CHEMICAL ENGINEERING TRANSACTIONS

VOL. 71, 2018

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Xiantang Zhang, Songrong Qian, Jianmin Xu

ISBN 978-88-95608-68-6; **ISSN** 2283-9216

DOI: 10.3303/CET1871044

Anomaly Detection of Petrochemical Process Engineering Based on Neural Network

Xiangrui Chen

College of Computer Science and Technology, Changchun Normal University, Changchun 130032, China
c_x_r1973@163.com

The petrochemical industry is large in production scale, it uses a large number of toxic and harmful devices in the process, and intensively relies on the safe, effective and long-term operation of these devices. Chemical accidents are often caused by abnormal conditions in the process, so the anomaly detection is an important tool to ensure the safe production of petrochemicals. The traditional anomaly detection methods based on mechanism model and knowledge model are limited by the complexity of the model and the artificial subjective experiences, and it is difficult to quickly apply them to new process systems. The anomaly detection methods based on mathematical statistics have problems such as slow calculation speed and inconsistent with actual parameters. Aiming at the problems of traditional anomaly detection methods, this paper proposes a petrochemical anomaly detection method based on the neural network, this algorithm uses the echo state network (ESN) to construct the feature data model, which can simplify the complexity of feature extraction and combine with the support vector machine (SVM) algorithm for the judgement and detection of abnormal data. At the same time, combining with the process of petrochemical industry, this paper proposes application examples of anomaly detection in petrochemical industry based on the neural network. This anomaly detection method can quickly perform feature data extraction and model construction, and can be applied to different processes, and has good versatility and practicability.

1. Introduction

The petrochemical industry has great danger, and it completely depends on the safe, efficient, long-term operation of devices. As the production technology is improving year by year, the production process gradually presents the trend of larger scale, more dangerous, and more refined, at the same time, the devices which are used in the petrochemical process are becoming more and more complex and refined as well. The production process is accompanied by a large number of uncertain factors, which brings enormous challenges to the reliable operation of the petrochemical process system. Anomaly detection technology (Mehra et al., 1971) is a key technology which can ensure the normal, safe and stable operation of modern production processes. The anomaly detection technology can locate the position, type, and time of the abnormality by judging, detecting and preventing the abnormality or failure of the system. It can detect the abnormal situation in time before the catastrophic failure happens to the system and avoid the occurrence of production accidents, which guarantees the personal safety and property safety to the greatest extent, and reduces economic losses caused by the accident. It greatly improves the safety of chemical process, and it is one of the key technologies essential to the chemical industry.
The anomaly detection methods are mainly divided into the following three types (Frank, 1990): anomaly detection method based on mechanism model, anomaly detection method based on knowledge rules, and anomaly detection method based on data-driven. The anomaly detection method based on mechanism model is the first anomaly detection method to be studied. At first, it needs to establish a physical model for the system, based on which, it constructs a complete mathematical model to describe the system, and then performs anomaly detection through the analysis model of abnormal and normal conditions which is constructed by the observable input and output variables. The anomaly detection method based on knowledge rules does not need to establish a complex mathematical model, instead, it introduces expert systems and knowledge libraries to form a series of knowledge rules, and simulates human judgments about anomalies or

Please cite this article as: Chen X., 2018, Anomaly detection of petrochemical process engineering based on neural network, Chemical Engineering Transactions, 71, 259-264 DOI:10.3303/CET1871044

faults through the expert system. For example, the literature (Bruton et al., 2015) proposes an anomaly detection system and model based on the expert system, and used it to perform anomaly detection on the energy problems caused by air processer in buildings. In recent years, with the in-depth research and development of big data technology, the anomaly detection method which is based on the data-driven has made breakthrough research progress. The anomaly detection method based on data-driven does not need to establish a complex mathematical model, nor does it need to introduce an expert system or a knowledge library, it directly analyzes and processes the system data, and it realizes anomaly detection by collecting system related data, using algorithms for feature extraction and abnormal judgment. Common data-driven anomaly detection methods include principal component analysis (PCA) (Wise et al., 2015), partial least squares (PLS) (Kresta et al., 2010), etc. The literature (He et al., 2010) proposed a non-parametric anomaly detection method based on KNN. This method relies on the similarity of the distance between the sample and its neighboring k samples to make abnormal judgments, which reduces the influence of nonlinearity and multi-modality in the process system. The literature (Lee et al., 2004) proposes a PCA method based on the kernel theory, it can effectively solve nonlinear problems. The literature (Gawehns et al., 2016) proposes an anomaly detection method based on Independent Component Analysis (ICA) that can identify hidden factors from multidimensional statistical data. Reference [Ma et al, 2012] proposed a local neighbor normalized strategy anomaly detection method, which uses the mean and standard deviation of the local neighbor sets of the samples to standardize the samples, it can translate the data centers of each stage to the origin, meanwhile by adjusting the degree of dispersion of each stage to make them approximately the same, so that the multi-stage process anomaly detection capability is improved.

