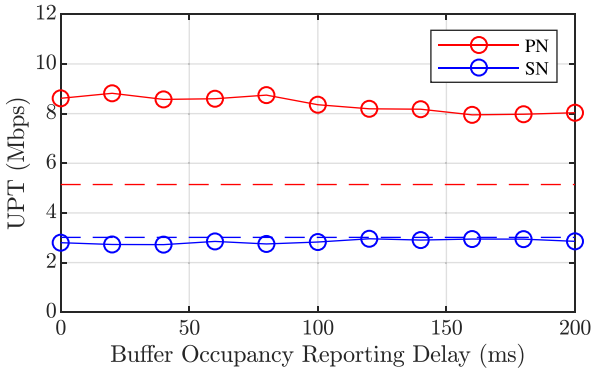
Fig. 19. Performance of PN and SN with adaptive versus fixed EDTs.

choose parameters such as γ_1 to balance the possible losses in hidden node cases with the possible gains in exposed node cases. In this particular simulation, the PN is in a high traffic state with an average queue length of 38,294 bytes. The γ_1 parameter for the Q-learning is chosen to be 75,000 bytes. By choosing a higher γ_1, it would be possible to achieve a greater UPT for the Adaptive EDT in this example at the possible expense of generality for the algorithm, tailoring it too much for one specific scenario.

3) Adapting to PN Load: In the next set of simulations, we have four nodes with two operators in a shared-carrier topology, as shown in Fig. 18. In this scenario, the PN load is effectively doubled, as each SN needs to defer to two

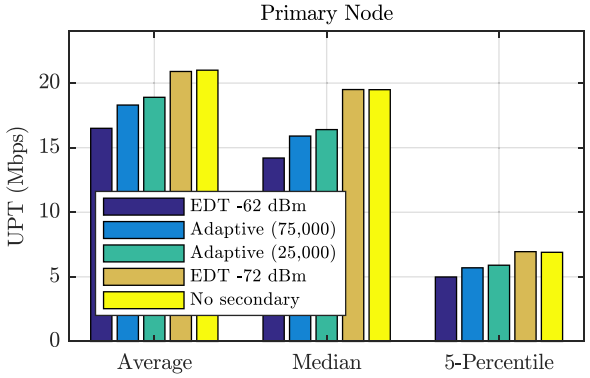


(a) Normal distribution of UEs simulation. The solid line with markers represents the results with the adaptive EDT using Q-learning, and the dashed line represents the baseline, static -62 dBm EDT without Q-learning.

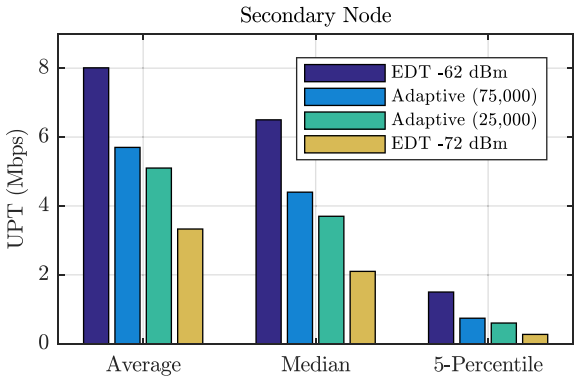


(b) Hidden node simulation. The solid line with markers represents the results with the adaptive EDT using Q-learning, and the dashed line represents the baseline, static -62 dBm EDT without Q-learning.

Fig. 20. Spectrum sharing with a simulated network latency in PN reports. As the delay in the reporting of PN statistics increases, there is only a minor drop in performance of the adaptive EDT algorithm showing the resilience of the Q-learning to latency in PN buffer occupancy and queue length reports. Even with significant network delays, the PN performance with an adaptive EDT is shown to be significantly better than the static -62 dBm EDT.



(a) PN UPT with differing γ_1 values



(b) SN UPT with differing γ_1 values.

Fig. 21. Spectrum sharing with different L_m threshold values.

different PNs. The UE distribution for each node is randomized in proximity around each node. We present the simulation results in Fig. 19. We can see that there is not a significant change in the PN UPT for any scheme, as they all approach the upper bound. For the case with no secondary node and the case with a -72 dBm fixed EDT, this is to be expected, as the SNs will be able to sense the PN and defer for nearly every PN transmission. However, when using an adaptive EDT, the SN UPT is increased by a factor of four. This is because the reinforcement learning can adaptively shrink when buffer occupancy remains low at the PN, by taking advantage of momentarily light traffic loads and/or transmissions to UEs located further away from the SN.

