Machine Learning Accelerated Antenna Selection for an Underlay Spectrum Sharing System

Saurav Nath, Rimalapudi Sarvendranath *Member, IEEE* and Salil Kashyap *Member, IEEE*

***Abstract*—Machine learning (ML) techniques can help reduce the computational complexity of antenna selection (AS) for an underlay spectrum sharing system, which enables efficient spectrum utilization. We illustrate this by proposing two ML approaches: i) A supervised learning based classification approach with instantaneous channel power gains as input and antenna index as output. It predicts antenna index directly with low complexity; ii) A model-based regression approach with system parameters such as the number of antennas, fixed transmit power, and channel statistics as inputs and a parameter required to implement the conventional AS rule as output. It substitutes the obtained parameter in the conventional AS rule to select the antenna index and exploits the structure of the optimal AS rule. Furthermore, it generalizes to different system parameters and reduces training overhead significantly. We show that both neural network and random forest models trained for the above approaches are near-optimal. We also show that these models work for different channel models including 3GPP adapted tap delay line models while requiring orders of magnitude lower amount of time.**

***Index Terms*—Machine learning, spectrum sharing, underlay, antenna selection, model based.**

I. Introduction

Spectrum sharing is crucial to support large number of users and high data rates in the limited wireless spectrum [1]. Due to its promising nature, it has become a part of many wireless standards such as 5G new radio unlicensed, IEEE 802.11.be [2]. In underlay spectrum sharing, the low priority secondary users transmit concurrently in the same spectrum as the high priority primary users, while ensuring interference caused to the primary receiver (PRx) is below a threshold. These interference constraints degrade performance of the secondary system [3], [4].

Multiple-antennas have been employed at the secondary transmitter (STx) to recover the performance of the secondary system [3], [5]. For example, transmit beamforming, which requires a dedicated radio frequency (RF) chain for each antenna, is explored in [5]. Antenna selection (AS) technique is developed to reduce the cost of the RF chains and to exploit the diversity benefits of multiple antennas. In it, an antenna is selected based on the channel conditions and the available RF chain is connected to it [3]. AS is part of LTE and IEEE 802.11n wireless standards [6].

Conventional AS: Transmit AS for a spectrum sharing system has been widely studied using conventional techniques [3], [4], [7]. Their computational complexity increases as the number of antennas increase [8] and low-complexity suboptimal algorithms are considered in practice. Optimal AS rule that minimizes the average symbol error probability (SEP) for an average interference constrained secondary system is developed in [3]. The optimal selection metric of each antenna involves numerical integration and requires a parameter that is obtained by solving an equation with double integral. Similarly, the optimal AS rules developed for interference outage constraint in [7] also requires numerical integration.

Machine learning (ML) models, which can learn complex relations in a given data, are shown to perform well for AS while transferring the complexity to the offline training phase [9]–[11]: In [9], K-nearest neighbors and support vector machine based AS is shown to be feasible while reducing online computational complexity. In [10], two cascaded convolutional neural networks (CNNs) are employed to perform AS. A residual neural network (ResNet) based deep unsupervised approach is employed in [11] to perform AS. Here, the dimensions of the dataset depends on the system parameters and models require retraining as the system parameters change. Moreover, ML techniques have not been well studied for AS in an underlay spectrum sharing system.

Focus and Contributions: We consider an underlay spectrum sharing system in which a secondary transmitter (STx) equipped with multiple antennas and single RF chain performs AS to communicate to a single user. Our goal is to develop ML models that select antenna index to reduce SEP of the secondary system while satisfying practically relevant average interference constraint [3] and applicable to any channel model. We focus on reducing the computational complexity and generalization to different system parameters. Existing ML techniques developed for traditional communication systems [9]–[11] need not satisfy the interference constraint.

