

# A Spectrum Anomalies Diagnosis Method Based on Two - Dimensional Hidden Markov Model

Cheng Cheng<sup>1, a)</sup>, Yunfeng Jia<sup>2, b)</sup>

<sup>1</sup>School of Public Administration, Beihang University, Beijing 100191, China.

<sup>2</sup>School of Electronics and Information Engineering, Beihang University, Beijing 100191, China.

<sup>a)</sup>buaachengc@163.com

<sup>b)</sup>jiayunfeng@buaa.edu.cn

**Keywords:** Spectrum, Diagnosis, Hidden, Markov, Model

**Abstract.** Radio monitoring system can cover a range of area, where we can obtain the electromagnetic mode information, and the information we get is the result of a mixture of various radio transmission signals. We can't get the intrinsic characteristics of the electromagnetic spectrum. It cannot determine whether the spectrum interferes, and the interference type. This paper studied the two-dimensional spectrum occupancy modeling based on the two-dimensional hidden Markov (2D-HMM). The usual HMM method is unable to fully express and reveal the relationship between different time and frequencies. We proposed to construct a two-dimensional hidden Markov based model library of the electromagnetic spectrum to achieve the self-identification of abnormal electromagnetic spectrum method; Through simulations and experiments, we have verified the feasibility of the spectral abnormalities identification method based on two-dimensional hidden Markov model, it can recognize the co-channel interference, the source super level emission, the ultra-high-power adjacent channel emission and other typical type of exception in a large change interval and with a rather high accuracy.

## INTRODUCTION

Radio frequency monitoring in the spectrum of independent diagnosis theory and method research is spectrum management and radio monitoring areas urgently need to carry out the key issues. In the radio monitoring, the spectrum is abnormally autonomous. Many general of spectrum anomalies detection are presented for traffic of computer networks relying on data mining and information-theoretic techniques in [1][2]. According to the data of GSM900/1800 band and TV band spectrum, the Auto Regressive AR model is established. In [3-4] an auto regressive moving average (Auto Regressive Average Average) ARMA) model is proposed. In [5] the time-domain correlation of the occupied state of the spectrum is studied. By constructing the equivalent model of the state change of the spectrum occupancy state, the method based on the time series analysis requires that the original time series is a stationary time series. Non-stationary spectral occupancy model is usually used in the method of differential stabilization, such a smoothing method is not very good, which also directly affect the prediction accuracy and fit. In [6] a 5-continuous state Markov modeling method is proposed under the premise of coal quality access layer and Poisson distribution hypothesis. In [7] a two-state Markov chain model with spectral occupancy is established to describe the statistical law of the spectrum, but this method is in need of prior knowledge of the frequency spectrum usage of a frequency point. Need to specify the threshold. In [8-9], a method of modeling the spectrum sensing problem is described. The decision center can obtain a more accurate spectral state, but does not take into account the potential safety problems, and the spectrum anomaly problem is rarely involved.

This paper focuses on the research of spectral anomaly recognition based on hidden Markov (HMM). Based on the analysis of the requirements of complex frequency station monitoring and traditional one-dimensional spectrum occupancy, a two-dimensional Markov representation model is proposed, and a spectrum occupancy model based on two-dimensional hidden Markov model is

constructed and the implementation process In order to solve the shortcomings of the passive spectrum detection in the existing spectrum management, we propose to construct the spectrum usage pattern library based on the 2D hidden Markov model to realize the automatic detection and classification of the spectrum anomaly.

### SYSTEM MODEL

The nature of the spectrum anomaly recognition problem is a problem of pattern classification. In the exception recognition, different recognition status, representing the different patterns. To identify, first of all, to establish a variety of standard reference state, that is, the so-called normal and abnormal spectral state of the model, and then to identify the state with the comparison, and identify which reference model. The process of establishing a general standard state is called a classifier design, that is, a large number of samples according to certain rules for classification; identification process, the resulting classifier to identify the data for decision-making classification. The figure 1 is a general procedure for using the pattern recognition for abnormal diagnosis.

