IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 11, 2017, accepted September 5, 2017, date of publication September 22, 2017,
date of current version October 12, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2754447

Fragment Anomaly Detection With Prediction and Statistical Analysis for Satellite Telemetry

DATONG LIU[1], (Senior Member, IEEE), JINGYUE PANG[1], GE SONG[2], WEI XIE[3], YU PENG[1], (Member, IEEE), AND XIYUAN PENG[1]

[1]Department of Automatic Test and Control, Harbin Institute of Technology, Harbin 150080, China
[2]SAIC Motor Corporation, Ltd., Shanghai 200041, China
[3]School of Business Administration, South China University of Technology, Guangzhou 510640, China

Corresponding author: Datong Liu (liudatong@hit.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61571160, Grant 61771157, and Grant 61301205.

ABSTRACT In aerospace engineering, condition monitoring is an important reference for evaluating the performance of complex systems. Especially, effective anomaly detection, based on telemetry data, plays an important role for the system health management of a spacecraft. With the advantages of easy-to-use, high efficiency, and data-driven, the predicted model has been applied for anomalous point detection for monitoring data. However, compared with the point abnormal mode, fragment anomaly is more attractive and meaningful for the system identification. Therefore, the detection strategy of fragment anomaly is proposed based on the uncertainty estimation of least square support vector machine and statistical analysis. Moreover, some effective estimation indicators are presented to evaluate the performance of the detection method. Experimental validations are also carried out for some typical simulation data sets and open source data sets. In particular, relied on the analysis of fragment anomaly modes, experiments are conducted with the real satellite telemetry data and different anomaly modes are injected to examine the applicability of the proposed framework.

INDEX TERMS Satellite, anomaly detection, fragment anomaly, LS-SVM, uncertainty.

I. INTRODUCTION

With the improvement of system complex on structures and functions, it is meaningful to estimate the status and performance of an equipment or a system based on condition monitoring data [1], [2]. In addition, to improve system safety and reliability, effective and in-time anomaly detection may help avoid catastrophes and serious faults [3], [4]. Consequently, anomaly detection based on condition monitoring data has attracted considerable attentions from the researchers in the field of reliability, testing, signal processing, data mining, prognostics and health management, etc. [5], [6]. Particularly, in aerospace engineering, the telemetry data is the only basis for the ground staffs to judge the condition of spacecraft in-orbit [7]. The anomalies in the telemetry may indicate data error, communication link failure, sensor fault, equipment failure, and even the degradation of system performance. Given the high reliability requirement of spacecraft, telemetry-data-based anomaly detection for the equipment or system becomes necessary and important to save the high failure cost [8].

Three methods are usually implemented for anomaly detection with monitoring data, i.e., threshold-based method, expert experience method and data-driven method. Specifically, threshold-based anomaly detection is the easiest and fastest method used in different areas. However, the threshold should be set by experience and design requirement and is relatively large which may miss some anomalies [9]. The experience-based method is more effective than the threshold-based one, because it has more injected prior knowledge. While, the completeness of experience is limited by knowledge which may restrict the detection ability [10]. At present, the data-driven method is becoming increasingly popular in anomaly detection, with the advantages of easy-to-extend, independent of area knowledge, and strong learning ability, intensive research has been conducted based on this type of method.

Generally, according to the dissimilarity of measure indicators, data-driven anomaly detection can be classified as similarity-based method, deviation-based method, and probability-based method. For the similarity-based methods,

we mainly referred to the methods of K-nearest neighbor (KNN), brute force [11], and clustering [12], which label the sample anomalous that has low similarity with the normal samples. Thus, the accuracy of this method highly depends on the selection of the similarity measure function. The deviation-based methods include statistics [13], prediction [14] and several classification [15] functions, which construct a model to depict the normal data. Then, one can compare the deviation between the model output and the real sample to detect the anomalies when the test data is available. The advantage of deviation-based anomaly detection is that it can realize high testing efficiency. The last method is based on the probability and it is also referring to frequent items and association rules method [16]. This method judges the probability of the test data from the normal data. Discretized representation needs to be applied for this method where some detailed information may be missed. Given that the telemetry data have the characteristics of pseudo-period, strong regularity, and high requirement on detecting efficiency, the deviation-based method is more feasible to realize online detection for telemetry data. In particular, the predicted-based method is perhaps the most suitable choice in aerospace application with online monitoring, i.e., detecting anomalies with strong learning ability and does not rely on expert experience.

Predicted-based method has been applied for anomaly detection by comparing the normal range and the real value. Namely, the model trained by the normal samples can give the confidence interval for the testing sample. If the new test sample exceeds the confidence interval, it will be highlighted. Thus far, Naïve predictor, Single-layer linear network, Multilayer perception [17], Gaussian Process Regression (GPR) [2], AR [18], ARMA [19], ARIMA [20], Kalman Fitter (KF) [21], Support Vector Machine (SVM) [22] and Least Square Support Vector Machine (LS-SVM) [23] have been applied for anomaly detection. It is worth mentioning that, due to strong nonlinear features and high performance, LS-SVM is an important and effective prediction algorithm that has been adopted for telemetry point prediction [24]. In this work, we focus on the implementation of LS-SVM for the anomaly detection of satellite telemetry data.

However, compared with the anomalous points, fragment anomaly, the anomalous subseries formed by several points, is more useful in aerospace application, because it can indicate meaningful anomaly modes, such as communication link failure, sensor failure, component failure, etc. Hence, this paper aims at proposing an anomaly detection framework that concentrates on the fragment anomaly with LS-SVM algorithm. Applying the LS-SVM algorithm, the predictive intervals can be calculated with the marginalized model parameters. As a result, the anomaly detection capability can be obtained with the predicted value as well as the interval values. Moreover, the evaluation can be implemented by considering multiple continuous point detection results at the same time. Therefore, the fragment anomaly detection can be realized and the confidence probability is also involved.

This paper is organized as follows. Section II will briefly introduce and analyze the preliminary work, including basic LS-SVM algorithm, time series prediction, and LS-SVM-based anomaly detection. In Section III, the proposed framework of anomaly detection, especially the anomaly detection strategy for fragment anomaly, will be discussed in detail. Experimental results based on open source data sets will be presented to validate the proposed framework in Section IV. In Section V, a case study of anomaly detection for aerospace application is described, in which fragment anomalies are injected into the real normal telemetry series. Finally, Section VI concludes the work and provides some future research directions.

II. PRELIMINARY WORK

A. LS-SVM ALGORITHM

LS-SVM [25] algorithm is proposed as an improved SVM to solve the optimization problem with large data set.

Given the training set $D = \{(x_i, y_i)\}_{i=1}^{N}$, where the training input $x_i \in R^n$, the training target $y_i \in R$, and N is the sample size. Then, the associated regression function can be written as

$$f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b \tag{1}$$

where $\mathbf{w}^T \in \mathbb{R}^n, b \in \mathbb{R}$, and $\phi(\cdot)$ is the setting kernel function that used to solve nonlinear problem. For the LS-SVM model, the optimization problem associated with the regression function is shown in (2).

$$\begin{aligned} &\min_{\mathbf{w},b,e} J(\mathbf{w}, b, e) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{\gamma}{2}\sum_{i=1}^{N} e_i^2 \\ &s.t.\ y_i = \mathbf{w}^T\phi(\mathbf{x}_i) + b + e_i, i = 1, 2, \ldots, N \end{aligned} \tag{2}$$

where e_i is the error between the estimated value and the real one and γ is the regularization parameter. Compared with the SVM model, LS-SVM introduces the sum of error squares which converts the quadratic programming problems into linear equations. Consequently, the Lagrangian relaxation of the optimization problem can be expressed as (3).

$$L(\mathbf{w}, b, e, \alpha) = Q(\mathbf{w}, b, e) - \sum_{i=1}^{N} \alpha_i[\mathbf{w}^T\phi(\mathbf{x}_i) + b + e_i - y_i] \tag{3}$$

where α_i are Lagrange multipliers. By solving partial derivatives with respect to w, b, e, α, the optimality conditions of the above Lagrangian function are

$$\begin{cases} \dfrac{\partial L}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w} = \sum\limits_{i=1}^{N} \alpha_i \phi(\mathbf{x}_i), \\ \dfrac{\partial L}{\partial b} = 0 \Rightarrow \sum\limits_{i=1}^{N} \alpha_i = 0, \\ \dfrac{\partial L}{\partial e_i} = 0 \Rightarrow \alpha_i = \gamma e_i \quad i = 1, 2, \cdots, N, \\ \dfrac{\partial L}{\partial \alpha_i} = 0 \Rightarrow \mathbf{w}^T\phi(\mathbf{x}_i) + b + e_i - y_i = 0 \\ \qquad\qquad i = 1, 2, \cdots, N. \end{cases} \tag{4}$$

The above equalities can be converted into a matrix form as:

$$\begin{bmatrix} 0 & \vec{\mathbf{1}}^T \\ \vec{\mathbf{1}} & \Omega + \gamma^{-1}\mathbf{I} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \tag{5}$$

where $\vec{\mathbf{1}} = [1, 1, \ldots, 1]$, $\Omega_{kj} = K(x_k, x_j)$, $\alpha = [\alpha_1, \ldots, \alpha_N]$, and $\mathbf{y} = [y_1, \ldots, y_N]$.