Our specific contributions are as follows:

i) First, we focus on a supervised learning based classification approach. Here, we separately train a neural network based classifier and a random forest based classifier to learn the relationship between the instantaneous channel power gains and the antenna index. This approach performs well with lower complexity. However, retraining would be needed for any change in the system parameters.

ii) Second, we develop a model-based regression approach that exploits the optimal structure of the conventional AS rule. Here, we input all the key system parameters such as transmit power, number of antennas, interference threshold, constellation size, and the channel statistics to the ML model

The work of Rimalapudi Sarvendranath was supported in part by the Science and Engineering Research Board (SERB) Research under Grant SRG/2023/000070, and in part by the Department of Science and Technology (DST) INSPIRE Research Grant. Emails: sarvendranath@iittp.ac.in, sauravjamun@gmail.com, salilkashyap@gmail.com

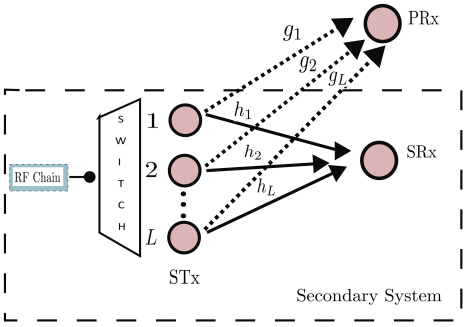


Fig. 1: System model with L antennas and an RF chain at the STx and one antenna at both SRx and PRx.

to obtain a Lagrangian parameter as output. We then plug-in this parameter into the conventional AS rule to select the antenna index. This approach is more general and does not need retraining as the system parameters change.

iii) Our numerical results show that the performance of the proposed ML approaches matches with the conventional AS rule while satisfying the interference constraint. Furthermore, they show that the trained models also work for 3GPP adopted outdoor channel models.

Outline: Section II presents our system model and optimal AS rule. The classification and model-based ML approaches are discussed in Section III. Our numerical results and conclusions are presented in Sections IV and V, respectively.

II. System Model and Conventional AS Rule

Our system model is shown in Figure 1. In it, an STx with L transmit antennas and one RF chain performs AS to communicate to a single antenna secondary receiver (SRx) while sharing spectrum in an underlay manner in the presence of single antenna PRx. AS is done once for each coherence interval and not at the symbol level. Let h_i, for $i \in \{1, 2, \ldots, L\}$ denote the instantaneous channel power gain from the i^{th} antenna of the STx to the SRx and g_i denote the instantaneous channel power gain from the i^{th} antenna of STx to the PRx. The STx to SRx and STx to PRx channel gains are assumed to be independent and identically distributed (i.i.d.). This can be achieved by ensuring sufficient spacing between antennas Let μ_h and μ_g denote the average values of h_i and g_i, respectively. Let $\mathbf{h} \triangleq [h_1, h_2, \ldots, h_L]$ and $\mathbf{g} \triangleq [g_1, g_2, \ldots, g_L]$ denote the channel power gain vectors. We assume that STx knows $\mathbf{h}$ and $\mathbf{g}$, which can be obtained using methods described in [3].

The STx employs on-off power adaptation [3], [7]. It transmits symbol x ($|x|^2 = 1$), which is drawn from a constellation of size M, with a fixed symbol power E_s in the on state or with zero power in off state to reduce interference to the PRx. Let s denote the antenna selected at the STx. We consider the zero transmit power option as selecting a virtual antenna and denote it by $s = 0$. The channel power gains corresponding to $s = 0$ are defined as $h_0 = g_0 = 0$. Hence, $s \in \{0, 1, \ldots, L\}$. The signal r received at the SRx and the interference i_p at the PRx are given by [3]

$$r = \sqrt{E_s}\sqrt{h_s}e^{j\theta_{h_s}}x + n + w_{ps}, \tag{1}$$

$$i_p = \sqrt{E_s}\sqrt{g_s}e^{j\theta_{g_s}}x, \tag{2}$$

where $n + w_{ps}$ is additive white Gaussian noise plus interference, which is assumed to be $\mathcal{CN}\left(0, \sigma^2\right)$, and θ_{h_s} and θ_{g_s} are the phases of the STx to SRx and STx to PRx channel gains.

From (2) the fading averaged interference seen by the PRx is given by $E_s\mathbb{E}_{\mathbf{h},\mathbf{g}}[g_s]$. The average interference constraint that limits the fading averaged interference to be below the threshold I_{ave} is given by $E_s\mathbb{E}_{\mathbf{h},\mathbf{g}}[g_s] \leq I_{\text{ave}}$. Let $\text{SEP}(h_s)$ denote the instantaneous SEP when option s is selected. For M-ary phase shift keying (PSK), we know [3]

$$\text{SEP}(h_s) = \int_0^{\left(1-\frac{1}{M}\right)\pi} \exp\left(\frac{-h_s E_s \sin^2\left(\frac{\pi}{M}\right)}{\sigma^2 \sin^2\theta}\right) d\theta. \tag{3}$$

