Importance Measurement of Parameters for Satellite Attitude Control System Fault Diagnosis Based on DBN

Mingjia Lei[1], Yuqing Li[1](✉), Guan Wu[2], and Junhua Feng[3]

[1] Deep Space Exploration Research Center, Harbin Institute of Technology, Harbin 150080, China
bradley@hit.edu.cn
[2] Key Laboratory for Fault Diagnosis and Maintenance of Spacecraft in-Orbit, Xi'an 710000, China
[3] National Key Laboratory of Astrodynamics, Xi'an 710000, China

Abstract. Efficient and accurate fault diagnosis of satellite attitude control system has an important role and significance to ensure the reliability of satellites in orbit. Recent researches on satellite fault diagnosis focuses on diagnosis methods, but less on the importance of telemetry parameters. Since the satellite itself is a highly complex nonlinear system, there are many types of telemetry parameters that can be used for fault diagnosis. The importance of different parameters has a greater impact on fault diagnosis. Aiming at the above problems, a new DBN-based parameter importance measurement method (PIM-DBN) is proposed by constraining the DBN network structure and fixing some part of weights during the update process. The proposed method can automatically solve the importance weights of the telemetry parameters by training network with the CD-K divergence algorithm. This method was applied to the fault diagnosis of the momentum wheel of a satellite with 10 telemetry parameters. In order to verify the effectiveness of PIM-DBN, three diagnosis method (SVM, ANN and DBN) were used to classify the data set. The accuracy of there methods above achieved 87.54%, 68.83% and 93.45% respectively with using the weighted data. These results show that the proposed importance measurement method is effective for the data-driven fault diagnosis field and health assessment field.

Keywords: Importance measurement · Deep belief network (DBN) · Satellite fault diagnosis · Attitude control system

This research was supported by the project: 2019001TJ20172A04097-4, the Key Laboratory Opening Funding of Harbin Institute of Technology of Deep Space Exploration Landing and Return Control Technology (HIT.KLOF. 2018.076, HIT.KLOF. 2018.074), and the pre-research projects of equipment development department of China Central Military Commission (JZX7Y20190243001201).

Q. Wu et al. (Eds.): WiSATS 2020, LNICST 358, pp. 61–73, 2021.
https://doi.org/10.1007/978-3-030-69072-4_6

1 Introduction

The satellite attitude control system is an important subsystem to guarantee the on-orbit operation and mission execution of the satellite. Therefore, it is of great significance to improve the health monitoring ability and fault diagnosis accuracy of the satellite attitude control system to ensure the reliability of the satellite [1, 2]. Generally, the satellite attitude control system has a variety of telemetry signals and a large amount of data, which provides a data basis for the attitude control system fault diagnosis and health assessment.

There are more and more studies on satellite health assessment. From the perspective of data, the current research can be divided into two categories: research on a single data source [3, 4] and research on multi-source data [5, 6]. With the deepening of research, the advantages of multi-source data in information expression completeness and accuracy are prominent. In recent years, the research on satellite fault diagnosis and health assessment gradually tends to focus on multi-source data.

With more and more various of telemetry parameters used for fault diagnosis and health assessment, the traditional methods based on threshold and expert system are limited in practical application. In order to make up for the shortcomings of traditional methods, many machine learning algorithms have been applied in the field of satellite fault detection [7, 8]. Although there are lots of researches on satellite fault detection methods, the importance of parameters used in fault detection is seldom studied. However, as the satellite system itself is a complex system with high coupling, there are intricate relationships between satellite telemetry data, and these relationships often have a great impact on satellite fault diagnosis results [9]. Hence, the importance of telemetry parameters in the process of satellite fault diagnosis can be analyzed to assist computers or field experts to make further judgment on satellite health status.

At present, the research on parameter importance measurement mainly focuses on two kinds of methods: based on physical model and based on data-driven. Considering the complexity of the satellite attitude control system, the data-driven method can be adopted to design the measurement method of parameter importance for the satellite.

In this paper, a parameter importance measurement method based on DBN is proposed to measure the abundance of satellite telemetry signals. This method enhances the number constraint of the underlying nodes and the connection weight constraint of DBN and trains the network based on the contrast divergence algorithm, and then obtains the parameter importance weight matrix based on RBM. The test on the satellite momentum wheel telemetry data set proves that this method can effectively improve the accuracy of fault diagnosis.

