Regret Based Learning for UAV assisted LTE-U/WiFi Public Safety Networks

Dasun Athukoralage[1], Ismail Guvenc[2,1], Walid Saad[3], and Mehdi Bennis[4]

[1]Department of Electrical and Computer Engineering, Florida International University, Miami, FL, USA
[2]Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA
[3]Wireless@VT, Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA
[4]CWC - Centre for Wireless Communications, University of Oulu, Finland
Emails: dwick008@fiu.edu, iguvenc@ncsu.edu, walids@vt.edu, bennis@ee.oulu.fi

Abstract—Broadband wireless communication is of critical importance during public safety scenarios as it facilitates situational awareness capabilities for first responders and victims. In this paper, the use of LTE-Unlicensed (LTE-U) technology for unmanned aerial base stations (UABSs) is investigated as an effective approach to enhance the achievable broadband throughput during emergency situations by utilizing the unlicensed spectrum. In particular, we develop a game theoretic framework for load balancing between LTE-U UABSs and WiFi access points (APs), based on the users' link qualities as well as the loads at the UABSs and the ground APs. To solve this game, we propose a regret-based learning (RBL) dynamic duty cycle selection (DDCS) method for configuring the transmission gaps in LTE-U UABSs, to ensure a satisfactory throughput for all users. Simulation results show that the proposed RBL-DDCS yields an improvement of 32% over fixed duty cycle LTE-U transmission, and an improvement of 10% over Q-learning based DDCS.

Index terms— 5G, drone, LTE-U, public safety communications, regret based learning, unmanned aerial base station.

I. Introduction

The ubiquitous availability of broadband wireless connectivity carries critical importance during public safety scenarios. It can enable real time situational awareness for first responders, and help in protecting lives, property, and critical infrastructure. On the other hand, existing broadband communication networks can get damaged in the aftermath of a disaster scenario, as seen during the 2011 tsunami in Japan [1], which can eliminate the most important form of communications between victims, first responders, and other personnel during critical emergency situations. Due to their mobility and self-organization features, aerial platforms such as unmanned aerial base stations (UABSs) [2], [3] are uniquely suited for delivering broadband connectivity during public safety scenarios.

A critical bottleneck during public safety incidents is the limited availability of the wireless spectrum to satisfy the data rate requirements needed for real-time situational awareness applications. In addition to the licensed spectrum, the use of unlicensed frequency bands can increase the achievable

throughput for public safety networks (PSNs). In particular, LTE-Unlicensed (LTE-U) is a promising technology that can alleviate wireless network congestion by smartly off-loading the wireless traffic from the licensed spectrum, into the unlicensed spectrum [4]. However, the use of unlicensed bands by the LTE technology may severely affect the performance of WiFi networks due to the difference of medium access control (MAC) mechanisms of both radio access technologies (RATs). LTE radio resource allocation is controlled by LTE base stations (BSs) in a centralized manner. In contrast, WiFi uses carrier sense multiple access/collision avoidance (CSMA/CA) to sense whether the medium is occupied before sending any data.

The use of airborne platforms to maintain wireless connectivity during disaster scenarios has recently attracted attention in the existing literature. In [5], the authors consider the performance of 4G LTE-WiFi multimode base stations deployed on airborne platforms which provide coverage for first responders during an emergency. The feasibility of using an unmanned aerial vehicle (UAV) based WiFi network during disaster response scenarios has been studied in [6], while a UAV network with underlaid device-to-device communications has been explored in [7]. Coexistence studies between LTE-U and WiFi has also been investigated in the literature [8]–[10]. For example, a Q-Learning based dynamic duty cycle selection (DDCS) technique is proposed in [8] for configuring LTE transmission gaps, and allow transmission opportunities for WiFi. It is recently shown in [9], [10] that without any carefully designed coexistence schemes, WiFi performance can drop significantly while the LTE-U system performance is only slightly affected.

The main contribution of this paper is to propose an LTE-U based UAV-assisted heterogeneous network to sustain ubiquitous broadband connectivity in the aftermath of a natural disaster, as shown in Fig. 1. Such a scenario, to our best knowledge, has not yet been considered in the literature. In particular, LTE-U UABSs are deployed to fill coverage gaps due to the damaged infrastructure. The studied system
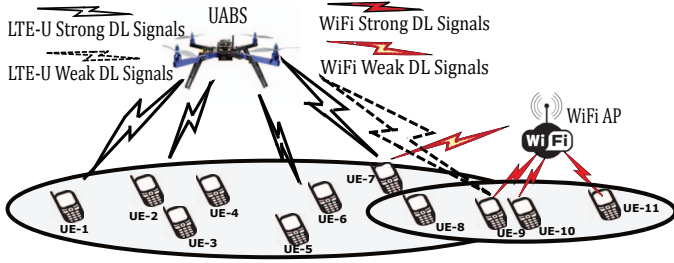
Fig. 1: UAV assisted LTE-U/WiFi heterogeneous network. Here UE-8 receives very strong signals both from UABS and AP, UE-7 receives strong signals from UABS and very weak signals from AP. Similarly UE-9 receives strong signals from AP but weak signals from UABS.

considers the interference between the RATs and aims to reduce the degradation in WiFi communications due to LTE-U transmissions. To make the association decision for a user equipment (UE) between LTE-U and WiFi, we consider the loads and link qualities of both LTE-U and WiFi networks. We formulate the problem as a game between WiFi APs and LTE-U UABSs. In this game, the goal of LTE-U UABSs is to maximize their capacity while keeping minimum interference towards WiFi transmissions. In particular, we propose a regret-based learning (RBL) dynamic duty cycle selection (DDCS) technique for configuring LTE-U transmission gaps, so that a satisfactory throughput is maintained both for UABSs and WiFi access points (APs). Simulation results show that the proposed RBL-DDCS yields an improvement of 32% over fixed duty cycle based LTE-U transmission, and an improvement of 10% over Q-learning based DDCS.

