MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Anomaly Detection and Diagnosis Using Pre-Processing and Time-Delay Autoencoder

Liu, Bryan; Guo, Jianlin; Koike-Akino, Toshiaki; Wang, Ye; Kim, Kyeong Jin; Parsons, Kieran; Orlik, Philip V.; Hashimoto, Shigeru; Yuan, Jinhong

TR2021-107 September 16, 2021

Abstract

This paper proposes an anomaly detection algorithm for a factory automation system, which jointly performs data pre-processing and time-delay autoencoder (TDAE) with a hybrid loss function. The source data are pre-processed by digital filters before feeding into a TDAE for anomaly detection. The digital filters extract analog signals from a variety of frequency bands to facilitate identifying anomalies. The pre-processed data then takes time-delay reform to explore temporal relationship of data signals. In addition, two anomaly diagnosis algorithms, a statistical based method and an autoencoder based method, are presented. Numerical results show that time-delay reform can improve the anomaly detection accuracy compared to the conventional autoencoder. Data pre-processing can further improve the anomaly detection accuracy. Moreover, we confirm that our anomaly diagnosis algorithms outperform traditional method that does not perform data pre-processing and time-delay reform.

IEEE International conference on emerging technologies and factory automation 2021

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Anomaly Detection and Diagnosis Using Pre-Processing and Time-Delay Autoencoder

Bryan Liu*†, Jianlin Guo*, Toshiaki Koike-Akino*, Ye Wang*, Kyeong Jin Kim*,
Kieran Parsons*, Philip Orlik*, and Jinhong Yuan†
*Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139, USA
†Electrical Engineering and Telecommunication, University of New South Wales, Sydney, Australia

***Abstract*—This paper proposes an anomaly detection algorithm for a factory automation system, which jointly performs data pre-processing and time-delay autoencoder (TDAE) with a hybrid loss function. The source data are pre-processed by digital filters before feeding into a TDAE for anomaly detection. The digital filters extract analog signals from a variety of frequency bands to facilitate identifying anomalies. The pre-processed data then takes time-delay reform to explore temporal relationship of data signals. In addition, two anomaly diagnosis algorithms, a statistical based method and an autoencoder based method, are presented. Numerical results show that time-delay reform can improve the anomaly detection accuracy compared to the conventional autoencoder. Data pre-processing can further improve the anomaly detection accuracy. Moreover, we confirm that our anomaly diagnosis algorithms outperform traditional method that does not perform data pre-processing and time-delay reform.**

***Index Terms*—Factory automation, Deep learning, Anomaly detection, Fault diagnosis**

I. Introduction

The performance of factory automation (FA) systems is typically determined by a factor of downtime, which describes the period of time that system stops its production. Manufacturing machines in the FA system might stop working due to undesired faults, intrusions and system failures [1]. Therefore, it is crucial to detect the anomalies, where a machine is in an abnormal condition whenever anomalies occur [2], [3]. Traditional machine learning algorithms such as isolation forest (IF), one-class support vector machine (OC-SVM) and local outlier factor (LOF) have been widely used for anomaly detection [4]–[8]. Without introducing a large computation latency, these unsupervised machine learning techniques are easy to be implemented but the detection performance is limited since a complex data structure is involved for analysis.

Modern deep learning techniques have been applied to anomaly detection for various manufacturing systems due to its excellent performance to learn a complex data structure [9]. Among numerous deep learning techniques, autoencoder (AE), as a powerful technique to learn the underlying representation of the data [10], is the core technology in deep anomaly detection [2], [11]. The attractiveness of employing AE on factory anomaly detection is that it can inherit various neural network architecture to find a best-fit compressed representation of the original data. As such, AE has shown a robust and better anomaly detection performance in many real-world problems [12], [13]. Nevertheless, the AE-based algorithms are designed relying on the complex structure of neural networks, where the source data are usually not pre-processed. After an off-line training, the overall AE performance can be degraded by disruption that contains undesirable information.

In the meantime, as an extension of anomaly detection, several anomaly diagnosis methods have been studied [12], [14], [15]. Once an anomaly has been detected, anomaly diagnosis is to identify the exact sources that have caused the anomaly. In FA system, a source can be a machine, a hardware part, or a software program. Since many sensors are typically placed to collect data in FA system, the anomaly diagnosis aims at finding specific sensors that have captured anomalous events. The anomaly score of an anomaly detection algorithm describes the status of sensor signal at a specific time and the individual anomaly contribution of each sensor signal needs to be characterized according to the specific anomaly detection algorithm.

Bryan Liu conducted this research during his internship at MERL.

This paper firstly introduces a pre-processing based time-delay autoencoder (Prep-TDAE) structure for the anomaly detection in FA systems. The proposed Prep-TDAE structure pre-processes the sensor signals before feeding into a TDAE. In specific, digital filters are used to process the sensor signals to extract the data features among a preset frequency range. By filtering the signals, we expect the TDAE to represent the signals more precisely during the decoding, which results in an improved anomaly detection performance. In addition to the proposal of the anomaly detection algorithm, this paper then proposes anomaly diagnosis schemes based on the proposed Prep-TDAE algorithm, in which we introduce a statistics-based anomaly diagnosis method named Prep-TD-SD and an AE-based anomaly diagnosis method named Prep-TDAE-AD to demonstrate the effectiveness of the time-delay reform and data pre-processing. We test the proposed algorithms based on the data that is obtained from a real FA system. Numerical results show that Prep-TDAE outperforms the traditional machine learning anomaly detection algorithms and both anomaly diagnosis methods outperform the traditional diagnosis method, which does not perform time-delay reform and data pre-processing.