2 Parameter Importance Measurement Based on DBN (PIM-DBN)

2.1 Restricted Boltzmann Machine (RBM)

Restricted Boltzmann Machine (RBM) is a generative stochastic neural network that contains a visible layer and a hidden layer. Both the visible layer variable and the hidden layer variable of the network are binary variables, that is, their states are taken as $\{0,1\}$. The entire network is a bipartite graph. There is no mutual connection between neurons (nodes) on the same layer, and there is only mutual connection between neurons between adjacent layers. The connection between connectable neurons is bidirectional and symmetric, that is, the information can flow between the two connected neurons during the network training process, and the connection weight is independent of the direction. Usually the visible layer is used as the input layer of the data, and the hidden layer is the output layer of the generated data. Figure 1 shows the structure of the RBM.

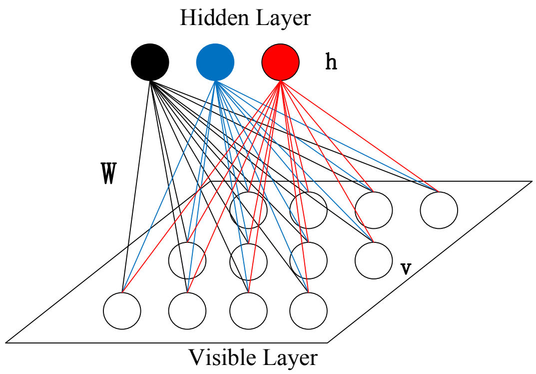

Fig. 1. The structure of RBM

Objective Function

RBM is an energy-based model, denoting its visible layer as V and hidden layer as h. In a Restricted Boltzmann Machine (RBM), visible element Vi is real number or binary number, hidden element hj is binary number, the energy expressions of visible variable V and implied variable h are shown in Eqs. (1) and (2):

$$E(v, h; \theta) = -h^T Wv - a^T v - b^T h \tag{1}$$

$$E(v, h; \theta) = -\sum\nolimits_{ij} W_{ij} v_i h_j - \sum\nolimits_i a_i v_i - \sum\nolimits_j b_j h_j \tag{2}$$

Where θ is a parameter of RBM $\{W, a, b\}$, W is the weight between visible cell (visible node) and hidden cell (hidden node), a, b are the offset vector (bias) of visible cell and hidden cell respectively. The learning process of a RBM is to find the optimal value of θ in order to fit the given input data.

According to the joint allocation energy of V and h, the joint allocation probability of V and h can be obtained as shown in Eq. (3).

$$P(v, h; \theta) = \frac{1}{Z(\theta)} exp(-E(v, h; \theta)) \tag{3}$$

According to Eq. (2), the above equation can be further written as:

$$P(v, h; \theta) = \frac{1}{Z(\theta)} exp\left(\sum\nolimits_{ij} W_{ij} v_i h_j + \sum\nolimits_i a_i v_i + \sum\nolimits_j b_j h_j\right) \tag{4}$$

The likelihood function of maximized observed data $P(v;\theta)$ is as shown in Eq. (5), which can be obtained from the edge distribution of h shown in Eq. (4).

$$P(v;\theta) = \frac{1}{Z(\theta)}\sum\nolimits_h exp\left(v^T Wh + a^T v + b^T h\right) \tag{5}$$

Where $Z(\theta)$ is a partition function, also known as a normalization factor, and $Z(\theta)$ is shown in Eq. (6).

$$Z(\theta) = \sum\nolimits_v \sum\nolimits_h \exp(-E(v, h; \theta)) \tag{6}$$

If the training data set is assumed to contain N samples, then by maximizing $P(v;\theta)$ to get the parameters of the RBM, while the maximization of $P(v;\theta)$ is the maximization of the objective function in Eq. (7):

$$L(\theta) = \sum\nolimits_{n=1}^{N} \log P\left(v^{(n)};\theta\right) \tag{7}$$

The learning objective of RBM is to find the parameter set that can maximize the logarithmic likelihood function on the training data set, and these parameter sets mainly include the weight matrix W connecting the visible layer and the hidden layer and the bias vector of the two layers.

