Research of Spectrum Monitoring Based on Signal Lifetime

Fuyin Zhao [a], Xueling Zhang, Tan Wang and Liang Chen

The State Radio Monitoring Center, Beijing 100037, China.

[a]zhaofuyin@srrc.org.cn

Abstract. In view of the fact that the traditional spectrum occupancy parameter cannot reflect the transmission of multiple signals within the same frequency, in order to master the more detailed usage of radio frequency, this paper takes radio signals as the object and puts forward the concept of the signal lifetime which presents the regularity of signal activity in a more detailed manner. Based on actual monitoring data, a signal lifetime automatic extraction process is designed, and big data technology is applied to predict signal future transmission time. Actual tests show that the prediction of signal future transmission time has a fairly high accuracy. This paper demonstrates a technical exploration of the fine management of spectrum resources and provides technical reference for the automation and intellectualization of radio monitoring.

Keywords: Radio monitoring, signal, signal lifetime, prediction, big data.

1. Introduction

In recent years, the rapid development of wireless communication technologies such as Internet of things and 5G has brought challenges to the management of radio spectrum resources: on one hand, it is necessary to plan appropriate frequency bands for new radio services from the scarce spectrum resources; on the other hand, it is important to guarantee the order of radio waves emitted by various radio services and avoid harmful interference. As an important technical measure of spectrum resource management, spectrum monitoring not only provides necessary data support for spectrum planning, but also plays a key role in radio interference analysis [1].

At present, the two parameters, frequency band occupancy and frequency channel occupancy, are mainly used to evaluate the use of the spectrum in the spectrum monitoring domestic and abroad, and they respectively analyze the utilization efficiency of the spectrum from the perspective of frequency domain and time domain [2]. As the physical carrier of spectrum resources, radio signal is the electromagnetic wave emitted or received by radio service and the direct object of radio interference analysis. However, the two parameters of frequency band occupancy and frequency channel occupancy cannot well reflect the radio signal transmission. At the same time, existing spectrum monitoring is mostly based on historical and current monitoring data, so it is impossible to determine the future use of signals, while predicting the future transmission state of signals is of great significance for spectrum monitoring.

Therefore, this paper takes radio signal as the basic object, puts forward the concept of signal lifetime, analyzes the regularity of signal's historical activity and predicts the regularity of signal's future transmission. Based on monitoring data, the automatic data extraction process for signal lifetime is designed, and the accuracy of predicting the signal's future transmission time is tested with typical radio service signals as an example.

2. Signal Lifetime

In the computer domain, lifetime refers to the existing time of an object, for example, from the time the object is allocated to memory to when it was removed from memory by the garbage collector. In the field of medicine, lifetime refers to the predicted survival time of a patient after suffering from a disease. In terms of the concept of lifetime described in the field of computer and medicine, lifetime means a time duration, which includes both the existing time and the possible existing time in the future. In the field of spectrum monitoring, it is very important to trace the radio signal's transmission history and predict the radio signal's future transmission time.

In reference to the concept of lifetime in computer field and medical field, this paper defines the signal lifetime as: the existence and predicted existence time of radio signal. On the basis of traditional spectrum monitoring and signal analysis, signal lifetime can be displayed in a more fine-grained manner, and can be used to predict the future use of signals, which is beneficial to maintain radio wave order and make spectrum planning.

3. Signal Lifetime Generation Process

The signal lifetime describes the radio signal's historical transmission regularity and future transmission regularity, and the main content of the research is the radio signal has appeared in what time period and may appear in what time period in the future. The signal lifetime stems from a large number of historical monitoring data, statistical signal's historical transmission regularity, and machine learning is used to predict the signal's future transmission regularity. According to the data processing procedure, the signal lifetime generation process can be divided into five steps: monitoring data acquisition, signal search and discovery, signal survival state recording, signal's historical transmission regularity statistics and signal's future transmission regularity prediction, as shown in figure 1.
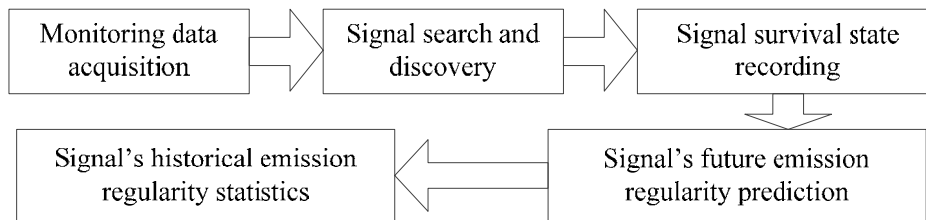

