Chapter 54
Health Condition Assessment of Hydraulic System Based on Cloud Model and Dempster–Shafer Evidence Theory

Shuaijie Mei, Mei Yuan, Jin Cui, Shaopeng Dong, and Juanru Zhao

Abstract Hydraulic transmission systems are widely used in industry because of their high output power and compact structure. To cope with the ambiguity and uncertainty in the process of hydraulic system health monitoring, this paper adopts the combination of cloud model and Dempster–Shafer evidence theory for multi-sensor data fusion from three levels: data layer, feature layer, and decision layer, which effectively avoids the problem of high conflict of evidence in Dempster–Shafer theory and completes the assessment of health status of a complex hydraulic system. Firstly, the cloud parameters are calculated to establish the expert knowledge base. Secondly, the membership matrix is used to obtain the basic probability assignments of the evidence. Then, the fusion decision of combining the same type of sensors with evidence iterations is used to improve the efficiency of fusion, and finally, Dempster's rule is performed to obtain the hydraulic system health status assessment results. The feasibility and effectiveness of this method are verified on a real data set.

S. Mei · M. Yuan · S. Dong · J. Zhao
School of Automation Science and Electrical Engineering, Beihang University, Beijing, China
e-mail: shuaijiemei@buaa.edu.cn

M. Yuan
e-mail: yuanm@buaa.edu.cn

S. Dong
e-mail: dspsx@buaa.edu.cn

J. Zhao
e-mail: audrey_zhao@buaa.edu.cn

J. Cui (✉)
Research Institute for Frontier Science, Beihang University, Beijing, China
e-mail: jincui@buaa.edu.cn

M. Yuan · J. Cui · S. Dong
Ningbo Institute of Technology, Beihang University, Ningbo, China

G. Liu and F. Cen (eds.), *Advances in Precision Instruments and Optical Engineering*, Springer Proceedings in Physics 270, https://doi.org/10.1007/978-981-16-7258-3_54

54.1 Introduction

Hydraulic systems have become indispensable transmission systems in many fields under their high stability, high transmission ratio, and adaptability to complex operating conditions [1]. To ensure the reliability of hydraulic systems operating in harsh environments, it is necessary to monitor the condition health of hydraulic system components [2]. The first is a model-based approach based on the modeling of the physical and structural information of the hydraulic system, which usually has poor monitoring results due to the inability to obtain enough detailed information about the complex structure. The second is a statistical approach based on historical measurement data and fault characteristic information [3], including the deep learning approach [4], which has become popular in recent years, the shortcoming of this approach is that it requires a large amount of historical data and has high requirements for data completeness and certainty [5]. Therefore, considering the problems of uncertainty and randomness in the operation of hydraulic systems, the above methods are not better applicable. Dempster–Shafer (D–S) evidence theory is a means of decision-making based on expert experience and has good advantages in the problem of ambiguity [6]. However, in the process of multi-sensor fusion, the evidence theory often suffers from the problem of conflicting or even contradictory evidence from multiple sources [7]. In this paper, we use the improved D-S evidence theory combined with the cloud model, which organically combines the fuzziness and randomness in the concept of uncertainty using cloud model [8], and calculates the cloud parameters of each state parameter as well as the affiliation degree; then, the basic probability assignment matrix of each evidence in D–S theory is obtained from the affiliation degree of each state parameter; further, to solve the problem of high conflict of evidence in D–S evidence theory, the isomorphic sensor evidence is averaged and iterated, and finally, dempster rule evidence synthesis is performed for heterogeneous sensors to obtain the evaluation results of hydraulic system health status.