The optimal transmit antenna index s^* that minimizes the average SEP of the secondary system under the average interference constraint is derived in [3]. It is given by $s^* = \arg\min_{i\in\{0,1,\ldots,L\}}\{\text{SEP}(h_i)+\lambda g_i\}$, where $\lambda > 0$ is computed numerically such that the average interference $I_{\text{opt}}(\lambda) = E_s\text{E}_{\mathbf{h},\mathbf{g}}[g_{s^*}]$ is equal to I_{ave}, i.e., $I_{\text{opt}}(\lambda) = I_{\text{ave}}$. The average SEP is an important performance metric of a communication system and has been well studied in the literature [3], [7], [12]. For Rayleigh fading channel model, the expression for $I_{\text{opt}}(\lambda)$ is given by [3, Sec II C]

$$I_{\text{opt}}(\lambda) = \frac{E_s L\alpha}{\Omega m^{\frac{\alpha}{\Omega}}\mu_g}\int_0^{m\pi}\int_0^{\frac{m-y_1}{\lambda}} g_1 y_1^{\frac{\alpha}{\Omega}-1} e^{\frac{g_1}{\mu_g}} \times\left(1 - e^{-\frac{y_1+\lambda g_1}{\lambda\mu_g}}\beta\tilde{\gamma}\left(\frac{\alpha}{\Omega}+1, \frac{y_1+\lambda g_1}{\lambda\mu_g}\right)\right)^{L-1} dg_1 dy_1, \tag{4}$$

where $m = 1 - \frac{1}{M}, \beta = (\lambda\mu_g/m)^{\frac{\alpha}{\Omega}}, \alpha \triangleq \csc^2\left(\frac{\pi}{M}\right), \Omega \triangleq E_s\mu_h/\sigma^2, \tilde{\gamma}(s,x) \triangleq \int_0^x t^{s-1}e^t dt$.

Complexity of the Conventional Optimal AS Rule: Firstly, the selection metric of each antenna, i.e., $\text{SEP}(h_i) + \lambda g_i$, requires numerical integration to compute $\text{SEP}(h_i)$. Secondly, we need to obtain parameter λ. For statistical channel models, we can derive an expression for $I_{\text{opt}}(\lambda)$ and solve for $I_{\text{opt}}(\lambda) = I_{\text{ave}}$ numerically. As shown in (4), this involves solving equations with multiple integrals. Furthermore, for empirical channel models, we need to employ line search based methods to find λ. Obtaining λ accurately is critical to ensure that the interference constraint is not violated.

We note that solving $I_{\text{opt}}(\lambda) = I_{\text{ave}}$ requires to implement AS, which is computationally complex. We illustrate this below. Let the domain of integration be a polygon Ω consisting of $\mathcal{T}$ triangles. From [13], the error ϵ is bounded by $|\epsilon| \leq K_n e^{n+1}|\Omega|$, where $|\Omega|$ is the area of the polygon, e is the maximum edge length of any triangle, and n denote the degree of exactness of the quadrature rule, K_n is a positive constant independent of e. For trapezoidal rule, with $n = 1$, the maximum length of any edge of the triangles in $|\Omega|$ for

desired error ϵ is $\sqrt{\epsilon/(|\Omega|K_1)}$, i.e. $\mathcal{O}(\sqrt{\epsilon/|\Omega|})$. But, given that Ω is broken into $\mathcal{T}$ triangles, the edge length is $\mathcal{O}(\sqrt{|\Omega|/\mathcal{T}})$. Therefore, the computational complexity to compute any two dimensional area by triangulating it into $\mathcal{T}$ triangles and taking a weighted sum of all the triangles is $\mathcal{O}(|\Omega|^2/\epsilon)$. We note that the computational complexity to solve $I_{\text{opt}}(\lambda) = I_{\text{ave}}$ to solve λ could explode for larger $\mathcal{T}$ and smaller ϵ, making it infeasible for deployment in practical scenarios.

III. Accelerating AS in Underlay Spectrum Sharing Using Machine Learning

With a motivation to reduce the above computational complexity to implement the optimal AS rule, we now propose two ML approaches: i) Supervised learning based classification approach and ii) Model based regression approach. We train a deep neural network (DNN), which is known for its universal function approximation, and random forest, which is an ensemble of decision trees. We note that the other ML models such as support vector machines (SVM), K-nearest neighbors (KNN) can also be studied for the given problem. However, we focus on random forest and DNN due to interpretabilty and low computational complexity.

