CHEMICAL ENGINEERING TRANSACTIONS
VOL. 71, 2018
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet
Guest Editors: Xiantang Zhang, Songrong Qian, Jianmin Xu

ISBN 978-88-95608-68-6; **ISSN** 2283-9216
DOI: 10.3303/CET1871215

Application of Improved Rough Set Reduction Algorithm in On-line Fault Diagnosis of Chemical Equipment

Chengzhang Ji, Shanqun Lu

School of software of the Weifang University of Science and Technology, Shouguang 262700, China
chengzhangji2112@163.com

The application of improved rough set reduction algorithm in Chemical Equipment on-line fault diagnosis is discussed. The rough set association degree is applied to attribute association degree matrix, and the improved rough set reduction algorithm is obtained, and the application scope of this method in Chemical Equipment online fault diagnosis is studied. The results show that the improved rough set reduction algorithm is a new data mining method, which can effectively solve the problem of attribute classification of equipment in online fault diagnosis. Therefore, the improved rough set reduction algorithm can quickly and effectively diagnose on-line faults in Chemical Equipment system and eliminate the corresponding faults in time.

1. Introduction

In recent years, with the rapid development of computer technology and automation technology, human information has increased and data mining theory has become one of the indispensable tools accordingly. The scientific operation of the Chemical system also makes it become the main research work about how to deal with organizational data.

Therefore, this paper explores the application of rough set and cluster analysis methods and finds similar clusters in the data. By analysing the source of fault information in the distribution network, a dynamic hierarchical fault diagnosis method using multi-source information from distribution automation system, wide-area measurement system and fault information system is proposed. The method includes a switch layer, a feeder layer and a substation layer, and the diagnostic strategies can be dynamically adjusted according to the characteristics of the fault information. The purpose of this study is to enrich the analysis algorithm, promote the improvement of automation application level, and create research application value.

2. Literature review

The fault detection and diagnosis (FDD) methods have evolved along three major paradigms: physics-based, knowledge-based and data-driven approaches (Pattipati et al., 2008). The physics-based approach uses the underlying physical principles of a system in the form of mathematical models. This method exploits consistency checks (residuals) between the outputs of a physics-based model and the recorded measurements. There are two main methods to generate the consistency checks: 1) observers (e.g., Kalman filters, reduced-order unknown input observers, particle filters, interacting multiple models), and 2) parity relations (dynamic consistency checks among measured variables) (Patton et al., 2013). It is expected that anomalies result in inconsistencies between the observed and the expected system behaviors. If the model is satisfactory, then its behavior resembles the actual system; indeed, such models can be beneficial in anticipating incipient faults at the design stage. This approach is useful for small- and medium-scale systems where system complexity is fairly moderate. The knowledge-based approach incorporates qualitative knowledge about the system and an inference mechanism for FDD. It uses rule-based expert systems, case-based reasoning, Bayesian networks, and graphical models, such as Petri nets, multisignal [multifunctional flow graphs, dependence graphs (digraphs)], cause–effect directed graphs, and symptom trees for diagnostic knowledge representation and diagnostic inference, (Luo et al., 2005). This method is applicable where expert knowledge about the system is available. If the system monitoring data for both normal and degraded

Please cite this article as: Ji C., Lu S., 2018, Application of improved rough set reduction algorithm in on-line fault diagnosis of chemical equipment, Chemical Engineering Transactions, 71, 1285-1290 DOI:10.3303/CET1871215

conditions are available, a data-driven approach is preferred. This method does not require a mathematical model of the system.