The rest of this paper is organized as follows. Section II presents the system model for coexisting LTE-U UABS and WiFi AP networks. In Section III, we describe the load-balancing RAT selection technique among LTE-U UABSs and WiFi APs, while in Section IV we introduce the DDCS framework for LTE-U and WiFi coexistence. Section V develops our proposed RBL-DDCS technique and provides convergence analysis for it. Section VI presents various simulation results on the performance of the proposed RBL-DDCS approach, and the last section provides some concluding remarks.

II. SYSTEM MODEL

A. Hybrid UABS-WiFi Architecture

Consider a hybrid cellular network such as in Fig. 1 that consists of UABSs with LTE-U capability and ground-based WiFi APs that provide network coverage in the aftermath of a diaster. The LTE-U UABSs can be rapidly deployed to form an on-the-fly cellular network to fulfill the throughput requirements of PSN. We assume that the UABSs are dynamically positioned in order to maximize the combined coverage with the WiFi network. Both WiFi and LTE-U RATs transmit in the 5 GHz unlicensed band, and they share a common 20 MHz unlicensed spectrum for transmission. In this model, our key goal is to operate the LTE-U based UABSs to fill the coverage

gaps without jeopardizing the performance of ground WiFi users.

We focus on the downlink transmission which consists of a set $\mathcal{U} = \{1, ..., U\}$ of LTE-U UABSs and a set $\mathcal{W} = \{1, ..., W\}$ of WiFi APs (hereinafter, both referred to as base stations (BSs)). Let $\mathbf{x}$ be a vector of two-dimensional coordinates measured with respect to the origin, and L_b be the coverage area of BS b such that any given UE at a given location $\mathbf{x}$ is served by BS b if $\mathbf{x} \in L_b$. For an arbitrary UE at location $\mathbf{x} \in L_b$, let the nearest UABS $u \in \mathcal{U}$ at a distance $\|\mathbf{x}_u\|$ be its UABS of interest (UOI) and the nearest WiFi AP $w \in \mathcal{W}$ at a distance $\|\mathbf{x}_w\|$ be its WiFi AP of interest (WOI). Let the transmitted power of UABS and AP be denoted as P_{tx}^u and P_{tx}^w, respectively, and let h_u denote the altitude of the UABS u.

Based on these definitions, the average received signal power from the UOI and the WOI is respectively given by:

$$S_u(\mathbf{x}) = \frac{P_{\text{tx}}^u}{10^{(\text{PL}_{\text{avg}}/10)}} \ , \ S_w(\mathbf{x}) = \frac{P_{\text{tx}}^w K_w H_w}{\|\mathbf{x} - \mathbf{x}_w\|^{\beta}} \ , \qquad (1)$$

where K_w accounts for geometric parameters such as transmitter/receiver antenna heights of the WiFi AP, β is the path loss exponent (PLE), and the random variable $H_w \sim \exp(1)$ accounts for Rayleigh fading over the user-WiFi AP link.

The channels between the BSs and the UEs are modeled as additive white Gaussian noise (AWGN) channels with noise variance σ^2 for the downlink transmission. Let $\ _u$ be the transmission indicator of UABS u such that $\ _u = 1$ indicates the ON state while $\ _u = 0$ reflects the OFF state. Then, the received signal to interference plus noise ratio (SINR) values of a UE at location $\mathbf{x} \in L_b$ due to the UOI and the WOI can be respectively expressed as:

$$\eta_u(\mathbf{x}) = \frac{S_u(\mathbf{x}) \ _u}{\displaystyle\sum_{\forall u' \in \mathcal{U} \backslash u} S_{u'}(\mathbf{x}) \ _{u'} + \sum_{\forall w' \in \mathcal{W}} S_{w'}(\mathbf{x}) + \sigma^2} \ , \qquad (2)$$

$$\eta_w(\mathbf{x}) = \frac{S_w(\mathbf{x})}{\displaystyle\sum_{\forall u' \in \mathcal{U}} S_{u'}(\mathbf{x}) \ _{u'} + \sum_{\forall w' \in \mathcal{W} \backslash w} S_{w'}(\mathbf{x}) + \sigma^2} \ . \qquad (3)$$

