A Novel Data-Driven Fault Diagnosis Algorithm Using Multivariate Dynamic Time Warping Measure

Jiangyuan Mei, Jian Hou, Hamid Reza Karimi, and Jiarao Huang

1 Department of Computer Science, University of California, Santa Barbara, CA 93106, USA
2 School of Information Science and Technology, Bohai University, No. 19, Keji Road, Jinzhou 121013, China
3 Department of Engineering, Faculty of Engineering and Science, University of Agder, 4898 Grimstad, Norway
4 Research Institute of Intelligent Control and Systems, Harbin Institute of Technology, Harbin 150080, China

Correspondence should be addressed to Jiangyuan Mei; meijiangyuan@gmail.com

Received 31 March 2014; Accepted 1 May 2014; Published 15 May 2014

Copyright © 2014 Jiangyuan Mei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Process monitoring and fault diagnosis (PM-FD) has been an active research field since it plays important roles in many industrial applications. In this paper, we present a novel data-driven fault diagnosis algorithm which is based on the multivariate dynamic time warping measure. First of all, we propose a Mahalanobis distance based dynamic time warping measure which can compute the similarity of multivariate time series (MTS) efficiently and accurately. Then, a PM-FD framework which consists of data preprocessing, metric learning, MTS pieces building, and MTS classification is presented. After that, we conduct experiments on industrial benchmark of Tennessee Eastman (TE) process. The experimental results demonstrate the improved performance of the proposed algorithm when compared with other classical PM-FD classical methods.

1. Introduction

With the development of automation degrees in industrial area, the modern system has become more and more complicated. Process monitoring and fault diagnosis is one of important building blocks of many operation management automation systems. To strengthen safety as well as reliability in the industrial process, PM-FD has been a hot area of research in industrial field in the past decades [1, 2]. Recently, most of the PM-FD methods could be divided into two categories [3]. They are model-based PM-FD approaches and data-driven PM-FD approaches. The model-based PM-FD techniques include observer-based approach [4], parity-space approach [5], and parameter identification based methods [6]. In these approaches, some models of automation systems are built to detect the occurrence of fault. One advantage of these PM-FD approaches is that these models are independent of the application. They can predict as well as diagnose the impacts of faults. However, these system models are always based on human expertise or a priori knowledge [7]. Different from hand-built model, data-driven PM-FD algorithms detect and diagnose fault directly by measuring data. Thus, the data-driven PM-FD approaches catch the attention of numerous scholars in application and research fields [8, 9].

There are many data-driven PM-FD approaches in literatures. One of the most basic and famous methods is principal component analysis (PCA) [10, 11]. PCA is a statistical method which preserves the principal components and reduces the less important dimensions. The principal components are converted from original data using orthogonal transformation. PCA is regarded as a powerful tool. And it is widely used in various practical applications due to its efficiency in processing data with high dimensions. Many PM-FD approaches are derived from the basic PCA algorithm. In [12], the authors propose a modified PCA method which can deliver an optimal fault detection under given confidence level. In [13], a dynamic PCA (DPCA) is proposed to deal with autocorrelation of process variables. Another improved PCA algorithm is named as multiscale PCA (MSPCA) [11]. This method uses wavelets to decompose the individual sensor signals into approximations and details at different scales. Then, a PCA model is constructed to extract correlations
at each scale. Another powerful statistical method in fault
detection and diagnosis is partial least squares (PLS) [14–
16]. The main idea of PLS is to identify a linear correlation
model by utilizing covariance information to project the
observable data as well as predicted data into a new space.
Standard PLS often requires numerous components or latent
variables, which may be useless for prediction. To reduce
false alarm and missing alarm rates, the total projection to
latent structure (TPLS) approach [15] improves the PLS by
further decomposing the results of standard PLS algorithm
on certain subspaces. Another alternative modified version
is the so-called MPLS [16]. This method firstly estimates the
correlation model in the least-square sense and then further
performs an orthogonal decomposition on measurement
space. It is worth noting that PCA, PLS, and their modified
versions are all based on the assumption that the measure-
ment signals follow multivariate Gaussian distribution. When
dealing with the signals with non-Gaussian distribution,
independent component analysis (ICA) [17] is a good choice.
The basic idea of ICA is to decompose the measurement sig-
nal as a linear combination of non-Gaussian variables which
are named as ICs. A modified version (MICA) [18] gives a
unique solution of ICs. At the same time, the MICA algorithm
improves the computation efficiency compared with the
standard ICA. There are also many other PM-FD approaches,
including fisher discriminant analysis (FDA) and subspace
aided approach (SAP). Please refer to [19, 20] for more details.