The anomaly detection method based on the mechanism model relies too much on the establishment of the model. When the model of the system is too complicated, it is difficult to make adaptive adjustment according to the changes of the system. The anomaly detection method based on knowledge rules relies too much on the establishment of expert knowledge, which is greatly influenced by human factors, and it is difficult to quickly form expert knowledge due to the generation of new technologies. At present, the data-driven anomaly detection method does not depend on mathematical models or expert knowledge, but there are still problems such as the current algorithm relies on the extraction of feature parameters, the implementation algorithm is complex, the operation speed is low, and the algorithm is not universal, etc.

Therefore, based on the ESN in the deep learning neural network, this paper proposes an anomaly detection method combining ESN with SVM. ESN has the characteristics of fast operation speed and high model applicability. It can quickly establish a feature model for the new system, combining with SVM, it can perform detection and judgement on the abnormal data, and improve the operation speed and method versatility of the anomaly detection method.

2. Anomaly detection method based on ESN and SVM

2.1 ESN

ESN (Jaeger, 2001), the echo state network, is a new type of neural network model, which is the optimization and improvement of the traditional neural network model. ESN only needs to train the output layer, and its training process degenerates to the solving of linear regression, which overcomes the disadvantages of complex algorithm and slow operation speed of the traditional neural network.

The hidden layer of the traditional recurrent neural network (RNN) is composed of neurons with full connections, while the ESN introduces a reserve pool to replace the hidden layer of the traditional neural network. The dynamic reserve pool of ESN is a sparse network composed of a large number of neurons. The connection state between sparse neurons in the reserve pool is random, which can imply the operating state of the system and has the function of short-term training memory. At the same time, unlike the traditional neural network using the gradient descent algorithm to update the weight, the connection weight in the reserve pool of ESN is fixed. Therefore, ESN can greatly reduce the computation amount of the model training process, meanwhile, to some extent, it can avoid the situation of local optimal solution which appears when using gradient descent algorithm to update the weight.

The structure of ESN is shown as Figure 1. The left side is the input layer, the middle is the reserve pool, and the right side is the output layer. The reserve pool accepts input from two directions, one from the input layer, and the other from the output of the previous state of the reserve pool. The state feedback weights do not need training, and are determined by the random initial state, so W^W^ is a large sparse matrix, in which the non-zero elements indicate the activated neurons in the reserve pool.

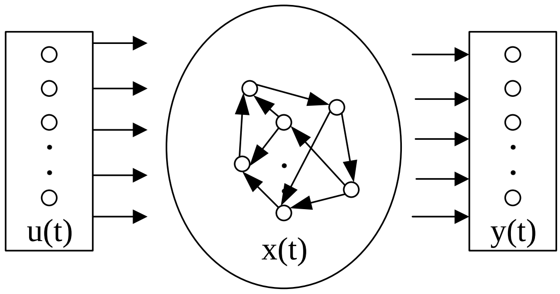

Figure 1: ESN structure

Suppose the input layer has K nodes, the reserve pool has N nodes, and the output layer has L nodes. The input value at time t is u(t), the state of the reserve pool is x(t), and the output value is y(t), which are:

$$\mathrm{u(t)} = [u_1(t), u_2(t), u_3(t), \cdots u_K(t)]^T$$

$$\mathrm{x(t)} = [x_1(t), x_2(t), x_3(t), \cdots x_N(t)]^T$$

$$\mathrm{y(t)} = [y_1(t), y_2(t), y(t), \cdots y_L(t)]^T$$

The connection weight from the input layer to the reserve pool is a matrix W_{in} of N*K orders, the connection from the reserve pool to the reserve pool of next moment is a matrix W of N*N orders, and the connection weight from the reserve pool to the output layer is a matrix W_{out} of L*(K+N+L) orders. The connection weight from the output layer of the previous moment to the reserve pool of the next moment is a matrix W_{back} of N*L orders. In ESN, both W and W_{in} are generated by random initialization when the network is initially established, and both are fixed.
The basic equations for an ESN are as follows:

