

## The Study of Cognitive Radio Prediction Based on Big Data

Haoxiang Sun <sup>a</sup>, Changxing Chen, Yunfei Ling, Jiyao Huang and  
Xiangyang Lin

Graduate school, Air Force Engineering University, Xian 710051, China.

<sup>a</sup> 593264062@qq.com

**Keywords:** Big data, Prediction, Spectrum resources, Cognitive radio, Application advantages

**Abstract.** With the advent of the age of big data, the application of big data to predict is a powerful way to solve various problems. However, for the increasingly tense spectrum resources, the prediction of cognitive radio based on big data is an inevitable trend. In this paper, the advantages of big data prediction technology in each process of cognitive radio are discussed, and the technical process and common prediction algorithms of cognitive radio based on big data are briefly described.

### 1. Introduction

With the progress of the era, people begin to realize spectrum resources in the communication transmission increasingly tension, due to the restrictions of the characteristics in the spreading of radio, the spectral range which is easy to use still is too small in the current technical conditions, the emergence of cognitive radio technology have greatly improved the utilization of electromagnetic spectrum.

At the same time, today is not only the era of information, but also the era of data. With the improvement of computer performance and the rapid development of network technology, the application of big data in people's daily life is also increasingly widespread. Known as the "new oil of the future", big data can be extract visual conclusions from a large number of complicated and seemingly unrelated things through a series of analytical methods, such as the connection between things and the law of their development.

The use of radio is increasingly complicated in daily life. Therefore, the application of big data in the field of cognitive radio will become an inevitable trend. Data mining through different means will greatly help the whole process of cognitive radio.

The establishment of the Radio Environment Map (or REM for short) is a typical application of big data. Through the continuous collection of the data of electromagnetic environment in a certain spatial range, and combining with the electronic map, to get the integration of time information, frequency information and Geographic information and many other elements of the large database shown in Fig.1. It can be used to establish mathematical models to explore the inter-relationship between the various elements and to achieve the forecast in a certain period of time for the radio environment.

