Research on equipment fault prediction expert system based on big data dimension reduction

Ningning ZHU1, Bin SONG2, Yan LU1 and Yingtang LI1 and Zhongzheng ZHU1

1School of Computer and Communication, Lanzhou University of Technology, Lanzhou, 730050, China;
2Lanzhou Chemical Research Center, China Petroleum and Petrochemical Research Institute, Lanzhou, 730050, China
E-mail: 307516638@qq.com

Abstract: Due to the complex structure and massive volume of large equipment in petrochemical enterprises, it was very difficult to carry out inspection and maintenance work. The traditional expert system used a single knowledge base and reasoning machine, so the processing efficiency and prediction accuracy were low. In this paper, the dimensionality reduction which is used to process a large amount of physical information collected by the sensors, remove redundant components and extract main characteristic parameters. Meanwhile, combined with the powerful pattern recognition and judgment ability of fuzzy neural network, the forward reasoning mechanism and error back propagation keeping continuous training correction until meeting the accuracy requirements, the expert system gave the equipment fault prediction conclusion. Finally, taking main fan oil station circulating oil pump and motor of heavy oil catalytic cracking unit and nine stages of bearing damage as monitoring objects, the fault predictions was implemented by the following steps: data collection, feature extraction, dimension reduction of big data and expert system with the fusion of fuzzy neural network, and it is proved that the expert system with fuzzy neural network based on dimension reduction can greatly improve the convergence speed.

1. Introduction
With the rapid development of big data and artificial intelligence technology, factory production is gradually transformed to the direction of scale, efficiency and intelligence. The structure of various mechanical equipment as the core of production is increasingly complex, the utilization rate is more frequent, so equipment failures often occur with the intricate evaluation indicators, the random running state of equipment and particular signal processing[1]. The expert system, in addition, can make full use of the existing expert experience, and deduce by simulating the expert's logical thinking mode, so as to solve the problem of relying solely on human experts for the fault prediction of complex equipment subject to the limitation of space and time or other empirical factors[2]. In literature 3, the fault diagnosis of rolling bearings based on generalized fine composite multi-scale sample entropy and optimized support vector machine is proposed, which improves the classification accuracy, but the speed is slow. In literature 4, an improved wavelet neural network is proposed to extract the characteristics of signals by using the second generation wavelet analysis. In reference 5, a one-dimensional convolutional neural network is used to extract fault diagnosis features of analog circuits.

Big data technology which plays an increasingly important role in the production of large enterprises such as electric power, petrochemical industry can analyze various types of data information in
production more efficiently[6]. The physical information collected by the sensor is integrated through dimensionality reduction, which pre-processes the data entering the expert system.

Combined with the experience threshold of the existing historical data and the law of function approximation of fuzzy reasoning rules to simulate the human mind’s fuzzy logic thinking and neural network’s self-learning ability and update the knowledge base and reasoning machine, the expert system based on dimension reduction and fuzzy neural network sets up equipment classification of anomalies and fault prediction data modeling to simulate the expert reasoning methods and speculate the location and cause of failure. The purpose of building an equipment fault prediction system based on dimension reduction and fuzzy neural network is to provide a more efficient, safer and more economical guarantee for large-scale and intelligent production. This paper also proves its superiority in speed.

2. Equipment fault prediction method based on condition monitoring

2.1. Condition monitoring and fault prediction

Condition monitoring and fault prediction, as two different levels of equipment maintenance in large enterprises, can reflect the real-time change of physical information in the operating state of the equipment, and diagnose the equipment in normal or abnormal state. While in abnormal parts, it will predict the future state of the equipment[7].

Condition monitoring can be carried out by the sensors or related electronic devices, and the information collected will be monitored, transformed, analyzed and processed, combined with historical data predicting the current state or evaluating, recording or displaying amplitude, frequency, temperature, pressure, sound, such as quantity, and transmitting into the signal system to get the corresponding eigenvectors and judge the equipment running status. Condition monitoring of big data laid a foundation for implementing the scientific fault prediction[8].

Fault prediction is based on a large number of characteristic parameters acquired in real time by condition monitoring, with the specific working condition, process, environment and historical data on the site to determine the diagnosis scheme or predict the future operation status of the equipment by analyzing fault types, causes and specific locations for the existing faults[9].

