Satellite Telemetry Data Anomaly Detection with Hybrid Similarity Measures

Datong Liu, Jingyue Pang,
Department of Automatic Test and Control
Harbin Institute of Technology
Harbin, Heilongjiang, China,
liudatong@hit.edu.cn

Ben Xu, Zan Liu, Jun Zhou, Guoyong Zhang
Shanghai Institute of Satellite Engineering,
Shanghai, China

***Abstract*—Anomaly detection based on telemetry data can improve the operating safety for spacecrafts. Most of the anomaly detection methods in this domain are based on Euclidean distance for similarity measure of monitoring parameters. However, the Euclidean distance has many limitations on telemetry data similarity measure and may affect the detecting performance. Therefore, improved distance measures and combined distance measures are applied in telemetry data analysis. An improved anomaly detection framework with different similarity measures are presented for multiple monitoring parameters of satellite in this paper. Then, the proposed anomaly detection approach based on the k-Nearest Neighbor (KNN) classification with improved similarity measures are applied into the actual satellite telemetry data. Experimental results show that the presented anomaly detection method can achieve satisfied performance on the actual satellite telemetry data sets.**

Keywords-satellite telemetry data; anomaly detection; Mahalanobis distance; Dynamic Time Warping; KNN

I. INTRODUCTION

Nowadays, analyzing telemetry data is an effective way to understand the operating condition and health status of the spacecraft in orbit. Especially, the data mining, refer to automatic analysis, anomaly detection, diagnosis and prognosis has important significance to the satellite reliability and safety [1]. Furthermore, the monitoring data, which reflects the important condition of satellite, should be monitored and analyzed to detect the anomalies [2][3]. The anomaly detection can help carry out fault diagnosis and prognosis and prevent the occurrence of potential failures and other operating problems.

Distance measure is used to measure the similarity of the monitoring parameters, which is the elementary step in data analysis and data-driven anomaly detection [4]. The suitable similarity measure can effectively reflect the gradual and tiny change of the monitoring series, and thus, can improve the performance of abnormal detection.

The complexity of the telemetry series (i.e., noise, drift, pseudo-period, sparse) makes the similarity measure more important for data analysis. The traditional similarity measures include Minkowski distance [5], normalized inner product, Mahalanobis distance [6], Pearson correlation coefficient, etc. For the growing of complexity of data mining, a lot of novel similarity measures are proposed, such as Dynamic Time Warping (DTW) [7][8], Piecewise Linear Representation (PLR), Symbolic distance, transformation based pattern distance, and other measures.

In this work, we apply the novel similarity measures to fully represent the satellite telemetry parameters. The comparison and the comprehensive evaluation are involved to find the most suitable distance measure to improve the anomaly detection on multiple monitoring parameters.

II. METHODOLOGY

A. Mahalanobis Distance

Mahalanobis distance is used to represent the convariance distance of the data. Given $X_i = (x_{i1}, x_{i2}, ..., x_{in})^T$ and $X_j = (x_{j1}, x_{j2}, ..., x_{jn})^T$, the Mahalanobis distance D_M between X_i and X_j is defined by

$$D_M(X_i, X_j) = \sqrt{(X_i - X_j)^T C^{-1} (X_i - X_j)} \tag{1}$$

where C denotes the covariance matrix.

