Research on Diagnosis Data Fusion of Aero-engine based on Improved K-means Cluster and D-S Evidence Theory

Mingming Zhang[1, a], Xiaobo Liu[1,b]

[1]Aeronautical Manufacturing Engineering College, Nanchang Hangkong University,

Nanchang 330063, China

[a]709104871@qq.com, [b] liuxiaobo2000@21cn.com

Keywords: aero-engine, D-S evidence theory, data fusion

Abstract. The data fusion method combined improved K-means clustering algorithm with D-S evidence theory was used for vibration fault data fusion of aero-engine in this paper. The later calculated amount was reduced by the improved K-means clustering algorithm. On the basis of the improved K-means clustering algorithm, the basic belief function of vibration data was determined. D-S evidence theory was used for fusion of fault vibration data of aero-engine which had been processed through the improved K-means clustering analysis. The results of diagnostic instance show that the method can improve the diagnosis rates of aero-engine fault effectively.

Introduction

Aero-engine, as the power device for the aircraft, its stability and reliability will directly affect the safety of flight. Therefore, in order to find the aero-engine fault quickly and accurately, several sensors were installed in multiple parts of engine for signal acquisition[1]. However, when multiple sensors support the same functions, data fusion for these sensors is needed to find out fault feature, improve the certainty and reliability of information and realize the engine fault diagnosis. When each of sensors is not entirely sure of their own judgment, we can use D-S evidence method based on statistic method for data fusion, which was widely used in data fusion of multi-sensor containing uncertain information[2-3].

In this paper, the improved K-means clustering analysis was combined with D-S evidence theory, and the output of the improved K-means clustering algorithm was regard as the evidence of D-S evidence theory. Then D-S evidence theory was used for fusion of fault signal data which had been processed through the improved K-means clustering analysis.

Improved K-means clustering algorithm

D-S evidence theory, which can deal with the uncertainty led by randomness, also can handle multiple fault features, is widely used in data fusion[4-5]. D-S evidence theory has a good judgment on the fault symptom of aero-engine, however, the result is often not very precise in the processing of large amount of data. Therefore, in the paper, clustering analysis was performed on a large number of data as the early treatment before D-S data processing, it reduced the later calculated amount. Determining the initial clustering center of K-means clustering algorithm is a difficulty, the previous K-means clustering algorithm assigned the initial clustering center randomly. The improved K-means clustering algorithm is used in the paper. Firstly, the k value were obtained by the hierarchical algorithm, and then the K-means clustering analysis was done with the obtained k value, that will make the calculated fault symptom data more accurately. For details, see in literature [6].

D-S evidence theory

In equipment diagnosis, the equipment operation data transmit the information by multi-sensor, the information will be regarded as the data evidence of some possible fault, each fault occurred in the evidence under a certain probability. In the D-S evidence theory, the mass function used to

describe the probability, the transmission data of each sensor formed into the mass function of each fault. D-S evidence theory began to fusion, got the new function after the fusion of the mass function, and then calculated the plausibility and belief, finally determined the fault types.

(1) Frame of discernment

Defining that the parameter a is an event and U is the set of a_n, named the frame of discernment. The elements of U are assumed to be mutually exclusive. 2^U denotes the power set of U, then

U=(a_1,a $_2$,a_3,…,a_n);

2^U={{Φ}, {a_1}, …, {a_n}, {a_1, a_2}, {a_1, a_3}…, {a_1, a_n}, …, {a_{n-1}, a_n}, {a_1, a_2, a_3}, … ,{a_1, a_2, … , a_n} } (1)

(2) The basic probability assignment function

The function m: 2^U ~[0,1] is the basic probability assignment function, then 1) m(Φ)=0; 2)

$$\sum_{A \subseteq 2^U} m(A) = 1$$

.Where ϕ is the null set, A is the focus element and m(A) is the basic probability value and represents the support degree of the event A. The former equation expresses the support degree of the impossible proposition ϕ is 0, the latter represents the sum of the support degree of all elements in the set 2^U which contains all subsets is 1.It basically reflects the belied degree of the event A itself.

(3) Plausibility function

$$Pl(A)=1-Bel(\overline{A}) \tag{2}$$

$$PI(A) = \sum_{A \cap B \neq \phi} m(B) \tag{3}$$

Where Pl is the plausibility function, Pl(A) can be regarded as the total amount of belief that could be potentially placed in A.

