A Novel Machine Learning Approach to Estimating KPI and PoC for LTE-LAA-based Spectrum Sharing

Susanna Mosleh[†§], Yao Ma[‡], Jacob D. Rezac[‡], and Jason B. Coder[‡]

[†]Associate, Communications Technology Laboratory, National Institute of Standards and Technology, USA
[§]Department of Physics, University of Colorado, Boulder, Colorado, USA
[‡]Communications Technology Laboratory, National Institute of Standards and Technology, USA

***Abstract*—Machine learning (ML) approaches have been extensively exploited to model and to improve wireless communication networks in the past few years. Nonetheless, the estimation of key performance indicators (KPIs) and their uncertainties in Long Term Evolution License Assisted Access (LTE-LAA) based coexistence systems is not adequately addressed. For example, it is not clear if an ML method can accurately predict achievable KPIs (e.g. throughput) and the probability of coexistence (PoC) of LTE-LAA coexistence systems based on partial or no information of MAC and physical layer protocols and parameters. In this paper, we develop a novel ML method by combining a neural network with a logistic regression algorithm to track and estimate KPIs and PoC of coexisting LTE-LAA and wireless local area network (WLAN) links. This ML method can be applied when KPI samples at the base stations (BSs) and access points (APs) are available, without using knowledge of MAC and physical layer parameters. Comparison between the ML and simulation results indicate that the proposed ML method can track the system KPIs and predict the system PoC with good accuracy.**

***Index Terms*—Artificial neural network, LTE-LAA, logistic regression, MAC layer, machine learning, PHY layer, wireless coexistence, WLAN.**

I. Introduction

Wireless communications are tightly integrated in our daily lives. Laptops, tablets, smartphones, and online social networking applications make a level of connectivity to the world available that we have never experienced in the past. This trend continues to dramatically increase wireless network dimensions in terms of subscribers and data throughput [1], especially in the realm of the Internet of Things (IoT). As a consequence, wireless device protocols are beginning to transition from an exclusively-licensed spectrum environment to a shared one, and utilizing the unlicensed spectrum bands seems to be inescapable.

Long Term Evolution (LTE) operating in unlicensed bands, such as license assisted access (LAA), was introduced to improve the spectral efficiency and to help the cellular industry deal with the shortage of the spectrum [2]. However, there are many challenges to overcome in order for multiple networks to constructively share a spectrum. Hence there is a need to accurately evaluate spectrum sharing performance among operators; ensuring the effectiveness of coexistence requires careful consideration.

Recently, artificial intelligence (AI) and machine learning (ML) are having a transformative effect in almost every industry. ML is an important tool for the support of next-generation of wireless device networks. Future 5G and beyond mobile terminals are expected to access the spectral bands using highly developed spectrum learning and inference. However, owing to the involved network topologies, coordination schemes, and the various end-user applications, future networks will be immensely more intricate. Hence, obtaining many optimal key performance indicators (KPIs) might be computationally infeasible or undesirable. Moreover, due to nonlinearity in underlying wireless channels, modeling the end-to-end system's behavior analytically is not easily achieved. It gets even more difficult when it comes to managing networks efficiently in a coexistence scenario where different types of networks need to share a given section of spectrum. ML algorithms can mitigate the underlying unknown non-linearities and can reduce the network complexity so as to be tractable and useful while keeping up ambitious performance goals.

In line with standardization efforts on the evaluation of wireless coexistence [3], [4], in this paper, we develop an ML method to track and estimate the probability of coexistence (PoC) of WLAN and LTE systems in an intelligent and adaptive way. Accurately quantifying the PoC in a given shared spectrum is principal to the evaluation of wireless coexistence, as discussed in the ANSI C63.27 standard [3]. Our proposed method builds up an effective sharing of spectrum, provides an accurate coexistence performance evaluation, and helps to design future radio technologies (e.g., 5G new radio in unlicensed spectrum (5G NR-U) [5]).

The main contribution of this study is to check whether or not ML algorithms can track and provide reliable estimates of KPIs and PoC of coexisting networks. We aim to develop an ML model which provides reliable PoC estimates of various MAC and physical layer parameters, and apply this model to coexistence systems where analytical KPI formulas are not available. These results can support future versions of the ANSI C63.27 standard, and provide insight on developing new ML methods to support KPI uncertainty evaluation in 5G coexistence systems.