B. Path Loss Model

For the UABS channel, we use the air-to-ground propagation model as described in [11] and [12] from UABS to UE. Therefore line of sight (LoS) and non-line of sight (NLoS) path loss (PL) from a given UABS to a given UE can be expressed (in dB) as:

$$\text{PL}_{\text{LoS}} = L_{\text{FS}} + \beta \log[d_u(\mathbf{x}, h_u)] \ , \qquad (4)$$

$$\text{PL}_{\text{NLoS}} = L_{\text{FS}} + \beta \log[d_u(\mathbf{x}, h_u)] + \log L_{\text{NLoS}} \ , \qquad (5)$$

where L_{FS} is the free space PL given by $20 \log(d_u(\mathbf{x}), h) + 20 \log(f) + 20 \log(4\pi/c)$, and $d_u(\mathbf{x}, h_u) = \sqrt{\|\mathbf{x} - \mathbf{x}_u\|^2 + h_u^2}$ is the distance between UE and UABS u, and L_{NLoS} is an additional attenuation factor due to the NLOS connection. The probability of LoS connection is given by [12]:

$$P(\text{LoS}) = \left(1 + a \exp(-b[\theta - a])\right)^{-1} \qquad (6)$$

where a and b are constants which depend on the environment (rural, urban, dense urban, or others) and $\theta = \sin^{-1}\left(h_u/d_u(\mathbf{x}, h_u)\right)$ is the elevation angle. The average path loss can then be expressed as (in dB):

$$\text{PL}_{\text{avg}} = P(\text{LoS}) \times \text{PL}_{\text{LoS}} + P(\text{NLoS}) \times \text{PL}_{\text{NLoS}} , \quad (7)$$

where $P(\text{NLoS}) = 1 - P(\text{LoS})$.

III. RAT SELECTION AMONG LTE-U AND WiFi

As shown in Fig. 1, even when the received signal strength (RSS) is good from both UABS and WiFi APs, one RAT might be highly loaded over other. Therefore, considering the proper balance between RSS and cell load is critical for UE association. We use the α-optimal user association and cell load balancing as described in [13] for the RAT selection between WiFi and LTE-U. We define the traffic load density $\gamma(\mathbf{x})$ of any UE at location $\mathbf{x} \in L_b$ as $\gamma_b(\mathbf{x}) = \frac{\lambda(\mathbf{x})}{\mu(\mathbf{x})}$, where $\lambda(\mathbf{x})$ is the file transfer requests arrival rate per unit area. Hence, the load follows an inhomogeneous Poisson point process and file sizes are independently distributed with mean $\frac{1}{\mu(\mathbf{x})}$ at location $\mathbf{x} \in L_b$. Let $R_b(\mathbf{x})$ be the data rate of any UE at location $\mathbf{x}$. The fraction of time BS b needs to serve the traffic $\gamma_b(\mathbf{x})$ from BS b to location $\mathbf{x}$ is defined as $P_{\text{tx}}^u = \gamma_b(\mathbf{x})/R_b(\mathbf{x})$. Thus, the *system-load density* of BS b is given by:

$$\rho_b = \int_{\mathbf{x} \in L_b} q_b(x)\mathrm{d}x. \quad (8)$$

In some of the existing works such as [14] and [15], UE association is typically used based on the received SINR values while ignoring the load of BSs. This may lead to overloading of BSs since a UE may associate to a BS which gives high SINR even though that particular BS is already overloaded. Moreover, in [13] and [16], the authors study load balancing; however, they only consider the RSS instead of the received SINR for UE association. Thus, we propose a new UE association criteria by considering both load balancing and received SINR aspects.

Under such a scheme, UEs must be aware of the current load of BSs, and each BS needs to transmit their estimated *system-load density* $\rho_b^*(t)$ via a broadcast control message. Then, a UE at location $\mathbf{x} \in \{L_u \cap L_w\}$ selects one of UOI and WOI according to the following criteria [16]:

$$b(\mathbf{x}, t) = \underset{b \in \{u, w\}}{\arg\max} \ \{(1 - \rho_b^*(t))^\alpha \ \eta_b(\mathbf{x}, t)\}, \quad (9)$$

where $\alpha \geq 0$ is a parameter specifying the desired degree of load balancing. For $\alpha = 0$, classical SINR based UE association is achieved, while for $\alpha > 0$, the BS load affects the UE association. The $\eta_b(\mathbf{x}, t)$ is the received SINR value as defined in (2),(3) at UE in location $\mathbf{x} \in \{L_u \cap L_w\}$ at time instant t. Each BS will compute its estimated *system-load density* $\rho_b^*(t)$ as a moving-time average as follows:

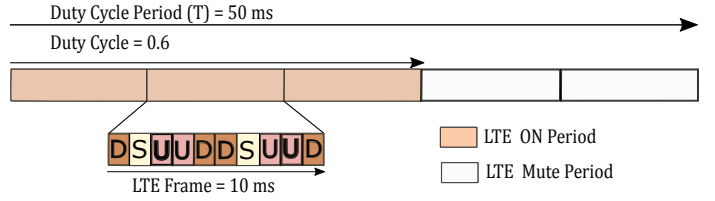$$\rho_b^*(t) = v(t)\rho_b(t-1) + (1 - v(t))\rho_b^*(t-1), \quad (10)$$

Fig. 2: LTE-U frame representation for a duty cycle period of 0.6.

where $v(t)$ is the learning rate of the load estimation, and $v(t)$ is selected such that the load estimation procedure is slower than the UE association process in (9).

IV. COCHANNEL LTE-U AND WiFi CO-EXISTENCE

In this section, we use ON/OFF strategy of LTE-U UABS transmissions to maintain coexistence with WiFi APs. In the proposed approach, the UABSs need to autonomously and dynamically select their transmission duty cycle (DDCS) in order to maximize their utility functions. However, the achievable throughput of a UABS depends not only on its own choice of action but also on remaining UABSs due to interference. Therefore, we formulate a noncooperative game [17], in particular, a U player normal-form game. Here we assume that UABSs are the players of this noncooperative game $\mathcal{G}$, which can be defined as: $\mathcal{G} = (\mathcal{U}, \{\mathcal{S}_i\}_{i \in \mathcal{U}}, \{\Phi_i\}_{i \in \mathcal{U}})$, where Φ_i is the utility of player i, and $\mathcal{S}_i$ is the strategy space of player i. To enable LTE-U transmission in the unlicensed band we consider the LTE frame configuration shown in Fig. 2.