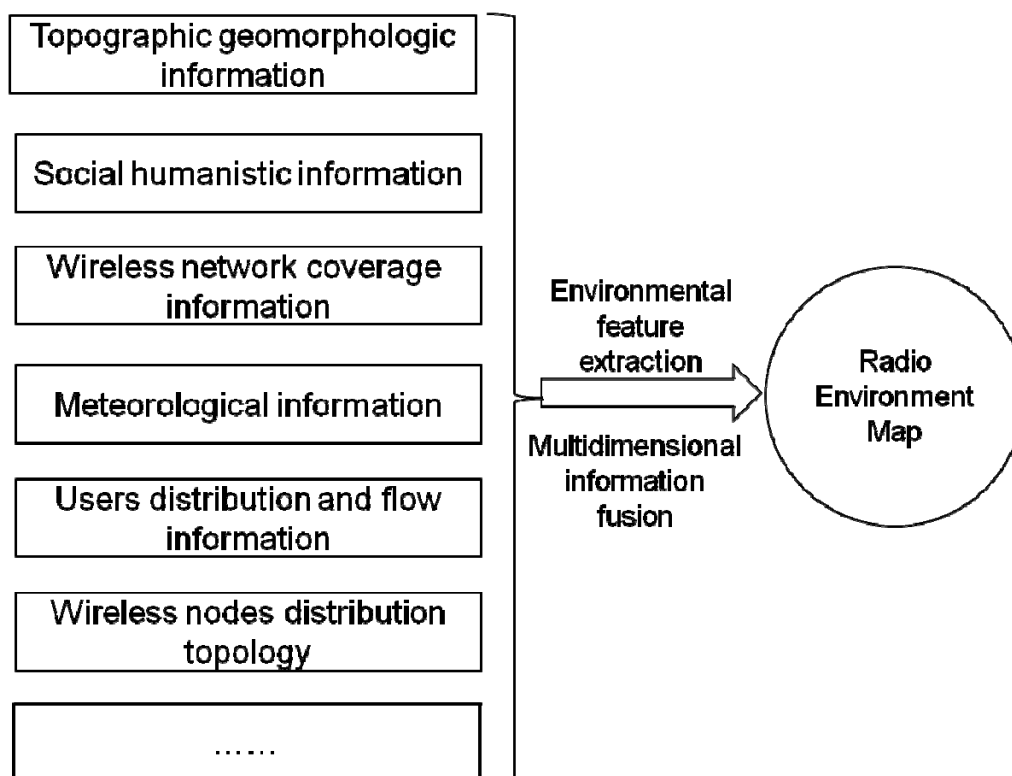


Fig. 1 The construction process of the radio environment map

This article will mainly talk about the application advantages of big data and the corresponding methods in data mining from three aspects as follows: spectrum sensing prediction of cognitive radio, interference perception prediction, spectrum analysis learning and optimization of spectrum decision.

## 2. Application Advantages of Spectrum Sensing Prediction Based Big Data

Spectrum sensing refers to sensing whether a primary user signal exists in a channel of different frequency bands within a certain spatial range, so as to discover whether there is radio "spectrum hole" in the radio environment that is not utilized or not fully utilized in the current time and space.

Prediction is based on the existing information, according to some certain methods and laws to measure things in the future, in order to understanding the process and results of the development in advance.

Spectrum sensing prediction refers to establishing a corresponding mathematical model based on the Radio Environment Map and a certain algorithm to analyze the usage rules of radio frequency spectrum in the current time and space, so as to be able to predict the development of the radio environment in a certain period of time in the future and predict the state of the tested channel, free bandwidth and many other information.

Traditional radio spectrum sensing technologies usually detect the target channel through certain sensor devices in real time or sends inquiring information to the target channel to obtain the target channel state by obtaining the feedback. When it is detected that the channel is free, the secondary user is allowed to access at this moment, otherwise, keeping detection and not accessed. On the one hand, the shortcomings of these technologies are not good enough at real time which will affect the communication rate; on the other hand, when the secondary users are accessing which detect the access of the primary users, signals, the exit of the secondary users may be inevitably collide with the primary users which will interfere with primary users, signals. Using the spectrum sensing prediction based on big data can look for a certain regularity of the access of primary users to make full preparation before the access and exit of secondary users and combine with the traditional detection methods to make full use of the channels, free state in the spectrum sensing process and it will be not easy to causes interference for the primary users, signals.

### 3. Application Advantages of Interference Perception Prediction

In addition to detecting the channel status and judging whether the primary user exists, the main task of spectrum sensing is also extremely important for the perception and judgment of interference in cognitive radio. The existence of electromagnetic interference affects spectrum perception on the one hand and affect the radio transmission on the other hand. According to the sources, the current interference need to be solved mainly includes two kinds, one is the inherent interference of the environment existing in the electromagnetic space, and the other is the collision interference between the secondary user and the primary user in the cognitive radio working due to failure to coordinate in time.

The inherent interference of the environment existing in the electromagnetic environment is an important part of the cognitive radio process that needs to be studied and grasped in depth. It exists almost along with the discovery of electromagnetic phenomena and has been affecting the applications of electromagnetic communications all along. In general, the inherent interference of the environment is mainly divided into two types of natural interference and human interference. Natural interference includes electromagnetic noise from the atmosphere and radiant fields from the universe. Human interference mainly comes from some kinds of artificial devices out of the electromagnetic energy interference, both radio, television, mobile phones and other communications equipment and radar and other active sources of intentional emission of electromagnetic energy sources of interference. And there are some unintentional interference sources that radiate electromagnetic energy due to their large power such as power lines, lighting, appliances and public transport, etc.

The application of big data technology through the acquisition of the radio environment of the current environment within a certain period of time by a number of sensor nodes to establish the radio environment map can effectively analyze the existence of environmental inherent interference and its variation over time. On the one hand, it can effectively reduce the shadowing effect of interference and reduce the false alarm probability and missed detection probability when it is perceived. On the other hand, it can eliminate the interference pertinently in the modulation-transmission-demodulation of radio signals.