A. Classification Based Approach

Here, we train the above mentioned ML models to predict the optimal antenna index directly for a given set of instantaneous channel power gains. We now describe the dataset generation and the training process of the DNN and random forest models.

Dataset Generation: For a given channel model, number of antennas L, symbol power E_s at the STx, channel statistics μ_h, μ_g and for a given I_{ave} value, we compute $\lambda > 0$ using line search method if $E_s\mu_g \geq I_{\text{ave}}$. Else $\lambda = 0$. Next, we generate channel vectors $\mathbf{h}$ and $\mathbf{g}$ based on the channel model and obtain the optimal antenna index s^*. We repeat this for N_t number of channel vectors. The above steps are repeated for a range of different I_{ave} values to generate the complete dataset. The concatenated vector of $\mathbf{h}$, $\mathbf{g}$, and I_{ave}, with $2L+1$ features forms the input vector and selected antenna index $s^* \in \{0, 1, \ldots, L\}$ form the output label of each sample. Among all the system parameters, we have only included I_{ave} in the input vector for two reasons. Firstly, to limit the dataset size and secondly to ensure interference constraint is satisfied, which is crucial for an underlay spectrum sharing system.

DNN Based Classification: We first consider a fully connected feed forward DNN with $2L + 1$ neurons in the input layer, multiple hidden layers, and an output layer with $L + 1$ neurons, where L neurons correspond to the antenna indices and one additional neuron correspond to $s = 0$. We employ $Relu(z) = \max(0, z)$ as the activation function for all the hidden layers and soft-max as the activation function for the output layer. Cross-entropy is chosen to be the loss function to obtain weights and biases.

Random Forest Based Classification: It combines the output of multiple independent decision trees to reach a single result. Each decision tree has a root node with a random subset of the total training samples. To build a tree, we split these samples into two nodes recursively until we reach the leaf nodes. The samples are split such that the Gini index is minimized.

The above models trained for a given set of system parameters perform well for those parameters. However, retraining is needed if any of the system parameters change. For example, dimensions of the input vector changes as the number of antennas L change. Furthermore, in a wireless communication system, the change of certain parameters is unavoidable. For example, the channel statistics μ_h and μ_g will change as the relative distance between transmitter and receiver changes. In such a scenario, we would need to retrain the classifiers for the new set of system parameters. we develop a new approach in the next section to address this issue.

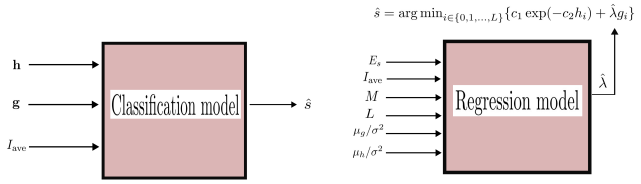
B. Model Based Learning Approach

We will now develop an ML approach to address generalization issue with the classification based approach introduced earlier. We shall refer to this as model-based approach as it exploits the structure of the optimal AS rule derived using conventional model-based techniques. We develop regression models, which takes different system parameters as inputs and outputs estimate of λ. Then, we substitute this estimate in the selection metric to select the antenna index. Figure 2b illustrates the operation of model-based approach and compares it with classification based approach in Figure 2a. To further reduce the complexity, we simplify the selection metric by replacing $\text{SEP}(h_i)$ with its approximation, i.e., $\text{SEP}(h_i) \approx c_1 \exp(-c_2 h_i)$, where c_1 and c_2 are modulation specific constants. This approximation avoids numerical integration and is shown to be accurate for many modulations [7]. The simplified AS rule becomes $\arg\min_{i\in\{0,1,\ldots,L\}}\{c_1 \exp(-c_2 h_i) + \lambda g_i\}$. This replacement avoids numerical integration in (3) and reduces computational complexity.

Dataset Generation for Model Based Approach: Unlike the classification based models, the model-based regression approach takes in the system parameters E_s, I_{ave}, M, L, μ_h/σ^2 and μ_g/σ^2 as inputs. First, we generate different combinations of system parameters in their valid ranges. For each set of parameters, we compute $\lambda > 0$ by using line search method, which is applicable to any channel model. We generate the dataset with system parameters as input features and corresponding λ, which satisfies the constraint, as the output. This is repeated for all N_m combination of system parameters.