Let $\mathbf{A} = \Omega + \gamma^{-1}\mathbf{I}$, then,

$$b = \frac{\vec{\mathbf{1}}^T \mathbf{A}^{-1}\mathbf{y}}{\vec{\mathbf{1}}^T \mathbf{A}^{-1}\vec{\mathbf{1}}}, \quad \alpha = \mathbf{A}^{-1}(\mathbf{y} - b\vec{\mathbf{1}}) \tag{6}$$

Finally, we can obtain the prediction function of LS-SVM as:

$$f(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i K(\mathbf{x}, \mathbf{x}_i) + b \tag{7}$$

B. TIME SERIES PREDICTION WITH LS-SVM

From the introduction in subsection II.A, LS-SVM can realize function regression and prediction. The main procedures of time series prediction based on LS-SVM are shown in Fig. 1 and detailed description is listed as follows.

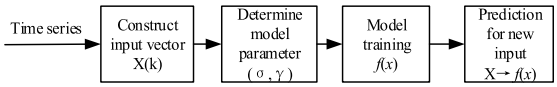


FIGURE 1. Time series prediction based on LS-SVM.

(1) Constructing the input vectors
For time series $x(t), t = 1, 2, \ldots, N$, the phase-space reconstruction series is (8).

$$X(t) = F\{x(t), x(t-\tau), \cdots, x[t-(m-1)\tau]\} \tag{8}$$

where τ and m are the delay interval and embedded dimension, respectively. Thus, the reconstructed phase space can be achieved with the observed value as well as its m-dimension delayed value.

(2) Setting the model parameters
The key parameters for LS-SVM are the regularization parameter γ and the parameters of kernel function which are determined by the setting type of kernel function.

(3) LS-SVM model training
Based on the input vectors construction and parameters setting, we can train the LS-SVM model as described in subsection II.A. First, construct the kernel function matrix and solve the linear equation equations of N dimensions. Then, compute the Lagrange multipliers α and deviant b and determine the decision function $f(x)$.

(4) Prediction with new input
One-step prediction can be realized with the decision function $f(x)$ and the new input vector.

C. ANOMALY DETECTION BASED ON UNCERTAINTY ESTIMATION OF LS-SVM

Unfortunately, the LS-SVM model cannot provide predictive intervals. Alternatively, one can marginalize the model parameters to compute the variance of the predicted value. As a result, the predicted intervals with confidence level P can be estimated with the variance, and uncertainty representation capability of the LS-SVM will be improved. This estimated uncertainty interval can be used as the normal range for the detected value. The updated value which exceeds the intervals can be treated as anomalous point with confidence level P. While, if the updated value locates in the intervals, it can be labeled as normal point.

The whole LS-SVM-based anomaly detection framework is shown in Fig. 2 [24].

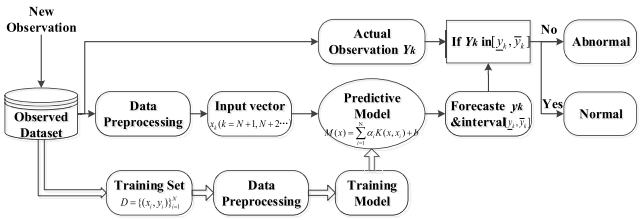


FIGURE 2. The framework of anomaly detection based on LS-SVM time series prediction.

From Fig. 2, the predicted interval can be obtained by computing the predicted errors. If the prediction error exceeds the normal range, the new point will be labeled as anomalous. For processing the anomalies, there are two strategies, i.e., anomaly detection and mitigation (ADAM) and anomaly detection only (AD) [17]. To keep the accuracy of a long-period detection, the AD strategy is applied in this work which merely highlights the anomalous sample without adding the predicted point into the input vector.

Note that the key step to realize anomaly detection based on the above framework is to calculate the normal range of prediction error. In the related papers, the normal range of error is defined by multiple times of the maximum error during the training process and it is significantly influenced by the performance of training. In [17], the mean and variance of the predicted error are computed by 10-fold cross validation with the cost of time-consuming computation. Thus, to effectively estimate the normal interval of the prediction error, the measurement error is estimated by considering both error level of training data and model estimation. The detailed description is presented as follows.

First, we assume that the model parameter w and the error vector e obey Gaussian distribution.

$$\mathbf{w} \sim N(0, \frac{1}{\mu}) \tag{9}$$

$$e \sim N(0, \frac{1}{\zeta}) \tag{10}$$

where $\frac{1}{\mu}$ and $\frac{1}{\zeta}$ are the variances of $w_j (j = 1, 2, \ldots, n_h)$ and $e_i (i = 1, 2, \ldots, N)$, respectively.

By equal variance processing w and e, the optimized questions in (2) can be transformed to

$$\min_{\mathbf{w},b,e} J(\mathbf{w}, b, e) = \mu E_{\mathbf{w}} + \zeta E_d \tag{11}$$

where

$$E_{\mathbf{w}} = \frac{1}{2}\mathbf{w}^T\mathbf{w} \tag{12}$$

$$E_d = \frac{1}{2}\sum_{i=1}^{N} e_i^2 \tag{13}$$

Here μ and ζ are the regularized parameters.

Then, the optimal solution of LS-SVM regression is given by

$$y^* = \mathbf{w}^{*T}\phi(\mathbf{x}) + b^* \tag{14}$$

The new predicted value y_{N+1} corresponding to a new update input vector x_{N+1} can be computed with (14). The observation value of y_{N+1} is

$$Y_{N+1} = y_{N+1} + e_{N+1} \tag{15}$$

where $E(e_{N+1}) = 0$ and $D(e_{N+1}) = 1/\zeta_{N+1}$. Due to the independence of y and e, the mean and variance of observation Y_{N+1} can be derived as

$$E(Y_{N+1}) = E(y_{N+1}) + E(e_{N+1}) = y^*_{N+1} \tag{16}$$

$$D(Y_{N+1}) = D(y_{N+1}) + 1/\zeta_{N+1} \tag{17}$$

Based on (16), the predicted value can be viewed as the mean of true value. Hence, the variance can be obtained as

$$\begin{aligned} D(y_{N+1}) &= E[(y - y^*)^2] \\ &= E\{[(\mathbf{w}^T\phi(\mathbf{x}) + b) - (\mathbf{w}^{*T}\phi(\mathbf{x}) + b^*)]^2\} \\ &= \varphi(\mathbf{x})^T Q \varphi(\mathbf{x}) \\ &= \varphi(\mathbf{x})^T \mathbf{H}^{-1} \varphi(\mathbf{x}) \end{aligned} \tag{18}$$

where $\varphi(\mathbf{x}) = [\phi(\mathbf{x}); \vec{\mathbf{1}}]$ and $Q = convar(w, b)$ is the covariance matrix.

$$Q = H^{-1} = \begin{bmatrix} \dfrac{\partial^2 J}{\partial w^2} & \dfrac{\partial^2 J}{\partial b \partial w} \\ \dfrac{\partial^2 J}{\partial w \partial b} & \dfrac{\partial^2 J}{\partial b^2} \end{bmatrix}^{-1} \tag{19}$$

The detailed solution of $D(y_{N+1})$ can be found in [26]. Finally, one has $Y_{N+1} - y_{N+1} \sim N(0, \sqrt{D(Y_{N+1})})$. By setting the confidence level as $P = 100\%(1 - \alpha)$, the confidence interval of variance is

$$\left[t_{\alpha/2,N-1}\sqrt{D(Y_{N+1})}, t_{\alpha/2,N-1}\sqrt{D(Y_{N+1})}\right] \tag{20}$$

and the normal range of the prediction is

$$\left[y_{N+1} + t_{\alpha/2,N-1}\sqrt{D(Y_{N+1})}, y_{N+1} + t_{\alpha/2,N-1}\sqrt{D(Y_{N+1})}\right] \tag{21}$$

Thus, anomaly detection based on uncertainty estimation of LS-SVM can be realized by comparing the new point and the confidence interval.