However, one common point of these methods is that
they all utilize the static features to detect and diagnose
the faults. These static features are extracted from specific
points of measurement signals. Compared with static data,
time series which comprises dynamic features varying with
time can provide more information about the changing of
measurement signals in internals. Thus, some scholars pay
their attention on the fault diagnosis using dynamic analysis.
In [15], the authors review recently the dynamic trend analysis
methods, which represents the measurement signals as the
combination of several basic primitives. This method has
some advantages when compared with traditional static
methods. However, the drawback of this method is also
obvious. In PM-FD problems, more than one measurement
signal is observed using various sensors. It is worth noting
that each fault will reveal different dynamic process on
different sensors. Some sensors may have almost the same
dynamic signals compared with the normal signals while
other sensors may reveal very different signals. Besides,
some sensors suffer from a bad work environment, and
the collected signals may contain lots of noise and outliers.
Therefore, we can conclude that some of variables play
important roles in detecting and diagnosing the fault while
others have weak or no effect. Another problem is that
some industrial processes are very complex and sensors
can not get consistent, repeatable signals, especially for fault
signals. Thus, there may be no one-to-one correspondence
between the test instances and training measurement signals
when measuring their similarity. The dynamic trend analysis
methods are not robust to these disturbance.

In this paper, we use multivariate time series (MTS) to
represent the dynamic features of the measurement signals.
Considering the above problems, this paper proposes a
multivariate dynamic time warping measure which is based
on Mahalanobis distance. The contribution of the paper
can be summarized as follows. (1) We use the proposed
multivariate dynamic time warping measure to compute the
similarity of multivariate signals. The Mahalanobis distance
over the feature space is obtained by using metric learning
algorithm to learn the static feature vectors in measurement
signals. And the obtained Mahalanobis distance is utilized
to compute the distances of local feature vectors in MTS.
After that, the DWT algorithm is applied on these local
distances. The MTS will be aligned with the same phase and
the similarity between two MTS instances can be measured.
(2) The paper applies the proposed multivariate dynamic time
warping measure on PM-FD problem and presents a novel
PM-FD framework, including data preprocessing, metric
learning, MTS pieces building, and MTS classification. (3)
The proposed framework is applied on the benchmark of TE
process. The experimental results demonstrate the improved
performance of the proposed method.

The remainder of this paper is organized as follows. In
Section 2, we illustrate the proposed multivariate dynamic
time warping measure which is based on Mahalanobis
distance. Then, the application of the proposed method on
Tennessee Eastman (TE) process is presented in Section 3.
Section 4 reports the experimental results on the benchmark
of TE process to demonstrate the effectiveness of the pro-
posed algorithm. Finally, we draw conclusions and point out
future directions in Section 5.

2. Multivariate Dynamic Time
Warping Measure

In this section, we mainly presents a novel measure for
dynamic measurement signals. We use MTS $X$ and $Y$ to
represent two measurement signals,

\[
X = [x_1(t) \ x_2(t) \ \ldots \ x_n(t)]^T,
\]

\[
Y = [y_1(t) \ y_2(t) \ \ldots \ y_n(t)]^T,
\]

where $n$ is number of features and $t = 1, 2, \ldots, T$ represents
the time step. Besides, $x_k(t)$ is used to stand for the $k$th feature
(variable) in the measurement signals while $X'$ stands for all
the features in $X$ at the $i$th time step.

In time series analysis, DTW is a common used algo-
rithm measuring similarity between temporal sequences with
different phases and lengths. The main objective of DTW is
to find the optimal alignment by stretching or shrinking the
linearly or nonlinearly warped time series. Using this optimal
alignment, these two time series will be extended to two new
sequences which have one-to-one correspondence. And the
distance between these two extended time sequences is the
minimum distance between the original time series.