2.2. Data collection and feature extraction

2.2.1. Data collection. Since the signals of the diagnosed equipment usually involve sound, temperature, vibration and pressure, etc., it is necessary to collect the required data from various physical fields. The initial state signal is selected by corresponding performance of different operating states of the equipment in sound, temperature, vibration, pressure and other aspects, with the diagnostic purposes and requirements of different equipment, which is the basis for subsequent fault information analysis and prediction.

2.2.2. Sensor. For data collection and extraction, appropriate sensors should be selected. This system first needs to collect all signals of the operating state of the predicted equipment through various sensors on the site, and establish a concise operating state information table.

When the sensor is used for engineering measurement, its selection is very important. It is usually selected in terms of the required parameters and frequency selection range. If the relevant value of acceleration needs to be obtained, the acceleration sensor is adopted. While it is necessary to measure the speed, a speed sensor is adopted. Generally, because the acceleration sensor is small in size, easy to install and able to measure certain dynamic parameters such as amplitude and frequency, it’s usually selected[10].

2.2.3. Feature extraction. A large number of collected initial signals often need to be extracted some important features through analog-to-digital conversion and transformation domain analysis, etc. Sometimes the signal waveform in the time domain cannot provide all the information, so it is often
necessary to convert the signal into the frequency domain for spectrum analysis. When some of the extracted features are redundant information, the dimension can be compressed by deleting or reducing the weight. Feature extraction is the basis of the dimensionality reduction algorithm.

A. Analog to digital conversion

Generally, the signals collected by sensors are analog, while the computer system is digital. Therefore, it is necessary to digitize the analog signals, namely A/D conversion. The digitization process is divided into three steps: sampling, quantization and coding\[11\].

(1) sampling: a series of discrete numbers are selected from the continuous time signal by the instantaneous pulse.

(2) quantization: the sampled signal is transformed into a discrete signal on the amplitude.

(3) coding: the quantized amplitude of the discrete signal is converted into a binary number "0" or "1".

B. Spectrum analysis

The main task of digital signal processing is to perform spectrum analysis, including correlation analysis and traditional analysis, the core of which is the discrete Fourier transform and the fast Fourier transform.

In order to achieve signal segmentation on the time axis, the formula of discrete Fourier transform after windowing is:

\[
X(k) = \sum_{n=0}^{N-1} x(n)e^{-\frac{2\pi nk}{N}}
\]

Compared with the discrete Fourier transform, the fast Fourier transform significantly improves the computation speed and saves time when the computation amount is large.

The method we adopted in the comparison and analysis of multiple sample data is shown in figure 1.

Figure 1. block diagram of signal analysis and processing of multiple sampled data.

3. Dimension reduction

In terms of the performance of different operating states of the equipment in sound, temperature, vibration, pressure and other aspects, after a comprehensive analysis of the collected physical information, the corresponding feature parameters are extracted by signal processing technology, but some features are irrelevant or have little correlation with equipment failure, so the dimensionality
reduction algorithm is used to reduce the extracted features. Dimensionality reduction refers to the reduction of big data in the high-dimensional space converted to the data information in the low-dimensional space through some mapping relation, that is, the establishment of a mapping function \( f : x \rightarrow y \), where \( x \) is the amount of input big data, and the vector form can be used. \( y \) is the low-dimensional vector after mapping; \( f \) can be linear or non-linear. After the conversion, the large amount of redundant information and noise components contained in the original big data are removed, which reduces the operational error and greatly improves the operation speed.

Dimensionality reduction deletes other features to select important feature subsets and turns original features into a small number of major new features. In this paper, the dimensionality of the real-time big data transmitted to the cloud is firstly reduced, and the collected data is analyzed by PCA (principal component analysis) to find the intrinsic structural characteristics of the data, and the data type with large weight is extracted and analyzed as the main influencing factor of fault prediction, as shown in figure2.

![Figure 2. schematic diagram of big data preprocessing.](image)

Dimensionality reduction, as a preprocessing method to feed data into fuzzy neural network and expert system, can reduce system complexity and improve system performance.

4. **Fuzzy neural network and expert system design**

4.1. **Fuzzy neural network structure and algorithm**

Because the controlled object in multivariate analysis is generally difficult to describe accurately. In this paper, the multi-input and multi-output fuzzy neural network is used to establish the mathematical model, which includes the four-layer structure: input layer, membership function layer, rule layer and output layer.

\[ x(n) \text{ and } x(n-1) \text{ represent the input value of the current moment and the previous moment; } \]
\[ x_{n+1} \text{ is the output value of the next moment after the fuzzy neural network training; } \delta^i \text{ is the network weight value.} \]

In this model, the fastest descent BP method is adopted, and the error function is defined as:

\[ e = \frac{1}{2} \sum_{i=1}^{n} (x_i - y_i)^2 \]  

(2)

The error back propagation mode of weight is:

\[ \Delta\delta^i = \eta \alpha^i \alpha_i^j \]  

(3)

\( \eta \) is the learning rate, \( \alpha^i \) is the local gradient, and \( \alpha_i^j \) is the output of the upper layer.