We propose a novel PoC estimator to provide an improved assessment of concurrent operation of WLAN and LTE networks in the unlicensed band. Specifically, we take the operations of both networks in the MAC and physical layers into account, and employ a novel machine learning algorithm by leveraging neural network with a logistic regression method that utilizes all MAC and physical layers' parameters as inputs (such as contention window size, maximum back-off stage, slot durations, and link signal-to-noise ratio (SNR)) and generates PoC as an output. To do this, we use a neural network to estimate KPIs from input MAC and physical layer parameters. A logistic regression model is then used to estimate PoC from estimated KPIs. The proposed algorithm enables both networks to evaluate wireless coexistence and guarantees a constructive coexistence among operators in an intelligent and a well-planned course of action. It is worth noting that the proposed technique we develop here could be incorporated into many other spectrum sharing systems and we use LTE simply as an example.

The remainder of this paper is organized as follows: Section II describes the system model and assumptions required for our analysis. Section III presents the problem formulation and introduces our proposed intelligent PoC estimator. The impacts of the MAC and physical layer parameters in evaluating a coexistence scenario is also explained in Section III. Simulation results are shown and discussed in Section IV. Finally, in Section V, an overview of the results and some concluding remarks are presented.

II. System Model

Consider a downlink coexistence scenario where two mobile network operators (MNOs) share the same unlicensed bands for operation in an industrial, scientific, and medical (ISM) radio band. Note that we are primarily interested in the operation of cellular base stations in an unlicensed band. However, the LTE base stations may have permission to utilize a licensed band as well. We assume each unlicensed band can be shared between the MNOs in a time sharing fashion. The LAA network consists of L eNodeBs indexed by the set $\mathcal{L} \triangleq \{n_\ell | \ell = 1, 2, \ldots, L\}$ while the Wi-Fi network is composed of W APs indexed by the set $\mathcal{W} \triangleq \{n_w | w = 1, 2, \ldots, W\}$ APs. Each transmission node $i \in \{\mathcal{L}, \mathcal{W}\}$ serves one single antenna user. The eNodeBs and APs are randomly distributed over a particular area while LAA user equipment (UEs) and Wi-Fi clients are respectively distributed around each eNodeB and AP independently and uniformly. We assume the transmission node i transmits with power p_i and the user association is based on the received power. We also assume (*i*) both Wi-Fi and LAA are in the saturated traffic condition, (*ii*) there is no hidden node problem in the network, *i.e.*, every transmission node i is able to hear one another, (*iii*) a successful transmission happens if only one link transmits at a time, *i.e.*, exclusive channel access (ECA) model is considered, (*iv*) the channel knowledge is ideal (perfect channel sensing among the links), *i.e.*, the only source of packet failure (unsuccessful transmission) is collision, and (*v*) each link is subject to Rayleigh fading and Log-normal shadowing.

Wi-Fi networks utilize a contention-based medium access with a random back-off process, also known as carrier sense multiple access with collision avoidance (CSMA/CA) [6]. In order to discover whether the channel is idle or busy, the station that accesses the medium should sense the channel by performing a clear channel assessment (CCA). The distributed coordination function (DCF) operation progresses if the channel is found out to be idle. Otherwise, the transmitting station refrains from transmitting data until it senses the channel is available. Similarly, LTE uses a listen before talk (LBT) channel access mechanism to maintain fair coexistence with Wi-Fi. Among different LAA-LBT schemes, Category 4 LBT is based on the same Wi-Fi CSMA/CA scheme and is well-suited in a coexistence scenario [2]. Although LTE and Wi-Fi technologies follow the same channel access procedure, they select different carrier sense mechanisms and threshold levels for sensing a channel, and they use different channel contention parameters leading to different unlicensed channel access probabilities which can result in different throughputs.

Our aim is to investigate the feasibility of machine learning algorithms to learn and track system measurement equations (when there is only partial knowledge of MAC/PHY parameters available), or even a system model for which the system equations are not available. To be specific, our goal is to develop a method to map PHY and MAC layers' parameters to PoCs, as depicted in Fig. 1. The PoC here is assessed in terms of the normalized network throughput. Here, we will briefly derive the normalized network throughput of both systems, a quantity that will be used in calculating the PoC later. Conforming with the analytical model in [7], the probability of transmitting a packet by a transmitting node i in a randomly-chosen time slot on an unlicensed channel can be written as

$$p_{\mathrm{tr},i}^{(k)} = \frac{2(1-2p_{c,i}^{(k)})}{(1-2p_{c,i}^{(k)})(1+\mathrm{CW}_{\min,i}^{(k)})+p_{c,i}^{(k)}\mathrm{CW}_{\min,i}^{(k)}(1-(2p_{c,i}^{(k)})^{m_i^{(k)}})}, \quad (1)$$

where $\mathrm{CW}_{\min,i}$ and m_i are the minimum contention window size and the maximum back-off stage of the transmitting node i, respectively, and $p_{c,i}$ is the probability of collision experienced by the i-th transmitting node. The probability of collision experienced by the n_w AP and the n_ℓ eNodeB can be expressed as