III. PROPOSED FRAMEWORK FOR ANOMALY DETECTION

Compared with the anomalous points which are generally caused by the influence of noise, error code, and instantaneous events, more meaningful and effective anomalies always appear in the fragment mode. The former framework shown in Fig. 2 makes no condition on the relationships of different points, for which the system may face some unnecessary false-alarms. In the real application, unnecessary false-alarms may significantly affect the performance of anomaly detection method. In this work, the detection framework for fragment anomaly is proposed based on statistics theory.

A. FRAGMENT ANOMALY DETECTION STRATEGY

Fragment anomaly is an extension of point anomaly, in which the time series fragment is composed of time series points. One example of fragment anomaly is shown in Fig. 3.

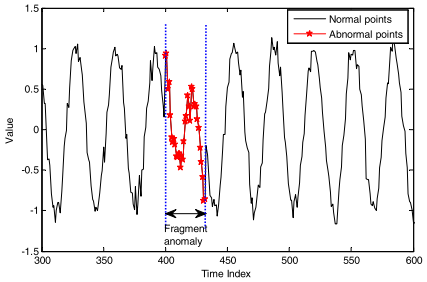


FIGURE 3. One example of fragment anomaly.

In Fig. 3, the normal series is a sinusoidal series with noise, and one fragment anomaly appears at the time index from 400 to 432 with the changed period. In this work, we define a time series observation as $O(t_0)$ at time instant t_0. Thus, a time fragment (TF) can be written as $TF(l, t_0)$ with the ending time instant t_0 and interval l, i.e., $TF(l, t_0) = [O(t_0 - l + 1), O(t_0 - l + 2), \cdots, O(t_0)]$. The purpose of fragment anomaly detection is to discover the anomalous degree of the corresponding time subseries

In aerospace application, the anomalies in the telemetry data can show the degraded or abrupt characteristics. In what follows, we present two kinds of detection strategies for fragment anomaly.

Strategy I:

Applying the point anomaly detection method, we can detect each potential anomalous point with confidence probability P. If the observation value at time instant t is detected as normal, then $|O(t)| = 0$. Otherwise $|O(t)| = 1$.

We denote the number of anomalies in the time fragment $TF(l, t_0)$ as

$$\begin{aligned} |TF(l, t_0)| &= |O(t_0 - l + 1), O(t_0 - l + 2), \cdots, O(t_0)| \\ &= \sum_{i=1}^{l} |O(t_0 - l + i)| \end{aligned} \tag{22}$$

For the time series fragment with length l, the anomaly probability for $|TF(l, t_0)|$ can be calculated as

$$P(|TF(l,t_0)|) = \begin{pmatrix} l \\ |TF(l,t_0)| \end{pmatrix} (1-P)^{|TF(l,t_0)|} P^{l-|TF(l,t_0)|} \tag{23}$$

It is easy to know that the confidence probability of the normal state equals P for each observation point and the probability of one point exceeding this range is 1-P. We assume that the observation points are independent to each other. Thus, $|TF(l, t_0)|$ follows the Bernoulli distribution, which allows us to calculate the probability of anomalies with the number of $|TF(l, t_0)|$ in the fragment. The higher the value of $|TF(l, t_0)|$ is, the lower the probability $P(|TF(l, t_0)|)$ remains (which means that $P(|TF(l, t_0)|)$ is small in general). Once this situation comes out, we can treat it as the anomalous fragment. Without loss of generality, we can set the lower bound (LB) for the threshold of $|TF(l, t_0)|$ and label the fragment anomalous as $|TF(l, t_0)| \geq LB$.

$$\begin{aligned} &P(TF(l,t_0), LB) \\ &\quad = P(|TF(l,t_0)| = LB) + P(|TF(l,t_0)| = LB+1) \\ &\qquad + \ldots + P(|TF(l,t_0)| = l) \\ &\quad = \sum_{|TF(l,t_0)|=LB}^{l} \begin{pmatrix} l \\ |TF(l,t_0)| \end{pmatrix} (1-P)^{|TF(l,t_0)|} P^{l-|TF(l,t_0)|} \end{aligned} \tag{24}$$

where l and LB should be selected appropriately to make $P(TF(l, t_0), LB)$ stay in a very low level, e.g., $P(TF(l, t_0), LB) \leq 1\%$. This means the probability of fragment anomaly equals $1 - P(TF(l, t_0), LB) \geq 99\%$ when the number of anomalous points is larger than the LB in a single fragment. Namely, once the number of anomalies exceeds the LB being detected, it is a fragment anomaly with confidence level 99%.

Therefore, the confidence level c should be considered based on l and LB as $1 - P(TF(l, t_0), LB) \geq c$.

Strategy II:

On the other hand, we can also use the error series for anomaly detection. We assume that the errors follow independent normal distributions such that no trend for the random error series exists. For the error series with number h showing monotonic feature, the associated probability is

$$P(h) = \frac{1}{2} \cdot \frac{1}{3} \cdot \ldots \cdot \frac{1}{h} \cdot 2 = \frac{2}{h!} \tag{25}$$

When $h = 7$, we can obtain $P(h = 7) \geq 2/(7!) \approx 0.04\%$ which is a very low probability. Thus, there is a small probability of fragment anomaly. However, this strategy is very sensitive to the assumption. Once the error series is not generated from normal distributions, we may need to set a much larger h for this strategy.

As shown in Fig. 4, sliding window is applied to obtain the detected segmentation of time series. In the detection, the point anomaly and the error series of the former observation points (with length l-1 in the sliding window) are considered. The procedure of fragment anomaly detection is introduced as follows.

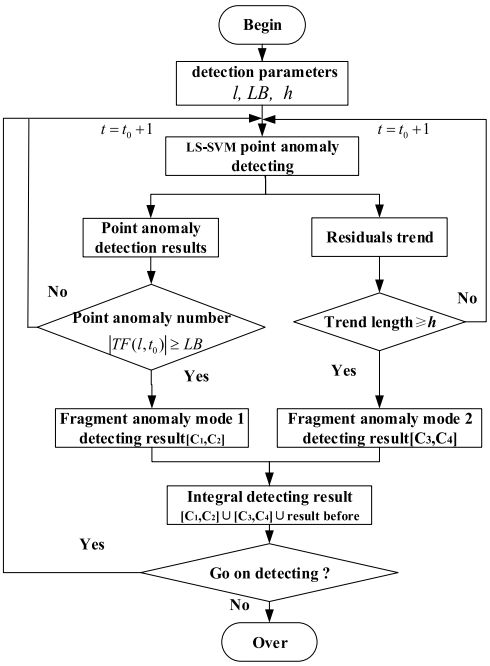


FIGURE 4. The algorithm flow of fragment anomaly detection.

Step 1: Checking the parameters setting for the model

The key parameters, i.e., the length of detected time series fragment l, the lower bound of the number of anomalous points LB, and the continuous monotonic length h, are required to be set in advance. In addition, all of the parameters should satisfy

$$\begin{cases} 1 - P(TF(l,t_0), LB) \geq c, \\ P(h) = \frac{2}{h!} \geq c. \end{cases} \tag{26}$$

In the real application, the users can determine the parameter of l and h by controlling the probability bound of c. In this work, false alarm is one of the most important factors restricting the applicability of the method. So c is considered larger than 99%.

Step 2: Point anomaly detection with LS-SVM algorithm

According to the time series prediction with LS-SVM, the predicted error can be computed to determine the detected result.

Step 3: Fragment anomaly detection with strategy I

We can compare the number of point anomalies $|TF(l, t_0)|$ with the setting threshold LB. If $|TF(l, t_0)| \geq LB$, the detected fragment is labeled as anomalous and the anomalous location is marked as [C1, C2] which represents the time range from $t_0 - l + 1$ to t_0. If no anomaly is detected, [C1,C2] is defined as a null set and the algorithm continues to the next time instant.

Step 4: Fragment anomaly detection with strategy II

In this step, we need to combine the current error and former errors of length h-1 to judge the monotonic feature. If the monotonic situation occurs, the detected fragment is determined as anomalous, and the anomalous location is marked as [C3, C4] which represents the time range

from $t_0 - h + 1$ to t_0. If no anomaly is detected, [C3, C4] is defined as a null set and the algorithm continues to the next time instant.

Step 5: Fusion of the two anomalous results

The union of two types of detected results is calculated to obtain the fusion result, i.e., [C1, C2] ∪ [C3, C4].

Step 6: If all of the detections cannot be ended, then repeat Steps 1 to 5. Otherwise, fuse all the detected anomalous fragments and end the detection.