RBM Training Method

According to Eqs. (2)–(7) and Gibbs sampling process, Eq. (7) can be further derived as

$$L(\theta) = \sum\nolimits_{n=1}^{N} (\log \sum\nolimits_h \exp\left(-E\left(v^{(n)}, h; \theta\right)\right) - \log \sum\nolimits_{v,h} \exp(-E(v, h; \theta))) \tag{8}$$

In order to sample effectively and efficiently, Hinton et al. put forward a new algorithm, that is, the CD(Contrastive Divergence) algorithm, whose basic thinking diagram is shown in Fig. 2.

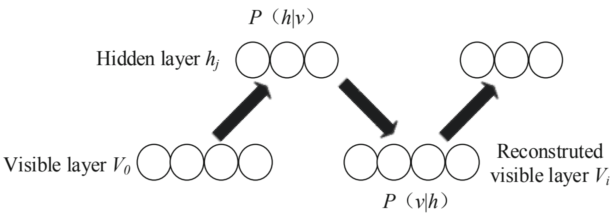

Fig. 2. Schematic diagram of basic idea of CD algorithm

After using the training data to initialize the visible layer where v_0, step on m (usually m = 1) Gibbs sampling. At the beginning of a contrast divergence algorithm, set the visible state of each node to a training sample, and use the Eq. (9) to calculate values of all the hidden layer nodes {0, 1}. When calculating all of the hidden layer unit of binary, according to the Eq. (10), the probability of each node of the visible layer being 1 is calculated, so as to obtain the first reconstruction of the visible layer. Then, the reconstructed visible layer is taken as the real model and put into the formula of RBM parameter update to calculate the updated RBM parameter.

$$P(h|v) = \prod\nolimits_j P\big(h_j|v\big)P\big(h_j = 1|v\big) = sigmoid\Big(b_j + \sum\nolimits_i W_{ij}v_i\Big) \quad (9)$$

$$P(v|h) = \prod\nolimits_i P(v_i|h)P(v_i = 1|h) = sigmoid\bigg(a_i + \sum\nolimits_j W_{ij}h_j\bigg) \quad (10)$$

The steps of training RBM with CD-K algorithm are as follows:

Step 1: Initialize

① Given training sample set X;
② Given the training period M, the learning rate η, and the parameter K of CD-k algorithm;
③ Specify Nk which is the number of cells in the hidden layer (Nv, which is the number of cells in the visible layer, is determined by the sample characteristic dimension);
④ Initialize the bias vectors a, b, and the weight matrix W;

Step 2: Training

⑤ Call CD-K (K, X, W, a, b) to calculate the network parameters variation ΔW, Δa, Δb;
⑥ Update parameters: $W_{ij} = W_{ij} + \eta\Delta W_{ij}$, $a_i = a_i + \eta\Delta a_i$, $b_j = b_j + \eta\Delta b_j$;
⑦ Repeat the above process ⑤–⑥ M times.

The flow of RBM algorithm is shown in Fig. 3.

2.2 Deep Belief Network (DBN)

Deep Belief Network (DBN) proposed by Geoffrey Hinton in 2006 is a generative model developed from logistic belief network. By training the weights among its neurons, the entire network is allowed to generate training data according to the maximum probability. Deep Belief Network (DBN) consists of multiple layers of neurons, which are divided into visible (dominant) neurons and hidden (recessive) neurons. Visible (dominant) neurons used to receive input data, hidden (recessive) neurons are used to implement feature extraction, which are also known as the hidden neurons characteristics detector (feature detectors). The top two layers of DBN structure are associative memory formed by undirected graph model. The other lower layers are connected by a directed connection between the upper and lower layers, forming a directed graph model. The hybrid model of deep belief network is shown in Fig. 4.