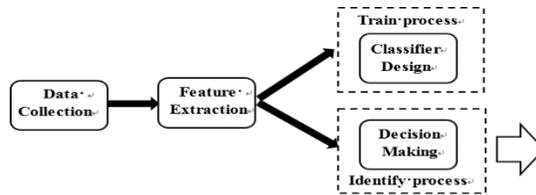


FIGURE 1. System model for spectrum anomaly detection

### Methodology

$N$ : The number of states of the external hidden Markov. The state is  $\theta_1 \dots \theta_N$ , At some point the external hidden Markov is in the state is  $q_t \in (\theta_1 \dots \theta_N)$ .

$\pi$ : The initial probability distribution vector of E-HMM is  $\pi = (\pi_1 \dots \pi_N)$ .

$$\pi_i = P(q_1 = \theta_i), 1 \leq i \leq N \tag{1}$$

$A$ : E-HMM state transition probability matrix is  $A = (a_{ij})_{N \times N}$

$$a_{ij} = P(q_{t+1} = \theta_j | q_t = \theta_i), 1 \leq i, j \leq N \tag{2}$$

$N^{(h)}$ : The number of states of the internal hidden Markov (I-HMM). The state at  $N^{(h)}$  is  $\eta_1 \dots \eta_{N^{(h)}}$ . The  $m$  parameter of I-HMM in the state is  $s_m^{(h)} \in (\eta_1 \dots \eta_{N^{(h)}})$ .

$\pi^{(h)}$ : The initial probability distribution vector of I-HMM.  $\pi^{(h)} = (\pi_1^{(h)} \dots \pi_{N^{(h)}}^{(h)})$ .

$$\pi_i^{(h)} = P(s_1^{(h)} = \eta_i), 1 \leq i \leq N^{(h)} \tag{3}$$

$A^{(h)}$ : The state transition probability matrix of I-HMM,  $A^{(h)} = (a_{ij}^{(h)})_{N^{(h)} \times N^{(h)}}$

$$a_{ij}^{(h)} = P(s_{m+1}^{(h)} = \eta_j | s_m^{(h)} = \eta_i), 1 \leq i, j \leq N^{(h)}$$

$B^{(h)}$ : Observation probability matrix is  $B^{(h)} = \{b_i^{(h)}(O_{t,m})\}$ .

Discrete I-HMM has  $w^{(h)}$  observations. The observed sequence is  $(V_1 \dots V_{w^{(h)}})$ .

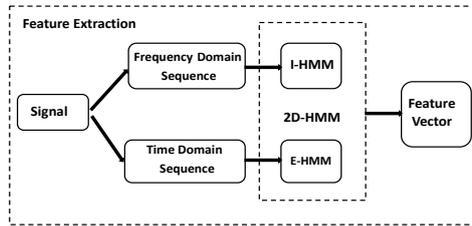
$$b_i^{(h)}(O_{t,m}) = \sum_{r=1}^{R^{(h)}} c_{ir}^{(h)} N(O_{t,m}, \mu_{ir}^{(h)}, U_{ir}^{(h)}), 1 \leq i \leq N^{(h)} \tag{4}$$

$c_{ir}^{(h)}$  Indicates that E-HMM is in state  $n$ , I-HMM The weight of the first Gaussian model in the state,  $N(O_{t,m}, \mu_{ir}^{(h)}, U_{ir}^{(h)})$  denotes a Gaussian function model with mean  $\mu_{ir}^{(h)}$  and variance  $U_{ir}^{(h)}$ .

A 2D-HMM parameter can be expressed as  $\lambda = (\pi, A, \Delta)$ ,  $\Delta = \{\Delta^1, \Delta^2, \dots, \Delta^N\}$   $\Delta^{(h)} = (\pi^{(h)}, A^{(h)}, B^{(h)})$

**Feature Extraction**

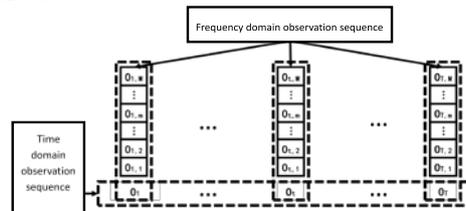
Based on 2D-HMM feature extraction, the process can be represented by the figure 2.