DNN Based Regression: We consider a feed forward DNN with six neurons in the input layer, one neuron in the output layer, and multiple hidden layers. Each neuron in the input layer correspond to the system parameters E_s, I_{ave}, M, L, μ_h/σ^2 and μ_g/σ^2. The single neuron in the output layer yields the estimate of λ. We employ $Relu(z) = max(0, z)$ as the activation function of hidden layers and linear function as the activation function at the output layer. Mean squared error (MSE) is used as the loss function.

Random Forest Based Regression: Each decision tree in the random forest regressor algorithm starts with all the training



(a) Classification based approach (b) Model based approach

Fig. 2: Schematic of classification and model-based learning approaches

samples at the root node and splits them recursively. Here, the optimal split of input samples is decided by minimizing the MSE given by $\frac{1}{N_m}\sum_{i=1}^{N_m}(\lambda_i - \hat{\lambda}_i)^2$, where λ_i is the true output of the i^{th} training sample and $\hat{\lambda}_i$ is the output estimated by the learning model. The random forest model, which is based on crowd wisdom theory, aggregates $\hat{\lambda}$ estimated by independent decision trees. The final estimate $\hat{\lambda}$ is the average of all the estimates of all the decision trees.

Once the regression models are trained, we can estimate λ for different system parameters by just changing the input to the trained model. We need not generate a new dataset and retrain our learning models every time a system parameter changes, which makes our model-based approach general. Furthermore, in the next Section, we show that this approach achieves a similar performance as classification based approach with much smaller training dataset.

Comparison of Operations Performed by Three Approaches to Select Antenna Index: To highlight the benefits of ML approaches, we now discuss the operations performed by the conventional optimal AS approach, classification approach and model-based approach to select antenna index.

1) *Conventional Approach:* For antenna i, the conventional optimal AS rule computes $\text{SEP}(h_i) + \lambda g_i$, which involves computationally complex numerical integration, each with a computational complexity of $\mathcal{O}(L|\Omega|^2/\epsilon)$. Then, it computes antenna index s^* by comparing $L+1$ values.

2) *Classification Based Approach:* DNN performs basic operations such as multiplications and additions in the hidden layers. The number of floating point operations performed for a given dense layer is $n_{in}n_{out}$, where n_{in} and n_{out} are the number of input and output neurons, respectively. The output layer computes probabilities using softmax function and compares $L+1$ values to select antenna index. Random forest traverses each tree and selects the antenna based on the leaf nodes reached. Both ML models avoid complex numerical integration required by the conventional approach.

3) *Model Based Approach:* Here, model takes in system parameters as inputs and outputs the estimate of λ. Therefore, the run time of the ML models are independent of L and have constant time complexity. Once the λ is estimated, for $i \in \{0, 1, 2, \ldots, L\}$ it computes closed-form expression $c_1 \exp(-c_2 h_i) + \lambda g_i$ and performs $L+1$ comparisons to select antenna.

Table I lists the number of operations performed by conventional AS rule, and ML approaches. Both approaches can be deployed in resource-limited devices due ot limited number of operations and small memory requirement.

TABLE I: Comparison of number of operations

Operations	Conventional	Classification	Model based
Numerical integration	L	0	0
Comparisons	$L+1$	$L+1$	$L+1$

IV. Numerical Comparison

In this Section, we compare the performance of the proposed supervised learning based classification approach and the model-based regression approach with the conventional optimal AS rule using Monte Carlo simulations.

1) Classification Approach: For one value of I_{ave}/σ^2, we generated a dataset with $N_t = 10^5$ samples and repeated this process for I_{ave}/σ^2 ranging from -20 dB to 20 dB in steps of 1 dB. Therefore, the total number of samples in the complete dataset is equal to 41×10^5. We use 75% randomly selected samples from the above dataset for training the DNN and random forest classification models. Remaining 25% samples are used for testing the performance of the trained models.

(i) Neural Network: The architecture of the neural network with 4637 trainable parameters is tabulated in Table II. We used Adam optimizer with 0.001 as the learning rate to optimize the DNN parameters. The network was trained for 1000 epochs and we observed a test accuracy of around 97.8%.

