Risk Assessment of Grid Communication System Based on MIC-FNN under SG and UPIOT

Yunli Yue\textsuperscript{1*}, Xuefeng Jia\textsuperscript{2} and Cunbin Li\textsuperscript{2}

\textsuperscript{1}Economic and technological research institute, State Grid Jibei Electric Power Corporation, 21 East Yangfangdian Rd., Haidian District, Beijing, 100038, China

\textsuperscript{2}School of Economic and Management, North China Electric Power University, Beijing, 102206, China

*Yunli Yue's e-mail: 39213654@qq.com

Abstract. The communication system is the carrier of grid and information network coupling. The construction of reliable Smart Grid and Ubiquitous Power Internet of Things depends on the high coupling between the power grid and the information network. With the development of "the Grid and Internet", the risk assessment of the power grid communication system has the characteristics of various levels, complex indexes and strong correlation. To deal with the risk effectively, this paper proposes an index system for evaluating the operation risk of grid communication system, which includes the cable loss level, the communication network power supply risk, the equipment operation risk, the external environment risk, the business interruption risk and the management risk. At the same time, a risk assessment model is built based on a fuzzy neural network, which overcomes the problem that traditional risk assessment methods rely too much on expert experience and makes prior knowledge embedded in the network. Finally, the model is proved to be scientific and practical by an example analysis.

1. Introduction

The development of Smart Grid (SG) and Ubiquitous Power Internet of Things (UPIOT) make the power system become a Cyber-Physical System (CPS). The security and stability of the power system depend on the smooth operation of the communication system. Therefore, the risk assessment of the Smart Grid Communication System (SGCS) has received a lot of attention. Risk assessment can reduce the risk loss of the grid and maintain the safe and stable operation of the grid. SGCS research methods mainly include the following two.

First, traditional SGCS risk studies focus on estimating the state of the power system. Reference [1] proposed a new fuzzy taxi method for risk assessment of communication networks. Reference [2] proposed an improved method based on FAHP to evaluate the risk of the system. In reference [3], graph theory is applied to study the defense mechanism to prevent the attack of wrong data injection. There are many other methods used to detect erroneous data injection attacks, such as tracking the dynamics of measurement changes [4], $\chi^2$ detector and cosine similarity matching methods [5], and nonlinear grid models [6]. However, such methods rely too much on the empirical judgment of experts and cannot avoid the existence of subjective bias.

Secondly, with the development of Machine learning (ML), more and more in-depth learning algorithms are used for risk assessment of SGCS. Reference [7], [8] provides a strategy at the system level to assess and mitigate cascading outages of power systems, taking into account the possibility of...
hidden faults in the protection system that affect the risk of cascading outages. References [9] analyzed and compared conventional risk index prediction methods, compared the advantages and disadvantages of various risk index prediction methods, combined with the actual situation of smart grid, LVQ neural network prediction method is used to classify the risk level of this grid. In reference [10], BP neural network is used to cause mountain fires on transmission lines of 220 kV and above. Reference [11] proposed a method based on the attribute reduction theory of rough set and multi-class SVM classification. However, such methods require a large amount of data and are difficult to collect.

In this paper, a fuzzy neural network model for SGCS risk assessment is proposed. This paper relies on a neural network technology to weaken the bias of expert subjectivity. At the same time, the method of fuzzy mathematics is adopted to quantify the scores of experts and increase the data volume of machine learning. The rest of this paper is as follows: section 2 lists the index system of risk assessment in a communication network; Section 3 introduces the theory of fuzzy set and BP neural network. Section 4 explains how to establish a fuzzy neural network and carry out case analysis. Section 5 draws a conclusion.

2. SGCS Risk Assessment Index System

Based on previous literature studies [12], this paper proposes six types of operational risk index including the cable loss level, the communication network power supply risk, the equipment operation risk, the external environment risk, the business interruption risk, and the management risk.

![SGCS Risk Assessment Index System](image)

(1) Cable loss level
A survey shows that most of China's optical cable service years are more than ten years, the longest up to 30 years. The aging fault will make the communication network information delay or information block, affect the safety and stability of the power system.

(2) Communication network power supply risk
The increasing of new equipment in the power communication network will cause the power load of the communication network to be too large, exceeding the power carrying capacity. Besides, whether the power supply of the communication network takes into account the emergency plan of failure is also a basis for judging the reliability of power supply.

(3) Equipment operation risk
With the progressive development of the smart grid, Synchronous Digital Hierarchy (SDH) and pulse code modulation (PCM) devices have become increasingly important in the communication system. If these two devices fail, the data in the power communication network cannot be converted into form, which will cause the impact of communication interruption.