$$\mathrm{x(t+1)} = f(W_{in} \times \mathrm{u(t+1)} + W_{back} \times \mathrm{x(t)})$$

$$\mathrm{y(t+1)} = f_{out} \times (W_{out} \times (\mathrm{u(t+1)}, \mathrm{x(t+1)}))$$

Where u(t+1) is the input at time t+1, x(t) is the state of the reserve pool at the previous moment t, and y(t+1) is the output at time t+1. $f$and f_{out} are the activation functions of the reserve pool layer and the output layer, such as sigmoid and tanh activation functions. The training process of ESN is to determine W_{out} according to the target output value y(target). The goal is that the difference between y(t+1) and y(target) is as small as possible, and the training process of ESN is the training of W_{out}, The training of ESN is divided into two phases: the sampling phase and the weight calculation phase. In the sampling phase, the training samples are added to the reserve pool through the input connection weight matrix Win, and the system state and output y(t) are calculated and collected according to the above two state equations. The weight calculation is to calculate the connection weight Wout according to system state matrix and sample data collected during the sampling phase. Because the state variable x(t) and the predicted output have a linear relationship, the goal that needs to be achieved is to use the predicted output to approximate the expected output y(t), which will degenerate into a simple linear regression problem.

2.2 One-class SVM (1-SVM)

SVM is a machine learning algorithm based on the principle of structural risk minimization (Andrew, 2001). The basic model of basic SVM is to find the best separating hyperplane in the feature space so that the positive and negative sample interval is the largest on the training set. SVM is an effective supervised learning algorithm for solving the binary classification problems. However, in anomaly detection, it only needs to identify the abnormal positions, so the SVM-based anomaly detector is often a 1-SVM.
1-SVM separates all data points from the zero-point in the feature space and maximizes the distance between the separating hyperplane and the zero-point, which generates a binary function that captures the probability density region of the data in the feature space. When in the training data point area, it returns +1, while in other areas it returns -1. The optimization objectives of 1-SVM are as follows:

$$\underset{w,\zeta_i,\rho}{min} \frac{1}{2}\|w\|^2 + \frac{1}{\nu n}\sum_i^n(\zeta_i - \rho)$$

$$\mathrm{s.t.} \left(w^T\phi(x_i)\right) > \rho - \zeta_i, \zeta_i \geq 0$$

The hyperplane equation is $\mathrm{f(x)} = \mathrm{w}x_i - \rho$, ζ_i represents the slack variable, v sets an upper limit for the fraction of the abnormal value, and it's also the lower limit of the number of examples of the support vector in the training data set, $\phi(x_i)$ is a nonlinear function.
The Lagrangian function is established as:

$$\mathrm{L}(w,\zeta,\rho) = \frac{1}{2}\|w\|^2 + \frac{1}{vn}\sum_i^n(\zeta_i - \rho) - \sum_i^n a_i[\langle w, \phi(x_i)\rangle - \rho + \zeta_i] - \sum_i^n \beta_i \zeta_i$$

By finding partial derivatives for each variable we can get:

$$\mathrm{w} = \sum_{i=1}^n a_i \phi(x_i)$$

Substituting the parameters into the original Lagrangian function, meanwhile by using the Gaussian kernel function $K(x_i, x_j)$, the dual problem of the original problem can be obtained as:

$$min\ \frac{1}{2}\sum_{i,j=1}^n a_i a_j K(x_i, x_j)$$

$$\mathrm{s.t.}\ 0 \le\ a_i \le 1, \sum_{i=1}^n a_i = vn$$

After solving this quadratic programming problem, the optimal solution of the quadratic programming problem can be solved:

$$\mathrm{f(x)} = \sum_{i=1}^n a_i K(x_i, x) - \rho$$

After that, by using the kernel function and KKT condition in the traditional SVM, the optimal solution of the quadratic programming problem can be solved.

