Measurement 133 (2019) 182–192

Contents lists available at ScienceDirect

Measurement

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

Deep belief network-based internal valve leakage rate prediction approach

Shen-Bin Zhu [a], Zhen-Lin Li [a,*], Shi-Min Zhang [a], Ying-Yu [a], Hai-Feng Zhang [b]

[a] *College of Mechanical and Transportation Engineering, China University of Petroleum-Beijing, 102249 Beijing, China*
[b] *PetroChina Pipeline R&D Center, 065000 Langfang, China*

ARTICLE INFO

Article history:
Received 7 June 2018
Received in revised form 27 August 2018
Accepted 6 October 2018
Available online 9 October 2018

Keywords:
Internal valve leakage
Acoustic emission
Deep belief network
Deep learning
Leakage rate prediction

ABSTRACT

If a leak occurs for a valve in a natural gas station, it will first cause economic loss. Second, the gas leakage may also lead to the pollution of other pipeline systems as well as to environmental pollution. Under extreme conditions, it may even lead to an explosion, endangering the safety of the staff. Therefore, we urgently need a means to solve these problems. At present, acoustic emission (AE) detection technology is the most widely used method of diagnosing the valve leakage. The effects of gas leakage mainly depend on the valve leakage rate. However, the internal valve leakage rate is a multivariable, nonlinear, and time-varying process. Therefore, the accurate prediction of the valve leakage rate is an important challenge. Recognising this challenge, a novel prediction method, namely, regression-based deep belief network (DBN), which substitutes the linear regression (LR) layer for the linear softmax classification layer at the top of the general DBN's structure, has been proposed to predict the internal leakage rates of a valve in a natural gas pipeline. The internal leakage signals of a ball valve and a plug valve were collected using the AE system. The time–frequency features of the signals, inlet pressure of the pipe, and the valve type are used as the input variables to predict the leakage rates with the DBN model. The ball valve leakage data, plug valve leakage data, and mixed leakage data of both are used to establish and test the proposed models separately. At the same time, the back-propagation neural network (BPNN), support vector regression using linear (L-SVR), polynomial (P-SVR) and Radial basis function (RBF-SVR) kernels and the proposed DBN were all developed and compared to check the performance. After analysing the prediction results of these models, we found that the nonlinear and unstable features of the valve internal leakage signals could be well studied by using the DBN model. In addition, the performance of the DBN model was superior to that of the traditional prediction models for the three types of data. Therefore, it can be proven that the proposed model has huge practical application value for predicting the gas leakage rates of a valve in a natural gas pipeline system and has a guiding significance for predicting the other fluid leakage rates of a valve in other pipeline systems.

1. Introduction

The internal natural gas leakage of a valve refers to the leakage of natural gas through a closing member (valve) to the downstream. Valves in a natural gas pipeline generally cause internal leakage because of the complex operating environment. There are many reasons for internal valve leakage, include corrosion of sealing surface caused by non-drying and non-anticorrosive treatment when leaving factory, impurity entering valve seat caused by valve not infused with sealing grease, sphere damage caused by nonstandard installation or welding, scratching of sealing surface caused by welding slag and other construction residues, damage of sealing face caused by improper cleaning and so on [1]. When there is an internal leak in the valve, whether the natural gas leakage rate is within the permitted range is a key to replace or repair the damaged valve in time. Therefore, it is necessary to predict the internal valve leakage rate accurately.

The leakage rate of a valve is affected by many factors, mainly the inlet pressure, leak hole size and shape, valve size and valve type. Because the size and the shape of the leak hole cannot be measured, many scholars have studied the leakage rate on the basis of the inlet pressure, valve size and valve type [2–4], and these factors can reflect the degree of leakage in the valve intuitively. At present, many scholars are focused on solving the problem of predicting the valve leakage rate from the features of the

* Corresponding author.
E-mail address: zhenlinli@263.net (Z.-L. Li).