Each player $i \in \mathcal{U}$ has $\mathcal{S}_i^t = \{s_{i,1}^t, ..., s_{i,|\mathcal{S}_i^t|}^t\}$ set of actions where an action of UABS i, s_i^t, is composed of its transmission duty cycle. For our game we used 0.2, 0.4, 0.6 and 0.8 as s_i^t values. Let $\pi_i^t = \left[\pi_i^t(1), ..., \pi_i^t(|\mathcal{S}_i^t|)\right]$ be a probability distribution in which UABS i selects a given action from $\mathcal{S}_i^t$ at time instant t, i.e. π_i^t is UABS i's mixed strategy where s_i^t is the action of player i at time t.

For the utility function, we use the average cell throughput of the UABS i for one duty cycle period (50 ms). Since we model this as a noncooperative game, every player (UABS) tries to maximize its own cell throughput selfishly. Let UABS $i \in \mathcal{U}$ serve N UEs. According to the frame structure in Fig. 2, UABS i schedules its users on each downlink-subframe (DL-SF). Let one duty-cycle period consist of a set $\mathcal{C} = \{1, ..., C\}$ of DL-SFs. As an example, according to our frame structure one duty cycle period consists of $\mathcal{C} = \{1, ..., 30\}$ DL-SFs, with $\|\mathcal{C}\| = 30$.

Let $N_c(\kappa)$ be the number of allocated resource blocks (RBs) for user κ on DL-SF c and ω as the bandwidth of one RB which is equal to 180 kHz in LTE. Then, $\eta_i^c(\mathbf{x}_\kappa)$ is the received SINR value (2) of user-κ (i.e UE-κ) from UABS i at the location $\mathbf{x}_\kappa$ on the DL-SF c. Then we can write down the average cell throughput of the ith UABS for one duty cycle

period (50 ms) as follows:

$$\phi_i^t(s_i^t, \boldsymbol{s}_{-i}^t) = \frac{1}{\|\mathcal{C}\|} \sum_{c=1}^{\|\mathcal{C}\|} \sum_{\kappa=1}^{N} \mathcal{I}_i^c \ N_c(\kappa)\omega \log_2(1 + \eta_i^c(\mathbf{x}_\kappa)),$$
(11)

where $\phi_i^t(s_i^t, \boldsymbol{s}_{-i}^t)$ is the utility function of each player and $s_i^t \in \mathcal{S}_i^t$ is the selected strategy by a player i during the tth duty cycle period. The $\mathcal{I}_i^c$ is an indicator function and $\mathcal{I}_i^c = 1$ indicates that UABS i is in ON state on c^{th} DL-SF. The UEs are scheduled by transmission time intervals (TTIs) which is 1 ms in LTE. Therefore, the achievable DL throughput for UABS i for one duty cycle period has been taken as the utility function of UABS i. We can see that (11) depends on the received SINR value of player i, which also depends on the received interference from other players (2). Hence potentially, this represents the player i utility function which depends on other players' actions as well thus motivating a game-theoretic approach.

We can see that the number of non zero $\mathcal{I}_i^c$ terms equal to the number of ON state DL-SFs within one duty cycle period (T). However we know that the action of a player i, s_i^t (the duty cycle of player i) is proportional to the number of ON state DL-SFs within T. Therefore (11) represents player i's utility as a function of players' actions. Game $\mathcal{G}$ is played as seen later, to find the solution of this game, the players must engage in a learning algorithm over discrete time slots $t \in \{1, 2, ...\}$. In this game, one time slot is equal to one duty cycle period which is 50 ms for this frame structure.

V. CORRELATED EQUILIBRIUM AND LEARNING

Our goal is to develop a distributed mechanism to solve the LTE-U ON/OFF game and reach the correlated equilibrium (CE). To this end, we introduce a class of no-regret learning algorithm called regret-matching [18], which converges to the CE and has been used widely in resource allocation problems of wireless networks such as in [19] [20].

A. Correlated Equilibrium

The concept of CE generalizes the idea of a Nash equilibrium to enable correlated strategy choices across the players. The idea is that a strategy profile is chosen randomly according to a joint distribution. Given the recommended strategy, it is in the players' best interests to conform with this strategy.

Definition 1: For the proposed game $\mathcal{G}$, define π as the probability distribution over the joint strategy space $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2 \times ..... \times \mathcal{S}_U$. The set of correlated equilibria $\mathcal{C}_e$ is the convex polytope given by:

$$\mathcal{C}_e = \Big\{ \pi : \sum_{\boldsymbol{s}_{-i} \in \mathcal{S}_{-i}} \pi(s_i, \boldsymbol{s}_{-i})[\phi_i(s_i', \boldsymbol{s}_{-i}) - \phi_i(s_i, \boldsymbol{s}_{-i})] \leq 0,$$
$$\forall i \in \mathcal{U}, s_i, s_i' \in \mathcal{S}_i \Big\} . \quad (12)$$

B. Regret Based Learning for Co-Channel Coexistence

In this subsection, we propose a learning algorithm based on the notion of *regret matching* [18]. In order to reach the CE, we need to select a suitable probability distribution over the strategy space of the players. In [18], it is shown that if we select the probability distributions according to the proposed regret matching algorithm, it will reach the CE.

