Machine Learning-Based Energy-Spectrum Two-Dimensional Cognition in Energy Harvesting CRNs

YONGJIAN FAN1, WENJUN XU2, (Senior Member, IEEE), CHIA-HAN LEE3, (Member, IEEE), SILEI WU2, (Graduate Student Member, IEEE), FAN YANG2, (Member, IEEE), AND PING ZHANG4

1School of Information and Electrical Engineering, Hebei University of Engineering, Handan 056038, China
2Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China
3Institute of Communications Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan
4State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

Corresponding author: Wenjun Xu (wjxu@bupt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61629101 and Grant 61771066, and in part by the Scientific and Technical Research Project of the Higher Education Institutions of Hebei under Grant ZD2016017.

ABSTRACT Energy harvesting cognitive radio network (EH-CRN) is a promising approach to address the shortage of spectrum resources and the increase of energy consumption simultaneously in wireless networks. In this article, we propose a novel machine learning (ML)-based energy-spectrum two-dimensional (2D) cognition technology to improve the sensing accuracy as well as the network throughput in EH-CRNs, which consists of sensing, prediction and decision modules. More specifically, we first study the 2D sensing module which is achieved by a carefully constructed dynamic Bayesian network (DBN) to effectively exploit the coupling between spectrum usage and energy harvesting in EH-CRNs. Then we propose a deep neural network (DNN) based 2D transmission decision module to optimize the transmission energy of secondary users (SUs). With our proposed novel 2D cognition scheme, SUs can characterize the energy-spectrum correlation and transmit data with optimal transmission energy. The proposed ML-based 2D cognition is evaluated via extensive simulations in terms of sensing accuracy, prediction accuracy, and network throughput, and simulation results indicate that our proposed scheme significantly outperforms the conventional one-dimensional (1D) cognition scheme working in spectrum or energy dimension only.

INDEX TERMS Machine learning, 2D cognition, energy-spectrum correlation, probability graph model, energy harvesting CRNs.

I. INTRODUCTION

With the rapid increase of service demands in wireless communications, the spectrum resource limitation has become a bottleneck problem. However, studies reveal that the utilization efficiency of most licensed bands is considerably low [1]. To tackle this apparent contradiction between the limited spectrum resource and the low spectrum utilization, researchers have proposed cognitive radio networks (CRNs) [2], which allow unlicensed secondary users (SUs) to cognize the occupancy status of licensed spectrum and access the idle spectrum to the primary users (PUs) to improve the spectrum efficiency. Meanwhile, along with the fast growth of wireless services, the energy consumption has also become a critical issue for wireless communication networks because of the increasing concerns on energy cost and carbon emission [3].

Aiming to conquer the double challenges in improving spectrum efficiency and reducing energy consumption, the energy harvesting CRN (EH-CRN) driven by renewable energy has received increasing attention from both academia and industry [3]–[8]. EH-CRNs are usually defined as CRNs wherein SUs perform energy harvesting powered by electromagnetic radiation, light, thermal gradients, or fluid flow from environment. For some EH-CRNs deployed in the harsh environment, where the grid-electricity-based power supply to PUs and SUs is unavailable, PUs will also harvest energy from environment to transmit data [6]. Hence both PUs and
SU{s}\textsuperscript{s} can perform energy harvesting in EH-CRN{s}, which may find interesting and critical applications in many practical scenarios. In this case, the licensed spectrum status (occupied or idle), is not only determined by whether PUs have data to transmit, but also depends on whether PUs have harvested enough energy. To this end, SUs shall carefully utilize the information in spectrum and energy dimensions jointly to estimate the licensed spectrum status with more efficient spectrum and energy utilization. This new problem is very challenging, which is the focus of our paper.

There are some traditional methods focusing on sensing the spectrum status in EH-CRN{s} \cite{4, 5}. For example, Chung et al. consider a fixed-length temporal sequential sensing system, which takes advantage of temporal correlation within the fixed duration in EH-CRN{s}, and the average throughput of the secondary network is maximized by balancing the sensing duration and the sensing threshold \cite{4}. In a paper by Olawole et al. \cite{5}, SUs at different locations implement a cooperative spectrum sensing scheme, which takes into account the spatial correlation among different SU{s}, and the network throughput is enhanced by employing a near-optimal mixed integer nonlinear sensing. However, traditional methods of using temporal correlation in data transmission of PUs and SUs separately \cite{4} cannot fully exploit the spatial correlations between PUs and SUs, while other methods exploiting spatial correlation \cite{5} are difficult to adapt to continuous changes with correlation in the time domain. It is noted that a lack of consideration for temporal-spatial correlations between PUs and SUs may lead to inaccurate sensing results. Therefore, novel schemes that can address this difficulty in characterizing the temporal-spatial correlation are called for to achieve more efficient EH-CRN{s}.

Recently, machine learning (ML) techniques have been applied to conventional CRNs and EH-CRN{s} to exploit the temporal-spatial correlations between PUs and SUs. Some ML-based temporal-spatial sensing methods are discussed in conventional CRNs to promote the sensing performance. For example, Qiu et al. apply recurrent neural network (RNN) over temporal-spatial model to improve the cognition accuracy and traffic prediction performance of CRNs \cite{9}. In a paper by Han et al. \cite{10}, the licensed spectrum status is obtained from both spatial and temporal domains by using Bayesian network (BN). Moreover, there are some emerging works focusing on temporal-spatial sensing in EH-CRN{s}. Specifically, some reinforcement learning (RL) based energy harvesting resource allocation schemes are proposed to enhance the capacity of the EH-CRN{s} \cite{7, 8}. However, the works in \cite{7} and \cite{8} do not make use of the potential correlation between PUs and SUs in the energy dimension. Note that such correlation is an inherent but possible hidden behavior in the process of energy harvesting, which requires an in-depth mining to grasp the essential mechanism. It is worth pointing out that the licensed spectrum access of PUs is constrained by their harvested energy, and hence the spectrum cognition capability and spectrum utilization can be improved by finely mining the correlation of harvested energy between PUs and SUs. However, most existing ML-based sensing methods \cite{7}–\cite{10} are not capable of jointly characterizing temporal-spatial correlations in both energy and spectrum dimensions. To overcome this difficulty, dynamic Bayesian network (DBN) is introduced in EH-CRN{s}, which bears a potential to jointly characterize the information from spectrum and energy dimensions \cite{11}. In the work of \cite{11}, Zhang et al. first integrate the correlations between PUs and SUs in energy and spectrum dimensions to improve the sensing accuracy in EH-CRN{s} by applying DBN. However, the superiority in spectrum status prediction and transmission decision for the future slot of cognition is not considered. In fact, the spectrum prediction and the transmission decision for the future slot are critical for efficient spectrum utilization. To elaborate further, if the SU{s} transmit data only based on the sensing results, the duration of transmission time will be limited in the current slot and a large amount of sensing overhead is not fully utilized. In addition, it is assumed that SU{s} in \cite{11} can only use the energy harvested at the moment to transmit data without energy buffer to restore energy, which is not practical and may cause extra energy waste as compared to energy storage scenarios. To sum up, the existing works with ML technique in EH-CRN{s} cannot achieve comprehensive energy-spectrum two-dimension (2D) cognition, which should consider the energy storage and the spectrum access in the forthcoming transmission slots for enhanced spectrum utilization with smart and proactive decision. To this end, there is an urgent need to investigate a comprehensive ML-based energy-spectrum 2D cognition scheme that integrates spectrum status sensing and prediction for transmission decision of SU.

In this article, we consider an EH-CRN scenario, where both PUs and SU{s} perform energy harvesting equipped with energy buffer for practical energy storage. In particular, we apply ML technique in energy-spectrum 2D cognition, which can effectively exploit the correlation between spectrum usage and energy harvesting in EH-CRN{s}. Our proposed energy-spectrum 2D cognition scheme integrates spectrum status sensing and prediction for transmission decision of SU, which consists of the sensing, prediction and decision modules, respectively, to improve the sensing and prediction accuracy and the network throughput. To be specific, the contributions of this article are as follows.

1) We design a new energy-spectrum 2D cognition scheme, which contains not only the spectrum sensing module but also the state prediction module and the transmission energy decision module. Different from the existing cognition schemes that only contain sensing module \cite{7, 8, 10, 11}, our proposed 2D cognition scheme is more comprehensive and practical. In particular, the prediction module and the decision module are complementary to spectrum sensing for predicting the licensed spectrum status by exploiting the observations of SU{s} and estimations of PUs, thus making a smart and proactive decision on the transmission energy of SU{s} for enhanced performance.
We propose a DBN-based energy-spectrum 2D sensing module, which exploits 2D correlations between PUs and SUs, and estimates the harvested energy and the transmission energy of PUs at the current slot. Note that the correlations are not mined in traditional 1D sensing [4], [5] or existing ML-based temporal-spatial sensing [7]–[10] schemes. Compared with these existing works, our proposed DBN-based 2D sensing characterizes the energy-spectrum 2D correlations by using the hidden Markov model (HMM), a graphic realization of DBN, which effectively reveals the inherent behaviors of this complicated correlations. Empowered by DBN-based 2D sensing, the sensing accuracy can be significantly improved.

To further enhance network throughput, we propose a 2D prediction module to predict the transmission energy of PUs, the sensing signal of SUs, and the harvested energy of PUs and SUs in the forthcoming slots, and further propose a ML-based 2D decision module to determine the transmission energy of SUs by using prediction results. Different from the previous works that SUs always transmit data with a greedy strategy to exhaust the harvested energy in each slot [11], SUs in our scheme can dynamically optimize energy storage and transmission in the forthcoming slots to achieve better spectrum utilization.

The remainder of this article is organized as follows. Section II describes the system model of EH-CRN. In Section III, problem formulation in an energy-spectrum 2D cognition scheme of EH-CRN is analyzed, which contains the sensing, prediction, and decision modules. In Section IV, the proposed ML-based energy-spectrum 2D cognition algorithm is studied. Simulation results are shown in Section V. In Section VI, conclusions and future research directions are discussed.

