VOLUME - 13, ISSUE - 01, JANUARY - 2024 • PRINT ISSN No. 2277 - 8160 • DOI : 10.36106/gjra

Original Research Paper

Engineering

STIRRED-TANK HEATING SYSTEM FAULT DIAGNOSTIC MECHANISM BASED MAHALANOBIS DISTANCE

Hanh Chieu Vu	Faculty of Electrical & Electronics Engineering, Ly Tu Trong College, Ho Chi Minh City, Viet Nam.
Cam Hue Tang	Faculty of Electrical & Electronics Engineering, Ly Tu Trong College, Ho Chi Minh City, Viet Nam.
Hieu Trinh Tran	Faculty of Electrical & Electronics Engineering, Ly Tu Trong College, Ho Chi Minh City, Viet Nam.

ABSTRACT Predictive maintenance of the plants can be performed using multivariate sensor data gathered from the manufacturing and process sectors. This data represents actual operation behaviors. The intricate behaviors of industrial systems, sensor interactions, control system corrections, and variability in aberrant behavior make anomaly identification and diagnosis—a crucial component of predictive maintenance— to be more and more challenging. Specific chemical processes necessitate extra stringent requirements in addition to high-precision actuator functioning. Even slight changes in the outcome product's quality can result from chemical interactions. Thus, in addition to the requirement for a high-performance integrated control system, monitoring operations must be quick and accurate enough to identify and isolate defects when system issues arise. This research investigates a data-driven estimation based process fault diagnostic and detection approach. According to this approach, the discrepancy between the process response and the process model response is used to identify the process failure. For fault diagnostic purposes, errors are classified using the Mahalanobis distance. The technique is validated in this study using the stirred-tank heating process. The outcomes of the simulation show how effective the suggested strategy is.

KEYWORDS : Faults diagnostic, mahalanobis distance, industrial process, model prediction.

INTRODUCTION

There are several options available now for data processing algorithms that seek to examine data as effectively as possible. Every technique has benefits and drawbacks of its own. They can be broadly separated into two groups: qualitative methods and quantitative methods. The qualitative method is very challenging to use in actual processes and requires developers to have a thorough understanding of the system. Consequently, the focus of this study will be on quantitative methods using statistical algorithms (PCA, KPCA, MD, etc.)."Point anomalies" and "contextual anomalies" are the two categories into which abnormalities in industrial data fall.

According to Chandola et al. (2009), "contextual anomalies" are those that are aberrant in a particular context but not elsewhere, such as delayed buildup of material in equipment, while "point anomalies" are those that can be regarded anomalous with respect to the rest of the data (e.g., bias in sensor readings). Industrial system anomaly identification and diagnosis is a difficult undertaking because of the intricate behavior of machinery and processes, sensor interactions corrective actions of control systems and variability in anomalous behavior.

The literature describes several data-driven methods for finding anomalies in industrial data. Several anomaly detection methods were examined by Chandola et al. (2009) based on fundamental strategy used in each technique and evaluated the efficiency of these methods. Goldstein and Uchida (2016) conducted a similar study to assess the performance of unsupervised anomaly detection systems for multivariate data. Qin (2009) used Hoteling's T2 and Q statistics along with principal component analysis (PCA) and its variations to assess flaws found in complicated industrial processes. PCA was utilized by Mujica, Rodellar, Fernandez, and Guem (2010) to assess damage to aircraft structures. For fault identification and process monitoring, Chen, Zang, Yuri, Shardt, Ding, Yang, Yang, and Peng (2017) compared the T2 and Q statistics. d on clustering-derived thresholds. Mahalanobis Taguchi System (MTS), based on Mahalanobis Distance (MD), was proposed by Soylemezoglu and Jagannathan (2011) for a fault detection, isolation, and prognostics scheme is described for centrifugal pump failures. They used thresholds determined from clustering to perform anomaly detection and created fault clusters based on MD values.

This study will propose a general process to develop fault diagnostic model Mahalanobis Distance based on data-driven estimation.

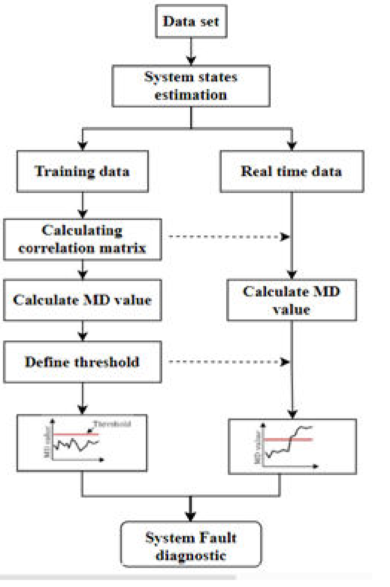

Figure 1. Mahalanobis distance monitoring process

2. METHOD

System identification is typically the first stage in looking into the features of a particular system that is thought of as a "black box" model. The system output is collected simultaneously with all projected data in order to perform a statistical algorithm-based correlation analysis between two data sets. Therefore, during actual operation, the fault diagnostic mechanism's quality will be determined by the correctness of the preceding stage. *Figure 1* describes system monitoring method based on MD with threshold value [7,8,9].

