Spectrum Availability Prediction in Cognitive Aerospace Communications: A Deep Learning Perspective

Lixing Yu, Qianlong Wang, Yifan Guo, and Pan Li
Department of Electrical Engineering and Computer Science
Case Western Reserve University
Cleveland, OH, USA
{lxy257, qxw204, yxg383 and lipan}@case.edu

***Abstract*—Cognitive Radio (CR) technology enables secondary users (SUs) to opportunistically access unused licensed spectrum owned by the primary users (PUs). Therefore, it can potentially significantly enhance communication capacity, and hence is very encouraging in aerospace communications and deserve thorough study. One of the key problems in cognitive aerospace communications is to determine spectrum availability. In the past, many researchers have proposed to employ spectrum sensing to address this issue, which, however, consumes considerable energy and time. In this paper, we develop a deep learning system to predict spectrum availability, which does not require a priori knowledge of the activities of PUs. The performance of the proposed system is analyzed through extensive simulations.**

Keywords—Cognitive aerospace communication; spectrum availability prediction; deep learning.

I. Introduction

Since Federal Communications Commission (FCC) opens the discussion on intelligently share licensed spectrum, cognitive radio (CR) have emerged as a promising technology. It can release the spectrum from shackles of authorized licenses of primary users (PUs), and enable unlicensed users or secondary users (SUs) to opportunistically access under-utilized licensed spectrum in either temporal or spatial domain, as long as their usage does not significantly impact the PUs' service provisioning [1] [9] [4], [5]. Due to the increasing demand on bandwidth in aerospace communications and the dynamic nature of spectrum in both spatial and temporal domains therein, CR technology has great potential to significantly improve the communication capacity in aerospace communications.

One of the most critical problems in CR communications is to determine spectrum availability. Most previous works [6] [15] address this problem by developing complicated spectrum sensing schemes, which scan the whole spectrum to check if any of the channels is occupied by PUs or not. Such schemes obviously consume tremendous energy and time, and hence are not appropriate for aerospace communications.

Recently, some researchers propose machine learning based methods to help carry out CR communications. Lunden et al. [6] and Sutton et al. [12] develop reinforcement learning algorithms to find better spectrum sensing policies. Cui et al. [2], Huang [3], and Ramon et al. [11] design support vector machine (SVM) based learning systems to classify PUs and determine wireless communication parameters. However, how to take advantage of machine learning to efficiently and effectively determine spectrum availability is still an open and challenging problem.

In this paper, we propose to utilize a deep learning technique, called long short-term memory (LSTM) network, to predict spectrum availability. The main idea is to learn from the past spectrum availability data, and exploit the intrinsic spectral-temporal correlation among them to predict future spectrum availability. There are a couple of works like [10] [13] which try to predict spectrum with artificial neural networks. These works can only make predictions based on temporal correlations among historical data, while our system can explore the correlation in the spectral-temporal domain and hence is more effective.

The rest of this paper is organized as follows. Section II describes the proposed system model for spectrum availability prediction. Section III details the proposed LSTM based scheme, which is followed by simulation results and discussions in Section IV. Finally, we conclude the paper in Section V.

II. System Model

In this work, we propose to make predictions on the spectrum availability at an arbitrary location by exploit the spectral-temporal correlation. We divide the spectrum of interests into spectrum bands or channels. We denote the system status representing the spectrum availability at time t by a "multi-hot" vector $\mathbf{x_t}$. Each element in $\mathbf{x_t}$ is either 1, standing for the corresponding channel is occupied, or 0, standing for the corresponding channel is available to use. Denote by M as the

This work was partially supported by the U.S. National Science Foundation under grants CNS-1602172 and CNS-1566479.

L. Yu, Q. Wang, Y. Guo, and P. Li are with Department of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, OH 44106. Email: {lxy257, qxw204, yxg383, lipan}@case.edu.