A. Notations

We use small boldface letters for vectors. For any real matrix $\mathbf{A}$, its transpose is denoted by $\mathbf{A}^{\mathrm{T}}$. $\mathrm{tr}\{\mathbf{A}\}$ indicates the trace of a matrix $\mathbf{A}$. $\mathbf{I}_N$ refers to the $N \times N$ identity matrix. The probability density function (PDF) of the multivariate normal distribution is denoted as $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. A uniform PDF is defined to be $\mathcal{U}(a, b)$, where a and b are the minimum and maximum values of the uniform interval, respectively. $|\mathbf{x}|$ denotes the cardinality of the vector $\mathbf{x}$. We let abs$(\cdot)$ denote the element-wise absolution function. We use $\Upsilon(\mathbf{x}) = \{i | x_i \in \mathbb{F}_2\}$ to denote a function which finds the set of indices of the binary variables in $\mathbf{x}$ and $\Phi(\mathbf{x}) = \{0, ..., |\mathbf{x}| - 1\} \setminus \Upsilon$ as the set of indices of the non-binary variables in $\mathbf{x}$. In general, we use a symbol with prime, e.g., $\mathbf{x}'$ to indicate the test data and a symbol without prime to indicate the training data.

II. Preliminaries

A. System Model and Data Type

In this paper, we consider an FA system that uses N sensors to collect data. Mixture of binary and analog sensors are considered. Since the analog sensors have different ranges of values, the sensors source signals are normalized by min-max normalization, where the resulting signals are in the range from 0 to 1. We use $\mathbf{x}^t = [x_0^t, x_1^t, ..., x_{N-1}^t]^{\mathrm{T}}$ to denote the min-max normalized sensor signals at the tth time index. Correspondingly, $(\mathbf{x}')^t$ indicates the test data and $(y')^t \in \{0, 1\}$ denotes the labelling of the test data, where 0 indicates that the sensor collected normal data at the tth time index and 1, otherwise. The time index indicates the sampling index during the consecutive data collection.

Using real FA devices such as robot and conveyor, a testbed has been built for data collection and functional test. The testbed uses about 100 sensors for data collection. The sensors include both binary sensors and analog sensors. The data are sampled at a sampling rate of 100 Hz. Normal and abnormal data are collected separately. Normal data are collected with a duration of 3232s and the abnormal data are collected multiple rounds, where each round has a duration of 205s. During abnormal data collection, the anomaly occurs at around 170s after the start of the data collection. Therefore, the abnormal data are labelled as 1 after 170s. A downsampling factor f can be considered to reduce the total number of training and test samples. The corresponding sampling rate becomes $100/f$ Hz.

B. TDAE Design

TDAE is one of the autoencoder structures to leverage neural networks for the task of representation learning [10]. TDAE has been proposed to learn acoustic-phonetic features in [16]. It learns the temporal relationships between the signals over time to reveal critical information in FA systems. As shown in Fig. 1, given a time-delay window size w, the source data are successively concatenated to a time-delay sequence. We define the concatenated sequence as follows.

Definition 1 *(Time-delay form):* The sensor signals are concatenated in w time samples, where $\mathbf{l}^t = [\mathbf{x}^t, ..., \mathbf{x}^{t+w-1}]$ denotes the tth concatenated sequence. □

In Fig. 1, the encoder and decoder refer to the neural network layers in the AE structure that compresses the data and recover the input

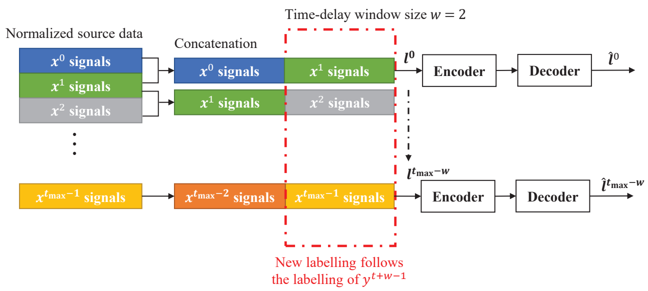

Fig. 1. An example of a conventional TDAE with a time-delay window size of $w = 2$.