It is worth noting that the traditional DTW can only deal
with univariate time series (UTS), which is not appropriate
to deal with PM-FD problems. In literature [21], the authors
proposed a multidimensional DTW algorithm which regards
time points \( X^i \) and \( Y^j \) in MTS \( X \) and \( Y \) as the primitive. Then, the work uses square Euclidean metric to measure the local distance between \( X^i \) and \( Y^j \)

\[
d(X^i, Y^j) = \left( X^i - Y^j \right)^T \left( X^i - Y^j \right). \tag{2}
\]

And these local distances \( d(X^i, Y^j) \) are aligned with traditional DTW algorithm. One weak point of this method is that it assigns the same weight to each variable, which is not practical in PM-FD problems. As mentioned above, different variables play different roles in the fault diagnosis process. Besides, some signals measured by different sensors may be coupled with each other. It is obvious that the Euclidean distance can not measure the difference among these local vectors accurately. Thus, we should find an appropriate similarity metric which can build the relationship between feature space and the labels (normal or abnormal) of measurement signals to measure the divergence among local vectors. In this paper, the metric is selected as Mahalanobis distance which is parameterized by a positive semidefinite (PSD) matrix \( M \). The square Mahalanobis distance between local vectors \( X^i \) and \( Y^j \) is defined as

\[
d_M(X^i, Y^j) = \left( X^i - Y^j \right)^T M \left( X^i - Y^j \right). \tag{3}
\]

In the case that \( M = I \), where \( I \) is a identity matrix, the Mahalanobis distance \( d_M(X^i, Y^j) \) degenerates into the Euclidean distance \( d(X^i, Y^j) \). The main difference between Mahalanobis distance and Euclidean distance is that the Mahalanobis distance takes into account the correlations of the data set and is scale-invariant. If we apply singular value decomposition to the Mahalanobis matrix \( M \), we can get \( M = H \Sigma H^T \), where \( H \) is a unitary matrix which satisfies \( HH^T = I \) and \( \Sigma \) is a diagonal matrix which consists of all the singular values. Therefore, (3) can be rewritten as

\[
d_M(X^i, Y^j) = \left( X^i - Y^j \right)^T H \Sigma H^T \left( X^i - Y^j \right) = \left( H^T X^i - H^T Y^j \right)^T \Sigma \left( H^T X^i - H^T Y^j \right). \tag{4}
\]

From (4), we can see that the Mahalanobis distance has two main functions. The first one is to find the best orthogonal matrix \( H \) to remove the couplings among features and build new features. The second one is to assign weights \( \Sigma \) to the new feature. These two functions enable Mahalanobis distance to measure the distance between instances effectively.

In this paper, we assume that the optimal warp path \( W \) between MTS \( X \) and \( Y \) is expressed as

\[
W = w_1, w_2, \ldots, w_K,
\]

while the \( k \)th element in \( W \) is \( w_k = (i, j) \), which means that the \( i \)th element of \( X \) corresponded to the \( j \)th element of \( Y \). \( K \) represents the length of the path and it is not less than the length of \( X \) or \( Y \) but not greater than their sum. The literature [22] has pointed out that there are two constraints when constructing the warp path \( W \). One constraint is that the warp path \( W \) should contain all indices of both time series. Another constraint is that the warp path \( W \) should be continuous and monotonically increasing. That is to say, the starting point of \( W \) is restricted as \( w_1 = (1, 1) \) while the ending point should be \( w_K = (T, T) \). Meanwhile, the adjacent points \( w_k = (i, j) \) and \( w_{k-1} = (i', j') \) should also satisfy the fact that

\[
i - 1 \leq i' \leq i, \tag{6}
\]

\[
j - 1 \leq j' \leq j.
\]

Thus, there are only three choices for \( w_{k-1} \); they are \((i - 1, j)\), \((i, j - 1)\), and \((i - 1, j - 1)\).

In order to find the minimum distance warp path, we assume that \( Dist(w_k) = Dist(i, j) \) is the minimum warp distance of two new MTS \( X_i \) and \( Y_j \). \( X_i \) represents a sub-MTS which contains the first \( i \) points of the MTS \( X \) while \( Y_j \) contains the first \( j \) points of the MTS \( Y \). And the \( Dist(i, j) \) can be computed using the following equation:

\[
Dist(i, j) = \sum_{l=1}^{k} d_M(X^{w_l(1)}, Y^{w_l(2)}). \tag{7}
\]

In this equation, we know that the \( k \)th point in the \( W \) is \( w_k = (i, j) \). Thus, the above equation can be rewritten as

\[
Dist(i, j) = d_M(X^i, Y^j) + \sum_{l=1}^{k-1} d_M(X^{w_l(1)}, Y^{w_l(2)}) = d_M(X^i, Y^j) + Dist(w_{k-1}). \tag{8}
\]

As mentioned above, there are three choices for \( Dist(w_{k-1}) \), including \( Dist(i - 1, j - 1) \), \( Dist(i - 1, j) \), and \( Dist(i, j - 1) \). At the same time, we require to choose the minimum warp distance for \( X_i \) and \( Y_j \). Therefore, our multivariate DWT algorithm is expressed as

\[
Dist(i, j) = d_M(X^i, Y^j) + \min \left\{ \begin{array}{l}
Dist(i - 1, j - 1) \\
Dist(i - 1, j) \\
Dist(i, j - 1),
\end{array} \right. \tag{9}
\]

where \( Dist(1, 1) = d_M(X^1, Y^1) \).

