

Fault detection Based Bayesian network and MOEA/D applied to Sensorless Drive Diagnosis

Qing Zhou^{1,a}, Ling He² and PengFei Lu¹

¹Wuhan University of Technology, School of Automation, 122 Luoshi Road, Wuhan, China

²Wuhan University of Technology, School of Management, 205 Xiongchu Road, Wuhan, China

Abstract. Sensorless Drive Diagnosis can be used to assess the process data without the need for additional cost-intensive sensor technology, and you can understand the synchronous motor and connecting parts of the damaged state. Considering the number of features involved in the process data, it is necessary to perform feature selection and reduce the data dimension in the process of fault detection. In this paper, the MOEA / D algorithm based on multi-objective optimization is used to obtain the weight vector of all the features in the original data set. It is more suitable to classify or make decisions based on these features. In order to ensure the fastness and convenience sensorless drive diagnosis, in this paper, the classic Bayesian network learning algorithm-K2 algorithm is used to study the network structure of each feature in sensorless drive, which makes the fault detection and elimination process more targeted.

1 Introduction

The complexity of automation systems is steadily increasing, especially in process engineering. These are usually made up of individual autonomous subsystems, and their failure can have a serious impact on the overall system. Usually, this can lead to loss of production, which can lead to considerable economic losses. The sensorless monitoring of synchronous motors is based on the analysis of the current data in the process of operation, so as to obtain the damage state of synchronous motors and connecting components [1]. Considering the number of features involved in the process data, if the characteristics of all measurements are taken directly, the amount of calculation will be larger in the process of fault detection. Moreover, due to the interference of redundant features, the classification of fault states is not very accurate, so feature selection (FS) is needed to reduce the data dimensionality. There have been some studies in this field. For example, the generalised Rayleigh coefficient [2] computed using LDA is applied to choose suitable feature combinations among all possible combinations optimised with respect to compactness and separability of the classes [3]. But most feature selection algorithms focus only on selection rather than selection and weighting, where the later has the potential to further enhance the classifier performance [4].

In this paper, we use the Dataset for Sensorless Drive Diagnosis to select and weigh the 48 features in the data set based on the MOEA/D algorithm [5]. The feature vectors are selected and weighted or scaled simultaneously to project the data points to such a hyper space, where the distance between data points of non-

identical classes is increased, thus, making them easier to classify.

At the same time, in order to quickly troubleshoot in the fault detection problem, you need to know the relationship between the variables. There are many methods for this, such as neural networks [6], Petri nets [7,15], fault tree [8], decision tree [9], etc., but there are some deficiencies in all respects. In this paper, the classical Bayesian network learning algorithm (K2 algorithm [11]) for fault detection [10] is applied to learn the sensorless drive network structure according to the data after the feature selection and weighting. The network structure is closely related to the causal relationship between the various features, and more targeted in the process of sensorless drive.

2 Dataset

The data set has 48 consecutive features, is about the current signals of the motor, shaft, bearing and load sensors in the motor running process. The data set also includes a category column, which has 11 possible values, describing the different failure modes of the various components of sensorless drive [1].

Features are extracted from electric current drive signals. The drive has intact and defective components. This results in 11 different classes with different conditions. Each condition has been measured several times by 12 different operating conditions, this means by different speeds, load moments and load forces. The current signals are measured with a current probe and an oscilloscope on two phases.

^a Qing Zhou: qingzhou@whut.edu.cn

The Empirical Mode Decomposition (EMD) was used to generate a new database for the generation of features. The first three intrinsic mode functions (IMF) of the two phase-currents and their residuals (RES) were used and broken down into sub-sequences. For each of those sub-sequences, the statistical features mean, standard deviation, skewness and kurtosis were calculated. More information about the measurement procedure and a detailed description of the demonstrator's mechanical part can be found in [1].

3 Fault Detection Based Bayesian Network and MOEA/D

3.1. MOEA/D Algorithm

The fewer features are required for distinction, the more efficient is the decision making. A dimension reduction results in a better interpretability as well as in a reduced complexity of the information to be analysed. This leads to less learning effort for classification and a faster reaction time which is an important criterion for future implementations in industrial applications.