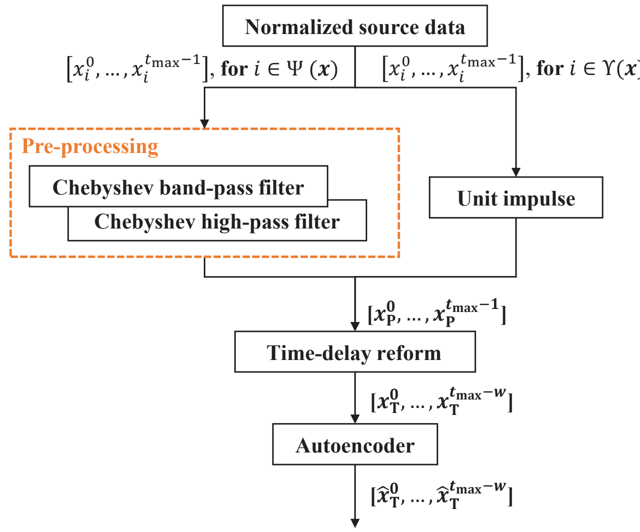

Fig. 2. Flowchart of the proposed Prep-TDAE.

sequence $\mathbf{l}^t$ to an estimate $\hat{\mathbf{l}}^t$. A loss function $L(\mathbf{l}^t, \hat{\mathbf{l}}^t)$ is preset to manipulate the reconstruction loss of the autoencoder, where the backpropagation is performed to tune the trainable parameters by stochastic gradient descent algorithms. The TDAE is trained by using normal data. When the reconstruction loss $L(\mathbf{l}^t, \hat{\mathbf{l}}^t)$ becomes large, it indicates that an abnormal data sequence (anomaly) is encountered by the anomaly detector. The one-to-one labelling of the tth concatenation is determined according to the last concatenated sequence y^{t+w-1}. Since the TDAE considers the sensor signals in both the time and space domains, it is able to achieve a better anomaly detection performance compared to the conventional AE, which directly processes $\mathbf{x}^t$ without the time-delay reform.

III. Anomaly Detection

In FA systems, the sensor signals are sensitive to the corruptions such as noise and inconsistent operations of the machines. Therefore, quickly and efficiently sensing the erroneous information and processing the information becomes the main target of anomaly detection.

In order to magnify the difference in signals between normal and abnormal data and improve the accuracy of anomaly detection, we introduce the Prep-TDAE algorithm, which pre-processes the normalized analog sensor signals by applying digital filters. The digital filters are chosen to filter the signals in a preset frequency band. As a result, the magnitude of each sensor's constituent frequencies that are out of the designed frequency bands are diminished. The TDAE is expected to learn the frequency characteristics better after digital filtering.

Fig. 2 depicts the flowchart of the proposed Prep-TDAE algorithm, where a Chebyshev type-1 band-pass and high-pass filters are employed to pre-process the normalized analog sensor signals [17]. The essential parameters, the order of the filters and the cutoff frequencies are chosen to minimize the TDAE's reconstruction loss of the normal data.

Let $\mathbf{x}_{\mathrm{F}}^t = \{\mathbf{x}_i^t | \ i \in \Upsilon(\mathbf{x}^t)\}$ and $\mathbf{x}_{\mathrm{A}}^t = \{\mathbf{x}_i^t | i \in \Psi(\mathbf{x}^t)\}$ be the set of binary sensor signals and analog sensor signals, respectively. The signals $\mathbf{x}_{\mathrm{A}}^t$ will be convoluted with the predefined Chebyshev type-1 filters. The output of the band-pass and high-pass filters of the analog sensors are defined as $\mathbf{x}_{\mathrm{B}}^t = [x_{\mathrm{B},0}^t, \ldots, x_{\mathrm{B},|\Psi(\mathbf{x}^t)|-1}^t]^{\mathrm{T}}$ and $\mathbf{x}_{\mathrm{H}}^t = [x_{\mathrm{H},0}^t, \ldots, x_{\mathrm{H},|\Psi(\mathbf{x}^t)|-1}^t]^{\mathrm{T}}$. As shown in Fig. 2, the unity impulse function indicates that the binary sensor signals are preserved. In the following, we define two forms of the processed data.

Definition 2 (Pre-processed form): The concatenation of the pre-processed analog signals and the binary signals at time index t, i.e., $\mathbf{x}_{\mathrm{P}}^t = [\mathbf{x}_{\mathrm{F}}^t, \mathbf{x}_{\mathrm{B}}^t, \mathbf{x}_{\mathrm{H}}^t]$, which has a cardinality of $N_{\mathrm{P}} = |\Upsilon(\mathbf{x}^t)| + 2|\Psi(\mathbf{x}^t)|$. □

Definition 3 (Time-delay pre-processed form): The concatenation of the pre-processed signals over a time-delay window of size w, i.e., $\mathbf{x}_{\mathrm{T}}^t = [\mathbf{x}_{\mathrm{P}}^t, \ldots, \mathbf{x}_{\mathrm{P}}^{t+w-1}]$, which has a cardinality of $N_{\mathrm{T}} = wN_{\mathrm{P}}$. □