At the same time, in view of human interference, by using of big data can more effectively predict the timing of the primary user signal to enter and prepare in advance.

### 4. Big Data and Spectrum Analysis Learning

When it comes to big data, people's first impression is the huge size of data. So, what is the big data? This does not have a definite numerical limit, but the use of big data there is a threshold, when the numerical value is greater than this threshold, it can be used to more effectively solve the problem, and less than this threshold will bring a lot of uncertain error, the analysis does not have enough reliability. Therefore, in order to ensure that the amount of data can reach this ambiguous threshold, the measure we usually take is constantly expanding our database. So, we need to continuously enrich and accumulate the data of radio environment map in the cognitive radio process.

However, as increasing of the amount of data, the data processing work what we need to do also becomes more and more difficult and complex, therefore it is not absolutely better for the amount of data is greater, so we also need do some constraints and weight-loss of data. At the same time, we have to ensure that the amount of data can guarantee the reliability of the result. Therefore, another important attribute of the data, value density, must be used here.

When we simply record the data, the amount of data is growing, but the value density what is the proportion of useful information hidden in the massive information has not changed, so we need to find ways to filter useful information and increase the value density.

This is mainly used in the spectrum analysis stage of cognitive radio. In this stage, the process of transforming the time domain signal to the frequency domain signal of the spectrum state obtained in the spectrum sensing stage is mainly used to obtain the energy distribution, the amplitude value distribution, frequency distribution and other distributions of the current spectrum. Through spectrum

analysis, quantitative interpretation of spectrum sensing results can be obtained. At the same time, we can more clearly determine the weight of each factor in the big data prediction process and make clear the fitting degree of the separate information in the radio environment map and the correct prediction result, and then filter useful information to remove useless information.

This process is also called the process of learning and training. The training is to adjust the weight of each factor in the prediction model by testing the test data. At the same time, the improved model can be used to test the sample data and judge the fitting degree of the sample data to remove useless data. In this way, the learning process is realized, so that the prediction algorithm of big data can be continuously improved during the process of using, thereby further improving reliability and improving efficiency.

## **5. Big Data and Optimization of Spectrum Decision**

There are mainly two kinds of spectrum allocation strategies, one is cooperative spectrum allocation, mainly taking into account the mutual interference between the various communication node, which will reduce interference and improve spectrum utilization and communication rate, but the disadvantages are the algorithm is complex and requires considering the impacts of many factors, and due to the coexistence of communication users will cause additional overhead; the other is non-cooperative spectrum allocation, that is, only consider the best available communication nodes, regardless of whether it will affect other nodes, which will undoubtedly have a great interference, and lead to a decline in spectrum utilization.

In the spectrum decision-making stage, the advantage of using big data is that it can realize the global coordination in terms of time latitude and spatial extent based on its predicted superiorities. In the time dimension, it can predict the free spectrum of a channel and the time length of the spectrum and distribution in time axis, so that it is possible to design the transmission process of radio signals, to find available communication channels at different time based on the availability of spectrum, and to design the length of the transmission data blocks for the predicted available time of the channel, so secondary users can exit in time to avoid conflict with the primary users when the predicted time reached. At the same time, other channels which is frequency band will be in free state according to the prediction result, so it may continue to transmit signals, thereby avoiding collisions and improving communication rates and spectrum utilization.

At the same time, there is an extremely significant advantage in spectrum allocation. By using the radio environment map, it is possible to macroscopically predict whether each node will generate mutual interference, interference degree and interference sources when signals are transmitted on different channels. Based on these information We can more conveniently select the appropriate cooperative nodes when implementing cooperative spectrum allocation, and we can also take measures to reduce the interference so that it can be transmitted under the non-cooperative spectrum allocation mode in the condition of certain transmission efficiency and cost to reduce interference.