Rough intensiveness is a new theory of data analysis, which is often used to deal with fuzzy and uncertain problems. Based on the analysis of classical knowledge acquisition methods in rough set theory, an improved attribute importance heuristic reduction algorithm is proposed. The algorithm corrects the difference matrix and redefines the calculation method of attribute importance. Through experimental verification, the algorithm can mine high reliability diagnosis rules from the existing historical diagnosis knowledge and expert knowledge. This method can provide a reasonable basis for fault diagnosis (Zhao et al., 2017). Yao et al. proposed a rough set attribute reduction algorithm based on class discrimination matrix and improved attribute importance. Considering the correlation between different condition attributes and class attributes, an attribute weighted naive Bayesian classifier algorithm based on entropy weight method is proposed. The accuracy of fault classification is improved. Rolling bearing fault data is compared. The proposed hybrid approach has certain advantages in the accuracy and rapidity of fault diagnosis (Yao et al., 2015). The passenger train door is the key system for the operation and maintenance of urban rail vehicles. Lin et al. used the method of parameter estimation to find the physical parameters of the door under different working conditions. Then, the principal component analysis and rough set theory are used to conduct fault diagnosis experiments on the passenger door. The accuracy of fault diagnosis under different opening and closing time settings was verified by the test bench (Lin et al., 2014). Che et al. proposed a data mining distribution network fault location correlation analysis model based on rough set (RS) theory and immune algorithm (IA) to solve the correlation analysis process of distribution network fault location, due to the feeder terminal unit (FTU) The problem of variation of fault information caused by poor operating environment, damaged components or lost information (Che et al., 2017).

In order to deal with the rich information in fuzzy clustering, Yang et al. proposed a fault diagnosis method based on rough clustering and fuzzy clustering. In the fuzzy equivalence relation matrix, the number of iterations is reduced and the number of clusters is reduced. Finally, the effectiveness of the method is verified by an example (Yang et al., 2014). Rough set theory and grey theory have the same advantages in dealing with inaccurate data. Therefore, a fusion algorithm based on rough set theory and grey theory is proposed. The attribute reduction algorithm of rough set theory can simplify the fault diagnosis decision table. Grey theory can be used to predict faults based on a new reduction decision table. Through the verification of a certain type of aeronautical radio equipment, the results show that the accuracy of fault prediction is high, which provides a basis for improving the reliability and maintainability of the equipment (Song et al., 2014). Gan et al. introduced variable-precision rough set theory into Chemical transformer fault diagnosis. The implicit information in the fault data of the Chemical transformer is reduced by the reduction method of the variable precision rough set. Information that plays a major role in fault classification was obtained. The accuracy of the fault diagnosis was improved. The proposed method has better diagnostic performance (Gan et al., 2013). In order to solve the fault diagnosis of oil-filled transformers, a fast fault cause identification method based on directed acyclic graph-support vector machine (DAG-SVM) differential rough set attribute reduction is proposed. Zhu et al. proposed a fault diagnosis method based on C4.5 combined decision tree is proposed. The C4.5 decision tree is used as the basic unit to construct the fault network model for fault diagnosis of track circuits. The rough set attribute reduction method is given to simplify the condition attributes for each network element. Therefore, the C4.5 training efficiency is improved to obtain the diagnosis rule. The step-by-step compression diagnosis range can finally identify 14 common failure modes. The simulation test results show that not only the test accuracy of the coarse fault range is higher than the BP neural network method, the fault mode. The diagnostic accuracy after segmentation has also improved significantly (Zhu et al., 2018).

The example analysis shows that the method has feature vectors, which can effectively improve the efficiency of fault diagnosis and improve the accuracy of fault diagnosis in system detection (Meng et al., 2017). Luan et al. established r based on historical data and corresponding fault types. Then, the state attribute data is discretized by the equal-frequency segmentation method, and the decision matrix is distinguished by the difference matrix in the rough set theory. The low-recognition redundancy attribute is removed by establishing a diagnostic rule between each type of fault. Finally, the support vector machine is used to construct the DAG-SVM fault of the multi-class diagnostic classifier (Luan et al., 2016). The rough set fuzzy neural network is used to monitor the overcurrent fault of the motor drive system of the mining scraper conveyor. The phase current signal is input to the neural network. Then, the fuzzy logic set theory is used to process the current signal for optimization. The results show that this method can improve the accuracy and reliability of the motor drive system of mining scraper conveyor (Zhang et al., 2013). Through the reasonable division of attribute intervals, the correlation of various attributes of the disease was found (Luo, 2015).

In summary, rough intensive applications are widely used in fault diagnosis in various fields. There are also related application studies in Chemical Equipment. Thus, aiming at the problem of low efficiency of vegetable leaf image data mining, the rough set algorithm is further processed. Based on previous studies, the rough intensive algorithm is improved and applied to the field of Chemical Equipment online fault diagnosis to complete the research work.

3. Methods

3.1 System architecture

The system architecture is shown in Figure1.