3. Research Plant

To implement the suggested approach, a Heating Tank Process will be examined in this study. We'll start by summarizing the input-output relationship for the model. Second, a Matlab Simulink model is created to replicate the input-output reaction of a process. Ultimately, the numerical model is incorporated into the suggested methodology to locate, identify, and categorize process errors.

$$d_{MD} = \sqrt{(x_b - x_a)^T * C^{-1} * (x_b - x_a)}$$

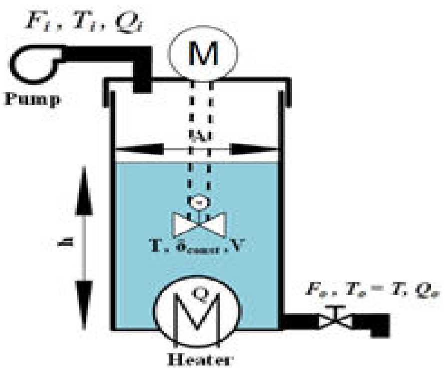

Figure 2. Heating tank system

Where: δ: density of liquid
F_i, F_o : input and output flow
T_i, T_o : input and output temperature
Q_i, Q_o: represent for the enthalpy of the inlet and outlet stream(s).
Q : heater energy
h : water level height
A : tank area

$$cov_{x,y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{N-1}$$

Table 1. Single tank's parameters

Parameters	Value
Tank area (A)	1400 cm2
Exhausted coefficient (C)	0.65
Exhausted valve area (b)	6.5 cm2
Power coefficient of pump (K)	1450 cm3/Vs
Gravity coefficient (g)	9.81 m/s2

In this study, behaviors of heating tank system based on below assumptions:

1. Material of liquid is homogeneous
2. System's parameters are constant with time.
3. Perfect mixing, the exit temperature T is also the temperature of the tank content.
4. Heat losses are negligible .

Because changes in internal energy are greater than changes in potential or kinetic energy, changes in these energy types can be disregarded.

4. Mahalanobis Distance

The distance in multivariate space between two points could be considered as the Mahalanobis distance.This is comparable to the Euclidian distance, with the exception that the MD technique also takes into account a correlation between variables [9]. Among the MD's most frequent goals is in searching the multivariate space for outliers. This is frequently utilized in applications like some measuring tools' fault detection. Equation (1) defines the Mahalanobis distance between two objects as follows: (download file)

(1)where C is the sample covariance matrix and x_a, x_b, and refer to a pair of objects.

The Mahalanobis distance has the benefit of being able to solve for the the Euclidian distance's limit. The Mahalanobis distance measures the distance between the point and the distribution itself, which resolves the issue of utilizing the Euclidian distance in a multivariate space, which is incorrect because it only deals with distances between points [10].

Equation (2) illustrates how this is accomplished by utilizing the covariance of the matrix distribution in the computation:

(2) where *x* and *y* denote the distributions' means, and x_i and y_i stand for the data values of the distributions. A contour map covering a scatterplot of 100 randomly selected draws from a bivariate normal distribution with a zero mean, unit variance, and 50% correlation is an example of a Mahalanobis distance plot, as shown in *Figure 3*. A blue square denotes the centroid determined by the marginal means. It should be mentioned that outliers in the bivariate space can be found using the Mahalanobis distance.

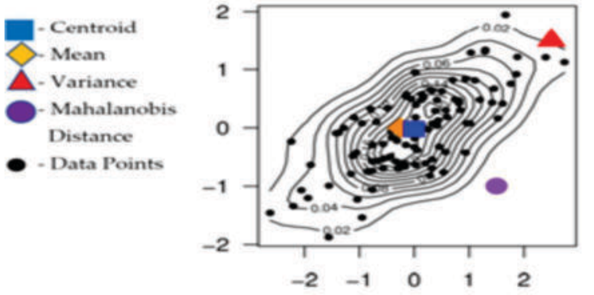

Figure 3. *Illustration of the Mahalanobis distance in a bivariate space adapted from [11]; the contours represent the behavior of the MD in the bivariate space. The Mahalanobis distance can detect outliers in the bivariate space.*

5. Experemental Results

System will be investigated under normal an abnomal condition to verify proposed fault detection mechanism. *Figure 4* illustrates MD of system respond during operation.

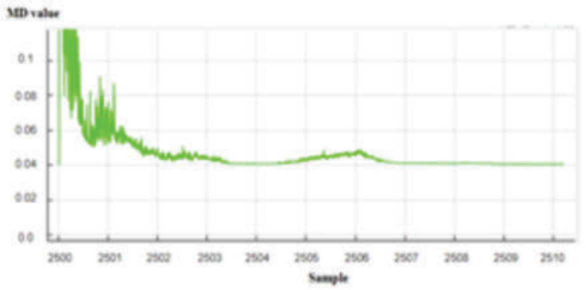

Figure 4. System monitoring with MD

Based on normal condition, threshold value will be calculated to detect any outliers (or abnomal samples in this situation) based real-time and estimation responds of system.