II. SYSTEM MODEL
As illustrated in Fig. 1, we consider an energy-spectrum joint cognitive system in an EH-CRN. In the following part of this section, the network structure, spectrum sensing model, energy harvesting correlation, prediction and decision for transmission of SU are presented.

A. NETWORK STRUCTURE
As shown in Fig. 1, there is a primary network composed of a primary base station (BS) and an energy harvesting PU, and a secondary network composed of a secondary BS and an energy harvesting SU. The primary BS and secondary BS receive the data transmitted by PU and SU, separately. Both PU and SU perform energy harvesting to collect energy for transmission. The spectrum is licensed to the PU, thus SU needs to detect spectrum status to access the idle spectrum. Besides, both PU and SU are equipped with energy batteries as energy buffer to restore the harvested energy. Within our considered scenario, the licensed spectrum is occupied, which is not only related to whether the PU has data to transmit, but also to whether the harvested energy of PU is enough for transmission. In this case, the SU has to sense both the transmitted signal of PU in the spectrum dimension and the harvested energy of PU in energy dimension, which motivates the energy-spectrum 2D cognition of SU.

Figure 1 illustrates the working process of an EH-CRN with energy-spectrum 2D cognition. It is noted that both PU and SU harvest energy from the environment for data transmission. If there is a transmission mission in the primary network, the PU will transmit data as soon as enough energy is harvested. For SU, when there is a transmission mission in the secondary network, it will transmit data only if the licensed spectrum status is idle. Therefore, the SU performs 2D sensing to jointly estimate the spectrum utilization and the harvested energy of PU at the current slot. Armed with the sensing results, the prediction and decision for the future slot are possible, which can facilitate the transmission of SU.

| Table 1. Notations. |
|---------------------|
| Sym. | Definition | Sym. | Definition |
| $s_t$ | transmitted signal of PU | $\lambda_1$ | threshold of the spectrum sensing |
| $r_t(n)$ | the $n$th sample of received signal of SU | $\lambda_2$ | threshold of the energy sensing |
| $r_t(n)$ | noise of the $n$th sample of received signal of SU | $\epsilon_1$ | convergence threshold of DBN |
| $h_{1,1}$, $E_{1,1}^{PU}$, $E_{1,1}^{SU}$ | channel coefficients of PU to SU, PU to BS, SU to BS | $c_2$ | convergence threshold of DNN |
| $Y_t$ | detection of transmitted signal from PU | $P_d$ | detection probability |
| $H_{0,1}$, $H_{1,1}$ | binary hypothesis test of spectrum status | $P_f$ | false alarm probability |
| $E_{t,1}^{PU}$, $E_{t,1}^{SU}$ | transmission energy at PU or SU, PU, and SU | $P_m$ | miss detection probability |
| $P_{1,1}^{PU}$ | harvested energy of PU and SU | $R_1$, $R_{1,1}^{PU}$, $R_{1,1}^{SU}$ | transmission rates of PU or PU, PU, and SU |
| $d_{1,1}^{PU}$, $d_{1,1}^{SU}$ | spectrum access indexes of PU and SU | $H_{1,1}^{PU}$, $H_{1,1}^{SU}$ | hidden states of PU in $X_s$, where $s = 1, 2$ |
| $\Delta t$ | duration of data transmission | $O_{1,1}^{PU}$, $O_{1,1}^{SU}$ | observed states of SU in $X_s$, where $s = 1, 2, 3$ |
| $B$ | system bandwidth | $D_T$ | decision on transmission energy of SU |
| $N_0$ | thermal noise power density | $A$, $B$, $H_{1,1}^{PU}$, $C$, $D$, $H_{1,1}^{SU}$ | state transition matrices of DBN |
| $K_t$ | size of data packet of SU at slot $t$ | $a_{i,j}$, $b_{i,j,m}$, $c_{i,j}$, $d_{i,j,m}$ | transition probabilities of DBN |
| $X_{1,1}^r$ | the $r$th node in the HMM graph | $N_d$ | discretization levels of the decision of SU $D_T$ |
| $N_1$ | discretization levels of the $n$th node in the HMM graph | $W_{DBN}$, $W_{DNN}$ | set of parameters of DBN, DNN |
| $B_{1,1}$ | initial distribution of the states of DBN | $K_1$, $K_2$, $K_3$ | number of nodes of DBN in input, hidden, and output layers |
| $R$ | iteration time in the training of DBN | $l$ | learning rate of DNN |
B. SPECTRUM SENSING MODEL

Spectrum sensing is based on the detection of the signals transmitted from the PU through the local observations of SU, wherein matched filter detection, energy detection, and cyclostationary feature detection are among the popular solutions [2]. Energy detection is a method with simple operation and high detection accuracy [12], which is adopted in our spectrum sensing module of SU. The energy detection of the transmitted signal from PU is compared with a given detection threshold to make a judgement on whether the PU transmits data at the current slot. Specifically, let $s_t$ denote the transmitted signal of PU and $r_t(n)$ denote the received signals as observations with $N$ samples at slot $t$, where $n = 1, 2, \ldots, N$. The detection of the transmitted signal at SU can be presented as

$$Y_t = \sum_{n=1}^{N} \|r_t(n) - u_t\|^2 = \sum_{n=1}^{N} \|h_t \cdot s_t + n_t(n) - u_t\|^2, \quad (1)$$

where $n_t(n)$ is the additive white Gaussian noise (AWGN) of the $n$th sampling point at slot $t$, $u_t$ is the transmitted signal of SU that is known by SU, and $h_t$ is the channel coefficient from PU to SU that is not varying across the $N$ sampling points. Assuming the number of sample $N$ is sufficiently large, the sum energy of the detection signal at SU $Y_t$ follows a Gaussian distribution, and the mean of distribution is related to the transmitted signal of PU $s_t$. It is worth noting that the detection energy is different from the harvested energy. The amount of detection energy is very small and cannot be used for the transmission of SU [12].

C. ENERGY HARVESTING CORRELATION

Both PU and SU need to harvest energy from the environment for data transmission. Therefore, both PU and SU are equipped with energy harvesting devices to perform energy harvesting. The harvested energy at the geographically adjacent PU and SU are strongly correlated, since the environment presents a great deal of continuity both in time and space.

In the time domains, given the fact that the amount of energy harvested is closely related to the ambient environment, e.g., solar energy depends on weather conditions and wind energy depends on wind speed, the amount of energy harvested varies across multiple timescales. The temporal correlation in energy dimension is well exploited by using time-series data from the previous slots. For example, the harvested energy may follow Markov distribution for time horizon from several milliseconds up to a few minutes [13].

The spatial correlation between PU and SU in energy dimension can be further observed by using some existing datasets. For example, one dataset [14] from four photovoltaic (PV) sites is utilized here for illustrative purpose. Similar to the energy harvesting PUs and SUs, the PV sites are used for energy harvesting, the spatial correlation patterns among these PVs may show similarity to those between the PUs and SUs, thus offering valuable insights to our work. More specifically, we have analyzed the relationships on the harvested energy between each two geographically neighboring PV sites. Enumerating all the pair-wise relationships among the four PV sites, there are six pairs of energy harvesting relationships, and the approximately linear correlations are observable as shown in Fig. 2. The above observations offer valuable insights. The temporal and spatial correlations of energy harvesting can be exploited in EH-CRNs since PU and SU are adjacent to each other. The correlations will provide additional information about PU activities to enable SU to utilize the licensed spectrum efficiently.
where $d_t^{PU}$ and $d_t^{SU}$ are the spectrum access indexes of PU and SU, which indicate whether the PU or SU access the licensed spectrum and can only take the value 0 or 1.

From (3), the instantaneous rate and the transmission energy of SU show a logarithmic relationship, which means that the increase of the instantaneous rate will gradually slow down as the transmission energy keeps increasing. Meanwhile, reserving the energy of SU that is not used for transmission at the current slot can increase the available energy for the forthcoming transmission of SU to improve the network throughput. Note that greedily increasing the transmission energy of SU can maximize the instantaneous rate of SU but may decrease the network throughput. Therefore, a smart and proactive decision on the transmission energy of SU is necessary to balance the increase of the transmission energy and the reservation of more available energy for the future transmission.

To achieve the transmission decision of SU, the spectrum status should be accurately sensed and predicted, and the transmission energy should be optimized. The problems are formulated in a comprehensive cognition model in the next section.

### III. PROBLEM FORMULATION IN 2D COGNITION SCHEME

In this section, we formulate the problem in our new energy-spectrum 2D cognition scheme of EH-CRN, which contains the sensing, the prediction and the decision modules, respectively. In particular, the sensing and prediction accuracy, which is evaluated by detection probability and false alarm probability, is enhanced in our proposed sensing and prediction modules, respectively, and the network throughput of EH-CRN is improved by optimizing transmission energy of SU in the decision module.

#### A. 2D SENSING MODULE

Sensing module is to detect whether the licensed spectrum is accessed by PU, which can be expressed as a binary hypothesis test $H_0(\text{idle})$ or $H_1(\text{occupied})$. In order to evaluate the sensing accuracy of spectrum status, the detection probability $P_d$ and the false alarm probability $P_f$ are introduced. The detection probability $P_d$ represents the probability that PU accesses the licensed spectrum and meanwhile SU detects the existence of the signal of PU in the spectrum. The false alarm probability $P_f$ represents the probability that PU does not access the licensed spectrum while SU detects the existence of the signal of PU in the spectrum. In a sensing problem, it is expected that $P_d$ is as large as possible when $P_f$ is a constant, and $P_f$ is as small as possible when $P_d$ is constant.

According to (1), the sensing signal in the licensed spectrum $Y_t$ is obtained. The 1D spectrum sensing problem can be formulated by maximizing the conditional probability of the transmitted signal of PU $s_t$ under the condition of the sensing
signal $Y_0, \ldots, Y_t$ as

$$P_1 : \max_{s_t \geq 0} P(s_t|Y_0, \ldots, Y_t).$$

(5)

Considering the 1D spectrum sensing accuracy, the detection probability $P_d$ and the false alarm probability $P_f$ can be expressed as

$$P_d^S = P(s_t > \lambda_1|H_1),$$

(6)

$$P_f^S = P(s_t > \lambda_1|H_0),$$

(7)

where $\lambda_1$ is the threshold of the spectrum sensing.