The algorithm will dynamically select the LTE-U duty cycle for UABSs transmission. In each time slot t (t^{th} duty cycle period), player i chooses strategy s_i^t according to the probability distribution $\pi_i^t(k)$; $k \in \mathcal{S}_i$. We can define $\pi_i^t(k)$ as follows:

$$\pi_i^t(k) = \frac{1}{|\mathcal{S}_i|}; \ \ t = 1 \ \ (\text{Initialization}), \quad (13)$$

$$\pi_i^t(k) = \begin{cases} \frac{1}{\mu^{t-1}} \ \mathcal{R}_i^{t-1}(j,k) \ ; \ \ k \neq j; \ t > 1, \\ 1 - \sum_{h \neq j, h \in \mathcal{S}_i} \pi_i^t(h) \ ; \ \ k = j; \ t > 1. \end{cases} \quad (14)$$

Here, in slot t ($t > 1$), given the previously chosen strategy j in slot ($t-1$), the probability of $\pi_i^t(k)$ of choosing k given by above is expressed as a function of the regret of having used j instead of k in the past ($j \neq k$) ($\mathcal{R}_i^{t-1}(j,k)$). The probability of staying with the previous strategy j (i.e., choosing $k = j$), is simply the remaining probability, which is given by:

$$1 - \sum_{h \neq j, h \in \mathcal{S}_i} \pi_i^t(h). \quad (15)$$

At time slot t, the regret of selecting strategy j instead of k ($j \neq k$) is defined as follows:

$$\mathcal{R}_i^t(j,k) \triangleq \ \max\{\zeta_i^t(j,k), 0\} , \quad (16)$$

$$\zeta_i^t(j,k) = \frac{1}{t} \sum_{\tau=1}^{t} (\phi_i^\tau(k, \boldsymbol{s}_{-i}^\tau) - \phi_i^\tau(j, \boldsymbol{s}_{-i}^\tau)) . \quad (17)$$

The expression $\mathcal{R}_i^t(j,k)$ in (16) can be viewed as a measure of the average regret. The μ^{t-1} is the summation of regrets over all actions $k \in \mathcal{S}_i$ in the previous time slot (i.e. $t-1$):

$$\mu^{t-1} = \sum_{k \in \mathcal{S}_i} \mathcal{R}_i^{t-1}(j,k). \quad (18)$$

We can see that the regret learning algorithm requires user i to know not only its own payoffs and the strategy history, but also the payoff of strategy k which was not realized in the past. Therefore, we follow the following procedure to overcome this problem. When $t = 1$, the probability distribution of choosing an action for player i is uniform (12). Thus, player i selects an action arbitrarily, e.g., $s_i = 0.6$. According to the utility function formula, player i's capacity/utility is calculated per DL-SFs. Therefore, even if player i selects the duty cycle as 0.6, player can calculate its utility for $s_i = 0.2$ and $s_i = 0.4$ as well. The problem arises when determining the utility for $s_i = 0.8$, which has not been chosen by the player i. In this case player i predicts the value based on the number of allocated RBs. Then, player i can easily compute its regret values for

all the actions and based on that values player can determine an action which has the maximum regret to play for next state (i.e., $t = 2$).

On the other hand, when $t > 1$, player i chooses an action which is determined based on the previous state regrets. Now, consider the utility functions calculations at t^{th} duty cycle period. Let's say according to the regret values of $(t-1)^{\text{th}}$ duty cycle period, player i determines that it should select 0.4 as the duty cycle (action) for the tth duty cycle. That means player i can compute utility function for $s_i = 0.2$ and $s_i = 0.4$ but unable to calculate for $s_i = 0.6$ and $s_i = 0.8$ at t^{th} duty cycle period. In this case it uses previous state $((t-1)^{\text{th}})$ utility function values for $s_i = 0.6$ and $s_i = 0.8$ to compute regret values at t^{th} duty cycle period.

Player i will select its action (duty cycle value) at each duty cycle period according to the previously described procedure until its all regrets values are less than some threshold γ. Procedure for properly selecting γ, is described in the following section such that the convergence is achieved. The complete algorithm is described in Algorithm 1.

C. Convergence Analysis

In [18, Theorem A], the authors show that each player's regrets $R_i^t(j,k)$ converge to zero almost surely. Also [18] provides the speed of convergence for the expectation of the regret E$[R_i^t(j,k)]$ as $\mathcal{O}(1/\sqrt{(t)})$ which is an upper bound for the rate of convergence.

However, in our model, players (UABSs) can not wait too long in practice such that all its regrets become zero. Therefore we need to guarantee the convergence within a certain amount of time. We define a threshold γ for regrets and each player perform regret calculations and actions selection

until the maximum regret of each player is greater than γ. If $\sup\{R_i^{t'}(j,k)\} \leq \gamma$ when $t = t'$, then we consider that the algorithm has come to its point of convergence. Then this will lead to the following theorem.