$$\begin{aligned}
C &= E\left\{\left[X_i - E(X_i)\right]\left[X_j - E(X_j)\right]^T\right\} \\
&= E\begin{bmatrix} \left[X_{i1} - E(X_i)\right] \\ \left[X_{i2} - E(X_i)\right] \\ \vdots \\ \left[X_{in} - E(X_i)\right] \end{bmatrix}\left[\left[X_{j1} - E(X_j)\right]\left[X_{j2} - X_{j2}\right]E(X_j)\cdots\left[X_{jn} - E(X_j)\right]\right] \\
&= \begin{bmatrix} E\left[X_{i1} - E(X_i)\right]\left[X_{i1} - E(X_j)\right] & \cdots & E\left[X_{i1} - E(X_i)\right]\left[X_{in} - E(X_j)\right] \\ E\left[X_{i2} - E(X_i)\right]\left[X_{i1} - E(X_j)\right] & \cdots & E\left[X_{i2} - E(X_i)\right]\left[X_{in} - E(X_j)\right] \\ \vdots & \ddots & \vdots \\ E\left[X_{in} - E(X_i)\right]\left[X_{i1} - E(X_j)\right] & \cdots & E\left[X_{in} - E(X_i)\right]\left[X_{in} - E(X_j)\right] \end{bmatrix} \\
&= \begin{bmatrix} \delta_{11}^2 & \delta_{12}^2 & \cdots & \delta_{1n}^2 \\ \delta_{21}^2 & \delta_{22}^2 & \cdots & \delta_{2n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{n1}^2 & \delta_{n2}^2 & \cdots & \delta_{nn}^2 \end{bmatrix}
\end{aligned} \tag{2}$$

978-1-5090-4020-9/17 $31.00 © 2017 IEEE
DOI 10.1109/SDPC.2017.116

When the covariance matrix is unknown, we can use Eq. (2) to calculate, where δ_{st}^2 refers to the covariance of the s th component in X_i and t th component in X_j. The larger C means higher correlation between two variables.

Mahalanobis distance introduces the covariance matrix, and considers the correlation between the components of the two variables. Thus, the correlation between samples can be excluded. Mahalanobis distance reduces the effects of the different components that contains the same feature [6].

B. DTW Distance

DTW searches the optimized path based on dynamic time warping method, and can simultaneously measure the series of both equal length and unequal length. Moreover, it can keep relatively stable to the discontinuity and achieve the sequence of asynchronous measure.

DTW can catch the time series pattern better by bending and warping the time axis dynamically. $DTW(Q,C)$, the DTW distance, is the minimum distance measure between two time series: $Q=\{q_1,q_2,\cdots,q_m\}$ and $C=\{c_1,c_2,\cdots,c_n\}$. Warping path [9] can be described as $P=\{p_1,p_2,\cdots,p_k\}$, where p_k indicates the corresponding relationship between q_{ik} and c_{jk}. $d(p_k)$ is the warping cost of q_{ik} and c_{jk}, and usually $d(p_k)=d(i_k,j_k)=(q_{ik}-c_{jk})^2$ for the DTW based on Euclidean distance, where $i_k=1,2,\cdots,m$, and $j_k=1,2,\cdots,n$. The $DTW(Q,C)$ is,

$$DTW(Q,C)=\min_{p}\sum_{k=1}^{K}d(p_k) \tag{3}$$

The above equation can be solved by constructing cost matrix R shown as follows.

$$R(i,j)=d(i,j)+\min\{R(i,j-1),R(i-1,j-1),R(i-1,j)\} \tag{4}$$

where $i=1,2,\cdots,m$, and $j=1,2,\cdots,n$, and $R(0,0)=0$, and $R(i,0)=R(0,j)=+\infty$. $R(m,n)$ indicates the minimum distance of time series Q and C, that is $DTW(Q,C)=R(m,n)$.

C. KNN Classification

KNN classification is the most commonly used algorithm in data mining. It labels a sample by the majority vote in the class labels of its K nearest neighboring training samples. Since the classification process only involves K nearest neighboring samples, thus, KNN can be better applied in applications which the class distribution has the overlapped characteristics [10][11].

The KNN classification works as follows.

1) Choose training and test samples, recording the class label of the training samples, setting the parameter K;
2) Compute the distance between the test samples and every training sample according to the chosen distance metric;
3) Find the K nearest neighboring training samples;
4) Label the test sample by the majority vote of the class labels for its K nearest neighboring.

The traditional distance metric in KNN classification is the Euclidean distance. However, other distance metrics (e.g., Hamming distance and Mahalanobis distance) can also be used according to the features of the data samples.