The proposed multivariate DWT measure has two main advantages when compared with traditional MTS measure. The first one is that all the variables of MTS stretch or shrink along time axis integrally rather than independently. The MTS is treated as a whole and it will not break the relationship among variables. Another advantage is that a good Mahalanobis distance will build an accurate relationship among variables. The important signals will be highlighted while noise and outliers in some variables will be suppressed, which will be a benefit for precise MTS classification. The time complexity of the multivariate DWT measure is \( O(n^2 d^2) \), where \( d \) is the dimension of the feature space and \( n \) is the length of the MTS. The measure is a little time consumption. We can use some fast techniques, including SparseDTW [23] and the FastDTW [22], to accelerate the DWT algorithm.

3. Fault Diagnosis on TE Process

The previous section illustrated the novel multivariate dynamic time warping measure which is based on Mahalanobis distance. This section presents some details on how
to use this measure in fault diagnosis on TE process [25]. TE process is a realistic simulation program of a chemical plant which has been widely studied as a benchmark in many PM-FD methods. In our paper, the data sets of TE process are downloaded from (http://depts.washington.edu/control/LARRY/TE/download.html). In this database, there are 21 faults, named as IDV(1), IDV(2), \ldots, IDV(21) (please refer to [24, 26] for more details). IDV(1–20) are process faults while IDV(21) is another value fault. There are 22 training sets and 22 testing sets in the database, including 21 process measurement signals and one normal data set. Measurement variables or features. Each training data file contains 480 rows which record the 52 measurement signals for 24 operation hours. Meanwhile, the data in each testing data set is collected via 48-hour plant operation time, in which the faults occur at the beginning of the 8th operation hour. That is to say, each testing data file contains 960 rows, while the first 160 rows represent the normal data and the following 800 points are abnormal data.

The proposed fault diagnosis framework consists of 4 steps, that is, data preprocessing, metric learning, MTS pieces building, and MTS classification. First of all, all the measurement signals in training data and testing data are normalized in the data preprocessing. After that, we use the training data to learn a Mahalanobis distance. Then, the MTS pieces are built in both the training set and testing set. Finally, the testing MTS pieces are compared with training MTS pieces and we use KNN algorithm to determine the label (normal or abnormal) of the testing MTS pieces.

In this paper, we assume that the normal measurement signals follow multivariate Gaussian distribution. In fault signals, some data but not all data are out of the confidence interval. The normal data is always in the obviously centralized distribution while abnormal data is lowly concentrated, even dispersive. There are two main problems when measuring the similarity between fault signals. The first one is that the amplitude between training data and testing data might be different. Figures 1(a) and 1(b) show the 46th measurement signals of IDV(0) and IDV(13). When considering the IDV(13), we can see that the amplitude of the testing data is obviously larger than that of training data. The second one is that the difference between abnormal signals may be larger than the difference between abnormal signals and normal signals. From Figures 1(e) and 1(f), we can see that some data of 19th measurement signal in the training data of IDV(21) is out of the upper limit of the confidence interval while some data in the testing data is out of the lower limit of confidence interval. This will result in two much false fault missing results. Therefore, an appropriate data preprocessing method should deal with these two problems at the same time. In our method, we apply the Gaussian kernel to normalize the original measurement signals. Consider

\[
\overline{x}_i(t) = \exp \left( -\frac{(x_i(t) - \mu_i)^2}{2\sigma_i^2} \right),
\]

where \(x_i(t)\) represent the \(i\)th original measurement signal and \(\overline{x}_i(t)\) is the normalized measurement signal. In this equation, the parameters \(\mu_i\) and \(\sigma_i\) are the mean value and variance evaluated using the \(i\)th normal measurement signal. Figures 1(c), 1(d), 1(g), and 1(h), respectively, show the normalized measurement signals of normal data set and fault data set. We can see that the normal signals concentrate nearly 1 while some data in fault signals is much less than 1. Besides, all large amplitude in the original fault signals will map to small value nearby 0 in the normalized signals and the difference between testing data and training data is very small. At the same time, the fault data which are out of the upper limit or lower limit of the confidence interval are almost the same in the normalized signals, which will be easy to measure their similarity.