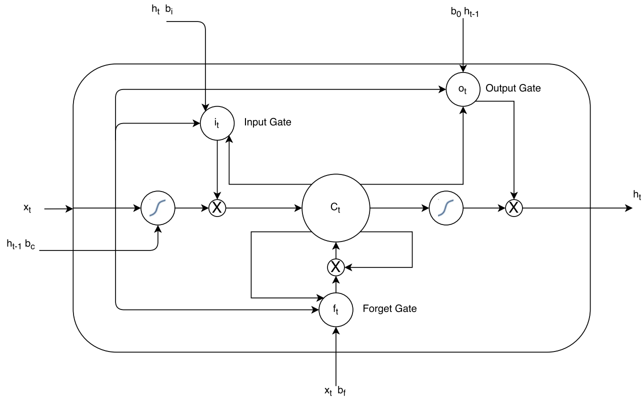


Fig. 1. An LSTM memory cell.

total number of channels. Then $\mathbf{x_t}$ is a vector of dimension $M\times 1$, where the elements equal to 1 are called "hot elements". We aim to predict the next spectrum availability status $\mathbf{x_{t+1}}$ based on the current and the previous spectrum availability status, i.e., $\mathbf{x_t}$, $\mathbf{x_{t-1}}$, $\cdots$, etc.

III. Spectrum Availability Prediction

The main idea of our system is to build an LSTM network to predict the spectrum availability by learning the spectral correlation among the spectrum availability in the past time slots. In what follows, we first introduce the fundamentals of the LSTM network and then elaborate how to develop our LSTM based system.

A. The Basic LSTM Architecture

LSTM was proposed based on the traditional RNN aiming at addressing its long-term temporal dependency issue. Particularly, although RNN can theoretically well deal with data that are temporally correlated, it cannot handle long-term time dependency in practice. LSTM addresses this problem by having a memory cell, which is featured with gates as shown in Fig.1. The gates are mathematically formulated as:

$$
\begin{aligned}
\mathbf{i}_t &= \sigma(\mathbf{W}_{ix}\mathbf{x}_t + \mathbf{W}_{ih}\mathbf{h}_{t-1} + \mathbf{W}_{ic}\mathbf{c}_{t-1} + \mathbf{b}_i)\\
\mathbf{f}_t &= \sigma(\mathbf{W}_{fx}\mathbf{x}_t + \mathbf{W}_{fh}\mathbf{h}_{t-1} + \mathbf{W}_{fc}\mathbf{c}_{t-1} + \mathbf{b}_f)\\
\mathbf{c}_t &= \mathbf{f}_t \circ \mathbf{c}_{t-1} + \mathbf{i}_t \circ \phi(\mathbf{W}_{cx}\mathbf{x}_t + \mathbf{W}_{ch}\mathbf{h}_{t-1} + \mathbf{b}_c)\\
\mathbf{o}_t &= \sigma(\mathbf{W}_{ox}\mathbf{x}_t + \mathbf{W}_{oh}\mathbf{h}_{t-1} + \mathbf{W}_{oc}\mathbf{c}_t + \mathbf{b}_o)\\
\mathbf{h}_t &= \mathbf{o}_t \circ \phi(\mathbf{c}_t)
\end{aligned}
$$

Here, $\mathbf{i}$, $\mathbf{f}$, $\mathbf{o}$ and $\mathbf{c}$ denotes the input gate, forget gate, output gate, and cell state, respectively. These gates are all of the same dimension as the hidden vector $\mathbf{h}$ which is assumed to be of $N\times 1$ dimension. σ is a sigmoid function, and ϕ is a nonlinear function which maps the input to $[-1, 1]$. $\mathbf{W}_{ic}$, $\mathbf{W}_{fc}$, and $\mathbf{W}_{oc}$ are the peephole connection matrices, which connect cell state to input gate, forget gate, and output gate, respectively. Similarly, $\mathbf{W}_{ix}, \mathbf{W}_{fx}, \mathbf{W}_{ox}$ and $\mathbf{W}_{cx}$ are the weight matrices connecting between input vector $\mathbf{x}_t$ and input gate, forget gate, output gate and cell state, respectively. Besides, since the gates and the input vector $\mathbf{x}_t$ have the dimension of $N\times 1$ and $M\times 1$ respectively, we can have that the dimensions of matrices $\mathbf{W}_{ih}, \mathbf{W}_{ic}, \mathbf{W}_{fh}, \mathbf{W}_{fc}, \mathbf{W}_{ch}, \mathbf{W}_{oh}, \mathbf{W}_{oc}$ are all the same, which is $N\times N$, and the dimensions of matrices $\mathbf{W}_{ix}, \mathbf{W}_{fx}, \mathbf{W}_{cx}, \mathbf{W}_{ox}$ are $N\times M$.

B. Our LSTM System for Spectrum Availability Prediction

1) The Main Framework: In our system, the input first goes through the LSTM layer. The output of the LSTM layer goes into the dense layer, i.e., a fully connected neural network, where a dropout process is applied in training to avoid overfitting. Finally, we employ an activation function, which is set to the Softmax function, and obtain the prediction result $\mathbf{y_t}$ in the time slot t. The system architecture is presented in Fig. 2. Where h_t^L is the output of the LSTM memory cell in the time slot t.