(4) External environment risk
Power communication networks need to be tested by real-world conditions, including natural problems such as thunder, lightning, rain, snow, and strong winds. Besides, there are also human-made...
environmental problems to consider. Therefore, the environment may have an impact on the power communication network.

5) Business interruption risk
The failure of crucial communication components in the power communication network will make the operation monitoring system unable to obtain the operation information of the power grid or unable to timely and reliably issue control orders, resulting in the loss of the observability and controllability of the power grid and the occurrence of power failure.

6) Management risk
The power communication network is a kind of hierarchical network, divided into three levels of communication. Each level of the network contains an increasing number of nodes. The number of nodes in the electric power communication network increases, resulting in a gradual increase in the difficulty of management. If it cannot be managed effectively, it will affect the efficiency of the communication network and endanger the stability of the power grid.

3. MIC-FNN Theory

3.1. Maximum Information Coefficient
The maximum Mutual Information Coefficient (MIC) is developed based on Mutual Information (MI), which has substantial fairness and universality [13]. MI is a method to calculate the degree of nonlinear dependence between variables. MIC overcomes the disadvantage of MI not being able to calculate continuous variables and can better reflect the correlation between attributes.

For two attributes $X, Y$ of binary data set $D$, MIC to differentiate the grid $G$ of $X \times Y$, calculate the probability for each cell in the grid $G$, get the probability distribution of grid $D$, its most significant MI value $\text{Max} \left( I[D(x,y)] \right)$, save it as $I^* [D(x,y)]$, such as Eq (1).

$$I^* [D(x,y)] = \text{Max} \left( I[D(x,y)] \right)$$

Standardize this mutual information and calculate MIC, as shown in Eq (2) and (3).

$$M(D)_{x,y} = \frac{I^*[D(x,y)]}{\text{in min}(x,y)}$$

$$F(D)_{\text{MIC}} = \max_{xy<B(n)} \left\{ M(D)_{x,y} \right\}$$

Where, $n$ is the number of samples, and $B(n)$ is the function to represent the sample size, which marks the grid $G$ partition and puts constraint of total $xy$, generally $B(n) = n^{0.6}$. MIC is the mutual information data after normalization, and the value range is $[0,1]$. The larger the MIC value between the two variables, the stronger the correlation, and vice versa.

In this paper, out of 150 sample data, 100 groups were randomly selected to calculate the MIC of each attribute and the final risk level, and the MIC of each attribute and the final risk level of 100 groups is obtained after 30 repetitions. This is used as the basis of indicator screening, and the MIC value is eliminated if it was small.

3.2. Fuzzy Neural Network
FNN is an intelligent data processing system that combines the subjective fuzziness of human evaluation language with the intelligent classification system of the neural network [14]. FNN has the self-learning ability and adaptive ability of the primary neural network, and the fault tolerance rate is relatively high. At the same time, it can also make up for the shortcoming of the neural network that the learning rules are not clear enough. Since fuzzy language can express the uncertainty of information, the training rules of the neural network can be established through the subjective experience of experts to improve the ability of information processing and problem-solving.

FNN includes an input layer, hidden layer, fuzzy layer, analytic layer and defuzzy layer. Based on the risk assessment index system of the SGCS, the topology of FNN is shown in Fig 2:

1) Input layer. Set $x = \{x_1, x_2, ..., x_n\}$ is the input variable, representing the evaluation value of transmission risk, power quality risk, power grid frequency stability risk, island effect risk, protection and control risk, and planning and design risk. The five-level semantic scale is shown in Table 1.
Table 1 Risk scoring criteria

| Risk | VL | L  | M  | H  | VH |
|------|----|----|----|----|----|
| Score | >0.8 | 0.6-0.8 | 0.4-0.6 | 0.2-0.4 | <0.2 |

(2) Hidden layer. The hidden layer uses nodes to express language variables and membership. The set of linguistic risks is very low risk (VL), low risk (L), middle risk (M), high risk (H) and very high risk (VH). Denotes the scale value of a language variable by $\mu_1, \mu_2, \ldots, \mu_n$, where $i_1, i_2, \ldots, i_n \in \{1, 2, \ldots, 5\}, j = 1, 2, \ldots, 12$. There are 12 units on the third floor. In Fig 3, $B_i^k$ represents the scale of language variable project risk, $\mu_{B_i}^k$ represents membership, and $W_{ij}$ represents the connection weight between nodes $i$ and $j$ in the two levels. The weight size is generated randomly and modified by network training.

