Non-cooperative OFDM Spectrum Sensing Using Deep Learning

Qingqing Cheng, Zhenguo Shi, Diep N. Nguyen, Eryk Dutkiewicz
School of Electrical and Data Engineering, University of Technology Sydney, Sydney, Australia
{Qingqing.Cheng, Zhenguo.Shi}@student.uts.edu.au, {Diep.Nguyen, Eryk.Dutkiewicz}@uts.edu.au

***Abstract*—Although spectrum sensing, a key technique in dynamic spectrum access, has been widely investigated, conventional methods suffer from carrier frequency offset (CFO), timing delay and noise uncertainty, which can significantly degrade the sensing performance. In this paper, we aim to tackle those challenging issues by developing a stacked autoencoder based spectrum sensing approach (SAE-SS). The SAE architecture is employed to effectively learn useful and hidden information from the original received signals. Compared to the existing sensing methods, our approach is more robust to CFO, noise uncertainty and timing delay. Unlike the traditional feature-based detection approaches, the proposed framework does not require the prior knowledge or specific features of incumbent users (IUs). Moreover, in comparison with machine learning based sensing approaches, our solution does not need any external feature extraction algorithms to extract specific features (that is essential for ML-based ones). Through extensive experimental results, our proposed method is demonstrated to achieve notably higher sensing accuracy, e.g., 29% reduced probability of miss detection, than that of state-of-the-art approaches.**

***Index Terms*—Dynamic spectrum access, OFDM, spectrum sensing, stacked autoencoder, deep learning.**

I. Introduction

In recent years, dynamic spectrum access (DSA) has received paramount research interest. In its context, secondary users (SUs) are allowed to opportunistically utilize the licensed spectrum resource of incumbent users (IUs) without inducing any adverse interference [1]. For that purpose, before using the licensed frequency channels, SUs must detect IU's activity states on those channels [2]. In this work, we study the non-cooperative sensing methods using orthogonal frequency division multiplexing (OFDM) signals.

The energy detection (ED) is the simplest non-cooperative OFDM detection approach, while is particularly susceptible to the noise uncertainty [3]. The feature-based OFDM sensing methods are more robust to the noise uncertainty by leveraging the unique structural features of OFDM signals, e.g., the cyclic prefix (CP) features [4], covariance matrix (CM) features [5], etc. However, their sensing performance degrades dramatically due to practical drawbacks, e.g., carrier frequency offset (CFO) and timing delay [6].

Although various machine learning (ML)-based methods for spectrum sensing have been proposed, most existing works focus on cooperative spectrum sensing, e.g., using ML as a tool at a fusion center to aid the decision making process [7]. Unlike the learning-based spectrum sensing in cooperative systems, in this work, we investigate the spectrum sensing method using deep learning networks in a non-cooperative manner. In the context of non-cooperative spectrum sensing, only a few works leverage ML to extract specific features of received signals, e.g., the energy [8] and cyclostationary features [4]. However, the accuracy of extracted features is heavily impacted by communications impairments (e.g., noise uncertainty, timing delay and CFO). Besides, extracting specific features can only retrieve partial knowledge from the received signals, while those hidden but helpful features/information (playing a vital role for spectrum sensing) are inevitably ignored/lost.

In this paper, we aim to address all the aforementioned drawbacks using deep learning networks [9]. There are numerous ML network architectures, including stacked autoencoder (SAE), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), etc [10], [11]. Among those models, RNN can effectively deal with generative problems, while it suffers from the "vanishing gradient" problem [12], hence how to train RNN properly is an extremely difficult task. For typical CNN, it is inevitable to lose features/information in the process of feature extraction because of its partial connection model. In contrast, SAE is conceptually simple and easy to be trained [13]. Moreover, SAE can effectively extract features of the input data and then fully reconstruct the input data using the extracted features [13]. These features of SAE allow us to facilitate accurate sensing performance with the simple training process. We summarize our major contributions as follows.