3.1.1 Algorithm principle

MOEA/D is a feature selection and weighting method aided with the decomposition based evolutionary multi-objective algorithm. The formulated inter-class and intra-class distance measures are maximized and minimized simultaneously by using a Multi-Objective Evolutionary Algorithm based on Decomposition (MOEA/D). The population of MOEA/D comprises of vectors containing the candidate feature weights, which are varied according to the algorithm to optimize the inter-class and intra class distance measures discussed subsequently. A repair mechanism is used to increase the probability of choosing lesser number of features. Also, a penalty function is augmented with the fitness functions to reduce the number of features selected. After convergence to optima using MOEA/D, a fuzzy membership based technique [12] is incorporated to choose only a single feature subset (representing the best compromise solution) out of the Pareto optimal solution set.

In the process of multi-objective evolutionary iteration, there are two objective functions, which are the distance between classes and the distance within the class. The individual of the evolutionary population is the weight vector of the weight of all the features.

3.1.2 Algorithm flow

The goal of evolution is to find a feature weight vector which maximizes the inter-class and minimizes the intra-class and the specific algorithm flow is as follows:

- (1) Input the training dataset;
- (2) Normalize the training dataset;
- (3) Compute inter-class and intra-class distance;
- (4) Randomly initialize Feature Selecting-Weighting vector (W) which comprise the population of MOEA/D,

Repeat (4) to (8) for every individual of the MOEA/D population;

(5) Crossover and Mutation to create new feature selecting-weighting vectors;

(6) Repair the newly generated selecting-weighting vectors;

(7) Compute the Objective values;

(8) Update solutions;

(9) Judgment of iteration termination criteria If termination criterion not reached, return to (4), else go to (10) end if;

(10) Use Best Compromise Solution technique to find the optimal feature selection-weighting vector (W^*) from the Pareto optimal set.

In the process of obtaining the optimal feature selection-weighting vector (W^*), according to its processing of the data, we can get the data with selection and weighting, as the input data to learn the relationship between the various characteristics.

3.2. Fault detection base Bayesian network

A Bayesian Network (BN) is a probabilistic graphical model [13], which provides an elegant way to express probabilistic relationships between random variables. Edges of the graph represent dependence between linked nodes. A formal definition is given here:

A Bayesian network is a triplet $\{G, E, D\}$ where:

$\{G\}$ Is a directed acyclic graph, $G = (V, A)$, with V the set of nodes of G , and A the set of edges of G ,

$\{E\}$ Is a finite probabilistic space (Ω, Z, p) , with Ω a non-empty space, Z a collection of subspace of Ω , and p a probability measure on Z with $p(\Omega) = 1$,

$\{D\}$ Is a set of random variables associated to the nodes of G and defined on E . Such as:

$$p(V_1, V_2, \dots, V_n) = \prod_{i=1}^n p(V_i | C(V_i)) \quad (1)$$

with $C(V_i)$ the set of causes (parents) of V_i in the graph G .

In the network, although the relationship between the parent node and the child node does not require causality, this representation can accurately reflect the dependency relationship between them in the fault diagnosis. Bayesian network structure (directed graph) expresses the relationship between the characteristics and the degree of influence. The node variables represent features, the directed edge of the connected node represents the association between the features, the conditional probability represents the degree of influence between the features. The Bayesian network has the following characteristics:

- (1) Bayesian network is an uncertain causal relationship model. The Bayesian network is a kind of probabilistic knowledge representation and reasoning model which visualizes the multivariate knowledge, and

more closely implies the causal relationship and the conditional relation between the node variables [14].

(2) Bayesian networks have powerful ability to deal with uncertain problems. The Bayesian network uses conditional probabilities to represent the correlations among the information elements. It can learn and reason under limited, incomplete and uncertain information.

(3) Bayesian network can effectively perform multi-source information representation and fusion. The Bayesian network can integrate various information related to fault detection into the network structure, and can be unified according to the way of nodes, and can be effectively fused according to the correlation of the information.