4) Delayed Feedback Scenario: One concern when using PN network statistics to drive the adaptive EDT algorithm is the effect of non-real-time buffer occupancy reports on Q-learning training. In any real system, sharing of data cannot be instantaneous, so we investigate the effect of system delay in PN buffer occupancy reporting in this simulation and measure the performance degradation. Fig. 20 shows changes in primary and secondary node UPT as system delay increases for both hidden nodes and normally-distributed UE topologies. We do not evaluate delays greater than 200 ms, as we consider this to be a rough upper bound of the intra-network latency experienced by a real deployment. However, even when reaching the upper end of this range, standard user mobility would not drive considerable changes to the network topology or traffic load in such a short time window. Therefore, these results are more a reflection of the increased convergence time to a static scenario rather than measuring adaptability to change. With that in mind, we initialized this simulation with a more aggressive EDT of -62 dBm and limited the experiment to 5 seconds to amplify the impact of the initial convergence time.

In the results, we can see that the normally-distributed topology converges to acceptable EDT values quickly even with delay, resulting in no perceived UPT loss for SN or PN. In the hidden node topology, however, the primary UPT experiences a slight inverse relationship with the system delay as a result of the extended convergence time. These results show that the Q-learning based adaptive EDT is resilient to cases where the PN buffer occupancy reports may not be exactly real time.

5) Effect of the State Transition Boundary, γ_1: In the next set of simulations, we have two nodes with two operators, a subset of the scenario in Fig. 18. We perform the simulation with two different γ_1 settings for the Q-learning, 25,000 bytes and 75,000 bytes. The choice in γ_1 is notable in that it governs the division between states 1 and 2 in the Q-learning algorithm.

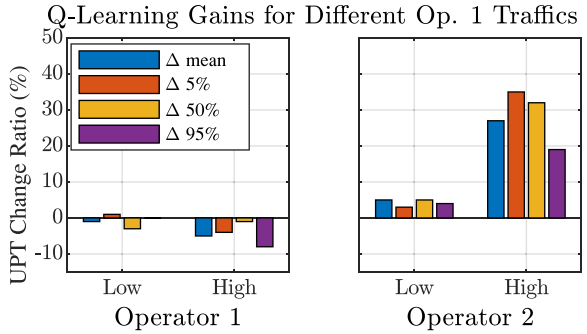


Fig. 22. Change in UPT for Op. 1 and Op. 2 when Op. 1 has a 0.125 MB/s (Low) and a 1.05 MB/s (High) average traffic arrival rates while Op. 2 always has a 1.05 MB/s traffic arrival rate. Here, each bar shows the change in the mean or corresponding percentile when Q-learning is used to alter the EDT of the SN.

A transition from state 1 to 2 causes a negative reward for the action that caused the state transition.

Fig. 21 shows these simulation results. In Fig. 21a, we show the UPT of the PN. In the case of a fixed EDT of -62 dBm, our highest considered EDT, we see the lowest performance for the PN as the SN will not as readily defer to the PN. Inversely, for a fixed EDT of -72 dBm, our lowest considered EDT, the SN will defer heavily to the PN, resulting in a performance similar to the case where there is no secondary node. When using an adaptive EDT, by changing the value of γ_1, we can balance the performance of the SN and PN. In Fig. 21b, we can see the complementary performance of the SN. This result shows that tuning of the γ_1 parameter can be used to balance the tradeoff between PN interference and SN channel access.

6) Multi-Node Scenario: In our final set of simulations, we use the LAA indoor scenario as outlined previously (shown in Fig. 4). Fig. 22 shows the benefit of Q-learning for the case where Op. 1 has at first a low, 0.125 MB/s average traffic arrival rate and then a high, 1.05 MB/s one while the SN traffic is kept high at 1.05 MB/s. For the low-traffic case, there is only about a 5% gain in UPT when using Q-learning because the SNs were already exploiting the many spectrum holes created by the limited traffic activity at PNs. However, when Op. 1 has a higher traffic load, the SN significantly benefits from Q-learning where the adaptive EDT leads to a median UPT improvement of over 30%.