Fig 1. generation process of signal lifetime

Step 1: monitoring data acquisition

The original spectrum monitoring data is from radio receiver or spectrometer receiving air-wave signal directly through the antenna feeder, which is the most authentic and direct reflection of electromagnetic environment. Monitoring data is the raw data for signal lifetime extraction, and is also the direct data source of signal searching. In order to ensure the accuracy of tracing the history state of the signal in detail, in this paper, the monitoring data acquisition must follow two principles below: 1) It must be based on feature of radio services and use appropriate resolution in different frequency bands not to miss any signal in the searching phase, at the same time reducing the amount of calculation; 2) Ensure that the continuity of data meets the long-term continuous acquisition requirement. After the monitoring equipment finish scanning one frequency band, it forms a frame of electrical level, such as $\{x_1, x_2, x_3, x_4, …, x_m\}$, which is continuously collected to form a frame-by-frame monitoring data stream.

Step 2: signal search and discovery

The signals need to be accurately extracted from the original monitoring data stream. The commonly used methods are absolute threshold level and dynamic threshold level based on background noise. As background noise fluctuates with time and external environment [3], the absolute threshold level extraction method requires manual adjustment of the absolute threshold level according to the variation of environment. In this paper, a dynamic threshold level method based on background noise is used to extract signals in order to improve automation. Compare the electrical level of each frame in real-time spectrum monitoring data stream with the calculated threshold data, and take the level data as a signal when they are continuously greater than threshold data, such as $\{x_2,x_3,x_4,x_5,x_6\}$, and then calculate the signal center frequency f_i, bandwidth b_i, and the maximum level x_{max}, and finally, the signal data with time information t_j can be obtained, such as $\{t_j, f_i, b_i, x_{max}\}$.

Step 3: signal survival state recording

In this paper, a signal database is designed to record the signal survival states, including the central frequency f_n, the bandwidth b_n, the signal transmission time set $t_0, t_1……$, which forms $\{f_n, b_n, t_0,t_1\}$.

Since in signal lifetime study we focus on every single signal, it is necessary to distinguish different signals transmitted in the same frequency. Signal data contains two parameters, which is the center frequency f_i and bandwidth b_i. Considering the features of different signals in different services, two parameters are introduced, frequency redundancy Δf and redundancy bandwidth Δb, to differentiate services frequency band. At t_j, compare signal data $\{t_j, f_i, b_i, x_{max}\}$ and signal $\{f_n, b_n, t_0, t_1\}$ recorded in signal database, if f_i and f_n are both within range Δf, and b_i and b_n are both within range Δb, then f_i and f_n can be looked as the same signal at t_j; record t_j to f_n, and update f_n in signal database as $\{f_n, b_n, t_0, t_1, t_j\}$. Otherwise f_n can be looked as a new signal, and its information $\{f_i, b_i, t_j\}$ can be updated into signal database. After real-time calculation, the collected signal data with time information is continuously stored in the signal database, and the record of the signal historical survival state is completed. In addition, the signal database also completely covers data acquisition time $\{t_0, t_1, t_2, t_3......\}$, which can be readily used to generate signal historical transmission regularity statistics.

Step 4: signal's historical transmission regularity statistics

Signal survival state recording is the basic data source of signal's historical transmission regularity statistics. The amount of data increases with the ongoing time of signal acquisition. Signal processing and database storage efficiency are studied in this paper. Considering processing speed, storage efficiency, and accuracy requirement of statistics in different radio services, for different radio services, processing intervals ΔT are set accordingly for signal historical transmission regularity statistics, namely for each ΔT interval time, signal's survival state is counted and stored within the ΔT. For instance, with a certain service frequency $[f_k, f_l)$, and ΔT of 15 minutes, the statistical method for historical transmission regularity is illustrated as follow:

In every 15 minutes, when $f_k \leqq f_n < f_l$, the central frequency f_n in the signal database is taken out. According to the signal transmission time set $t_0, t_1, t_j......$ recorded in f_n, signal transmission time set within the 15 minutes such as $\{t_1, t_j\}$ is concluded, and the signal survival time Δt (ms) in the interval time ΔT can be counted.

$$\Delta t \ += \ (t_j \ - \ t_{j-1}) \tag{1}$$

In this formula, t_j is one unit in the signal's transmission time set within the interval time ΔT, t_{j-1} is the previous to t_j recorded in database for signal data acquisition. In this case, signal survival time $\Delta t = (t_1 - t_0) + (t_j - t_{j-1})$, and signal's historical survival time set such as $\{f_n, b_n, \Delta t_1, \Delta t_2......\}$ can be calculated every ΔT.