According to the temporal and spatial correlations in energy dimension, there is a linear relationship between the harvested energy of PU $E^{PU}_t$ and SU $E^{SU}_t$, and the previous harvested energy at SU $E^{SU}_{t-1}$ under condition of the harvested energy at SU $E^{SU}_0, \ldots, E^{SU}_t$. The 1D energy sensing problem can be formulated by maximizing the conditional probability of the harvested energy at PU $E^{PU}_t$ under condition of the harvest energy at SU $E^{SU}_0, \ldots, E^{SU}_t$ as

$$\max_{E^{PU}_t \geq 0} P(E^{PU}_t|E^{SU}_0, \ldots, E^{SU}_t).$$

(8)

Considering the 1D energy sensing accuracy, the detection probability $P_d$ and the false alarm probability $P_f$ can be expressed as

$$P_d^E = P(E^{PU}_t > \lambda_2|H_1),$$

(9)

$$P_f^E = P(E^{PU}_t > \lambda_2|H_0),$$

(10)

where $\lambda_2$ is the threshold of the energy sensing. To sum up, the above 1D spectrum sensing and 1D energy sensing can be achieved by using the observable information at SU in spectrum or energy dimension. However, both of the 1D spectrum sensing and the 1D energy sensing do not consider the correlation information from each other.

If the spectrum sensing and energy sensing are jointly considered, the SU will detect the licensed spectrum more accurately and comprehensively. Therefore, the energy-spectrum 2D sensing is proposed to jointly exploit the observable information at SU in spectrum and energy dimensions. More specifically, the transmitted signal of PU $s_t$ and harvested energy of PU $E^{PU}_t$ are jointly estimated by utilizing the sensing signal of SU $Y_0, \ldots, Y_t$ and the harvest energy of SU $E^{SU}_0, \ldots, E^{SU}_t$. The energy-spectrum 2D sensing problem can be formulated by maximizing the joint conditional probability as

$$\max_{s_t \geq 0, E^{PU}_t \geq 0} P(s_t, E^{PU}_t|Y_0, E^{SU}_0, \ldots, Y_t, E^{SU}_t).$$

(11)

Furthermore, the detection probability $P_d$ and the false alarm probability $P_f$ of 2D sensing can be expressed as

$$P_d = P(s_t > \lambda_1, E^{PU}_t > \lambda_2|H_1),$$

(12)

$$P_f = P(s_t > \lambda_1, E^{PU}_t > \lambda_2|H_0).$$

(13)

It is noted that the traditional optimization methods in 1D spectrum or energy sensing are difficult to solve the 2D sensing problem without the prior information of the spectrum-energy 2D correlation. Specifically, the 1D sensing methods can estimate the transmitted signal of PU $s_t$ and the harvested energy of PU $E^{PU}_t$ according to (5) and (8), respectively. However, the transmitted signal of PU $s_t$ in the spectrum dimension and the harvested energy of PU $E^{PU}_t$ in the energy dimension are not independent of each other. The harvested energy at SU $E^{SU}_t$ can provide energy-spectrum correlation information to PU $s_t$. Moreover, the correlation coefficient between the conditional probabilities $P(s_t|Y_0, \ldots, Y_t)$ and $P(E^{PU}_t|E^{SU}_0, \ldots, E^{SU}_t)$ is not a constant, but a multi-dimensional function based on $s_t$ and $E^{PU}_t$. Without the prior information of the coefficient, the traditional methods can only assume the correlation function in a specific form, which lacks of generalization ability.

Thanks to the ML method DBN, the 2D correlation can be characterized automatically and intelligently without any prior information of the correlation. DBN is based on the probability graph model, and uses the transition probabilities as parameters to characterize the relationships among the nodes in the graph, which is trained by statistics of historical data. On the basis of this feature, we utilize DBN to connect the observation information of $E^{SU}_t$ and $Y_t$ to the detected information of $E^{PU}_t$ and $s_t$ through the state transition matrices. The matrices consist of the conditional probabilities between the observation information at SU and the estimated information of PU. The conditional probabilities can be obtained by the training of DBN without the prior information about correlation. In this way, the transmitted signal of PU $s_t$ and the harvested energy of PU $E^{PU}_t$ can be jointly estimated by DBN in a more accurate and generalized way, which will be discussed in detail in Section IV. With the 2D sensing results of the unobservable information, all information of PU and SU at the current slot can be obtained. By using the known information at the current slot and the energy-spectrum correlation characterized in DBN, the information of PU and SU in the forthcoming slots can be predicted in the 2D prediction module. The licensed spectrum status can be estimated as well through further operations.

### B. 2D PREDICTION MODULE

In the 2D cognition scheme of EH-CRN, 2D prediction module is cascaded with 2D sensing module to achieve the prediction of the states of PU and SU in spectrum and energy dimensions in the forthcoming slots. Specifically, the sensing signal of SU $Y_t$ and the harvested energy of SU $E^{SU}_t$ are observed at SU as the observation information, and the transmitted signal of PU $s_t$ and harvested energy of PU $E^{PU}_t$ are estimated in the 2D sensing module as the sensing results. By using the observation information and the sensing results at the current slot, 2D prediction module aims to predict the harvested energy at PU $E^{PU}_{t+n}$, the transmitted signal by PU $s_{t+n}$, the spectrum sensing signal by SU $Y_{t+n}$, the harvested energy at SU $E^{SU}_{t+n}$, and the residual energy in the buffer of SU $E^{BU}_{t+n}$, respectively, where $n$ denotes the forthcoming time slot and $n > 0$. In particular, one-slot look-ahead prediction is applied when $n = 1$. 
On the other hand, the energy-spectrum 2D correlation has already been characterized by using the conditional transition probabilities in the sensing module. The forthcoming states of PU and SU can be predicted by maximizing the conditional transition probabilities under the condition of the states in the current slot. The problem of 2D prediction based on the conditional transition probabilities can be formulated as:

\[
P_2: \max_{s_{t+n}} P(s_{t+n}|s_t, Y_t, E_{t+1}^{PU}, E_{t+1}^{SU}),
\]

\[
P_3: \max_{E_{t+1}^{PU}} P(E_{t+1}^{PU}|s_t, Y_t, E_{t+1}^{PU}, E_{t+1}^{SU}),
\]

\[
P_4: \max_{Y_{t+n}} P(Y_{t+n}|s_t, Y_t, E_{t+1}^{PU}, E_{t+1}^{SU}),
\]

\[
P_5: \max_{E_{t+1}^{SU}} P(E_{t+1}^{SU}|s_t, Y_t, E_{t+1}^{PU}, E_{t+1}^{SU}),
\]

\[
P_6: \max_{E_{t+1}^{SU}} P(E_{t+1}^{SU}|s_t, Y_t, E_{t+1}^{PU}, E_{t+1}^{SU}).
\]

Note that all of the conditional transition probabilities in the above problems \((14) - (18)\) are obtained by DBN in the sensing module, which will be detailed in the Section IV later. Similar to the evaluation criteria of the sensing module, the evaluation criteria of the prediction module can also use the detection probability \(P_d\) and the false alarm probability \(P_f\) to represent the prediction accuracy. In particular, the detection probability \(P_d\) represents the probability that the prediction is correct when the licensed spectrum status is occupied, while the false alarm probability \(P_f\) represents the probability that the prediction is wrong when the licensed spectrum status is occupied, which can be expressed as follow:

\[
P_d = P(s_{t+n} > \lambda_1, E_{t+1}^{PU} > \lambda_2|H_1),
\]

\[
P_f = P(s_{t+n} > \lambda_1, E_{t+1}^{PU} > \lambda_2|H_0).
\]

It should be pointed out that the estimation of the future spectrum status only depends on the prediction results of PU \(s_{t+n}\) and \(E_{t+1}^{PU}\), while all of prediction results of PU and SU are used as the input for the spectrum-energy 2D transmission decision of SU in 2D decision module following.

### C. 2D DECISION MODULE

Through 2D sensing and prediction modules, the states of PU and SU in spectrum and energy dimensions are obtained. 2D decision module is the last step in the energy-spectrum 2D cognition scheme, which uses the prediction results to make a decision on how much energy the SU allocating to access the spectrum. In conventional CRN, SU is usually assumed to be stably powered, and thus SU transmits data with the maximum transmission energy when a spectrum hole is detected. In EH-CRN, the transmission energy of SU is limited by the amount of energy harvested. 2D transmission decision can improve the network throughput in two aspects. First, the better sensing performance of 2D sensing provides the more opportunity for SU to access the idle spectrum with the protection of PUs. Secondly, 2D transmission decision can take advantage of and adapt to the random energy arrivals more effectively, because it can make full use of historical observed information about both spectrum and energy.