IV. Conclusion

In this paper, we examine the challenges of using LBT for PAL–GAA spectrum sharing in CBRS networks by evaluating two different LBT schemes and showing that they can be used to improve significantly the SN UPT with a minor decrease in PN UPT. To reduce the negative consequences of spectrum sharing on the PN, we presented a novel, Q-learning algorithm that adjusts SN opportunistic access via learning an EDT for carrier sensing. We showed that by using average and differential PN buffer occupancy as the environmental observations, the SN can improve their throughput by up to 350% with only marginal losses to the PN UPT (4%). In future work, we can extend the intelligence globally from the local learning framework presented in this work, to jointly optimize within and across different shared-spectrum deployments, and examine how this work can scale to situations with multiple SNs.

Appendix
CBRS Background

The CBRS band is a promising new paradigm for managing spectrum in the United States. In this section, we present an overview of this unique architecture followed by a discussion of the use cases and main organizations setting the direction for the band and standard.

A. History

In 2010, the U.S. President called for an additional 500 MHz of wireless spectrum to be made available within ten years, recognizing the need for additional broadband spectrum [22]. It was later a finding of the President's Council of Advisors on Science and Technology (PCAST) that sharing the spectrum would be essential to meet the wireless challenges that are currently seen [23]. This recommendation was embraced by the Federal Communications Commission (FCC) in a notice of proposed rulemaking in which they suggest a three-tiered, database-managed, spectrum-sharing scheme for the $3550 - 3700$ MHz band, which would become the Citizens Broadband Radio Service (CBRS) [2]. Even before the presidential memorandum, these themes could be seen in other spectrum policies such as in TV White Space where unlicensed devices can utilize unused portions of ultra-high-frequency (UHF) bands that were licensed for TV stations [24]. The CBRS band is quickly developing and will see deployment in commercial devices soon with the launch of the Qualcomm X20 LTE modem, which supports operations on that band [25]. Under the new administration, many key terms to the licenses were altered leading to longer license lengths over a larger geographic area [6].

B. Architecture

The CBRS band is a three-tiered, shared-spectrum platform managed dynamically by the Spectrum Access System (SAS) as shown in Fig. 1. The incumbent users in the band include Department of Defense (DoD) radar systems as well as Fixed Satellite Service (FSS). The FCC establishes an Environmental Sensing Capability (ESC) to protect this tier. ESC nodes will monitor the band and notify the SAS in the case that an incumbent becomes present. The SAS will then move lower-tier users to unoccupied channels.

In the next tier, users may purchase spectrum in 10 MHz channels via an auction to become PALs. Unlike other traditionally licensed bands, purchasing a channel does not tie the licensee to a specific frequency. Instead, the SAS can dynamically assign the PAL to any channel in the band.

Initially, the PAL licenses were valid for only three years and for a census tract (a geographic area that is roughly sized according to a fixed population size of approximately 4,000 people) as opposed to longer terms and larger areas seen in other bands. However, this has recently changed in the Trump

administration. The PAL license now lasts for 10 years and is valid over an entire county, a larger geographic license area than the original census tract [6]. In any county, the FCC issues only 7 PAL licenses, guaranteeing that a minimum of 80 MHz of the band is always available for GAA use.

The bottom tier is designed to be similar to an unlicensed band to enable a low-cost, flexible solution for a large group of potential applications. The SAS may assign GAA users to any channel in the full 150 MHz band as long as they do not interfere with a user of higher priority.

All devices are managed by the SAS, which is the true innovation of this band. It is motivated by the TV White Space database system, but it is designed to be more dynamic. The primary functions of the SAS include determining available frequencies and assigning them to devices, determining maximum power levels for devices, enforcing exclusion zones, protecting PALs from GAA users, and facilitating coordination between GAA users. While dynamic, it is an explicit goal of the SAS not to micromanage the spectrum, leaving such fine-grained tuning to the operators.

C. Advantages and Use Cases

One of the main advantages of the CBRS band is the low barrier to entry, leading to many possible use cases. The GAA tier can be accessed in an almost unlicensed-like fashion at no cost by registering with the SAS which creates many opportunities for neutral-host or private LTE networks. For example, an arena could deploy a neutral-host infrastructure on 3.5 GHz using the GAA tier. Enterprise campuses could deploy private LTE networks that cover their entire campus. In addition, existing cell providers could augment their networks by utilizing this band as a GAA user. In select, competitive markets, users may choose to purchase a priority license to reduce interference and guarantee a minimum quality of service. The FCC also allows the CBRS band to be used for a fixed-wireless, infrastructure-type deployment where higher transmit powers are permitted.