(ii) Random Forest: We considered 50 decision trees each with a maximum depth of 12 for the random forest, which is trained with the same dataset as above. An individual tree was allowed to take only 60% of the total features of the training samples to reduce the variance of the model. Once trained, the random forest algorithm is able to predict the antenna index with an accuracy of around 96.1%.

TABLE II: Architecture of classification neural network

	Size	Parameters	Activation
Input	$2L$+1		
Layer-I (Dense)	48	480	relu
Batchnorm 1-D	48	96	
Layer-II (Dense)	48	2352	relu
Batchnorm 1-D	48	96	
Layer-III (Dense)	24	1176	relu
Batchnorm 1-D	24	48	
Layer-IV (Dense)	12	300	relu
Batchnorm 1-D	12	24	
Layer-V (Dense)	L+1	65	softmax

2) Model Based Learning Approach: For dataset generation, we have considered $L \in \{2, 4, 8, 16\}$ and $M \in \{4, 8\}$. For channel parameters, we considered $\mu_g/\sigma^2 \in [1, 12]$, and $\mu_h/\sigma^2 \in [1, 12]$. For the transmit power, we considered $E_s \in [20, 30]$ dBm in steps of 1 dBm, and $I_{\text{ave}}/\sigma^2 \in [-20, 20]$ dB in steps of 1 dB. For each combination of the system parameters, we computed λ that satisfies the interference constraint to generate the dataset. The total number of samples in the dataset

TABLE III: Architecture of regression neural network

	Size	Parameters	Activation Function
Input	6		
Layer-I (Dense)	48	288	relu
Layer-II (Dense)	48	2352	relu
Layer-III (Dense)	24	1176	relu
Layer-IV (Dense)	12	300	relu
Layer-V (Dense)	1	13	linear

of the model-based approach is $472,320$, where each sample has a input feature vector comprising of E_s, I_{ave}, M, L, μ_h/σ^2 and μ_g/σ^2. The variation in μ_g/σ^2 and μ_h/σ^2 captures the randomness of the user location. This dataset is split into training and test sets in the ratio of $80:20$.

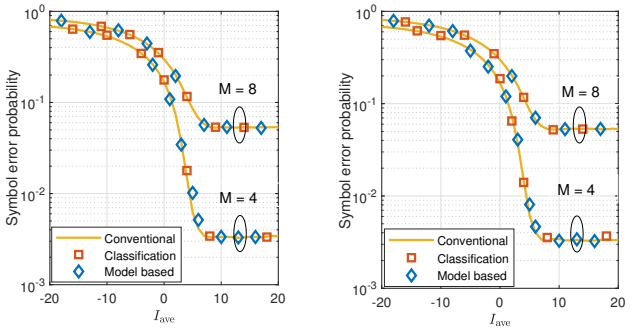
(i) Neural Network: For regression, we considered, a neural network with architecture shown in Table III. It has $4,129$ parameters. The MSE is used as the loss function. The Adam optimizer was chosen with a learning rate of 0.001, $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The MSE of the trained model is observed to be 7.0×10^{-5} after 100 epochs.

(ii) Random Forest: An ensemble of 30 trees each with maximum depth of 12 is considered for the random forest based regression model. We used MSE as the criterion to measure the quality of the splits. To increase the diversity of each tree, we allowed an individual tree to take only 50% of the total features of the training samples. The trained model achieved an MSE of 1.2×10^{-4}.

Figure 3 compares the performance of the classification, the model-based regression, and the conventional approaches. Figures 3a and 3b plot the average SEP as a function of I_{ave} for different constellation sizes for DNN and random forest models, respectively. For smaller value of I_{ave}, we see that the SEP decreases as I_{ave} increases for all the three approaches. This is due to relaxation in I_{ave} allows the STx to transmit with E_s for more time. For large value of I_{ave}, we see that the SEP saturates for all the three approaches. This happens as the STx enters interference unconstrained regime and transmits with E_s all the time. We see that the performance of the classification and model-based approaches match with the conventional approach. The model-based approach achieves it with smaller training dataset.