In our proposed multivariate dynamic time warping measure, how to learn an appropriate Mahalanobis distance is another key problem in our framework. Metric learning is a popular approach to accomplish such a learning process. In our previous work [27], we have proposed a Logdet divergence based metric learning (LDML) method to learn such a Mahalanobis distance function. After data processing, we use the normalized points in measurement signals to build the triplet labels \(\{x^i, x^j, x^k\}\), which means that \(x^i\) and \(x^j\) are in the same category while \(x^k\) is in another category. The objective of metric learning is to ensure that most of the triplet labels \(\{x^i, x^j, x^k\}\) satisfy the fact that

\[
d_M(x^i, x^j) - d_M(x^j, x^k) \leq -\rho
\]

while \(\rho > 0\). The updating formulation of the LDML algorithm is expressed as

\[
\begin{align*}
\Gamma &= M_t - \frac{\alpha M_t (x^i - x^j)(x^i - x^j)^T}{1 + \alpha (x^i - x^j)(x^i - x^j)^T M_t}, \\
M_{t+1} &= M_t + \frac{\alpha M_t (x^j - x^k)(x^j - x^k)^T}{1 - \alpha (x^j - x^k)(x^j - x^k)^T M_t}.
\end{align*}
\]

while \(M_0 = I\) and \(\alpha\) is the learning rate parameter (please refer to [27] for more details). This algorithm can accurately and robustly obtain the Mahalanobis distance which can distinguish fault and normal signals.

From Figures 1(c) and 1(d), we can see that the integral trends of fault signal in training data and testing data do not match exactly. However, some pieces in the testing data can be found in the training data. Therefore, we use MTS pieces instead of the whole measurement signals in the following classification process. Each MTS piece contains \(a\) continuous points in the training data or testing data. In the experiment, we have proven that when \(a\) is selected as 12–16, the proposed method will have the best performance. We measure the similarity between testing MTS pieces and all the training MTS pieces, including the normal data and fault data. Then, we use the KNN algorithm to classify MTS pieces as the normal ones or fault ones.
4. Experiments Results

In this section, we conduct experiments on the TE process to illustrate the performance of the proposed fault diagnosis framework. In the following experiments, all experiments are tested in MATLAB 2011b, and all tests are implemented on a computer with Intel (R) Core (TM) i3-3120 M, 2.50 GHz CPU, 4 G RAM, and Windows 7 operating system. In this paper, we use two general indices to test the performance to evaluate PM-FD performance. They are fault detection rate (FDR) and false alarm rate (FAR). Assume that \( t_p \), \( t_n \), \( f_p \), and \( f_n \) respectively, represent correct fault detection results, correct normal signal detection results, false fault alarm results, and false fault missing results. The definitions of these FDR and FAR are given as

\[
\text{FDR} = \frac{t_p}{t_p + f_n} \times 100\%,
\]

\[
\text{FAR} = \frac{f_p}{f_p + t_n} \times 100\%.
\] (13)
Another important performance index is chosen as equal error rate (EER), where the FAR and false fault missing rate \((1 - FDR)\) are equal to each other. As mentioned above, in our algorithm, the PM-FD process is regarded as a classification problem. In the KNN classification algorithm, different \(k\) will produce different FDR and FAR. In our experiment, the performance is chosen as the FDR and FAR which is the closest to the EER line.

The first experiment is to illustrate the relationship between the fault diagnosis performance and the length of the MTS pieces. We, respectively, use MTS pieces with 10, 12, 16, and 20 points to separately detect the IDV(11) and IDV(21). The experimental results with different \(k\) are illustrated in Figure 2. From the results, we can see that the FDR rises when the length of MTS pieces increases from 10 points to 16 points. However, when the length of MTS pieces rises to 20 points, the performance will degrade. The reason for the phenomena can be explained as follows. When the length of MTS pieces is too short, these MTS contain too little information to record the trends and variation of the measurement signals completely. The comparison of MTS will be inaccurate and the performance will degrade. In an extreme case, the length of MTS pieces is 1 and the algorithm degrades to the traditional methods which are based on measuring the similarity of static feature points. If the length of MTS pieces is too long, it will tend to measure the integral similarity of signals. As mentioned above, in some measurement signals, the integral trends can not match each other, but some pieces in the testing data are similar to those in the training data. Thus, the performance will also have a risk of declining. The experimental results suggest that the best length of the MTS pieces is 12–16.