D. Standards and Regulatory Bodies

The FCC developed the three-tier architecture, but other organizations have been formed to create specific standards for the band. Two of these major organizations are the Wireless Innovation Forum (WInnForum) and CBRS Alliance. The WInnForum is developing standards for certifying devices and developing a SAS protocol and is not tied to any specific wireless technology. The CBRS Alliance is developing standards for operating LTE on the CBRS band, which they recently have named OnGo [26].

The CBRS Alliance imposes an additional structure to be able to facilitate coordination amongst GAA users. In [21], a coexistence group (CxG) is created which is managed by a coexistence manager (CxM). Multiple GAAs can be a part of a CxG, and the SAS will allocate a pool of spectrum to the CxG instead of to individual CBSDs. The CxM can then enable more fine-grained sharing among the members of its group.

REFERENCES

[1] M. Tonnemacher *et al.*, "Opportunistic channel access using reinforcement learning in tiered CBRS networks," in *Proc. DySPAN*, Seoul, South Korea, Sep. 2018, pp. 1–10.

[2] *Amendment of the Commission's Rules With Regard to Commercial Operations in the 3550–3650 MHz Band*, document FCC 15–47, Federal Commun. Commission, Washington, DC, USA, Apr. 2015. [Online]. Available: https://apps.fcc.gov/edocs_public/attachmatch/FCC-15-47A1.pdf

[3] D. Flore, *LAA Standardization: Coexistence Is the Key*, 3GPP, Sophia Antipolis, France, Jul. 2016. [Online]. Available: http://www.3gpp.org/news-events/3gpp-news/1789-laa_update

[4] M. Chmaytelli, *LTE-U Forum: Ensuring LTE and Wi-Fi Fairly Coexist in Unlicensed Spectrum*, Qualcomm, San Diego, CA, USA, Mar. 2016. [Online]. Available: https://www.qualcomm.com/news/onq/2015/03/03/lte-u-forum-ensuring-lte-and-wi-fi-fairly-coexist-unlicensed-spectrum

[5] *MulteFire Release 1.0 Technical Paper*, MulteFire Alliance, Fremont, CA, USA, Jul. 2017. [Online]. Available: https://www.multefire.org/wp-content/uploads/MulteFire-Release-1.0-whitepaper_FINAL.pdf

[6] *Promoting Investment in the 3550–3700 MHz Band*, document FCC 18–149, Federal Commun. Commission, Washington, DC, USA, Oct. 2018. [Online]. Available: https://docs.fcc.gov/public/attachments/FCC-18-149A1.pdf

[7] J. Nie and S. Haykin, "A Q-learning-based dynamic channel assignment technique for mobile communication systems," *IEEE Trans. Veh. Technol.*, vol. 48, no. 5, pp. 1676–1687, Sep. 1999.

[8] A. Galindo-Serrano and L. Giupponi, "Distributed Q-learning for aggregated interference control in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1823–1834, May 2010.

[9] C. Chen, R. Ratasuk, and A. Ghosh, "Downlink performance analysis of LTE and WiFi coexistence in unlicensed bands with a simple listen-before-talk scheme," in *Proc. IEEE Veh. Technol. Conf.*, Glasgow, U.K., May 2015, pp. 1–5.

[10] N. Rupasinghe and I. Güvenç, "Licensed-assisted access for WiFi-LTE coexistence in the unlicensed spectrum," in *Proc. IEEE GC Wkshps*, Austin, TX, USA, Dec. 2014, pp. 894–899.

[11] H.-J. Kwon *et al.*, "Licensed-assisted access to unlicensed spectrum in LTE release 13," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 201–207, Feb. 2017.

[12] B. Jia and M. Tao, "A channel sensing based design for LTE in unlicensed bands," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, London, U.K., Jun. 2015, pp. 2332–2337.

[13] N. Rupasinghe and I. Güvenç, "Reinforcement learning for licensed-assisted access of LTE in the unlicensed spectrum," in *Proc. IEEE Wireless Commun. Netw. Conf.*, New Orleans, LA, USA, Mar. 2015, pp. 1279–1284.

[14] A. Galanopoulos, F. Foukalas, and T. A. Tsiftsis, "Efficient coexistence of LTE with WiFi in the licensed and unlicensed spectrum aggregation," *IEEE Trans. Cogn. Commun. Netw.*, vol. 2, no. 2, pp. 129–140, Jun. 2016.