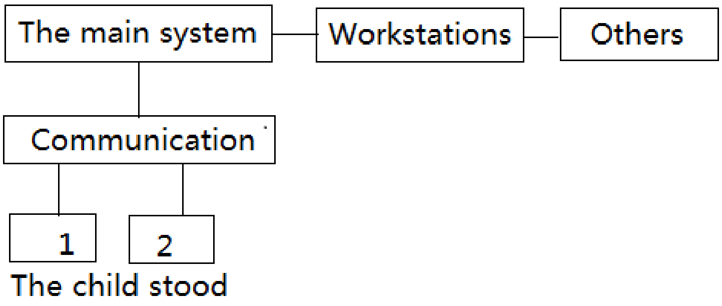

Figure 1: Automation system structure

3.2 Advanced application architecture pattern

In the expert system, expert knowledge and field experience in various fields are applied to form a knowledge base. Through designing inference control strategies for reasoning and judgment, the expert system simulates the decision process of human experts to solve complex problems that require expert decision. The typical application of expert system in fault diagnosis and load forecasting etc. in the monitoring system is based on the production rule using the depth-first strategy system. The pattern recognition method and the expert system complement each other to form an expert system that meets various requirements. The flow chart is shown in Figure 2.

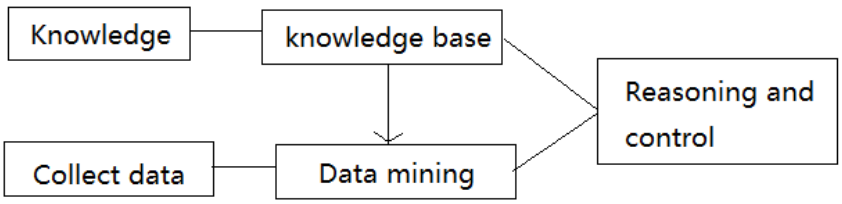

Figure 2: Expert system working mode

The reasoning technology of expert system is to use the computer for machine thinking, realize the intelligent operation of reasoning, and use the knowledge of the grid structure and the collected grid information to judge the faulty component or the Chemical outage area. Expert domain knowledge, network structure knowledge and system domain knowledge are expertly summarized to form an expert knowledge base, and data information is entered into the data warehouse through data extraction and data cleaning; then, the professional information of the expert knowledge base is used for data mining, and the results of data mining and the information of the expert knowledge base are comprehensively analysed; finally, by inference control, the reasoning conclusion is obtained.

3.3 Chemical Equipment fault diagnosis method

The main task of Chemical Equipment network fault diagnosis is to analyse the alarm signal generated by various intelligent devices in the distribution network, the state change information of switches and circuit breakers, and the characteristics of electrical quantity measurement values in the Chemical grid; it also infer possible fault location, faulty equipment and fault types according to the logic of protection actions and the experience of the operating personnel. The fault process is shown in Figure3.

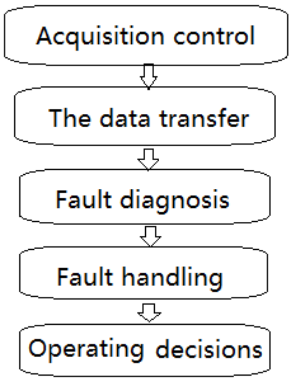

Figure 3: Chemical failure process

Chemical Equipment network fault diagnosis, isolation and recovery of Chemical supply are the important application functions of distribution automation and management system. It must monitor the status of feeder line section switch and tie switch, feeder current and voltage, and realize the remote opening and closing operation of switch. When the Chemical grid fails, the monitoring system receives the fault information, judges and removes the fault area in the shortest time, restores the Chemical supply to the non-faulted section, and minimizes the fault loss. After determining the location of the fault section, a fault analysis report is given to form a fault isolation operation sequence, so that the switches before and after the fault point are automatically separated to isolate the fault section. Based on the fault location, the Chemical supply is restored in the non-faulty section. This function can be completed at the primary station or by the distribution sub-station. After the sub-station completes the fault processing, it sends relevant information to the primary station.