We note that the input $\mathbf{X_t}$ in the time slot t is a matrix of dimension $M\times T$, which means that the system predicts the spectrum availability in the forthcoming time slot by exploiting the data in the most recent T time slots. In $\mathbf{X_t}$, each column represents the spectrum availability in a time slot as shown in Fig.3.

2) LSTM Layer: In the LSTM layer, there are T LSTM memory cells, one for each input vector x_j ($t-T+1 \le j \le t$). The ouput of the j_{th} memory cell at time t, i.e., h_{t-j} and c_{t-j}, is part of the input of the next, i.e., the $j-1_{th}$ memory cell. The output of the LSTM layer, i.e., the output of the last LSTM memory cell, goes to a dense network.

3) Dense Network with Dropout: Essentially, the dense network is a fully connected neural network as shown in the left part of Fig. 4. In the dense network, the nodes in each layer are fully connected to all the nodes in the previous layer. The reason for having a dense network here is that the output of the LSTM contains the feature information we need to make prediction. While the dense network aims to map the feature data to the prediction result. In this work, we build three layers in the dense network.

During the training phase, some randomly chosen nodes in the dense layer can be turned off given each training data as shown in the right part of Fig. 4. This is called dropout and can effectively help prevent the dense network from overfitting, and enhance the prediction performance in the future.

4) Activation Function and Loss Function: In order to calculate the final prediction output, we choose softmax as the activation function that takes the output of the dense networks as the output. Particularly, this activation function maps the output of the dense networks into a vector of elements between 0 and 1, each of which represents the probability of a channel being occupied and the sum of which equals 1. The softmax function can be calculated as:

$$\mathbf{y_t^m} = \frac{e^{\mathbf{z^m}}}{\sum_{i=1}^{M} e^{\mathbf{z^i}}}, \text{for } m = 1, \ldots, M$$

In the equation above, $\mathbf{z}$ stands for the output of the dense network $\mathbf{z^m}$ and $\mathbf{y_t^m}$ represent the m_{th} elements of the LSTM output vector $\mathbf{z}$, and that of the output $\mathbf{y_t}$. To map the

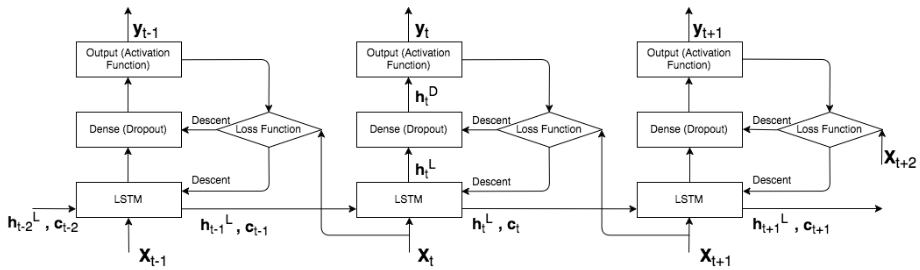


Fig. 2. Our LSTM system architecture with input $\mathbf{X}_t$ in the time slot t.

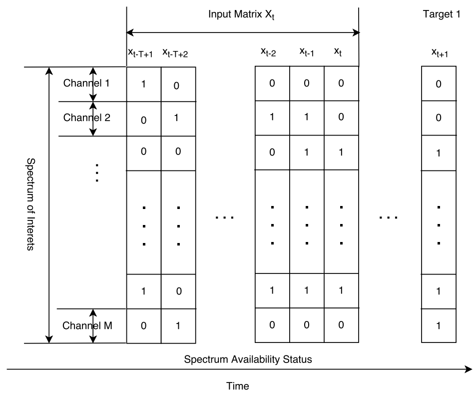


Fig. 3. The input matrix X_t.

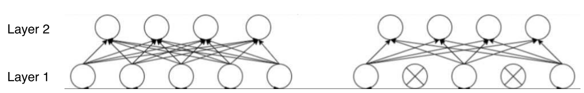


Fig. 4. Dropout in the dense network.

