Simulation Modelling Practice and Theory 80 (2018) 104–127

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat

Fault-diagnosis for reciprocating compressors using big data and machine learning

Guanqiu Qi[a,b], Zhiqin Zhu[a,*], Ke Erqinhu[c], Yinong Chen[b], Yi Chai[d], Jian Sun[e,d]

[a] Collaborative Innovation Center for Industrial Internet of Things, College of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[b] School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ 85287, USA
[c] Research Center of China National Offshore Oil Corporation, Beijing 100010, China
[d] College of Automation, Chongqing University, Chongqing 400044, China
[e] College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China

ARTICLE INFO

Article history:
Received 15 September 2017
Revised 16 October 2017
Accepted 30 October 2017
Available online 3 November 2017

Keywords:
Reciprocating compressor
Big data
Cloud computing
Deep learning
RPCA
SVM

ABSTRACT

Reciprocating compressors are widely used in petroleum industry. A small fault in reciprocating compressor may cause serious issues in operation. Traditional regular maintenance and fault diagnosis solutions cannot efficiently detect potential faults in reciprocating compressors. This paper proposes a fault-diagnosis system for reciprocating compressors. It applies machine-learning techniques to data analysis and fault diagnosis. The raw data is denoised first. Then the denoised data is sparse coded to train a dictionary. Based on the learned dictionary, potential faults are finally recognized and classified by support vector machine (SVM). The system is evaluated by using 5-year operation data collected from an offshore oil corporation in a cloud environment. The collected data is evenly divided into two halves. One half is used for training, and the other half is used for testing. The results demonstrate that the proposed system can efficiently diagnose potential faults in compressors with more than 80% accuracy, which represents a better result than the current practice.

1. Introduction

Reciprocating compressors are widely used in petroleum industry. It is important to keep reciprocating compressors working properly. Reciprocating compressors used in offshore oil and gas production usually operate in a high-temperature, high-pressure, flammable, explosive, corrosive working environment. Due to the harsh sea working environment, it is difficult to perform maintenance, fault detection, and reparation on reciprocating compressors. Comparing with other types of equipment used on land, offshore reciprocating compressors have higher reliability requirements [1]. In history, the explosion of offshore oil platform caused a huge and irreparable loss, such as British North Sea Piper Alpha platform [2], and Deepwater Horizon platform [3]. It is important to identify and repair any defects of equipment in time. Any small fault that is not repaired in time may finally result in a disaster.

* Corresponding author.
E-mail addresses: guanqiuq@asu.edu (G. Qi), zhuzq@cqupt.edu.cn (Z. Zhu), keerqh@cnooc.com.cn (K. Erqinhu), yinong@asu.edu (Y. Chen), chaiyi@cqu.edu.cn (Y. Chai), cq_jsun@163.com (J. Sun).

https://doi.org/10.1016/j.simpat.2017.10.005
1569-190X/

Regular maintenance and fault diagnosis after a system failure are the two traditional methods that are still used today. Unfortunately, they cannot efficiently detect faults in advance to avert the disaster. Periodical human inspection is lack of timeliness in fault detection. Human inspection often focuses on exterior and cannot easily detect the interior faults. Thus, it is difficult for current methods to predict the potential faults in advance. Fault diagnosis is usually a post-processing method.

The installation and working condition of reciprocating compressor vary, and many parameters affect the working status of reciprocating compressor. In most cases, one type of signal cannot point out the potential faults. Actually, the interaction of different factors and parameters causes the faults of reciprocating compressor. Therefore, it is always difficult to map related parameters to corresponding faults on reciprocating compressor. With the development of online testing technology [4–6], real-time information needs to be processed by hundreds of engineers in offshore oil fields. The analysis costs are expensive.

Fault diagnosis of reciprocating compressors often uses a real-time signal-processing model [7–9]. Reciprocating compressors in offshore oil platforms usually work in a salty and foggy environment, and the harsh environment makes traditional solutions not accurate. In addition, the real-time information often contains different parameters, such as vibration, temperature, and displacement. The relationships among these parameters are complex as they affect each other. The external environment, such as temperature, currents, and so on, also affects the working status of reciprocating compressor. Thus, it is difficult to determine the root cause for a system failure. It is important to identify the relevant information in a large number of data from heterogeneous multiple sources [10,11].

As the heart of a drill platform, even any inefficient operation and unplanned shutdown of a reciprocating compressor are not acceptable, which can seriously harm the crude oil extraction. Traditional fault-diagnosis methods focus on monitoring, and they can only detect faults after the compressor fails to work. We cannot rely on the traditional measurement technology alone to guarantee the early detection of potential failures to prevent damages. This paper proposes a real-time fault-diagnosis system using the big-data and self-learning approach to minimize failures and their high repair costs. Based on a data-driven classification method, an automatic recognition model is integrated into the proposed fault-diagnosis system. The main contributions of this paper are as follows [12]:

1. A new fault-diagnosis model of reciprocating compressor is proposed, which integrates and customizes existing machine learning techniques;
2. A large size of data from reciprocating compressors is processed in cloud environment to perform the real-time data analysis;
3. A multiple-category SVM (Support Vector Machine) is developed for recognizing normal and faulty data to identify potential faults;
4. The proposed model is robust to the impact of changes in external environment, and can determine the normal working status of reciprocating compressor to ensure the high accuracy of fault identification; and
5. The effectiveness and efficiency of the proposed fault-diagnosis model are evaluated and verified by using the data collected from real reciprocating compressors.

The rest of the paper is structured as follows: Section 2 introduces the proposed framework; Section 3 presents the fault-diagnosis process; Section 4 discusses the robustness of fault diagnosis; Section 5 evaluates the proposed framework and analyzes experiment results; and Section 6 concludes the paper.

2. Architecture

Oil exploration along a long coastline with a large number of oil and gas reserves at the bottom of the sea requires many drilling platforms as the geographic conditions of the diverse ocean can vary a lot from surface water to deep oceanic trenches. Thus, compressors in different ocean areas face different conditions.

2.1. Data analysis architecture

Real-time fault diagnosis system is used to reduce the maintenance costs and improve working efficiency of equipment [12–15]. Until now, a large number of methods are based on pressure, vibration, and acoustic emission (AE) signals, and they have been used to diagnose faults in reciprocating compressors [1].

A large amount of data needs to be processed to monitor the status of compressors as each compressor generates about 3GB of data per hour. A drilling platform often has hundreds of compressors, but it is often located far away from land and does not have enough computing capacity to process large-size data. Thus, large volumes of data are sent from platforms to a cloud environment for processing [16].

Two types of data can be obtained from compressors: structured and unstructured data. Structured data is related to the status of compressors such as temperature, speed, and acceleration. Unstructured data is from video surveillance. This paper focuses on analyzing structured data. The proposed data analysis system has two parts:

1. **Learning part:** It analyzes data to develop a model for the prediction of future working status;
2. **Analysis part:** It uses the generated model to predict the status of compressors and identify potential faults.