$$\begin{aligned} p_{c,n_w} &= 1 - \Big(\prod\nolimits_{\acute{w}\neq w}(1 - p_{\mathrm{tr},n_{\acute{w}}})\Big)\prod\nolimits_{\ell}(1 - p_{\mathrm{tr},n_\ell}), \\ p_{c,n_\ell} &= 1 - \Big(\prod\nolimits_{\acute{\ell}\neq \ell}(1 - p_{\mathrm{tr},n_{\acute{\ell}}})\Big)\prod\nolimits_{w}(1 - p_{\mathrm{tr},n_w}), \end{aligned} \quad (2)$$

respectively, where $\acute{w} = 1, \ldots, W$ and $\acute{\ell} = 1, \ldots, L$ [8]–[10]. The probability of collision can be split into three parts: the

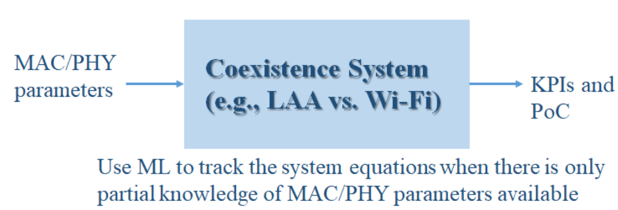

Fig. 1. System Model

probability of collision due to the collision among the Wi-Fi transmissions, among the LAA transmissions, and between the Wi-Fi and the LAA transmissions, respectively given by

$$p_{c,\mathcal{W}} = (1 - p_{tr,\mathcal{L}})\big[p_{tr,\mathcal{W}} - \sum_{w} p_{\text{tr},n_w} \prod_{\acute{w}\neq w}(1 - p_{\text{tr},n_{\acute{w}}})\big],$$
$$p_{c,\mathcal{L}} = (1 - p_{tr,\mathcal{W}})\big[p_{tr,\mathcal{L}} - \sum_{\ell} p_{\text{tr},n_\ell} \prod_{\acute{\ell}\neq \ell}(1 - p_{\text{tr},n_{\acute{\ell}}})\big],$$

and $p_{c,\mathcal{W},\mathcal{L}} = p_{tr,\mathcal{L}} \cdot p_{tr,\mathcal{W}}$. Here $p_{tr,\mathcal{L}} = 1 - \prod_{\ell}(1 - p_{\text{tr},n_\ell})$ and $p_{tr,\mathcal{W}} = 1 - \prod_{w=1}^{W}(1 - p_{\text{tr},n_w})$ denote the LAA's and Wi-Fi's probability of transmission [8], [9], respectively. Moreover, the probability of a successful transmission by the n_w AP and the n_ℓ eNodeB can be respectively written as [8]–[10]

$$p_{s,n_w} = p_{\text{tr},n_w}\big(\prod_{\acute{w}\neq w}(1 - p_{\text{tr},n_{\acute{w}}})\big)\prod_{\ell}(1 - p_{\text{tr},n_\ell}),$$
$$p_{s,n_\ell} = p_{\text{tr},n_\ell}\big(\prod_{\acute{\ell}\neq \ell}(1 - p_{\text{tr},n_{\acute{\ell}}})\big)\prod_{w}(1 - p_{\text{tr},n_w}). \quad (3)$$

Hence, the average length of a time slot can be calculated as

$$T_{\text{avg}} = (1 - p_{tr})\mathbb{E}\{T_{\text{idle}}\} + p_{s,\mathcal{W}}\mathbb{E}\{T_{s,\mathcal{W}}\} + p_{s,\mathcal{L}}\mathbb{E}\{T_{s,\mathcal{L}}\}$$
$$+ p_{c,\mathcal{W}}\mathbb{E}\{T_{c,\mathcal{W}}\} + p_{c,\mathcal{L}}\mathbb{E}\{T_{c,\mathcal{L}}\} + p_{c,\mathcal{W},\mathcal{L}}\mathbb{E}\{T_{c,\mathcal{W},\mathcal{L}}\}$$

where p_{tr} is the probability of occupation of the unlicensed channel, and $p_{s,\mathcal{W}}$ and $p_{s,\mathcal{L}}$ denote the successful transmission probability of the entire Wi-Fi and LAA network, respectively. Moreover, $T_{s,\mathcal{L}}$, $T_{s,\mathcal{W}}$, $T_{c,\mathcal{W}}$, $T_{c,\mathcal{L}}$, and $T_{c,\mathcal{W},\mathcal{L}}$ indicate the time that the channel is being occupied by an LAA successful transmission, a Wi-Fi successful transmission, a collision among the Wi-Fi transmissions, a collision among the LAA transmissions, and a collision between the Wi-Fi and the LAA transmissions, respectively [10]. The network throughput of LAA and Wi-Fi systems as a function of both MAC and physical layers' parameters can be expressed as