- We propose a stacked autoencoder based spectrum sensing method (SAE-SS) to infer the IU's presence/absence by extracting and leveraging all the hidden and useful information of the received signals. That helps SAE-SS to probe the IU's activity more correctly.
- Compared to the existing OFDM sensing methods (e.g.,[3], [4], [5], [8]), our developed SAE-SS is more robust to CFO, timing delay and noise uncertainty.
- Unlike the traditional feature-based OFDM detection methods (e.g., [3], [4], [5]), SAE-SS is capable of sensing IU's activity states using the received data only, without any prior information of IU's signals. That makes SAE-SS more suitable for practical implementation, e.g., sharing the military radar bands in the FCC's Spectrum Access System (SAS).
- SAE-SS is capable of automatically extracting the hidden information from the original received signals without requiring any external feature extraction methods (often required in ML-based methods, e.g.,[8]). That improves the sensing accuracy and makes SAE-SS more suitable for practical applications.
- We carry out extensive simulations to evaluate SAE-SS's performance. The results show that compared with the

This work has been supported in part by the Australian Research Council (Discovery Early Career Researcher Award DE150101092).

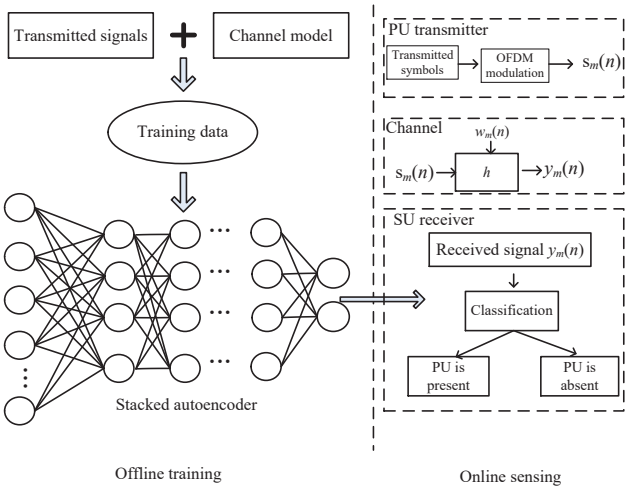


Fig. 1. Architecture of proposed SAE-SS

typical OFDM sensing approaches, the probability of miss detection of SAE-SS is significantly smaller, especially under poor communication conditions.

II. System Model

Consider a typical OFDM DSA system, where the IU transmits OFDM signals, and SUs can opportunistically utilize the IU's spectrum when that IU is absent. Let $y_m(n)$ represent the discrete time received signals, where $m = 0, \ldots, M-1$, and M denotes the number of received OFDM blocks in total. $n = 0, \ldots, N_c + N_d - 1$, N_c is the length of CP and N_d is the data block size. Then $y_m(n)$ can be expressed as

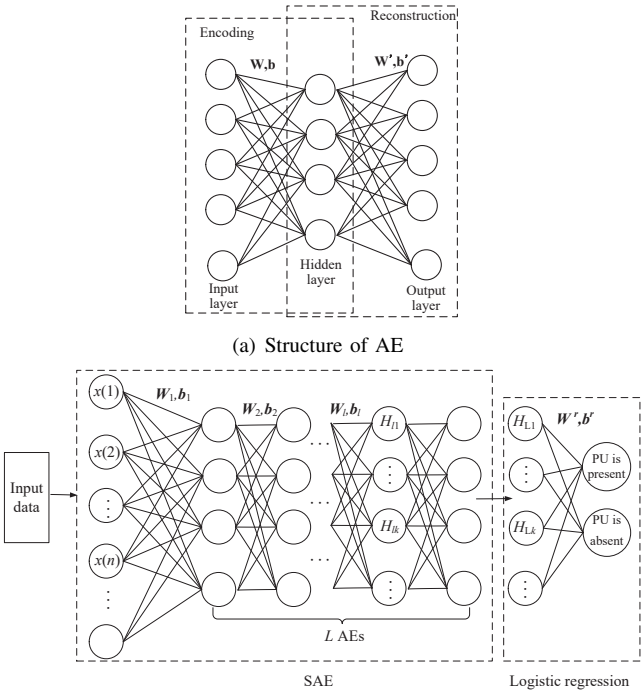
$$\begin{aligned} &\mathcal{H}_0 : y_m(n) = w_m(n), \\ &\mathcal{H}_1 : y_m(n) = e^{-j\frac{2\pi f_q (n-\delta)}{N_d}} \sum_{i=0}^{L_p-1} h_i s_m(n-\delta-i) + w_m(n), \end{aligned} \quad (1)$$