54.2 Materials and Methods

54.2.1 Cloud Model Characteristics

Let U be a quantitative domain of arbitrary dimension expressed through exact numerical values, C be a qualitative concept within this domain, x be a quantitative value, a random realization of C, $x \in U$, and x be a random number with a stable tendency for the determinacy $\mu(x) \in [0, 1]$ of C.

$$\mu : U \rightarrow [0, 1] \forall x \in U x \rightarrow \mu(x) \tag{54.1}$$

Then the distribution of x over the theoretical domain U is called a cloud, and each x is called a cloud drop. The normal cloud model expresses the numerical characteristics of a qualitative concept in terms of a set of mutually independent parameters that together reflect the uncertainty and wholeness of the concept, thus enabling better quantitative analysis [9]. The numerical characteristics of a cloud usually contain three parameters: expectation E_x, entropy E_n, and superb entropy H_e. λ is a constant value determined according to the ambiguity and randomness of specific different parameters.

$$E_x = \overline{x} = \frac{\sum_{m=1}^{k} x}{k} \tag{54.2}$$

$$E_n = \sigma_x = \sqrt{\frac{1}{k}\sum_{m=1}^{k}(x - \overline{x})^2} \tag{54.3}$$

$$H_e = \lambda \tag{54.3}$$

54.2.2 D–S Evidence Theory

D–S evidence theory is an imprecise inference theory approach that addresses uncertainty due to lack of knowledge, and uses the "identification fram" Θ to represent the set of data to be fused, and gives a function $m : 2^{\Theta} \to [0,1]$ if it satisfies

$$m(\varnothing) = 0, \sum_{A \subset \Theta} m(A) = 1 \tag{54.5}$$

Then m is called the set of basic credibility of such identification frame Θ, if A is contained in the identification frame Θ, $m(A)$ is called the basic credibility function of A. The basic credibility function $m(A)$ represents the magnitude of the credibility of A itself.

For an arbitrary set, D–S evidence inference gives a notion of credibility function

$$Bel(A) = \sum_{B \subset A} m(B) \tag{54.6}$$

Suppose there exists an A contained in this recognition framework Θ, then give the following definition:

$$pl(A) = 1 - Bel\left(\overline{A}\right)Dou(A) = Bel\left(\overline{A}\right) \tag{54.7}$$

where pl is called $Bel's$ likelihood function, and Dou is called $Bel's$ doubt function. Then according to this definition, we can call $pl(A)$ the seeming truth of A, and $Dou(A)$ can be called the doubtfulness of A. According to the plausibility synthesis law proposed by Dempster, then we give the synthesis law for two plausibility degrees

$$m(A) = m_1 \oplus m_2 = \frac{\sum_{A_i \bigcap B_j = A} m_1(A_i) m_2(B_j)}{1 - \sum_{A_i \bigcap B_j = \Theta} m_1(A_i) m_2(B_j)} \tag{54.8}$$

Using $m_1, m_2, \ldots, m_n$ to represent the credibility distribution function of n data, and these n data are independent of each other, then multiple credibilities can be written in the following form after fusion:

$$m(A) = m_1 \oplus m_2 \oplus \cdots \oplus m_n = \frac{\sum_{\cap A_i = A} \prod_{i=1}^{m} m_i(A_i)}{1 - \sum_{\cap A_i = \Theta} \prod_{i=1}^{m} m_i(A_i)} \tag{54.9}$$

54.2.3 Improved D–S Evidence Theory Based on Cloud Model

The flow chart of improved D–S evidence theory based on the cloud model for hydraulic system health condition assessment is shown in Fig. 54.1. It mainly includes the calculation of cloud model feature parameters, the establishment of cloud model knowledge base, the calculation of membership function and basic probability assignment, the fusion of D–S evidence, and the fusion result by fusion decision.

Suppose that there are n classes of faults in the expert system knowledge base of the system: F_1, F_2, F_3, ...F_n, each class of faults has m characteristic parameters: x_{i1}, x_{i2}, ..., x_{im}, where $x_{ij}(j = 1, 2, ..m)$ denotes the jth parameter of the ith class of faults. In industry, there are two main types of parameters for industrial equipment,