Our objective is to maximize the network throughput, which drives SU to increase the transmission energy for the improvement of the instantaneous throughput, and meanwhile to reserve more available energy in the buffer for the future transmission. Therefore, the optimal transmission energy of SU needs to balance the improvement of the instantaneous throughput and the future transmission. Assume that the size of data packet to be transmitted by SU at slot \((t+n)\) is \(K_{t+n}\). When the instantaneous transmission rate is sufficient to complete the data packet transmission, the transmission rate does not need to be increased. The optimal energy allocation problem at each forthcoming slot can be formulated as:

\[
\max_{d_{t+n}^{PU}, d_{t+n}^{SU}} d_{t+n}^{PU} R_{t+n}^{PU} + d_{t+n}^{SU} R_{t+n}^{SU},
\]

s.t.

\[
C_1: d_{t+n}^{PU} R_{t+n}^{PU} + d_{t+n}^{SU} R_{t+n}^{SU} \leq K_{t+n}. \tag{22}
\]

\[
C_2: E_{t+n}^{SU} \leq E_{t+n}^{PU} + E_t^{R}. \tag{23}
\]

where

\[
R_{t+n}^{PU} = B \log_2 \left( 1 + \frac{E_{t+n}^{PU} \|h_{t+n}^{PU}\|^2}{N_0 B \Delta t + E_{t+n}^{SU} \|h_{t+n}^{SU}\|^2} \right), \tag{24}
\]

\[
R_{t+n}^{SU} = B \log_2 \left( 1 + \frac{E_{t+n}^{SU} \|h_{t+n}^{SU}\|^2}{N_0 B \Delta t + E_{t+n}^{PU} \|h_{t+n}^{PU}\|^2} \right). \tag{25}
\]

Here, \(E_{t+n}^{PU}\) and \(E_{t+n}^{SU}\) are the transmission energy of PU and SU, \(d_{t+n}^{PU}\) and \(d_{t+n}^{SU}\) are the spectrum access indexes of PU and SU, which indicate whether the PU or SU access the licensed spectrum and can only take the value 0 or 1, \(R_{t+n}^{PU}\) and \(R_{t+n}^{SU}\) are the instantaneous rate of PU and SU, and \(E_t^{R}\) is the residual energy at the beginning of the slot. Constraint (22) shows that SU will not increase the transmission energy as the rate is sufficient for data packet transmission in the slot. Constraint (23) is that the amount of the transmission energy is constrained by the sum of the harvested energy in the slot and the residual energy at the beginning of the slot.

In 2D transmission decision, the spectrum access index of PU \(d_{t+n}^{PU}\) cannot be accurately known and it is only inferred through sensing and prediction results. Furthermore, although the size of transmitted data packets \(K_{t+n}\) and the channel coefficients \(h_{t+n}^{PU}\) and \(h_{t+n}^{SU}\) can be measured and recorded at slot \((t + n)\), they are difficult to be obtained accurately in advance for the future decision at the current slot. Due to the fact that \(h_{t+n}^{PU}\), \(h_{t+n}^{SU}\), and \(K_{t+n}\) cannot be accurately obtained as prior information, the traditional convex optimization method cannot solve the optimal energy allocation problem. ML methods have the ability to connect input and output without knowing the prior information from input to output, but using the neural network parameters to fit the transition process. To reduce the impact of uncertainty on the spectrum access index \(d_{t+n}^{PU}\), the size of data packets \(K_{t+n}\), and the channel coefficients \(h_{t+n}^{PU}\), \(h_{t+n}^{SU}\), DNN-based transmission decision module is proposed, which establishes
the relationship between the prediction results and the optimal transmission energy without additional prior information about $a_{t+n}^{PU}$, $K_{t+n}^{PU}$, $h_{t+n}^{SU}$. The DNN-based transmission decision is detailed in the following section to maximize the network throughput.

**IV. ML-BASED ENERGY-SPECTRUM 2D COGNITION ALGORITHM**

To solve the problems of each module in 2D cognition, ML-based energy-spectrum 2D cognition algorithm is proposed. Specifically, energy-spectrum 2D correlations are characterized by DBN, the transmitted signal of PU $s_t$ and the harvested energy of PU $E_t^{PU}$ are jointly estimated by DBN in the sensing module, the states of PU and SU in the forthcoming slots are predicted in the prediction module, and the transmission energy of SU are optimized by DNN in the decision module.

**A. OVERVIEW ON ML-BASED ENERGY-SPECTRUM 2D COGNITION ALGORITHM**

In this subsection, the ML-based energy-spectrum 2D cognition is described in details. Fig. 3 illustrates the structure of energy-spectrum 2D cognition, which contains the interactive relationship among the offline training modules and the online cognition modules.

![FIGURE 3. Structure of ML-based energy-spectrum 2D cognition.](image)

Historical information consists of PU and SU states, sizes of file and channel conditions in EH-CRN is used for the training of DBN in the sensing module and DNN in the decision module. The historical information is recorded at PU, SU and BS with real data at consecutive time slots. To elaborate further, the harvested energy and the transmission energy of PU are recorded at PU, the spectrum sensing signal, the harvested energy, the size of data packets SU are recorded at SU, and the channel conditions of PU and SU are recorded at BS. It is noted that the recorded information of PU cannot be used as the live information at SU in the online cognition, because the shortage of spectrum resources and the lack of stable energy supply limit the information communication between PU and SU.

Both offline training and online cognition based on DBN and DNN are carried out at SU. On the offline side, DBN is trained by the historical data of PU and SU states, and DNN is trained by all of the historical data consists of the states of PU and SU, the sizes of transmitted data packets and the channel coefficients of PU and SU in EH-CRN. With the trained DBN and DNN, energy-spectrum 2D correlations and the transmission decision policy of SU are obtained, which are further input into the online cognition. On the online side, the spectrum sensing signal and the harvested energy of SU at the current slot are detected by SU as the current observation. DBN based 2D sensing module estimates the hidden states of the harvested energy and the transmission energy of PU through the current observation. The future states of PU and SU are predicted based on the current observation and the current estimation of hidden states in 2D prediction module. The future states are input into the 2D decision module, and the transmission energy of SU are output. Finally, the SU will transmit the data packet with the optimal transmission energy output from DNN.

**B. 2D CORRELATIONS CHARACTERIZED BY DBN**

In this subsection, we adopt DBN to exploit the energy-spectrum 2D correlations during offline training and achieve 2D sensing during online computing. We first adopt a HMM structure which is the probabilistic graphical model of DBN to model the process of energy-spectrum 2D cognition. The 2D correlations are represented by the state transition matrices in the HMM structure. Note that the transition probabilities in the state transition matrices are the parameters of DBN, which are trained based on historical data. Based on the trained DBN, we propose an online algorithm to solve the 2D sensing problem.

1) **THE PROBABILISTIC GRAPHICAL MODEL OF DBN**

We adopt DBN to model energy harvesting and spectrum access of PU, the correlation between the harvested energy of SU and PU, and the correlation between transmitted signals of PU and spectrum detection of SU. Markovian behavior of PU is widely assumed in sensing the spectrum hole (e.g., in [11], [13]), and the energy arrivals of both PU and SU are modeled as finite-state Markov chains since finite-state Markov chains strike good balance between model complexity and accuracy for renewable energy harvesting. HMM structure is adopted as the probabilistic graphic of DBN to divide the states in the finite-state Markov chains into hidden states and observed states, as shown in Fig. 4(a).

In EH-CRN, the SU occupies the licensed spectrum only when the PU does not access. However, the harvested energy and the transmission energy of PU cannot be directly observed at SU to confirm the spectrum status directly. Furthermore, the shortage of spectrum resources and the lack of stable energy supply limit information interaction between
PU and SU. Therefore, the SU could only estimate the hidden states of PU with observations at SU. Specifically, in the proposed HMM structure, the harvested energy and the transmission energy at PU are the hidden information of SU, while the spectrum sensing signal, the harvested energy and the residual energy at SU are the observations, and the detailed modeling is discussed as follows.

The proposed HMM structure is composed of the pair of distribution $(B_0, B_→)$, where $B_0$ is an initial distribution as shown in Fig. 4, and the residual energy at SU is assumed to be empty at the beginning. $B_→$ is a two-slice temporal Markovian chain as shown in Fig. 4(b), which defines the transition probabilities $P (X_t|X_{t-1})$ by means of a directed acyclic graph with the corresponding transition probabilities as

$$P (X_t|X_{t-1}) = \prod_{s=1}^{N} P (X_s^t|\text{Pa}(X_s^t)).$$

In (26), $X_s^t \in \{H_t^T, O_t^R, O_t^F, O_t^E\}$ is the $s$th node from the top to the bottom of the Markovian chain in $B_→$ at slot $t$, where $s = 1, 2, 3, 4, 5$. $H_t$ denotes the states of the hidden information and $O_t$ denotes the states of the observed information. To elaborate further, $H_t^T$ and $H_t^E$ denote the harvested energy at PU and the transmission energy at PU at time slot $t$, which are the hidden states from the perspective of SU. Here $O_t^E, O_t^F, O_t^R$ denote the spectrum sensing at SU, the harvesting energy at SU, and the residual energy at SU, which are the observed states of SU. To ensure that the number of states is finite, the values of states $X_s^t$ are discretized into $N_s$ different levels from 0 to $(N_s - 1)$. HMM can well represent the correlations between PU and SU in both energy and spectrum dimensions by using the transition probabilities among the states. $\text{Pa}(X_s^t)$ is the parent of $X_s^t$ in Fig. 4(b). Particularly, both the state of transmission energy at PU and the state of harvested energy at SU have two parents, and hence their current states depend on not only their own states at the previous slots but also the parent states at the current slot.

The transition matrices contain the spatial and temporal correlation information in energy and spectrum dimensions. The transition matrix of the harvested energy state at PUs $A \in \mathbb{R}^{N_1 \times N_1}$, the transition matrix of the transmission energy state at PUs $B(H_t^E) \in \mathbb{R}^{N_1 \times N_2}$, the transition matrix of the spectrum sensing power state at SUs $C \in \mathbb{R}^{N_2 \times N_3}$, and the transition matrix of the harvested energy state at SUs $D(H_t^E) \in \mathbb{R}^{N_3 \times N_4}$ are given as

$$A = \{a_{i,j} = P(H_{t-1}^E = i|H_t^E = j)|0 \leq a_{i,j} \leq 1, i, j = 0, 1, \ldots, N_1\},$$

$$B(H_t^E) = \{b_{i,j,m} = P(H_{t-1}^T = i|H_t^E = j, H_{t-1}^E = m)|0 \leq b_{i,j,m} \leq 1, i, j, m = 0, 1, \ldots, N_2\},$$

$$C = \{c_{i,j} = P(O_t^F = i|O_t^E = j)|0 \leq c_{i,j} \leq 1, i, j = 0, 1, \ldots, N_3\},$$

$$D(H_t^E) = \{d_{i,j,m} = P(O_t^E = i|O_t^E = j, H_t^E = m)|0 \leq d_{i,j,m} \leq 1, i, j, m = 0, 1, \ldots, N_4\},$$

where $A, B(H_t^E), C, D(H_t^E)$ represent temporal correlation in energy dimension, the correlation between energy and spectrum dimensions, the spatial correlation in spectrum dimension, and the temporal-spatial correlation in energy dimension, respectively. It is noted that the transition matrix of the transmitted power state at PUs $B(H_t^E)$ and the transition matrix of the harvested energy state at SUs $D(H_t^E)$ are the function of $H_t^E$, because the state transition of the transmission energy at PU $H_t^E$ and the state of the harvested energy at SU $O_t^E$ depend on multiple parent nodes and cannot be represented by one constant matrix. Here, $a_{i,j}, b_{i,j,m}, c_{i,j}, d_{i,j,m}$ are the specific conditional transition probabilities in the transition matrices, which are also the parameters of DBN $W_{DBN} = \{a_{i,j}, b_{i,j,m}, c_{i,j}, d_{i,j,m}\}$. Next, these parameters will be calculated by offline training algorithm of DBN.