where $\mathcal{H}_0$ stands for the absence of the IU and $\mathcal{H}_1$ represents the presence of the IU. $w_m(n)$ denotes the complex additive white Gaussian noise (AWGN) with zero-mean and variance σ_w^2 (i.e., $w_m(n) \sim CN(0, \sigma_w^2)$). $s_m(n)$ stands for the transmitted OFDM signals from the IU. f_q represents the normalized CFO. δ is the timing delay. h_i refers to the channel gain of the ith channel component, which does not change in the sensing duration, and L_p is the total number of multi-path components between the SU and the IU. Without loss of generality, we assume that $s_m(n)$, h_i and $w_m(n)$ are independent with each other. Under the central limit theorem, $s_m(n)$ approximately obeys the complex Gaussian distribution with zero-mean and variance σ_s^2, given sufficiently large length of the received signals. Thus, SNR of $y_m(n)$ under $\mathcal{H}_1$ is defined as SNR $= \sigma_s^2 \sum_{i=0}^{L-1} |h_i|^2 / \sigma_w^2$.

For typical OFDM sensing methods, Take the ED method [3] as an instance, its test statistic is

$$T_{ED} = \frac{1}{M(N_c + N_d)} \sum_{m=0}^{M-1} \sum_{n=0}^{N_c+N_d-1} |y_m(n)|^2. \quad (2)$$

From equation (2), it is clear that the ED sensing method infers the presence of the IU utilizing the energy of the received signals only. Thus it is sensitive to noise uncertainty, and a small mismatch in the estimated and actual noise power would lead to significant performance degradation, especially in low SNR conditions. Although the feature-based sensing methods are superior to the ED method regarding the robustness to noise uncertainty, they are vulnerable to timing delay and CFO. Consequently, their performance will deteriorate dramatically even with a small error between the transmitted and received signals in the frequency-domain or time-domain [5], [6].



(a) Structure of AE

(b) Structure of SAE with logistic regression classifier

Fig. 2. The structure of spectrum sensing based on deep learning network

III. Proposed Stacked Autoencoder Based Spectrum Sensing Method

The system architecture for the proposed SAE-SS consists of two main stages: offline training and online sensing, as shown in Fig. 1. For the former, it trains the SAE network before the spectrum sensing stage(i.e., offline stage), which aims to learn all the important information from the received OFDM signals. For the latter, it detects the IU's activity states in the online spectrum sensing stage by utilizing the extracted features. In the next, we present the proposed SAE-SS in detail, which includes pre-training, fine-tuning of SAE, and classification.

A. Pre-training of Stacked Autoencoder

For the pre-training of SAE, it learns the hidden information from the received OFDM signals in two main steps. First, divide the whole SAE architecture into several independent autoencoders (AEs) which are three-layer networks that include the input layer, the hidden layer and the reconstruction layer (e.g., the output layer), as shown in Fig. 2(a). Second, train all the independent AEs one by one, and stack all trained AEs together by connecting the input and hidden layers of each AEs, as shown in Fig. 2(b).

We set the number of input units for the first AE as the integral multiple of OFDM blocks. Notably, AE can work with non-complex numbers. As such, we divide all the received OFDM signals into imaginary parts and real parts, respectively.