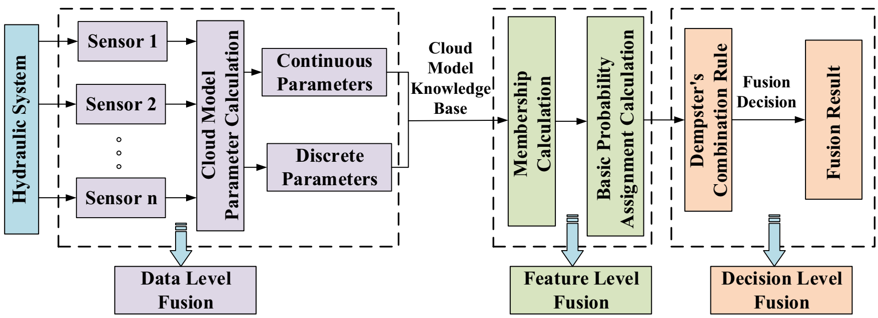

Fig. 54.1 Block diagram of multi-sensor data fusion in hydraulic system

discrete and continuous parameters, so these two can be modeled separately. The required a priori knowledge is obtained through historical information to construct the fault knowledge base.

For the modeling of continuous parameters, since the values of variables are different under different fault conditions, the cloud with the main action region as the bilateral constraint region can be used to approximate the modeling. If the value interval in the fault signal measured under the rth fault mode is $[C_{\min}(r), C_{\max}(r)]$, the median value of the constraint can be adopted as the expected value, and the specific cloud model parameters are calculated as follows:

$$E_{xij}(r) = \frac{(C_{\min}(r) + C_{\max}(r))}{2} \tag{54.10}$$

$$E_{nij}(r) = \frac{(C_{\min}(r) + C_{\max}(r))}{6} \tag{54.11}$$

$$H_{eij} = \lambda \tag{54.12}$$

For the modeling of discrete parameters, the cloud model of the expert system fault knowledge base can be established directly by experimentally measuring the mathematical expectation and standard deviation of the variable parameters (see 54.2–54.4). For the characteristic variables of discrete parameters, the membership degree is calculated as follows, which is the same as continuous parameters.

$$\mu_{ij} = e^{\frac{-\left(x_j - E_{xij}\right)^2}{2\left(E'_{nij}\right)^2}} \tag{54.13}$$

where $\mu_{ij}(k)$ is the membership degree of the jth characteristic parameter of a fault signal obtained by the measurement with respect to the jth characteristic pattern of the i-th class of faults in the expert knowledge base, E_{xij} denotes the expectation value previously obtained in the expert knowledge base, and E'_{nij} is a normal random number generated with the entropy E_{nij} as the expectation and the superentropy H_{eij} as the standard deviation and the normal random number generated. This leads to the affiliation matrix

$$R_{m\times n} = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{n1} \\ \mu_{12} & \mu_{22} & \cdots & \mu_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{1m} & \mu_{2m} & \cdots & \mu_{nm} \end{bmatrix} \tag{54.14}$$

To improve the credibility and accuracy of the fusion results, the membership degree matrix $R_{m\times n}$ is normalized.

$$\gamma_{ij} = \mu_{ij} / \sum_{i=1}^{n} \mu_{ij}, \ j = 1, 2, \ldots m \tag{54.15}$$

The uncertainty of the actual measurement signal due to the errors caused by the circumstances such as the measurement environment and the measurement method in the actual project is represented by the variable θ. Where $max(\mu_{i1}, \mu_{i2}, \ldots, \mu_{nm})$ denotes the maximum value of each element in each row of the membership degree matrix.

$$\theta_j = 1 - \max(\mu_{i1}, \mu_{i2}, \ldots, \mu_{nm}), \ j = 1, 2, \ldots m \tag{54.16}$$

Thus, the basic probability assignment function can be computationally determined as

$$\begin{cases} m(\Theta_j) = \theta_j, \ j = 1, 2, \ldots m \\ m(F_{ij}) = (1 - \theta_j)\gamma_{ij}, \ i = 1, 2, \ldots n \end{cases} \tag{54.17}$$

where $m(\Theta_j)$ denotes the basic probability assignment of the jth evidence uncertainty in the test sample, and $m(F_{ij})$ denotes the basic probability assignment of the jth characteristic parameter of the fault signal obtained from the measurement compared to the jth characteristic value of the ith fault in the expert knowledge base. The basic probability assignment matrix $M_{m \times (n+1)}$ for m rows and $n + 1$ columns can be obtained after considering both measurement data and uncertainty

$$M_{m \times (n+1)} = \begin{bmatrix} m(R_{11}) & \cdots & m(R_{n1}) & \theta_1 \\ m(R_{12}) & \cdots & m(R_{n2}) & \theta_2 \\ \vdots & \ddots & \vdots & \vdots \\ m(R_{1m}) & \cdots & m(R_{nm}) & \theta_m \end{bmatrix} \tag{54.18}$$