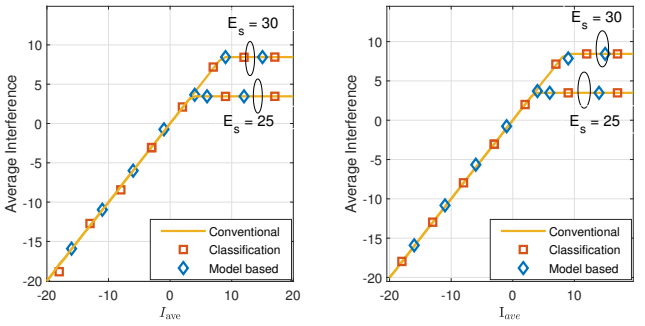
Figure 4 plots the average interference at the PRx normalized by noise power as a function of I_{ave} for $E_s = 25$ dBm and 30 dBm. We see that both DNN and random forest based classification and model-based approaches meet the interference constraint. It shows that the average interference increases initially as the I_{ave} increases and then saturates. The increase happens as the STx transmits with E_s more often as I_{ave} increases and saturation happens as the average interference reaches $E_s\mu_g$, which is the maximum interference possible.

Figure 5 plots the SEP as a function of I_{ave} for $L = 2$ and 4. We again see that the performance of the ML approaches match with the conventional. This shows that the proposed ML approaches work well for different system parameters. In the



(a) Neural network (b) Random forest

Fig. 3: SEP as a function of I_{ave} for different values of M ($E_s = 30$ dBm, $L = 4, \mu_g/\sigma^2 = 7$, and $\mu_h/\sigma^2 = 9$).

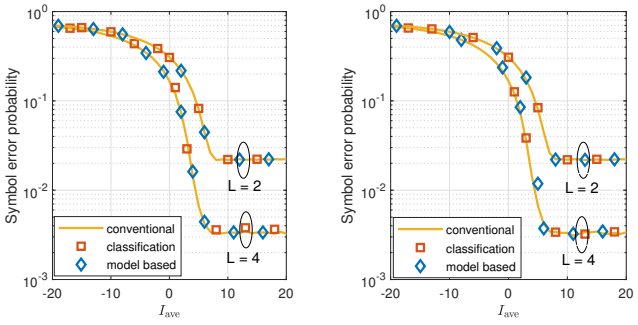


(a) Neural network (b) Random forest

Fig. 4: Average interference as a function of I_{ave} for different values of E_s ($L = 4$, $M = 4, \mu_g/\sigma^2 = 7$, and $\mu_h/\sigma^2 = 9$).

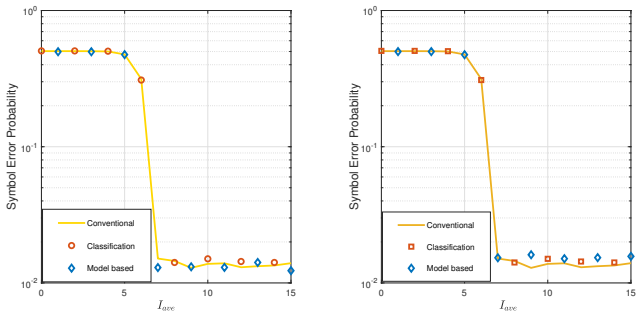
classification based approach, we retrain the model with new dataset as L changes. However, in model-based approach, we just need to change the input to the trained regression model. to obtain the new estimate of λ, which is then used to obtain the selected antenna index. This approach is generalizable across different data distributions.

Tapped Delay Line (TDL) Channel Model: We now demonstrate that our ML models trained for Rayleigh fading



(a) Neural network (b) Random forest

Fig. 5: SEP as a function of I_{ave} for different values of L ($E_s = 30$ dBm, $M = 4, \mu_g/\sigma^2 = 7$, and $\mu_h/\sigma^2 = 9$).



(a) Neural network (b) Random forest

Fig. 6: SEP as a function of I_{ave} for TDL-D channel model ($L = 2$, $E_s = 40$ dBm, $M = 4, \mu_g/\sigma^2 = 1$, and $\mu_h/\sigma^2 = 10$).

model work well with practical channel models adopted by 3GPP [14]. These channels are widely used for link level and system level simulations. As TDL models are wideband channel models with multiple taps, we use the channel power gain of the strongest tap for selection. We employ an OFDM receiver at the SRx and the SEP averaged over all the sub-carriers is shown. Figure 6 plots this average SEP as a function of I_{ave} for the TDL-D channel model described in [14, pp. 66]. Among the different scenarios such as urban microcell, urban macrocell, indoor office and rural macrocell, supported by this model, we have chosen urban microcell, where the base-stations are mounted at rooftop levels of the buildings. We have chosen carrier frequency to be 2 GHz with maximum Doppler shift $f_d = 98$ Hz, and delay spread $\tau = 59$ ns. The trend observed in the average SEP is similar to that of the Rayleigh fading channel model. Though there is a degradation in the performance due to frequency selective channel, we see that both the DNN and random forest models of the classification and model-based approaches perform close to the conventional selection rule. This shows that our ML models trained for the stochastic models are robust and generalize to other channel models.