Since the source data contain both binary and analog signals and the binary signals are preserved by the pre-processing, a hybrid loss function that combines cross-entropy and squared error is proposed. We define the sigmoid cross-entropy loss function as:

$$L_{\text{c}}(\mathbf{x}_{\text{T}}^t, \hat{\mathbf{x}}_{\text{T}}^t) = \sum_{i \in \Upsilon(\mathbf{x}_{\text{T}}^t)} \text{CE}(x_{\text{T},i}^t, \hat{x}_{\text{T},i}^t), \tag{1}$$

where

$$\text{CE}(x_{\text{T},i}^t, \hat{x}_{\text{T},i}^t) = -x_{\text{T},i}^t \log\Big(\frac{1}{1+\exp(-\hat{x}_{\text{T},i}^t)}\Big) - (1-x_{\text{T},i}^t)\log\Big(\frac{1}{1+\exp(-\hat{x}_{\text{T},i}^t)}\Big), \tag{2}$$

and the squared error loss function is defined as

$$L_{\text{s}}(\mathbf{x}_{\text{T}}^t, \hat{\mathbf{x}}_{\text{T}}^t) = \frac{1}{2} \sum_{i \in \Psi(\mathbf{x}_{\text{T}}^t)} (x_{\text{T},i}^t - \hat{x}_{\text{T},i}^t)^2. \tag{3}$$

Then, the final loss function $L(\mathbf{x}_{\text{T}}^t, \hat{\mathbf{x}}_{\text{T}}^t)$ at the tth time index is define to be

$$L(\mathbf{x}_{\text{T}}^t, \hat{\mathbf{x}}_{\text{T}}^t) = \frac{1}{N_{\text{T}}}\Big(\lambda L_{\text{c}}(\mathbf{x}_{\text{T}}^t, \hat{\mathbf{x}}_{\text{T}}^t) + (1-\lambda) L_{\text{s}}(\mathbf{x}_{\text{T}}^t, \hat{\mathbf{x}}_{\text{T}}^t)\Big), \tag{4}$$

where λ is a scaling factor to balance the loss contribution of each loss function. The larger of the reconstruction loss $L(\mathbf{x}_{\text{T}}^t, \hat{\mathbf{x}}_{\text{T}}^t)$, the more likely an abnormal data is observed at the tth time index. In the training phase of the autoenocder, the data that are obtained when a machine works normally are used to train the tunable parameters in the neural network. During the testing phase, the same steps are performed as in Fig. 2 but the data input becomes the test data. It will be shown in the numerical results that performing the pre-processing to the analog sensors is capable to improve the anomaly detection performance.

IV. Anomaly diagnosis

Once an anomaly is detected, anomaly diagnosis is to find the exact sources that work abnormally. Other than detecting the exact time index that a machine works abnormally, as described in Sec. I, anomaly diagnosis aims at identifying the sensors that captured abnormal data. In other words, each sensor's contribution to the anomaly needs to be measured. We present our statistics-based anomaly diagnosis scheme Prep-TD-SD and AE-based anomaly diagnosis scheme Prep-TDAE-AD herein.

A. Prep-TD-SD

Notice that the sensor signals can be naturally correlated and a large number of sensors can be used in an FA system. Limited by the computational complexity, variables substitution is not applicable to determine the anomalous sensors if the number of anomalous sensors is large [12]. Regarding to the Prep-TD-SD, suppose $\tilde{\mathbf{x}}_{\text{T}}^t$ is the standard normalized sequence of $\mathbf{x}_{\text{T}}^t$, where the entries of $\tilde{\mathbf{x}}_{\text{T}}^t$ have zero mean and unit variance, we assume that $\tilde{\mathbf{x}}_{\text{T}}^t$ follows a model of:

$$\tilde{\mathbf{x}}_{\text{T}}^t = \mathbf{\Omega}\mathbf{s}^t + \mathbf{e}^t, \tag{5}$$

where $\mathbf{\Omega}$ is the matrix that correlates the unknown independent and identically distributed random variables $\mathbf{s} \in \mathbb{R}^{N_{\text{T}} \times 1}$ that has $s_i \sim \mathcal{N}(0,1)$. Here, $\mathbf{e} \in \mathbb{R}^{N_{\text{T}} \times 1}$ refers to the error sequence, where the non-zero elements are the error signals of the anomalous sensors. If the FA system works normally, all the entries of $\mathbf{e}$ are zeros. All the three variables at the right-hand side are unknown. Our motivation is to de-correlate the time-delay pre-processed signals and transform it back to the anomaly contribution of each sensor. The Prep-TD-SD is summarized into two phases, the training phase and the testing phase. The algorithm is summarized in Algorithm 1. In the training phase, the whitening matrix is found and stored before the testing phase. In Algorithm 1, we use $\mathbf{p}^t$ to denote the anomaly score in a *pre-processed form*, which has $\mathbf{p}^t = [\mathbf{p}_{\text{F}}^t, \mathbf{p}_{\text{B}}^t, \mathbf{p}_{\text{H}}^t]$. The final anomaly scores of the sensors are denoted by $\mathbf{a}^t$.

Note that to avoid the covariance matrix to be noninvertible, a shrunk covariance estimator for the covariance matrix $\Sigma_{\tilde{\mathbf{x}}_{\text{T}}}$ can be computed. The shrunk covariance is obtained by

$$\Sigma_{\text{shurnk}} = (1-\alpha)\Sigma_{\tilde{\mathbf{x}}_{\text{T}}} + \alpha \frac{\text{tr}\{\Sigma_{\tilde{\mathbf{x}}_{\text{T}}}\}}{N_{\text{T}}}\mathbf{I}, \tag{6}$$

where α is a shrinkage coefficient to balance the bias and variance of the estimation in the covariance matrix. Since the shrunk coavriance matrix is used, the signals after whitening transform are standard normalized, where the standard normalizations are performed in both the training and testing phases.