In order to solve the problems of low sensitivity among fault features, high conflict among fused evidence and large uncertainty, this study determines the weights of fused evidence by two aspects and reallocates the weights by the uncertainty coefficient ω_j^{σ} and the overall support coefficient ω_j^{s} of the evidence, respectively, so as to mitigate the conflict problem among the evidence, and after that, use Dempster's rule for evidence fusion. Let ω_j be the weight coefficient of the jth fault feature measured after fusing the evidence, then ω_j should satisfy the condition that

$$\sum_{j=1}^{m} \omega_j = 1, \omega_j \geq 0 \tag{54.19}$$

$$\omega_j = 0.5\omega_j^{\sigma} + 0.5\omega_j^{s}, \ j = 1, 2, \ldots, m, 0 \leq \omega_j \leq 1 \tag{54.20}$$

In the actual industry, there will be errors when measuring and collecting data due to the layout location of heterogeneous sensors, environmental conditions, and other factors, and the weight coefficient determined by the uncertainty brought by the sensor measurement and the fusion evidence is defined as the uncertainty coefficient. Let the relative measurement error of the sensor be χ_j, then

$$\chi_j = \sqrt{\sigma_j^2}/E(x_j), \ j = 1, 2, \ldots, m \tag{54.21}$$

$$\omega^{\sigma} = \frac{1}{\chi_j + \theta_j} / \sum_{k=1}^{m} \frac{1}{\chi_j + \theta_j}, \ j = 1, 2, \ldots, m \tag{54.22}$$

where $E(x_j)$, σ_j^2 denote the mean and variance of the jth fault characteristic parameter, respectively. The overall support coefficient indicates the mutual support between the evidence and the evidence, and the overall support of the evidence is determined by the distance between the evidence, assuming the existence of two pieces of evidence m_j and m_d, and defining the distance function as

$$d(m_j, m_d) = \frac{\|m_j - m_d\|}{\sqrt{s}} = \sqrt{\frac{1}{s}\sum_{i=1}^{s}(m_{ji} + m_{di})^2 j}, \ d = 1, 2, .., m \tag{54.23}$$

Then the overall support of the evidence is

$$\eta(m_j) = \sum_{d=1, d \neq j}^{m} (1 - d(m_j, m_d)) j = 1, 2, \ldots, m \tag{54.24}$$

where the larger $\eta(m_j)$ indicates that the higher the support of the evidence in the overall evidence, the less conflict with other evidence, and thus the greater the weight of the evidence in the final fusion. The overall support coefficient of the evidence is calculated as

$$\omega_j^s = \eta(m_j) / \sum_{j=1}^{m} \eta(m_j) j = 1, 2, \ldots, m \tag{54.25}$$

In order to reduce the number of evidence in the final fusion, reduce the running time and improve the fusion efficiency, after obtaining the basic probability assignment matrix and the fusion weight coefficients, the combined evidence iterations are performed on the homogeneous sensor information, and then the final fusion is performed by the Dempster combination rule after the iterations to obtain the decision results. Let there be j pieces of evidence generated by the feature parameters obtained from the homogeneous sensors, the average iterative evidence is calculated as

$$m(\Theta) = \sum_{j=1}^{m} \omega_j \theta_j, m(F) = \sum_{j=1}^{m} \omega_j m\big(F_{ij}\big) i = 1, 2, \ldots, n \tag{54.26}$$

54.3 Results and Discussion

This work uses a publicly available real dataset of complex hydraulic systems, which has been publicly released by the UC Irvine Machine Learning Repository [2, 10]. The author has developed a hydraulic test bench to measure the state data of this hydraulic system through multiple real and virtual sensors, from which the characteristics of the hydraulic system under different faults are analyzed. This study is illustrated with one of the cooling state health states, which are divided into three operating states: Close to Total Failure (CTF), Reduced Efficiency (RE), and Full Efficiency (FE). For each condition, 150 sets of data are selected for the calculation to obtain a priori knowledge, and 60 sets of data are selected as tests to verify the results.