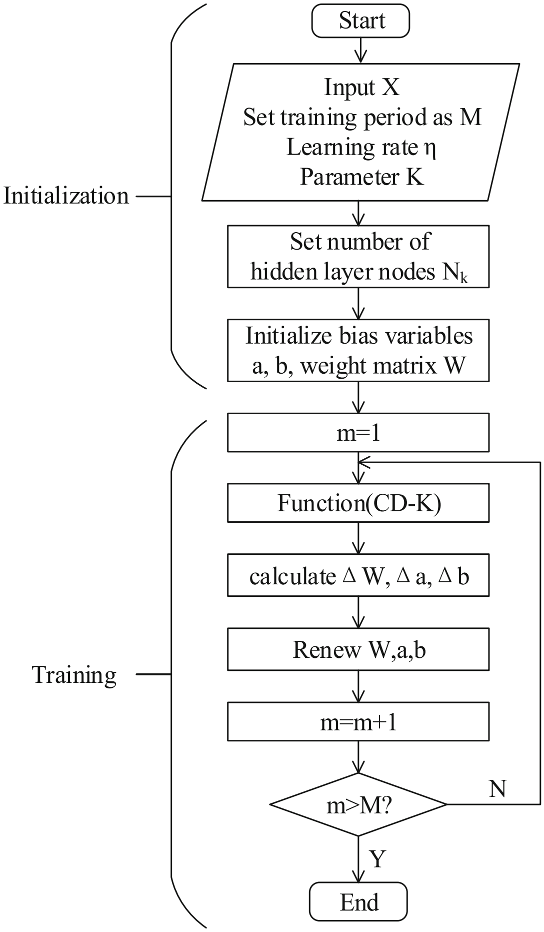

Fig. 3. The flow chart of training RBM with CD-K algorithm

DBN is essentially a stack of Restricted Boltzmann Machines, and the training process of DBN is carried out layer by layer. In RBM composed of two adjacent layers, the lower layer is the visible layer of RBM, and the data vector of the visible layer is used to infer the data of the hidden layer, and then the inferred hidden layer is used as the data vector of the next visible layer of RBM.

2.3 Parameters Importance Measurement Based on DBN (PIM-DBN)

The input data dimension is n, and the input layer node number of DBN network is n. This paper proposes a parameter importance measurement method based on DBN. By constraining the number of nodes and connecting weight matrix between the input layer and the first hidden layer of DBN, this method realizes the parameter importance

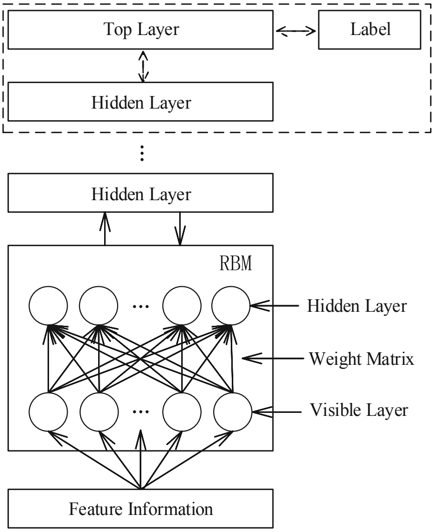

Fig. 4. Deep Belief Network (DBN) model structure diagram

measurement based on DBN structure and training method. The Eq. (11) and Fig. 5 represent the constraints.

$$\begin{cases} Node(InputLayer) = Node(HiddenLayer1) \\ w_{ij} = 0, (i \neq j) \end{cases} \tag{11}$$

Where $Node(InputLayer) = Node(HiddenLayer1)$ represents the number of nodes of input data equal to that of the first hidden layer of DBN; w_{ij} is the connection weight between the input layer node v_i and the hidden layer node h_j, and if $i \neq j$, then $w_{ij} = 0$.

The importance coefficient of input data $\{w_{11}, w_{22}, \cdots, w_{nn}\}$ can be obtained through network training. The parameter importance measurement method based on DBN is shown in the Fig. 6 as follows.

