

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Traffic Learning: a Deep Learning Approach for Obtaining Accurate Statistical Information of the Channel Traffic in Spectrum Sharing Systems

OGEEN H. TOMA¹, (Student Member, IEEE), AND
MIGUEL LÓPEZ-BENÍTEZ^{1,2}, (Senior Member, IEEE)

¹Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, U.K

²ARIES Research Centre, Antonio de Nebrija University, 28040 Madrid, Spain

Corresponding author: Ogeen H. Toma (e-mail: ogeen.toma@liverpool.ac.uk).

ABSTRACT In recent works, the statistical information of the channel traffic has been increasingly exploited to make effective decisions in spectrum sharing systems. However, these statistics can not be obtained perfectly under (realistic) Imperfect Spectrum Sensing (ISS). Therefore, in this work we study comprehensively the approaches in the literature that correct the estimation of the channel traffic statistics under ISS, namely the closed-form expression approach and the algorithmic reconstruction approach. Then, we introduce a novel approach named Traffic Learning as a Deep Learning (DL) approach for providing accurate estimation of the channel traffic statistics under ISS. For this novel approach, deep neural networks using Multilayer Perceptron (MLP) models are found for the estimation of several statistical metrics. In addition, we show that utilising effective features from spectrum sensing observations can lead to a considerable improvement in statistics estimation for each, mean, variance, minimum and distribution of the channel traffic under ISS, outperforming the existing approaches in the literature, which are based on either closed-form expressions or reconstruction algorithms.

INDEX TERMS Spectrum sharing, dynamic spectrum access, cognitive radio, channel traffic statistics, spectrum sensing, machine learning, deep learning.

I. INTRODUCTION

THE advancement of Deep Learning (DL) in computer vision, speech recognition and natural language processing domains has inspired a large community of experts in the communications field to exploit the potential of this technology for solving a wide range of problems in communication systems. Such problems are either difficult to represent with tractable mathematical models or impractical to be solved by following the classical methods and algorithms. In this context, there has been an increasing interest in exploiting DL in wireless communications, in particular, Spectrum Sharing (SS) systems. This is due to the demonstrated improvements that DL has brought to several applications of SS such as spectrum management, spectrum sensing, spectrum prediction, network security and so on. These applications are crucial for the ongoing deployment of 5G technology, including but not limited to, 5G New Radio

Unlicensed (NR-U) [1], unlicensed LTE [2] and License Assisted Access (LAA) [3].

In recent works, the statistical information of the channel traffic has increasingly been exploited as input features to the neural network of DL models. These statistical features can make significant improvement in the performance of DL for solving particular problems in SS systems. For instance, in [4] traffic statistics (mean, variance and kurtosis) have been exploited as features for a neural network to recognise user-level applications such as YouTubeTM and WhatsAppTM. On the other hand, in [5] the accuracy of spectrum sensing in cognitive radio has remarkably been enhanced by exploiting traffic statistics as input features to a DL model used to sense the spectrum. Moreover, [6] has employed the historical samples of the channel traffic statistics to train a DL to predict the future channel occupancy ratio. Obtaining accurate statistical information of the channel traffic can also find a wide range

of applications in enhancing the performance of cooperative spectrum sensing systems operating under realistic environmental conditions [7]–[10] as well as in cognitive radio for Vehicular Ad Hoc Network (VANET) [11].

From the above discussion, it is evident that traffic statistics play an effective role in the performance of various applications in SS systems which apply DL technique. The majority of these works, however, assume perfect estimation of these statistics, such that they can smoothly be exploited in DL models. In practice, however, these statistics can be corrupted due to signal detection errors as discussed in [12]. Inaccurate traffic statistics, consequently, can worsen the training process of a DL model and thus provide inaccurate results. Therefore, in order to exploit channel traffic statistics in SS systems it is essential to estimate these statistics accurately especially under a realistic, imperfect spectrum sensing (ISS), scenario.

In the literature, the estimated traffic statistics under ISS have been corrected through two approaches: 1) reconstruction algorithms [13]–[15], where the observed idle/busy periods under ISS of the channel traffic are reconstructed to provide accurate statistics. 2) closed-form expressions [12], [16]–[18], where mathematical expressions are derived for the original statistics as a function of their corresponding statistics observed under ISS, probability of sensing error and sensing period. Reconstruction methods can provide some accuracy improvements but are typically based on heuristic algorithms and therefore sub-optimal. Although closed-form expressions would be the most attractive solution to correct the estimation of traffic statistics under ISS, it is challenging sometimes to find these expressions for higher statistical moments such as variance, skewness and kurtosis under ISS (whereas the mean, duty cycle and distribution have been found in [12]). In addition, although these expressions provide accurate estimations, they may still show some considerable estimation errors when a short sensing period T_s is employed [12, Section VIII]. In some cases, closed-form expressions are known or can be obtained but they are unable to lead to accurate estimations of the true traffic statistics under ISS, like for example the minimum period as analysed in [13]. In this work, therefore, we consider a DL approach to provide an accurate estimation of the channel traffic statistics under ISS and evaluate the performance of such approach with respect to the previous approaches, showing that the DL approach proposed in this work can provide significant performance improvements.

The contribution of this work can be highlighted as follows:

- We propose Traffic Learning (TL) as a DL approach to learn from the channel traffic under realistic ISS scenario in order to provide accurate statistical information about channel traffic activity in SS systems.
- Deep Neural Networks (NNs), namely Multilayer perceptron (MLP) models, are found to provide accurate estimation for the moments of the channel traffic statistics (mean, variance and minimum period) based on the

TABLE 1. List of acronyms and abbreviations

Acronym	Description
5G	Fifth Generation
CDF	Cumulative distribution function
DL	Deep Learning
E	Exponential
ED	Energy Detection
G	Gamma
GP	Generalised Pareto
ISS	Imperfect Spectrum Sensing
KS	Kolmogorov-Smirnov
LAA	License Assisted Access
LTE	Long Term Evolution
MAE	Mean Absolute Error
MLP	Multilayer Perceptron
MoM	Method of Moments
MSE	Mean Squared Error
NN	Neural Network
NR-U	New Radio Unlicensed
PSS	Perfect Spectrum Sensing
SNR	Signal-to-Noise Ratio
SS	Spectrum Sharing
TL	Traffic Learning
VANET	Vehicular Ad Hoc Network
W	Weibull

observations of ISS.

- The estimation of channel traffic distribution can then be achieved in two stages, first by classifying the distribution type of the actual channel traffic into one of a possible set of candidate distributions using a deep NN model, second finding the parameters of the classified distribution using the Method of Moments (MoM) technique [19] based on the previously obtained statistical parameters (mean, variance and minimum period).
- Finally, the performance of the proposed DL approach to estimate channel traffic statistics under ISS is validated numerically and compared with the previous approaches in the literature namely, algorithmic and closed-form expression methods. The obtained results show that the proposed approach outperforms the previous approaches with remarkable accuracy improvement, providing accurate channel traffic statistics under ISS.

The remainder of this work is organised as follows. First, Section II formulates the problem of channel traffic statistics estimation and introduces the system model considered in this work. Then Sections III and IV respectively discuss the algorithmic and closed-form expression approaches considered in the literature to correct the estimation of the channel traffic statistics. Section V proposes the novel DL approach for channel traffic statistics estimation under ISS. The NN models for the estimation of the traffic mean, variance and minimum period are discussed in Section VI, while the NN model for the classification and estimation of the traffic distribution is given in Section VII. The performance improvements of the proposed approach are demonstrated in Section VIII. Finally, Section IX concludes the paper. A list of acronyms and abbreviations used throughout the paper is given in Table 1.

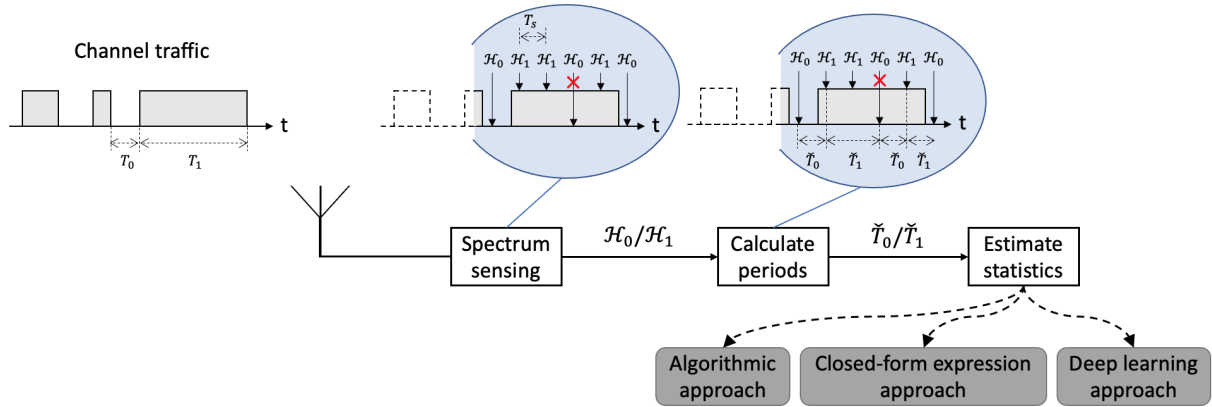


FIGURE 1. Channel traffic statistics estimation in spectrum sharing system.

