Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 5, pp. 3580–3586, DOI 10.12694/scpe.v25i5.3173

APPLICATION OF CLUSTER ANALYSIS ALGORITHM IN SUPPLY CHAIN RISK IDENTIFICATION

QINGPING ZHANG* AND YI HE†

Abstract. The risk control model of the power supply chain system is established. A fault information identification method based on fuzzy clustering is proposed. This method fully considers the power grid's characteristics and uses terrible data. A risk assessment model based on fuzzy set theory is established by the COWA operator weight method and grey cluster evaluation method. The security risk identification model of power grid enterprises uses insufficient data. The security risk identification data are normalized and classified. Empirical analysis determines various risk factors that may appear in power projects. The applicability and feasibility of the index system and evaluation model are verified.

Key words: Fuzzy clustering; Power supply chain; Security risk monitoring; Risk identification

1. Introduction. The rapid development of the electric power industry can benefit the people and greatly promote the development of the national economy. In the early stage of development, because the power industry has a high degree of monopoly, many operating entities such as power generation, electricity sales and transmission are concentrated in one company. This results in an industrial governance model similar to the corporate governance model, which has a high degree of monopoly but also causes a lot of resource waste, low efficiency, and high operating costs. At the same time, there are problems such as rent-seeking, price discrimination and network barriers. China has launched a series of power system reforms and market-oriented plans by introducing market competition to change the monopoly situation of the whole industry. This can reduce the operating cost and realize the rationality of resource allocation. It can not only effectively promote market competition but also reduce the waste of energy. The domestic power grid has separated power plants from the grid and established a perfect quotation model for power generation and other industrial sectors. It has broken through the monopoly model of the past and established a complete power grid supply chain. However, some potential security risks cannot be ruled out in the domestic power supply chain. Against this background, it is an important research direction for power network security risk monitoring.

The FMEA model was analyzed in the literature [1]. A complex diffusion network model of fault mode is established by using a complex network analysis method. The influence of the correlation between fault modes and fault modes is studied. Taking a typical supply chain as the research object, FMEA and complex network methods are studied to test their application value in fault-type evaluation. Literature [2] proposed the generation mechanism of risk factors in the construction stage of hydropower projects based on the perspective of the supply chain. The sub-indexes of 5 categories and 18 categories were established. The final risk assessment value is consistent with the actual risk status of the project. The effectiveness and feasibility of this method are proven. Literature [3] constructs the framework of an inter-provincial power trading system based on RPS. Clarify power supply and demand issues among various market participants. The customer's subjective choice is introduced, and the utility function describes the customer's purchasing behavior. The optimal decision problem of multiple trade participants in the supply chain based on maximum return is constructed. By using the method of reverse derivation, the optimal trade decision problem of each trader is solved [4]. It has particular reference significance to the cross-province power market in our country. However, the power supply in these ways is unstable. Therefore, this paper uses the fuzzy clustering method to monitor the security risks of each

*Business School, Shunde Polytechnic, Foshan, Guangdong, 528333, China (Corresponding author, 10370@sdpt.edu.cn)
†School of Economics and Management, Guangdong Vocational College of Post and Telecom, Guangzhou, Guangdong Province, 510630, China