Thus the received OFDM signals in equation (1) can be rewritten as

$$\mathbf{y}_m := \{\Re(y_m(0)), \Im(y_m(0)), \Re(y_m(1)), \Im(y_m(1)), \ldots, \Re(y_m(N_c+N_d-1)), \Im(y_m(N_c+N_d-1))\}, \quad (3)$$

where $\Im(.)$ and $\Re(.)$ denote the imaginary part and real part, respectively. We express the input vector of the first AE as

$$\mathbf{x} := \{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{M-1}\}^T, \quad (4)$$

where the length of $\mathbf{x}$ is $N_{input} = 2M(N_c + N_d)$. Note that $\mathbf{x}$ is OFDM signals containing many unique features (e.g., CP and PT structure), which is beneficial for feature extraction and sensing IU's activities. We then feed $\mathbf{x}$ into the network to extract the hidden features through training all AEs.

We express the training process of AEs using the study case of lth AE. Let H_{lp}^{out}, H_{lk} and H_{lp}^{in} stand for the pth output unit, the kth hidden unit and pth input unit of the lth AE, respectively. The relationship among them can be given by [14]

$$H_{lk} = f(\sum_{p=1}^{P_l} W_{lpk} H_{lp}^{in} + b_{lk}), \quad (5)$$

$$H_{lp}^{out} = f(\sum_{k=1}^{K_l} W_{lkp}^{'} H_{lk} + b_{lp}^{'}), \quad (6)$$

where P_l represents the number of input units at the lth AE. K_l stands for the number of hidden units in the lth AE. W_{lpk} is the weight between the pth input unit and the kth hidden unit in the lth AE. $W_{lkp}^{'}$ stands for the weight between the pth output unit and the kth hidden unit of the lth AE. For simplicity, we assume $W_{lpk} = W_{lkp}^{'}$. b_{lk} and $b_{lp}^{'}$ are the biases of the kth hidden unit and the pth output unit in the lth AE, respectively. $f(.)$ stands for the activation function. The sigmoid function [13] is selected as $f(.)$ in this work, given by

$$f(H_{lp}^{in}) = \frac{1}{1+e^{-H_{lp}^{in}}}. \quad (7)$$

Notably, when $l = 1$, $\mathbf{H}_1^{in} = \mathbf{x}$, which is

$$\mathbf{H}_1^{in} := \{H_{11}^{in}, H_{12}^{in}, \ldots, H_{1P_1}^{in}\}^T, \quad (8)$$

where P_1 represents the number of input units in the first AE, $P_1 = N_{input} = 2M(N_c+N_d)$. When $l > 1$, $\mathbf{H}_l^{in}$ is the hidden units of $(l-1)$th AE.

The core task of training the lth AE is to minimize the error between H_{lp}^{in} and H_{lp}^{out}, by continuously updating the values of W_{lpk}, b_{lk} and $b_{lp}^{'}$. In this work, we adopt the cross-entropy method to measure the difference between H_{lp}^{in} and H_{lp}^{out}, which is given by

$$\chi = \sum_{p=1}^{P_l} [H_{lp}^{in} \log(H_{lp}^{out}) + (1 - H_{lp}^{in}) \log(1 - H_{lp}^{out})]. \quad (9)$$

Let $\Omega = \{W_{lpk}, b_{lk}, b_{lp}^{'}\}$, then the objective function is

$$\Omega = \arg\min_{\Omega} \chi(H_{lp}^{in}, H_{lp}^{out}) \quad (10)$$

To optimize Ω, we adopt the gradient descent method [15], and the rules are given by

$$W_{lpk} \leftarrow W_{lpk} - \kappa \frac{\partial \chi(H_{lp}^{in}, H_{lp}^{out})}{\partial W_{lpk}}, \quad (11)$$

$$b_{lk} \leftarrow b_{lk} - \kappa \frac{\partial \chi(H_{lp}^{in}, H_{lp}^{out})}{\partial b_{lk}}, \quad (12)$$

$$b_{lp}^{'} \leftarrow b_{lp}^{'} - \kappa \frac{\partial \chi(H_{lp}^{in}, H_{lp}^{out})}{\partial b_{lp}^{'}}, \quad (13)$$

where κ is the learning rate.