54.3.1 Cloud Model Parameters Calculation

The actual industry faces the problem of inconsistent sensor sampling rate, to unify the data length, this paper adopts the unified data length utilizing time–frequency domain feature extraction, 24 common time–frequency domain features are selected in this paper [11], and finally, the cloud model feature matrix data of the cooling condition of the hydraulic system of 3*24 is obtained, and the value of super entropy is taken as a constant value of 0.1, as shown in Table 54.1. Due to space limitation, only the first five parameters of a set of test data are shown in all the following tables.

Table 54.1 Cloud characteristics of partial time–frequency features

Condition type		Feature 1	Feature 2	Feature 3	Feature 4	Feature 5
Cloud expectation	CTF	19.8686	0.2613	19.8678	19.8704	0.5098
	RE	27.7732	0.2889	27.7724	27.7748	0.5411
	FE	47.1203	0.2618	47.1199	47.1210	0.5123
Cloud entropy	CTF	0.1921	0.0637	0.1920	0.1923	0.1151
	RE	0.2339	0.0747	0.2339	0.2340	0.1222
	FE	0.3311	0.0726	0.3311	0.3311	0.1167

Table 54.2 Partial basic probability assignment matrix

Time and frequency characteristics	CTF	RE	FE	Uncertainty
1	0.0000	0.0000	0.8047	0.1953
2	0.3225	0.3301	0.3365	0.0109
3	0.0000	0.0000	0.7231	0.2769
4	0.0000	0.0000	0.7928	0.2075
5	0.3239	0.3288	0.3448	0.0025

Table 54.3 Partial final fusion evidence and results

Evidence	CTF	RE	FE	Uncertainty
1	0.0077	0.0093	0.0117	0.0031
2	0.0083	0.0071	0.0094	0.0062
3	0.0108	0.0060	0.0128	0.0015
4	0.0096	0.0108	0.0105	0.0010
5	0.0064	0.0065	0.0087	0.0183
Θ	0.1750	0.1441	0.6809	0.0000

54.3.2 Calculation of the Basic Probability Assignment Matrix

According to the obtained cloud model parameters, calculate the membership degree of each parameter relative to the corresponding parameter of the cooling state of the hydraulic system, calculate the uncertainty of each piece of evidence, to obtain the basic probability assignment matrix of the hydraulic system as shown in Table 54.2.

54.3.3 Iteration of Homogeneous Sensor Merging Evidence

The same sensors in this dataset are iterated to merge evidence, 8 heterogeneous sensor evidence information are obtained, and finally, these 8 shreds of evidence are fused by applying Dempster's rule to obtain the cooling condition health assessment of this hydraulic system, as shown in Table 54.3, and the final fusion result of this test data is FE.

54.3.4 Experimental Results

The 60 sets of data of the cooling condition of each type of hydraulic system are subjected to the above experimental calculation, and the final operating condition

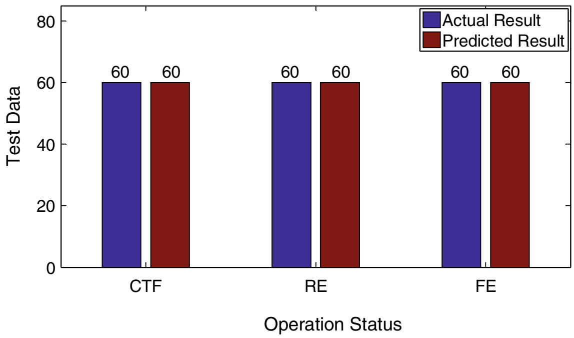

Fig. 54.2 Experimental results

classification results are shown in Fig. 54.2. It can be seen that the classification accuracy of the three operating states of the cooling condition of this hydraulic system reached 100%, and achieved quite good classification results, which proved the feasibility and effectiveness of the improved D–S evidence theory based on the cloud model.