2) OFFLINE TRAINING OF DBN
After establishing the probabilistic graph of DBN, we propose an offline training algorithm to calculate the parameters of DBN. Note that the parameters of DBN $W_{DBN} = \{a_{i,j}, b_{i,j,m}, c_{i,j}, d_{i,j,m}\}$ contain the transition probabilities of each state in the probability graph model HMM, and the
values of the parameters all range from 0 to 1. The parameters can be trained by using the iterative parameter estimation [11]. To be more specific, let $W_{DBN}^k$ denote the estimated parameters in the $k$th iteration. The initial parameters within $W_{DBN}^0$ in the iteration $k = 0$ are randomly initialized. Then, in each iteration, the newly estimated parameters $W_{DBN}^k$ are calculated by historical states and previous parameters $W_{DBN}^{k-1}$. The iteration process continues until the parameters converge.

In particular, the expectation of log-likelihood function $L(W_{DBN}^k; W_{DBN}^{k-1})$ is used to evaluate the gap between the parameters of the $k$th and the $(k-1)$th iteration, and the expected log-likelihood is non-decreasing during the iteration algorithm. The estimated parameters in $W_{DBN}^k$ are obtained by maximizing the expectation of log-likelihood function $L(W_{DBN}^k; W_{DBN}^{k-1})$ in each iteration when the index of iteration $k > 1$, which can be expressed as

$$
max_{W_{DBN}} L(W_{DBN}^k; W_{DBN}^{k-1}) = \max_{W_{DBN}} \sum_{i=1}^T \sum_{j=1}^T P(H_i^E, H_i^T, O_i^S, O_i^F | W_{DBN}^{k-1}) \times \log P(H_i^E, H_i^T, O_i^S, O_i^F | W_{DBN}^k)
$$

(31)

where $W_{DBN}^k = \{a_{i,j}^k, b_{i,j,m}^k, c_{i,j}^k, d_{i,j,m}^k\}$, and $a_{i,j}^k, b_{i,j,m}^k, c_{i,j}^k, d_{i,j,m}^k$ are the conditional probabilities in the state transition matrices $A^k, B^k(H_i^E, H_i^T)$, $C^k, D^k(H_i^E)$, respectively, which can range from 0 to 1. From (32), $P(H_i^{k-1}, H_i^T, O_i^S, O_i^F | W_{DBN}^k)$ is a constant for slot $t$, and the maximum optimization problem can be separated into four subproblems, namely the last polynomial additions, to optimize parameters $a_{i,j}^k, b_{i,j,m}^k, c_{i,j}^k, d_{i,j,m}^k$, respectively. Following [11], the optimized values of parameters in each transition matrices at the iteration $k$ can be calculated as

$$
da_{i,j}^k = \frac{\sum_{t=1}^T P(\tilde{H}_i^E = j, H_i^{T-1} = i | W_{DBN}^{k-1})}{\sum_{t=1}^T P(\tilde{H}_i^E, H_i^{T-1} = i | W_{DBN}^{k-1})},
$$

(33)

$$
b_{i,j,m}^k = \frac{\sum_{t=1}^T P(\tilde{H}_i^T = j, H_i^{T-1} = i, \tilde{H}_i^T = m | W_{DBN}^{k-1})}{\sum_{t=1}^T P(\tilde{H}_i^T, H_i^{T-1} = i, \tilde{H}_i^T = m | W_{DBN}^{k-1})},
$$

(34)

$$
c_{i,j}^k = \frac{\sum_{t=0}^T P(\tilde{H}_i^E = j, O_i^S = i | W_{DBN}^{k-1})}{\sum_{t=0}^T P(\tilde{H}_i^E = i, O_i^S | W_{DBN}^{k-1})},
$$

(35)

$$
d_{i,j,m}^k = \frac{\sum_{t=1}^T P(\tilde{H}_i^E = m, O_i^E = j, O_i^F = i | W_{DBN}^{k-1})}{\sum_{t=1}^T P(\tilde{H}_i^E = i, O_i^E, O_i^F | W_{DBN}^{k-1})},
$$

(36)

where $\tilde{H}_i$ denotes the transition states calculated by the parameters $W_{DBN}^{k-1}$ and the historical data contains $H_{i-1}, O_i, O_{i-1}$. Based on the calculations in (33)–(36), the parameter values of the transition matrices are continuously iterated by the new value until the optimization objective log-likelihood function $L(W_{DBN}^k; W_{DBN}^{k-1})$ converges. Specifically, let $\varepsilon_1$ denote the convergence threshold which is a small number. When $L(W_{DBN}^k; W_{DBN}^{k-1}) - L(W_{DBN}^1; W_{DBN}^{k-2}) \leq \varepsilon_1$, the log-likelihood function reaches convergence.

Note that the iterative algorithm is only guaranteed to converge at a stationary point, not at a global maximum point, which is related to the initialization of estimated parameters. This is because the problem (31) is solved in each iteration based on the condition parameters, and the global optimal parameters in each iteration are obtained based on the parameters in the previous iteration. It is known that the parameters in the initial iteration is determined randomly, so that the convergence point of the iteration algorithm is related to the initial parameters. In this case, the iterative algorithm needs to be repeated several times with different initial values. Then the parameters corresponding to the maximum value of log-likelihood function is selected. Assume that the number of iterations is $R$. In summary, the offline training algorithm of DBN is given as Algorithm 1.

Algorithm 1 The Proposed Offline Training Algorithm of DBN

Require: Historical data including the harvested energy at PU $H^E$, the transmission energy at PU $H^T$, the spectrum sensing power at SU $O^S$, and the harvested energy at SU $O^E$.

Ensure: Trained parameters $W_{DBN}$.

1: for $r = 1 : 1 : R$ do
2: Initialize DBN parameters $W_{DBN}^0$ consisting of the conditional transition probabilities $a_{i,j}^0, b_{i,j,m}^0, c_{i,j}^0, d_{i,j,m}^0$, which are randomly assigned from 0 to 1.
3: $L(W_{DBN}^0; W_{DBN}) \leftarrow \varepsilon_1, L(W_{DBN}^0; W_{DBN}) \leftarrow -\varepsilon_1, k \leftarrow 0$
4: while $L(W_{DBN}^k; W_{DBN}^{k-1}) - L(W_{DBN}^{k-1}; W_{DBN}^{k-2}) \geq \varepsilon_1$ do
5: Calculate $a_{i,j}^k, b_{i,j,m}^k, c_{i,j}^k, d_{i,j,m}^k$ by using historical data $[H^E, H^T, O^S, O^E]$ according to (33) to (36).
6: According to (32), calculate the expectation of log-likelihood $L(W_{DBN}^k; W_{DBN}^{k-1})$
7: $k \leftarrow k + 1$
8: end while
9: Record the expectation of log-likelihood value and the parameters as $L(W_{DBN}^m; W_{DBN}^{m-1})$ and $W_{DBN}^m$, respectively.
10: end for
11: Output the DBN parameters with maximum expectation of log-likelihood value from the record $W_{DBN}^m$.

After the training process of DBN in Algorithm 1, the conditional transition probabilities $a_{i,j}, b_{i,j,m}, c_{i,j}, d_{i,j,m}$ are calculated, which are the parameters of DBN. In other words, the energy-spectrum correlations characterized by the
optimal state transition matrices \( \mathbf{A}, \mathbf{B}(\mathbf{H}_E), \mathbf{C}, \mathbf{D}(\mathbf{H}_E) \) are obtained. With the correlations, the hidden states of PU can be estimated in the 2D sensing module.

3) ONLINE 2D SENSING BY TRAINED DBN

Based on the HMM structure established for EH-CRN, SU can use the observed information, i.e., spectrum sensing of received signals \( O_S^t \) and SU’s harvested energy \( O_E^t \), to infer the hidden information, i.e., PU’s harvested energy \( H_E^t \) and transmission energy \( H_T^t \). The 2D sensing process could be achieved by maximizing the condition probability of the hidden states according to the problem formulation in (11). Hence, the estimation values of the hidden states under 2D sensing are calculated as

\[
\hat{H}_E^t, \hat{H}_T^t = \arg \max_{(H_E^t, H_T^t)} P(H_E^t = h_E^t, H_T^t = h_T^t | O_S^0, \ldots, O_E^t, O_E^0),
\]

(37)

where \( \hat{H}_E^t, \hat{H}_T^t \) represent the sensing results of the harvested energy and the transmission energy at PU, respectively. Since the observed states \( O_S^0, O_E^0 \) are the existing deterministic information, the joint probability \( P(O_S^0, O_E^0) \) is a constant. According to the Bayes rules, maximizing the objective conditional probability is equivalent to maximizing the joint probability of hidden states and observed states, which can be expressed as \( P(H_E^t, H_T^t, O_S^0, O_E^0, O_S^1, O_E^1, \ldots, O_S^{t-1}, O_E^{t-1}) \).