Following the model assumption in (5), when Λ perfectly de-correlates and whitening transforms $\tilde{\mathbf{x}}_{\text{T}}^t$, the term $\Lambda\Omega\mathbf{s}^t$ can be reduced to $\mathbf{s}^t$, which is considered as N_{T} variables that are normally distributed, and the error term $\Lambda\mathbf{e}^t$ is added.

Algorithm 1 Prep-TD-SD Algorithm

Phase 1 – Training

Step 1: Compute the statistical mean $\boldsymbol{\mu}_{\mathbf{x}_\mathrm{T}} \in \mathbb{R}^{N_\mathrm{T}-1}$ and variance $\boldsymbol{\sigma}^2_{\mathbf{x}_\mathrm{T}} \in \mathbb{R}^{N_\mathrm{T}-1}$ based on the data $\mathbf{x}_\mathrm{T}$.

Step 2: Standard normalize the signals by $\tilde{x}^t_{\mathrm{T},i} = (x^t_{\mathrm{T},i} - \mu_{\mathbf{x}_\mathrm{T},i})/\sigma_{\mathbf{x}_\mathrm{T},i}$ for $i \in \{0, ..., N_\mathrm{T}-1\}$ and compute the statistical covariance matrix $\Sigma_{\tilde{\mathbf{x}}_\mathrm{T}}$.

Step 3: Given the covariance matrix $\Sigma_{\tilde{\mathbf{x}}_\mathrm{T}}$, compute the whitening matrix Λ.

Step 4: Whitening transform $\tilde{\mathbf{x}}^t_\mathrm{T}$ by $\dot{\mathbf{x}}^t_\mathrm{T} = \Lambda\tilde{\mathbf{x}}^t_\mathrm{T}$.

Step 5: Find the mean $\mu_{\dot{\mathbf{x}}_\mathrm{T},i}$ and variance $\sigma^2_{\dot{\mathbf{x}}_\mathrm{T},i}$ of the ith whitened signal $\dot{\mathbf{x}}_{\mathrm{T},i}$, where $i \in \{0, ..., N_\mathrm{T}-1\}$.

Phase 2 – Testing

Step 1: Standard normalize the test data $(\tilde{x}'_{\mathrm{T},i})^t = ((x')^t_{\mathrm{T},i} - \mu_{\mathbf{x}_\mathrm{T},i})/\sigma_{\mathbf{x}_\mathrm{T},i}$.

Step 2: Whitening transform $(\dot{\mathbf{x}}'_\mathrm{T})^t = \Lambda(\tilde{\mathbf{x}}'_\mathrm{T})^t$.

Step 3: Standard normalize $(\dot{\mathbf{x}}'_\mathrm{T})^t$ by $(\acute{x}'_{\mathrm{T},i})^t = \frac{(\dot{x}'_{\mathrm{T},i})^t - \mu_{\dot{\mathbf{x}}_\mathrm{T},i}}{\sigma_{\dot{\mathbf{x}}_\mathrm{T},i}}$ for $i \in \{0, ..., N_\mathrm{T}-1\}$.

Step 4: Find $\mathbf{p}^t$ by $\mathbf{p}^t = \sum_{m=t}^{t+w-1} \left((\acute{\mathbf{x}}_\mathrm{P}')^m\right)^2$, where $\left((\acute{\mathbf{x}}_\mathrm{P}')^m\right)^2$ is the pre-processed form of the anomaly score at the mth time index.

Step 5: Find the summation of the anomaly scores of the analog sensors by $\mathbf{a}^t_\mathrm{A} = [\mathbf{p}^t_\mathrm{B} + \mathbf{p}^t_\mathrm{H}]$. The final anomaly scores of the sensors are returned as $\mathbf{a}^t = [\mathbf{p}^t_\mathrm{F}, \mathbf{a}^t_\mathrm{A}]$.

When the error sequence $\mathbf{e}^t$ is sparse, we expect to accurately detect the anomalous sensors by element-wisely comparing $\mathbf{a}^t$ to a threshold, where the anomalous sensor is detected if $\mathbf{a}^t$ is greater than the threshold. Notice that when the time-delay pre-processed form is not applied and the Mahalanobis whitening matrix is used, the algorithm reduces to the complete decomposition contribution (CDC) to T^2 as discussed in [15].

B. Prep-TDAE-AD

Instead of collecting the information of covariance as in previous section, Prep-TDAE-AD is designed based on the proposed AE structure in Sec. III. The loss function of an AE assumes that the reconstruction errors of the signals are

Algorithm 2 Prep-TDAE-AD Algorithm

Phase 1 – Training

Step 1: Find the residual error $\mathbf{r}^t_\mathrm{T}$ of each element in $\mathbf{x}^t_\mathrm{T}$.