4.2. **Expert system**

Expert system which belongs to the category of artificial intelligence is a computer software and a kind of advanced fault prediction technology. The expert system inputs historical data and original expert experience into the knowledge base, and sends real-time detection data to the reasoning machine for processing and comparison after state recognition, so as to simulate the reasoning thinking mode of human experts to form a decision and reach a diagnostic conclusion[12]. The expert system owning the
ability of relearning can display, delete and store the analyzed feature vectors and prediction results, and then sends the previous diagnosis back to the knowledge base for training and correction, which improves the accuracy and effectiveness of prediction. The task of equipment fault prediction expert system based on state monitoring is to identify the future trend of fault prediction by monitoring the running state of mechanical equipment, finally it gives suggestions.

4.2.1. Knowledge base. Knowledge base, one of the key design of expert system, stores a lot of history data and expert experience as knowledge input, making corresponding processing can according to the actual functional requirements. Knowledge can be encoded data representation, which can apply text production rules to express the character symbol. Production rule reflects the eigenvalues and the relevant causal relationship, in virtue of the knowledge expression for its effective management.

\[ \alpha_1, \alpha_2, \ldots, \alpha_n \] and \[ \beta_1, \beta_2, \ldots, \beta_n \] are respectively the outliers and conclusions judged by experts when the failure occurs, constituting the decision table. 0 and 1 are used to represent the combination of no abnormal and abnormal features, as shown in table 1.

| Expert system Predict conclusion | Abnormal value | \( \alpha_1 \) | \( \alpha_2 \) | \( \ldots \) | \( \alpha_n \) |
|----------------------------------|----------------|-------------|-------------|----------|------------|
| \( \beta_1 \)                   | 1              | 0           | 0           | 0        |
| \( \beta_2 \)                   | 0              | 1           | 0           | 0        |
| \( \ldots \)                    | \( \ldots \)   | \( \ldots \) | \( \ldots \) | \( \ldots \) |
| \( \beta_n \)                   | 1              | 1           | 1           | 1        |

The knowledge base contains the data of the training sample and the threshold information of the neural network, as shown in figure 3.

![Figure 3. construction process of knowledge base.](image)

4.2.2. Reasoning machine. The inference machine builds a matching model to complete the reasoning process relying on the relevant data and experience values in the knowledge base. The reasoning machine which first extracts some feature vectors from the database as initial values and matches them through forward inference rules, records the numerical solution results and dynamically adjusts the network parameters until the error accuracy is met, and finally the causes and locations of faults will be obtained.

The output results of the fault prediction system based on dimensionality reduction and fuzzy neural network can also be input to the knowledge base as expert experience so as to implement the continuous updating of the knowledge base and improve the accuracy of fault judgment, as shown in figure 4.
5. Design and analysis of fault prediction system based on dimensionality reduction and fuzzy neural network

5.1. System prediction results
When a large number of physical quantities detected by the sensor are taken as input factors, there is a certain nonlinear relationship in each dimension. Dimensionality reduction is applied to the data preprocessing of the system. The main components (amplitude, frequency, phase, spectrum, speed, temperature, vibration) are taken as the input layer of the expert system, and the data are continuously trained and learned in the system. Finally, the output layer of the expert system outputs the prediction results.

In this paper, a fault prediction system based on dimensionality reduction and fuzzy neural network was established on the common equipment in large refinery enterprises, including main fan oil station circulating oil pump and motor of heavy oil catalytic cracking unit and nine stages of bearing damage.

Figure 5. system interfacemotor.
Figure 6. monitoring and prediction results of the common equipment in large refinery enterprises.

Figure 7. monitoring and prediction results of nine stage of bearing damage.

It can be seen from the fault prediction conclusion that the fault prediction system based on
dimensionality reduction and fuzzy neural network can give advice by analyzing machine sound, temperature, vibration, the components (amplitude, frequency) of the waveform and spectrum. Based on the analysis of the corresponding components in the nine stages of bearing failure, the bearing life is predicted and the Suggestions for repair or replacement are given.