B. EVALUATION CRITERIA FOR FRAGMENT ANOMALY DETECTION

For point anomaly detection, FPR (False positive rate), FNR (False negative rate), and ACC (Accuracy) are used to compare and evaluate the detection results. However, these indicators cannot be applied to estimate the performance of fragment anomaly detection. Therefore, for fragment anomaly detection, we will compare the whole detection results with the actual anomalous range in the series, from which the coincidence degree is addressed.

First, the coincidence range is computed by comparing the detected anomalous range and the actual anomalous range. We record the results in Table 1.

TABLE 1. Coincidence degree for fragment anomaly detection.

Detection Results for Each Segmented Series	True Anomalous Range (TA)
Detected anomalous range (DA)	Intersection of True Negative (ITN)
Detected normal range (DN)	Intersection of False Positive (IFP)

Then, we define the following new anomalous detection evaluation criteria to evaluate and verify the performance of fragment anomaly detection.

(1) True Negative True Rate (TNTR)

$$TNTR = \frac{ITN}{TA} \tag{27}$$

TNTR represents the ratio of the detected range of an actual fragment anomaly. This means that the larger TNTR is, the higher the detected ratio is. For example, TNTR = 100% implies that all of the anomalous samples are highlighted. When TNTR =0 (or TNTR ≠ 0), no anomalous sample is detected and the missing rate is 100% (some of the anomalous samples are detected).

(2) True Negative Detected Rate (TNDR)

$$TNDR = \frac{ITN}{DA} \tag{28}$$

TNDR is the ratio of the true negative detected range of the overall detected results, which indicates that the larger TNDR is, the lower the false detected ratio is. When TNDR = 0, all of the actual anomalous samples are missed and the complete false detection happens.

According to the two proposed evaluation criteria, we can summarize the detection mechanisms as follows.

(1) Anomaly detection achieves better performance when TNTR and TNDR are large.

(1) The larger TNTR is, the higher the detected ratio is. When TNTR = 100%, all of the anomalous samples are discovered.

(2) The larger TNDR is, the lower the false detection ratio is. When TNDR = 100%, no false detection happens.

(3) When TNTR = 0, no anomalous samples are highlighted.

(4) When TNDR = 0, complete false detection happens.

(6) If there are n actual anomalous fragments, we need to calculate TNTR n times for each single segment. If there are m detected anomalous fragments, we need to calculate TNDR m times for each single detected series.

(7) When two or more detected fragments have the same intersection with an actual anomalous fragment, the final TNDR can be obtained by fusing several different TNDRs.

Next, we will show some instances to assist the interpretation of the above evaluation criteria.

As shown in Fig. 5, the segments marked in orange are the actual anomalous subsequence (6 in total), with the index ranges [a, c], [f, k], [m, o], [p, q], [u, x], and [y, z], respectively. Correspondingly, the segments with blue mark are the detected anomalous fragments, which locate at [b, d], [e, g], [h, l], [n, r], [s, t], and [v, w], respectively. Note that different intersections represent different detection cases.

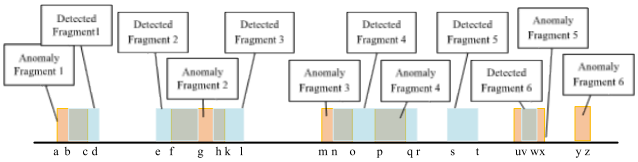


FIGURE 5. Some cases of fragment anomaly detection.

(1) The first detected fragment, referred to [b, d], is intersecting with one actual anomaly fragment (the first anomalous fragment), and its indicators TNTR and TNDR are computed as follows.

$$TNTR_1 = \frac{ITN}{TA} = \frac{c - b + 1}{c - a + 1} \tag{29}$$

$$TNDR_1 = \frac{ITN}{DA} = \frac{c - b + 1}{d - b + 1} \tag{30}$$

(2) Because the second detected fragment of [e, g] intersects with the actual anomaly fragment [f, k] (so does the third detected fragment [h, l]), the corresponding TNTR and TNDR are calculated as

$$TNTR_2 = \frac{g - f + 1}{k - f + 1} + \frac{k - h + 1}{k - f + 1} \tag{31}$$

$$TNDR_2 = \frac{g - f + 1}{g - e + 1} \tag{32}$$

$$TNDR_3 = \frac{k - h + 1}{l - h + 1} \tag{33}$$

$$TNDR[(2\&3)2] = \frac{(g - f + 1) + (k - h + 1)}{(g - e + 1) + (l - h + 1)} \tag{34}$$

(3) The fourth detected fragment intersects with the third and the fourth true anomaly fragments simultaneously, thus one has

$$TNTR_3 = \frac{o-n+1}{o-m+1} \tag{35}$$

$$TNTR_4 = \frac{q-p+1}{q-p+1} = 100\% \tag{36}$$

$$TNDR_4 = \frac{(o-n+1)+(q-p+1)}{(r-n+1)} \tag{37}$$

(4) [s, t] has no intersection with the actual anomaly, so $TNDR_5 = 0$.

(5) The real anomaly fragment [u, x] overlays the detected anomalous fragment [v, w], hence

$$TNDR_6 = \frac{w-v+1}{w-v+1} = 100\% \tag{38}$$

$$TNTR_5 = \frac{w-v+1}{x-u+1} \tag{39}$$

(6) The sixth anomaly fragment is not detected, so $TNTR_6 = 0$.

It is worth pointing out that the above evaluation criteria can only been applied after knowing the anomaly labels. Therefore, in this work, the anomalies in the actual telemetry are simulated with expert experience to verify the performance of the proposed framework.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

To evaluate performance of the proposed framework for anomaly detection, in this section, two types of public data sets are adopted to conduct experiments as follows:

(1) Open-source simulated data sets of MA time series are first applied to verify the performance of the proposed method.

(2) Open-source benchmark data set of NASA electromagnetic valve time series from space shuttle is used to test the detection ability for the real data set.

A. DATA SETS DESCRIPTION

(1) Simulated data sets

Simulated data of MA are generated from the stochastic process that used to test SVR algorithm in [27]. Specifically, the normal data samples are generated from specific distribution, in which we define anomalies as the data samples that do not fit the distribution. The anomalous samples are added into the normal data set randomly and the normal data series is given by (40).

$$x_0 = \sin(\frac{30\pi}{N} \cdot t) + n_0 \tag{40}$$

where $N = 1200$ and n_0 is the Gaussian noise with zero mean and variance 0.1.

Two types of abnormal series are presented as:

$$\begin{aligned} x_1 &= x_0 + e_1(t) \\ x_2 &= x_0 + e_1(t) + e_2(t) \end{aligned} \tag{41}$$

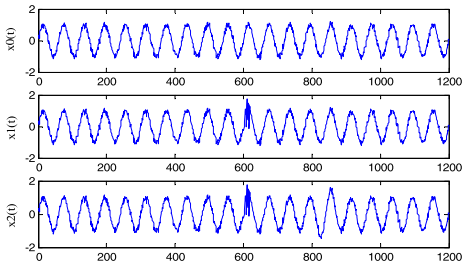


FIGURE 6. MA time series.

where e_1 is the Gaussian noise with mean 0 and variance 0.5 and the abnormal index of t ranges from 580 to 600. Besides, $e_2(t) = \sin(\frac{30\pi}{N} \cdot t)$ with the anomaly is located between 800 and 850. Fig.6 shows the normal MA time series and the simulated anomalous MA time series.

(2) Open-source benchmark Marotta data set

The Marotta series was obtained from NASA open-source benchmark data set, which is the state monitoring data of electromagnetic valve in space shuttle [28], [29]. In each cycle, the raw data includes 1000 samples and the normal and abnormal samples are marked by the engineers of NASA. To improve the operating efficiency of the algorithm, we re-sample the raw data set and allow each cycle to have 250 samples in the experiments.

Fig.7 shows the Marotta data set involving the normal and abnormal samples.

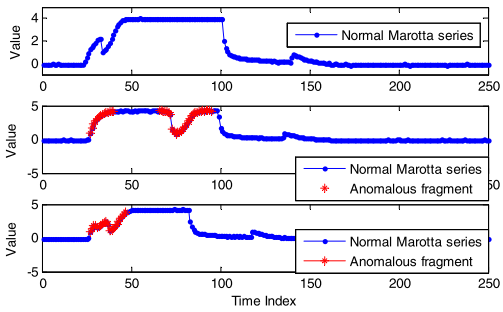


FIGURE 7. Abnormal series of Marotta data set for space shuttle.