54.4 Conclusions

In this paper, a hydraulic system health assessment method based on the cloud model and D–S evidence theory is proposed to address the problem of ambiguity and randomness of each assessment state quantity in hydraulic system health state assessment. To cope with the problem of the inconsistent sampling rate of hydraulic system acquisition sensors in the industry, a time–frequency domain feature analysis is performed to unify the data length; then, a multi-source information uncertainty fusion method for hydraulic system condition monitoring is constructed using the cloud model from quantitative to qualitative modeling and using D–S evidence theory to obtain the assessment results of hydraulic system health status. The validity and feasibility of the method are verified by real data sets to provide a basis for the condition maintenance of the hydraulic system.

Acknowledgements The research is supported by Beijing Natural Science Foundation (grant number L212033), and the National Key Research and Development Program of China (grant number 2019YFB1705502).

References

1. P. Guo, J. Wu, X. Xu, Y. Cheng, Y. Wang, Health condition monitoring of hydraulic system based on ensemble support vector machine, in *2019 Prognostics and System Health Management*

Conference (PHM-Qingdao) (2019), pp. 1–5. https://doi.org/10.1109/PHM-Qingdao46334.2019.8942981
2. N. Helwig, E. Pignanelli, A. Schuetze, Ieee, *Condition Monitoring of a Complex Hydraulic System using Multivariate Statistics* (2015 IEEE International Instrumentation and Measurement Technology Conference, 2015), pp. 210–215
3. T.M.A. Manghai, R. Jegadeeshwaran, G. Sakthivel, R. Sivakumar, D.S. Kumar, IEEE, Condition monitoring of hydraulic brake system using rough set theory and fuzzy rough nearest neighbor learning algorithms, in *2019 Ieee International Symposium on Smart Electronic Systems* (IEEE Computer Soc, Los Alamitos (in English), 2019), pp. 229–232
4. C. Konig, A.M. Helmi, Sensitivity analysis of sensors in a hydraulic condition monitoring system using CNN models. Sensors **20**(11), 3307. https://doi.org/10.3390/s20113307
5. L. Wang, K. N. Teng, W. M. Lv, System level health condition assessment method of complex equipment under uncertainty based on D-S evidence theory, in *2014 International Conference on Management Science & Engineering*, H. Lan Ed., (International Conference on Management Science and Engineering-Annual Conference Proceedings, IEEE, New York, 2014), pp. 435–441
6. G.Z. Zhao, A.G. Chen, G.X. Lu, W. Liu, Data fusion algorithm based on fuzzy sets and D-S theory of evidence, (in English). Tsinghua Sci. Technol. **25**(1), 12–19. https://doi.org/10.26599/tst.2018.9010138
7. F.Y. Xiao, A new divergence measure for belief functions in D-S evidence theory for multisensor data fusion (in English). Inf. Sci. **514**, 462–483. https://doi.org/10.1016/j.ins.2019.11.022
8. S. Kaparthi, A. Mann, D.J. Power, An overview of cloud-based decision support applications and a reference model (in English). Stud. Inform. Control **30**(1), 5–18. https://doi.org/10.24846/v30i1y202101
9. Y. Li, A. Wang, X. Yi, Based on normal cloud model and D-S evidence theory comprehensive evaluation of operation state of fire control system, in *2019 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)*, (2019), pp. 877–883. https://doi.org/10.1109/SDPC.2019.00167
10. S.S. Chawathe, *Condition Monitoring of Hydraulic Systems by Classifying Sensor Data Streams* (2019 IEEE 9th Annual Computing and Communication Workshop and Conference, 2019), pp. 898–904
11. Y. Lei, Z. He, Y. Zi, Q. Hu, Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAS. Mech. Syst. Signal Proc. **21**(5), 2280–2294 (2007). https://doi.org/10.1016/j.ymssp.2006.11.003