**FIGURE 2.** Feature Extraction Process

First, the signal needs to be preprocessed. The average power of the channel is calculated by the integral method, and the occupancy of the signal is calculated, the pre-processed signal is divided into frequency domain sequence and time domain sequence, The two-dimensional hidden Markov algorithm is used to extract the characteristics of the spectrum from the two angles of time-frequency domain.

The formation of the entire two-dimensional hidden Markov observation sequence can be represented by the following figure 3.

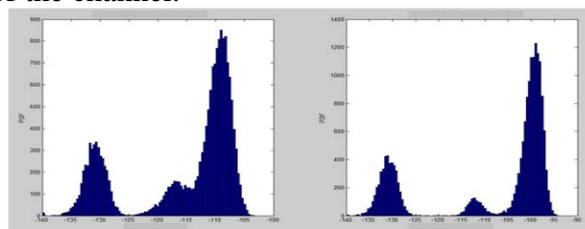


**FIGURE 3.** Formation of the two-dimensional HMM observation sequence

**Time Domain State Number**

The field strength data obtained by continuous monitoring of several typical channels in the radio service are obtained by calculating the Probability Density Function (PDF) of each channel is shown in fig 4.

The 861.7625MHz and 863.7875MHz PDFs consist of three Gaussian envelopes. The monitoring of the spectral field strength is the result of a mixture of multiple Gaussian models. The number of states of the E-HMM can thus be determined by the number of Gaussian models in the field strength distribution of the channel.



**FIGURE 4.** PDF of 861.7625MHz and 683.7875MHz

**Feature Extraction Based on 2D - HMM**

After two-dimensional hidden Markov classification. The original frequency domain data matrix is mapped to an area of the time-domain data is classified by the viterbi algorithm of two-dimensional hidden Markov, which can reflect the change of signal power with time. By using the external hidden Markov and the internal hidden Markov, the time and frequency dimensions of the signal can be obtained by two-dimensional classification of time and frequency dimension.

Through the two-dimensional hidden Markov method, the original spectral data will be mapped to a  $3 \times 3$  state space. The 2D-HMM model of the segment spectrum can be represented by the topology shown in Fig 5.

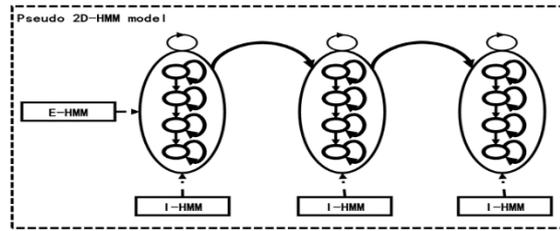


FIGURE 5. Pseudo dimensional hidden Markov model topology

### Feature Extraction Based on 2D - HMM

After two-dimensional hidden Markov classification. The original frequency domain data matrix is mapped to an area of the time-domain data is classified by the viterbi algorithm of two-dimensional hidden Markov, which can reflect the change of signal power with time. By using the external hidden Markov and the internal hidden Markov, the time and frequency dimensions of the signal can be obtained by two-dimensional classification of time and frequency dimension.

Through the two-dimensional hidden Markov method, the original spectral data will be mapped to a  $3 \times 3$  state space. The 2D-HMM model of the segment spectrum can be represented by the topology shown in Fig 5.

Where E-HMM is in the state of the moment is  $q_t \in \{1,2,3\}$ , The state of the first element of the I-HMM is  $s_m \in \{1,2,3\}$ , That is, the whole spectrum is divided into 9 categories. Each element can be represented by a binary array  $Q_{t,m} = (q_t, s_m^{(t)})$ . These nine cases are (1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3). (1, 3) represents that the channel is completely unoccupied, (3,3) represents the channel time domain is high occupancy, (2,2) on behalf of the channel frequency domain are low occupancy. This can be represented by Fig 6.