$$S_{\mathcal{L}} = p_{s,\mathcal{L}} T_{P,\mathcal{L}} R_{\mathcal{L}} / T_{\text{avg},\mathcal{L}},$$
$$S_{\mathcal{W}} = p_{s,\mathcal{W}} T_{P,\mathcal{W}} R_{\mathcal{W}} / T_{\text{avg},\mathcal{W}}, \quad (4)$$

where $T_{\text{avg},\mathcal{W}}$ ($T_{\text{avg},\mathcal{L}}$) denotes the average time duration to assist a successful transmission in the Wi-Fi (LAA) network, $T_{P,\mathcal{W}}$ ($T_{P,\mathcal{L}}$) indicates the Wi-Fi (LAA) payload duration, and $R_{\mathcal{W}}$ ($R_{\mathcal{L}}$) refers to the Wi-Fi's (LAA's) physical data rate.

Here, we modify the PoC metric described in [11] and define coexistence in terms of the ability to maintain throughput above a certain threshold. Based on the throughput of both LAA and WLAN systems, the PoC metrics that quantify the coexistence performance of these two systems can be calculated as follows

$$\text{PoC}(\eta_{\text{LAA}}, \eta_{\text{WiFi}}) = \mathbb{P}(\text{S}_{\mathcal{L}} > \eta_{\text{LAA}}, \text{S}_{\mathcal{W}} > \eta_{\text{WiFi}}) \quad (5)$$

where equation (5) shows how the joint throughput of Wi-Fi and LAA networks can be mapped to the PoC.

III. Predicting PoC using ML approaches

As mentioned above, an evaluation of the wireless coexistence performance can be determined by the probability of coexistence metric, defined in [11]. This metric will identify the ability of both wireless networks to successfully perform their desired functionality in a given shared spectrum band.

As we discussed earlier, if the physical and MAC layers' parameters are appropriately selected, the throughput of both systems, given by (4), increases, leading to a higher PoC. To be specific, if transmitter i selects the unlicensed spectrum when it is not utilized by the other transmission nodes, then the probability of successful transmission by the i-th node increases, according to (3), leading to a higher throughput. Moreover, if each unlicensed band is selected such that interference is avoided (or at least minimized) among the transmission nodes then there will be a higher SNR, leading to a higher physical data rate (*i.e.,* $R_{\mathcal{L}}$ and $R_{\mathcal{W}}$) and thus higher throughput.

In order to estimate the PoC of Wi-Fi and LAA networks in the unlicensed band, we now develop a machine learning algorithm by leveraging a neural network with a logistic regression method, as depicted in Fig. 2. The proposed model aims to track and estimate KPIs of coexisting LTE-LAA and WLAN links, and evaluate the PoC of these networks. To be specific, the neural network maps the physical and MAC layers input parameters to the KPIs, such as throughput, and the low-cost training logistic regression algorithm conducts a regression analysis to map KPIs to PoC.

A neural network can be thought of as a tracking system used to predict a quantity. It approximates a mapping function from input variables to output variables. In this context, we will use it to track (approximate) the mapping function, i.e., Eq. (4), and determine the LAA and Wi-Fi KPIs based on MAC layer parameters (e.g., contention window size, maximum back-off stages, slot duration) and physical layer characteristics (such as link SNR, link distances, fading parameters). We aim to develop a model which provides reliable approximation of this mapping function of various MAC and physical layer parameters, so we can apply this model to more complex coexistence systems where analytical KPI formulas are not available. The neural network consists of processing nodes organized into three layers, input layer, hidden layer(s), and output layer. These nodes are densely interconnected. In the model that we consider in this paper, known as feed-forward, data moves through the layers in one direction. Each node in a hidden/output layer is connected to several nodes and receives data from them. Moreover, each node in a hidden layer is connected to several nodes and sends data to them. Each connection between nodes is assigned a number named a coefficient or weight. At the training phase, training data is fed to the input layer, it passes through the hidden layer(s), and arrives at the output layer. The data get multiplied by the weights, added together, and go through the activation function of each node. During the training phase, the weights are adjusted until a predefined goal is achieved, such as the mean squared error (MSE) of the training data falling below a pre-set threshold. After this goal is achieved, the trained neural network is applied to test (unseen) data.