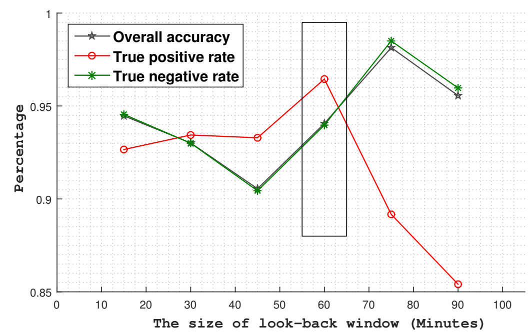


Fig. 5. Performance comparison under different look back window sizes.

probabilities into binary series, we find an optimal threshold during the training process which minimizes the difference between real data and the predicted ones, which are either 0s or 1s.

Besides, in order to enable the system to learn and evolute, we need to define a loss function. We employ the cross-entropy as the system loss function, since the problem in this work is basically a classification into variant labels and the cross-entropy is widely applied and proved to be effective. Particularly, cross-entropy can be calculated as:

$$\xi(\mathbf{x_{t+1}}, \mathbf{y_t}) = -\sum_{i=1}^{M} \mathbf{x^i_{t+1}} \log \mathbf{y^i_t},$$

Where, $\mathbf{x^i_{t+1}}$ and $\mathbf{y^i_t}$ are the i_{th} element in $\mathbf{x_{t+1}}$ and in $\mathbf{y_t}$, respectively.

IV. Performance Evaluation

A. Data preprocessing

In our study, the simulation data was sampled and gathered from Share Spectrum Company [7] [8]. In particular, we sampled the spectrum occupancy data from two spectrum reports obtained at two locations: New York City, NY and Vienna, VA. We study the spectrum from 3MHz to 5.4MHz and divide the spectrum into 26 channels. According to the spectrum report, if the signal strength detected in one channel is lower than -100 dbm, this channel will be considered as “idle”. Otherwise, this channel will be regarded as“busy”.

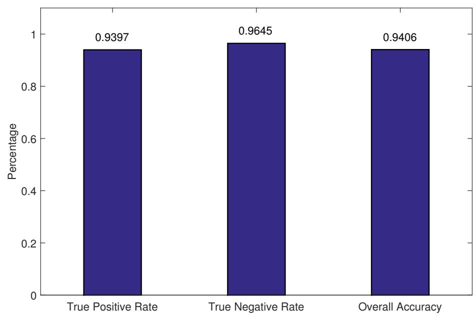
B. Simulation Settings

Our simulation settings are as follow. In the time domain, we set a time slot to 1 minute. We set the look back window from 15 time slots to 90 time slots. In the spectrum domain, we predict all the channels availability, instead of predicting only one channels availability [14]. We set the number of nodes at the output of the LSTM layer to 128. In the dense network, we have three dense layers, which contains 512, 256, 128 nodes respectively. Besides, there are 26 nodes at the output layer. The reason why we set relatively large number of nodes is that we would like the system to better characterize the dynamic states in the spectrum occupancy and better capture the features of the input. As mentioned before, we choose Softmax as activation function and employ RMSprop as the optimizer.

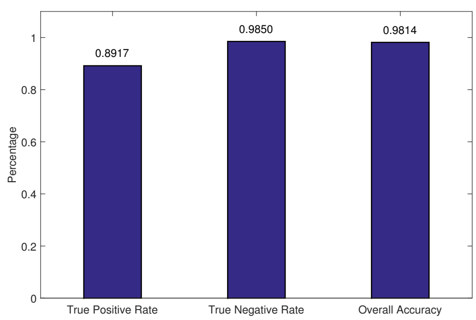
C. Simulation Results

With the system setting above, we evaluate the spectrum availability prediction performance. We conduct the simulations by varying our look back window length from 15 minute to 30, 45, 60, 75, 90, respectively. We would like to find out that with which look back window size, the system could obtain more accurate prediction result.

We present the simulation results in Fig. 5 and Fig. 6, For example, when the look back window size is 60, the system



(a) Look back window of 60 minutes



(b) Look back window of 75 minutes

Fig. 6. Performance when the look back window size is 60 and 75.

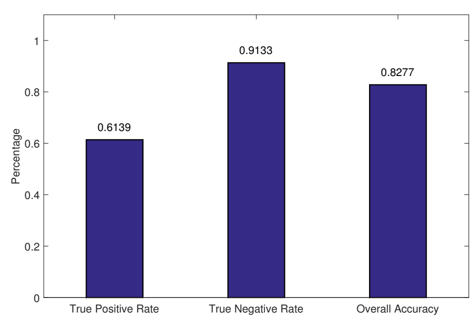


Fig. 7. Performance of a simple ANN model.

could achieve a good performance in predicting spectrum availability. The true positive rate is 93.97% the true negative rate is 96.45%, and the over all prediction accuracy is 94.06%. Moreover, when the look back window size is 75, the system achieves even higher overall accuracy. Particularly the true positive rate is 89.17%, the true negative rate is 98.5% and the over all prediction accuracy is 98.14%. Finally, we compare our performance with that of the basic ANN prediction model. We choose a three-layer ANN model with 512, 256 and 128 nodes respectively. We apply the same training and test processes with the same data sets, and show the performance results in Fig.7. The true positive rate is only 61.39% and the overall accuracy is only 82.77%. We can clearly see that our system well outperforms the ANN model.