D. Anomaly Detection with KNN Classification

The basic principle to detect anomaly by classification algorithm is as follows. Firstly, the time series containing anomaly is labeled to a class with a classification algorithm. Then, the matching degree between the time series and the class is evaluated. There are many criteria to measure this matching degree. In this research, we evaluate the matching degree by comparing the minimum distance between time series and samples in the class with the average distance between samples inside the class. At last, the anomaly detection result is given according to the matching degree.

Given training data set $X=\{x_1,x_2,\cdots,x_n\}$ and corresponding labels $L=\{l_1,l_2,\cdots,l_n\}$, x_i is the i_{th} element of X and its class label is l_i. $X_j=\{x_{j1},x_{j2},\cdots,x_{jn}\}$ is all the elements in class j, and the average distance $\overline{s}_j$ inside class j can be got from Eq. (5).

$$\overline{s}_j=\begin{cases}\dfrac{2}{n_j(n_j-1)}\sum\limits_{x_{jk}\in X_j}\sum\limits_{x_{jt}\in X_j}dist(x_{jk},x_{jt}), & n_j>1\\ \min(\overline{S}) & ,n_j=1\end{cases} \tag{5}$$

where n_j is the number of samples in class j, x_{jk} is the k_{th} sample in class j, x_{jt} is the t_{th} sample in class j, and $dist(x_{jk},x_{jt})$ is the distance between x_{jk} and x_{jt}. $\overline{S}$ is the series of average distance inside each class, and $\overline{S}=\{\overline{s}_1,\overline{s}_2,\cdots,\overline{s}_c\}$ with c as the total class numbers.

The time series to be detected is x', and D is the distance series between x' and X. If x' is classified to class l', then the minimum distance between x' and samples in class l' is $d_{\min}$. P is defined as a sensitive parameter for anomaly detection.

With above definition, the anomaly detection framework with KNN classification is show in Figure 1. The detailed steps are as follows.

1) Choosing the time series x' to be detected and labeled time series X, setting the parameters K and P;
2) Calculating $\overline{S}$ which is the average distance between samples inside each class in X;

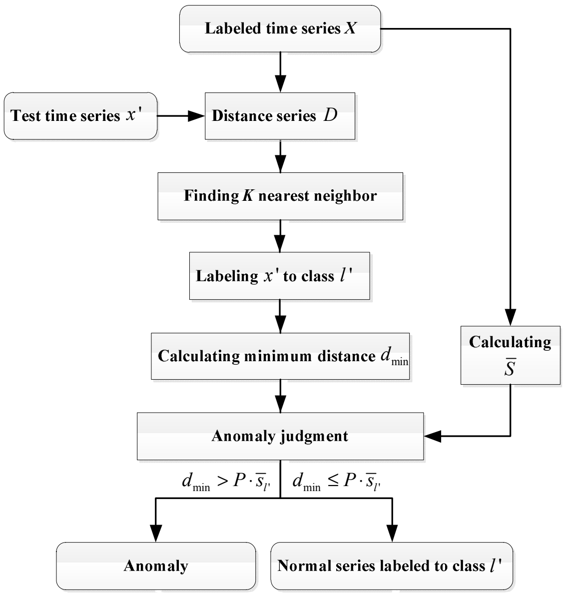

Figure 1. Anomaly detection framework with KNN classification

3) Calculating D which is the distance series between x' and X;
4) Finding K nearest neighbor sample of x' by choosing K minimum value in D;
5) Labeling x' to class l' by majority vote in the class labels of samples found in Step (4);
6) Computing the $d_{\min}$ which is the minimum distance between x' and samples in class l';
7) Anomaly judgment. If $d_{\min} > P \cdot \overline{s}_{l'}$, then x' is containing anomaly. And vise versus.