The appeal Bayesian network characteristics are consistent with the requirements of fault detection. It can be seen that the Bayesian network-based fault detection method is more superior and qualitatively and quantitatively represents the relationship between variables and Visualization. Here we use the classic K2 algorithm, based on data learning to obtain the relationship between the various characteristics of the Bayesian network.

4 Experimental results

Before the feature selection, the size of the data set is 50000 * 49, where the first 48 columns are different for the synchronous motor operation. The last column is the label of each sample, indicating which class the sample belongs to, a total of 11 classes. Since the data given is about the synchronous motor's various current characteristics, it is continuous data. Firstly, the data is pretreated and normalized and discretized. Each feature is divided into 15 classes, each of which is valued as an integer between [1 15]. Without the application of feature selection, the 40000 sets of data are based on the classical K2 algorithm to study its network structure, and the resulting Bayesian network structure is shown in Figure 1.

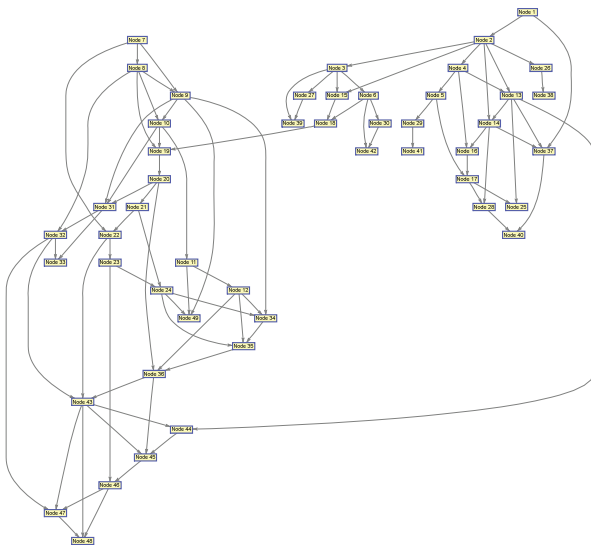


Figure 1. Bayesian network structure before feature selection.

Based on the first 10000 sets of samples, the MOEA/D algorithm is used to iterate the 150 generation to get the feature weight vector W^* with the highest accuracy, as shown in Table 1. According to the eigenvector W^* , the data is dimensioned down to obtain a data set with a size of 50000 * 29, and then the normalized and discretized processing is carried out. The Bayesian network structure obtained by K2 algorithm is shown in Figure 2.

Table 1. feature weight vector W^* .

W							
10	0	1.668	0	8.702	1.211	9.995	10
10	9.474	9.884	10	0	0	6.100	0
0	0	0	0	0	0	1.201	0
10	10	10	10	0	0	10	1.381
3.282	1.424	10	1.470	5.559	0	6.247	5.835
0	2.512	0	0	2.575	0	3.808	0

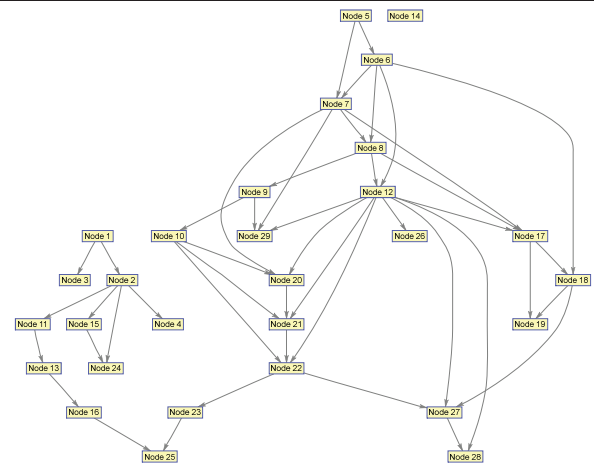


Figure 2. Bayesian network structure after feature selection.

Before and after the FS, the parameters of the network structure learning are shown in Table 2, where the accuracy is calculated as

$$\eta = \frac{N_{true}}{N_{test}} \times 100\% \tag{2}$$

Where N_{test} is the number of samples in the test data and N_{true} is the number of samples correctly classified. It can be seen that the time required for learning the Bayesian network structure (T_{train}) is significantly shortened, and the accuracy of classification is not reduced. In this way, the feature selection of the data reduces the computational complexity and reduces the time required for training. The fault detection method based on Bayesian network can visualize the relation among each feature, so that the fault can be eliminated quickly.