Notation: Subscript i denotes the state of the channel to which a symbol belongs ($i = 0$ for idle and $i = 1$ for busy state). For periods of type i , T_i represents the period length; μ_i , m_i , and v_i represent their sample minimum, sample mean and sample variance, respectively; $F(T_i; \mu_i, \lambda_i, \alpha_i)$ denotes their cumulative distribution function with location parameter μ_i , scale parameter λ_i and shape parameter α_i . The true value of parameter is denoted as x_i and its accented versions \hat{x}_i , \check{x}_i , \tilde{x}_i represent the corresponding PSS observation, ISS observation and final estimation, respectively. $\mathbb{E}(\cdot)$ and $\mathbb{V}(\cdot)$ denote expected value and variance, respectively. The non-linear activation function of a NN is denoted by $\sigma(\cdot)$ and the loss function by $\mathcal{L}(\cdot, \cdot)$. \mathbb{R} represents the set of real numbers and $\|\cdot\|^2$ denotes the 2-norm.

II. PROBLEM FORMULATION AND SYSTEM MODEL

We consider the channel traffic in a particular frequency channel as shown in Fig. 1. This traffic is generated by the activity of the licensed users within their allocated frequency channel. Channel traffic can be represented as a sequence of idle/busy periods in the time-domain, hence, the duration of these periods can be modelled to follow a particular distribution. In the literature, and based on the practical measurements and observations, these periods are best described as Generalised Pareto (GP) distribution [20]. In this work, however, the distribution of channel traffic will be considered unknown to the SS system. In such system we assume to have a single unlicensed user, which monitors the activity of the channel traffic to find and exploit any opportunistic unoccupied duration in the frequency channel without causing harmful interference to the licensed users. This monitoring mechanism is achieved by performing periodic spectrum sensing at the unlicensed user. There have been significant research efforts in the last few years to develop high accuracy methods/algorithms for spectrum sensing, where the simplest and widely known method is Energy Detection (ED) [21]. Despite its variant forms and accuracy, spectrum sensing objective and output is the same, which is to provide binary decisions on the state of the channel, \mathcal{H}_0 for idle and \mathcal{H}_1 for

busy state. These sensing decisions can then be exploited to compute the durations of the idle T_0 and busy T_1 periods of the channel traffic which in turn are used to calculate channel traffic statistics.

Under high SNR conditions Perfect Spectrum Sensing (PSS) can be achieved. In practice, however, spectrum sensing is imperfect due to the presence of sensing errors caused by the wireless channel impairments and low SNR conditions, thus Imperfect Spectrum Sensing (ISS) is a more realistic scenario. Sensing errors occur as false alarms, when an idle state of the channel is sensed as busy, and missed-detections, when a busy state is sensed as an idle. These sensing errors can be represented as independent and identically distributed (i.i.d.) random variables with P_{fa} and P_{md} probabilities, respectively, which is a common modelling approach in the literature. Unfortunately under ISS, the presence of sensing errors corrupts the calculation of the idle/busy periods of the channel traffic such that they are observed as shorter fragments (\tilde{T}_0/\tilde{T}_1) of the original periods (T_0/T_1). These fragments, as a result, provide significantly corrupted channel traffic statistics.

As it was highlighted in the previous section, there are two approaches in the literature to correct the estimation of the channel traffic statistics under ISS, namely reconstruction algorithms [13]–[15] and closed-form expressions [12], [16]–[18]. The target of the first approach is to infer the position of potential sensing errors in the sequence of idle/busy periods observed under ISS and correct them in order to reconstruct the likely original sequence of idle/busy periods to provide accurate statistics. The second approach, on the other hand, derives mathematical expressions that can provide accurate estimation for the original traffic statistics from the ones observed under ISS. In this paper we propose a new approach based on Deep Learning (DL) to provide accurate statistical information of channel traffic under ISS, which will be compared with respect to the previous approaches. Therefore, we can illustrate these three approaches as shown in Fig. 1, which will be discussed in the next sections.

III. CLOSED-FORM EXPRESSION APPROACH

This approach formulates the problem of estimating a statistical parameter of the channel traffic under ISS in a closed-form expression, based on which a method can be derived to improve the accuracy of estimation. Consider the idle/busy periods T_i (where $i = 0$ for idle state and $i = 1$ for busy state) shown in Fig. 2. These periods are observed as \hat{T}_i under PSS and \check{T}_i under ISS. As it can be noticed, the observations under PSS (i.e., without sensing errors) provide a reasonable degree of accuracy for the original periods T_i (where the accuracy is only affected by the time resolution of the sensing period T_s [16]). On the other hand, the observed periods under ISS are significantly corrupted since sensing errors divide the observations of the original periods into shorter fragments \check{T}_i . As a result, the estimation of the channel traffic statistics based on the observed periods \check{T}_i under ISS is highly inaccurate with respect to the original statistics of T_i periods. The work in [12] formulated and provided closed-form expressions for some of the statistical parameters (e.g., mean, duty cycle and distribution) observed under ISS as a function of the original ones. For example, the mean $\mathbb{E}(\check{T}_i)$ of the observed periods under ISS is found in closed-form expression as a function of the original mean $\mathbb{E}(T_i)$, probabilities of sensing error P_{fa} and P_{md} , and sensing period T_s as [12]:

$$\mathbb{E}(\check{T}_i) = \frac{\mathbb{E}(T_i) - (-1)^i \mathbb{E}(T_0)P_{fa} + (-1)^i \mathbb{E}(T_1)P_{md}}{1 + \left(\frac{\mathbb{E}(T_0)}{T_s} - 2\right) \dot{P}_{fa} + \left(\frac{\mathbb{E}(T_1)}{T_s} - 2\right) \dot{P}_{md}} \quad (1)$$

where \dot{P}_{fa} and \dot{P}_{md} are defined in equations (13) and (14) of [12]. The original mean can then be estimated by solving (1) for $\mathbb{E}(T_i)$ as shown in (2). This method, in general, provides accurate estimation, however, some considerable error might still exist when short sensing period T_s is employed as explained in [12, Section VIII-B]. Therefore, we will use this method for comparison with other approaches in Section VIII. Note that the estimation of higher moments statistics (e.g., variance) under ISS is challenging to find in closed form expressions. Therefore, other approaches might be considered for such statistics.

IV. ALGORITHMIC RECONSTRUCTION APPROACH

In this approach reconstruction algorithm is used to correct the estimation of channel traffic statistics under ISS. Simple reconstruction algorithms were first proposed in [13] and then developed in [14], [15]. Therefore, we consider the latest reconstruction algorithm in the literature given by [15] and illustrated here in Algorithm 1. This algorithm reconstructs the periods in an iteration process and in each iteration the shortest periods will be reconstructed as $(\check{T}_{i,n-1} = \check{T}_{i,n-1} + \check{T}_{i,n} + \check{T}_{i,n+1})$, where n denotes the sequence of the periods, then the mean of the reconstructed periods will be calculated. This iteration will continue until the mean of the reconstructed periods reaches the value of the mean estimated using (2), i.e., this algorithm exploits the mean expression obtained from the previous approach as an indicator to determine when the periods are correctly reconstructed, however

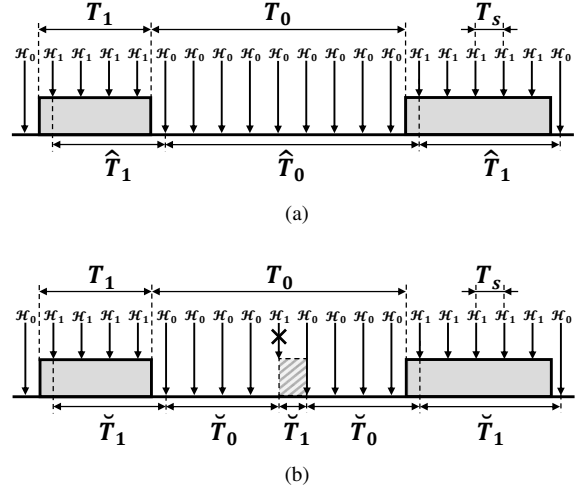


FIGURE 2. The observed idle/busy periods (a) under Perfect Spectrum Sensing (PSS), (b) under Imperfect Spectrum Sensing (ISS) [13].

Algorithm 1 Reconstruction algorithm [15]

Input: (\check{T}_i) The observed periods under ISS

Output: (\bar{T}_i) The reconstructed periods

- 1: Calculate the mean (\check{m}_i) of the periods under ISS
 - 2: Estimate the mean (m_i) of the periods using (2)
 - 3: $k = 0$
 - 4: $\bar{T}_i = \check{T}_i$
 - 5: **while** $\check{m}_i < m_i$ **do**
 - 6: $k = k + 1$
 - 7: **for each** $\check{T}_{i,n} = kT_s$ **do**
 - 8: $\bar{T}_{i,n-1} = \bar{T}_{i,n-1} + \check{T}_{i,n} + \check{T}_{i,n+1}$
 - 9: **end for**
 - 10: $\check{m}_i = \mathbb{E}(\bar{T}_i)$ Calculate the mean of the reconstructed periods
 - 11: **end while**
 - 12: **return** (\bar{T}_i)
-

once the process is finished, other statistics (not only the mean) can also be estimated. Therefore, this algorithm will be used to compare the performance of the estimation of channel traffic statistics under ISS with respect to other approaches in Section VIII.