5.2. Comparative analysis of the output prediction results of the three systems

By comparing the expert system, the expert system based on fuzzy neural network and the expert system based on dimensionality reduction and fuzzy neural network in the aspect of operation speed, the superiority of the design is verified and its practical significance is proved.

From the results of dimension reduction used in the system clearly as you can see, dimension reduction algorithm which analyzes principal components makes the system efficiency greatly increasing. With the steady accumulation of expert experience and the increasing of main component elements, the running result of the use of dimensionality reduction and fuzzy neural network has more obvious advantages, which is very suitable for the calculation problem of fault prediction. Meanwhile, the method can save computer memory and reduce resource waste.

6. Conclusion

In this paper, we have made detailed analysis on fault prediction expert system based on dimensionality reduction and fuzzy neural network from data acquisition, feature extraction, spectrum analysis, data dimension reduction, fuzzy neural network to expert system for the entire process, it is concluded that the superiority of fault prediction expert system based on dimensionality reduction and fuzzy neural network is verified.

In the big data problem of expert system used for fault prediction and status monitoring, good results can be obtained by preprocessing the big data first, and multi-dimensional data compression can be implemented by principal component analysis, which can produces the following results: (1) the time and space complexity of the system is reduced; (2) noise from data sets is removed; (3) the cost of extracting unnecessary features is saved; (4) the simpler model has stronger robustness on small data sets; (5) big data can be represented by fewer features; (6) the visualization of data is carried out. In this system, dimensionality reduction as the preprocessing method unites with the strong recognition and judgment ability of fuzzy neural network, and the feasibility of the scheme is verified by experiments.

Acknowledgements

This paper has been funded by NSFC project 51667014.

Reference

[1] Strite S and Morkoc H 1992 J. Vac. Sci. Technol. B 10 1237
[2] Nakamura S, Senoh M, Nagahama S, Iwase N, Yamada T, Matsushita T, Kiyoku H and Sugimoto Y 1996 Japan. J. Appl. Phys. 35 L74

[3] Cheng Y, Chen BY, Zhang WH 2019 Adaptive Multipoint Optimal Minimum Entropy Deconvolution Adjusted and application to fault diagnosis of rolling element bearings. IEEE Sensors Journal. 19 12153-12164.

[4] Berredjem T, Benidir M 2018 Bearing faults diagnosis using fuzzy expert system relying on an Improved Range Overlaps and Similarity method. Expert Systems With Applications. 108 134-142.

[5] Wang ZY, Yao LG, Cai YW 2020 Rolling bearing fault diagnosis using generalized refined composite multiscale sample entropy and optimized support vector machine. Measurement. 156 107574.

[6] Peng B, Liu ZQ 2007 Fault diagnosis of rotating machinery based on improved wavelet neural network. Vibration, Test and Diagnosis. 27 32-35.

[7] Yang HH, Chen M, Wang C 2020 Data-driven feature extraction for analog circuit fault diagnosis using 1-D convolutional neural network. IEEE Access. 8 18305-18315.

[8] Belhadi A, Zkik K, Cherraf K, et al. 2019 Understanding big data analytics for manufacturing processes: insights from literature review and multiple case studies. Computers & Industrial Engineering. 137 106099.

[9] Zeng J, Chen YF, Yang P, et al. 2018 A review of fault diagnosis of large-scale wind turbine. Power Grid Technology. 42 849-859.

[10] Waghen K, Ouali MS 2019 Interpretable logic tree analysis: A data-driven fault tree methodology for causality analysis. Expert Systems With Applications. 136 376-391.

[11] Li CK, Zhao DF, Wang CL, et al. 2018 Research on an automatic method of fault diagnosis knowledge in chemical process. Journal of Chemical Engineering University of China. 32 393-400.

[12] Ramamohan KN, Comesañ DF, Uniaxial GL. 2018 Acoustic vector sensors for direction-of-arrival estimation. Journal of Sound and Vibration. 437 276-291.

[13] Chen JP, Zou WW, Wu GL, et al. 2018 Key technology and application of photon analog-to-digital conversion. Journal of Shanghai Jiao tong University. 52 1220-1225.

[14] Cheng K, Lu ZZ 2020 Hierarchical surrogate model with dimensionality reduction technique for high-dimensional uncertainty propagation. International Journal For Numerical Methods In Engineering. 121 2068-2085.

[15] Si JP, Ma JC, Niu JH 2017 Intelligent fault diagnosis expert system based on fuzzy neural network. Vibration and Impact. 36 164-171.