As the Markovian property shown in Fig. 4(b), the joint conditional probability can be calculated by the recursive derivation process as

\[
P(H_E^t, H_T^t | O_S^0, O_E^0, O_S^1, O_E^1, \ldots, O_S^{t-1}, O_E^{t-1}) = \sum_{H_E^{t-1}, H_T^{t-1}} \sum \sum P(O_S^t | H_T^{t-1}) P(O_E^t | H_E^t, O_E^{t-1}, O_S^{t-1}) P(H_E^t | O_S^0, O_E^0, O_S^1, \ldots, O_S^{t-1}, O_E^{t-1}),
\]

(38)

Note that there are five probabilities in (38), and the values of the first four probabilities are the parameters of DBN \( a_{i,j}, b_{i,j,m}, c_{i,j}, d_{i,j,m} \), respectively, which have been trained in Algorithm 1. Through the online observation of SU \( O_E^t, O_S^t \), the hidden states to be estimated \( H_E^t, H_T^t \), and the previous information \( H_E^{t-1}, H_T^{t-1}, O_S^{t-1}, O_E^{t-1}, O_S^{t-1}, O_E^{t-1} \), the indexes of the parameters \( i, j, m \) are specifically determined. The last probability is the objective function at the previous slot, which is obtained by the recursive derivation. In the initial slot, the objective function in (38) can be transformed as

\[
P(H_E^0, H_T^0 | O_S^0, O_E^0) = P(O_S^0 | H_E^0) P(O_E^0 | H_E^0) P(H_T^0 | H_E^0) P(H_T^0).
\]

(39)

Similarly, the three first probabilities are the parameters of DBN \( b_{i,j,m}, c_{i,j}, d_{i,j,m} \), respectively. The index \( i, j, m \) are determined by the online observation of SU \( O_S^0, O_E^0 \) and the hidden states to be estimated \( H_E^0, H_T^0 \). The last probability is decided by the initial probability distribution of the harvested energy at PUs. Therefore, the joint conditional probability in (37) are obtained by the recursive calculation of (38) and (39), and the values of the hidden states \( \hat{H}_E^t, \hat{H}_T^t \) in (37) are estimated by maximizing the joint conditional probability. Online 2D sensing algorithm is summarized in Algorithm 2.

Algorithm 2 The Proposed DBN-Based Online 2D Sensing Algorithm

Require: Observed states of SU \( O_S^t, O_E^t \) at the current slot.
Ensure: Estimation of the hidden states of PU at the current slot \( \hat{H}_E^t, \hat{H}_T^t \).

1: for \( t = 0 : 1 : T \) do
2: \( n_1 = 1 : 1 : N_1 \) do
3: for \( n_2 = 1 : 1 : N_2 \) do
4: Get the online observation \( O_S^t, O_E^t \).
5: \( H_E^t \leftarrow n_1, H_T^t \leftarrow n_2 \)
6: Calculate the joint conditional probability in (37) by using (38) and (39).
7: end for
8: end for
9: Select the estimation of the hidden states of PU \( \hat{H}_E^t, \hat{H}_T^t \) with the maximum joint conditional probability calculated in step 6.
10: end for

Algorithm 2 enables the 2D sensing module, which estimates the hidden states of PU by using the observation of SU at each slot. The estimations will be further applied in the prediction and transmission modules to achieve network performance improvement.

C. TRANSMISSION OF SU REALIZED BY DNN

After obtaining the trained DBN, 2D correlations are represented by the state transition matrices, and the hidden information at the current slot \( \hat{H}_E^t, \hat{H}_T^t \) are estimated. In order to access the licensed spectrum in the forthcoming slots, the SU needs to know the spectrum status in advance. To this end, the information of PU and SU in the forthcoming time slots need to be predicted at the current slot. Based on the prediction results, SU can make a smart and proactive transmission decision by applying DNN, which is trained to improve network throughput.

1) 2D PREDICTION ALGORITHM

The problems of 2D prediction are initially characterized by (14)-(18), which maximize the joint conditional probabilities to obtain the states of PU and SU in the future. In the sensing module, the joint conditional probabilities of the harvested energy of PU \( H_E^t \), the transmission energy of PU \( H_T^t \), the sensing signal of SU \( O_S^t \), and the harvested energy of SU \( O_E^t \) are represented in HMM as the parameters of DBN \( a_{i,j}, b_{i,j,m}, c_{i,j}, d_{i,j,m} \), respectively, which are calculated by offline training of DBN in Algorithm 1. The hidden states of PU are estimated by online 2D sensing algorithm in Algorithm 2. Based on the joint conditional probabilities and the estimation
of the hidden states, 2D prediction algorithm is proposed as follows.

We first consider one-slot look-ahead 2D prediction. The specific mathematical solution of one-slot look-ahead 2D prediction is presented as

$$\hat{H}^E_{t+1} = \arg \max_{H^E_{t+1}} P(H^E_{t+1} = h^E_{t+1} | H^E_{t} = \hat{H}^E_{t}), \quad (40)$$

$$\hat{H}^T_{t+1} = \arg \max_{H^T_{t+1}} P(H^T_{t+1} = h^T_{t+1} | H^T_{t} = \hat{H}^T_{t}, H^E_{t} = \hat{H}^E_{t}), \quad (41)$$

wherein the hidden states of harvested energy $\hat{H}^E_{t+1}$ and the transmission energy of PU $\hat{H}^T_{t+1}$ at the next slot can be predicted. Here, $\hat{H}^E_{t}$ and $\hat{H}^T_{t}$ are the 2D sensing results from (37), and the joint conditional probabilities in (40) and (41) are the parameters of DBN $a_{i,j}, b_{i,j,m}$ in the state transition matrices $A, B(H^E_t)$, respectively. Similarly, the following states of the spectrum sensing signal $\hat{O}^S_{t+1}$, the harvested energy $\hat{O}^E_{t+1}$ and the residual energy $\hat{O}^R_{t+1}$ of SU can be expressed as

$$\hat{O}^S_{t+1} = \arg \max_{O^S_{t+1} \geq 0} P(O^S_{t+1} = o^S_{t+1} | H^T_{t+1} = \hat{H}^T_{t+1}), \quad (42)$$

$$\hat{O}^E_{t+1} = \arg \max_{O^E_{t+1} \geq 0} P(O^E_{t+1} = o^E_{t+1} | O^E_{t} = \hat{O}^E_{t}, H^E_{t+1} = \hat{H}^E_{t+1}), \quad (43)$$

$$\hat{O}^R_{t+1} = \min(O^R_{t+1} + O^E_{t} - D^T_{t}, N_s), \quad (44)$$

where $D^T_t$ denotes the actual value of the transmission energy of SU at the slot $t$, which is discretized into $N_d$ different energy levels. The joint conditional probabilities in (42) and (43) are the parameters of DBN $c_{i,j}, d_{i,j,m}$ in the state transition matrices $C, D(H^E_t)$, respectively.

After solving the one-slot look-ahead prediction problem, we now study the multi-slot look-ahead prediction, which can be considered as carrying out one-slot look-ahead prediction multiple times. The specific implementation of $N$-slot look-ahead 2D prediction algorithm can be summarized in Algorithm 3.

**Algorithm 3 N-Slot Look-Ahead 2D Prediction Algorithm**

**Require:** Observed states of SU $O^S_t, O^E_t$ and estimation of hidden states of PU at the current slot $\hat{H}^E_t, \hat{H}^T_t$.

**Ensure:** Prediction results $\hat{H}^E_{t+n}, \hat{H}^T_{t+n}, \hat{O}^S_{t+n}, \hat{O}^E_{t+n}, \hat{O}^R_{t+n}$ at the slot $(t+n)$.

1. for $n = 1 : 1 : N_d$ do
2. Use $\hat{H}^E_{t+n-1}, \hat{H}^T_{t+n-1}$ to predict the value of the hidden states next slot $\hat{H}^E_{t+n}, \hat{H}^T_{t+n}$ according to (40), (41).
3. Use $\hat{H}^E_{t+n}, \hat{H}^T_{t+n}$ to predict the value of SU states next slot $\hat{O}^S_{t+n}, \hat{O}^E_{t+n}, \hat{O}^R_{t+n}$ according to (42), (43), (44).
4. end for

The results of N-slot look-ahead 2D prediction $\hat{H}^E_{t+n}, \hat{H}^T_{t+n}, \hat{O}^S_{t+n}, \hat{O}^E_{t+n}, \hat{O}^R_{t+n}$ are used as the input of the decision module to obtain the optimal transmission energy of SU, which is implemented by DNN below.

2) **OFFLINE TRAINING OF DNN**

**a: NETWORK STRUCTURE**

As shown in Fig. 5, DNN is proposed to build connection from the future states of PU and SU $\hat{H}^E, \hat{H}^T, \hat{O}^S, \hat{O}^E, \hat{O}^R$ to the transmission energy of SU $D^T$, which performs data interaction with communication environment to enhance the network throughput. In the proposed DNN, there are three layers in the neural network, which are the input layer, hidden layer and output layer. In the input layer, the prediction results are input to the neural network, while in the output layer, the transmission energy of SU are output after the computation of DNN. We define $W_{DNN} = \{w_{in}, w_{out}\}$ as the parameters of the DNN. The parameters need to be trained by using historical dataset, which will drive the DNN to output the optimal transmission energy.

**b: DATASET AND LABELING**

To train the parameters of the proposed DNN, input dataset and output dataset which is the label of the input are needed. The information of PU and SU in energy and spectrum dimensions $\hat{H}^E, \hat{H}^T, \hat{O}^S, \hat{O}^E, \hat{O}^R$ are used as the input dataset, which are recorded from the real EH-CRN environment. The output dataset needs to be prepared by labeling the input dataset. Specifically, we use the optimal transmission energy at SU $D^T_t$ as the label of the input dataset, which can be obtained by maximizing the instantaneous throughput under the condition of energy reservation detailed in (21). As the channel coefficients $h_{PU}^T, h_{SU}$ and the size of transmitted data packet of SU $K_t$ are recorded at PU and SU for offline training, the optimization problem can be easily solved, which aims to increase the network throughput and reserve more available energy. In a word, the input dataset at the current slot is $\{\hat{H}^E_t, \hat{H}^T_t, \hat{O}^S_t, \hat{O}^E_t, \hat{O}^R_t\}$ and the output dataset is $\{D^T_t\}$ as the label.