III. Proposed Framework for Anomaly Detection of Multiple Telemetry Data

By analyzing the parameters of telemetry data, we can find that high correlation relationship exists among the parameters. Due to the limitation on the correlation for the Euclidean distance, the Mahalanobis distance can be used considering the correlation relationship. Moreover, we combine the Mahalanobis distance and DTW to define the DTW distance based on Mahalanobis distance. From the perspective of the principle, this improved distance measure can eliminate the correlation between each dimension, and can realize the asynchronous measure, so it is also applied in the anomaly detection. As a result, Mahalanobis distance, DTW distance and DTW distance based on Mahalanobis distance are all used for verification and evaluation.

The framework of anomaly detection based on KNN and improved distance measure is shown in Fig. 2. Firstly, the subsequences of satellite historical telemetry data is acquired by the segmentation according to the periodic characteristic and argument of perihelion. Then, hierarchical clustering is conducted on the unlabeled satellite telemetry time series to realize anomaly detection.

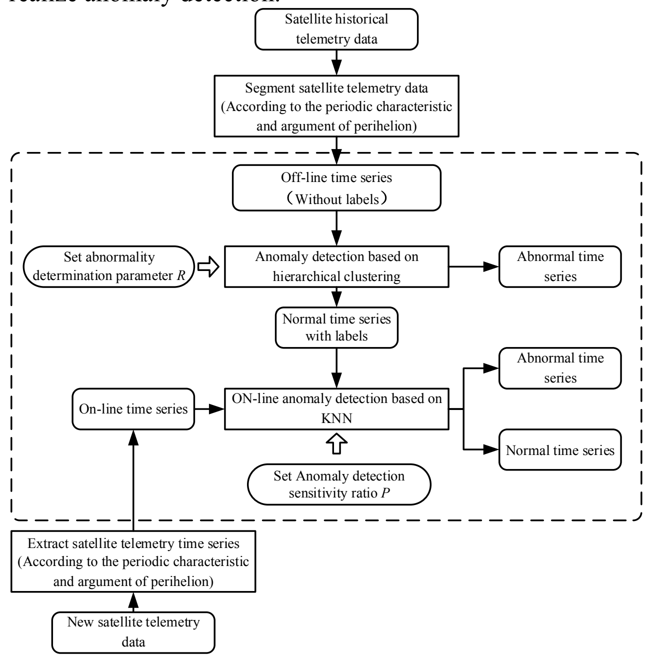

Figure 2. Framework of anomaly detection for telemetry data

As an offline operating mode, if the anomalous data exist in the data of hierarchical clustering [12], the clustering results will contain the class with few members. In this case, in order to achieve anomaly detection, we can introduce the abnormal determination parameter R to determine the clusters. If $R > \frac{n_i}{N}$, the ith class is anomalous, where n_i indicates the number of members in ith class, and N is number of all members. We may delete the anomalous class to extract the normal patterns.

It is noted that with this labeled clusters, we use classification method -KNN, to realize on-line anomaly detection. In the framework, the improved distance measure is more effective for the distance calculation of the satellite telemetry data.

IV. Experimental Results

A. Experiments on Open Source Data Sets

In the experiments, we first use the open-source data sets that are close to the actual satellite data. The Wafer data sets by Carnegie Mellon University (CMU) and ECG data sets, RobotFailure data sets by University of California at Irvine (UCI) are applied [13][14] which are described in Table I.

The experimental results are shown in Tables II and III. Here the result of the RobotFailure is the average value. Experimental results of DTW distance and DWT distance based on the included angle under the boundary conditions are shown in Tables IV and V.