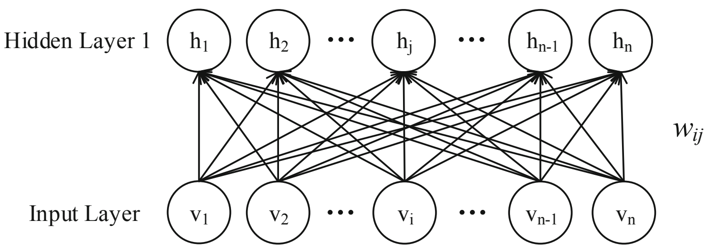

Fig. 5. The fix structure of the first RBM

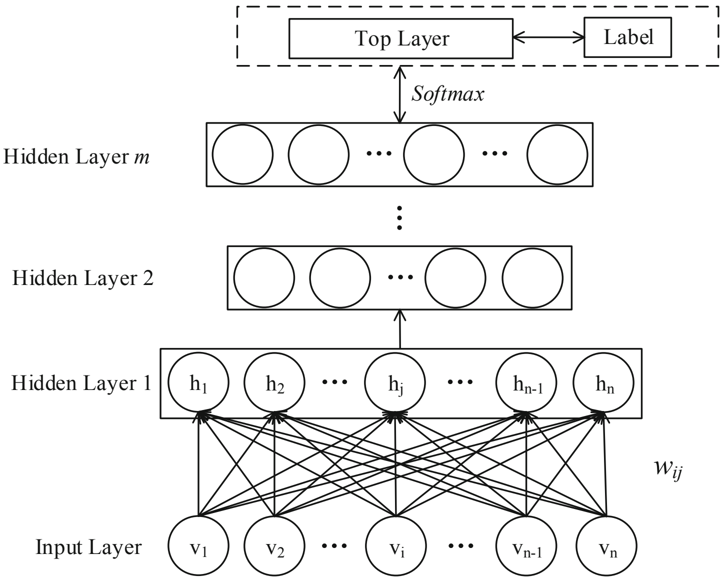

Fig. 6. The structure of DBN for parameters importance measurement

3 Fault Diagnosis of Attitude Control System Based on PIM-DBN

3.1 Parameter Set of Satellite Attitude Control System

There are many types of telemetry data for the satellite attitude control system. in this paper, 10 kinds of parameters are selected as the fault diagnosis parameter set: rotate speed; voltage; current; voltage/current; bearing temperature; case temperature; line temperature; axle temperature; friction moment; acceleration. Because satellite telemetry data often has some missing data and abnormal data noise, it is necessary to preprocess the original telemetry data. For the problem of large data noise, this paper uses median filtering to reduce noise; for the problem of missing data, the mean replacement method is used to supplement the data. In addition, to learn the importance of parameters, dimensionless and normalized data should be processed first.

3.2 The Structure of PIM-DBN

Because the dimension of the fault diagnosis parameter set of the satellite attitude control system selected in this paper is 10, the PIM-DBN structure shown in the Table 1 is designed.

Table 1. The structure parameter of PIM-DBN for satellite attitude control system

	Input layer	Hidden layer 1	Hidden layer 2	Hidden layer 3	Hidden layer 4	Hidden layer 5	Output layer
Number of nodes	10	10	64	128	64	32	N

Where N represents the number of categories of the sample data label.

3.3 Training Algorithm

PIM-DBN training algorithm is designed based on RBM training algorithm, and the specific process is as follows:
Step 1: Unsupervised pre-training

① Train the first restricted Boltzmann machine (RBM);
② The weight matrix W and offset vectors a, b of the first RBM are fixed, and then the state of the hidden layer neuron is used as the input vector of the second restricted Boltzmann machine. When the visible layer of RBM is the input layer, the weight matrix is $w_{ij} = 0, (i \neq j)$;
③ After the second RBM is fully trained, stack the second RBM on top of the first RBM;
④ Repeat the above three steps;

Step 2: Supervised fine-tuning of parameters

⑤ The output feature vector of top-level RBM is the input feature vector of fine-tuning BP network;
⑥ The final weight matrix W and the bias vector a, b are determined according to the error of label data and the calculated output result of fine-tuning BP network.

The flow chart of PIM-DBN algorithm is shown as follows (Fig. 7).

4 Analysis of Calculation Examples Based on Telemetry Data

4.1 Telemetry Data of Satellite Attitude Control System

This paper takes momentum wheel telemetry parameter of a satellite as an example to verify the effectiveness of the proposed method. The original telemetry parameters used in this paper are shown in Fig. 8.