Step 2: Find the mean of each signal $\boldsymbol{\mu}_{\mathbf{r}_\mathrm{T}} \in \mathbb{R}^{N_\mathrm{T}-1}$ and the variance $\boldsymbol{\sigma}^2_{\mathbf{r}} \in \mathbb{R}^{N_\mathrm{T}-1}$.

Phase 2 – Testing

Step 1: Compute the residual errors $(\mathbf{r}'_\mathrm{T})^t$ of the test sequence $(\mathbf{x}'_\mathrm{T})^t$.

Step 2: Standard normalize $(\acute{\mathbf{r}}'_{\mathrm{T},i})^t = \frac{(r'_{\mathrm{T},i})^t - \mu_{\mathbf{r},i}}{\sigma_{\mathbf{r},i}}$, where $(\acute{\mathbf{r}}'_\mathrm{T})^t$ has its *time-delay pre-processed form*: $(\acute{\mathbf{r}}'_\mathrm{T})^t = [(\acute{\mathbf{r}}'_\mathrm{P})^t, ..., (\acute{\mathbf{r}}'_\mathrm{P})^{t+w-1}]$.

Step 3: Find the anomaly score of the *pre-processed form*, where $\mathbf{p}^t = \sum_{m=t}^{t+w-1} \mathrm{abs}\big((\acute{\mathbf{r}_\mathrm{P}}')^m\big)$.

Step 4: Find the summation of the anomaly score of the analog sensors by $\mathbf{a}^t_\mathrm{A} = [\mathbf{p}^t_\mathrm{B} + \mathbf{p}^t_\mathrm{H}]$. The final anomaly scores of the sensors are returned as $\mathbf{a}^t = [\mathbf{p}^t_\mathrm{F}, \mathbf{a}^t_\mathrm{A}]$.

independent as the anomaly detection score is the linear combination over all the signals. Therefore, define $\mathbf{r}^t_\mathrm{T} \in \mathbb{R}^{N_\mathrm{T}}$ as the residual errors of a Prep-TDAE. The residual errors of binary sensors are computed by $r^t_{\mathrm{T},i} = \mathrm{CE}\big(x^t_{\mathrm{T},i}, \hat{x}^t_{\mathrm{T,i}}\big)$, for $i \in \Upsilon(\mathbf{x}^t_\mathrm{T})$ and analog sensor's residual error is computed by $r^t_{\mathrm{T},i} = \frac{1}{2}\big(x^t_{\mathrm{T},i} - \hat{x}^t_{\mathrm{T},i}\big)^2$, for $i \in \Psi(\mathbf{x}^t_\mathrm{T})$. Once all the signals' residual errors are found, by standard normalizing $\mathbf{r}^t_\mathrm{T}$ of the Prep-TDAE, we can assume that the normalized reconstruction errors are samples from an identical and independent Gaussian distribution. Finally, we compute the anomaly contribution of each sensor based on the normalized reconstruction errors.

The main difference between Prep-TDAE-AD and Prep-TD-SD is that Prep-TD-SD performs a linear transformation to find the anomaly contribution of each sensor, whereas Prep-TDAE-AD relies on the target loss function of Prep-TDAE.

In a practical application, a threshold is preset to determine the anomalous sensors. It can be observed that if the residual errors of Prep-TDAE are independent, when an anomaly occurs, a large reconstruction error of the corresponding sensors will be induced. The anomaly contribution of anomalous sensors will exceed the preset threshold and as a result, anomaly is diagnosed.

V. Numerical Results

A. Anomaly Detection

We use a downsampling factor of $f = 10$ to generate the training data and test data. For the TDAE and conventional machine learning algorithms, the time-delay form of data are used as the training and test sequences. We evaluate the performance of anomaly detection methods in terms of area under the curve (AUC) of receiver operating characteristics (ROC). Note that a worse AUC-ROC is obtained if the time-delay form is not performed. The hyper-parameters to train the Prep-TDAE and TDAE are listed in Table I.

TABLE I
Hyper-parameters for the TDAE and Prep-TDAE training

Hyper-parameters	Default value
Time-delay window size w	10
Number of epochs	80
Learning rate	0.5
Hidden layer size	[800, 700]
Momentum factor	0.9
Loss weighting λ	0.5
Mini-batch size	30
Optimizer	Momentum optimizer
Shrunk estimator factor α	0.005
High pass filter cutoff frequency	0.6 (normalized)
Band pass filter cutoff frequencies	[0.25, 0.9] (normalized)

We use normalized anomaly score to show the robustness of our Prep-TDAE anomaly detection algorithm. Fig. 3 shows that even conventional AE detects anomaly at around 180s, where a large value of anomaly score is obtained, but it has false alarms and big end transient. Both TDAE and Prep-TDAE are able to detect anomaly with much smaller end transient. From Fig. 3, we can observe that Prep-TDAE has much more stable and smaller anomaly scores during the normal region compared to the TDAE algorithm. This directly results in a better AUC-ROC obtained by the Prep-TDAE algorithm. The AUC-ROC results of anomaly detection algorithms are plotted in Fig. 4. It can be seen that at a small false alarm probability (false positive rate), Prep-TDAE has the highest true positive rate. We observe that employing the pre-processing is efficient to improve the AUC-ROC of an FA system that contains both binary and analog sensors. Since binary sensors are contained, conventional machine learning anomaly detection algorithms may lack in high moment statistics that result in worse performance than the Prep-TDAE algorithm.

