An automatic update framework for fault diagnosis machine on BOG compressor

Guangyao Xie¹, Yuyun Zeng¹, Wenbo Zhu², Zhongkun Wu² and Jingquan Liu¹ *

¹Department of Engineering Physics, Tsinghua University, Beijing, 100084, China
²Qingdao LNG Co. Ltd., Sinopec, Qingdao, 266400, China
Corresponding author’s e-mail: jingquan@tsinghua.edu.cn

Abstract. The fault diagnosis on boil-off gas (BOG) compressor is necessary for the safety and economic benefits in operation. This paper proposed an automatic update framework for diagnosis machine on BOG compressor. Data discretization is applied to transform data through expert experience to Bayesian network. This framework is able to update diagnosis models from expert experience provided by operators.

1. Introduction
The boil-off gas (BOG) compressor is widely used for recycling the excessive boil-off gas of liquefied natural gas (LNG). The extra-low suction temperature of BOG compressor brings about great challenges to operational safety[1]. Research on fault diagnosis of compressors is necessary for the safety and economy of compressors. Once the machines are abnormal and fault, it will affect the normal production and bring huge economic loss. Recent research on compressors is based on data-driven methods which relies on historical failure data. K-means clustering algorithm with respect to the characteristics of manifold distribution of fault data was applied for reciprocating compressor fault diagnosis[2]. The information entropy SVM method was used in compressor valve fault diagnosis but have difficult to handle large data[3]. A method using Teager–Kaiser energy operation and deep belief networks (DBNs) is applied to classify the faults of compressor valves[4].

In this paper, an automatic update framework is proposed to diagnose compressor failure. A discrete fault knowledge base is established to solve the lack of valid fault data. Bayesian network using discretized data is applied to estimate probabilities of different fault modes. The probability of failure is used to determine whether current fault occurs.

2. Automatic update framework for fault diagnosis machine
An automatic updating framework based on expert experience for fault diagnosis machine is proposed to solve the problem of the lack of industrially valid fault data. A discrete fault knowledge base through expert experience is established to support for the construction of Bayesian network structure. By updating the expert knowledge base from industrial operators, the framework combines experience and diagnostic models. The framework of fault diagnosis machine is shown as in Figure 1.
2.1 Fault knowledge base construction
The main failure modes and performances of the BOG compressor are obtained by expert experience and literature research. Fault knowledge base is established by these failure modes and performances. Each fault knowledge includes fault component, fault mode, fault parameter performance and fault mechanism. A total of 17 failure modes named from F1 to F17 and 10 failure parameter performances named from S1 to S10 are used as 28 fault nodes to construct a Bayesian network.

The Compressor operating parameters are continuously distributed while Bayesian network needs discrete input value. According to fault knowledge base, parameters in a particular fault are discretized into five intervals: normal, high, low, extra high, extra low. Parameter is assumed to follow a normal distribution. Normalized value of parameter offers threshold of five intervals as shown in Table 1.

Table 1. Normalized values of parameter for five discretized parameter intervals.

| Discretized parameter intervals | Normalized value of parameter |
|---------------------------------|------------------------------|
| Extra low                       | $(-\infty, -4.5)$            |
| Low                             | $(-4.5, -1.5)$               |
| Normal                          | $(-1.5, 1.5)$                |
| High                            | $(1.5, 4.5)$                 |
| Extra high                      | $(4.5, +\infty)$             |

2.2 Bayesian network construction
Bayesian network can effectively solve difficulties due to randomness and uncertainty of power transformer fault diagnosis data [5]. Fault knowledge base offers information to construct a Bayesian network, including nodes, edges, and discrete input probability values. The construction of Bayesian network is composed of structure construction and probability calculation.

2.2.1 Bayesian network structure construction
The Bayesian network structure is constructed based on the fault nodes and parameters as shown in Figure 2. The nodes from F1 to F17 represent 17 fault modes in fault knowledge base. The nodes from S1 to S10 represent fault performance parameters in fault knowledge base. According to fault knowledge record, the directed edge in Bayesian network connects the fault mode node to the fault parameter performance. For example, the fault knowledge record is that the leakage of the first-stage intake valve leads to the first-stage suction temperature rise, the first-stage suction pressure rise and the secondary
exhaust pressure decrease. F1 node stands for “leakage of the first-stage intake valve” while S1, S2, S5 stand for “first-stage suction temperature”, “first-stage suction pressure” and “second-stage exhaust pressure” respectively. Three directed edges connect F1 and S1, S2, S5 indicate that fault mode F1 causes significant changes of fault parameter performance S1, S2, S5. Basic structure of “failure-parameter” Bayesian network model for BOG compressor is established.