Following the rules mentioned above, all the AEs can be trained. After that, we stack the input and hidden layers of trained AEs together, layer by layer, to obtain the trained SAE. We express the output of SAE's pre-training (also the hidden units of Lth AE) as

$$\mathbf{H}_L := \{H_{L1}, H_{L2}, \ldots, H_{LK_L}\}^T, \quad (14)$$

where K_L stands for the number of hidden units in the Lth AE. Note that, $\mathbf{H}_L$ contains hidden information/features of received signals, which will be utilized to detect the activity states of IUs.

B. Fine-Tuning of Stacked Autoencoder and Classification

The pre-training process of SAE can be seen as the stage of unsupervised feature extraction of IU's signals. To leverage the property of SAE for the spectrum sensing, the pre-trained SAE needs to be fine-tuned. In this work, we implement the logistic regression classifier to fine-tune the whole SAE architecture, as illustrated in Fig. 2(b).

Notably, the logistic regression classifier can be treated as a neural network with a single layer, which applies the softmax function as the activation of the output layer. Its input is $\mathbf{H}_L$ that is also the output features of the pretraining SAE. For its output, U, it is a set of conditional probabilities of $\mathbf{H}_L$, $\mathbf{W}^r$ and $\mathbf{b}^r$, where $\mathbf{b}^r$ and $\mathbf{W}^r$ are biases and weights of the logistic regression layer, respectively. Note that SAE is used to detect whether the IU is present or not, which can be regarded as a binary classification (i.e., the number of classes is 2). In this case, the conditional probability of U can be written as

$$P(U=\tau|\mathbf{H}_L, \mathbf{W}^r, \mathbf{b}^r) = \frac{e^{\mathbf{W}_\tau^r \mathbf{H}_L + \mathbf{b}_\tau^r}}{\sum_{j=0}^{1} e^{\mathbf{W}_j^r \mathbf{H}_L + \mathbf{b}_j^r}}, \quad (15)$$

where τ is 0 or 1. $\tau = 1$ stands for the presence of the IU, and $\tau = 0$ denotes the absence of the IU. The logistic regression classifier can be trained by applying the backpropagation method [13]. It is noteworthy that when training the logistic regression classifier, the fine-tuning of SAE is also completed at the same time.

After pre-training and fine-tuning of the SAE network, which are conducted in the offline phase, SAE-SS is able to detect IU's presence/absence leveraging the received signals only. Compared to the conventional feature-based OFDM sensing methods (e.g., [3], [4], [5], [8]), SAE-SS does not need external feature extraction algorithms to process the received signals, such as calculating the energy of received signals. That significantly reduces the complexity of SAE-SS. Moreover, SAE-SS detects the IU's presence/absence using the received

signals only, without requiring any prior information of the IU (such as transmitting power or the features of IU's signals). That makes SAE-SS more practical than conventional OFDM sensing methods.

C. Complexity Analysis

In this section, we compare the computational complexity of our proposed SAE-SS during the online sensing phase with CP [4], CM [5], and Artificial neural network (ANN) methods [8], as shown in Table I. We select the number of the real multiplication and complex multiplication as metrics because they are the most computationally expensive. Note that, one complex multiplication can be regarded as four times of real multiplication, hence we use the total number of real multiplication to illustrate the computational complexity. In Table I, K_l and K_L stand for the numbers of units in the hidden layer of the lth and Lth layer, respectively. P_l denotes the number of input units in the lth layer. S_δ means the set of consecutive indices for which $x(n) = x(n + N_d)$, given the synchronization mismatch δ.

From this table, it is worth noting that no complex multiplication is involved in SAE-SS because we divide complex signals into the real and imaginary parts (as shown in equation (3)). The online computational complexity of the proposed SAE-SS is intermediate, which is less than the CM and CP methods but higher than the ANN method. When $N_c = 8$, $N_d = 64$, $M = 2$, $L = 2$, $N_{S_\tau} = 7$, $K_1 = 100$, and $K_2 = 50$, the total number of real complication of the proposed SAE-SS is 33850. In contrast, the numbers of CM and ANN methods are 53568, 43392 and 7754, respectively.