[15] M. Palola *et al.*, "Field trial of the 3.5 GHz citizens broadband radio service governed by a spectrum access system (SAS)," in *Proc. DySPAN*, Piscataway, NJ, USA, Mar. 2017, pp. 1–9.

[16] (May 2018). *You Don't Need High Grade Navy Radar Systems to Spot Which Companies Just Achieved Another Industry Milestone for Customers*. [Online]. Available: http://www.verizon.com/about/news/you-dont-need-high-grade-navy-radar-systems-spot-which-companies-just-achieved-another

[17] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas, "LTE-advanced: Next-generation wireless broadband technology [invited paper]," *IEEE Wireless Commun.*, vol. 17, no. 3, pp. 10–22, Jun. 2010.

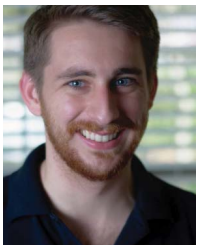
[18] "Study on licensed-assisted access to unlicensed spectrum, version 13.0.0," 3GPP, Sophia Antipolis, France, Rep. TR 36.889, Jun. 2015. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2579

[19] L. Wang, K. Wu, and M. Hamdi, "Combating hidden and exposed terminal problems in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 4204–4213, Nov. 2012.

[20] C. J. C. H. Watkins, "Learning from delayed rewards," Ph.D. dissertation, King's College London, London, U.K., 1989. [Online]. Available: https://www.cs.rhul.ac.uk/home/chrisw/new_thesis.pdf

[21] *CBRS Coexistence Technical Specification*, CBRS Alliance, Beaverton, OR, USA, 2018. [Online]. Available: https://www.cbrsalliance.org/wp-content/uploads/2018/04/CBRS-Coexistence-Technical-Specification.pdf

[22] B. Obama. (Jun. 2010). *Presidential Memorandum: Unleashing the Wireless Broadband Revolution*. [Online]. Available: https://obamawhitehouse.archives.gov/the-press-office/presidential-memorandum-unleashing-wireless-broadband-revolution
[23] *Realizing the Full Potential of Government-Held Spectrum to Spur Economic Growth*, Executive Office President, President's Council Advisors Sci. Technol., Washington, DC, USA, Jul. 2012. [Online]. Available: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast_spectrum_report_final_july_20_2012.pdf
[24] *Unlicensed Operation in the TV Broadcast Bands; Additional Spectrum for Unlicensed Devices Below 900 MHz and in the 3 GHz Band*, document FCC 08–260, Federal Commun. Commission, Washington, DC, USA, Nov. 2008. [Online]. Available: https://apps.fcc.gov/edocs_public/attachmatch/FCC-08–260A1.pdf
[25] (Mar. 2018). *Snapdragon X20 Modem With Category 18 Gigabit Class LTE*. [Online]. Available: https://www.qualcomm.com/products/snapdragon/modems/x20
[26] K. Mun. (May 2018). *OnGo: New Shared Spectrum Enables Flexible Indoor and Outdoor Mobile Solutions and New Business Models*. [Online]. Available: https://www.cbrsalliance.org/wp-content/uploads/2018/04/Mobile-Experts-OnGo.pdf



Chance Tarver (S'13) received the B.S. degree in electrical engineering from Louisiana Tech University, Ruston, LA, USA, in 2014, and the M.S. degree in electrical and computer engineering from Rice University, Houston, TX, USA, in 2016, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering. His current research interests include spectrum sharing for unlicensed communications and physical layer impairment corrections.

Matthew Tonnemacher received the B.S. and M.S. degrees in electrical engineering from Southern Methodist University, Dallas, TX, USA, in 2011 and 2013, respectively, where he is currently pursuing the Ph.D. degree. He has been a Wireless Systems Engineer with Samsung Research America, Plano, TX, USA, since 2016. His current research interests include spectrum sharing, software defined radio, and next generation wireless technologies.

Vikram Chandrasekhar received the Ph.D. degree from the University of Texas at Austin, TX, USA. He is currently with Samsung Research America, Mountain View, CA, USA, where he leads a team developing Samsung's Network Analytics and automation tools for 5G cellular systems. His past works include leading PHY-MAC systems work for Cisco's small cell product and developing a cloud-hosted radio resource management platform for managing Wi-Fi networks. He has also participated in the 3GPP standardization of the physical layer (3GPP-RAN1) for LTE (Releases 8–13) and New Radio (Release 15) technologies. His research interests include wireless network optimization, information theory, and cellular system design.