![Figure 2. Bayesian network structure for BOG compressor fault diagnosis](image)

2.2.2 Probability calculation

After the network structure is established, joint probability distribution should be calculated to provide value for edge weight. Joint probability distribution is defined as $P(X|Y_{i=1,2,\ldots})$ where $X$ represents fault parameter performance and $Y$ represents different fault modes. Discretization of fault modes takes two values: 0 indicates normal state and 1 indicates fault state. Discretization of fault parameter performance takes five values consistent with fault knowledge base: 0 indicates extra low, 1 indicates low, 2 indicates normal, 3 indicates high and 4 indicates extra high.

The prior distribution of fault mode is determined by expert experience. For each fault mode, the prior distribution is $P(Y = 0) = 0.95, P(Y = 1) = 0.05$ where probability of normal state is 0.95 and probability of fault state is 0.05. Conditional probability distribution of single fault mode is calculated by discretization. Parameter is assumed to different normal distribution for five parameter performances with equal variance. The mean values of five parameter performance distributions are -6, -3, 0, 3, 6 respectively. The five intervals discretized to represent distribution are $(-\infty,-4.5), (-4.5,-1.5), (-1.5,1.5), (1.5,4.5)$ and $(4.5,\infty)$. In each parameter performance distribution, the probability of discretized intervals is calculated by summing the probability values of corresponding interval distribution. As shown in Figure 3, mean of parameter performance “Low” is -3. The probability values for five intervals are the sum of probability density and the values are 0.066, 0.866, 0.066, $3.39 \times 10^{-6}$ and $3.18 \times 10^{-14}$. The parameter probability distributions of single fault mode through calculation of different parameter performances are shown in Table 2.

| Parameter performances | Discretized probability distribution of parameter intervals |
|------------------------|----------------------------------------------------------|
|                        | $(-\infty,-4.5)$ | $(-4.5,-1.5)$ | $(-1.5,1.5)$ | $(1.5,4.5)$ | $(4.5,\infty)$ |
| Extra low              | 0.933           | 0.066         | $3.39 \times 10^{-6}$ | $3.18 \times 10^{-14}$ | 0             |
| Low                    | 0.066           | 0.866         | 0.066         | $3.39 \times 10^{-6}$ | $3.18 \times 10^{-14}$ |
| Normal                 | $3.39 \times 10^{-6}$ | 0.066         | 0.866         | 0.066         | $3.39 \times 10^{-6}$ |
| High                   | $3.18 \times 10^{-14}$ | $3.39 \times 10^{-6}$ | 0.066         | 0.866         | 0.066         |
| Extra high             | 0              | $3.18 \times 10^{-14}$ | $3.39 \times 10^{-6}$ | 0.066         | 0.933         |
Figure 3. The probability distribution of parameter performance “Low”

In this paper, fault mode is assumed to be independent. Joint conditional probability distribution of parameters in multiple fault combinations can be calculated by summing the distributions of single fault mode. For parameter $X_1$ “first-stage suction temperature” with set of corresponding fault modes $\{Y_1, Y_9\}$, the joint distribution are calculated as in Table 3. $X_{1,0}$ to $X_{1,4}$ represents the five parameter performances. $Y_{1,0}$ represents normal state and $Y_{1,1}$ represents the first fault mode. After the edge weight calculation is completed, network structure is checked to make sure that there are no loops in the network.

| Fault mode $Y_9$ | $X_{1,0}$ | $X_{1,1}$ | $X_{1,2}$ | $X_{1,3}$ | $X_{1,4}$ |
|------------------|----------|----------|----------|----------|----------|
| $X_{1,0}$        | $3.39 \times 10^{-6}$ | 0.933    | $4.31 \times 10^{-26}$ | 0.066    | $3.39 \times 10^{-6}$ |
| $X_{1,1}$        | 0.066    | 0.066    | $3.19 \times 10^{-17}$ | 0.021    | 0.001    |
| $X_{1,2}$        | 0.066    | 0.001    | 0.066    | 0.021    | 0.066    |
| $X_{1,3}$        | $3.39 \times 10^{-6}$ | 0.001    | 0.933    | 0.001    | 0.933    |