TABLE I. OPEN SOURCE DATA SETS DESCRIPTION

Data sets		Length of series	Dim ensio ns	No. of classes	No. of training samples	No. of test samples
ECG		54	2	2	100	100
Robot Failure	LP1	15	6	4	38	50
	LP2	15	6	5	17	30
	LP3	15	6	4	17	30
	LP4	15	6	3	42	75
	LP5	15	6	5	64	100
Wafer		114	6	2	298	896

TABLE II. ANOMALY DETECTION ACCURACY

Data sets	Euclidean	Mahalanobis	DTW	DTW based on Mahalanobis
ECG	82.00%	83.00%	79.00%	82.00%
RobotFailure	68.77%	59.65%	69.47%	63.86%
Wafer	88.50%	96.32%	98.66%	97.88%

TABLE III. OPERATING EFFICIENCY COMPARISON (TIME: S)

Data sets	Euclidean	Mahalanobis	DTW	DTW based on Mahalanobis
ECG	0.9986	6.5923	52.6713	232.0522
RobotFailure	0.5262	1.3307	5.1446	22.0741
Wafer	12.2632	265.1067	6196.4017	25686.8514

TABLE IV. EXPERIMENTAL RESULTS OF DTW DISTANCE AND DWT DISTANCE BASED ON THE INCLUDED ANGLE UNDER BOUNDARY CONDITIONS

Data sets	DTW distance under the boundary conditions		DWT distance based on the included angle under the boundary conditions	
	Accuracy	*w*	*Accuracy*	*w*
ECG	82.00%	1	83.00%	1
RobotFailure	**70.18%**	4	63.86%	5
Wafer	**98.88%**	7	98.10%	5

TABLE V. EXPERIMENTAL RESULTS OF DTW DISTANCE AND DWT DISTANCE BASED ON THE INCLUDED ANGLE UNDER BOUNDARY CONDITIONS

Data sets	DTW distance under the boundary conditions		DWT distance based on the included angle under the boundary conditions	
	Operating time	*w*	*Operating time*	*w*
ECG	4.5807	1	19.0010	1
RobotFailure	2.7422	4	13.2860	5
Wafer	823.9824	7	2431.2206	5

Among the three distance measures, Mahalanobis distance and DTW distance with boundary conditions show good performance. The DTW distance based on Mahalanobis distance with boundary conditions does not indicate better performance as expected. The reason may be that the multiple time series do not obey the strict Gaussian distribution, thus the covariance matrix does not fully eliminate the correlation relationship between the variables.

B. Experiments on Satellite Telemetry Data Set

The actual satellite telemetry data is applied to further evaluate and verify the proposed method. Considering that the actual satellite data generally contains fewer anomalous samples, we inject the anomalous samples into the actual telemetry data.

Four types of anomalous from the engineering experience, refer to Pulse anomaly, Step anomaly, Graded anomaly and Periodic anomaly, are focused here. The anomalous definitions are shown in TABLE VI.

TABLE VI. MATHEMATIC MODEL OF ANOMALY TAPE

Anomalous type	Mathematic Model
Pulse anomaly	$F_1(\tau)=\Phi(\tau)+A\cdot\sigma(t)$
Step anomaly	$F_2(\tau)=\Phi(\tau)+A\cdot\delta(t)$
Graded anomaly	$F_3(\tau)=\Phi(\tau)+(A\cdot t+B)\cdot\delta(t)$
Periodic anomaly	$F_4(\tau)=\Phi(\tau)+A\cdot\cos(\omega t+\varphi)+B$

where $\Phi(\tau)$ is the normal value, and $\sigma(t)$ is the random pulse function, $\sigma(t)$ is the random step function, and A is the amplitude, ω is angular velocity, φ is bias of angular velocity, and B is bias of amplitude.

In the experiment, for pulse anomaly and step anomaly, A equals to ΔY , 0.5 ΔY , 0.25 ΔY , - ΔY , -0.5 ΔY , -0.25 ΔY respectively, and t equals to 0.25 T , 0.5 T , 0.75 T , respectively. ΔY is the difference of maximum and minimum test parameters, and T is the length of time series. Therefore, 18 types of pulse anomalies and step anomalies are generated for evaluation respectively.

