Research on Pump Product Intelligent Fault Diagnosis System Based on Big Data

Yimei Xu1,*, Xiaoqing Xu2
12 School of Information Engineering, Xi’an Eurasia University, Xi’an, Shaanxi, 710065, China
*Corresponding author’s e-mail: xuyimei@eurasia.edu

Abstract. With the integration of information technology and manufacturing industry and the improvement of equipment monitoring data and computer computing ability, equipment fault diagnosis has entered the era of "big data". Using big data analysis method for fault diagnosis, the fault diagnosis model can learn its own characteristics and complete fault identification, so that the fault diagnosis is more intelligent and automatic on the basis of high identification accuracy. This paper mainly studies the application of fault diagnosis method and big data analysis method in pump product fault diagnosis.

1. Introduction
Today, the world has set off a new round of industrial transformation with the transformation and upgrading of manufacturing industry as the primary task. The deep integration of the new generation of information technology and manufacturing industry will promote the service transformation and product upgrading in the industrial field and reshape new competitive advantages[1].

As an important part of water conservancy and hydropower system, pump products have higher and higher requirements for safety and reliability. Most of the relevant national maintenance standards are very conservative. With the improvement of equipment design and manufacturing technology, material technology and operation scheduling technology, the relevant specifications and standards seriously do not adapt to the actual situation of the project. The implementation of the specifications leads to a large number of "over maintenance" problems in the industry. Frequent unnecessary disassembly and assembly not only destroys various gaps between dynamic and static running in of the unit, Moreover, most of the maintenance and disassembly are not as professional and standard as the construction and installation, resulting in inadequate reassembly and gap guarantee, forming potential safety hazards, affecting the service life of the unit, unplanned shutdown, resulting in major economic losses and even casualties[2].

Relying on big data and integrating high-quality upstream and downstream resources of the industrial chain, study fault analysis and diagnosis methods based on big data, and strive to establish an intelligent fault diagnosis system with independent learning ability, which is an effective means to prevent accidents and unplanned downtime, the development direction of equipment maintenance, an important part of enterprise intelligent operation and maintenance, and the implementation of "Made in China 2025" Development planning.
2. Failure diagnosis method

2.1. Knowledge driven fault diagnosis method
Knowledge-driven fault diagnosis method does not need quantitative mathematical model. It is based on mechanical principle knowledge and human experience knowledge. It can not only introduce more information about the diagnosis object, but also make full use of the experience knowledge of human experts. Therefore, this method is suitable for fault diagnosis of non-linear systems and complex large-scale systems. Depending on the structure, it can be divided into qualitative model-based approach and symptom-based approach. Methods based on qualitative model include sign digraph method, fault tree method and qualitative simulation. Symptom-based methods include expert system method, fuzzy theory method, etc[3].

Knowledge-driven fault diagnosis method requires a great deal of mechanical principle knowledge and expert knowledge as support, and most of them combine reasoning mechanism formed by expert system with parameter estimation and state estimation. However, there are obvious shortcomings:
- It is difficult to update the knowledge base.
- Inference mechanism has poor flexibility and is difficult to adapt to the updated knowledge base.
- Manual operation is required and human-machine interaction is complicated.

2.2. Data Driven Fault Diagnosis Method
Because the I/O signal model of the system has certain connection with the fault source in terms of amplitude, phase, frequency and correlation, when fault occurs, some signal processing methods and feature extraction methods (such as various spectrum analysis methods, correlation analysis, time series feature extraction methods, adaptive signal processing methods and information fusion) can be applied for fault detection and separation. It does not need to establish an accurate model of the object, avoids the difficulties in establishing the model, so it has strong adaptability.

![Figure 1. Classification of data driven fault detection methods](image)

2.3. Value-driven fault diagnosis method
With the continuous increase of industrial monitoring data and the continuous improvement of computer
computing ability. In the development of industrial 4.0, the deep learning method has been well inherited and developed in the field of fault diagnosis. Many scholars attempt to use the deep learning model to evaluate the operation status, diagnose faults and forecast the life of industrial equipment.

The method of equipment fault diagnosis based on deep learning mainly trains the deep learning model in the way of supervised learning according to the data monitored in different states during the equipment working process, and then uses the trained model to diagnose the fault. The general process includes data acquisition, data preprocessing, deep learning model building, deep learning model training and fault diagnosis using the model. During data acquisition, one-dimensional data such as vibration, temperature and other data can be collected, as well as two-dimensional data in the form of pictures. Data preprocessing mainly includes data cleaning, data integration, data protocol and data conversion. In order to obtain sufficient data training model, data enhancement may also be included. In the process of constructing deep learning model, network structures such as NN, RNN and ANN are widely used at present. In the process of model training, select the appropriate optimization function and adjust the corresponding data parameters to get an accurate diagnosis model[4].