2.3 Anomaly detection method model

Based on this, this paper proposes an anomaly detection method based on ESN and SVM. First, this method trains the system model by ESN and extracts the feature data that can be used for anomaly detection, and then enters the anomaly detector of SVM to judge the feature data and realize anomaly detection. This method can quickly train the feature data suitable for a certain system through ESN, and then perform anomaly detection by SVM algorithm. In general, the method runs fast, and it can perform training according to different systems. At the same time, through feature extraction, it can effectively reduce the dimension of the feature data, and the model of this method is shown in Figure 2.

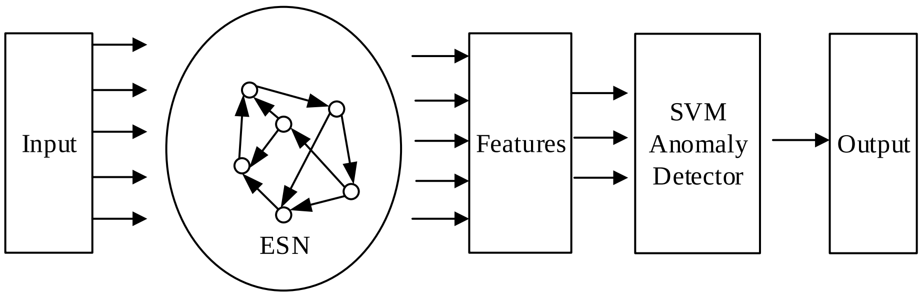

Figure 2: Anomaly detection model based on ESN and SVM

The anomaly detection method based on ESN and SVM is divided into the following steps:
(1) ESN parameter initialization: according to ESN model, initialize the parameters, including the number of nodes in the input layer and output layer, the number of nodes in the reserve pool, and so on.
(2) ESN training: the training process is the process of calculating the weight of the reserve pool to the output layer. Sampling first, namely to calculate the reserve pool matrix and the output layer matrix, and then linearly fit the sampled data.
(3) SVM training: it includes the selection of kernel function and loss function, and using sample data to train the SVM model.
(4) Anomaly detection: input the parameter values which are output by ESN to the anomaly detector of SVM for abnormality judgment.

3. Examples of oil process anomaly detection

The petrochemical process involves a large number of equipment facilities and processes. Based on the proposed anomaly detection method which combines with ESN and SVM, this paper gives a few anomaly detection methods during in the petrochemical process (Kong and Zhao, 2018).

3.1 Drilling process anomaly detection

There are many uncertain factors in the oil drilling process, meanwhile, during the drilling process, many variables will vary in magnitude as the state of the system changes. The variations in these drilling processes are used as the input parameters of the above-mentioned anomaly detection method which combines with ESN and SVM, and the anomaly detection and judgment can be performed based on the values of these variations.

For example, the following feature parameters can be selected as input feature variables for drilling process detection: well depth, weight-on-bit (WOB), drilling rate, drilling time, stand pipe pressure (SPP), total pool volume, outlet flow, outlet temperature, inlet flow, inlet temperature, hook load, pump stroke, casing pressure, rotary speed, etc. Since the correlations among these variables are quite different, therefore, at first, it’s necessary to extract these features using ESN and form simplified feature parameters. After that, using sample data to train SVM, it is necessary to set the threshold value for the feature parameters in different stages of the drilling process, such as the process of run in hole and pull out of hole, and this threshold value is used as the training sample data of SVM. After both ESN and SVM have completed the corresponding training, the anomaly detection can be performed according to different features in the drilling process in real time, so as to avoid the occurrence of accidents.

3.2 Rotating machinery anomaly detection

Rotating machinery is one of the most commonly used mechanical equipment in petrochemical industry. Many petrochemical accidents are caused by the failure of rotating machinery. The most important problem is the aging caused by long-term usage of the rotating machinery. Common rotating machinery failures include rotor imbalance, rotor misalignment, generating set resonant vibration, mechanical looseness, shaft bending and cracking, oil whirl, oil whip, etc. Therefore, we need to apply the above-mentioned anomaly detection method to the anomaly detection of rotating machinery.

The first is still the selection of feature variables, as for the rotating machinery failure, the feature variables can be selected include: radial vibration, rotating speed, temperature, misalignment kurtosis, shaft crack kurtosis, eccentric shaft kurtosis and other indicators. Record sample data for each indicator under normal and abnormal conditions. These features are then extracted by ESN to form a simplified feature parameter model. And then, use the sample data to train the SVM. After ESN and SVM both finished the corresponding training, the anomaly detection can be performed according to different feature indicators of the rotating machinery, so that the aging condition of the rotating machinery can be judged in advance to avoid the occurrence of chemical accidents.