In Fig. 7, the subseries marked with red star are the abnormal data samples recorded in two cycles.

B. EXPERIMENTAL RESULTS WITH MA DATA SET

As introduced in Section IV-A, the artificial anomalies are injected into the MA series located in $[580, 600] \cup [800,850]$. The first 400 data samples are used to train the model, and the latter 800 data samples are adopted as testing samples. In the experiments, the embedded dimension is set to 20 and the RBF kernel function is applied. The hyper-parameters are optimized as $C = 1.57532$ and $\sigma = 10.3102$.

The experimental results of x_2 are shown in Fig. 8 and Fig. 9.

For the fragment anomaly detection with strategy I, l is set as 6, LB is 3, and the detected anomalous fragment $t = [581, 602] \cup [800, 817] \cup [826, 848]$.

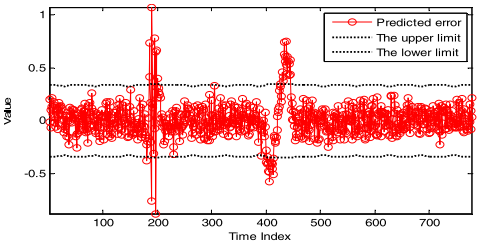


FIGURE 8. The error series and the normal range for MA predicted errors with simulated anomalies.

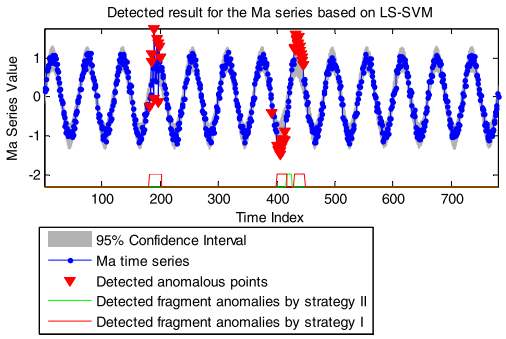


FIGURE 9. The fragment anomaly detection of MA series.

For the fragment anomaly detection with strategy II, h is set as 7 and the detected result is $t = [817, 827]$.

By integrating the two anomalous ranges, the final detected result is $t = [581, 602] \cup [800, 848]$, which is very close to the actual anomalous location $[580, 600] \cup [800, 850]$. In particular, the evaluated indicators TNTR and TNDR for the first detected anomaly fragment are

$$TNDR_1 = \frac{600 - 581 + 1}{602 - 581 + 1} = \frac{20}{22} = 90.9\%,$$
$$TNTR_1 = \frac{600 - 581 + 1}{600 - 580 + 1} = \frac{20}{21} = 95.2\%,$$

and TNTR and TNDR for the second anomaly fragment are

$$TNTR_2 = \frac{848 - 800 + 1}{850 - 800 + 1} = \frac{49}{51} = 96.1\%,$$
$$TNDR_2 = \frac{848 - 800 + 1}{848 - 800 + 1} = 100\%.$$

C. EXPERIMENTAL RESULTS WITH MAROTTA DATA SET FROM SPACE SHUTTLE

The experimental results with Marotta data set from space shuttle are shown in Fig. 10 and Fig. 11.

The detected result based on strategy I with $l = 6$ and $LB = 3$ is $t = [257, 279] \cup [309, 325] \cup [507, 524]$, while the detected results based on strategy II with $h = 7$ are $t = [296, 303] \cup [520, 527]$ and $t = [266, 280] \cup [305, 316]$. By fusing the two types of anomalous ranges, the final detected result is $t = [257, 280] \cup [296, 303] \cup [305, 325] \cup [507, 527]$, which is very close to the actual fragment anomalies.

According to the results shown in Fig. 11, we can find that the detected results are very close to the marked anomalies

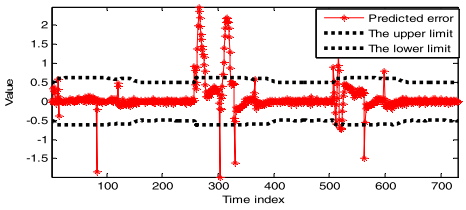


FIGURE 10. Error series and the normal range for Marotta predicted errors with simulated anomalies.

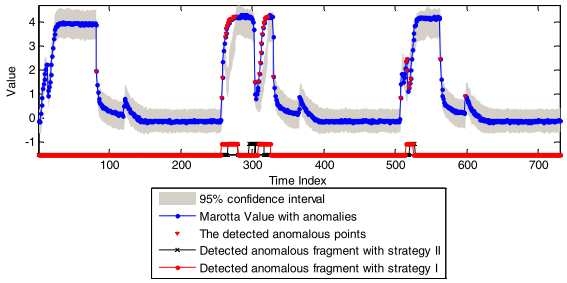


FIGURE 11. Anomaly fragment detection of Marotta value.

by experts. Basically, all of the anomalous fragments can be detected with the accurate locations and ranges. Because the accurate range is not provided by experts, the quantitative results of TNTR and TNDR are not given here.

V. CASE STUDY: ANOMALY DETECTION FOR AEROSPACE APPLICATION

A. FRAGMENT ANOMALY DESCRIPTION IN TELEMETRY DATA

Among the telemetry series, some aperiodicity telemetry series are generally influenced by the remote controlling which is relatively easy to detect with expert rules. Also, the degradation characteristics is hard to appear in short-term test with the low degradation speed. Given that the satellite orbit is relatively regular and the working condition has the periodic specialty, the analog telemetry series with the pseudo-periodic phenomenon are the focus of current work. In detail, we select the temperature and pressure series of catalytic bed from the actual satellite data sets to evaluate the proposed method. Actually, abnormal data samples are rare during the normal satellite in-orbit operation. Thus, we simulated artificial anomalies according to the expert experience and failure analysis of the actual satellites.

Three types of anomalies are defined as follows.

(1) Amplitude and trend anomaly

This type of anomaly may reflect the error of data lost, error code, and messy code. It can reveal the short-term changes in amplitude and trend with the following specific forms.

Type I anomaly: continuous unchangeable state in a short term.

Type II anomaly: sequence shocks in the time series.

Type III anomaly: unexpected concave side or convex side, including sine/cosine type and triangular pulse type.

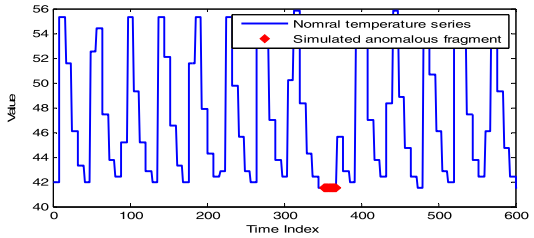


FIGURE 12. Type I amplitude anomaly in temperature telemetry.

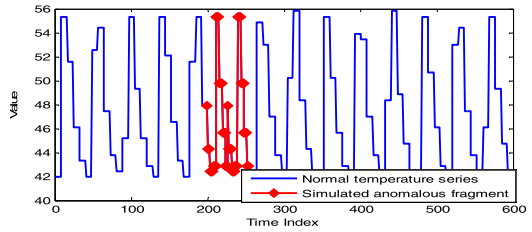


FIGURE 13. Type I time axis anomaly in temperature telemetry.

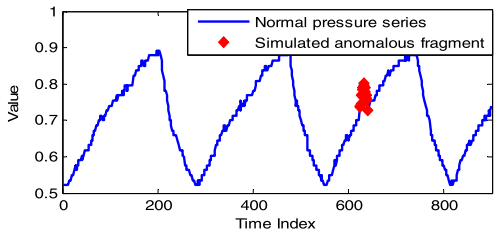


FIGURE 14. Sine convex anomaly in pressure series.

(2) Anomaly in time axis
The anomaly type represents mode switch error in the telemetry data. In what follows, we present two detailed sub-anomalies examples.
Type I anomaly: period compression.
Type II anomaly: period extension.

(3) Anomaly in noise
Short-term high power noise may be used to simulate the influence of the changes in spatial environments.

The above simulation anomalies are added in temperature and pressure series of catalytic bed, respectively. Fig. 12 shows the Type I amplitude anomaly I located in [353, 368], and Fig. 13 shows the Type I time axis anomaly located in [199, 254].

Fig. 14 represents the Type III amplitude anomaly located in [624,649], in which the sine convex anomaly is added into the actual pressure data set. Fig. 15 depicts the triangular pulse type added in the raw data set.

In this work, due to the similarity of the experiments as well as the limited paper space, we only describe the results of some typical experiments on different types of anomaly fragments.

B. DETECTION FOR AMPLITUDE AND TREND ANOMALY

(1) Experiments on temperature series of catalytic bed with Type I amplitude and trend anomaly.