3. Fault diagnosis system based on big data
Knowledge driven and data-driven fault diagnosis methods mainly rely on the richness of domain knowledge, the accuracy of diagnosis model and the completeness of data samples. These methods have the advantages of simplicity, interpretability and easy development, but they are easy to be disturbed by the environment. In the face of large-scale nonlinear and non-stationary complex systems, their effect is not satisfactory, and the interaction between data signals of complex coupled systems is variable. At the same time, with the exponential growth of monitoring data, the research on fault diagnosis is facing the great challenge of processing industrial big data.

The fault diagnosis method based on value driven industrial big data background can reduce the use of expert experience and reduce the cost of manual participation in the diagnosis process, improve intelligence and enhance the accuracy of fault diagnosis.

3.1. System architecture design
This topic builds an end-to-end fault detection system. The framework structure of pump product fault diagnosis system is shown in Figure 2.

![Figure 2. Framework structure of pump product fault diagnosis system](image-url)
3.2. Design of fault diagnosis module

Using the data integration and fusion technology and deep learning technology in big data analysis technology, find the relationship between multidimensional data by analyzing one-dimensional time domain, two-dimensional time domain, operation environment and inherent attribute data, so as to judge the type, location and severity of fault, find the cause of fault, provide remote guidance for equipment maintenance for maintenance parties, and shorten the maintenance time, Reduce maintenance cost; Before the failure, predict the remaining life of the product, the parts and devices that may fail in the short term, and do preventive maintenance to reduce the production delay and other losses caused by the failure[5][6].

1. Data preprocessing technology
   Data sources are diverse, including operation environment data, inherent attribute data, real-time operation status data, and data types and formats are also different, including structured, semi-structured and unstructured data; Historical discrete data and real-time continuous data acquisition; The data quality caused by sensor failure or network transmission is poor, which requires preprocessing such as de duplication, replenishment, de abnormality and so on.

2. Data analysis and modeling
   Data screening: in the face of a large number of data, select the data that can characterize the key factors. On this basis, take the average, maximum, minimum, variance, range, median, etc. to generate more feature quantities. Reduce the dimension by PCA principal component analysis. The data is divided into training set samples and test set samples, and establish the training classifier model.

3. Deep learning algorithm
   Industrial big data analysis methods include traditional statistical analysis algorithms, general machine learning algorithms, vibration analysis algorithms for rotating equipment, time series algorithms for time series data, text mining algorithms for unstructured data, etc.
   The system adopts the two-dimensional convolution neural network fault diagnosis model of signal. For the sampled samples, the video image is obtained through continuous wavelet transform, and then the time-frequency image is compressed as the input of the two-dimensional neural network model for training in the way of supervised learning. Avoid that only one-dimensional time domain will lack valuable information.

4. Conclusion
   The data fusion technology and deep learning technology of big data analysis technology are applied to
the fault diagnosis and analysis of industrial equipment, which can not only judge the type, location and severity of the fault, find the fault cause, provide remote guidance for equipment maintenance for the maintenance party, but also predict the fault and reduce accidents and unplanned downtime, And reduce the huge economic losses and casualties. The research results of this subject can help pump enterprises establish pump equipment fault diagnosis and prediction system, realize intelligent upgrading of pump equipment, improve product added value, improve service level and user experience of pump enterprises, and reshape new competitive advantages.

Acknowledgments
Sponsor: School Scientific Research Project of Xi'an Eurasia University: Research on Fault Diagnosis Technology Based on Big Data (2021XJZK02)

References
[1] Wei X.B., Lu X.J., Li T.B. (2021) Fault Analysis System Design of Real-time Monitoring of Large Pump Motor State, Electrotechnical Application, 09:43-47.
[2] Song G.D. (2021) Research on Fault Diagnosis and Preventive Management Method of Coal Mine Drainage Pump. Automation & Instrumentation, 06: 105-105.
[3] LI F., WANG Q., Hu J.X, Tian Y. (2021) Combined Data-driven and Knowledge-driven Methodology Research Advances and Its Applied Prospect in Power Systems, Proceedings of the CSEE,41: 4377-4390
[4] Dai J.Y., Tang J., Huang S.Z. (2018) Signal-based intelligent hydraulic fault diagnosis methods: Review and prospects. Mechanical Engineering, 32:1-22
[5] SIANO D, PANZA M A.(2018) Diagnostic method by using vibration analysis for pump fault detection, Energy Procedia,148:10-17.
[6] Chao Q., Tao J.F., Wei X.L. (2020) Identification of cavitation intensity for high-speed aviation hydraulic pumps using 2D convolutional neural networks with an input of {RGB}-based vibration data. Measurement Science and Technology, 31:105102