https://doi.org/10.1016/j.measurement.2018.10.020
0263-2241/

vibration signals produced by the gas leakage [4–9]. When the valve is leaking, the gas will be ejected from the leak hole into the downstream pipeline at a high speed, which will produce jet noise and cause pipeline vibration. The signals produced during gas leak process are mostly non-linear and non-stationary [10]. The key of the leakage rate prediction is to establish a corresponding model between the features of valve leakage signals and leakage rates, thereby achieving the purpose of predicting the valve leakage rates by detecting the internal leakage signals. In [4], physical theoretical models for the prediction of liquid leakage rates through ball valves and the globe valves have been established. The effects of leakage rates, inlet pressure levels, valve sizes and valve types on root mean square (RMS) have been studied. The results demonstrated that the AE signal power was significantly correlated with the affecting factors of the leakage rate. However, the physical theoretical model can effectively solve the problem of simple internal relationships, but it does not perform well when dealing with complex nonlinear problems. In [11], the degrees of valve leakage were predicted from the spectral amplitude of the signal, including non-leakage, moderate leakage, and severe leakage. This method is a simple classification of the degree of leakage. Therefore, it is not representative. In recent years, with the development of machine learning, the support vector machine (SVM) model based on statistical theory has been applied to the classification of the natural gas leakage flow level of a ball valve. Nine features, namely the mean standard deviation, root mean square value, energy and entropy of the time-domain signal and the root variance frequency, peak value and frequency centre of the frequency-domain signal, are used as the input parameters. The accuracy of prediction model is more than 95% [5]. In [12], a model based on factor analysis and k-medoids clustering was used to recognize recognition internal valve leakage rates, the recognition accuracy of the model is 96.28%. These classification models have a stronger ability to deal with complex nonlinear problems compared with the physical model, but the prediction result is only a judgment for a specific range of leakage flow levels. It lacks the ability to accurately quantitatively analyse the leakage rates and reduces the applicability of the model under different detection objects. In addition, these models are based on classification models of leakage flow levels in a single valve type, and the performance of the models can be greatly reduced while dealing with and analysing multi-type and multi-leakage-sized valve leakage signals. Therefore, we urgently need a stable and accurate method for predicting valve leakage rates to solve the abovementioned problems.

Along with the data explosion that resulted from the use of smart metering and various sensors, machine learning techniques are highlighted in the field of metering-based forecasting. SVR is a commonly used regression forecasting method [13]. It can be used to map the input data from a low-dimensional feature space to a high-dimensional feature space by using nonlinear mapping and then performing linear regression in the high-dimensional space [14,15]. In the field of energy, it is often used to deal with regression forecasting. In [16], the SVR forecasting model with the historic electric load and weather data of four large commercial office buildings was used to forecast the demand response baseline. The high degree of prediction accuracy and stability were shown in short-term load forecasting. In [17], the SVR model is also constructed for forecasting crude oil prices.

Another well-known prediction method is the neural network, where BPNN is the most widely used prediction method. In [18], the BPNN was applied and tested to forecast daily air pollutant concentrations. In [19], BPNN was applied to predict a ball bearing's remaining useful life. The above results all showed that the BPNN exhibited better forecasting performance.

The above-described learning methods are generally a shallow structure algorithm with one or no hidden layers [20]. Consequently, the limitation of these models lies in the limited representation ability of complex functions in the case of finite samples and computational units, and they cannot effectively explore the regularity of features [21]. Therefore, these shortcomings of the shallow models encourage us to re-examine the regression prediction problem based on deep learning. The concept of deep learning originates from the research of artificial neural networks. By combining lower-level features, a more abstract high-level model is formed to mine the deeper feature representations for the data itself [22–24]. A DBN is an important model in deep learning. It not only has the advantages of a traditional neural network but also has the strong ability of information fusion. The DBN model can make the objective function achieve global optimisation by pre-training and fine-tuning, which solved the shortcoming of the traditional neural network, which was the ease of reaching the local optimum [25]. DBN is mainly used for the modelling, feature extraction and recognition of images, documents, speech and other objects [24]. In recent years, DBN has attracted increasing attention in the field of regression prediction. In [23], a deep WSF framework and an intelligent approach based on DBN were investigated. The DBN model can enhance the WSF performance and the prediction efficiency. In [25], the regression-based DBN approach was applied to predict the sound quality of a vehicle interior's noise. Experimental verification and comparisons demonstrated that the DBN model exhibited better prediction accuracy and stability than the multiple linear regression (MLR), BPNN and SVM models.

The above literatures show that the DBN has some advantages over the traditional shallow prediction models. Nevertheless, thus far, DBN applied for predicting internal valve leakage rates has not yet been considered in the published literature. Therefore, based on the strong feature learning ability of DBN, we propose a novel regression based DBN method that replaces the top classification layer of DBN by a linear regression layer. It not only solves the problem of the poor accuracy of a physical model but also makes up for the limited ability of a shallow model to deal with complex data.

The purpose of this study was to evaluate the performance of DBN as a predictive tool for valve leakage rates. The performance of the model is directly related to the data structure, and for the different data, the performance of different models should be compared and analysed. Therefore, ball valve leakage data, plug valve leakage data and mixed leakage data of both are used to establish and test the proposed models. In addition, we compare the DBN model with the traditional prediction models such as SVR and BPNN. The experimental results show that the proposed method tested using the three types of data exhibited excellent performance in predicting valve leakage rates and is superior to the traditional methods in all three cases.

2. Methods

2.1. Restricted Boltzmann machine (RBM) architecture

RBMs are undirected probabilistic graphical models based on a bipartite graph which contain a single layer of observable variables and a single layer of latent variables. There are no intra-layer connections among the visible units or among the hidden units. For a given configuration (v,h), the energy function can be defined as follows:

$$E_\theta(v,h) = -\sum_{i=1}^{n} a_i v_i - \sum_{j=1}^{m} b_j h_j - \sum_{i=1}^{n}\sum_{j=1}^{m} v_i w_{i,j} h_j \tag{1}$$