V. DEEP LEARNING APPROACH

In this section we propose a novel approach for the estimation of the channel traffic statistics under ISS based on DL technique. The DL model in this work aims to provide an accurate estimation for the original statistical parameters of the channel traffic based on their corresponding (inaccurate) statistics observed under ISS. It is widely known that DL can solve various problems through formulating them as either classification or regression problems. The estimation of the statistical parameters mean, variance and minimum period is considered as a regression problem, while the estimation of the channel traffic distribution is solved by first classifying

$$\mathbb{E}(T_i) \approx \mathbb{E}(\tilde{T}_i) = \frac{\left(\mathbb{E}(\check{T}_i)(1 - P_{md}^{1-i} P_{fa}^i) - \mathbb{E}(\check{T}_{1-i})P_{md}^{1-i} P_{fa}^i\right) \left(1 - 2\dot{P}_{fa} - 2\dot{P}_{md}\right)}{\left(1 - P_{fa} - \frac{\dot{P}_{fa}}{T_s} \mathbb{E}(\check{T}_0)\right) \left(1 - P_{md} - \frac{\dot{P}_{md}}{T_s} \mathbb{E}(\check{T}_1)\right) - \left(\frac{\dot{P}_{fa}}{T_s} \mathbb{E}(\check{T}_1) - P_{fa}\right) \left(\frac{\dot{P}_{md}}{T_s} \mathbb{E}(\check{T}_0) - P_{md}\right)} \quad (2)$$

the type of the distribution, then finding its parameters. The estimation of these statistics can be solved using Multilayer Perceptron (MLP) fully-connected feedforward Neural Network (NN) [22].

An MLP with L (dense) layers maps the input layer x to the output layer y through one or more hidden layers in between. This mapping function can be written as $y = f(x; \theta)$, where θ denotes the NN parameters given by the weights W and basis b . Each layer of the NN consists of one or more neurons n , hence the output of the ℓ -th layer can be written as [23]:

$$f_\ell(x_{\ell-1}; \theta_\ell) = \sigma_\ell(W_\ell x_{\ell-1} + b_\ell), \quad \ell = 1, \dots, L \quad (3)$$

where $W_\ell \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ is the weight matrix, $b_\ell \in \mathbb{R}^{n_\ell}$ is the bias vector (note that n_ℓ denotes the number of neurons at the ℓ -th layer), and $\sigma_\ell(\cdot)$ represents the non-linear activation function which can be given by, e.g., ReLU, sigmoid, softmax, etc. The output of the ℓ -th layer $f_\ell(x_{\ell-1}; \theta_\ell)$ is based on the input $x_{\ell-1}$ from the previous layer and the parameter $\theta_\ell = \{W_\ell, b_\ell\}$ at the ℓ -th layer. In general, a NN is trained based on a labelled training dataset, which is an input-output (x, y) vector pairs of data. In our scenario, the input vector is the observations of a statistical parameter \check{s} under ISS (e.g., mean, variance, etc.) and the output vector is the corresponding original statistical parameter s . Therefore, this input-output (\check{s}, s) dataset is used to train a NN to find θ^* that minimises the loss function $\mathcal{L}(\check{s}, s)$:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \mathcal{L}(\check{s}, s) \quad (4)$$

For example, Mean Squared Error (MSE) loss function can be used as $\|s - f(\check{s}; \theta)\|^2$ to find θ that minimises the error. By selecting the appropriate hyper-parameters of the NN (e.g., number of layers, neurons, loss function) along with the useful input features, a DL model can be achieved to provide an accurate estimation for the statistical parameters of the channel traffic under ISS as it will be discussed next.

VI. MEAN, VARIANCE AND MINIMUM ESTIMATION BASED ON DL

Let us first consider the estimation of the original mean m_i of the idle/busy periods (where i can be 0 referring to idle periods, or 1 referring to busy periods). A DL model using MLP NN is built to find the accurate estimation of the mean of the channel traffic from the corresponding mean observed under ISS. Therefore, the inaccurate means \check{m}_0 and \check{m}_1 of the idle/busy periods observed under ISS are used as inputs to the DL model to provide the accurate estimation of the mean period \tilde{m}_i (where $\tilde{m}_i \approx m_i$). Since under ISS the presence of sensing errors corrupts the observation of the

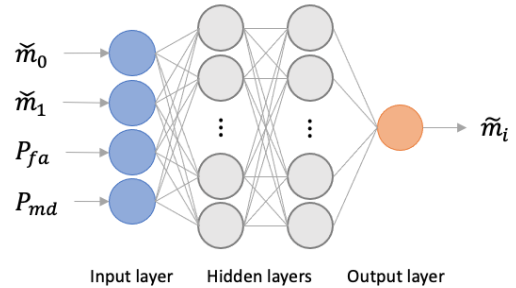


FIGURE 3. Deep Learning model for mean of periods estimation under ISS.

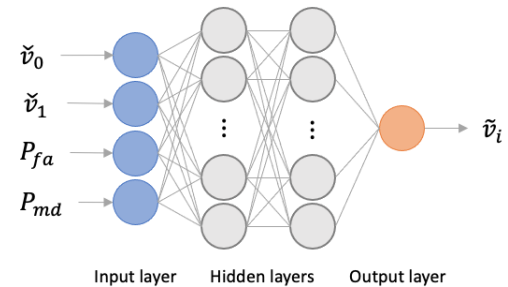


FIGURE 4. Deep Learning model for variance of periods estimation under ISS.

idle/busy periods as discussed in Section II, the mean of these periods would be significantly inaccurate depending on the probabilities of sensing error (i.e., P_{fa} and P_{md}). These probabilities can be pre-defined based on the employed sensing algorithm at the end terminal [12]. Therefore, P_{fa} and P_{md} can also be exploited as input features to the DL model along with \check{m}_0 and \check{m}_1 observed under ISS. P_{fa} and P_{md} can assist a NN to learn from how these features affect the observation of \check{m}_0 and \check{m}_1 under ISS, which in turn will help predicting the actual mean value at the output as shown in Fig 3. Note that when $P_{fa} = P_{md} = 0$, the observed mean will be equal to the original one [16]. A similar concept can also be applied to find a DL model for estimating higher statistical moments under ISS. In this work, we consider the second moment (variance v_i) of the idle/busy periods, which can similarly be found as shown in Fig 4. As it can be noticed, the observed statistics of both idle and busy periods are always considered as input features because they both are affected by false alarms and missed detections as it can be observed from (2) and therefore considering only the observed statistics for the same type of periods being estimated (idle or busy) would not provide complete input information.

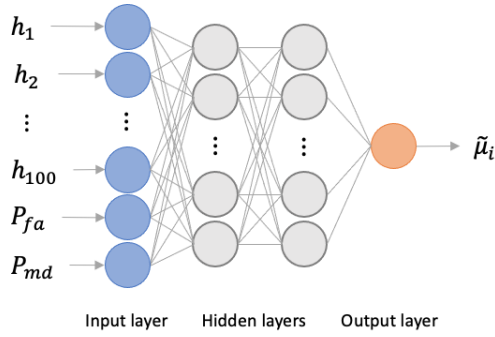


FIGURE 5. Deep Learning model for minimum of periods estimation under ISS.

On the other hand, the accurate estimation of the minimum period μ_i of the channel traffic under ISS is more challenging to find compared to the previous statistical parameters. This is because for any non-zero probability of sensing error ($P_{fa} > 0$ and $P_{md} > 0$) the observed minimum period $\tilde{\mu}_i$ under ISS is always equal to the duration of a single sensing error, which is same as the duration of the sensing period T_s (i.e., $\tilde{\mu}_i \neq \mu_i$ and $\tilde{\mu}_i = T_s, \forall P_{fa}, P_{md} > 0$) [13]. Therefore, a NN can not learn anything from the observed minimum idle/busy periods $\check{\mu}_0/\check{\mu}_1$ under ISS (unlike the previous statistical parameters) since they are always equal to the sensing period T_s , no matter how high or low the probability of sensing error is. In order to utilise a feature that can help a NN to predict the actual minimum period μ_i from the observations of the ISS, it is useful to look at the distribution of the observed periods under ISS. The observed periods under ISS have a discrete distribution with a bin size of T_s and starting at T_s as well. This distribution is distorted by the presence of sensing errors, however, it forms a distinguished pattern corresponding to a particular combination of probabilities of sensing error (P_{fa} and P_{md}). A NN can be trained to learn from these patterns of the observed distributions under ISS in order to locate the actual minimum period. As a result, it is found that by using the first h -th histogram bins of the observed periods under ISS along with the probabilities of sensing error (P_{fa} and P_{md}) it is possible to train a NN to provide an accurate estimation for the actual minimum period under ISS. The MLP NN in Fig. 5 shows an example of using 100 histogram bins of the observed periods under ISS as input features along with P_{fa} and P_{md} , where h_1 refers to the number of the observed periods under ISS within the first bin, while h_2 refers to the number of the observed periods under ISS within the second bin and so on. The number of bins was selected here after conducting several evaluations on the estimation accuracy of the minimum period under ISS while considering several scenarios of probabilities of sensing errors (P_{fa} and P_{md}), for which 100 bins were found to be sufficient to provide accurate results under any scenario of sensing errors. The output of this NN provides the accurate estimation $\tilde{\mu}_i$ for the actual minimum period μ_i (where $\tilde{\mu}_i \approx \mu_i$).