Table 2. Before and after FS.

	Feature number	Training data number	Training time	Test dataset number	accuracy
Before FS	48	38000	467.099	2000	88.25%
After FS	28	38000	110.688	2000	88.30%

5 Conclusion

In this paper, the MOEA/D algorithm based on multi-objective evolution is applied to the feature selection and weight calculation of Sensorless Drive. Under the premise of accurate classification and decision-making, we choose the least possible features to reduce the computational complexity. The classical K2 algorithm is applied to study from the data before and after FS, to get the network structure which characterizes the relationship between features. Through the above experiments, we can find that under the action of MOEA/D algorithm, the number of features is obviously reduced, the time required for training is reduced, and the accuracy of classification is not reduced. This kind of network relationship is more practical, in the actual fault detection and elimination process more quickly and efficiently.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (Grant No. 71603197, 71371148), Junior Fellowships for CAST Advanced Innovation Think-tank Program (DXB-ZKQN-2016-013), China Postdoctoral Science Foundation (2016M592403, 2017T100588), Special Project of Technology Innovation of Hubei Province -Soft Science (No. 2016ADC094) Fundamental Research Funds for the Central Universities (175211006).

References

1. Lohweg, Volker, Paschke, Fabian, Bayer, Christian. Sensorlose Zustandsüberwachung an Synchronmotoren[J]. 2013.
2. Bator M, Dicks A, Mönks U, et al. Feature Extraction and Reduction Applied to Sensorless Drive Diagnosis[C]// 22. Workshop Computational Intelligence. 2012.
3. Duda, R. O.; Hart, P. E.; Stork, D. G.: Pattern classification. New York: Wiley, 2 Aufl. ISBN 9780471703501. 2001.
4. Raymer M L, Punch W F, Goodman E D, et al. Dimensionality reduction using genetic algorithms[J]. IEEE Transactions on Evolutionary Computation, 2000, 4(2):164-171.
5. Paul S, Das S. Simultaneous feature selection and weighting - An evolutionary multi-objective optimization approach[M]. Elsevier Science Inc. 2015.
6. Mechefske C K, Li L. Induction Motor Fault Detection and Diagnosis Using Artificial Neural Networks[C]// ASME 2005 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. 2005.
7. Tazaki E, Yoshida K. A fuzzy Petri net model for approximate reasoning and its application to medical diagnosis[C]// IEEE International Conference on Systems, Man and Cybernetics. IEEE, 1992:627-631 vol.1.
8. Andow P K. Difficulties in Fault-Tree Synthesis for Process Plant[J]. IEEE Transactions on Reliability, 2009, R-29(1):2-9.
9. Zhao Y, Yang L, Lehman B, et al. Decision tree-based fault detection and classification in solar photovoltaic arrays[C]// Applied Power Electronics Conference and Exposition. IEEE, 2012:93-99.
10. Matsuura J P, Yoneyama T. Learning bayesian networks for fault detection[C]// Machine Learning for Signal Processing, 2004. Proceedings of the 2004, IEEE Signal Processing Society Workshop. IEEE, 2004:133-142.
11. Herskovits E, Herskovits E. A Bayesian Method for the Induction of Probabilistic Networks from Data[M]. Kluwer Academic Publishers, 1992.
12. Abido M A. A novel multiobjective evolutionary algorithm for environmental/economic power dispatch[J]. Electric Power Systems Research, 2001, 65(1):71-81.
13. Lenz H J, Finn V, Jensen, Thomas D, Nielsen (2007): Bayesian Networks and Decision Graphs[J]. Statistical Papers, 2011, 52(1):235-237.
14. Genesereth M R, Nilsson N J. CHAPTER 8—Reasoning with Uncertain Beliefs[J]. Logical Foundations of Artificial Intelligence, 1987:177-206.
15. Gao Meimei, Wu Zhiming. Fuzzy Reasoning Petri Net and its Application to Fault Diagnosis[J]. Acta Automatica Sinica, 2000, 26(5):677-680.