Logistic regression can be thought of as a classification algorithm used to assign observations to a discrete set of classes. In this context, we will use it to determine the probability that both LAA and WiFi KPIs are above a user-provided threshold, based on the outputs of the neural network. Logistic regression, when given a set of features, attempts to estimate the probability of success for some function of those features. We use it to estimate the probability of successful coexistence given LAA and Wi-Fi KPIs. In order to solve this prediction problem, we use the gradient descent optimization technique. Let us assume $\mathbf{x}_k$ is the n-dimensional feature vector (the above-mentioned physical and MAC layers' parameters) and y_k is the outcome (PoC of LAA and WLAN networks) of a given test-run $k \in \{1, \ldots, m\}$, where m denotes the total number of observations in the dataset. The feature matrix $\mathbf{X}$ and the outcome vector $\mathbf{y}$ can be written as

$$\mathbf{X}_{n\times m} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,m} \end{pmatrix},$$

and $\mathbf{y}_{1\times m} = \begin{pmatrix} y_1 & y_2 & \cdots & y_m \end{pmatrix}$. The primary assumption leading to logistic regression is that outcome y_k is Bernoulli-distributed with the success probability π_k. Logistic regression proceeds by estimating the probability of a constructive/destructive coexistence (indicated by the variable $\gamma = 1$ or $\gamma = 0$, respectively) given the training set $\mathbf{x}_k$, *i.e.,* $\hat{y}_k = \mathbb{P}(y_k = \gamma | \mathbf{x}_k) = \pi_k^{\gamma}(1 - \pi_k)^{1-\gamma}$.

In order to estimate this probability using the logistic regression method, we need to find a hypothesis $h_{\boldsymbol{\theta}_k}(\mathbf{x}_k) = \hat{y}_k$. The goal is to learn the optimum value of the regression coefficients $\boldsymbol{\theta}_k$ in the sense that $\hat{y}_k$ is approximately equal to the test target y_k. $\boldsymbol{\theta}_k$ is the set of weights corresponding to n features and the bias. In order to learn these weights, we need to define a cost function. A cost function is an estimator of how well our model predicts the known output. This cost function will be used to train the logistic regression model (prediction function) that could predict the PoC in unlicensed band. The logistic regression model can be given as

$$\hat{\mathbf{y}} = \mathrm{S}(\mathrm{diag}(\boldsymbol{\Theta}^{\mathrm{T}}\mathbf{X})), \tag{6}$$

where $\hat{\mathbf{y}}_{1\times m} \triangleq \begin{pmatrix} \hat{y}_1 & \cdots & \hat{y}_m \end{pmatrix}$, $diag(\mathbf{A})$ returns a vector of the main diagonal elements of $\mathbf{A}$, $\boldsymbol{\Theta}_{n\times m}$ is the weight matrix, subscript T is the transpose operator, and $S(z) = \exp(z)/(1 + \exp(z))$ is the so-called sigmoid (logistic) function. The sigmoid function $S(z)$ introduces non-linearity to the model and maps predicted values to probabilities. Then, the hypothesis of logistic regression for the training pair $(\mathbf{x}_k, y_k)$ can be written as

$$\hat{y}_k(\mathbf{x}_k) = h_{\boldsymbol{\theta}_k}(\mathbf{x}_k) = \frac{1}{1 + e^{-\boldsymbol{\theta}_k^T \mathbf{x}_k}}. \tag{7}$$

In order to calculate the weight matrix $\boldsymbol{\Theta}$, a cost function is needed for optimization. Cost functions are usually defined as MSE functions. However, it is known that when using this cost function, the optimization problem turns out to be non-convex and has many local minimums [12, Chapter 3]. Hence, in this paper we use a cost function called "cross-entropy", also known as log loss function, for each pair of training samples, *i.e.,* $(\mathbf{x}_k, y_k)$, as follows [13, Chapter 5]

$$\mathcal{L}(y_k, \hat{y}_k) = -y_k \log(h_{\boldsymbol{\theta_k}}(\mathbf{x}_k)) - (1 - y_k)\log(1 - h_{\boldsymbol{\theta_k}}(\mathbf{x}_k)),$$

which plays the same role as the MSE function, but now the optimization problem becomes convex in $\boldsymbol{\theta}$. This cost function turns the optimization problem into a convex one which is much easier to solve using standard computational techniques. $\mathcal{L}(y_k, \hat{y}_k)$ shows how well the prediction is in a single training example. The cost function of all training samples used in the logistic regression algorithm can be expressed as

$$\begin{aligned} \mathcal{J}(\boldsymbol{\Theta}) &= \frac{1}{m}\sum_{k=1}^{m} \mathcal{L}(y_k, \hat{y}_k) \\ &= \frac{1}{m}(-\mathbf{y}^T \log \hat{\mathbf{y}} - (1 - \mathbf{y})^T \log(1 - \hat{\mathbf{y}})). \end{aligned} \tag{8}$$