A. RAW DATASET CONSTRUCTION AND PREPROCESSING

In this work, data are obtained and prepared in two stages, in the first stage raw datasets are generated using MATLAB, then in the second stage the generated datasets are preprocessed using Python to train, validate and test the proposed DL model. Dataset generation using MATLAB can be achieved as follows:

- 1) First, a channel traffic is modelled by generating a large sequence of idle/busy periods (T_0/T_1) in a frequency channel drawn from a particular distribution such as GP distribution (which is one of the best representations of the channel traffic [20]).
- 2) Then spectrum sensing can be applied with periodic sensing period T_s , where T_s should be smaller than the minimum period of the channel idle/busy periods (i.e., $T_s < \mu_i$). In this work we consider to use a short $T_s = 1$ t.u. (time unit) when the minimum period $\mu_i = 10$ t.u. (i.e., 10% of the minimum period). This is to show how the estimation methods perform under the worse scenario of using such short sensing period since higher sensing periods (e.g., 90%) can provide more accurate estimations for traffic statistics under ISS [13].
- 3) Spectrum sensing is configured based on the selected probabilities of sensing error (i.e., P_{fa} and P_{md}), based on which a sensing threshold is adjusted to decide whether the channel is idle \mathcal{H}_0 or busy \mathcal{H}_1 . Sensing decisions are then used to calculate the duration of the idle/busy periods (\check{T}_0/\check{T}_1) observed under ISS.
- 4) The statistical parameters such as mean \check{m}_0/\check{m}_1 , variance \check{v}_0/\check{v}_1 or histogram $\{h_1, \dots, h_{100}\}$ can then be calculated from (\check{T}_0/\check{T}_1) periods observed under ISS in step 3. These statistics are saved into a `.mat` file along with the configured P_{fa} and P_{md} to represent the input vector (features). On the other hand, the corresponding original statistics m_0/m_1 for mean, v_0/v_1 for variance or μ_0/μ_1 for minimum of the idle/busy periods (T_0/T_1) generated in step 1 are also saved into the same `.mat` file to represent the output vector (labels).

The obtained features and labels in `.mat` file are then used to construct the required dataset for DL, 60% of which is for training, 20% is for validation and the remaining 20% is for testing as shown in Figs. 6 and 7. These raw datasets require some preprocessing before using them for DL training or testing. Python is used here, which offers numerous tools and advanced DL libraries (e.g., TensorFlow [24], Keras [25] and PyTorch [26]) that facilitate not only the preprocessing of the datasets, but also building, training and testing of the DL model. Therefore, the obtained dataset in `.mat` file is imported to Python for preprocessing, where the features and labels are extracted and stored into separate arrays. Since these data can hold any real values, it is a common practice to scale and normalise these values before learning from them. The `preprocessing.Normalization()` function from Keras library is used, which normalises its inputs

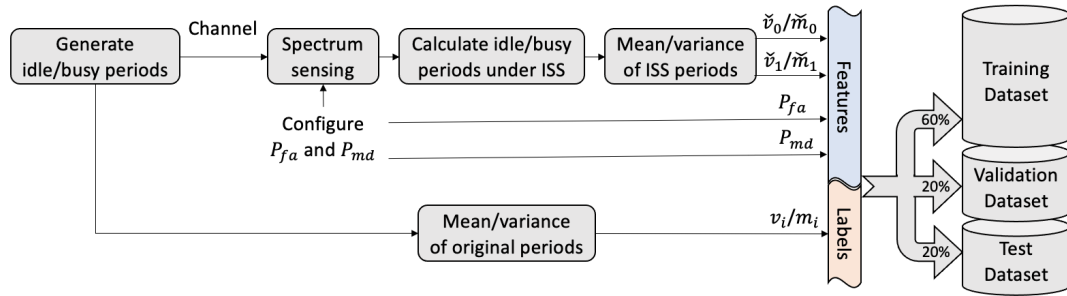


FIGURE 6. Datasets construction of DL for channel traffic mean and variance estimation.

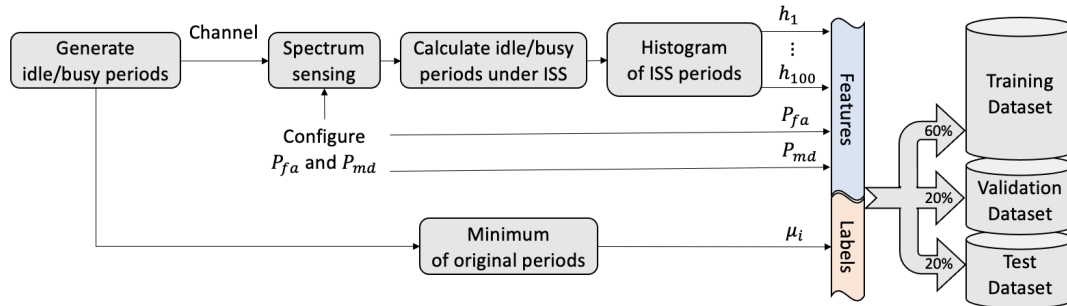
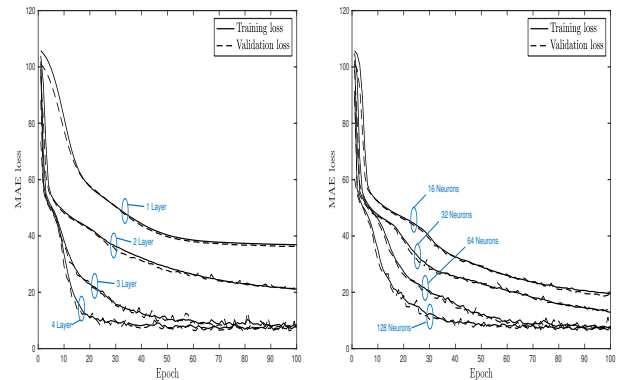


FIGURE 7. Datasets construction of DL for channel traffic minimum estimation.

into a distribution centred around zero with unit standard deviation. This is accomplished by applying the following normalisation relationship $(input - mean)/\sqrt{variance}$ to the input dataset.

B. TRAINING, VALIDATING AND TESTING THE DL MODEL

After preprocessing the datasets, they are ready to train, validate and test a DL model. An MLP NN has been examined using several hyper-parameter settings to build the required DL model for channel traffic statistics estimation under ISS. As shown in Fig. 8, different number of hidden layers {1, 2, 3 and 4} and neurons {16, 32, 64 and 128} are used to examine the accuracy of training based on Mean Absolute Error (MAE) loss function. It is found that a NN with 3 hidden layers can reach the same accuracy as a higher number of layers after 100 epochs of training. In the same way, 64 neurons per hidden layer can provide the same accuracy as a higher number of neurons after 100 epochs of training. As a result, the MLP NN shown in Table 2 is considered in this work to provide the accurate estimation of the channel traffic statistics under ISS. The output of this model would be either the accurate estimation of the mean \tilde{m}_i , variance \tilde{v}_i or minimum period $\tilde{\mu}_i$ when the input is the corresponding ISS mean \check{m}_i , variance \check{v}_i , or histogram bins $\{h_1, \dots, h_{100}\}$, respectively. This MLP NN model is trained based on the 60% of the preprocessed features and labels, while 20% of which is used to validate the training process. This validation is important to make sure that the NN can generalise to new data and avoid the overfitting problem. ReLU activation function is selected at each hidden layer, and Adam optimiser



(a) Using different number of hidden layers (64 neurons each). (b) Using different number of neurons per layer (for 3 hidden layers).

FIGURE 8. Training and validation accuracy of the MLP NN for the estimation of the mean, variance and minimum period.

TABLE 2. MLP NN model used for mean, variance and minimum period estimation under ISS.

Hyper-parameter	Settings
Number of hidden layers	3
Neurons per hidden layer	64
Activation function	ReLU
Optimiser	Adam
Learning rate	0.001
Loss function	MAE
Metric	Accuracy
Batch size	10
Epochs	100

is used with learning rate 0.001. After training and validating the DL model, it can now be tested based on the remaining 20% of the dataset to evaluate its estimation performance. Although the testing dataset has both features and labels, only features are fed to the NN to predict the accurate channel traffic statistics, while labels are used to quantify the accuracy of the estimation provided by the NN, which will be shown in the simulation results.

VII. DISTRIBUTION CLASSIFICATION AND ESTIMATION BASED ON DL

Having an accurate estimation for the distribution of the idle/busy period durations completes the whole picture of learning about the channel traffic activity (i.e., traffic learning). In the literature, different distribution models have been considered for the channel traffic. Exponential (E) distribution, for example, is one of the widely assumed models for channel traffic as in [27]–[29], which can simplify the mathematical analysis of the studies. However, field measurements in [20] have shown that the Generalised Pareto (GP) distribution is more realistic for channel traffic representation. In this work, however, we investigate the estimation of the channel traffic distribution under ISS using a DL approach without making any prior assumption about the original distribution type of the channel traffic. In addition, we compare this approach with previous methods for estimating the distribution under ISS. First, a DL model is used to classify the distribution type of the channel traffic based on the ISS observations. After classifying the distribution type, Method of Moments (MoM) inference technique [19] can then be used to estimate the distribution parameters (location μ , scale λ and shape α , if they all exist) from the sample moments obtained previously (i.e., mean, variance and minimum).

The classification problem can be solved using an MLP NN that selects a distribution class at the output based on the observations of the ISS for the channel traffic. Table 3 is considered for the list of the possible traffic distribution types that provides accurate representations for the empirical data [20], from which a NN can select the best match type for the channel traffic distribution. This list includes Exponential (E), Generalised Pareto (GP), Gamma (G) and Weibull (W) distributions (note that other distribution types can also be added to the list). Therefore, there is no particular type assumption for the channel traffic distribution (as often is assumed in the literature) since the list here can easily be extended to other distribution models. The input of the NN, as shown in Fig. 9, uses the first h -th histogram bins of the observed periods under ISS along with the probabilities of sensing error (P_{fa} and P_{md}) to predict the best classification for their distribution (the highest probability at the output). Note that the input of this NN is similar to the input of the NN used to find the minimum parameter μ in the previous section, however, the input here is used to solve a classification problem rather than a regression problem and as a result the NN has multiple outputs.