In the figure, nine sequence of states are ordered, 0.1 corresponds to (1,1), 0.2 corresponds to (2,1), 0.3 corresponds to (3,1), 0.4 corresponds to (1,2), 0.5 corresponds (2,3), 0.6 corresponds to (3,2), 0.7 corresponds to (1,3), 0.8 corresponds to (2,3), 0.9 corresponds to (3,3).

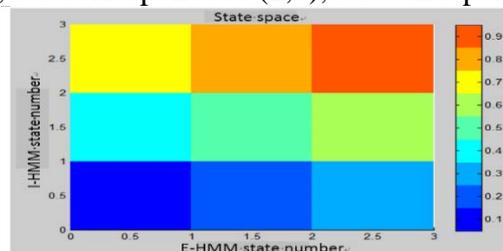


FIGURE 6. Dimensional state space color map

### Abnormal recognition

The abnormal behavior of the spectrum can be represented by a series of eigenvectors. The hidden Markov model is used to form the model, and the hidden Markov model is formed for different systems, and then these hidden Markov models are used to form a model base. When the system performs the anomaly recognition, the hidden vector of the unknown model is obtained, and the hidden Markov probability is obtained for the model in each model base. Then, according to the obtained probability, the system is judged to achieve the purpose of identifying the anomaly.

Generate HMM model library for typical exception models. Each exception type is marked as all kinds of abnormal and normal HMM model. These models form the model space  $\Omega = \{w_i\} (i = 0, 1, 2, \dots, M)$ . The feature vector  $X_t$  to be collected is input as input, which is input to the HMM model library. Find the conditional probability  $p(X_t | w_i)$  of each model of the input model in the model library. The posterior probability  $p(w_i | X_t)$  of each model is obtained. Finally, we compare the posterior probability, find the model corresponding to the maximum posterior probability, make the judgment of the abnormal model, and achieve the purpose of abnormal recognition.

## SIMULATION VERIFICATION

### Feature Extraction

The fig 7 is the original spectrum of the 851-866 MHz and the result of the 2D-HMM feature extraction. It can be seen that the original messy data becomes clear after the feature extraction.

After the feature extraction, the original messy data becomes clear, non-stationary data gradually stabilized, but also retains the original data feature information.

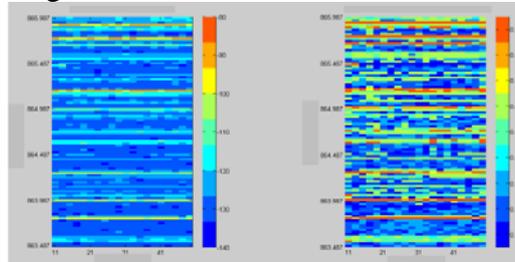


FIGURE 7. 2D-HMM feature extraction results of 863.4875MHz

### Abnormal recognition performance analysis

The selected mobile cluster communicates the data in the 851-866 MHz band and continuously monitors the 9-hour monitoring sample as experimental data. For the channel of the band, set the single source super-level transmission, two source super-level transmission, three-source super-level transmission, with the frequency interference (environment enhancement) and high-power adjacent frequency interference five abnormal types.

Suppose that the change in the Gaussian envelope of the channel probability distribution PDF is represented by  $\Delta E_{s_i}$

$$\Delta E_{s_i} = \mu_i + \sigma_i$$

(5)

Various exceptions can be expressed in TABLE 1.

TABLE 1. Typical spectral anomaly model parameter

Spectrum exception type	$\mu_i$ (dB)	$\sigma_i$ (dB)
Single source super - level emission(type1)	20	[0,15]
Interference with the same frequency(type2)	20	[0,15]
High power adjacent frequency interference(type3)	40	[0,15]

### Diagnostic results and algorithm performance analysis

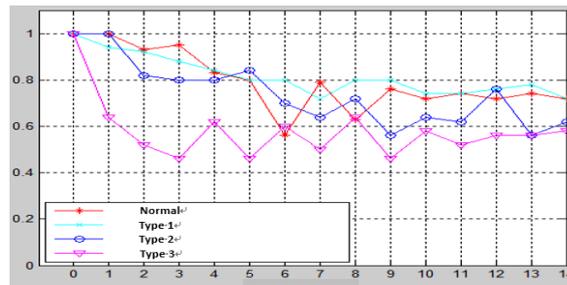
Training the normal spectrum at  $\sigma_i = 0$  and concentrating the 2D-HMM parameters of the typical anomalous spectral model and calculating the likelihood of each spectral pattern, as shown in table 2.