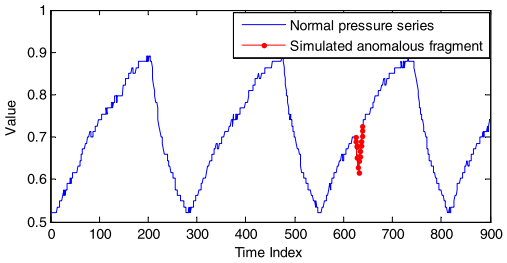


FIGURE 15. Triangular anomaly in pressure series.

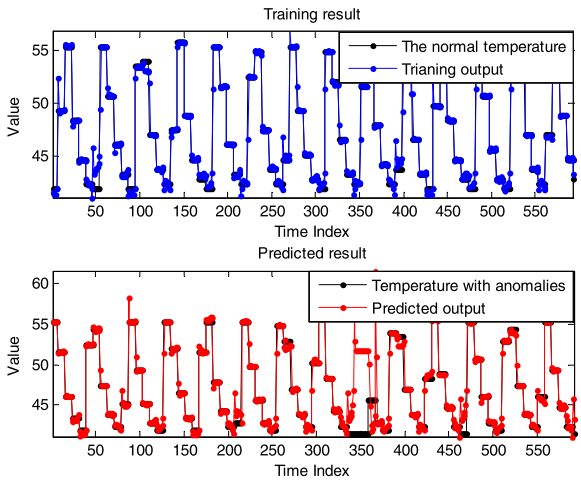


FIGURE 16. Training and predicted result for temperature series.

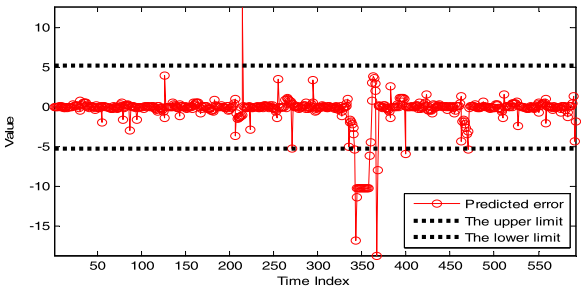


FIGURE 17. Error series and the normal range for temperature series with simulated anomalies.

Experimental setting: we set the embedded dimension as 10 and the anomaly detection parameters as $l = 6$, $LB = 3$, and $h = 8$. The optimized parameters searched by cross-validation are gam = 231.7315 and sig2 = 1.384763.

The detected results are presented in Figs. 16-18. Specifically, the training and prediction results are shown in Fig. 16, the predicted error series and the normal range of errors are given in Fig. 17, and the detected anomalous fragment is provided in Fig. 18.

The detected result is [342, 362] and the injecting anomalous range is [342, 358]. Thus, the detection is almost completely conforming to the real condition. In addition, TNTR and TNDR are

$$TNTR = \frac{358 - 342 + 1}{358 - 342 + 1} = 100\%$$

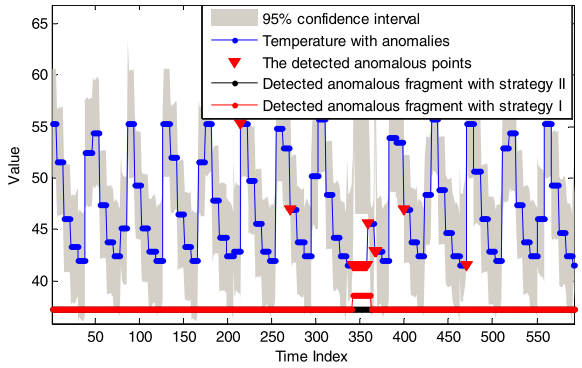


FIGURE 18. Detected anomalous fragment of temperature series with anomalies.

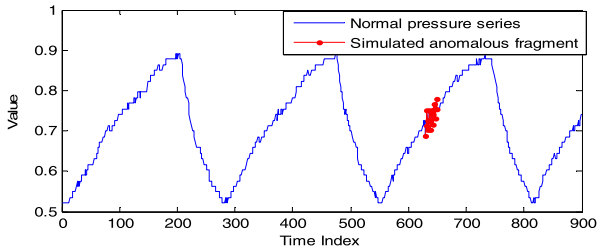


FIGURE 19. Pressure series of catalytic bed with amplitude and trend anomaly of Type II.

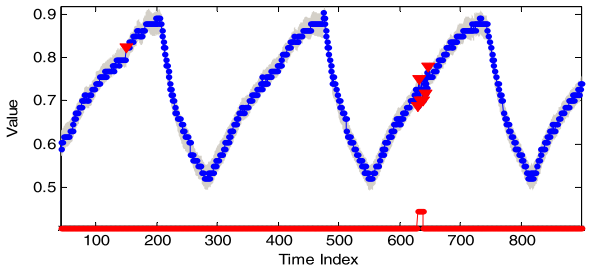


FIGURE 20. Detected anomalous fragment of pressure series.

and

$$TNDR = \frac{358 - 342 + 1}{362 - 342 + 1} = \frac{17}{21} = 81.0\%,$$

respectively.

(2) Anomaly detection for pressure series with Type II amplitude and trend anomaly.

For the pressure series, we add sequence shocks to it. Then, the Type II amplitude and trend anomaly can be observed in Fig.19, where the anomalies are injected in [630, 650]. Fig. 20 shows the detected anomalous fragment. With the same meanings, the detailed legends of curves are omitted in the following detection figures.

The detected anomalies are located in [630, 650] and the actual anomalous happened at [631, 649]. The TNTR and TNDR are computed as follows.

$$TNDR = \frac{649 - 631 + 1}{650 - 630 + 1} = \frac{19}{21} = 90.48\%,$$
$$TNTR = \frac{649 - 631 + 1}{649 - 631 + 1} = 100\%.$$

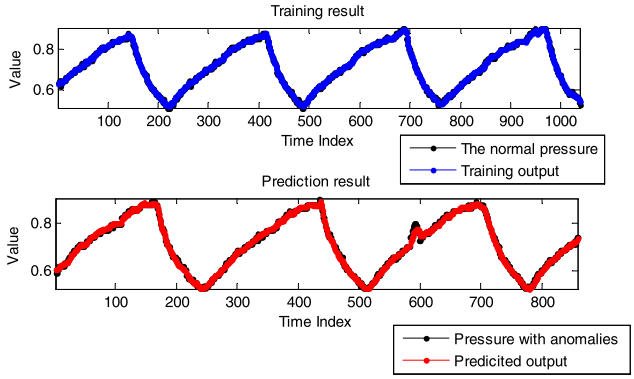


FIGURE 21. Training and predicted results for pressure series with simulated cosine anomalies.

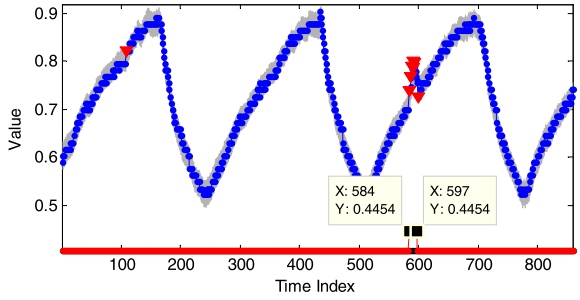


FIGURE 22. Detected fragment anomalies in pressure series with simulated cosine anomalies.

(3) Anomaly detection for pressure series with Type III amplitude and trend anomaly.

The amplitude anomaly of cosine type is injected in [583,599]. The embedded dimension is set as 40, gam = 20, sig2 = 0.215, and the other parameters remain unchanged. The detected results are shown in Fig. 21 and Fig. 22.

The detected range is [584,597] and the actual anomalous fragment is [583,599], thus, the interval is highly coincident. We also obtain

$$TNTR = \frac{597 - 584 + 1}{599 - 583 + 1} = \frac{14}{17} = 82.4\%,$$
$$TNDR = \frac{597 - 584 + 1}{597 - 584 + 1} = 100\%.$$

Hence, we can conclude that 82.4% actual anomalies are detected and no false detection happens.

(4) Anomaly detection for pressure series with Type III amplitude and trend anomaly.

A very small triangular anomaly is simulated and added into the normal pressure series, which locates in [597,613]. Embedded dimension equals to 40, gam = 20, sig2 = 0.215, $l = 7$, $LB = 3$, $h = 9$, and $p = 95\%$. The detected results are shown in Fig. 23 and Fig. 24.