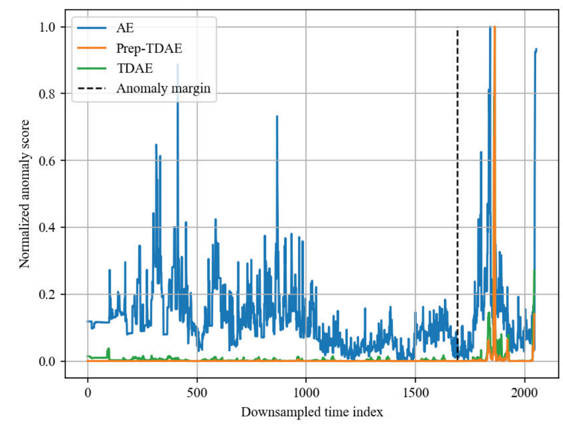

Fig. 3. Normalized anomaly score of TDAE and Prep-TDAE algorithms.

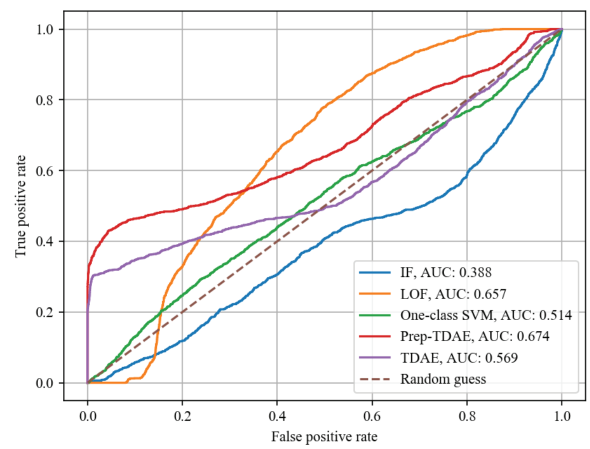

Fig. 4. AUC-ROC of anomaly detection algorithms.

B. Anomaly Diagnosis

Although the anomaly is observed during data collection, the exact sensors that cause the anomaly are unknown. Based on the normal data, in the anomaly diagnosis testing, we assume that an independent error event occurs during the normal operation. The process of testing anomaly diagnosis performance is shown in Fig. 5. We define a *binary error* as an erroneous flip of the sensor's value, i.e., $\{0 \to 1, 1 \to 0\}$. A *non-binary error* is introduced by replacing the original value of the ananlog

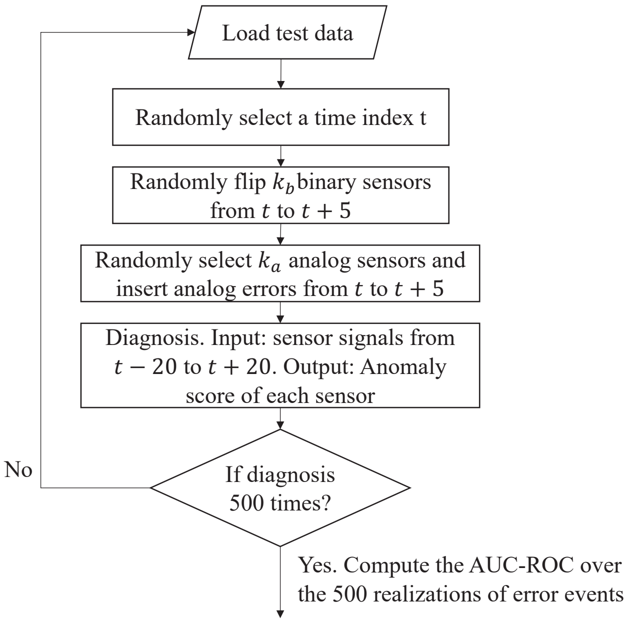

Fig. 5. Flowchart of the tests on anomaly diagnosis.

sensor with an outlier. For example, if jth sensor is assumed to be an anomalous sensor, where $j \in \Psi(\mathbf{x})$, an error e_j is randomly sampled from $\mathcal{U}(\mu_j + 3\sigma_j^2, \mu_j + 5\sigma_j^2)$ or $\mathcal{U}(\mu_j - 3\sigma_j^2, \mu_j - 5\sigma_j^2)$, where μ_j and σ_j^2 are the mean and variance values of the jth sensor, respectively. As shown in Fig. 5, for every error event, we assume that random k_b binary sensors and k_a analog sensors are operated abnormally, where $k_b, k_a \in \{1, ..., 4\}$. The erroneous value is assumed to be held for 5 consecutive time indices and we expect to detect the anomalous sensors by computing the anomaly score from $t - 20$ to $t + 20$. This time range is selected to sufficiently measure the possible false alarm and miss detection probabilities of the anomaly diagnosis. The random selection of the anomalous sensors are tested 500 times. The time index is randomly chosen from 50s to 130s after the start of data collection, which corresponds to the time duration that the machine works normally. The final anomaly score of each sensor is the summation of the anomaly score over the time interval from $t - 20$ to $t + 20$.