After classifying the distribution type of the channel traf-

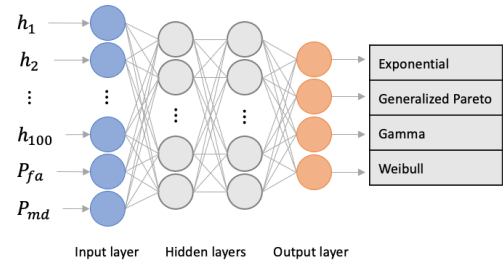


FIGURE 9. Deep Learning model for distribution type classification under ISS.

fic, MoM inference technique [19] is considered to estimate the distribution parameters (location μ_i , scale λ_i and shape α_i , if they all exist) from the sample moments obtained previously (i.e., mean, variance and minimum). The location parameter μ_i is the same as the minimum period estimated previously as $\tilde{\mu}_i$ using DL approach, while the scale λ_i and shape α_i parameters can be found from the mean and variance of the selected distribution model. Since the moments (mean and variance) can also be estimated accurately using the DL approach as discussed before, the scale λ_i and shape α_i parameters of the selected distribution can therefore be solved using MoM technique. For example, if the DL model shown in Fig. 9 classifies (with highest probability) the channel traffic observations as GP-distributed, their μ_i , λ_i and α_i parameters can then be found as [19, ch. 20]:

$$\mu_i \approx \tilde{\mu}_i \tag{5a}$$

$$\lambda_i \approx \tilde{\lambda}_i = \frac{1}{2} \left(1 + \frac{(\tilde{m}_i - \tilde{\mu}_i)^2}{\tilde{v}_i} \right) (\tilde{m}_i - \tilde{\mu}_i) \tag{5b}$$

$$\alpha_i \approx \tilde{\alpha}_i = \frac{1}{2} \left(1 - \frac{(\tilde{m}_i - \tilde{\mu}_i)^2}{\tilde{v}_i} \right) \tag{5c}$$

where $\tilde{\mu}_i$, \tilde{m}_i and \tilde{v}_i are the estimated minimum, mean and variance of the channel traffic using DL approach, respectively. Once the distribution parameters are found, the Cumulative Distribution Function (CDF) of the GP distribution F_{GP} can then be obtained from:

$$F_{GP} \approx \tilde{F}_{GP}(T_i; \tilde{\mu}_i, \tilde{\lambda}_i, \tilde{\alpha}_i) = 1 - \left[1 + \frac{\tilde{\alpha}_i(T_i - \tilde{\mu}_i)}{\tilde{\lambda}_i} \right]^{-1/\tilde{\alpha}_i} \tag{6}$$

In the same way we can find the expressions for other channel traffic distributions.

A. RAW DATASET CONSTRUCTION AND PREPROCESSING

As discussed before, distribution estimation is achieved by first classifying the distribution type using DL model, then estimating the distribution function using MoM technique. To solve the classification problem, datasets are required to be obtained. These datasets are constructed in the same way as step 1 to 4 in Section VI-A with some slight differences. In step 1, channel traffic is modelled 4 times using (E, GP, G and W) distributions. Then spectrum sensing and probability

TABLE 3. Considered probability distribution models for idle/busy period durations. Distribution names: E (Exponential), GP (Generalised Pareto), G (Gamma), and W (Weibull). Distribution parameters: μ_i (location), λ_i (scale), and α_i (shape). T_i represents the period length. $\mathbb{E}\{\cdot\}$ and $\mathbb{V}\{\cdot\}$ represent the mean and the variance of the distribution, respectively. $\gamma(\cdot, \cdot)$ is the lower incomplete Gamma function [30, 6.5.2] and $\Gamma(\cdot)$ is the (complete) Gamma function [30, 6.1.1]. (reproduced from [20]).

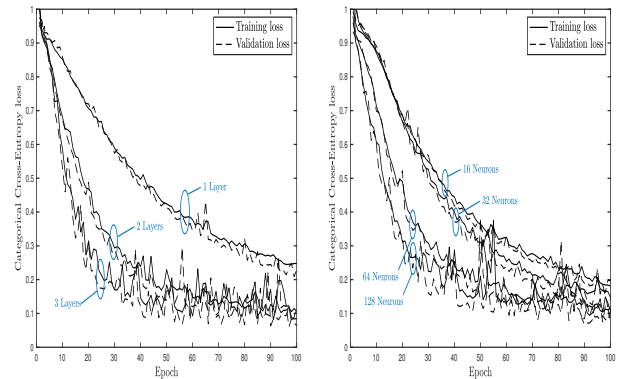
Distribution function	Parameters	Moments
$F_E(T_i; \mu_i, \lambda_i) = 1 - e^{-\lambda_i(T_i - \mu_i)}$	$T_i \geq \mu_i > 0$ $\lambda_i > 0$	$\mathbb{E}\{T_i\} = \mu_i + \frac{1}{\lambda_i}$ $\mathbb{V}\{T_i\} = \frac{1}{\lambda_i^2}$
$F_{GP}(T_i; \mu_i, \lambda_i, \alpha_i) = 1 - \left[1 + \frac{\alpha_i(T_i - \mu_i)}{\lambda_i}\right]^{-1/\alpha_i}$	$T_i \geq \mu_i$ ($\alpha_i \geq 0$) $T_i \in [\mu_i, \mu_i - \frac{\lambda_i}{\alpha_i}]$ ($\alpha_i < 0$) $\mu_i, \lambda_i > 0, \alpha_i < 1/2$	$\mathbb{E}\{T_i\} = \mu_i + \frac{\lambda_i}{1 - \alpha_i}$ $\mathbb{V}\{T_i\} = \frac{\lambda_i^2}{(1 - \alpha_i)^2(1 - 2\alpha_i)}$
$F_G(T_i; \mu_i, \lambda_i, \alpha_i) = \frac{\gamma(\alpha_i, \frac{T_i - \mu_i}{\lambda_i})}{\Gamma(\alpha_i)}$	$T_i \geq \mu_i > 0$ $\lambda_i > 0$ $\alpha_i > 0$	$\mathbb{E}\{T_i\} = \mu_i + \lambda_i \alpha_i$ $\mathbb{V}\{T_i\} = \lambda_i^2 \alpha_i$
$F_W(T_i; \mu_i, \lambda_i, \alpha_i) = 1 - \exp\left[-\left(\frac{T_i - \mu_i}{\lambda_i}\right)^{\alpha_i}\right]$	$T_i \geq \mu_i > 0$ $\lambda_i > 0$ $\alpha_i > 0$	$\mathbb{E}\{T_i\} = \mu_i + \lambda_i \Gamma\left(1 + \frac{1}{\alpha_i}\right)$ $\mathbb{V}\{T_i\} = \lambda_i^2 \left[\Gamma\left(1 + \frac{2}{\alpha_i}\right) - \Gamma^2\left(1 + \frac{1}{\alpha_i}\right)\right]$

of sensing error (P_{fa} and P_{md}) are configured in the same way as in step 2 and 3. In step 4, channel traffic statistics (histogram bins $\{h_1, \dots, h_{100}\}$) are computed from the ISS observations. These observations along with the configured P_{fa} and P_{md} represent the input vector (features) of the DL model, whereas the output vector (labels) is given by the classes of the original distribution used to model the channel traffic in step 1. Since we have 4 distribution classes (E, GP, G and W), they can be encoded as a *one-hot* vector $1_s \in \mathbb{R}^4$ (i.e., 4-dimensional vector, the s -th element of which is equal to one and zero otherwise [23]). These features and labels can then be saved into `.mat` file to be used later for training and testing.

However, preprocessing is required to be performed first on the produced dataset. Therefore, the obtained dataset in `.mat` file is imported to Python for preprocessing. Similar to section VI-A, `preprocessing.Normalization()` function from Keras library is used to normalise these datasets in order to be used for training and testing.

B. TRAINING, VALIDATING AND TESTING THE DL MODEL

After preprocessing the dataset, it can now be used to train, validate and test a DL model. An MLP NN with several settings has been examined to build the required DL model for classifying channel traffic distribution under ISS. As shown in Fig. 10, different number of hidden layers {1, 2, and 3} and neurons {16, 32, 64 and 128} are used to examine the accuracy of training based on Categorical Cross-Entropy loss function. It is found that a NN with 2 hidden layers can reach the same accuracy as a higher number of layers when 100 Epochs is used. In the same way, 64 neurons per hidden layer can provide the same accuracy as a higher number of neurons when 100 Epochs are used. As a result, an MLP NN shown in Table 4 is considered to provide accurate classification for the type of the channel traffic distribution under ISS. The output layer of this model has 4 neurons referring to the corresponding classes (E, GP, G and W). Therefore, by using Softmax activation function at this layer,



(a) Using different number of hidden layers (16 neurons per each). (b) Using different number of neurons per layer (for 2 hidden layers).

FIGURE 10. Training and validation accuracy of the MLP NN for distribution classification.