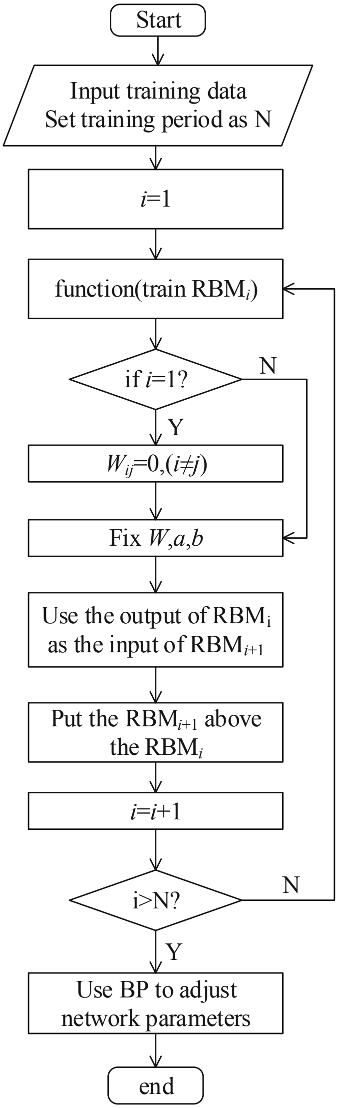

Fig. 7. The flow chart of PIM-DBN algorithm

In order to study the importance of the telemetry parameter of the momentum wheel in the fault diagnosis problem, momentum wheel telemetry data in a number of time periods are selected from the telemetry database, and the time periods include 6 modes: Acceleration Mode (Normal), Deceleration Mode (Normal), Stable Mode (Normal),

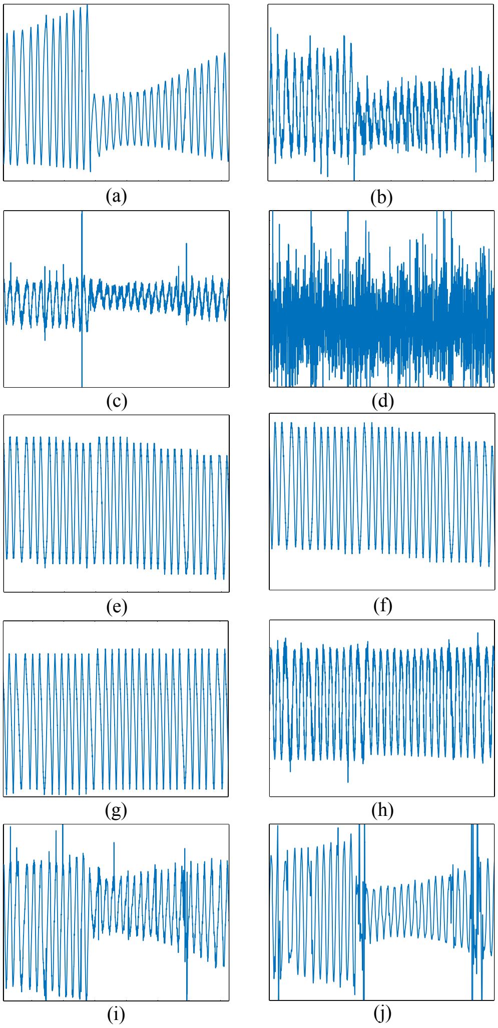

Fig. 8. Satellite momentum wheel telemetry data: (a) rotate speed; (b) voltage; (c) current; (d) voltage/current; (e) bearing temperature; (f) case temperature; (g) line temperature; (h) axle temperature; (i) friction moment; (j) acceleration

Acceleration Mode (Unnormal), Deceleration Mode (Unnormal) and Power Down, used for learning importance of parameters. The training set and testing set are shown in Table 2.

Table 2. The dataset for parameter importance measurement

State	Label	Training samples	Test samples
Acceleration Mode (Normal)	AN	3000	1000
Deceleration Mode (Normal)	DN	3000	1000
Stable Mode (Normal)	SN	3000	1000
Acceleration Mode (Unnormal)	AU	3000	1000
Deceleration Mode (Unnormal)	DU	3000	1000
Power Down	PD	3000	1000

4.2 Results and Analysis

In this paper, the PIM-DBN method is used to quantify the importance of the above 10 momentum wheel telemetry parameters:. The results are averaged by 10 trials to reduce the randomness. The importance weights of satellite momentum wheel telemetry parameters based on PIM-DBN are shown in Table 3.