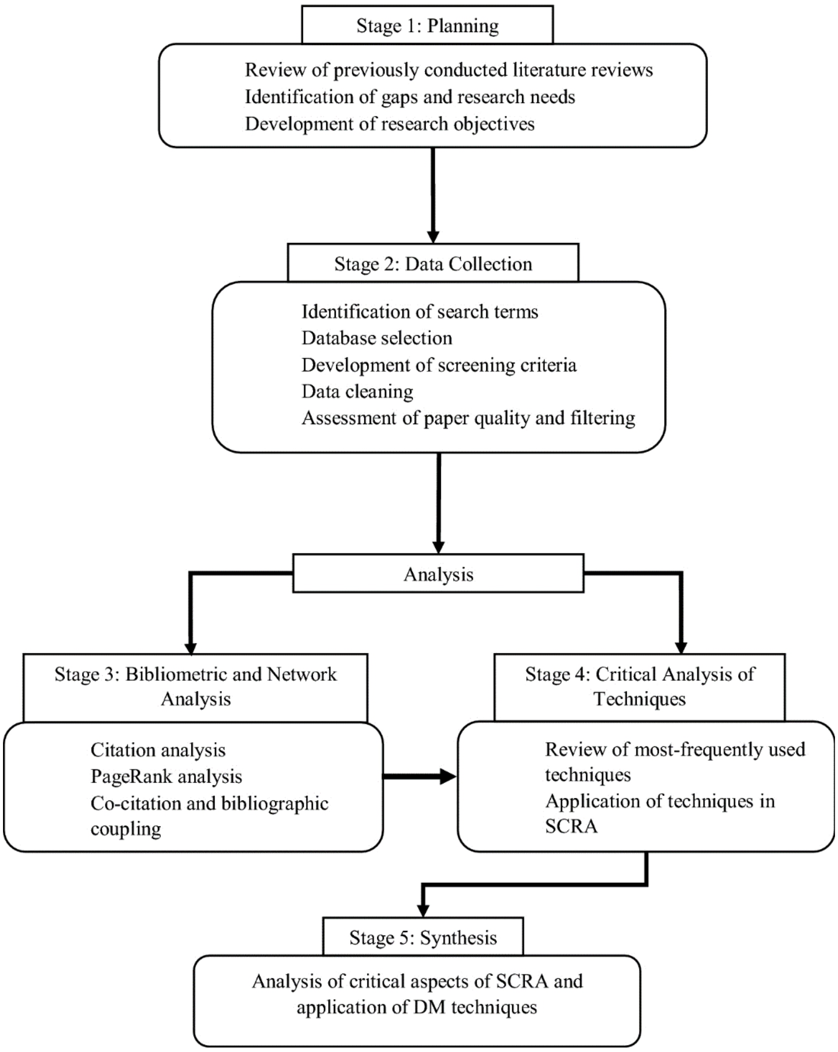

Fig. 2.1: *Specific construction process.*

link enterprise in the power grid to achieve the optimal allocation and supply chain management of each link enterprise in the power grid. The aim is to improve the stability of the power grid.

2. Fuzzy cluster monitoring in the power supply chain.

2.1. Security Risk Identification. The wrong data of the power grid system is analyzed according to the characteristics of the power grid itself based on the identification of power grid fault information. The security risks of power grid enterprises can be effectively identified [5]. The power grid operation's risk identification mechanism is obtained. First, the security of insufficient data must be collected and classified. The security risk index of the power supply chain system based on the network is constructed. Then, a preliminary framework of a fuzzy comprehensive evaluation system is constructed [6]. The detailed construction process is shown in Figure 2.1 (image cited in Annals of Operations Research, 2023, 322(2): 565-607). Through the identification of security risks of power grid enterprises, they are divided into internal and external categories. These two risk factors are listed in Table 2.1.

Table 2.1: *Safety risk factors.*

	Evaluation index	Level 1 security risk factor	Secondary security risk factors
Power supply chain security risk factors	Natural hazard index The closeness of national policy Industrial climate index Overall economic condition Customer satisfaction Inventory turnover Profit margin on sales Production cost Research and development phase Product flexibility Order fulfillment rate Qualified rate of finished product Order preparation time	External security risk of the power supply chain Internal security risk of the power supply chain	Ecological environment risk Policy and regulation risk The risk of public security hazards Economic and environmental risks Seller's security risks Manufacturer's safety risk Supplier security risk