TABLE 4. MLP NN model used for distribution classification under ISS

Hyper-parameter	Settings
Number of hidden layers	2
Neurons per hidden layer	64
Activation function	ReLU, Softmax (output)
Optimiser	Adam
Learning rate	0.001
Loss function	Categorical Cross-Entropy
Metric	Accuracy
Batch size	10
Epochs	100

the output of these 4 neurons will represent a probability of the corresponding distribution class. Hence, the output with the highest probability will indicate the best distribution class match for the observed channel traffic under ISS. After preprocessing the features and labels in the `.mat` file, 60% of these data is used to train this MLP NN model, while 20% is used to validate the training process. After training and validating the DL model, it can now be tested based on the remaining 20% of the dataset to evaluate its classification

performance. Although testing dataset has both features and labels, only features are fed to the NN to classify channel traffic distribution, while labels are used to quantify the accuracy of the classification provided by the NN, which will be shown in the simulation results.

VIII. RESULTS

A. MEAN, VARIANCE AND MINIMUM PERIOD ESTIMATION OF THE CHANNEL TRAFFIC UNDER ISS

In order to evaluate the estimation performance of the DL model proposed in Section VI to estimate the mean, variance and minimum period of the channel traffic under ISS, a large dataset is produced to train the DL model such that it can generalise a problem, i.e., to provide accurate estimation for the channel traffic statistics even when new data are observed under ISS. This can be achieved by repeating steps 1 to 4 in Section VI-A several times to remodel the original channel traffic to cover a wide variety of traffic statistics, and for each traffic model spectrum sensing is applied and configured in step 3 using different combinations of P_{fa} and P_{md} ranging from low (0.01) to high (0.1) probability of error. In the estimation of mean, for example, channel traffic in step 1 can be modelled repeatedly to have random mean values as $m_i \sim \mathcal{U}(10, 200)$ t.u., and for each traffic mean spectrum sensing is applied using several combinations of $P_{fa} \sim \mathcal{U}(0.01, 0.1)$ and $P_{md} \sim \mathcal{U}(0.01, 0.1)$ to observe the original mean under different scenarios of ISS. Similar procedures can also be followed to obtain the datasets for variance and minimum period statistics.

Then 60% and 20% of such datasets are used to train and validate the DL model, respectively, as discussed in Section VI-B, while the remaining 20% of the dataset is used to test the accuracy of the DL model. Fig. 11 shows the accuracy of estimating the mean of the channel traffic under ISS using different approaches (closed-form expression, reconstruction algorithm and DL). Each point in the figure represents the corrected estimation of the traffic mean observed under ISS for a particular P_{fa} and $P_{md} \sim \mathcal{U}(0.01, 0.1)$. As it can be noticed, DL approach outperforms the previous approaches for providing accurate estimation, in which all the points are distributed closely around the straight line that corresponds to the original mean value. It is worth mentioning that, the selected reconstruction algorithm in this work performs better than the closed-form expression because the algorithm itself exploits the closed-form expression to improve the estimation of the mean. It can also be noticed that, as the mean value increases the estimation performance degrades for all approaches. This is due to the fact that the longer the periods the higher the number of sensing errors occur within those periods, thus less accurate estimation can be achieved. In Fig. 12 and 13, on the other hand, the DL approach also provides higher accuracy for the estimation of the variance and minimum period, respectively. Variance estimation in Fig. 12 is only provided for DL and reconstruction approaches since, to the best of the authors' knowledge, no closed-form expression for such moment under ISS is

TABLE 5. Average error for statistics estimation using different approaches.

Approach	Mean (t.u.)	Variance (t.u.) ²	Minimum (t.u.)
Closed-form expression	10.83	–	106.75
Reconstruction algorithm	7.95	60.56	86.41
Deep Learning	5.22	10.33	2.42

available in the literature. In Fig. 13, on the other hand, even when a closed-form expression is provided for the estimation of the minimum period under ISS (which is simply given by $\tilde{\mu} = T_s$ [13]), it does not lead to accurate estimation of the true minimum period. Similarly, the reconstruction method also fails to provide accurate estimation for the minimum period under ISS, this is because even after reconstructing the corrupted idle/busy periods under ISS there will be still some short periods which have not been reconstructed properly, thus providing incorrect minimum period estimation. The distribution of estimation error for all approaches is also provided (in the middle plots), where it shows better performance for DL estimator as it is centred around zero with narrow standard deviation with respect to other approaches. This performance improvement can also be observed in the right hand side plots in terms of the Maximum Absolute Error (MAE) obtained within a 90% confidence interval. The performance shown in Figs. 11(a), 12(a) and 13(a) can also be presented in numerical form as shown in Table 5 by taking the average of the differences between the original values of these statistics and their estimations under ISS, for which it can be noticed that our proposed approach also, in average, provides less error in the estimation of the original statistics with respect to the previous approaches.

B. DISTRIBUTION CLASSIFICATION AND ESTIMATION OF THE CHANNEL TRAFFIC UNDER ISS

As discussed in Section VII, channel traffic distribution is estimated in two stages, first classifying the distribution type, second estimating the distribution parameters. To evaluate the performance of the DL model used to classify the distribution of the channel traffic, a large dataset of 4×10^5 histograms using 100 bins is produced by remodelling the channel traffic several times using (E, GP, G and W) distribution models. The corresponding observations of the channel traffic under ISS using random P_{fa} and $P_{md} \sim \mathcal{U}(0.01, 0.1)$ are obtained. Similar to the previous section, 60% and 20% of such dataset are used to train and validate the DL model, respectively, while the remaining 20% of the dataset is used to test the accuracy of classification. Fig. 14 shows the accuracy of classifying the distribution of the observed channel traffic under low (0.01) and high (0.1) probability of sensing error. Sensing errors can distort the shape of the observed traffic distribution. However, as it can be seen from the confusion matrix, even under high probability of sensing error the proposed DL model can still provide accurate classification for the observed channel traffic under ISS. To estimate the distribution parameters (μ_i , λ_i and α_i), MoM method can

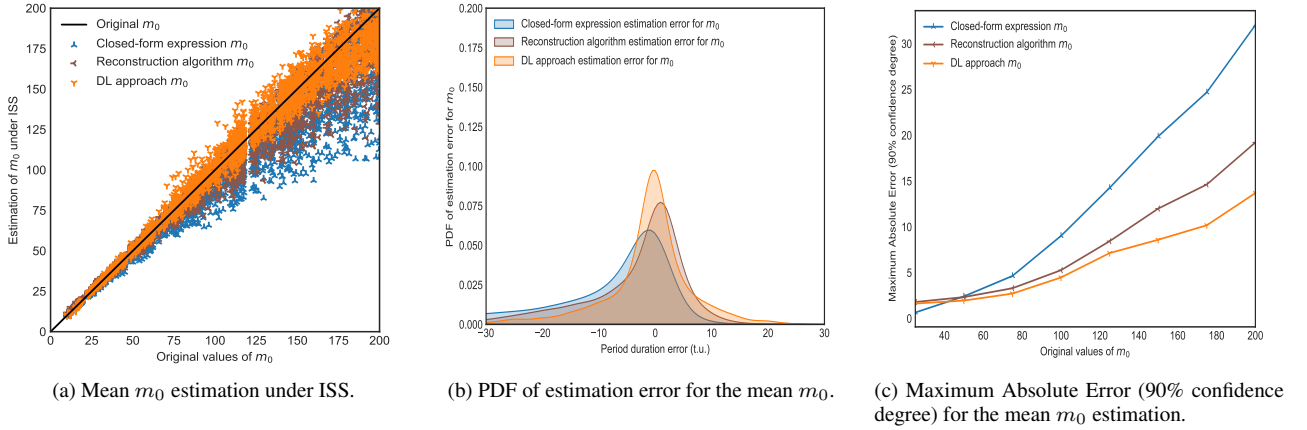


FIGURE 11. Simulation results for traffic mean estimation under ISS using different approaches, when P_{fa} and $P_{md} \sim \mathcal{U}(0.01, 0.1)$, $T_s = 1$ t.u..

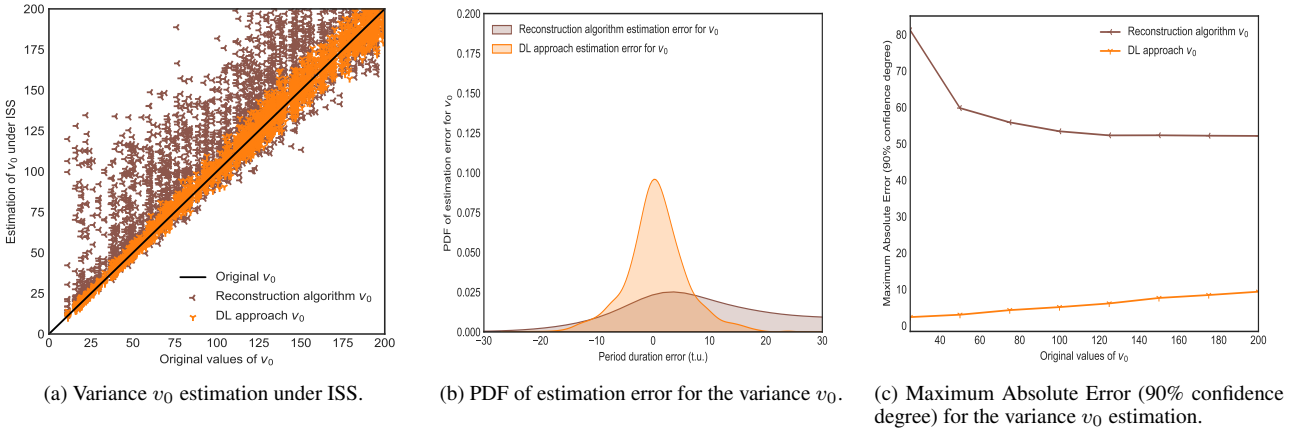


FIGURE 12. Simulation results for traffic variance estimation under ISS using different approaches, when P_{fa} and $P_{md} \sim \mathcal{U}(0.01, 0.1)$, $T_s = 1$ t.u..