Step 5: Signal's future transmission regularity prediction

The signal lifetime not only describes signal's historical transmission regularity, but also predicts signal's future transmission regularity according to its historical transmission regularity. Due to the continuous spectrum data acquisition, there would be a huge amount of data obtained from the signal's historical transmission regularity. In this paper machine learning technology is used on the big data platform to realize the prediction of signal's future transmission regularity. Machine learning generally consists of training stage and prediction stage [4], and its process is shown in figure 2.
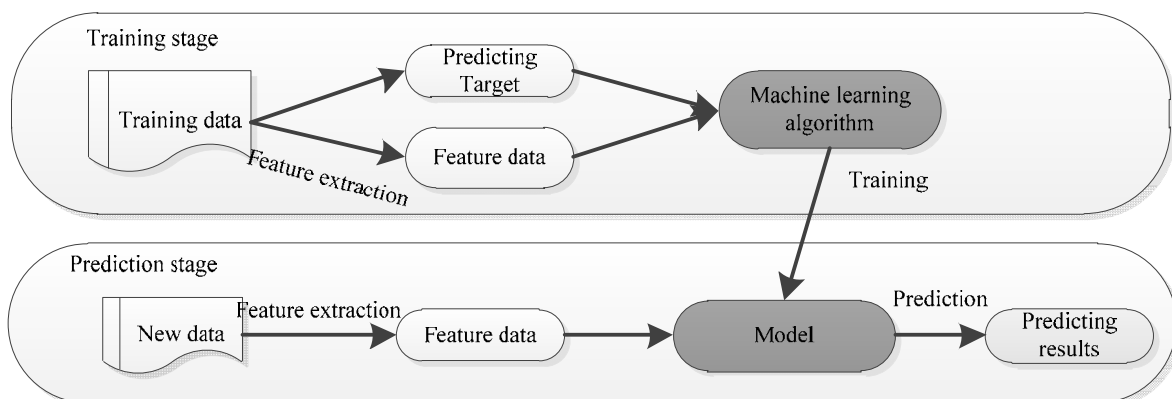

Fig 2. Two stages of machine learning

The training stage should go through the steps of feature extraction, generation of feature data and prediction of target samples. After the training of machine learning algorithm, the model will be generated and then enter the prediction stage. In the prediction stage, new feature data can be extracted and produced based on the newly input data, and then the model generated in the training stage can be used to predict and generate the prediction results.

The data obtained from the statistics of signal's historical transmission regularity are huge and dynamically expanding, which requires the machine learning algorithm to have good scalability and parallelism. Random forest algorithm is an integrated learning method of multiple decision trees, which can not only realize the parallelization of data processing, but also the parallelization of training unit, with good scalability and parallelism. At the same time, the random forest algorithm has a good accuracy rate in the prediction of future signal's transmission time which is one of the binary classification problems [5]. Therefore, random forest algorithm is selected in this paper to predict the signal's future transmission time.

4. Prediction of Signal'S Transmission Time

Signal's future transmission time is an integral part of signal lifetime. Workers and equipment are assigned to capture and analyze signal according to the occurrence time of predicted signal which can reduce the labor cost and improve the efficiency of spectrum monitoring.