2.2. Security risk monitoring. After obtaining the relevant information on the risk factors inside and outside the electricity market, the next step is to process the data [7]. The first step is to limit the scope of the relevant data. Monitoring the power grid security risk is to normalize and classify the collected information. The main contents include daily power threshold security risk monitoring, harmonic limit security risk monitoring, and parameter setting security risk monitoring. The security risk monitoring of the daily load threshold is mainly used to monitor the stability and security risks of the power grid load during the power grid operation. Power grid side security risk monitoring is mainly through the power grid voltage level and change amplitude to monitor the power grid operation [8]. If the voltage level and the range of change cannot meet the electricity demand, it must be modified. A power metering device is designed to measure the running state of the power system. If the parameter of the meter is set to 0.2, then the allowable deviation of the meter must be controlled within 0.2%. If the end indicator of the meter exceeds the control value, it indicates a security risk for the customer's meter in the entire grid. The power data collected during power grid operation is a vital link. It also includes the power meter precision, power rate, and other parameters set by the security risk monitoring function [9]. Prevent users from arbitrarily changing the user's power parameters.

3. A comprehensive risk assessment model of a power supply chain is established by combining C-OWA with grey clustering.

3.1. The COWA algorithm is used to calculate the weights of each evaluation index. The essence of the OWA algorithm is to sort the data incrementally, and the weighting depends only on the space. A ranked, weighted mean C-OWA algorithm is proposed. The detailed process is like this: 1) An expert is invited to evaluate the importance of the evaluation indicators at all levels, and the scores constitute the initial evaluation data set $(\eta_1, \eta_2, \cdots, \eta_i, \cdots, \eta_n)$ of the evaluation indicator H, and $\theta_0 \geq \theta_1 \geq \theta_2 \geq \cdots \theta_j \geq \cdots \theta_{n-1}$ is obtained from o in the order from high to low. 2) The weighting λ_{j+1} of data θ_i is directly determined by the number of combinations Z_{n-1}^{j} :

$$\lambda_{j+1} = \frac{Z_{n-1}^{j}}{\sum_{t=0}^{n-1} Z_{n-1}^{t}} = \frac{Z_{n-1}^{j}}{2^{n-1}}, j = 0, 1, 2, \text{ L}, n-1$$

Formula: $\sum_{j=0}^{n-1} \lambda_{j+1} = 1$. 3) Weight the evaluation data with weight vector λ to obtain the absolute weight $\bar{\delta}$ of index η_i :

$$\bar{\delta} = \sum_{j=1}^{n} \lambda_j \cdot \theta_j \in [0, 1], j \in [1, n]$$

4) Calculate the relative weight value δ_i for the exponential factor η_i

$$\delta_i = \frac{\bar{\bar{\delta}}}{\sum_{i=1}^{m} \bar{\delta}_i}, i = 1, 2, \cdots, m$$

3.2. Grey cluster analysis of the overall risk of the power supply chain. Many factors affect the integration risk of the power supply chain, and most of the existing evaluation methods rely on experts' subjective experience, knowledge level and subjective preferences [10]. So, the grey clustering method is used to study the risk of the power supply chain.

3.2.1. Grey category judgment and whitening weight. The complexity of risk assessment indicators will directly affect the refinement of grey categories [11]. It is assumed that the target to be assessed is divided into s grey categories, and the interval of its secondary index is also divided into s grey categories. It is divided into very low [2,0], low [4,2], average [6,4], high [8,6], and high [10,8]. People can get the point vector U= (9, 7, 5, 3, 1) by applying the traditional grey system theory.