IV. Simulation Results

In this section, we carry out extensive simulations to verify the performance of our proposed SAE-SS. For that purpose, we compare the results of SAE-SS and other four OFDM sensing methods (e.g., ED [3], CP [4], CM [5], and ANN [8]). We choose $N_d = 64$ as the data block size of the received OFDM signals, and $N_c = N_d/8$ as the length of CP. The signal bandwidth and the radio frequency are 5 MHz and 2.4 GHz, respectively. The values of SNR fall in the range from -20dB to -8dB. We set the timing delay as $\delta \in [0, N_c + N_d - 1]$. $f_q \in [0, 1]$ is the normalized CFO. There are two hidden layers in SAE-SS, and the number of units in the first and the second hidden layer are 100 and 50, respectively. $M = 2$ is the number of received OFDM blocks. The number of iterations for pre-training and fine-tuning are both 5000. Both the testing data and training data are $2 * 10^4$ samples.

Fig. 3 compares the probability of miss detection (P_m) of different sensing methods under the perfect conditions. The "perfect conditions" means that the SU has enough prior knowledge of IU signals, and there are no effects caused by noise uncertainty, timing delay or CFO. We set the probability of false alarm (P_f) as 0.05. It is clear that, P_m of SAE-SS is much smaller than ED [3], CP [4], CM [5], and ANN methods [8] under the perfect conditions. Additionally, when SNR increases, P_m of SAE-SS gets a greater decrease than those of the other four sensing methods.

To further evaluate the sensing performance, we show the Receiver Operating Characteristic (ROC) curves of various

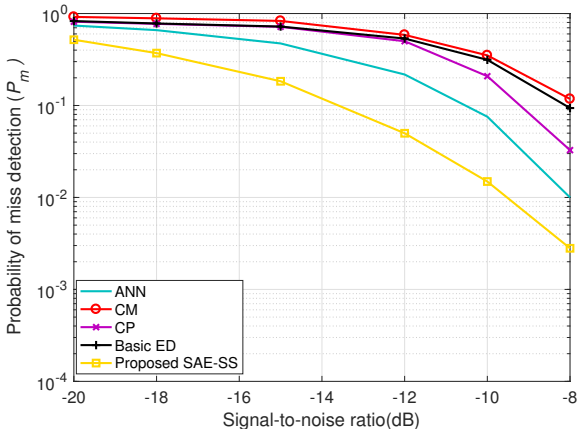


Fig. 3. Probability of miss detection among different spectrum sensing methods under perfect conditions

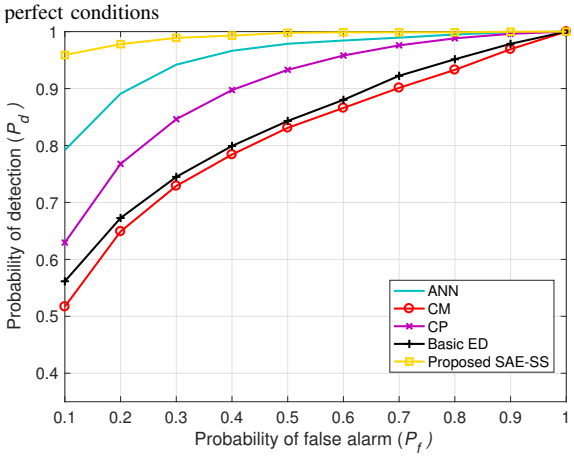


Fig. 4. Receiver Operating Characteristic (ROC) curves of different sensing methods with SNR=-12dB.

methods at SNR= -12dB, as illustrated in Fig. 4. Obviously, SAE-SS can achieve a much higher probability of detection (P_d, $P_d = 1 - P_m$) than the other four methods, with the same P_f. For instance, when $P_f = 0.1$, P_d of SAE-SS is 0.959, whereas those of the other four methods are all lower than 0.8.