By minimizing this cost function with respect to $\boldsymbol{\theta_k}$, the optimum value of the weight matrix $\boldsymbol{\Theta}$ can be found using following optimization

$$\min_{\boldsymbol{\Theta}} \quad \mathcal{J}(\boldsymbol{\Theta}). \tag{9}$$

In order to minimize the cost function $\mathcal{J}(\boldsymbol{\Theta})$ we apply the gradient descent method, which is the most popular approach to iteratively minimize the cost function. The update equation for the k-th observation in the data set can be written as

$$\begin{aligned} \boldsymbol{\theta}_k &= \boldsymbol{\theta}_k - \alpha \nabla_{\boldsymbol{\theta}_k} \mathcal{J}(\boldsymbol{\Theta}) \\ &= \boldsymbol{\theta}_k + \alpha(y_k - h_{\boldsymbol{\theta}_k}(\mathbf{x}_k))\mathbf{x}_k, \end{aligned} \tag{10}$$

where $\nabla \mathcal{J}$ denotes the gradient of the function $\mathcal{J}$, and $0 \leq \alpha \leq 1$ is the step size, also known as the learning rate, and determines how fast the learning happens. As is typically the case in learning algorithms, selecting α requires some care. A small value of α results in a long learning process (which could be detrimental in practice), while a large value of α could cause bouncing around the optimum point.

After calculating the optimal weight matrix $\boldsymbol{\Theta}$, the new label of an unseen sample can be estimated by using (7). To map the estimated label to a discrete class (constructive/destructive coexistence), the predefined threshold value η_{PoC} is selected above which we will classify values as constructive coexistence (high PoC) and below which we classify values as destructive coexistence (low PoC).

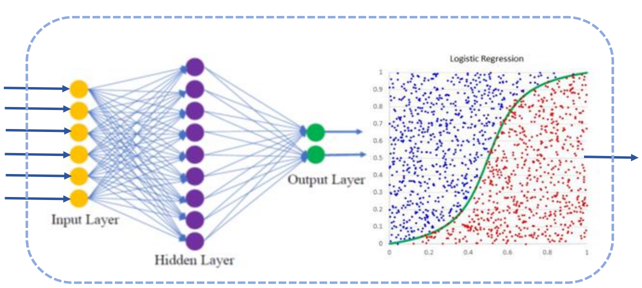

Fig. 2. Proposed Model

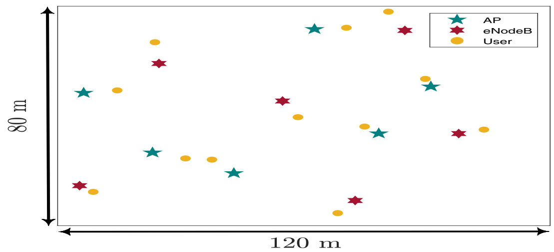

Fig. 3. Simulation Layout

TABLE I
MAC LAYER PARAMETERS

Parameter	value
LAA's packet payload duration	2 ms
Wi-Fi's packet payload duration	1 ms
MAC header	272 bits
PHY header	128 bits
ACK	112 bits + PHY header
SIFS	16 μs
DIFS	34 μs
Idle slot time	9 μs
Wi-Fi contention window size	16
LAA contention window size	16
Wi-Fi maximum backoff stage	6
LAA maximum backoff stage	3