Table IV compares the empirical computation time of the trained DNN models with the conventional optimal AS rule for $M = 4$ and $N_t = 10^4$ samples. The computations were performed on a computer with AMD Ryzen 7 5800H processor with a base clock speed of 3.2 GHz and 16 logical processors. For all values of L, we see that the conventional AS rule has very high computational time. This is due to the requirement of numerical integration to compute each selection metric. It also scales with the number of antennas.

We see that the computational times are almost constant for classification approach as number of operations in the DNN are almost same. These significantly lower computational times combined with the near-optimal performance illustrate the benefits of employing ML techniques.

V. Conclusion

We developed two ML based approaches for AS in an average interference constrained underlay spectrum sharing system. The classification based approach directly predicted the antenna index removing the requirement of performing numerical integration for computing selection metric and Lagrange parameter. It achieved very high classification accuracy at the cost of a large training overhead and required retraining for any change in the system parameter. Our model-based regression approach exploited the structure of the optimal AS rule and only estimated the Lagrange parameter. This approach yielded generalized learning models, which required smaller training dataset and need not be retrained when a system parameter changed. We showed that both the ML approaches achieved near-optimal performance while satisfying the interference constraint, which is crucial for an underlay spectrum sharing system.

TABLE IV: Computation time comparison (ms)

L	Conventional AS	Classification	Model based
4	22160	8.3	8.4
8	44580	8.3	8.6
16	87900	8.3	8.9

References

[1] C. Yang, J. Li, M. Guizani, A. Anpalagan, and M. Elkashlan, "Advanced spectrum sharing in 5G cognitive heterogeneous networks," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 94–101, Apr. 2016.

[2] E. Au, "IEEE 802.11be: Extremely high throughput standards," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 138–140, Sep. 2019.

[3] R. Sarvendranath and N. B. Mehta, "Antenna selection in interference-constrained underlay cognitive radios: SEP-optimal rule and performance benchmarking," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 496–506, Feb. 2013.

[4] K. Tourki, F. A. Khan, K. A. Qaraqe, H. Yang, and M. S. Alouini, "Exact performance analysis of MIMO cognitive radio systems using transmit antenna selection," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 425–438, Mar. 2014.

[5] R. Zhang and Y. C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 88–102, Feb. 2008.

[6] N. B. Mehta, S. Kashyap, and A. F. Molisch, "Antenna selection in LTE: From motivation to specification," *IEEE Commun. Mag.*, vol. 50, no. 10, pp. 144–150, Oct. 2012.

[7] R. Sarvendranath and N. B. Mehta, "Transmit antenna selection for interference-outage constrained underlay CR," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 3772–3783, Sep. 2018.

[8] Y. Gao, H. Vinck, and T. Kaiser, "Massive MIMO antenna selection: Switching architectures, capacity bounds, and optimal antenna selection algorithms," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1346–1360, Mar. 2018.

[9] J. Joung, "Machine learning-based antenna selection in wireless communications," *IEEE Commun. Lett.*, vol. 20, no. 11, pp. 2241–2244, Nov. 2016.

[10] A. M. Elbir and K. V. Mishra, "Joint antenna selection and hybrid beamformer design using unquantized and quantized deep learning networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1677–1688, Mar. 2020.

[11] Z. Liu, Y. Yang, F. Gao, T. Zhou, and H. Ma, "Deep unsupervised learning for joint antenna selection and hybrid beamforming," *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 1697–1710, Mar. 2022.

[12] R. Sarvendranath and N. B. Mehta, "Antenna selection with power adaptation in interference-constrained cognitive radios," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 786–796, Mar. 2014.

[13] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical Mathematics (Texts in Applied Mathematics)*. Berlin, Heidelberg: Springer-Verlag, 2006.

[14] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," 3GPP, Tech. Rep. TR 38.901 v16.1.0 (2020-11), 2020.