For graded anomaly, A equals to $\Delta Y/T$, 0.5 $\Delta Y/T$, 0.25 $\Delta Y/T$, - $\Delta Y/T$, -0.5 $\Delta Y/T$, -0.25 $\Delta Y/T$, and t equals to 0, 0.5 T , and B equals to 0, respectively. Thus, we can generate 12 types of graded anomalies. For periodic anomaly, A equals to ΔY , 0.5 ΔY , 0.25 ΔY , - ΔY , -0.5 ΔY , -0.25 ΔY , and ω equals to $4\pi/T$, $2\pi/T$, π/T , and φ equals to $\pi/2$, respectively. Thus, we can generate 18 types of periodic anomalies.

The samples of anomalies are shown in Fig. 3 to Fig. 6.

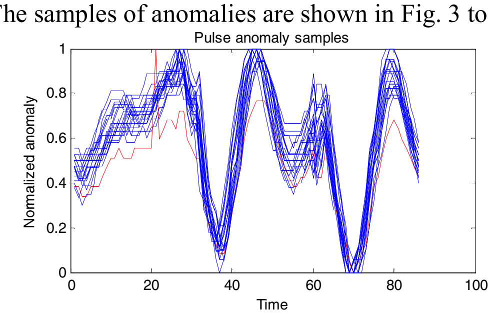

Figure 3. Anomalous samples with pulse anomaly injection for certain angle parameter with $A=0.5\Delta Y$ and $t=0.25T$

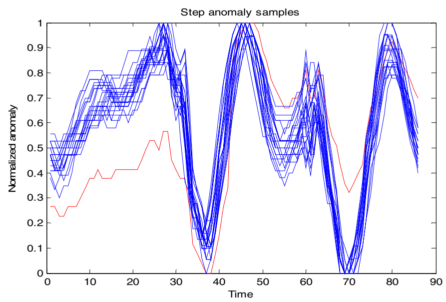

Figure 4. Anomalous samples with step anomaly injection for certain angle parameter with $A = 0.5\Delta Y$ and $t = 0.25T$

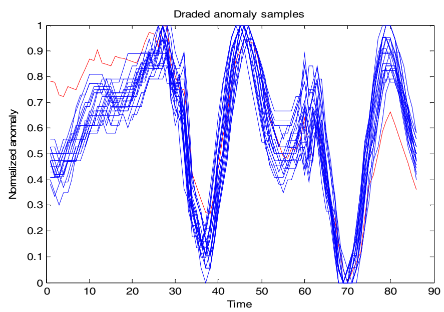

Figure 5. Anomalous samples with graded anomaly injection for certain angle parameter with $A = -0.5\Delta Y/T$ and $t = 0$

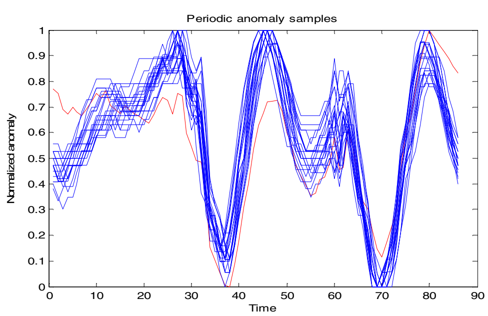

Figure 6. Anomalous samples with periodic anomaly injection for certain angle parameter with $A = 0.5\Delta Y$ and $\omega = 2\pi/T$

Respectively, the Mahalanobis distance and the DTW distance under boundary conditions are applied into hierarchical clustering and KNN classification to achieve the multiple satellite telemetry parameters anomaly detection, in which the boundary conditions of the DTW distance is set to 10. Three typical satellite parameters are applied for experiments, and the dimension of the parameters is 3. Angle test series: the number of off-line samples equals to 50, and the cluster equals to 3, and the abnormal determination parameter R equals to 0.03. The on-line samples are 25, and parameter K is set to 3, and parameter P is set to 1.

The other two parameters, referring to rotation speed test series and current test series, are set similarly.