The detected ranges are [598,608] and [610,616], and the actual anomalies are in [597,613]. The TNTR and

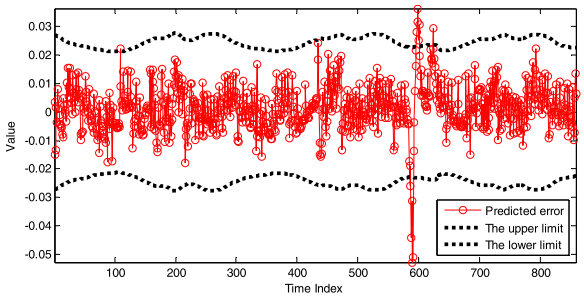


FIGURE 23. Predicted error for pressure series with simulated triangular anomalies.

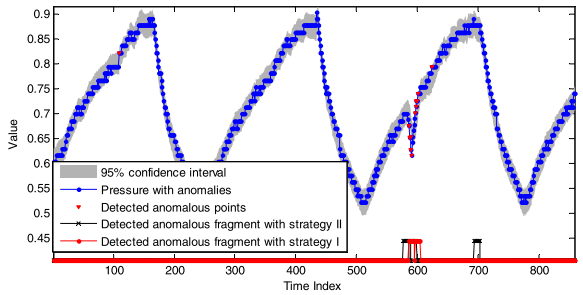


FIGURE 24. Detected anomalous fragment in pressure series with simulated triangular anomalies.

TNDR are

$$TNTR = \frac{(608-598+1)+(613-610+1)}{613-597+1} = \frac{11+4}{17} = 88.2\%,$$

$$TNDR_1 = \frac{608-598+1}{608-598+1} = 100\%,$$

$$TNDR_2 = \frac{613-610+1}{616-610+1} = \frac{4}{7} = 57.1\%.$$

The two detected fragments are matching the same actual anomalous fragment, so we can obtain

$$TNDR_1[(1\&2)] = \frac{15}{7+11} = 83.3\%.$$

C. ANOMALY DETECTION FOR THE TIME AXIS ANOMALY

In this experiment, the anomaly with period compression is injected into the temperature series. The embedded dimension is set as 10 and the anomaly detection parameters are $l = 6$, $LB = 3$, and $h = 8$. The optimized parameters searched by cross-validation are gam = 117.9015 and sig2 = 1.02117. Then, the detected results are given in Fig.25 and Fig.26.

The injected anomalies are located in [199, 254] and the detected anomalous fragment is $[211, 216] \cup [239, 244]$. In addition, the TNTR and TNDR are

$$TNTR_1 = \frac{(216-211+1)+(244-239)+1}{254-199+1} = \frac{13}{56} = 23.2\%,$$

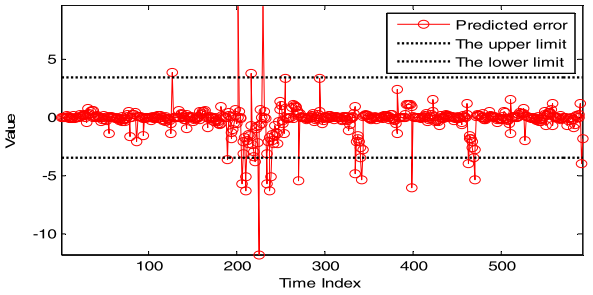


FIGURE 25. Predicted error for temperature series with period compression anomalies.

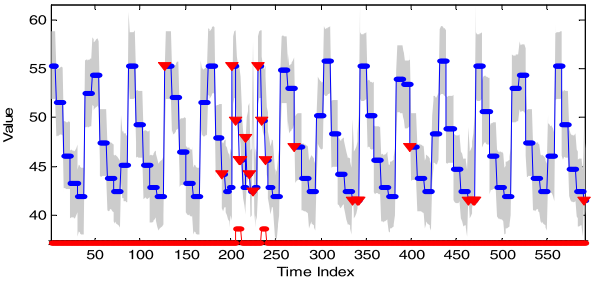


FIGURE 26. Detected fragment anomalies in temperature series with period compression anomalies.

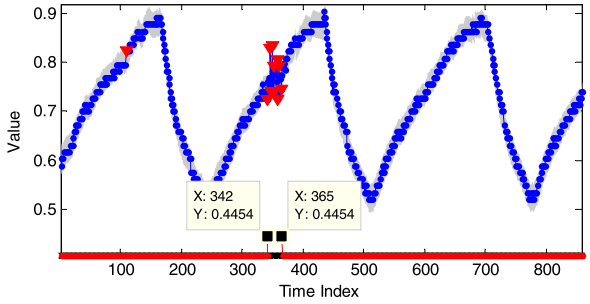


FIGURE 27. Detected fragment anomalies in pressure series with high power noise anomalies.

$$TNDR_1 = \frac{216-211+1}{216-211+1} = 100\%,$$

$$TNDR_2 = \frac{244-239+1}{244-239+1} = 100\%.$$

D. ANOMALY DETECTION FOR NOISE INTERFERENCE

An anomaly fragment is injected into pressure series with short-term high power noise. For the normal series, the training variance is 0.008. Thus, we simulated the high power noise with variance 0.025 (3 times to the normal value). The actual anomalous range is [340, 370] and the other parameters keep the same. One detected result is shown in Fig.27.

Given the influence of variance on the raw data, we implement 6 times detection to calculate the average values of TNTR and TNDR. Table 2 shows the results.

From Table 2, one can observe that, for each detection, the anomalous range can be closely detected.

TABLE 2. Anomaly detection for noise.

Index	ITN	TNTR	TNDR
1	38	92.68%	90.48%
2	38	87.80%	100%
3	27	65.85%	100%
4	30	73.17%	100%
5	35	85.37%	89.74%
6	35	85.37%	94.59%
average	33.83	82.52%	95.80%

E. EXPERIMENT ANALYSIS AND DISCUSSION

In this section, three types of anomaly fragment (which is unable to be detected by the threshold method), i.e., amplitude anomaly, time axis anomaly and noise interference, are injected into the real satellite telemetry data. In addition, some experiments are performed on the detailed subtypes of the above anomaly modes. The proposed framework can highlight all injected anomaly fragments. The only difference is that to label the whole or part of the anomaly fragments. This result is meaningful to the real aerospace application, because, with these highlights, we can accurately position the potential failures to improve the maintenance efficiency. Moreover, the false detection always happens near the actual anomalous fragment, which can provide early warnings to the ground staffs in the online application.

For the anomaly fragment which lasts a relative long time, missing detection may happen due to the influence of prediction window and the AD strategy. In particular, the TNTR and TNDR depend on the anomaly types and the model parameters. Therefore, in aerospace application, the characteristics of pseudo cycle and rare anomaly in telemetry data make the proposed framework effective to realize anomalous fragment detection.

VI. CONCLUSION

The contributions of this work can be concluded as follows: (1) a data-driven fragment anomaly detection method based on LS-SVM is proposed; (2) effective evaluation indicators are introduced to estimate the performance of fragment anomaly detection; (3) the predicted anomaly detection is applied to condition monitoring data sets (actual satellite with typical anomaly modes) to prove the effectiveness of the proposed anomaly detection framework.

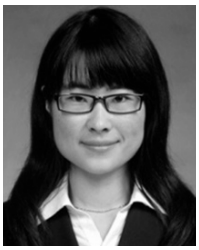
However, this framework has some limitations and further research is required in the future. For instance, most of the parameters are set by experiments and expert experience, e.g. the fragment length of l and h, which should be improved by combining the optimization algorithms and data characteristics. In addition, the adaptive updating strategy should focus more on the improvement of on-line detection performance. Moreover, anomaly detection for multivariate samples deserves a comprehensive study in the future.