3.4 Diagnosis and early warning

After the blackout in the United States and Canada in 2003, there were many Chemical outages in the Chemical system at home and abroad. The Chemical outage accidents evolved from local Chemical grid failures to large-scale Chemical outages. Therefore, research on the mechanism of preventing cascading failures and avoiding the deterioration of local faults to a large area is an important issue for the safety of Chemical grids. Through the analysis for the Chemical outages in the United States and Canada, in Italy and in Europe, it was found that the operating conditions under actual accident conditions have not been studied. The reason is that during the fault a large amount of data flooded into the dispatching end, making the dispatcher overwhelmed and unable to predict the dangers caused by the load transfer to the system, and then no effective scheduling measures are taken, so that the load capacity of the grid under the influence of multiple faults is deteriorated. In this situation, the Chemical flow is not balanced by the overload transfer, and the system runs to the critical point, which ultimately leads to the system voltage collapse. The fault diagnosis mode is shown in Figure4.

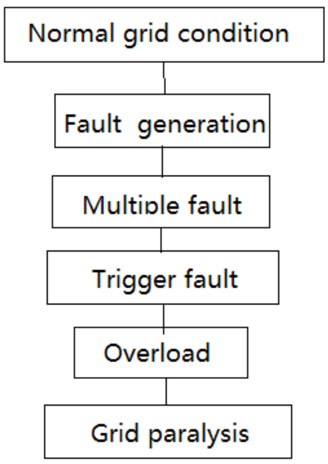

Figure 4: Outage process

4. Results and analysis

The decision table for constructing grid fault diagnosis and prediction functions, is similar to the construction of distribution network fault diagnosis decision table. The fault diagnosis decision information of the distribution network is relatively simple, and the condition attributes of the automation system are easily matched with the Chemical Equipment automation system. However, the problem of data source is inevitable in the fault diagnosis of large Chemical grids. The different construction of Chemical automation systems in different places makes the system diversified, resulting in complex fault information and transmission time problems, affecting the real-time requirements of the system, and making the condition attributes selection and processing complicated. Thus, the correlation between lines is difficult to define uniformly. The fault cause for grid cascading failure is taken as a condition, and the failure result is used as a decision to establish a fault diagnosis and predictive analysis model. These faults occur one after another, not accidentally. They are related to each other, and the degree of influence between the two faults depends on the degree of interaction between electrical equipment. In terms of electrical engineering, the degree of equipment components depends on two aspects, that is, the Chemical change of other components caused by the unit Chemical variation of one component, and the Chemical bearing degree of this component. The algorithm analysis structure is shown in Figure5.

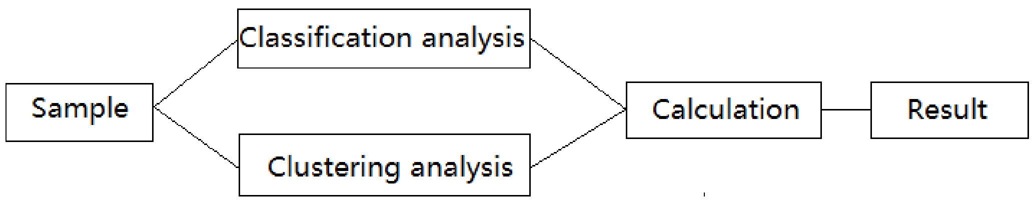

Figure 5: Methods the structure

The characteristics of ultra-short-term, short-term, medium-term and long-term load forecasting vary. The data of ultra-short-term load forecasting is usually used for real-time security analysis and real-time economic dispatching. Its data change law is consistent with the change of similar time period, so it is usually obtained by the method of similar day. Its changing factors include real-time meteorological conditions, the impact of accidents, and so on. The data of short-term load forecasting is usually used to formulate the daily operation mode, which has obvious periodicity in different periods of month, week and day. The long-term load forecasting data is used to design the Chemical supply point and the primary grid structure, and change the fuzzy situation of the Chemical grid planning. Besides, long-term load forecasting data is also required when determining the annual maintenance plan and annual operating mode. The characteristics of long-term load forecasting data are monotonous changes, with no periodicity, and the law of change is related to the local economic development, population, industrial structure, electricity price policy, and national economic status, e.g., the sample classification mode shown in Table 1.