In the prediction of signal's transmission time based on random forest algorithm, the signal's history survival time data set of a target signal f_n, such as { f_n, b_n, Δt_1, Δt_2......}, is preprocessed to extract feature data. For example, in case of $\Delta T=15$ minutes, the extracted feature data items are shown in table 1:

Table 1. feature data of signal's transmission time prediction

Feature	Value range	Sample
The current time of the day	0-95	8
Weekday	0-6	2
The probability of that moment in the past week	0-100%	71.4%
The sequence of which the moment comes up in the past week	seven bits by 0/1	1101101
The probability of the preceding moment in the past week	0-100%	71.4%
The probability of the last two preceding moment in the past week	0-100%	57.1%
The sequence of which of the last T time series in the current time	four bits by 0/1	0011
The interval from last time	≥0	0
The longest continuous time	≥0	5
Lable	0/1	1

In table 1, within a time interval ΔT, if a signal survival time Δt is greater than 60000 milliseconds, the signal is defined to exist in ΔT, represented by 1, otherwise by 0. Whether it shows up a label in the current time can be represented by 1 or 0.

History survival time data set is used to calculate and generate the feature data and target sample of f_n, which is used as a training data sample used by random forest algorithm to generate a binary classification model. Then the feature data of the current time is regarded as the new data to predict whether f_n exist in the current time. And then calculate the feature data of the next moment based on the current prediction to predict whether f_n exist in the next moment. Put the previous process in several cycles so that signal's transmission time for a period of time can be predicted. Figure 3 shows an example of the prediction result of a signal's transmission time in the next 24 hours. A column indicates the presence of a signal at this time.

Fig 3. Schematic diagram of the signal's transmission time prediction results

A part of radio signals on AM broadcast, FM broadcast, shortwave amateur service and maritime mobile service are chosen for acquisition and storage of signal lifetime, and the signal processing interval *ΔT=15*minutes. In this paper, Spark [6] was used on a Hadoop big data platform to realize the prediction of signal survival time. The next 24 hours of signal's transmission time is predicted every day, and the prediction is made for 7 consecutive days. The prediction results are compared with the actual results and the prediction accuracy is shown in figure 4.
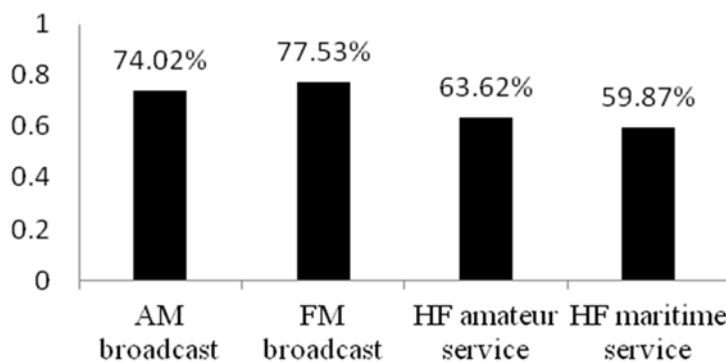

Fig 4. statistical diagram of prediction accuracy of signal's transmission time

It can be seen from figure 4 that all of the four kinds of radio services have a prediction accuracy over 50%, which has a certain practical significance. In particular, the prediction accuracy of signal transmission time of broadcast service is more than 70%. An alarm can be sent automatically when the predicted signal's transmission time is different from the actual signal's transmission time to help analyze whether there is any deviation in the transmission station, so as to improve the automation efficiency of radio spectrum monitoring.

5. Summary

This paper takes the radio signal as the object, proposes the concept of signal lifetime, which is used to display the spectrum usage in a more fine-grained manner, and designs the complete data extraction process of signal lifetime. In addition, the random forest algorithm is applied on the big data platform to predict the signal's future transmission time. The test concludes that the prediction results have a high accuracy and can improve the efficiency of spectrum monitoring. On the basis of the study in this paper, we can also carry out the research on signal classification, signal transmission correlation analysis and spectrum occupancy prediction etc. to further promote the efficient utilization of spectrum resources.

Acknowledgments

References

[1]. Zhou Hongshun. Handbook of spectrum monitoring. People's posts and telecommunications press. 2006, p.1-3.

[2]. Information on: www.itu.int.

[3]. Xia Feng, Xiaofeng Gong, Lidan Zhang, et.al. Research of background noise extraction based on texture feature. Acta electronica sinica. Vol.37(2009) No .9, p.2092-2095.

[4]. Dagui Lin. Hadoop and Spark big data massive analysis and machine learning. Tsinghua university press. 2017. p8-10.

[5]. Yanli Liu. The review of random forest, Master's thesis, Nankai university. China. 2008. p.25-28.

[6]. Sandy Ryza, Uri Laserson, Sean Owen, et.al. Advanced Analytics with Spark. People's posts and telecommunications press. 2015. p.4-10.