## **6. Technical Methods of Cognitive Radio Based on Big Data**

### **6.1 The Process of Cognitive Radio Based Big Data.**

By referring to the data and combining with the characteristics of cognitive radio, the process of cognitive radio based big data is shown in Fig.2.

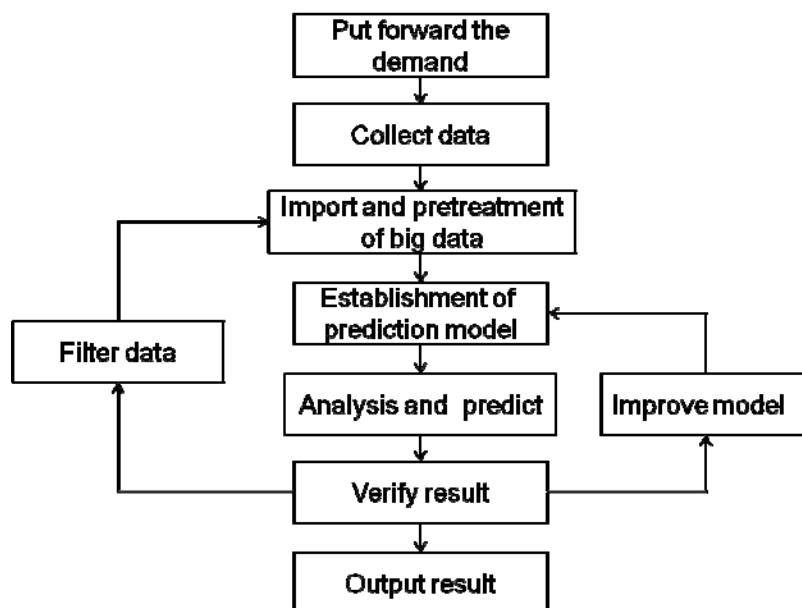


Fig. 2 The process of realizing cognitive radio prediction based on big data

First of all, we need to analyze the factors which affecting cognitive radio for the mission of cognitive radio that we need to complete, and to confirm the information we need to collect, such as geographical information, climate information, social environment, radio network type and channel characteristics of the selected area, wireless network topology and so on.

In this way, we can design the corresponding sensing system pertinently and obtain comprehensive and accurate scene information by means of geographic survey, meteorological telemetry, social survey, wireless channel sensors and environmental perception receivers.

Through the data collection of the perceptual system, we need to set up the data storage system to import the data from the front-end equipments to the centralized distributed database. However, in order to ensure that the stored data is neat and orderly, the data preprocessing needs to be completed before the data is stored. To analyze and find the relevance of the data to complete the data layering and classifying, so as to establish the corresponding data storage structure.

The core of cognitive radio prediction using big data lies in the selection of prediction algorithms and the establishment of prediction models. Choosing the appropriate algorithms for different cognitive goals and requirements is crucial for data mining and prediction of results.

After the building of the data storage system and the data processing system based on the predictive models, the existing data can be imported into the data processing system for analysis, and the prediction results can be obtained.

In order to ensure the accuracy of the prediction results, the results must be verified to some extent. The verification results can be used to screen the useful data on the one hand, and to improve the prediction model on the other hand.

## 6.2 Prediction Algorithm of Cognitive Radio Based on Big Data.

The cognitive radio based on big data refers to building a radio environment map based this large radio database and looking for effective information to complete the forecast target. The main methods are related analysis, classification analysis, cluster analysis and anomaly detection, etc. What, s more,it has developed a series of algorithm model on the basis of mathematics.

### 6.2.1 Link Prediction.

Static Link Prediction predicts the occurrence of a new link based on the current situation of radio propagation within a certain spatial range and analyzing the link information existing between nodes in the space which form a static link network.

However, the static link ignores the frequency characteristics of different situations at different time and cannot predict the repetitive occurrence of the old link. Therefore, in reference literature 2, there is a kind of prediction algorithm for a time series is proposed, which considers the effect of the link generation about the factor of time series. And at the same time, the weight of the static link in the

network diagram formed by the link will increase at each time to further predict the link generation at a certain time in the future and improve the prediction accuracy.