(4) The belief function

$$Bel(A) = \sum_{B \subseteq A} m(B) \qquad \forall A \subseteq U \tag{4}$$

Where Bel is the belief function, Bel(A) measures the total belief that the event A is true.

(5) Combination rule of sensor information

1) Combination rule of two-sensor information

Assuming that Bel$_1$ and Bel$_2$ are the two belief functions in the definition of frame of discernment U and m1 and m2 are the basic probability assignments corresponding to belief functions. Assuming that the corresponding focal elements of Bel1 are B_1、$B_2 \cdots B_k$, and the corresponding focal elements of Bel2 are C_1、$C_2 \cdots C_n$, then

$$\sum_{B_i \cap C_j = B} m_1(B_i) m_2(C_j) < 1 \tag{5}$$

Therefore, the belief probability assignment m: $2^U \rightarrow$ [0，1] after combination is shown as follows:

$$m(B) = \begin{cases} \dfrac{\sum\limits_{B_i \bigcap C_j = B} m_1(B_i) m_2(C_j)}{1 - \sum\limits_{B_i \bigcap C_j = \phi} m_1(B_i) m_2(C_j)} & B \neq \phi \\ 0 & B = \phi \end{cases} \tag{6}$$

The belief function of m is the intersection of Bel$_1$ and Bel$_2$. The belief function given by m becomes the direct sum of Bel$_1$ and Bel$_2$, denoted as $Bel_1 \oplus Bel_2$, if not satisfied, the direct sum does not exist.

2) Combination rule of multi-sensor information

The combination rule is similar to the combination rule of two-sensor, the difference is that we should overlying each of basic probability assignments obtained in the same frame of discernment.

The example analysis

The fault data of the aero-engine are large and complex. The acquisition points in different positions, the collected data will be different with each other. The aero-engine vibration fault diagnosis training data samples is adopted in this paper.

Using the improved clustering algorithm to preprocess each group of data, the steps of improving clustering algorithm: firstly, selecting one group of sample data, furthermore, using hierarchical clustering method to determine k value, finally using the K-means clustering algorithm after determining the k value, specific contents see in the literature [6]. Classifying the improved clustering data into three groups can get the initial state parameter matrix as follows: the first group of data S_1, S_2, S_3, the second group of data S_{12}, S_{13}, S_{23} and the third group of data S_{123}. Calculating S1 firstly:

Tab.1 Sample data

State model	point 1 mm/s	point 2 mm/s	point 3 mm/s	point 4 mm/s	point 5 mm/s	point 6 mm/s
	36.31	37.26	34.96	23.51	21.51	30.21
Misalignment	24.44	39.45	9.38	37.51	24.76	25.96
	6.89	30.28	11.29	40.01	16.89	33.91
	38.16	51.93	30.69	16.27	25.15	23.99
Imbalance	42.91	55.05	30.75	22.52	25.43	24.88
	38.16	51.93	30.69	16.27	25.15	23.99
	7.20	9.70	11.83	12.33	10.98	14.18
Rubbing	21.93	25.38	24.19	22.29	23.9	18.73
	28.17	22.89	20.25	25.62	24.16	19.28
	23.93	34.80	25.60	30.78	22.37	17.51
Fault-free	33.40	33.57	30.76	17.17	22.44	29.66
	34.10	34.34	32.34	19.90	22.20	30.81

$$S_1 = \begin{pmatrix} 34.96 & 30.69 & 11.83 & 25.60 \\ 23.51 & 16.27 & 12.33 & 30.78 \\ 21.51 & 25.15 & 10.98 & 22.37 \\ 30.21 & 23.99 & 14.18 & 17.51 \end{pmatrix}$$

Each group of initial state parameter matrix has the corresponding engine fault symptom warning quantization matrix E:

$$E_1 = \begin{pmatrix} 134.46 & 109.04 & 42.25 & 106.67 \\ 98.56 & 43.17 & 40.52 & 112.15 \\ 94.16 & 67.89 & 54.91 & 106.42 \\ 121.24 & 63.13 & 78.77 & 92.15 \end{pmatrix}$$

The corresponding belief assignment function B:

$$B_1 = \begin{pmatrix} 0.25 & 0.26 & 0.26 & 0.23 \\ 0.19 & 0.34 & 0.26 & 0.21 \\ 0.22 & 0.37 & 0.19 & 0.22 \\ 0.24 & 0.38 & 0.18 & 0.20 \end{pmatrix}$$

The D-S evidence fusion was used to the obtained probability assignments, we can get:

$$m_1 = \begin{bmatrix} 0.2176440 & 0.5595281 & 0.2096298 & 0.0131980 \end{bmatrix}$$

In like manner, the first group of data S1, S2, S3, the second group of data S12, S13, S23 and the third group of data S123. The corresponding belief assignment function can be obtained, D-S

evidence fusion was done for B, the results as shown in tab.2.