**c: TRAINING**

The back propagation (BP) algorithm is used to train the parameters of each layer, which consists of two processes, namely the forward propagation of the input and the
backward propagation of the error. As shown in Fig. 5, in the forward propagation process, the inputs of neural nodes are first multiplied by the parameters of the layer and added with a bias. Then the result of this linear mapping is fed to the activation function as nonlinear mapping, and the nonlinear mapping result is fed to the next layer. In the output layer, the result $\hat{D}^T_l$ of the nonlinear mapping is the final output of the neural network. The error between the expected output $D^T_l$ and the actual output $\hat{D}^T_l$ is calculated as

$$L(D^T_l, \hat{D}^T_l) = \frac{1}{2}(D^T_l - \hat{D}^T_l)^2. \quad (45)$$

Gradients descent method is applied during the back-propagation process. When back-propagating, the gradients of the output error for the parameters of each layer are reversely calculated from the output layer to the input layer. The gradients are used for correcting the weight of the nodes of each layer. After continuously adjusting the weights and bias of neurons nodes in each layer, the error of each node is reduced to the minimum. Let $W_{DNN} = \{w'_{in}, w'_{out}\}$ denote the new parameters updated by gradients. Parameters updating can be expressed as

$$w'_{in} = w_{in} + L \cdot \frac{\partial \sum_{t=0}^{T} L(D^T_l, \hat{D}^T_l)_{\frac{\partial w_{out}}{\partial w_{in}}}, \quad (46)$$

$$w'_{out} = w_{out} + L \cdot \frac{\partial \sum_{t=0}^{T} L(D^T_l, \hat{D}^T_l)_{\frac{\partial w_{out}}{\partial w_{in}}}, \quad (47)$$

where $L$ is the learning rate of the proposed DNN. Let $\varepsilon_2$ denote the threshold to judge whether the training of DNN has converged. The forward propagation of input and back-propagating of error process alternately, until the following relationship is satisfied:

$$\sum_{t=0}^{T} (L(D^T_l, \hat{D}^T_l | W_{DNN}) - L(D^T_l, \hat{D}^T_l | W_{DNN})) \leq \varepsilon_2, \quad (48)$$

where $T$ denotes the total number of time slots. When the condition in (48) is satisfied, the training of DNN is completed. The training algorithm of DNN is summarized in Algorithm 4.

Employing the Algorithm 4, the parameters of DNN are optimized to output the optimal transmission energy of SU, which can be applied in the online 2D decision module.

3) ONLINE 2D DECISION BY TRAINED DNN

Based on the proposed offline training algorithm of DNN, we obtained the parameters of DNN, and the optimal transmission energy of SU can be output in the forward propagation of DNN. The online 2D decision is proposed to output the optimal energy based on the online predicted results. Specifically, the SU performs $N$-slot look-ahead prediction online by Algorithm 3 to obtain the harvested energy at PU $H_{t+i+n}^E$, the transmission energy of PU $H_{t+i+n}^T$, the sensing signal at SU $\tilde{O}_{t+i+n}^S$, the harvested energy at SU $\tilde{O}_{t+i+n}^R$, and the residual energy of SU $\tilde{O}_{t+i+n}^R$. 2D decision module of SU uses the online prediction results to make the transmission energy decision $D^T_{t+i+n}$ by the trained DNN. The online DNN with $N$-slot look-ahead prediction process is expressed as

$$D^T_{t+i+n} = f_D(\tilde{H}^E_{t+i+n}, \tilde{H}^T_{t+i+n}, \tilde{O}^S_{t+i+n}, \tilde{O}^E_{t+i+n}, \tilde{O}^R_{t+i+n}), \quad (49)$$

where $f_D$ denotes the mapping from the prediction results to the transmission energy, which is achieved by trained DNN.

D. COMPLEXITY ANALYSIS

According to Algorithm 5, our proposed ML-based energy-spectrum 2D cognition algorithm are summarized, which consists of the offline training process of DBN and DNN, and the online cognition process. The complexity of the each process is analyzed as follows.

1) TRAINING OF DBN

DBN needs to repeat the iterative algorithm multiple times to find a maximum convergence value of objective log-likelihood function. In each iteration, the joint conditional probabilities of states of PU and SU are calculated with an enumeration method from slot 0 to slot $t$ in Algorithm 1. Therefore, the time complexity of the training of DBN is $O(RN^2)$, where $R$ denotes the iteration time till a maximum convergence value is found. Compared with the training algorithm in [11], the time complexity is not increased.

2) TRAINING OF DNN

Assume that the numbers of neural nodes in each layer of DNN are $K_1, K_2, K_3$. Neural nodes in adjacent layers are connected to each other, which leads to the complexity $O((K_1 K_2 + K_2 K_3)N)$ in the back-propagation training in Algorithm 4.

3) ONLINE 2D COGNITION

With the trained DBN and DNN, 2D cognition are realized by 2D sensing, 2D prediction and 2D decision in order. 2D
Algorithm 5: The Proposed ML-Based Energy-Spectrum 2D Cognition Algorithm at SU in EH-CRN

1: Train the parameters of DBN \(W_{DBN}\) by using the historical data, which includes the harvested energy at PU \(H^E\), the transmission energy at PU \(H^T\), the spectrum sensing power at SU \(O^S\), and the harvested energy at SU \(O^F\), as shown in Algorithm 1.
2: Train the DNN with the input dataset \(\{H^E_t, H^T_t, O^S_t, O^F_t\}\) and the output dataset \(\{D^T_t\}\), as shown in Algorithm 4.
3: for \(t = 0: 1: T\) do
  4:    Observe the states of SU \(O^S_t, O^F_t\) at the current slot and input to the 2D sensing module.
  5:    Estimate the value of the hidden states at current slot \(\hat{H}^E_t, \hat{H}^T_t\) according to Algorithm 2.
  6:    Input the sensing results \(\hat{H}^E_t, \hat{H}^T_t\) to 2D prediction module.
  7:    Predict the value of the forthcoming states \(\hat{H}^E_{t+n}, \hat{H}^T_{t+n}, \hat{O}^S_{t+n}, \hat{O}^F_{t+n}\) and \(\hat{O}^R_{t+n}\) according to Algorithm 3.
  8:    Input the prediction results into DNN to obtain the transmission energy at SU \(D^T_{t+n}\) in the decision module.
  9:    if \(t \geq n\) then
  10:       Transmit data with the transmission energy \(D^T_t\) calculated at slot \((t - n)\).
  11:    end if
  12: end for

sensing performs a recursive calculation of joint probability according to (37)–(39) in Algorithm 2, which leads to the complexity \(O(N^2)\). Note that 2D prediction in Algorithm 3 and 2D decision by DBN are all forward sequential operations, the complexity of the two modules are \(O(K_1N)\) and \(O(K_1K_2 + K_2K_3)N\), respectively.

To sum up, the time complexity of the proposed ML-based 2D cognition algorithm is \(O(RN^2 + (K_1 + K_1K_2 + K_2K_3)N)\), i.e., \(O(RN^2)\), accordingly.

V. PERFORMANCE EVALUATION

In this section, we present the performance evaluation of an EH-CRN to demonstrate the effectiveness of our proposed ML-based energy-spectrum 2D cognition.

A. SYSTEM SETUP

As the description of the network structure in Section II, an EH-CRN with one PU and one SU is considered, and the primary and secondary transmitters are assumed powered by harvested energy from the same renewable source, e.g., solar energy and wind energy.

We will illustrate the performance superiority of our proposed ML-based 2D cognition in sensing, prediction, and decision, respectively. Following the parameter setting in [11], the values of harvested energy of both PU and SU are quantized as \{0, 1, 2, 3\}, the channel noise level is set as \(\sigma^2 = 3\) dBW, and the channel path loss between primary and secondary transmitters is \(-4\) dB.

B. 2D SENSING

We use the receiver operating characteristic curve (ROC) to present the sensing performance under different sensing schemes. The ordinate of ROC is the missing detection probability \(P_m = 1 - P_d\), and the abscissa is \(P_f\). The smaller area under ROC denotes the better sensing performance. Using two detection SNR values 0 dB and 5 dB, respectively, Fig. 6 shows the relationship between the false alarm probability \(P_f\) and the miss detection probability \(P_m\) for both 1D sensing and 2D sensing. It is observed that 2D sensing achieves much better performance than 1D sensing, because SU takes full advantage of the spatial and temporal correlations of energy and spectrum dimensions, and exploits the correlation between two dimensions in our proposed 2D sensing scheme.

We also compare the sensing performance of 2D sensing with 1D energy sensing, in which the SU senses the spectrum hole only through its own harvested energy. As shown in Fig. 7, the 2D cognition scheme generates fewer false alarms with fixed miss probability \((P_m = 0.1)\). It is worth noting that there is a U-shape relationship between the false probability and PU loading of 2D cognition, where the PU loading...
means the probability that the PU has data to transmit when enough energy is harvested. This is because that when the PU loading is low, the extra information about the harvested energy interferes the inference of PU behavior. As PU loading increases, the importance energy sensing in 2D cognition increases gradually. However, at the same time, the role of spectrum sensing decreases gradually. Continuously increasing PU loading will bring imperfections in 2D cognition. Therefore, there will be a compromise between the role of PU energy sensing and the role of spectrum detection, resulting in the appearance of U-shaped 2D curves. On the other hand, 1D energy sensing does not include PU spectrum detection, and hence there is no compromise in 1D energy sensing, thus the performance is monotonous. Note that 1D energy sensing may nearly deliver the same performance as 2D cognition when both the spatial correlation between harvested energy of PUs/SUs and the PU loading are very high.