### **6.2.2 Markov Model.**

The Markov model is a statistical model, which means that the state obtained during the  $n$ th transition of a system's state transition is often determined by the result of the previous  $(n-1)$ th test and is therefore often used to analyze past of the data to predict the future results. The main application in the cognitive radio part of the Markov model can be established mathematical model about channels to predict the future state of the channels, which can get more optimized spectrum detection strategies, channel selection decisions and access into the strategies.

### **6.2.3 BP Neural Network.**

BP neural network is a multi-layer feedforward network based on error feedback. Therefore, BP neural network possesses strong nonlinear mapping ability and has the dynamic characteristics that can be adapted to the serious uncertain system through continuous learning. Therefore, it has achieved remarkable results in the field of cognitive radio for spectrum prediction. However, there are some problems in the BP neural network such as local minima, uncertain selection of initial weights, failure of prediction and slow convergence. In order to solve these problems, people are constantly looking for solutions without giving up the application advantages of BP neural network. For example, the improved BP neural network based on least square method (or LM for short) proposed by reference literature 7 and the differential evolution algorithm proposed in reference literature 6. Reference literature 12 proposed using an adaptive process to complete the weights, adjustment of neural network, which effectively improve the prediction accuracy.

### **6.2.4 Support Vector Machine Algorithm.**

Support vector machine (SVM) is a new type of machine learning algorithm developed on the basis of statistical learning theory. The basic idea of regression theory is that the data is projected to the high-latitude feature space for linear regression through nonlinear mapping. Therefore, there are significant advantages in solving the problems of nonlinear and high- dimensional models. So, it has good practicability, flexibility and accuracy for cognitive radio.

There are also many achievements for the improvement of SVM algorithms, such as least squares support vector machine (or LSSVM for short), which reduces the complexity of the algorithm by introducing the equality constraint and the least square loss function. Reference literature 9 proposes the use of gray relational to sample data and the least square support vector machine based on genetic algorithm is used to optimize the penalty coefficient and stone width by genetic algorithm.

In addition, the classification and hierarchical analysis of big data are also important analytical methods that can be used to help us find the correlation and causality of data, so that we can do the "cognitive" and "know why they are ", to do better in adjust the weights of influencing factors, and also adjust the intervention time and length of working time of each influencing factor in the time dimension.

## **7. Conclusion**

Through reading the literature, we can see that this technology still faces many problems. One is that it is difficult to acquire big data. For this point, we believe that with the development of the information age, the application of advanced technologies such as cloud storage of data will make us have more and more shared data; on the other hand, the prediction algorithms are complex, the technology is not mature and prone to error, so we still need to combine with traditional cognitive radio to complement each other and improve the accuracy and flexibility.

It is undoubted that the prediction technology based big data will play an important role in the application of cognitive radio, so that people can better manage the use of radio space and improve the utilization of spectrum resources.