In the experiments, the missed ratio and false ratio are applied to measure the detection performance. High performance of the anomaly detection means the lower missed ratio and the lower false ratio [15][16]. The detection results are shown in Table VII, VIII, and IX.

TABLE VII. ANOMALY DETECTION RESULTS FOR ANGLE PARAMETER

Metrics	Distance measures	Anomaly modes			
		Pulse	***Step***	***Graded***	***Periodic***
Missed ratio	Euclidean distance	83.33 %	44.44 %	83.33 %	44.44%
	Mahalanobis distance	83.33 %	38.89 %	66.67 %	33.33%
	DTW distance	50.00 %	11.11 %	50.00 %	11.11%
False ratio	Euclidean distance	0.00%	0.00%	0.00%	0.00%
	Mahalanobis distance	0.00%	0.00%	0.00%	0.00%
	DTW distance	0.00%	0.00%	0.00%	0.00%
Detection performance	Euclidean distance	16.67 %	55.56 %	16.67 %	55.56%
	Mahalanobis distance	16.67 %	61.11 %	33.33 %	66.67%
	DTW distance	50.00 %	88.89 %	50.00 %	88.89%

TABLE VIII. ANOMALY DETECTION RESULTS FOR ROTATION SPEED PARAMETER

Metrics	Distance measures	Anomaly modes			
		Pulse	***Step***	***Graded***	***Periodic***
Missed ratio	Euclidean distance	72.22 %	55.56 %	83.33 %	33.33%
	Mahalanobis distance	72.22 %	50.00 %	83.33 %	33.33%
	DTW distance	55.56 %	22.22 %	66.67 %	22.22%
False ratio	Euclidean distance	0.00%	0.00%	0.00%	0.00%
	Mahalanobis distance	0.00%	0.00%	0.00%	0.00%
	DTW distance	0.00%	0.00%	0.00%	0.00%
Detection performance	Euclidean distance	27.78 %	44.44 %	16.67 %	66.67%
	Mahalanobis distance	27.78 %	50.00 %	16.67 %	66.67%
	DTW distance	44.44 %	77.78 %	33.33 %	77.78%

Due to that we cannot use individual missed ratio or false ratio to evaluate the anomaly detection, we define the anomaly detection quality $Q = (1 - missed\ ratio)/(1 + false\ ratio)$ to represent the detection performance. Q locates between 0 and 1, and the larger of Q, the higher performance of the detection. Especially, when Q=1 , it achieves the ideal anomaly detection.

TABLE IX. ANOMALY DETECTION RESULTS FOR CURRENT PARAMETER

Metrics	Distance measures	Anomaly modes			
		Pulse	***Step***	***Graded***	***Periodic***
Missed ratio	Euclidean distance	88.89%	72.22%	91.67%	66.67%
	Mahalanobis distance	72.22%	55.56%	91.67%	44.44%
	DTW distance	66.67%	44.44%	75.00%	22.22%
False ratio	Euclidean distance	0.00%	0.00%	0.00%	0.00%
	Mahalanobis distance	0.00%	0.00%	0.00%	0.00%
	DTW distance	0.00%	0.00%	0.00%	0.00%
Detection performa nce	Euclidean distance	11.11%	27.78%	8.33%	33.33%
	Mahalanobis distance	27.78%	44.44%	8.33%	55.56%
	DTW distance	33.33%	55.56%	25.00%	77.78%

In the anomaly detection of the angle parameter, we can find that the detection performance is the best with the Mahalanobis distance and DTW distance among all of the distance measures. Similarly, we can conclude the same conclusion to the other two types of parameters.

V. CONCLUSION

We can conclude that the anomaly detection for the satellite telemetry data, with the Mahalanobis distance and the DTW distance, can obtain the satisfied results. The improved similarity distance can guarantee the same false detection ratio compared to the Euclidean distance. Moreover, it can detect the smaller difference between the normal samples and the anomalous samples, which shows the better anomaly identification capability.