Table 3. The weights of 10 parameters (P: parameters; V: voltage; C: current; T: temperature)

P	Rotate speed	V	C	V/C	Bearing T	Case T	Line T	Axle T	Friction moment	Acceleration
w	**0.635**	**0.798**	**0.721**	**0.872**	0.243	0.354	0.156	0.064	**0.578**	**0.697**

It can be seen from Table 3, rotate speed; voltage, current, voltage/current, friction moment and acceleration have higher weights,while the weight of the temperature-related parameters are lower. This result is consistent with the actual situation of the momentum wheel. Since the temperature-related parameters are greatly affected by lighting factors and less affected by the working state of the momentum wheel, the weight of the above 4 temperature remote measurements is relatively low in the fault diagnosis of the momentum wheel. On the contrary, the working state of the momentum wheel is closely related to its rotate speed, voltage, current, voltage/current, friction moment and acceleration, so the weight of the above 6 remote measurements is relatively high in the momentum wheel fault diagnosis.

In this paper, Support vector machine (SVM), Artificial neural network (ANN), DBN are used for fault diagnosis of the original data and the weighted importance data respectively. The results are averaged by 20 trials to reduce the randomness. The diagnostic accuracy is shown in Table 4. According to the data in Table 4, the accuracy of momentum wheel fault diagnosis based on DBN is higher than that based on SVM and ANN, and the fault diagnosis accuracy of the above 3 methods on the weighted original data is improved. The result shows that the parameter importance measurement method proposed in this paper can effectively improve the accuracy of fault diagnosis of satellite attitude control system.

Table 4. Comparison result of weighted data and raw data (%)

Methods	Mean accuracy	
	Weighted data	Raw data
SVM	87.54	76.34
ANN	68.83	53.64
DBN	93.45	82.79

5 Conclusion

This study presented a new DBN-based parameters importance measurement method. The main contributions of this study are proposing a new method PIM-DBN to calculate the importance of parameters and applying the PIM-DBN to the fault diagnosis field. The proposed method was tested on the telemetry data set of satellite attitude control system. In this paper, importance weights are learned for 10 types of telemetry data. The original data and weighted data are used for fault diagnosis based on SVM, ANN and DBN methods. All the method obtained higher accuracy by using weighted data. These results show the good potential of the proposed method in the data-driven fault diagnosis field and health assessment field.

References

1. Wang, D., Meng, L., Ye, P., et al.: Research of autonomous operation technology for deep space probe. Spacecraft Eng. **27**(6), 5–14 (2018)
2. Li, H., Dong, Y.: Intelligent development technology of satellite fault diagnosis system. Aerosp. Shanghai **34**(3), 52–59 (2017)
3. Ye, Z., Cheng, Y., Han, X., Jiang, B.: Fault location for attitude control systems of deep space exploration satellites. Control Theory Appl. **36**(12), 2093–2099 (2019)
4. Baldi, P., Castaldi, P., Mimmo, N., et al.: Combined singular perturbations and nonlinear geometric approach to FDI in satellite actuators and sensors. IFAC − PapersOnLine **50**(1), 7253–7259 (2017)
5. Jan, S.U., Koo, I.S.: Sensor faults detection and classification using SVM with diverse features. In: 2017 International Conference on Information and Communication Technology Convergence (ICTC), JejuIsland, Korea, pp. 576–578. IEEE (2017)
6. Ye, J., Zhen, C., Xuetao, W., et al.: A momentum wheel health ranking method based on fuzzy clustering model. Chin. Space Sci. Technol. **40**(3), 100–106 (2020)
7. Sheng, G., Wei, Z., Xu, H., et al.: Neural network-based fault diagnosis scheme for satellite attitude control system. In: 2018 Chinese Control and Decision Conference (CCDC). IEEE, Shenyang (2018)
8. Liu, Y., Yin, C., Hu, D., et al.: Communication satellite fault detection based on recurrent neural network. Comput. Sci. **47**(2), 227–232 (2020)
9. Du, Y., Budman, H., Duever, T.A.: Integration of fault diagnosis and control based on a trade-off between fault detectability and closed loop performance. J. Process Control **38**, 42–53 (2016)