## References

- [1]. Lu Benfu, Chen Jian. Research on Big Data Prediction and Related Issues [J]. Science and Technology for Development, 2014(1):60-65.
- [2]. Guo Jingfeng, Dai Junli, Ma Xin, et al. Time series link prediction algorithm for communication social networks [C]// China Database Academic Conference. 2009:552-559.
- [3]. Huang Chuan, Zheng Baoyu. Research on channel state prediction algorithm for multi-radio multi-channel environment in heterogeneous networks[J]. Journal of Electronics and Information Technology, 2010, 32(1): 178-182.
- [4]. Huang Chuan, Zheng Baoyu. A Novel Prediction Algorithm for Cognitive Radio Channel State[J]. Journal of Huaqiao University (Natural Science), 2010, 31(5):521-525.
- [5]. Tang Wanbin, Yu Huogen, Li Shaoqian. Cognitive radio hybrid spectrum switching algorithm based on channel prediction[J]. Computer Engineering and Applications, 2012, 48(27): 17-21.
- [6]. Li Shurong. Research on cognitive radio spectrum prediction based on DE-BP neural network[D]. Southwest Jiaotong University, 2013.
- [7]. Chen Binhua. Research on Spectrum Prediction Algorithm in Cognitive Radio System [D]. Beijing University of Posts and Telecommunications, 2011.
- [8]. Tumuluru V K, Wang P, Niyato D. Channel status prediction for cognitive radio networks[J]. Wireless Communications & Mobile Computing, 2012, 12(10):862-874.
- [9]. Deng Jun, Xiao Yao, Hao Yanjiao, et al. Radio interference prediction algorithm based on support vector machine[J]. Journal of Electrical Machines and Control, 2017, 21(8):18-24.
- [10]. Xu Yuan, Lu Huaxiang, Chen Xu. A Spectrum Prediction Method for Cognitive Radio Based on Support Vector Machines[J]. Telecommunication Science, 2014, 30(11):87-92.
- [11]. Chen Hao, Chen Wei, Gao Zhan. An Improved Weighted Support Vector Machine Prediction Algorithm in Cognitive Radio[J]. Electronic Quality, 2016(5):1-4.
- [12]. Chen Shuhua, Fu Qing, Ma Guilong, et al. Harmonic detection based on neural network adaptive prediction algorithm [J]. Chinese Electrotechnical Society, 2011, 26(s1):200-206.
- [13]. Liu Fen, Guo Yide. Fuzzy time series prediction algorithm based on interval similarity[J]. Journal of Computer Applications, 2013, 33(11): 3052-3056.
- [14]. Tang Zhoujin, Ren Feng, Peng Tao, et al. Least squares support vector machine predictor for chaotic time series based on iterative error compensation[J]. Acta Phys. Sin., 2014, 63(5):050505-10.
- [15]. Chen Yu, Yang Junan, Wu Yanhua. A Spectrum Blind Sensing Algorithm in Cognitive Radio [J]. Systems Engineering and Electronics, 2009, 31(6): 1311-1313.
- [16]. Cheng Yanyun, Zhang Shouchao, Yang Yang. Research and Application of Time Series Prediction Based on Big Data[J]. Computer Technology and Development, 2016, 26(6): 175-178.
- [17]. Zhang Wei, Huang Nannan. Research on Large-scale Antenna Channel Estimation Based on Rail Traffic Radio Environment Map[J]. China New Communications, 2016, 18(5):110-113.
- [18]. Wang Jiye, Ji Zhixiang, Shi Mengjie, et al. Research on demand analysis and application of intelligent large-scale R&D data[J]. Proceedings of the CSEE, 2015, 35(8):1829-1836.
- [19]. Haykin S. Cognitive radio: brain-empowered wireless communications[M]. IEEE Press, 2006.

- [20]. Xing X, Jing T, Cheng W, et al. Spectrum prediction in cognitive radio networks[J]. IEEE Wireless Communications, 2013, 20(2):90-96.
- [21]. Aslam S, Lee K G. Spectrum sharing optimization with QoS guarantee in cognitive radio networks , [J]. Computers & Electrical Engineering, 2013, 39(7):2053-2067.
- [22]. Iii J M. Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio, PhD Dissertation[J]. 2000.
- [23]. Wu X, Zhu X, Wu G Q, et al. Data Mining with Big Data[J]. IEEE Transactions on Knowledge & Data Engineering, 2014, 26(1):97-107. 24.
- [24]. Che D, Safran M, Peng Z. From Big Data to Big Data Mining: Challenges, Issues, and Opportunities[M]// Database Systems for Advanced Applications. Springer Berlin Heidelberg, 2013:1-15. 25.
- [25]. Rossi M, Brunelli D. Forecasting data centers power consumption with the Holt-Winters method[C]// Environmental, Energy and Structural Monitoring Systems. IEEE, 2015:210-214.