TABLE 2. Mobile Cluster 851-866MHz Spectral State Anomaly Recognition Maximum Likelihood Probability Calculation

Spectrum	normal	type1	Type2	Type3
normal	-8.47*e5	-8.54*e5	-8.50*e5	-8.50*e5
Type1	-8.55*e5	-8.54*e5	-8.56*e5	-8.55*e5
Type2	-8.55*e5	-8.57*e5	-8.49*e5	-8.56*e5
Type3	-8.50*e5	-8.56*e5	-8.50*e5	-8.50*e5

The algorithm can correctly identify the normal mode transmission, single source super-level transmission, the same frequency interference and high-power adjacent frequency interference.

In the actual measurement, due to radio wave propagation, multipath effect and other reasons, resulting in the detection of random signal is very large, abnormal mode will not be static. Set  $\sigma_i \in [1,15]$  to increase linearly. For each exception mode, 100 recognition is performed. We can get the spectrum anomaly recognition rate with the field intensity amplitude variance changes, as in fig 8.



**FIGURE 8.** The relationship between the anomaly recognition rate and the field mean amplitude variance

The discriminant algorithm has a relatively high recognition rate for the same frequency interference and for the field strength distribution variance is less than 5dB. The algorithm for the first three abnormal recognition rate can reach more than 80%.

## CONCLUSION

This chapter introduces the basic principle of spectrum anomaly recognition based on pattern recognition, and proposes a recognition method based on 2D-HMM, and gives the basic steps of using it to identify it. Then, based on the real measurement data 851-866MHz, combined with the common anomaly model, the model library of normal and several typical anomalies is established. Considering the problem of radio wave propagation and multipath effect, the accuracy of the algorithm is analyzed.

## ACKNOWLEDGMENTS

This work was financially supported by National Natural Science Foundation of China NO. 61371007.

## REFERENCES

- [1]. Adroulidakis, G. and Papavassiliou, S., "Improving network anomaly detection via selective flow-based sampling," *Communications, IET*, vol. 2, issue 3, pp. 399- 409, Mar. 2008.
- [2]. Adroulidakis, G., Chatzigiannakis, v., and Papavassiliou, S., "Using selective sampling for the support of scalable and efficient network anomaly detection," *Proc. IEEE GLOBECOM Conf.*, Washington, DC, USA, Nov.2007, pp.1 - 5.
- [3]. Wang Lei, Xie Shuguo, "A novel spectrum occupancy anomaly detection method based on time series analysis theory," *Proc. IEEE iWEM Conf*, Chengdu, China, Aug.2012, pp.1-4.
- [4]. WANG Zhe and Salous S. Time series ARIMA model of spectrum occupancy for cognitive radio [A]. *IET Seminar on Cognitive Radio and Software Defined Radios [C]*. London, IET, 2008. 1-4.
- [5]. Wang Lei. Analysis of Electromagnetic Spectrum Occupancy State Fluctuation Based on EGARCH Process [J]. *Journal of Electronics & Information Technology*, 2012, 34(11): 2767— 2773.
- [6]. Y. Xing, R. Chandramouli, and S. Mangold, Dynamic spectrum access in open spectrum wireless networks[J], *IEEE Journal on Selected Areas in Communications*, 2006, V24(3):626-637
- [7]. Lionel Gueguen and Berna Sayrac, Automatic Determination of Spectral States for Cognitive Radio[A], *IEEE Global Communications Conference (GLOBALCOM) [C]*, Institute of Electrical and Electronics Engineers, 2008:1-5
- [8]. T. Clancy and B. Walker, "Predictive dynamic spectrum access," in *Proc. SDR Forum Technical Conference*, 2006.