V. Conclusion

In this paper, we have proposed a spectrum availability prediction system based on LSTM neural networks. We employ the LSTM networks to exploit the spectral-temporal correlation among historical spectrum availability data in order to make predictions. We have also carried out simulations to evaluate the performance of the proposed system. Results show that our overall prediction accuracy can be up to about 98%.

Reference

[1] Ian F Akyildiz, Won-Yeol Lee, Mehmet C Vuran, and Shantidev Mohanty. Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey. *Computer networks*, 50(13):2127–2159, 2006.

[2] Yuanhao Cui, Xiao jun Jing, Songlin Sun, Xiaohan Wang, Dongmei Cheng, and Hai Huang. Deep learning based primary user classification in cognitive radios. In *Communications and Information Technologies (ISCIT), 2015 15th International Symposium on*, pages 165–168. IEEE, 2015.

[3] Yuqing Huang, Hong Jiang, Hong Hu, and Yuancheng Yao. Design of learning engine based on support vector machine in cognitive radio. In *Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on*, pages 1–4. IEEE, 2009.

[4] Ming Li, Pan Li, Xiaoxia Huang, Yuguang Fang, and Savo Glisic. Energy consumption optimization for multihop cognitive cellular networks. *IEEE Transactions on Mobile Computing*, 14(2):358–372, 2015.

[5] Ming Li, Sergio Salinas, Pan Li, Xiaoxia Huang, Yuguang Fang, and Savo Glisic. Optimal scheduling for multi-radio multi-channel multi-hop cognitive cellular networks. *IEEE Transactions on Mobile Computing*, 14(1):139–154, 2015.

[6] Jarmo Lundén, Visa Koivunen, Sanjeev R Kulkarni, and H Vincent Poor. Reinforcement learning based distributed multiagent sensing policy for cognitive radio networks. In *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*, pages 642–646. IEEE, 2011.

[7] George Lane-Roberts Shared Spectrum Company Mark A. McHenry, Dan McCloskey. Spectrum occupancy measurements location 4 of 6: Republican national convention, new york city, new york august 30, 2004 - september 3, 2004 revision 2. http://www.sharedspectrum.com/papers/spectrum-reports.

[8] Shared Spectrum Company Mark A. McHenry, Karl Steadman. Spectrum occupancy measurements location 2 of 6: Tysons square center, vienna, virginia april 9, 2004. http://www.sharedspectrum.com/papers/spectrum-reports.

[9] Joseph Mitola. Cognitive radio—an integrated agent architecture for software defined radio. 2000.

[10] Dusit Niyato and Ekram Hossain. Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion. *IEEE journal on selected areas in communications*, 26(1), 2008.

[11] Manel Martinez Ramon, Thomas Atwood, Silvio Barbin, and Christos G Christodoulou. Signal classification with an svm-fft approach for feature extraction in cognitive radio. In *Microwave and Optoelectronics Conference (IMOC), 2009 SBMO/IEEE MTT-S International*, pages 286–289. IEEE, 2009.

[12] Richard S Sutton, David A McAllester, Satinder P Singh, Yishay Mansour, et al. Policy gradient methods for reinforcement learning with function approximation. In *NIPS*, volume 99, pages 1057–1063, 1999.

[13] Jian Tang, Satyajayant Misra, and Guoliang Xue. Joint spectrum allocation and scheduling for fair spectrum sharing in cognitive radio wireless networks. *Computer networks*, 52(11):2148–2158, 2008.

[14] Vamsi Krishna Tumuluru, Ping Wang, and Dusit Niyato. A neural network based spectrum prediction scheme for cognitive radio. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1–5. IEEE, 2010.

[15] Kok-Lim Alvin Yau, Peter Komisarczuk, and Paul D Teal. Applications of reinforcement learning to cognitive radio networks. In *Communications Workshops (ICC), 2010 IEEE International Conference on*, pages 1–6. IEEE, 2010.