3.2.2. Grey cluster evaluation steps. 1) Construction of evaluation model. This paper classifies the risk degree of the power supply chain based on q class of power industry experts, supply chain research experts, mahagement experts and construction experts [12]. Then the evaluation matrix $S_i = [s_{ijt}]_{s\times q}$ is constructed according to the score of index $H_{ij}.s$ is the number of exponents of the matrix. 2) Construction of grey cluster weight matrix. The clustering factor of Grade H_{ij} and Grade e gray level is grade $U_{ije} = \sum_{n=1}^{q} g_e [s_{ijt}]$, and the overall level evaluation factor is grade $U_{ij} = \sum_{e=1}^{5} U_{ije}$. Then the weight vector of the gray cluster can be calculated as $c_{ije} = \frac{U_{ije}}{U_{ij}}$, and the weight matrix of the gray cluster can be obtained as follows

$$C_i = \begin{bmatrix} c_{i11} & c_{i12} & c_{i13} & c_{i14} & c_{i15} \\ c_{i21} & c_{i22} & c_{i23} & c_{i24} & c_{i25} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{ij1} & c_{ij2} & c_{ij3} & c_{ij4} & c_{ij5} \end{bmatrix}$$

3) Comprehensive cluster evaluation matrix. Evaluate each significant index cluster according to formula (3.5):

$$V_i = \delta \cdot C_i$$

Take $V_0 = [V_1, V_2, \cdots, V_n]^T$ as the comprehensive evaluation matrix of the above index, and then use formula (3.6) to conduct a comprehensive cluster evaluation of the index:

$$Y = \delta \cdot V_0 = [Y_1, Y_2, \cdots, Y_n]$$

4) Synthesize the evaluation values at all levels. Formula (3.7) is used to integrate weight Y and weight O to obtain the risk level of the power supply chain to prevent secondary losses in the evaluation process.

$$\Lambda = Y \cdot O^T$$

4. Experimental detection.

4.1. System Validity Check. A fault information identification method based on fuzzy clustering is proposed. Data was processed using the Xon (R) Server with Windows Server 2012R2 [13]. The network topology of the power system is established, and the total load is obtained. A risk identification model based on load level is proposed. Different test schemes are compared and analyzed [14]. The 46,890 labeled samples were classified. Among them, 70% are taken as training samples, 10% as confirmation samples and 20% as test samples (Table 4.1).

The correctness of the proposed algorithm is verified by comparing it with fuzzy clustering, k nearest neighbor classification, SVM, convolutional neural network and recurrent neural network. Each class of algorithms is based on a training set. Run ten times in one test set. Finally, the average of 10 samples is the final prediction

Table 4.1: Data set division unit.

Training set	Validation set	Test set
34191	4884	9769

Table 4.2: *Accuracy rates of different models.*

Model	Accuracy rate / %
Fuzzy clustering method	78.96
KNN	79.48
SVM	85.83
CNN	89.38
LSTM	90.21
CPE	93.54

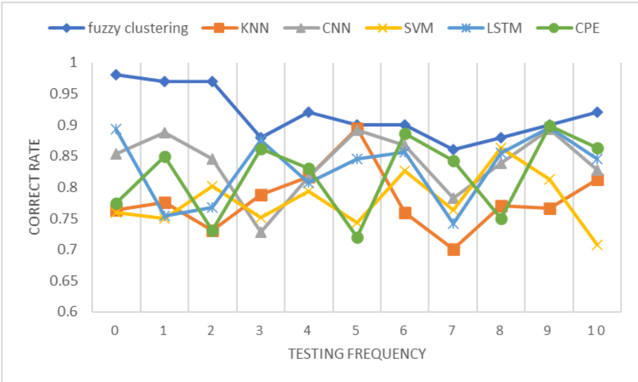

Fig. 4.1: *Test accuracy.*

accuracy [15]. Table 4.2 shows that the mathematical model proposed in this paper has good test results. It has advantages over classical machine learning algorithms such as fuzzy clustering, KNN, and support vector machines. Performance is better than CNN LSTM.