Table 1: Sample classification model based on experience

The serial number	1	2	3	4	5	6
Describe	New Year's Day	Christmas	Monday	Tuesday	...	Sunday

Classification and clustering are two important research areas in data mining. Classification requires related prior knowledge, and clustering often seeks its intrinsic features from the data itself according to certain similarity measure. In practical applications, there is usually an asymmetry between the classification results and the clustering results. For this problem, this paper gives the definition of harmonic matrix. On this basis, it proposes a harmonic clustering classification algorithm to keep the classification result and clustering in high consistency, and apply this method to load forecasting of Chemical system.

5. Conclusions

As an important research field, data mining has been used more widely. It involves the structure, storage and purpose of different data, and the existing theories and methods cannot effectively solve the practical problems. In the Chemical system load forecasting, there are various load forecasting model algorithms, which are inseparable from the load sample data, so selecting the load sample data is especially important for the accuracy of load forecasting. This paper analyses the basic load model of load forecasting, and uses the harmonic clustering and classification algorithm to select the sample space of daily load data, which improves the accuracy of sample space. In the selection of electrical load samples, this is more effective than conventional methods.

Reference

Che Y.B., Yu S.Y., Ge L.J., 2017, Fault locating model based on RS-IA data mining for distribution network, Electric Chemical Automation Equipment, 37(5), 122-128, DOI: 10.16081/j.issn.1006-6047

Gan S.C., Zhou A.H., Guo H., 2013, Application of Variable Precision Rough Set in Chemical Transformer Fault Diagnosis, Applied Mechanics & Materials, 23(14), 373-375, 824-828, DOI: 10.4028/www.scientific.net/amm.373-375.824

Pattipati K., 2008, An Integrated Diagnostic Process for Automotive Systems, Berlin, New York, NY, USA: Springer-Verlag, 191–218

Luan X.Y., Li Z.P., Liu T.Z., 2016, A novel attribute reduction algorithm based on rough set and improved artificial fish swarm algorithm, Neurocomputing, 174, 522-529, DOI: 10.1016/j.neucom.2015.06.090

Luo X.L., 2015, Improved Algorithm Based on Item Code and Rough Intensive Reduction Algorithm in the Application of Disease of Vegetable Leaf Image Mining, Applied Mechanics and Materials, 713-715, 4, DOI: 10.4028/www.scientific.net/amm.713-715.1733

Meng H., Haiyan X.U., Song X., 2017, Transformer Fault Diagnosis Based on Attribute Reduction of Rough Set and SVM, Journal of Nanjing University of Aeronautics & Astronautics, 49(4), 504-510, DOI: 10.16356/j.1005-2615.2017.04.008

Patton R.J., Frank P.M., Clark R.N., 2013, Issues of Fault Diagnosis for Dynamic Systems, New York, NY, USA: Springer-Verlag.

Song Z.Y., Zhang G.Y., Su Y.Q. 2014, One Fusion Algorithm of Equipment Fault Prediction Based on Rough Set Theory and Grey Model, Applied Mechanics & Materials, 687-691, 1377-1379, DOI: 10.4028/www.scientific.net/amm.687-691.1377

Yang S.J., Li J.H., Liu X.N., 2014, One Fault Diagnosis Method Based on Fuzzy Equal Relationship and Rough Set Theory, Applied Mechanics & Materials, 631-632, 175-179, DOI: 10.4028/www.scientific.net/amm.631-632.175

Yao C., Nan L., Feng Z., 2015, Fault Diagnosis Based on Rough Set Attribute Reduction and Bayesian Classifier, China Mechanical Engineering, 26(14), 1969-1977, DOI: 10.3969/j.issn.1004-132x2015.14.022

Zhang Y.Q., Ma X.M., Yang J.X., 2013, Rough Set Fuzzy Neural Network Fault Diagnosis for the over Current Detection of Coal Mining Scraper, Applied Mechanics and Materials, 321-324, 6, DOI: 10.4028/www.scientific.net/amm.321-324.2146

Zhao R., Li C., Tian X., 2017, A novel industrial multimedia: rough set-based fault diagnosis system used in CNC grinding machine, Multimedia Tools and Applications, 76(19), 19913-19926, DOI: 10.1007/s11042-016-3878-0

Zhu W.B., Wang X.M., 2018, Research on Fault Diagnosis of Railway Jointless Track Circuit Based on Combinatorial Decision Tree, Journal of the China Railway Society, 40(7), 74-79, DOI: 10.3969/j.issn.1001-8360