Fig. 5 shows P_m of the five sensing methods suffering from the noise uncertainty, η. As can be seen, η affects P_m of all the five sensing methods, but to different extents. Specifically, in comparison with the other four methods, SAE-SS gains the advantage regarding the robustness to noise uncertainty, and its P_m is much smaller.

Fig. 6 shows the impact of timing delay on different sensing methods. The timing delay is selected as 1 symbol and 5 symbols, respectively. Since timing delay does not affect the sensing performance of ED, we only express the sensing results of the other four sensing methods. It is clear that SAE-SS outperforms the other three sensing methods in terms of the robustness to timing delay. In addition, P_m of SAE-SS is the smallest among these four sensing methods even when suffering from timing delay.

Fig. 7 compares the impact of CFO on different sensing methods. The values of the normalized CFO f_q are 0.5 and 1, respectively. Since the ED is not affected by CFO, we do not present its figure here. As can be seen, the presented SAE-SS has an advantage over the other three sensing methods in terms of the robustness to CFO. Besides, P_m of the proposed SAE-SS is much lower than the other three sensing methods even with CFO.

TABLE I
ONLINE COMPUTATIONAL COMPLEXITY OF DIFFERENT SENSING METHODS

Method	Complex Multiplications	Real Multiplication
CP	$(N_c+N_d)(M+N_{S_\tau}+1)+M(N_c+N_d)^2$	$2(N_c+N_d)(N_c+N_d-N_{S_\tau})$
CM	$MN_d\log_2\frac{N_d}{2}+MN_d^2$	$2(N_d^2-N_d)$
ANN	$2M(N_c+N_d)$	$4K_1+\sum_{l=2}^{L}K_lP_l+K_L$
Proposed SAE-SS	None	$2MK_1(N_c+N_d)+\sum_{l=2}^{L}K_lP_l+K_L$

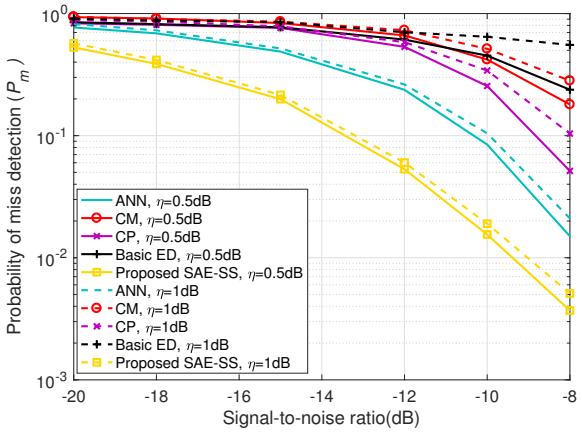


Fig. 5. Probability of miss detection of different sensing methods with noise uncertainty

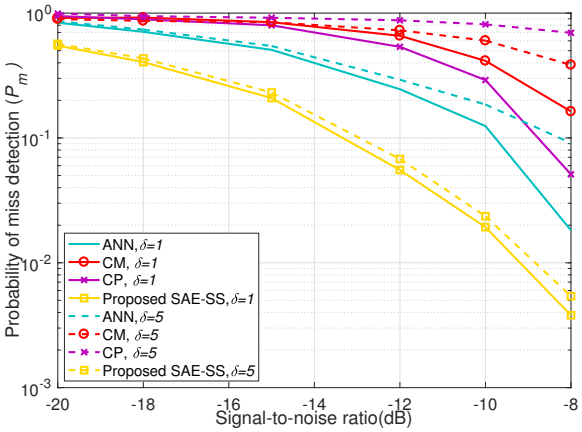


Fig. 6. Probability of miss detection of different sensing methods with timing delay

V. CONCLUSION

In this work, we proposed a Stacked Autoencoder Based Spectrum Sensing Method (SAE-SS) for sensing the presence/absence of IUs that use OFDM modulation. The proposed method is more robust to CFO, noise uncertainty and timing delay, in comparison with traditional OFDM sensing methods. Moreover, SAE-SS is able to detect IU's signals without any requirement for prior knowledge of IU's signals. Besides, the proposed method can complete the sensing task without requiring to conduct any arithmetical operations on the received signals.