For the Prep-TD-SD algorithm, the Mahalanobis whitening matrix is computed by a given statistical covariance matrix $\Sigma_{\tilde{\mathbf{x}}_\mathrm{T}}$. Table II shows that the proposed Prep-TD-SD algorithm outperforms the proposed Prep-TDAE-AD method. TDAE-AD indicates that the residual error of a TDAE is computed without the preprocessing reform. Table II shows that Prep-TDAE has a better AUC-ROC than the TDAE-AD. The CDC to T^2 without the time-delay form has the worst AUC-ROC score, while obtaining the time-delay form helps the diagnosis algorithm explore the time domain correlation and result in a better diagnosis performance.

TABLE II
COMPARISONS OF AUC-ROC OF THE ANOMALY DIAGNOSIS ALGORITHMS

Algorithms	AUC-ROC
Prep-TDAE-AD	0.869
Prep-TD-SD	**0.919**
TDAE-AD	0.848
CDC to T^2, $w = 10$	0.897
CDC to T^2, $w = 1$	0.743

VI. CONCLUSION

In this paper, we have proposed a data pre-processing based TDAE to detect anomaly. Digital filters have been used to process the analog sensor's signals and help the TDAE learn the data features. Moreover, based on the proposed Prep-TDAE algorithm, we have introduced a statistics-based method and an autoencoder-based method for anomaly diagnosis. Numerical results have shown that Prep-TDAE outperforms traditional machine learning detection algorithms and pre-processing is effective for improving the anomaly detection and diagnosis performances.

REFERENCES

[1] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection for discrete sequences: A survey," *IEEE Trans. Knowl Data Eng.*, vol. 24, no. 5, pp. 823–839, May. 2012.

[2] J. Liu, J. Guo, P. Orlik, M. Shibata, D. Nakahara, S. Mii, and M. Takáč, "Anomaly detection in manufacturing systems using structured neural networks," in *13th World Congress Intel. Control Aut. (WCICA)*, 2018, pp. 175–180.

[3] L. Martí, N. Sanchez-Pi, J. M. Molina, and A. C. B. Garcia, "Anomaly detection based on sensor data in petroleum industry applications," *Sensors*, vol. 15, no. 2, pp. 2774–2797, Jan. 2015.

[4] J. Ma and S. Perkins, "Time-series novelty detection using one-class support vector machines," in *IEEE Proc. Int. Joint Conf. Neural Networks*, vol. 3, 2003, pp. 1741–1745.

[5] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation-based anomaly detection," *ACM Trans. Knowl. Discov. Data*, vol. 6, no. 1, Mar. 2012.

[6] M. Amer and M. Goldstein, "Nearest-neighbor and clustering based anomaly detection algorithms for rapidminer," in *Proc. 3rd RapidMiner Community Meeting Conf. (RCOMM)*, S. Fischer and I. Mierswa, Eds. Shaker Verlag GmbH, Aug. 2012, pp. 1–12.

[7] M. Amer, M. Goldstein, and S. Abdennadher, "Enhancing one-class support vector machines for unsupervised anomaly detection," in *Proc. ACM SIGKDD workshop on outlier detection and description*, 2013, pp. 8–15.
[8] S. W. Choi, C. Lee, J.-M. Lee, J. H. Park, and I.-B. Lee, "Fault detection and identification of nonlinear processes based on kernel PCA," *Chemometrics Intel. Lab. Sys.*, vol. 75, no. 1, pp. 55 – 67, Jan. 2005.
[9] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," *CoRR*, vol. abs/1901.03407, 2019.
[10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
[11] T. Tagawa, Y. Tadokoro, and T. Yairi, "Structured denoising autoencoder for fault detection and analysis," in *Proc. Sixth Asian Conf. Machine Learning*, vol. 39. Nha Trang City, Vietnam: PMLR, Nov. 2015, pp. 96–111.
[12] K. Singh, "Anomaly detection and diagnosis in manufacturing systems: A comparative study of statistical, machine learning and deep learning techniques," in *PHM-CONF*, vol. 11, no. 1, Sep. 2019.
[13] Y. Fan, G. Wen, D. Li, S. Qiu, M. D. Levine, and F. Xiao, "Video anomaly detection and localization via Gaussian mixture fully convolutional variational autoencoder," *Comput. Vis. Image Understanding*, vol. 195, p. 102920, Jun. 2020.
[14] Y. Liang, S. Wang, W. Li, and X. Lu, "Data-driven anomaly diagnosis for machining processes," *Engineering*, vol. 5, no. 4, pp. 646 – 652, Aug. 2019.
[15] C. F. Alcala and S. Joe Qin, "Analysis and generalization of fault diagnosis methods for process monitoring," *J. Process Control*, vol. 21, no. 3, pp. 322 – 330, Mar. 2011.
[16] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. J. Lang, "Phoneme recognition using time-delay neural networks," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 3, pp. 328–339, Mar. 1989.
[17] L. D. Paarmann, *Design and Analysis of Analog Filters: A Signal Processing Perspective*. Springer Science & Business Media, 2006, vol. 617.