C. 2D PREDICTION

We further compare the prediction performances of 2D cognition and 1D cognition. HMM is used to predict the spectrum hole in the 1D cognition scheme. Consider that not only the spectrum prediction is a classification problem, but also the consequences of prediction errors are different when the licensed spectrum status is idle or occupied. The miss detection probability $P_m$ and the false alarm probability $P_f$ can separately represent the prediction effect of two different kinds of spectrum status. In Fig. 8, we also use ROC to present the one-slot look-ahead prediction performance under different cognition schemes. It is observed that there are smaller areas under the 2D cognition curves, which denote that 2D prediction renders much higher prediction accuracy than 1D prediction.

The historical sensing information is further exploited to predict the future $N_p$ slots of spectrum status. For instance, $N_p = 0$ means the case of spectrum inference with sensing, and $N_p = 1$ means the case of one-slot look-ahead prediction. As shown in Fig. 9, 2D cognition achieves considerably higher detection probability with fixed false alarm probability ($P_f = 0.1$), which means that the PU is better protected from the harmful interference caused by the SU’s transmissions.

D. TRANSMISSION DECISION BASED ON 2D COGNITION

Due to the high correlation of harvested energy between PU and SU, the PU has a high probability to occupy the spectrum when the SU harvests enough energy. Thus, a small energy buffer is equipped at the SU to help SU stagger the transmission to improve the energy and spectrum efficiency. In the simulations for Fig. 10, the energy buffer of SU is set as 10 J and the maximum transmit power of PU is 6 W. The greedy energy allocation policy and the case without buffering [11] are introduced for comparison, whose philosophy is to exhaust all the stored energy in the current slot if the current slot is detected as idle, in order to obtain the maximal instantaneous throughput.

It can be seen from Fig. 10 that 2D decision can obtain much better performance of the average throughput per slot than 1D decision with the same protection level for PU ($P_f = 0.8$). When the detection SNR is 5 dB, the performance gains are up to 30% and 132% for the 2D cognition combined with the greedy energy allocation policy and the optimal energy allocation policy, respectively. In addition, extra 24%
and 52% performance gains are obtained for the optimal 2D decision compared with the greedy policy and the scheme with no energy buffer.

VI. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

In this article, we have proposed a ML-based energy-spectrum 2D cognition scheme and discussed its potential applications to EH-CRNs. DBN is used to model the 2D cognition for EH-CRNs, representing the correlations between SUs and PUs in both energy and spectrum dimensions. In addition, the performance of ML-based 2D cognition scheme is evaluated in three aspects, namely sensing, prediction, and decision modules. The evaluation results illustrate the superiority of 2D cognition compared to 1D cognition. To encourage more research efforts on ML-based 2D cognition for EH-CRNs, the following research directions are envisioned.

A. MULTI-DIMENSIONAL COGNITION

There are other dimensions in EH-CRNs that may provide extra information to improve spectrum cognition performance in different scenarios. For example, different spectrum bands may jointly provide service to a PU, which causes the occupation probabilities of different spectrum bands to be correlated in multi-band CRNs. Moreover, when there are several SUs that need to use the licensed bands, the location distribution of SUs directly affects spectrum availability and spectrum sensing. Spectrum cognition that takes the spatial domain into consideration can obtain more information to improve the spectrum efficiency through multi-user sensing, which requires further study.

B. MACHINE LEARNING-BASED 2D COGNITION WITH MISSING DATA

In practical systems, it is difficult for energy harvesting SUs to obtain the complete energy and spectrum dimensional historical data. For example, SUs cannot detect the licensed spectrum without enough energy, and thus the information about spectrum status is missing when the SUs do not harvest enough energy. The incompleteness of historical data may degrade the performance of spectrum prediction. Therefore, developing efficient 2D cognitive schemes to handle missing data is an interesting topic for future study.

REFERENCES

[1] S. Force, “Spectrum policy task force report,” Federal Communications Commission, Washington, DC, USA, Tech. Rep., ET Docket, No.02-135, 2002.
[2] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, “A survey on spectrum management in cognitive radio networks,” IEEE Commun. Mag., vol. 46, no. 4, pp. 40–48, Apr. 2008.
[3] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, “Transmission with energy harvesting nodes in fading wireless channels: Optimal policies,” IEEE J. Sel. Areas Commun., vol. 29, no. 8, pp. 1732–1743, Sep. 2011.
[4] W. Chung, S. Park, S. Lim, and D. Hong, “Spectrum sensing optimization for energy-harvesting cognitive radio systems,” IEEE Trans. Wireless Commun., vol. 13, no. 5, pp. 2601–2613, May 2014.
[5] A. A. Olawole, F. Takawira, and O. O. Oyerinde, “Cooperative spectrum sensing in multichannel cognitive radio networks with energy harvesting,” IEEE Access, vol. 7, pp. 84784–84802, 2019.
[6] T. Shahjahi and K. V. Babu, “Cooperative scheme for wireless energy harvesting and spectrum sharing in cognitive 5G networks,” in Proc. Int. Conf. Innov. Electr., Electron., Instrum. Media Technol. (ICIEEMT), Feb. 2017, pp. 57–61.
[7] X. He, H. Jiang, Y. Song, X. Yang, and H. Xiao, “Optimal resource allocation for energy harvesting cognitive radio network with q learning,” in Proc. Int. Conf. Artif. Intell. Secur., New York, NY, USA, Jul. 2019, pp. 548–560.
[8] K. Pathak and A. Banerjee, “Harvest-or-transmit policy for cognitive radio networks: A learning theoretic approach,” IEEE Trans. Green Commun. Netw., vol. 3, no. 4, pp. 997–1011, Dec. 2019.
[9] C. Qiu, Y. Zhang, Z. Feng, P. Zhang, and S. Cui, “Spatio-temporal wireless traffic prediction with recurrent neural network,” IEEE Wireless Commun. Lett., vol. 7, no. 4, pp. 554–557, Aug. 2018.
[10] W. Han, H. Sang, X. Ma, J. Li, Y. Zhang, and S. Cui, “Sensing statistical primary network patterns via Bayesian network structure learning,” IEEE Trans. Veh. Technol., vol. 66, no. 4, pp. 3143–3157, Apr. 2017.
[11] Y. Zhang, W. Han, D. Li, P. Zhang, and S. Cui, “Power versus spectrum 2-D sensing in energy harvesting cognitive radio networks,” IEEE Trans. Signal Process., vol. 63, no. 23, pp. 6200–6212, Dec. 2015.
[12] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, “Sensing-throughput tradeoff for cognitive radio networks,” IEEE Trans. Wireless Commun., vol. 7, no. 4, pp. 1326–1337, Apr. 2008.
[13] M. He, L. Yang, J. Zhang, and V. Vittal, “A spatio-temporal analysis approach for short-term forecast of wind farm generation,” IEEE Trans. Power Syst., vol. 29, no. 4, pp. 1611–1622, Jul. 2014.
[14] Solar Radiation Database. Accessed: May 2018. [Online]. Available: http://solar.uq.edu.au/user/reports/Power.php

WENJUN XU (Senior Member, IEEE) received the B.S. and Ph.D. degrees from the Beijing University of Posts and Telecommunications (BUP), China, in 2003 and 2008, respectively. He is currently a Professor and the Ph.D. supervisor with the School of Information and Electrical Engineering, HUE, and a member of the Hebei Key Laboratory of Security and Protection Information Sensing and Processing. His research interests include wireless networks, data mining, privacy preservation, and so on.

YONGJIAN FAN received the B.S. degree from the Hebei University of Engineering (HUE), China, in 2009, and the Ph.D. degree from the Renmin University of China (RUC), China, in 2013. He is currently an Associate Professor and the B.S. supervisor with the School of Information and Electrical Engineering, HUE, and a member of the Hebei Key Laboratory of Security and Protection Information Sensing and Processing. His research interests include AI-driven networks, UAV communications and networks, cognitive communications, networking, and so on.

Y. Fan et al.: ML-Based Energy-Spectrum Two-Dimensional Cognition in EH-CRNs
CHIA-HAN LEE (Member, IEEE) received the B.S. degree from National Taiwan University in 1999, the M.S. degree from the University of Michigan, Ann Arbor, in 2003, and the Ph.D. degree from Princeton University, in 2008, all in electrical engineering. From 1999 to 2001, he served in the ROC Army as a Missile Operation Officer. From 2008 to 2009, he was a Postdoctoral Research Associate with the University of Notre Dame, USA. From 2010 to 2016, he was with Academia Sinica as an Assistant Research Fellow and then an Associate Research Fellow. Since 2016, he has been with National Chiao Tung University as an Associate Professor. His research interests include wireless communications and networks. He is an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE COMMUNICATIONS LETTERS.

SILEI WU (Graduate Student Member, IEEE) received the B.S. degree from the School of Telecommunication Engineering, Beijing University of Posts and Telecommunications (BUPT), China, in 2019. He is currently pursuing the M.S. degree with the Key Laboratory of Universal Wireless Communications, Ministry of Education, BUPT. He majors in the signal processing and wireless communications.

FAN YANG (Member, IEEE) received the B.S. degree from the School of Information Engineering, Chongqing University of Posts and Telecommunications (CQUPT), China, and the Ph.D. degree in the signal processing and wireless communications from the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications (BUPT).

PING ZHANG received the Ph.D. degree in electrical from the Beijing University of Posts and Telecommunications (BUPT), China, in 1990. He is currently a Professor with the School of Information and Communication Engineering, BUPT. He is an Executive Associate Editor-in-Chief on information sciences of Chinese Science Bulletin, Reviewer and a TCP Member of many magazines, journals, and conferences. He is a member of next-generation broadband wireless communication network in National Science and Technology Major Project, and 5th Advisory Committee of National Natural Science Foundation of China, the Chief Scientist of 973 National Basic Research Program of China, a member of the 11th Beijing Municipal Committee of Chinese People’s Political Consultative Conference, and the owner of the Special Government Allowance of State Council of China. He received the Second Award for National Science and Technology Prize twice, the Second Award for National Science and Technology Invention Prize once, the Provincial Science and Technology Awards many times, and the Title of Outstanding Science and Technological Workers, in 2010. His research interests include broadband wireless communication, new technologies on cognitive wireless networks, TD-LTE, MIMO, OFDM, and so on.

* * *