REFERENCES

[1] W. y. Lee and I. F. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 10, pp. 3845–3857, October 2008.

[2] D. Wang, N. Zhang, Z. Li, F. Gao, and X. Shen, "Leveraging high order cumulants for spectrum sensing and power recognition in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1298–1310, Feb 2018.

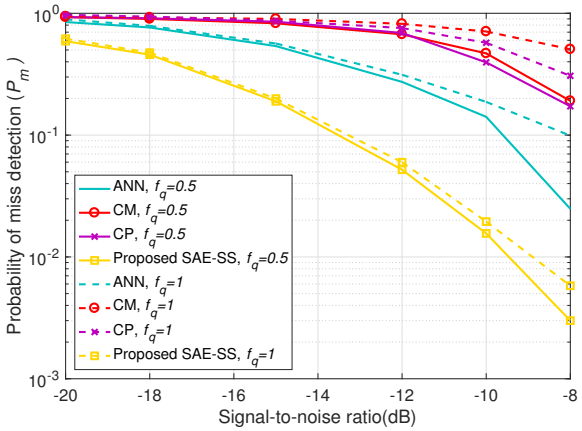


Fig. 7. Probability of miss detection of different sensing methods with CFO

[3] T. Riihonen and R. Wichman, "Energy detection in full-duplex cognitive radios under residual self-interference," in *2014 9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, June 2014, pp. 57–60.

[4] E. Axell and E. G. Larsson, "Optimal and sub-optimal spectrum sensing of ofdm signals in known and unknown noise variance," *IEEE Journal on Selected Areas in Communications*, vol. 29, pp. 290–304, Feb 2011.

[5] W. Xu, W. Xiang, M. Elkashlan, and H. Mehrpouyan, "Spectrum sensing of ofdm signals in the presence of carrier frequency offset," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 6798–6803, Aug 2016.

[6] Z. Lei and F. P. S. Chin, "Sensing ofdm systems under frequency-selective fading channels," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1960–1968, May 2010.

[7] K. M. Thilina, K. W. Choi, N. Saquib, and E. Hossain, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 11, pp. 2209–2221, November 2013.

[8] M. R. Vyas, D. K. Patel, and M. Lopez-Benitez, "Artificial neural network based hybrid spectrum sensing scheme for cognitive radio," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct 2017, pp. 1–7.

[9] O. Sigaud and A. Droniou, "Towards deep developmental learning," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 8, no. 2, pp. 99–114, June 2016.

[10] Z. Shao, L. Zhang, and L. Wang, "Stacked sparse autoencoder modeling using the synergy of airborne lidar and satellite optical and sar data to map forest above-ground biomass," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 12, pp. 5569–5582, Dec 2017.

[11] S. Lin and G. C. Runger, "Gcrnn: Group-constrained convolutional recurrent neural network," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 10, pp. 4709–4718, Oct 2018.

[12] D. P. Mandic and J. A. Chambers, "On the choice of parameters of the cost function in nested modular rnn's," *IEEE Transactions on Neural Networks*, vol. 11, no. 2, pp. 315–322, March 2000.

[13] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85 – 117, 2015.

[14] X. Zhang, G. Chen, W. Wang, Q. Wang, and F. Dai, "Object-based land-cover supervised classification for very-high-resolution uav images using stacked denoising autoencoders," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 7, pp. 3373–3385, July 2017.

[15] Y. Liu, W. Huangfu, H. Zhang, and K. Long, "An efficient stochastic gradient descent algorithm to maximize the coverage of cellular networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3424–3436, July 2019.