Hao Chen received the B.S. and M.S. degrees in information engineering from Xi'an Jiaotong University, Xi'an, in 2010 and 2013, respectively, and the Ph.D. degree in electrical engineering from the University of Kansas, Lawrence, KS, USA, in 2017. Since 2016, he has been a Research Engineer with the Standards and Mobility Innovation Laboratory, Samsung Research America. His research interests include dynamic spectrum access, network optimization, machine learning, and 5G cellular systems.

Boon Loong Ng (M'02) received the Bachelor of Engineering (Electrical and Electronic) degree and the Ph.D. degree in engineering from the University of Melbourne, Australia, in 2001 and 2007, respectively. He currently holds the position of Research Director with Standards and Mobility Innovation Laboratory, Samsung Research America, Plano, TX, USA. He had contributed to 3GPP RAN L1/L2 standardizations of LTE, LTE-A, LTE-A Pro, and 5G NR technologies from 2008 to 2018. He holds over 60 USPTO-granted patents on LTE/LTE-A/LTE-A Pro/5G and over 100 patent applications globally. Since 2018, he has been leading a research and development team that develop system and algorithm design solutions for commercial 5G and Wi-Fi technologies.

Jianzhong (Charlie) Zhang (F'16) received the Ph.D. degree from the University of Wisconsin, Madison. He is a VP and the Head of the Standards and Mobility Innovation Laboratory, Samsung Research America, where he leads research, prototyping, and standards for 5G and future multimedia networks. From 2009 to 2013, he served as the Vice Chairman for the 3GPP RAN1 Working Group and led development of LTE and LTE-Advanced technologies, such as 3-D channel modeling, UL-MIMO, CoMP, and carrier aggregation for TD-LTE.

Joseph R. Cavallaro (S'78–M'82–SM'05–F'15) received the B.S. degree in electrical engineering from the University of Pennsylvania, Philadelphia, PA, USA, in 1981, the M.S. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 1982, and the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY, USA, in 1988. From 1981 to 1983, he was with AT&T Bell Laboratories, Holmdel, NJ, USA. In 1988, he joined the faculty of Rice University, Houston, TX, USA, where he is currently a Professor of electrical and computer engineering and the Associate Chair. His research interests include computer arithmetic, and DSP, GPU, FPGA, and VLSI architectures for applications in wireless communications. From 1996 to 1997, he served at the U.S. National Science Foundation as the Director of the Prototyping Tools and Methodology Program. He was a Nokia Foundation Fellow and a Visiting Professor with the University of Oulu, Finland, in 2005, where he is an Adjunct Professor. He is currently the Director of the Center for Multimedia Communication, Rice University. He is an Advisory Board Member of the IEEE SPS TC on Design and Implementation of Signal Processing Systems and the Chair of the IEEE CAS TC on Circuits and Systems for Communications. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, and currently serves as an Associate Editor for the *Journal of Signal Processing Systems*. He was the General/Program Co-Chair of the 2003, 2004, and 2011 IEEE International Conference on Application-Specific Systems, Architectures and Processors, and the General/Program Co-Chair for the 2012, 2014 ACM/IEEE GLSVLSI conferences. At the IEEE SiPS workshop, he was the TPC Co-Chair in 2016 and the General Co-Chair in 2020. At the IEEE Asilomar Conference on Signals, Systems, and Computers, he was the TPC Chair in 2017 and the General Chair in 2020. He served on the IEEE CAS Society Board of Governors in 2014.

Joseph Camp (S'03–M'09) received the B.S. degree (Hons.) in electrical and computer engineering (ECE) from University of Texas at Austin, and the M.S. and Ph.D. degrees in ECE from Rice University. He joined the faculty of Southern Methodist University, Dallas, TX, USA, in 2009, where he is an Associate Professor of electrical and computer engineering. His research interests are wireless communications and networking, crowdsourcing, and drones, specifically focused on the deployment, measurement, and analysis of large-scale systems and development of embedded protocols. His research team has performed over 200 million in-field wireless measurements around the world via Android deployment and local characterization via drones, campus buses, vehicles, and buildings. He was the Chief Network Architect for the Technology for All Network, a mesh deployment in Houston, TX, USA, which served thousands of users in an under-resourced community. He was a recipient of the Ralph Budd Award for the best engineering thesis at Rice University in 2010, the National Science Foundation CAREER Award in 2012, and the Golden Mustang Teaching Award in 2014.