4. Conclusion

With the rapid development of the petrochemical industry, more precision equipment has been used in the petrochemical process. The overall petroleum process is highly dependent on the safe and long-term effective operation of the equipment. Abnormalities in the equipment will cause chemical accidents and even bring harm to personal and property safety. Anomaly detection is one of the most important safety production tools in the petrochemical industry. Traditional anomaly detection schemes based on knowledge structure and mechanism model rely too much on the establishment of mathematical models and expert knowledge libraries, which are limited by system complexity and human factors. The data-driven anomaly detection methods based on KNN have the problems of slow operation speed and multi-dimensions. With the rapid development of big data technology, aiming at the problems of traditional anomaly detection methods, this paper proposed a petrochemical anomaly detection method based on neural network and SVM, this method first used ESN to construct a feature data model, and then combined with 1-SVM to perform judgment and detection of abnormal data. At the same time, this paper combined with the process of petrochemical industry, and put forward application examples of this method in drilling process and rotating machinery detection. The anomaly detection method based on ESN and SVM can quickly construct anomaly detection model, and can adapt to new systems, no longer limited by system complexity, human factors and operation speed, and has good applicability and development and application prospects

References

Andrew A. M., 2001, An introduction to support vector machines and other kernel-based learning methods. Kybernetes, 32(1), 1-28. DOI: 10.1108/k.2001.30.1.103.6

Bruton K., Coakley D., Raftery P., Cusack D. O., Keane M. M., and O'Sullivan D. T. J., 2015, Comparative analysis of the ahu info fault detection and diagnostic expert tool for ahus with apar. Energy Efficiency, 8(2), 299-322. DOI: 10.1007/s12053-014-9289-z

Frank P. M., 1990, Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: a survey and some new results. Automatica, 26(3), 459-474. DOI: 10.1016/0005-1098(90)90018-d

Gawehns D., Wilderjans T. F., and Putten C. M. V., 2016, The statistical analysis of neuronal data: comparing algorithms for independent component analysis. Isme Journal, 2(7), 317-336. DOI: 10.1038/ismej.2008.44

He Q. P., and Wang, J., 2010, Large-scale semiconductor process fault detection using a fast pattern recognition-based method. IEEE Transactions on Semiconductor Manufacturing, 23(2), 194-200. DOI: 10.1109/TSM.2010.2041289

Jaeger H., 2001, The "echo state" approach to analysing and training recurrent neural networks. überwachtes lernen.

Kong X., Zhao D., 2018, Study on intelligent emergency management information system for petroleum andpetrochemical enterprises, Chemical Engineering Transactions, 66, 1015-1020. DOI:10.3303/CET1866170

Kresta J. V., Macgregor J. F., and Marlin T. E., 2010, Multivariate statistical monitoring of process operating performance. Canadian Journal of Chemical Engineering, 69(1), 35-47. DOI: 10.1002/cjce.5450690105

Lee J. M., Yoo C. K., Sang W. C., Vanrolleghem P. A., and Lee I. B., 2004, Nonlinear process monitoring using kernel principal component analysis. Chemical Engineering Science, 59(1), 223-234. DOI: 10.1016/j.ces.2003.09.012

Ma H., Hu Y., and Shi H., 2012, A novel local neighbourhood standardization strategy and its application in fault detection of multimode processes. Chemometrics & Intelligent Laboratory Systems, 118(7), 287-300. DOI: 10.1016/j.chemolab.2012.05.010

Mehra R. K., and Peschon J., 1971, Correspondence item: an innovations approach to fault detection and diagnosis in dynamic systems. Automatica, 7(5), 637-640. DOI: 10.1016/0005-1098(71)90028-8

Wise B. M., Gallagher N. B., Butler S. W., White D. D., and Barna G. G., 2015, A comparison of principal component analysis, multiway principal component analysis, trilinear decomposition and parallel factor analysis for fault detection in a semiconductor etch process. Journal of Chemometrics, 13(3-4), 379-396. DOI: 10.1002/(SICI)1099-128X (199905/08)13:3/4<379: AID-CEM556>3.0.CO