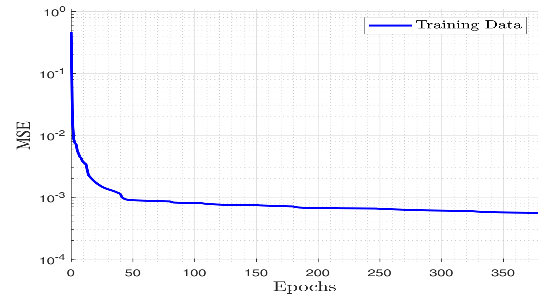

Fig. 4. Mean squared error on the training data in the NN

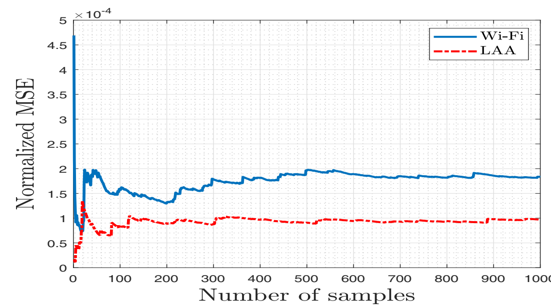

Fig. 5. Normalized mean squared error on the test data in the NN

IV. Simulation Results and Discussions

We evaluate the performance of the proposed algorithm in a coexistence scenario. We simulate a scenario in which 6 eNodeBs compete for an unlicensed channel with 6 APs. All transmitters are randomly distributed over an area of size 120×80 m^2 with a minimum distance of 20 meters, as shown in Fig. 3. All UEs and Wi-Fi clients are independently and uniformly distributed around each eNodeB and AP, respectively. We consider one UE (Wi-Fi client) per eNodeB (AP). Each UE (Wi-Fi client) is assigned to the eNodeB (AP) that provides it with the highest received power. The antenna height of the transmission nodes and users are 6 meters and 1.5 meters, respectively. The carrier frequency is 5 GHz and the bandwidth of each channel is 20 MHz. The path-loss and shadowing between transmission nodes and users are generated following [14] for the indoor scenario. The transmit power at each transmission node is fixed to 23 dBm while the noise figure and the thermal noise level at each user is set to 9 dB and -174 dBm/Hz, respectively [14]. Moreover, we assume the omni-directional antenna pattern with a 0 dBi antenna gain.

According to this geometry and propagation model, we compute the SNR of each link. Given the MAC layer parameters in Table I, we select a sample set of input parameters, and train the neural network to generate output KPIs of both LAA and WLAN networks. The data used for this simulation consists of 1000 feature vectors. Only 30% of the data are used for training and the rest are considered for test. Here, we consider a feedforward neural network consisting of one input layer with 16 nodes (Wi-Fi and LAA contention window sizes, Wi-Fi and LAA backoff stages, and 12 link SNRs), one hidden layer with 16 neurons, and one output layer with two nodes (LAA's and Wi-Fi's throughput). The network is trained and converges quickly, as shown in Fig. 4. After training the network, we applied the trained neural network on unseen (test) data. In order to evaluate the trained network, we calculated the MSE as an average of the squared error $(\mathbf{y}-\hat{\mathbf{y}})^2$, where $\mathbf{y}$ is the KPI values calculated using equation (4) and $\hat{\mathbf{y}}$ is the output of the trained neural network. Comparing the neural network's outputs with the analytical results, the MSE of the test data is equal to 0.0043. The small value of MSE on test data indicates that the neural network tracks the mapping function (system equations) well. We also plot the normalized MSE of both Wi-Fi and LAA KPIs in Fig. 5.

Having the LAA and WLAN throughput pairs as the outputs of neural network, now we map the KPI pair to PoC using the trained logistic regression model and then compare the results with the theoretical one found by Eq. (5). In order to train the logistic regression model, given the predefined thresholds η_{LAA} and η_{WiFi}, we first plot the probability of satisfactory quality of service (QoS) of LAA and Wi-Fi networks, *i.e.*, $\mathbb{P}(S_{\mathcal{L}} > \eta_{\text{LAA}})$ and $\mathbb{P}(S_{\mathcal{W}} > \eta_{\text{WiFi}})$, in Fig. 6 and Fig. 7, respectively. It is observed that the analytical and neural network results follow the same trend. As expected, by increasing the thresholds, the probability of satisfactory QoS of each system decreases. Moreover, the probability of coexistence versus the Wi-Fi's and LAA's throughput threshold is plotted in Fig. 8. The goal is to train the logistic regression model to accurately estimate PoC from estimated throughput and enable both networks to evaluate the wireless coexistence. We pass the test data to the trained logistic regression network.