In the future, more monitoring parameters from several different satellites will be considered to evaluate the proposed method. Also to combine different similarity measure together and fuse those with other type of anomaly detection algorithms can be attempted.

ACKNOWLEDGMENT

This work was supported in part by National Natural Science Foundation of China under Grant No. 61301205 and 61571160, New Direction of Subject Development in Harbin Institute of Technology under Grant No. 01509421.

REFERENCES

[1] Liu D, Wang H, Peng Y, Xie W, Liao H. “Satellite Lithium-ion Battery Remaining Cycle Life Prediction with Novel Indirect Health Indicator Extraction”, Energies, Vol. 6, No. 8, pp.3654-3668, 2013.

[2] Li L, Gariel M, Hansman RJ, Palacios R. “Anomaly detection in onboard-recorded flight data using cluster analysis.” in IEEE/AIAA 30th Digital Avionics Systems Conference (DASC), 2011.

[3] Tan S.C., Ting K.M. Liu T.F., “Fast Anomaly Detection for Streaming Data,” in Twenty-Second International Joint Conference on Artificial Intelligence, 2011.

[4] Lhermitte S, Verbesselt J, Verstraeten W W, et al. “A Comparison of Time Series Similarity Measures for Classification and Change Detection of Ecosystem Dynamics”, Remote Sensing of Environment, , Vol. 115, No. 12, pp. 3129-3152, 2011.

[5] Batyrshin I. “Constructing Time Series Shape Association Measures: Minkowski Distance and Data Standardization”. In Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence (BRICS-CCI) Porto, Galinhas, Brazil, 2013: 204-212.

[6] Maesschalck R D, Jouan-Rimbaud D, Massart D L. “The Mahalanobis Distance”, Chemometrics and Intelligent Laboratory Systems, Vol. 50, pp. 1-18, 2011.

[7] Bemdt D, Clifford J. “Using Dynamic Time Warping to Find Patterns in Time Series”, In Proceedings of AAAI-94 Workshop on Knowledge Discovery in Databases, Seattle, USA, 1994: 359-371.

[8] Wong T S F, Wong M H. “Efficient Subsequence Matching for Sequences Databases under Time Warping”, In Proceedings of the 7th International Database Engineering and Applications Symposium, 2003: 139-148.

[9] Keogh E, Pazzani M. “Derivative dynamic time warping”, Proceedings of the 1st SIAM International Conference on Data Mining, San Jose, CA, USA, 2001: 1-11.

[10] Altman, N S. “An Introduction to Kernel and Nearest-neighbor Nonparametric Regression”, The American Statistician, Vol. 46, No. 3, pp. 175–185, 1992

[11] Hosseinifard B, Moradi M H, Rostami R. “Classifying Depression Patients and Normal Subjects Using Machine Learning Techniques and Nonlinear Features from EEG Signal”, Computer methods and programs in biomedicine, , Vol. 109, No. 3, pp. 339-345, 2013.

[12] Murtagh F, Contreras P. “Algorithms for Hierarchical C lustering: an Overview”, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, Vol. 2, No. 1, pp. 86-97, 2012.

[13] Mustafa G B, George R. “Learning a Symbolic Representation for Multivariate Time Series Classification”, Data Mining and Knowledge Discovery, Vol. 29, pp. 400-422, 2015.

[14] Bache K, Lichman M. UCI machine learning repository[EB/OL]. University of California, School of Information and Computer Science, Irvine, CA, 2013. http://archive.ics.uci.edu/ml.

[15] Jingyue Pang , Liu D, Liao H, Peng Y, Peng X. “Anomaly detection based on data stream monitoring and prediction with improved Gaussian process regression algorithm”, In 2014 IEEE Conference on Prognostics and Health Management (PHM), 2014.

[16] Ioannidis J P A, Tarone R, Mclaughlin J K. “The False-positive to False-negative Ratio in Epidemiologic Studies”, Epidemiology, Vol. 22, No. 4, pp. 450-456, 2011.