The model was trained during the experiment. Each trained model is run ten times on the same data set. The method is predicted ten times, and the final prediction accuracy is obtained [16]. The following conclusions are drawn through the analysis of the experimental data: 1) The accuracy of the fuzzy cluster analysis model in this paper can reach 89.8% by comparing the accuracy of different models. This algorithm exceeds the conventional machine learning algorithm and is better than the widely used deep learning algorithms such as neural networks and LSTM. Because the initial value limits the selection of the KNN initial value, it isn't easy to obtain an ideal initial value. However, support vector machines have substantial limitations in selecting kernel functions and are unsuitable for large-scale data. Although CNN has a good feature extraction function, there is often an aggregation process after extraction. And through pooling, resulting in more significant information loss. The short-term memory method is ineffective in feature extraction [17]. The clustering method has the advantages of better feature extraction, not being easy to lose, not being affected by the initial value, and having a greater demand for data. 2) As seen in Figure 4.1, the variance of the cluster analysis model is the least, while the variance of other models is more significant, indicating that the clustering method is robust. 3) The weights of high-level hazard factors and low-level factors are shown in Figure 4.2. The results show that the factors causing harm are the largest in the distribution system, and the planning links occupy a large proportion.

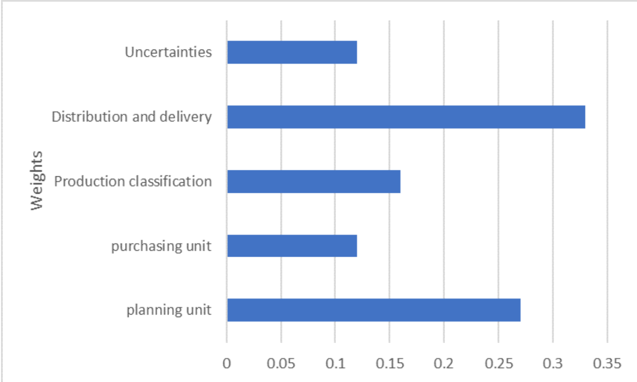

Fig. 4.2: *Risk weight values.*

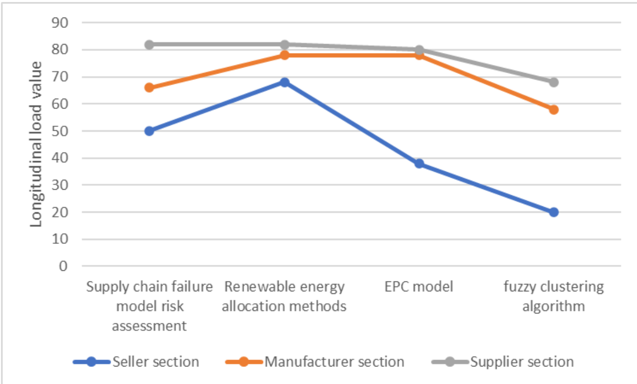

Fig. 4.3: *Comparison of power supply chain stability.*

4.2. System stability analysis. The supply chain failure mode risk assessment method, EPC mode method, renewable energy quota method and other methods are combined with this method, and the fuzzy cluster analysis method is compared [18]. The results show that the model in this paper can reflect the stability of the power grid well (Figure 4.3).

A risk assessment model based on fault types of supply chain is proposed. The resulting power grid stability is about 72.2% [19]. The simulation results show that the stability of the model is about 42.6%. The results show that the proposed algorithm has good stability. The fundamental reason is that the research idea proposed in this project is based on identifying insufficient data and integrating it with the characteristics of the power grid to identify the security risks of power grid enterprises. It guarantees the stability of the supply chain.

5. Conclusion. The fuzzy clustering model of power supply chain system risk is established to ensure the high stability of the power grid. This research result has important practical significance for developing China's power market. It is also relatively easy to implement. Its implementation process is simple and time-consuming is short, so it has good promotion value.

REFERENCES

[1] Shishodia, A., Sharma, R., Rajesh, R., & Munim, Z. H. (2023). Supply chain resilience: A review, conceptual framework and future research. The International Journal of Logistics Management, 34(4), 879-908.
[2] Babu, H., Bhardwaj, P., & Agrawal, A. K. (2021). Modelling the supply chain risk variables using ISM: a case study on Indian manufacturing SMEs. Journal of Modelling in Management, 16(1), 215-239.