Theorem 1: *Assume that, our proposed algorithm has the point of convergence when $t = t'$. Then, t' is bounded below by $1/\gamma^2$ asymptotically i.e. $t' = \Omega(1/\gamma^2)$.*

Proof. According to the Lemma 1, $\exists M > 0$ s.t.

$$\mathbb{E}\big[R_i^t(j,k)\big] \leq M\Big\{\frac{1}{\sqrt{t}}\Big\}, \tag{19}$$

Let $t = t'$ be our point of convergence. Then, we have

$$\sup\big(\mathbb{E}\big[R_i^{t'}(j,k)\big]\big) = M/\sqrt{t'} . \tag{20}$$

Also, according to the our definition of threshold γ, we have

$$\sup\big(R_i^{t'}(j,k)\big) \leq \gamma \implies \mathbb{E}\Big[\sup\big(R_i^{t'}(j,k)\big)\Big] \leq \gamma, \tag{21}$$

by using the property $\sup\big(\mathbb{E}[f(x)]\big) \leq \mathbb{E}\big[\sup(f(x))\big]$. Then we can write

$$\sup\Big(\mathbb{E}\big[R_i^{t'}(j,k)\big]\Big) \leq \mathbb{E}\Big[\sup\big(R_i^{t'}(j,k)\big)\Big]. \tag{22}$$

By using (20) and (21), we have

$$M/\sqrt{t'} \leq \mathbb{E}\big[\sup\big(R_i^{t'}(j,k)\big)\big] \leq \gamma, \tag{23}$$

which, after some manipulation, can be written as

$$t' \geq \big(M/\gamma\big)^2 , \ t' = \Omega\big(1/\gamma^2\big). \tag{24}$$

$\square$

Based on (24) we conclude that the convergence time is bounded below by $1/\gamma^2$ asymptotically. Therefore by properly selecting a sufficiently small γ we can achieve the convergence within the required amount of time, hence ϵ-CE. In our model, convergence means, player i has come to a final decision regarding his strategy (duty cycle), it is going to play. Because at the learning phase, UABSs switch their strategies (duty cycles) in each 50 ms duty cycle period.

VI. SIMULATION RESULTS

For our simulations, we consider a two layer cell layout for simplicity as shown in Fig. 3. Each layer consists of 7 cells. There are 10 WiFi users and 10 LTE-U users, associated with each WiFi AP and LTE-U UABS, respectively. WiFi users and LTE-U users move within the cell with a speed of 3 km/h, and their traffic arrival rates are given by $\lambda_{\text{WiFi}} = \lambda_{\text{LTE-U}} = 2.5$ in all the simulations. The MAC and PHY layers for LTE and WiFi are implemented as discussed in [8].

In Fig. 4, we present the aggregate LTE-U and WiFi capacity performance for different fixed duty cycles at LTE-U, Q-learning based DDCS, and RBL based DDCS. Results show that due to better frequency reuse, when a fixed duty cycle is used at LTE-U, aggregate capacity is maximized compared to other fixed duty cycle approaches. On the other hand, this
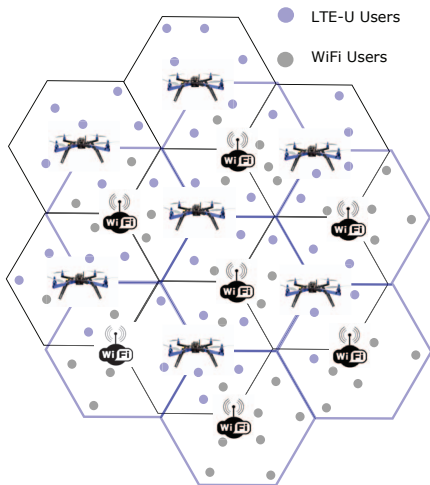
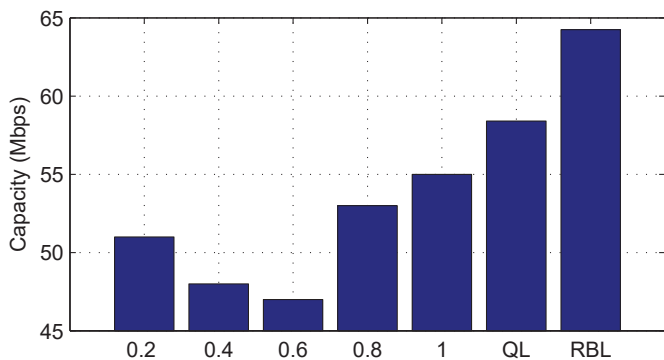Fig. 3: WiFi APs and LTE-U UABSs in a two-layer cell layout.

Fig. 5: WiFi DL SINR distribution with fixed and DDCS based LTE-U duty cycle configurations ($h = 50$ m, ISD $= 50$ m).

Fig. 4: Aggregate WiFi and LTE-U DL capacity different fixed duty cycles, Q-learning based DDCS, and RBL based DDCS ($h = 50$ m, ISD $= 50$ m).

Fig. 6: WiFi DL SINR distribution for different UABS heights with RBL ISD $= 50$ m.

also penalizes WiFi capacity, since LTE-U BSs transmit all the time. With Q-learning and RBL based DDCS, LTE-U BSs dynamically learn and adapt the duty cycle parameter. In particular, we can see that with RBL DDCS, there is a 32% improvement in aggregate capacity compared to always-on LTE transmission, and 10% improvement compared to Q-learning based DDCS.