We can known from tab.2, the more data the sensors collected, the fusion result is more ideal, that is to say, the multi-sensors fusion are more suitable for faults diagnosis than the single sensor fusion.

Tab.2 Improved results of clustering and D-S evidence theory fusion

	Misalignment	Imbalance	Rubbing	Fault-free
H_1	0.2176440	0.5595281	0.2096298	0.0131980
H_2	0.1843323	0.6230009	0.1810778	0.0115890
H_3	0.2746200	0.5306948	0.1700003	0.0249849
H_{12}	0.0990128	0.7968284	0.1039947	0.0001641
H_{13}	0.1958940	0.6736352	0.1285375	0.0019333
H_{23}	0.0948086	0.8427808	0.0621759	0.0002347
H_{123}	0.0308487	0.9438229	0.0253272	0.0000012

Tab.3 Direct D-S evidence theory fusion results

	Misalignment	Imbalance	Rubbing	Fault-free
m_1	0.2791580	0.4896528	0.2276047	0.0035845
m_2	0.2818619	0.3868995	0.3284825	0.0028162
m_3	0.2117826	0.6237558	0.1620435	0.0024181
m_{12}	0.2303774	0.5503280	0.2192657	0.0000289
m_{13}	0.1486302	0.7580466	0.0933019	0.0000212
m_{23}	0.1695803	0.6788457	0.1515550	0.0000189
m_{123}	0.1147646	0.8013694	0.0838658	0.0000002

In order to compare the improved clustering and D-S evidence theory fusion with the direct D-S evidence, we should calculate the results of direct D-S evidence, the results are shown in the tab.3. Known from the result of comparison, the improved clustering results are more accurate, the efficiency of the diagnosis fault has obvious improved.

Derived from the fusion results, the imbalance fault of the aero-engine vibration fault data has the highest probability of 0.9438229, therefore, the aero-engine vibration faults are mainly caused by imbalance. The data fusion method combined improved K-means clustering algorithm with D-S evidence theory can effectively improve the accuracy of engine diagnosis and increase the credibility of the results, which can prove the feasibility and effectiveness of the algorithm.

Conclusion

In this paper, the improved K-means clustering algorithm was combined with D-S evidence theory, and the fault symptom data was determined by the improved K-means clustering algorithm. We can regard the output of the improved K-means clustering algorithm as the evidence of D-S evidence theory, and then construct the basic belief allocation function. Furthermore, D-S evidence theory was used for fusion of fault signal data which had been processed through the improved K-means clustering analysis. The diagnosis examples show that this method can effectively improve the diagnosis rate.

Applying the improved K-means clustering algorithm to preprocess D-S evidence theory can reduce the later calculated amount of D-S evidence and make the fault symptom data more accurate.

References

[1]Randall Bickford, Donald Malloy. Development of a real time turbine engine diagnostic system[R], AIAA 2002-4306

[2] Wanhai Yang. The multi-sensor data fusion and its application. Xi'an: Xi'an Electronic and Science University press, 2004

[3] Yue Li, Yufa Xu, Guochu Chen, et al. Improvement and application of D-S evidence theory in multi-sensor fault diagnosis system[J].Journal of Southeast University(Natural Science Edition),2011,41s(9):102-106

[4]J.C.Helton, J.D. Johnson,W. L.Oberkampt, et al. A Sampling Based Computational Strategy for Representation of Epistemic Uncertainty in Model Predictions with Evidence Theory[J].Computer Methods in Applied Mechanics and Engineering, 2007,196 (37-40):3940-3998

[5] MacLean, W. James, Tsotsos, john K. Fast Pattern Recognition Using Normalized Grey-Scale Correlation in a Pyramid Image Representation[J], Machine Vision and Applications, 2008, 19(3):163-179

[6] Xiaobo Liu, Beibei Deng, Liangni Shen. The Improved K-means Cluster Analysis on Diagnosis Data Fusion of The Aero-engine [J]. Applied Mechanics and Materials, 2013, 328:463-467