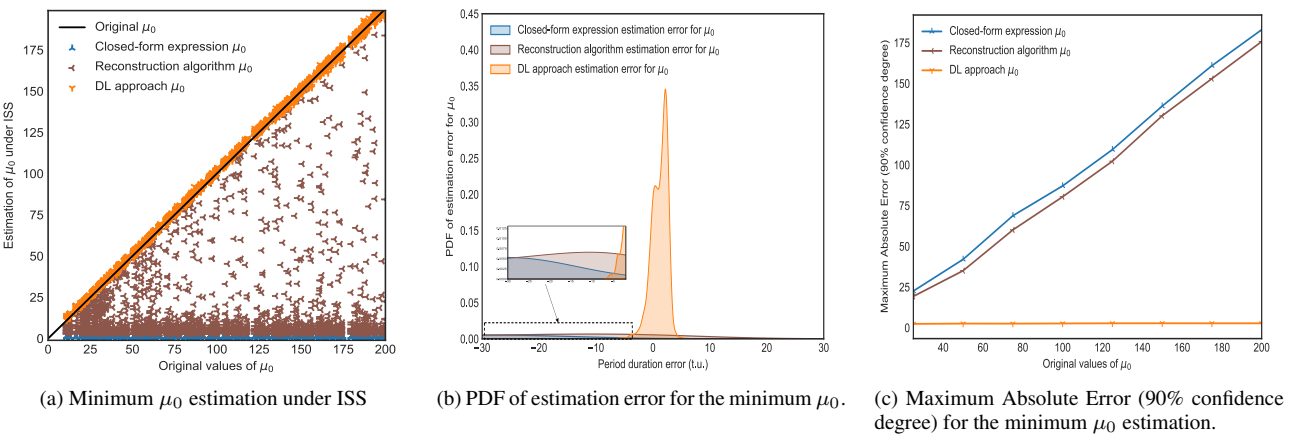


FIGURE 13. Simulation results for traffic minimum estimation under ISS using different approaches, when P_{fa} and $P_{md} \sim \mathcal{U}(0.01, 0.1)$, $T_s = 1$ t.u..

be applied according to the selected distribution type. Since the mean, variance and minimum period can be estimated accurately using DL approach as seen from the previous section results, accurate estimation can also be obtained for

$(\tilde{\mu}_i, \tilde{\lambda}_i$ and $\tilde{\alpha}_i)$, based on which the CDF of the channel traffic $\tilde{F}(T_i)$ can then be found as explained in Section VII. The accuracy of this estimation can be presented in terms of Kolmogorov-Smirnov (KS) distance [31], which is defined

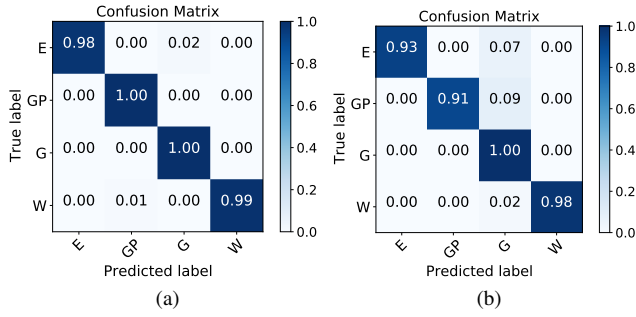


FIGURE 14. Channel traffic distribution classification under ISS when (a) $P_{fa} = 0.01$ and $P_{md} = 0.01$, (b) $P_{fa} = 0.1$ and $P_{md} = 0.1$.

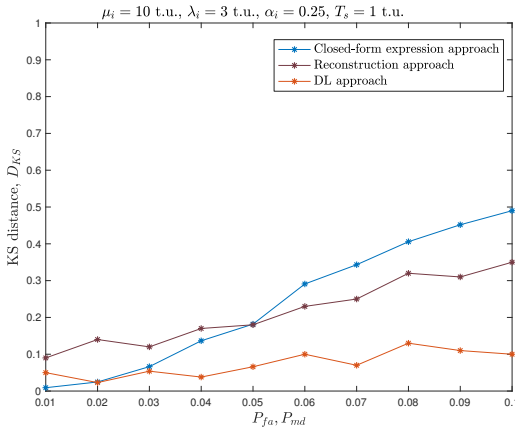


FIGURE 15. KS distance of the channel traffic CDF estimation under ISS.

as the maximum absolute difference between the estimated CDF $\tilde{F}(T_i)$ and the original CDF $F(T_i)$ of the channel traffic as:

$$D_{KS} = \sup_{T_i} |\tilde{F}(T_i) - F(T_i)| \quad (7)$$

where D_{KS} is the KS distance between the estimated distribution and the original one. Therefore, based on (7), the accuracy of estimating the distribution of the channel traffic under ISS is shown in Fig. 15 using DL, reconstruction algorithm and closed-form expression [12, eq. (45)] approaches when the original traffic distribution is drawn from GP with $\mu_i = 10$ t.u., $\lambda_i = 3$ t.u. and $\alpha_i = 0.25$ parameters. As it can be appreciated, the proposed DL approach achieves lower KS distance (i.e., higher accuracy of estimation) than the previous approaches for different values of P_{fa} and P_{md} . Since the estimation of the traffic distribution using DL approach is dependent on the estimations of the mean, variance and minimum period, its accuracy changes according to the accuracy of estimating those moments, which are also obtained using DL approach for the given P_{fa} and P_{md} . Similar observations can be obtained as well for the estimation of other types of distributions, showing significant improvement in the distribution estimation through using the proposed DL approach.

C. COMPUTATIONAL COMPLEXITY

The computational complexity of the different approaches used to estimate channel traffic statistics under ISS is an important aspect to investigate. Generally, closed-form expressions approach tends to be more attractive in terms of the complexity as it provides accurate estimations for the channel traffic statistics under ISS through using explicit mathematical equations. However, the accuracy of these equations tends to degrade as the sensing period T_s decreases, this is because decreasing the latter causes an increase in the number of the sensing events within an observed period, which in turn increases the occurrence of sensing errors as a result. In addition, regardless of being more attractive, closed-form expressions can be challenging sometimes to find for higher statistical parameters under ISS such as variance, skewness and kurtosis (where this can be noticed from the results of Fig. 12, the absence of the closed-form expression approach for variance estimation). The reconstruction algorithms approach, on the other hand, is less attractive in terms of the complexity as it performs heavily computational operations with several iterations in their algorithms in order to reconstruct the idle/busy periods corrupted by the sensing errors. In the reconstruction Algorithm 1, for example, each sensing error needs to be identified and then corrected using two arithmetic (addition) operations. These operations, therefore, increase significantly as the number of the sensing errors increases and they, even more, double for every iteration performed. In contrast, the complexity of the deep learning approach depends on the NN models used to perform estimation (i.e., number of layers, neurons, etc.). The computation requirements of this approach weighs more on the training process than on the prediction process of the DL models. However, this training operation does not take place often, in fact once a DL model is trained it can then be used to perform estimations for the channel traffic statistics.

Table 6 shows a comparison for the computational complexity of the considered approaches in this work in terms of the computation time taken to perform 100 samples of estimations for the channel traffic statistics under ISS. As it can be appreciated, the computational cost associated with the closed-form expressions approach is the most efficient one, while it is significantly higher for the reconstruction algorithms approach. On the other hand, the deep learning approach is considerably less complex than the algorithmic approach and reasonably more complex than the closed-form expression approach. It can also be noticed that the already trained DL models require significantly less computations than the resulting computations from the training process, however, as explained earlier, this training is not required to take place often to preform estimations for the channel traffic statistics. Therefore, considering the significant accuracy improvement with a reasonable increase in the complexity, the proposed DL approach can be considered an efficient solution for providing accurate estimation for the channel traffic statistics under ISS.

TABLE 6. Computation time (in seconds) required by each approach to provide 100 estimations for different statistical metrics.

Approach	Mean	Variance	Minimum	Distribution
Closed-form expressions	0.037	–	0.001	0.52
Reconstruction algorithms	200.7	243	266.8	280
Deep Learning	0.4	0.42	0.47	1.7
+ training	50.8	51.4	50.74	162.1

IX. CONCLUSION

The harmonious coexistence of several wireless communication systems in a shared frequency spectrum is highly dependent on making effective decisions for the utilisation of such spectrum. These decisions are usually based on the users' activity within the channel and their traffic statistical information. Therefore, it is crucial for a spectrum sharing system to obtain accurate estimation of the traffic statistics even under low SNR conditions (i.e., ISS). In this context, this work has studied the existing approaches in the literature that correct the estimation of the statistical parameters of the channel traffic under ISS, including both closed-form expression approach and the algorithmic reconstruction approach. In addition, a novel deep learning approach has been proposed, which can learn from the imperfect observations of the traffic statistics in order to predict their accurate estimations. Therefore, several estimation methods based on deep learning have been modelled and validated for the mean, variance, minimum and distribution of the channel traffic. It was demonstrated that the proposed approach outperforms the previous approaches widely used in the literature, which are based on closed-form expressions and reconstruction algorithms, under different scenarios of sensing error probabilities.