Fig. 5 shows the WiFi DL SINR distributions with different duty cycle adaptation strategies at LTE-U BSs. Under a fixed LTE-U duty cycle of 0.2, SINR values of WiFi users are generally greater than DDCS based procedures, since there is fairly low LTE-U air time and therefore less interference on WiFi transmissions. However, as was seen in Fig. 4, all fixed duty cycle scenarios resulted in lower aggregate throughput. With learning based DDCS procedures, WiFi users can receive fairly good SINR values, while maintaining improved system capacity as observed in Fig. 4.

In Fig. 6, the distribution of the WiFi DL SINR is presented for different heights 50 m, 100 m, and 200 m of the LTE-U UABSs. We observe that WiFi SINR improves at larger heights due to lower interference observed from LTE-U UABSs. On
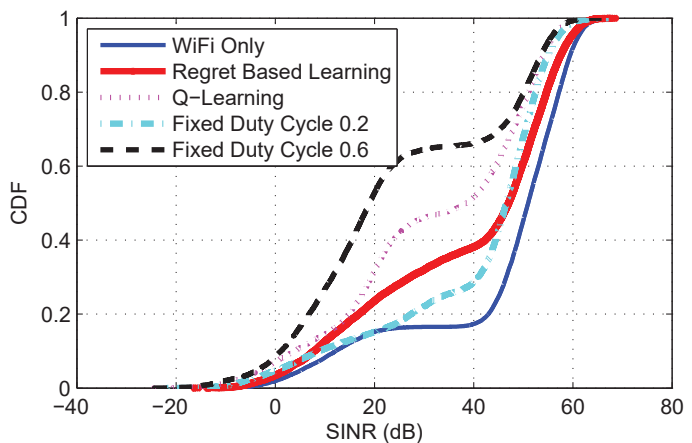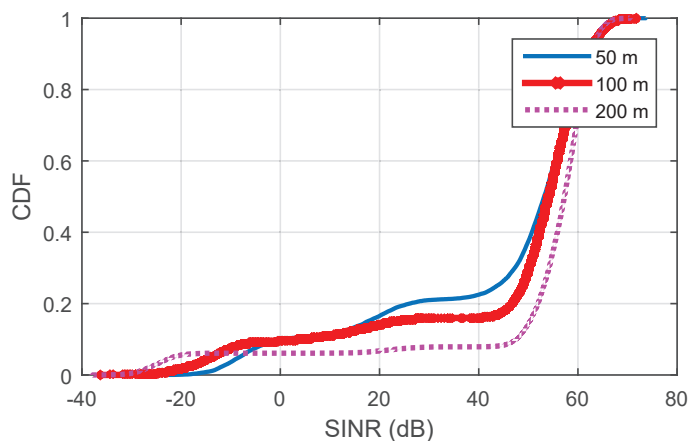
the other hand, larger UABS heights will also result in lower LTE-U capacity, as will be discussed in following figures.

In Fig. 7, considering coexistence scenario in Fig. 3, the RBL-DDCS capacity of LTE-U downlink, WiFi downlink, and aggregate LTE-U and WiFi downlink transmissions is presented for different UABS altitudes. When LTE-U UABS altitude is increased, aggregate LTE-U capacity decreases due to weaker signal strength, while WiFi capacity increases. On the other hand, the aggregate LTE-U/WiFi capacity is seen to be maximized at around 100 m UABS height.

In Fig. 8, we show the RBL-DDCS capacity of LTE-U downlink, WiFi downlink, and aggregate LTE-U and WiFi transmissions for different ISDs and different traffic arrival rates. From this figure, we can see that the capacity of LTE-U UABSs is significantly degraded for larger ISDs, while a relatively lower degradation is observed for WiFi. This is due to the longer transmission distances from LTE-U UABSs ($h = 50$ m) to the users, which, in turn, will result in a weaker received signal strength at the LTE-U users.
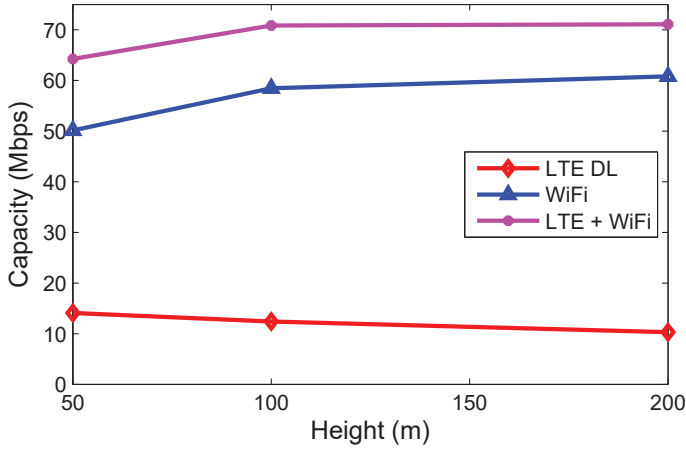
Fig. 7: RBL-DDCS capacity of LTE-U downlink, WiFi downlink, and aggregate LTE-U and WiFi downlink transmissions for different UABS altitudes in the coexisting LTE-U/WiFi scenario of Fig. 3.
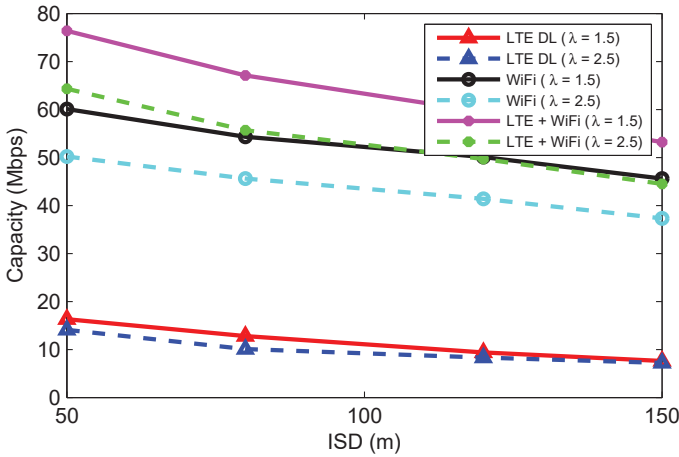