In the second experiment, we compare the proposed method with many other classical PM-FD methods, including PCA, DPCA, ICA, MICA, FDA, PLS, TPLS, MPLS, and SAP. The results of the proposed method are average values over 5 runs while the results of other classical PM-FD methods are reported by literature [26]. The testing results of various methods for all data sets are summarized in Table 1. The first 21 rows record the FDR for all the faults. Inspired by the literature [26], we also classified 21 faults as three categories. The first category consists of IDV(1-2), IDV(4–8), IDV(12–14), and IDV(17-18). These faults can be easily detected by all the methods. Almost all the methods have high FDR as well as low FAR in the fault diagnosing process. The faults in the second category, including IDV(10-11), IDV(16), IDV(19) and IDV(20-21), are not easy to detect. And the performance of methods on these faults will have obvious difference. In the last category, all methods perform bad results because these faults are very hard to detect. The last row illustrates the average FAR by detecting faults on the fault-free signals of IDV(0). From the results, we can see that the proposed method has good performance on most of the faults in the first and third categories. At the same time, the performance of the proposed method is not bad on the rest of faults, such as IDV(10–12), IDV(16), and IDV(20). Our method does not perform well on IDV(5) and IDV(19). The reasons can be explained as follows. The measurement signals of IDV(5) have obvious difference when the fault occurs. However, the IDV(5) and normal signals have almost the same trend after a certain time, which can
Table 1: FDRs (%) and FARs (%) comparison with classical PM-FD methods on TE datasets given in [24].

<table>
<thead>
<tr>
<th>Fault (free)</th>
<th>PCA</th>
<th>DPCA</th>
<th>ICA</th>
<th>MICA</th>
<th>FDA</th>
<th>PLS</th>
<th>TPLS</th>
<th>MPLS</th>
<th>SAP</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDV(1)</td>
<td>99.88</td>
<td>99.88</td>
<td>100</td>
<td>99.88</td>
<td>100</td>
<td>99.88</td>
<td>99.88</td>
<td>100</td>
<td>99.63</td>
<td>100</td>
</tr>
<tr>
<td>IDV(2)</td>
<td>98.75</td>
<td>99.38</td>
<td>98.25</td>
<td>98.25</td>
<td>98.75</td>
<td>98.63</td>
<td>98.88</td>
<td>98.88</td>
<td>98.88</td>
<td>97.88</td>
</tr>
<tr>
<td>IDV(4)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>87.63</td>
<td>100</td>
<td>99.50</td>
<td>100</td>
<td>100</td>
<td>99.88</td>
<td>100</td>
</tr>
<tr>
<td>IDV(5)</td>
<td>33.63</td>
<td>43.25</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>87.63</td>
<td>100</td>
<td>100</td>
<td>99.88</td>
<td>100</td>
</tr>
<tr>
<td>IDV(6)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>IDV(7)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>IDV(8)</td>
<td>95.38</td>
<td>95.38</td>
<td>95.25</td>
<td>95.00</td>
<td>95.63</td>
<td>95.25</td>
<td>95.50</td>
<td>95.50</td>
<td>95.50</td>
<td>95.50</td>
</tr>
</tbody>
</table>

Not detected by the proposed method. On the contrary, the measurement signals of IDV(19) are similar to normal signals at the beginning and then the difference became noticeable after a period of running time. Our method is not good at detecting this kind of fault too.

5. Conclusion

In this paper, we propose a novel framework for data-driven PM-FD problems. One contribution of this paper is that it firstly uses MTS pieces to represent the measurement signals as dynamic features in the measurement signals can provide more information than static features. Besides, this paper uses a Mahalanobis distance based DTW algorithm to measure the difference between two MTS. The method can build an accurate relationship between the variables and the labels of the signals, which is a benefit for the classification of measurement signals. Furthermore, we also present a new PM-FD framework which is based on Mahalanobis distance based DTW algorithm. The proposed algorithm is shown to be precise by experiments on benchmark data sets and comparison with classical PM-FD methods. One drawback of this algorithm is the heavy time consumption in measuring the similarity of MTS. In future, we will concentrate on further optimization of the proposed method with respect to computation efficiency and other issues.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The research leading to these results has received funding from the Polish-Norwegian Research Programme operated by the National Centre for Research and 24 Development under the Norwegian Financial Mechanism 2009–2014 in the frame of Project Contract no. Pol-Nor/200957/47/2013.

References


Submit your manuscripts at
http://www.hindawi.com