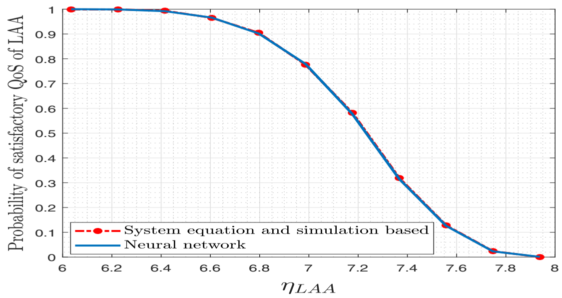

Fig. 6. Probability of satisfactory QoS of LAA versus different thresholds

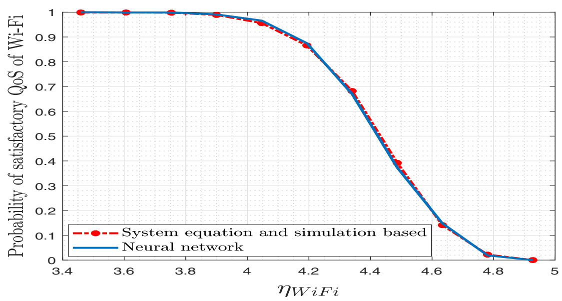

Fig. 7. Probability of satisfactory QoS of Wi-Fi versus different thresholds

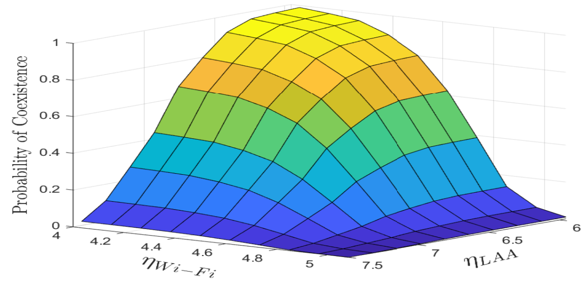

Fig. 8. PoC versus different Wi-Fi's and LAA's throughput threshold

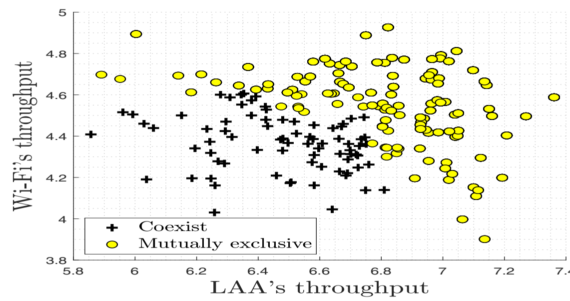

Fig. 9. Labeling Wi-Fi's and LAA's throughput pair using the trained LR

The logistic regression network labeled the unseen data as they can/cannot coexist with each other. Fig. 9 shows that by knowing the throughput of two networks we are able to decide whether or not these two networks can coexist with each other. Moreover, we compare the PoC calculated by equation (5) with the PoC output of the trained logistic regression network and compute the MSE. The MSE on the unseen test data is 0.0658. Furthermore, the logistic regression accuracy, which is defined as percentage of correct predictions, on the training and test data is calculated and given as 84.466% and 84.046%, respectively. In future work we develop further enhanced classification schemes to improve the accuracy.

V. Conclusion

In this paper, we have proposed a machine learning method to accurately track and estimate coexistence performance and its uncertainty between LTE-LAA and WLAN networks in unlicensed bands. The proposed method can work without knowledge of MAC and physical layer protocols and parameters of the system (except the KPI samples). We have also developed a system equation and simulation-based scheme to train and validate the performance of our method. Comparison of the proposed ML method and simulation results has demonstrated that our method can achieve a perfect KPI (i.e., throughput) tracking and a good PoC estimation performance. The proposed method can be extended to the cases where analytical results are not available. In future work, we will develop further enhanced ML methods for KPI and an uncertainty estimation, and generalize our proposed approach to more challenging coexistence scenarios.

References

[1] "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 20172022 White Paper," Cisco Systems, Febuary 2019.

[2] "Study on licensed-assisted access to unlicensed spectrum," 3GPP TR. 36.889 v13.0.0., Jun. 2015.

[3] "American National Standard for Evaluation of Wireless Coexistence," *in ANSI C63.27-2017*, pp. 1-77, May 2017.

[4] "IEEE Standard Definitions and Concepts for Dynamic Spectrum Access: Terminology Relating to Emerging Wireless Networks, System Functionality, and Spectrum Management," *in IEEE Std 1900.1-2008*, pp. 1-62, Oct. 2008.

[5] S. Verma and S. Adhikari, "3GPP RAN1 status on LAA and NR Unlicensed," *IEEE 802.11-18/0542r0*, Mar. 2018.

[6] E. Perahia and R. Stacey, "Next Generation Wireless LANs: 802.11n and 802.11ac," Cambridge, U.K.: Cambridge Univ. Press, June 2013.

[7] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," IEEE J. Sel. Areas Commun., vol. 18, no. 3, pp. 535-547, March 2000.

[8] S. Mosleh, Y. Ma, J. B. Coder, E. Perrins, and L. Liu, "Enhancing LAA co-existence using MIMO under imperfect sensing", *IEEE GLOBECOM Wkshps*, pp. 1 - 6, Waikoloa, HI, Dec. 2019.

[9] S. Mosleh, Y. Ma, J. D. Rezac, J. B. Coder, "Dynamic spectrum access with reinforcement learning for unlicensed access in 5G and beyond", accepted for publication in the *IEEE VTC Spring*, 2020.

[10] Y. Ma and D. G. Kuester, "MAC-layer coexistence analysis of LTE and WLAN systems via listen-before-talk," *IEEE CCNC*, pp. 534 - 541, Jan. 2017.

[11] Y. Ma, W. Young, E. Anderson, and J. Coder, "Probability of coexistence of LTE-LAA and WLAN systems based on delay constraints", *IEEE ICCCN*, pp. 1 - 9, Aug. 2018.

[12] M. Nielsen, "Neural Networks and Deep Learning," Determination Press, 2015.

[13] R. Klimberg, and B. D. McCullough, "Fundamentals of Predictive Analytics with JMP," SAS Institute, April 2012.

[14] "3GPP TSG RAN; Study on NR-based access to Unlicensed Spectrum; (Release 16)," 3GPP TR 38.889 V16.0.0, December 2018.