Fig. 8: RBL-DDCS capacity of LTE-U downlink, WiFi downlink, and aggregate LTE-U and WiFi transmissions for different ISDs and different traffic arrival rates (λ) in the coexisting LTE-U WiFi scenario of Fig. 3 ($h = 50$ m).

VII. CONCLUDING REMARKS

In this paper, we have shown that LTE-U can be used as a promising RAT for UABSs which provide coverage and capacity requirements in the aftermath of a natural disaster. We have proposed a regret based learning dynamic duty cycle selection algorithm to configure periodic transmission gaps of LTE-U. Simulation and analytical results have shown that due to the proposed regret based learning dynamic duty cycle selection procedure, UABS network and remaining WiFi network can effectively coexist in the unlicensed band with improved aggregate system capacity but without much degrading each RAT's performance.

REFERENCES

[1] Y. Shibata, N. Uchida, and Y. Ohashi, "Problem analysis and solutions of information network systems on East Japan Great Earthquake," in *Proc. Int. Conf. Advanced Information Networking and Applications Workshops (WAINA)*, 2012, pp. 1054–1059.

[2] A. Merwaday and I. Guvenc, "UAV assisted heterogeneous networks for public safety communications," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Mar. 2015, pp. 329–334.

[3] A. M. amd A. Tuncer, A. Kumbhar, and I. Guvenc, "Unmanned aerial base stations for public safety communications," IEEE Vehic. Technol. Mag., 2016.

[4] Qualcomm, "Extending LTE-Advanced to unlicensed spectrum," *white paper, https://www.qualcomm.com/media/documents/files/white-paper-extending-lte-advanced-to-unlicensed-spectrum.pdf*, Dec. 2013.

[5] K. Gomez, T. Rasheed, L. Reynaud, and S. Kandeepan, "On the performance of aerial LTE base-stations for public safety and emergency recovery," in *Proc. IEEE Globecom Workshops (GC Workshops)*, Dec. 2013, pp. 1391–1396.

[6] Y. Gu, M. Zhou, S. Fu, and Y. Wan, "Airborne WiFi networks through directional antennae: An experimental study," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2015, pp. 1314–1319.

[7] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 3939–3963, 2016.

[8] N. Rupasinghe and I. Guvenc, "Reinforcement learning for licensed-assisted access of LTE in the unlicensed spectrum," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2015, pp. 1279–1284.

[9] A. Cavalcante, E. Almeida, R. Vieira, F. Chaves, R. Paiva, F. Abinader, S. Choudhury, E. Tuomaala, and K. Doppler, "Performance evaluation of LTE and Wi-Fi coexistence in unlicensed bands," in *Proc. IEEE Vehic. Technol. Conf. (VTC)*, Jun. 2013, pp. 1–6.

[10] F. Abinader, E. Almeida, F. Chaves, A. Cavalcante, R. Vieira, R. Paiva, A. Sobrinho, S. Choudhury, E. Tuomaala, K. Doppler, and V. Sousa, "Enabling the coexistence of LTE and Wi-Fi in unlicensed bands," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 54–61, Nov. 2014.

[11] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2014, pp. 2898–2904.

[12] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569–572, Dec. 2014.

[13] H. Kim, G. de Veciana, X. Yang, and M. Venkatachalam, "Alpha-optimal user association and cell load balancing in wireless networks," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–5.

[14] A. Merwaday, S. Mukherjee, and I. Güvenç, "Capacity analysis of LTE-Advanced HetNets with reduced power subframes and range expansion," *EURASIP J. Wireless Commun. Netw. (JWCN)*, no. 1, pp. 1–19, 2014.

[15] S. Mukherjee, "Downlink SINR distribution in a heterogeneous cellular wireless network with max-SINR connectivity," in *Proc. IEEE Allerton Conf. Commun., Control, Computing*, Sep. 2011, pp. 1649–1656.

[16] S. Samarakoon, M. Bennis, W. Saad, and M. Latva-Aho, "Dynamic clustering and ON/OFF strategies for wireless small cell networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2164–2178, Mar. 2016.

[17] W. S. T. B. Z. Han, D. Niyato and A. Hjrungnes, "Game Theory in Wireless and Communication Networks: Theory, Models, and Applications," *Cambridge University Press*, 2012.

[18] S.Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, 2000.

[19] Z. Han, C. Pandana, and K. Liu, "Distributive opportunistic spectrum access for cognitive radio using correlated equilibrium and no-regret learning," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2007, pp. 11–15.

[20] L. Chen, "A distributed access point selection algorithm based on no-regret learning for wireless access networks," in *Proc. IEEE Vehic. Technol. Conf. (VTC)*, May 2010, pp. 1–5.