REFERENCES

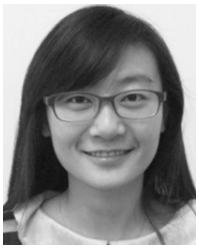
[1] Y. X. Cai, L. J. Cai, and Z. Lu, "Anomaly detection of online monitoring data of power equipment based on association rules and clustering algorithm," in *Proc. EEEIS*, Xi'an, China, Dec. 2016, pp. 289–298.
[2] Y. Song, D. Liu, C. Yang, and Y. Peng, "Data-driven hybrid remaining useful life estimation approach for spacecraft lithium-ion battery," *Microelectron. Rel.*, vol. 75, pp. 142–153, Aug. 2017.
[3] J. Ding, Y. Liu, L. Zhang, J. Wang, and Y. Liu, "An anomaly detection approach for multiple monitoring data series based on latent correlation probabilistic model," *Appl. Intell.*, vol. 44, no. 2, pp. 340–361, Mar. 2016.
[4] J. Pang, D. Liu, H. Liao, Y. Peng, and X. Peng, "Anomaly detection based on data stream monitoring and prediction with improved Gaussian process regression algorithm," in *Proc. IEEE PHM*, Cheney, WA, USA, Jun. 2015, pp. 1–7.
[5] D. K. Tolani, M. Yasar, A. Ray, and V. Yang, "Anomaly detection in aircraft gas turbine engines," *J. Aerosp. Comput., Inf., Commun.*, vol. 3, no. 2, pp. 44–51, Feb. 2006.
[6] S. Das, S. Sarkar, A. Ray, A. Srivastava, and D. L. Simon, "Anomaly detection in flight recorder data: A dynamic data-driven approach," in *Proc. ACC*, Washington, DC, USA, 2013, pp. 2668–2673.
[7] D. R. Azevedo, A. M. Ambrosio, and M. Vieira, "Applying data mining for detecting anomalies in satellites," in *Proc. EDCC*, Sibiu, Romania, 2012, pp. 212–217.
[8] B. R. Mohammad and W. M. Hussein, "A novel approach of health monitoring and anomaly detection applied to spacecraft telemetry based on PLSDA multivariate latent technique," in *Proc. REM*, 2014, pp. 1–6.
[9] T. Yairi, Y. Kawahara, R. Fujimaki, Y. Sato, and K. Machida, "Telemetry-mining: A machine learning approach to anomaly detection and fault diagnosis for space systems," in *Proc. SMC-IT*, Pasadena, CA, USA, 2006, pp. 468–476.
[10] S. Narasimhan and L. Brownston, "HyDE-A general framework for stochastic and hybrid model-based diagnosis," in *Proc. 18th Int. Workshop Principles Diagnosis DX-7*, 2007, pp. 162–169.
[11] Y. Bu, T.-W. Leung, A. W.-C. Fu, E. Keogh, J. Pei, and S. Meshkin, "WAT: Finding top-K discords in time series database," in *Proc. Siam Int. Conf. Data Mining*, Minneapolis, MN, USA, 2007, pp. 1–6.
[12] C. D. Truong and D. T. Anh, "An efficient method for motif and anomaly detection in time series based on clustering," *Int. J. Bus. Intell. Data Mining*, vol. 10, no. 4, pp. 356–377, Jan. 2015.
[13] D. Kifer, S. Ben-David, and J. Gehrke, "Detecting change in data streams," in *Proc. VLDB*, Toronto, Canada, 2004, pp. 180–191.
[14] V. Chandola, D. Cheboli, and V. Kumar. (Feb. 2009). "Detecting anomalies in a time series database," College Sci. Eng., Univ. of Minnesota, Minneapolis, MN, USA, Tech. Rep. 09-004. [Online]. Available: https://www.cs.umn.edu/research/technical_reports/view/09-004
[15] A. G. Rekha, "A fast support vector data description system for anomaly detection using big data," in *Proc. ACM SAC*, Salamanca, Spain, 2015, pp. 931–932.
[16] D. Pan, D. Liu, J. Zhou, and G. Zhang, "Anomaly detection for satellite power subsystem with associated rules based on kernel principal component analysis," *Microelectron. Rel.*, vol. 55, no. 9, pp. 2082–2086, Aug./Sep. 2015.
[17] D. J. Hill and B. S. Minsker, "Anomaly detection in streaming environmental sensor data: A data-driven modeling approach," *Environ. Model. Softw.*, vol. 25, no. 9, pp. 1014–1022, Sep. 2010.
[18] R. Fujimaki, T. Yairl, and K. Machlda, "An anomaly detection method for spacecraft using relevance vector learning," in *Proc. PAKDD*, 2005, pp. 785–790.
[19] B. Pincombe, "Anomaly detection in time series of graphs using ARMA processes," *Asor Bull.*, vol. 24, no. 3, pp. 2–3, Jan. 2005.
[20] H. Z. Moayedt and M. A. Masnadt-Shtrazi, "Arima model for network traffic prediction and anomaly detection," in *Proc. ITSim*, Kuala Lumpur, Malaysia, 2008, pp. 1–6.
[21] F. Knorn and D. J. Letth, "Adaptive Kalman flltering for anomaly detection in software appliances," in *Proc. INFOCOM*, Phoenix, AZ, USA, 2008, pp. 1–6.
[22] J. Ma and S. Perkins, "Online novelty detection on temporal sequences," in *Proc. ACM SIGKDD*, Washington, DC, USA, 2003, pp. 613–618.
[23] K. De Brabanter, J. De Brabanter, J. A. K. Suykens, and B. De Moor, "Approximate confidence and prediction intervals for least squares support vector regression," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 22, no. 1, pp. 110–120, Jan. 2011.

[24] G. Song, J. Liang, D. Liu, and Y. Peng, "Anomaly detection of condition monitoring with predicted uncertainty for aerospace applications," in *Proc. ICEMI*, Qingdao, China, 2015, pp. 248–253.
[25] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Process. Lett.*, vol. 9, no. 3, pp. 293–300, Jun. 1999.
[26] T. V. Gestel *et al.*, "Financial time series prediction using least squares support vector machines within the evidence framework," *IEEE Trans. Neural Netw.*, vol. 12, no. 4, pp. 809–821, Jul. 2001.
[27] K. P. Chan, W. C. Fu, and C. Yu, "Data structures and algorithms Haar wavelets for efficient similarity search of time series: With and without time warping," *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 3, pp. 686–705, Jun. 2003.
[28] Y. Chen *et al.* (2005). *The UCR Time Series Classification Archive*. [Online]. Available: http://www.cs.ucr.edu/~eamonn/discords/
[29] B. Farrell and S. Santuro. (2005). *NASA Shuttle Valve Data*. [Online]. Available: http://www.cs.fit.edu/pkc/nasa/data/

DATONG LIU (M'11–SM'16) received the B.Sc. degree in automatic test and control with a minor in computer science and technology, and the M.Sc. and Ph.D. degrees in instrumentation science and technology from the Harbin Institute of Technology (HIT), Harbin, China, in 2003, 2005, and 2010, respectively. He was a Visiting Scholar with The University of Arizona, Tucson, AZ, USA, from 2013 to 2014. He is currently an Associate Professor with the Department of Automatic Test and Control, School of Electrical Engineering and Automation, HIT. His current research interests include automatic test, system condition monitoring, machine learning-based test data analysis, data-driven prognostics and health management, and lithium-ion battery degradation modeling and prognostics.



JINGYUE PANG was born in Hebei, China, in 1988. She received the B.S. degree in measurement and control technology and instrumentation from the Chongqing University of Technology, Chongqing, China, in 2011, and the M.S. degree in instrument science and technology from the Harbin Institute of Technology, Harbin, China, in 2013, where she is currently pursuing the Ph.D. degree in instrument science and technology. Her research interest includes prognostics and health management, monitoring data processing, and anomaly detection on monitoring data.



GE SONG was born in Henan, China, in 1991. She received the B.S. and M.S. degrees from the Harbin Institute of Technology, Harbin, China, in 2014 and 2016, respectively. She is currently with SAIC Motor Corporation, Ltd., Shanghai, China. Her research interest includes intelligent data analysis, data mining, and statistics learning.

WEI XIE received the B.S. degree in applied mathematics from Shanghai University in 2010, and the Ph.D. degree in systems and industrial engineering from The University of Arizona in 2013. He is an Associate Professor with the School of Business Administration, South China University of Technology. His current research efforts focus on business analytics, reliability and service logistics, and operations management and marketing interfaces. His research work has been published in *IIE Transactions*, *Naval Research Logistics*, the *European Journal of Operational Research*, the IEEE TRANSACTIONS ON RELIABILITY, and the *International Journal of Production Research*.

YU PENG (M'10) received the B.S. degree in measurement technology and instrumentation, and the M.Sc. and Ph.D. degrees in instrumentation science and technology from the Harbin Institute of Technology (HIT), Harbin, China, in 1996, 1998, and 2004, respectively. He is currently a Full Professor with the Department of Automatic Test and Control, School of Electrical Engineering and Automation, HIT. His current research interests include automatic test technologies, virtual instruments, system health management, and reconfigurable computing.

XIYUAN PENG received the B.S., M.S., and Ph.D. degrees from the Harbin Institute of Technology (HIT), Harbin, China, in 1984, 1987, and 1992, respectively. He is currently a Professor and a Ph.D. Supervisor with the Department of Automatic Test and Control, School of Electrical Engineering and Automation, HIT. His current research interests include automatic testing technology and intelligent diagnostics and prognostics.