### 3. Study case

The automatic fault diagnosis machine for BOG compressor is built through the establishment of expert fault knowledge base and construction of Bayesian network. Algorithm verification case is a simulated failure of leakage of primary suction valve. The parameter changes when the fault occurs are shown in Figure 4. The first stage suction temperature rises from as low as -123.9°C to as high as -4.4°C. The first stage suction pressure rises from as low as 0.01MPa to as high as 0.07MPa while the second stage suction pressure is reduced from 0.875MPa to 0.671MPa. These changes are consistent with the fault knowledge that the leakage of the primary suction valve leads to the first-stage suction temperature rise, the first-stage suction pressure rise and the secondary exhaust pressure decrease.
Figure 4. Three parameters during the fault procedure. (a) First-stage suction temperature. (b) First-stage suction pressure. (c) Second-stage exhaust pressure.

The Bayesian network structure is constructed as described in 2.2.1 and the values of edges is calculated as described in 2.2.2. The simulated time series data is normalized and discretized to five integers from zero to four as the input to Bayesian network. The probability of 17 fault modes at every moment can be predicted. The probability of the leakage of first-stage suction valve rises from as low as 0.0029% to as high as 99.747% as shown in Figure 5. This indicates that automatic fault diagnosis machine can effectively identify faults.

Figure 5. Probability calculated from Bayesian network for leakage of first-stage suction valve.
When the predicted probability of leakage of primary suction valve is high, predicted probability of main fault modes is shown in Table 4. High probability of Leakage of primary suction valve provides a basis to determine the fault. Other main fault modes have a low fault probability while two fault modes share the probability 0.05: leakage of first-stage piston ring and leakage of second-stage piston ring. This illustrates that when leakage of first-stage suction valve occurs, there are certain probabilities that these two faults occur.

Table 4. Predicted probabilities of main fault modes when leakage of primary suction valve occurs.

| Fault Mode                        | Fault probability    |
|-----------------------------------|----------------------|
| Leakage of first-stage suction valve | 0.999993             |
| Blockage of first-stage suction valve | 5.67417e-13         |
| Leakage of first-stage piston ring | 0.0500031            |
| Leakage of second-stage piston ring | 0.0500031           |
| Bearing over temperature          | 9.93791e-13          |
| Crankshaft fracture               | 2.06403e-07          |
| Blockage of filter                | 2.13541e-18          |
| Leakage of Lubricating oil system | 4.10244e-18          |

4. Conclusion
An automatic update framework is proposed for fault diagnosis machine on BOG compressor. The framework combines expert experience with model analysis by building a fault knowledge base. Bayesian network is established based on expert fault knowledge without data. The framework is validated on a simulated case that the framework is able to discover fault and determine the mode of fault. The advantage of automatic update framework is the good explanation of faults.

Acknowledgments
This study is sponsored by the project “Research on health assessment and maintenance strategy for critical equipment in LNG terminals” of Qingdao LNG Co. Ltd., Sinopec.

References
[1] Zhao, B., Wang, T., Peng, X., & Feng J.(2015) Experimental study on performance of bog compressor. In: IOP Conference Series: Materials Science and Engineering, Macau, China, 90, 012022.
[2] Zhang, Z., Yang, Q., & An, D. (2018) An Improved K-Means Algorithm for Reciprocating Compressor Fault Diagnosis. In: 2018 Chinese Control And Decision Conference (CCDC), Shen Yang.
[3] Cui, H., Zhang, L., Kang, R., & Lan, X. (2009). Research on fault diagnosis for reciprocating compressor valve using information entropy and SVM method. Journal of Loss Prevention in the Process Industries, 22(6), 864-867.
[4] Tran, V. T., Althobiani, F., & Ball, A. (2014). An approach to fault diagnosis of reciprocating compressor valves using teager–kaiser energy operator and deep belief networks. Expert Systems with Applications, 41(9), 4113-4122.
[5] Zhao, W., Zhu, Y., & Wang, X. (2009). Combinatorial bayes network in fault diagnosis of power transformer. DianliZidonghuaShebei / Electric Power Automation Equipment, 29(11), 6-9.