Finally, the investigation of using other types, more powerful, neural networks, e.g., Convolutional Neural Network (CNN) and/or Recurrent Neural Network (RNN), to solve the problem of estimation of channel traffic statistics under ISS, and the potential of using multitask learning with a shared NN model to provide multi statistical parameters is suggested as a part of the future work. In addition, the complexity of these neural networks with respect to the ones considered in this work would be also important to investigate. A useful extension of this work, furthermore, would be the exploitation of the proposed estimation methods in various applications of spectrum sharing systems.

REFERENCES

- [1] M. Hirzallah, M. Krunz, B. Kecioglu and B. Hamzeh, "5G New Radio Unlicensed: Challenges and Evaluation," in *IEEE Transactions on Cognitive Communications and Networking*, 2020.
- [2] R. Zhang, M. Wang, L. X. Cai, Z. Zheng, X. Shen and L. Xie, "LTE-unlicensed: the future of spectrum aggregation for cellular networks," in *IEEE Wireless Communications*, vol. 22, no. 3, pp. 150–159, June 2015.
- [3] H. Lee, H. Kim, H. J. Yang, J. T. Kim and S. Baek, "Performance Analysis of License Assisted Access LTE with Asymmetric Hidden Terminals," in *IEEE Transactions on Mobile Computing*, vol. 17, no. 9, pp. 2141–2154, 1 Sept. 2018.
- [4] E. Testi, E. Favarelli and A. Giorgetti, "Machine Learning for User Traffic Classification in Wireless Systems," *2018 26th European Signal Processing Conference (EUSIPCO)*, Rome, 2018, pp. 2040–2044.
- [5] B. Soni, D. K. Patel and M. López-Benítez, "Long Short-Term Memory Based Spectrum Sensing Scheme for Cognitive Radio Using Primary Activity Statistics," in *IEEE Access*, vol. 8, pp. 97437–97451, 2020.
- [6] A. Al-Tahmeesschi, K. Umabayashi, H. Iwata, J. Lehtomäki and M. López-Benítez, "Feature-Based Deep Neural Networks for Short-Term Prediction of WiFi Channel Occupancy Rate," in *IEEE Access*, vol. 9, pp. 85645–85660, 2021.
- [7] P. Salvo Rossi, D. Ciuonzo and G. Romano, "Orthogonality and Cooperation in Collaborative Spectrum Sensing through MIMO Decision Fusion," in *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5826–5836, November 2013.
- [8] A. Patel, H. Ram, A. K. Jagannatham and P. K. Varshney, "Robust Cooperative Spectrum Sensing for MIMO Cognitive Radio Networks Under CSI Uncertainty," in *IEEE Transactions on Signal Processing*, vol. 66, no. 1, pp. 18–33, 1 Jan. 1, 2018.
- [9] I. Dey, D. Ciuonzo and P. S. Rossi, "Wideband Collaborative Spectrum Sensing Using Massive MIMO Decision Fusion," in *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5246–5260, Aug. 2020.
- [10] O. H. Toma and M. López-Benítez, "Cooperative Spectrum Sensing: A New Approach for Minimum Interference and Maximum Utilisation," *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2021, pp. 1–6.
- [11] M. A. Saleem et al., "Expansion of Cluster Head Stability Using Fuzzy in Cognitive Radio CR-VANET," in *IEEE Access*, vol. 7, pp. 173185–173195, 2019.
- [12] O. H. Toma, M. López-Benítez, D. K. Patel and K. Umabayashi, "Estimation of Primary Channel Activity Statistics in Cognitive Radio Based on Imperfect Spectrum Sensing," in *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2016–2031, April 2020.
- [13] M. López-Benítez, "Can primary activity statistics in cognitive radio be estimated under imperfect spectrum sensing?" in *Proc. 24th Annual IEEE Int'l. Symp. Pers., Indoor and Mobile Radio Comms. (PIMRC 2013)*, Sep. 2013, pp. 750–755.
- [14] A. Al-Tahmeesschi, M. López-Benítez, J. Lehtomäki and K. Umabayashi, "Improving primary statistics prediction under imperfect spectrum sensing," *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, Barcelona, 2018, pp. 1–6.
- [15] O. H. Toma, M. López-Benítez, D. K. Patel and K. Umabayashi, "Reconstruction Algorithm for Primary Channel Statistics Estimation Under Imperfect Spectrum Sensing," *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, Seoul, South Korea, 2020, pp. 1–5.
- [16] M. López-Benítez, A. Al-Tahmeesschi, D. Patel, J. Lehtomäki and K. Umabayashi, "Estimation of Primary Channel Activity Statistics in Cognitive Radio Based on Periodic Spectrum Sensing Observations," in *IEEE Trans. on Wirel. Comms.*, vol. 18, no. 2, pp. 983–996, Feb. 2019.
- [17] O. H. Toma, M. López-Benítez, D. K. Patel and K. Umabayashi, "Primary Channel Duty Cycle Estimation under Imperfect Spectrum Sensing Based on Mean Channel Periods," *2019 IEEE Global Communications Conference (GLOBECOM)*, Waikoloa, HI, USA, 2019, pp. 1–6.
- [18] O. H. Toma, M. López-Benítez and D. K. Patel, "Analysis of the Sample Size Required for an Accurate Estimation of Primary Channel Activity Statistics under Imperfect Spectrum Sensing," *2020 IEEE 31st Annual Intl. Symp. Pers., Indoor and Mobile Radio Comms. (PIMRC)*, London, 2020, pp. 1–6.
- [19] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous univariate distributions*, 2nd ed. Wiley, Nov. 1994, vol. 1.
- [20] M. López-Benítez, and F. Casadevall, "Time-dimension models of spectrum usage for the analysis, design, and simulation of cognitive radio networks," *IEEE Trans. Veh. Tech.*, vol. 62, no. 5, pp. 2091–2104, Jun 2013.
- [21] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, no. 4, pp. 523–531, Apr. 1967.
- [22] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.
- [23] T. O'Shea and J. Hoydis, "An Introduction to Deep Learning for the Physical Layer," in *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, Dec. 2017.
- [24] M. Abadi et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," 2015. [Online]. Available: <http://tensorflow.org/>.

- [25] F. Chollet. *et al.*, “Keras,” 2015. [Online]. Available: <https://github.com/fchollet/keras>.
- [26] Paszke, Adam *et al.*, “Pytorch: An imperative style, high-performance deep learning library,” in *Advances in neural information processing systems*, pp. 8026–8037, 2019.
- [27] W. Gabran, C. H. Liu, P. Pawelczak, and D. Cabric, “Primary user traffic estimation for dynamic spectrum access,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 3, pp. 544–558, March 2013.
- [28] P. Tehrani, L. Tong, and Q. Zhao, “Asymptotically efficient multichannel estimation for opportunistic spectrum access,” *IEEE Trans. Signal Processing*, vol. 60, no. 10, pp. 5347–5360, Oct. 2012.
- [29] Q. Liang, M. Liu, and D. Yuan, “Channel estimation for opportunistic spectrum access: uniform and random sensing,” *IEEE Trans. Mobile Computing*, vol. 11, no. 8, pp. 1304–1316, Aug. 2012.
- [30] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, 10th ed. New York: Dover, 1972.
- [31] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes: The art of scientific computing*, 3rd ed. Cambridge University Press, 2007.



MIGUEL LÓPEZ-BENÍTEZ (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees (Hons.) in telecommunication engineering from Miguel Hernández University, Elche, Spain, in 2003 and 2006, respectively, and the Ph.D. degree (*summa cum laude*) in telecommunication engineering from the Technical University of Catalonia, Barcelona, Spain, in 2011.

From 2011 to 2013, he was a Research Fellow with the Centre for Communication Systems Research, University of Surrey, Guildford, U.K. In 2013, he became a Lecturer (Assistant Professor) with the Department of Electrical Engineering and Electronics, University of Liverpool, U.K., where he has been a Senior Lecturer (Associate Professor) since 2018. His research interests include the field of wireless communications and networking, with special emphasis on cellular mobile communications and dynamic spectrum access in cognitive radio systems. He has been a member of the Organizing Committee of IEEE International Workshop on Smart Spectrum (IWSS) since 2015. He has been the Principal Investigator or Co-Investigator of research projects funded by the EPSRC, British Council, and Royal Society. He has also been involved in the European-funded projects AROMA, NEWCOM++, FARAMIR, QoS MOS, and CoRaSat. He is an Associate Editor of IEEE ACCESS, *IET Communications*, and *Wireless Communications and Mobile Computing*, and has served as a guest editor for other journal’s special issues.

...



OGEEN H. TOMA (Student Member, IEEE) received the B.Sc. degree in electrical and computer engineering from the University of Duhok, Iraq, in 2011, and the M.Sc. degree in wireless communications (Hons.) from the University of Southampton, U.K., in 2015. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering and Electronics, University of Liverpool, U.K., under the U.K.–India Education and Research Initiative (UKIERI) DST Thematic Partnerships Programme.

He also joined Liverpool John Moores University (LJMU) as a Research Assistant working on the development and testing of 5G network based on 60 GHz mmWave technology. He served as a Graduate Teaching Assistant with the Department of Electrical Engineering and Electronics, University of Liverpool since 2018. His research interests include 5G wireless communications, machine learning/deep learning for xG wireless communications, cognitive radio, dynamic spectrum access, spectrum sensing, statistics estimation, cooperative wireless communications, mmWave, MIMO and massive MIMO systems.