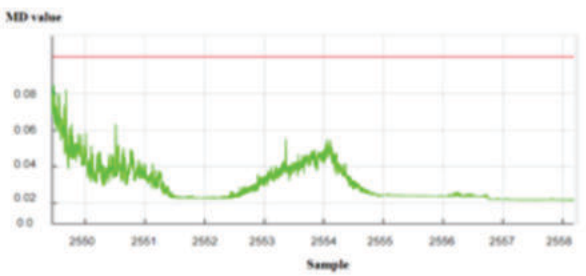

Figure 5. MD monitoring with thresold value

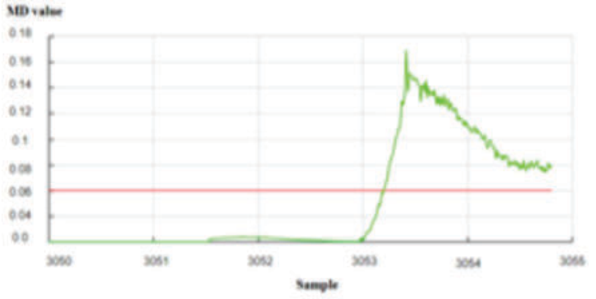

Figure 6. MD Fault detection mechanism

In normal operation, MD of all samples (system respond) project under threshold value (the red line) illustrated in *figure* 5. If errors or dangerous disturbances appear, MD of these samples will exceed the threshold (*figure 6*) and alert something wrong in this system.

6. CONCLUSIONS

This study proposes a method to develop faults diagnosing for industrial process based on statistical mahalanobis distance (MD). With a little information about new system, MD can detect almost all abnormal problems.

The simulation results are the evidence for the applicability and the effectiveness of the proposed approach in reality.

REFERENCES:

[1] Alcala, C. F., & Qin, S. J. (2010). "Reconstruction-based contribution for process monitoring with kernel principal component analysis". *Proceedings of the 2010 American Control Conference*, 29 July 2010, Baltimore, MD, USA, DOI: 10.1109/ACC.2010.5531315.

[2] Alcala, C. F., & Qin, S. J. (2011). "Analysis and generalization of fault diagnosis methods for process monitoring". *Journal of Process Control - J PROCESS CONTROL*. 21. 322-330. 10.1016/j.jprocont.2010.10.005.

[3] Amer, M., Goldstein, M. & Abdennadher, S. (2013). "Enhancing one-class support vector machines for unsupervised anomaly detection". *Proceedings of the ACM SIGKDD Workshop on Outlier Detection and Description*, August 11-11, New York, NY, USA, DOI: 10.1145/2500853.2500857.

[4] Chandola, V., Banerjee, A. & Kumar, V. (2009). "Anomaly Detection: A Survey". *ACM Computing Surveys*, 1-72 .

[5] Chen, Z., Zhang, K., Yuri A.W., Shardt, Ding, S. X., Yang, X., Yang C. & Peng, T. (2017). "Comparison of Two Basic Statistics for Fault Detection and Process Monitoring". *International Federation of Automatic Control*, Volume 50, Issue 1, July 2017, Pages 14776- 14781 .

[6] Choi, S. W., Lee, C., Lee, J. M., Park, J. H., & Lee, I. B. (2005). "Fault detection and identification of nonlinear processes based on kernel PCA". *Chemometrics and intelligent laboratory systems*, 75(1), 55-67, DOI: 10.1016/j.chemolab.2004.05.001.

[7] Xiaohang Jin. "Health Monitoring of Cooling Fans Based on Mahalanobis Distance With mRMR Feature Selection" , *IEEE Transactions On Instrumentation And Measurement, VOL. 61, NO. 8, AUGUST 2012.*

[8] Yu Wang. "Online Anomaly Detection for Hard Disk Drives Based on Mahalanobis Distance", *IEEE TRANSACTIONS ON RELIABILITY, VOL. 62, NO. 1, MARCH 2013.*

[9] Yue Cui, Lina Bertling Tjernberg. "An Anomaly Detection Approach Based on Machine Learning and SCADA Data for Condition Monitoring of Wind Turbines", *2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS).*

[10] Brereton, R.G.; Lloyd, G.R. Re-evaluating the role of the Mahalanobis distance measure. *J. Chemom.* **2016**, *30*, 134–143.

[11] Hill, T.; Lewicki, P.; Lewicki, P. *Statistics: Methods and Applications: A Comprehensive Reference for Science, Industry, and Data Mining*; StatSoft Inc.: Street Tulsa, OK, USA, 2006.

[12] Brereton, R.G.; Lloyd, G.R. Re-evaluating the role of the Mahalanobis distance measure. *J. Chemom.* **2016**, *30*, 134–143.