[3] Singh, S., & Kumar, K. (2021). A study of lean construction and visual management tools through cluster analysis. Ain Shams Engineering Journal, 12(1), 1153-1162.
[4] Han, Y., Yan, X., & Piroozfar, P. (2023). An overall review of research on prefabricated construction supply chain management. Engineering, Construction and Architectural Management, 30(10), 5160-5195.
[5] Zaridis, A., Vlachos, I., & Bourlakis, M. (2021). SMEs strategy and scale constraints impact on agri-food supply chain collaboration and firm performance. Production Planning & Control, 32(14), 1165-1178.
[6] Li, G., Kou, G., & Peng, Y. (2021). Heterogeneous large-scale group decision making using fuzzy cluster analysis and its application to emergency response plan selection. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 52(6), 3391-3403.
[7] Ivanov, D. (2023). The Industry 5.0 framework: Viability-based integration of the resilience, sustainability, and human-centricity perspectives. International Journal of Production Research, 61(5), 1683-1695.
[8] Yang, J., Xie, H., Yu, G., & Liu, M. (2021). Antecedents and consequences of supply chain risk management capabilities: An investigation in the post-coronavirus crisis. International Journal of Production Research, 59(5), 1573-1585.
[9] Chen, Z. S., Zhang, X., Rodríguez, R. M., Pedrycz, W., Martínez, L., & Skibniewski, M. J. (2022). Expertise-structure and risk-appetite-integrated two-tiered collective opinion generation framework for large-scale group decision making. IEEE Transactions on Fuzzy Systems, 30(12), 5496-5510.
[10] Akbari, M., & Do, T. N. A. (2021). A systematic review of machine learning in logistics and supply chain management: current trends and future directions. Benchmarking: An International Journal, 28(10), 2977-3005.
[11] Kang, H., & Jung, E. H. (2021). The smart wearables-privacy paradox: A cluster analysis of smartwatch users. Behaviour & Information Technology, 40(16), 1755-1768.
[12] Neuburger, L., & Egger, R. (2021). Travel risk perception and travel behaviour during the COVID-19 pandemic 2020: A case study of the DACH region. Current issues in tourism, 24(7), 1003-1016.
[13] Budianto, E. W. H. (2023). Research Mapping on Credit Risk in Islamic and Conventional Banking. AL-INFAQ: Jurnal Ekonomi Islam, 14(1), 73-86.
[14] Um, J., & Han, N. (2021). Understanding the relationships between global supply chain risk and supply chain resilience: the role of mitigating strategies. Supply Chain Management: An International Journal, 26(2), 240-255.
[15] Akın Ateş, M., Suurmond, R., Luzzini, D., & Krause, D. (2022). Order from chaos: a meta-analysis of supply chain complexity and firm performance. Journal of Supply Chain Management, 58(1), 3-30.
[16] Izaguirre, C., Losada, I. J., Camus, P., Vigh, J. L., & Stenek, V. (2021). Climate change risk to global port operations. Nature Climate Change, 11(1), 14-20.
[17] Xu, X., Zhu, D., Yang, X., Wang, S., Qi, L., & Dou, W. (2021). Concurrent practical byzantine fault tolerance for integration of blockchain and supply chain. ACM Transactions on Internet Technology (TOIT), 21(1), 1-17.
[18] Matiza, T., & Kruger, M. (2021). Ceding to their fears: A taxonomic analysis of the heterogeneity in COVID-19 associated perceived risk and intended travel behaviour. Tourism Recreation Research, 46(2), 158-174.
[19] Modgil, S., Gupta, S., Stekelorum, R., & Laguir, I. (2021). AI technologies and their impact on supply chain resilience during COVID-19. International Journal of Physical Distribution & Logistics Management, 52(2), 130-149.

Edited by: Zhigao Zheng
Special issue on: Graph Powered Big Aerospace Data Processing
Received: Jan 10, 2024
Accepted: Feb 18, 2024