(3) Fuzzy layer. Fuzzy layer uses node to describe vague rules, calculates the applying rules: $a_j = \min \{\mu_{i_1}^1, \mu_{i_2}^2, \ldots, \mu_{i_n}^n\}$, or $a_j = \mu_{i_1}^1, \mu_{i_2}^2, \ldots, \mu_{i_n}^n$, where $i_1, i_2, \ldots, i_n \in \{1, 2, \ldots, 5\}, j = 1, 2, \ldots, 12$. There are 12 units on the third floor. In Fig 3, $B_i^k$ represents the scale of language variable project risk, $\mu_{B_i}^k$ represents membership, and $W_{ij}$ represents the connection weight between nodes $i$ and $j$ in the two levels.

(4) Parsing layer. The parsing layer is the process of parsing the risk with the goal of mapping the running risk to five scales by weighting. $\mu_{B_i}^k$ represents the degree of membership. Similar to the previous parameter passing rules, took the form of weighting and scaling value transfer, membership degree with operations and set patterns, take in the collection, the largest is $\tilde{a}_j = a_j / \sum_i a_i$.

(5) Defuzzy layer. Follow the most significant membership degree principle, runs a risk for a layer of 5 scales for judgment, will be the most significant membership degree scale as the actual output of the risk state $y = \sum_{j=1}^m w_{ij} a_j$. Where, $w_{ij}$ represents the connection weight between layer 4 and layer 5, $k = 1, 2, \ldots, m$.

Fig 2 FNN Topology Diagram

FNN is a kind of multi-level feedforward neural network. $w_{ik}$ and $w_{kj}$ are the connection weights between the second and third layers and the third and fourth layers, respectively. The error function is shown in Eq (4).

$$E(W) = \frac{1}{2} \sum_{j=1}^m (d_j - y_j)^2$$

Where, $d_j$ is the expected output value and $y_j$ is the actual output value. Error backpropagation is used to calculate the gradient $\partial E / \partial W_{ij}$, and a stepped degree optimization algorithm is used to calculate the adjustment amount $\delta_j$ of $W_{ij}$. The process is shown as Eq (5) and (6).

$$\delta_j = \frac{\partial E}{\partial w_{jk}^{k+1}} = \frac{\partial E}{\partial o_j^{k+1}} \cdot \frac{\partial o_j^{k+1}}{\partial w_{jk}^{k+1}} = (O_i - d_i) f'(net_j^{k+1})$$

$$\delta_j = \frac{\partial E}{\partial w_k^{k+1}} = \frac{\partial E}{\partial net_j^{k+1}} \cdot \frac{\partial net_j^{k+1}}{\partial w_k^{k+1}} = -\frac{\partial E}{\partial net_j^{k+1}} O_i^k$$

Where, $\partial net_j^{k+1}$ is the input of the fourth layer, and $\partial O_i^{k+1}$ is the input of the activation function. The $\partial E / \partial y_j$ of the fourth layer and $\delta_j^{(4)}$ are calculated by the Eq (7) and (8).

$$\delta_j^{(4)} = -\frac{\partial E}{\partial y_j} = d_j - y_j$$
\[
\frac{\partial E}{\partial W_{kj}} = \frac{\partial E}{\partial y_j} \cdot \frac{\partial y_j}{\partial w_{kj}} = \delta^{(4)}_j z^{(3)}_k = -(d_j - y_j)z_k
\]  
(8)

Where, \(z^{(3)}_k\) is the output of the third layer.

The calculation for gradient adjustment in the third layer is as follows Eq (9) and (10).

\[
\delta^{(3)}_j = -\frac{\partial E}{\partial x_k} = -\sum_{j=1}^{3} \frac{\partial E}{\partial y_j} \cdot \frac{\partial y_j}{\partial x_k} = - \sum_{j=1}^{3} \delta^{(4)}_j W_{ki}
\]  
(9)

\[
\frac{\partial E}{\partial W_{ik}} = \frac{\partial E}{\partial z_k} \cdot \frac{\partial z_k}{\partial W_{ik}} = -\delta^{(3)}_k x^{(2)}_i = - \left( \sum_{j=1}^{3} \delta^{(4)}_j \partial W_{ik} \right) x_i
\]  
(10)

Where \(x^{(2)}_i\) is the output of the second layer.

\(W_{ik}(0)\) and \(W_{ki}(0)\) are random values in backpropagation.

### 4. Simulation

#### 4.1. Basic Data

This paper proposes a scale of the evaluation for a single risk. As shown in Table 1, the risk scale is divided into five grades: very low risk (VL), low risk (L), middle risk (M), high risk (H) and very high risk (VH). The average value of the evaluation summary given by the expert group is the risk score value of the risk indicator.

#### 4.2. Index Selection

This paper calculates the MIC between each index and the final risk, and the results are shown in Fig 3. It can be seen that the MIC of equipment operation risk, management risk and final risk are the largest, both above 0.9. Therefore, we mainly take these indexes as the input of the evaluation.

#### 4.3. Analysis

This paper investigated the power system in a district of Tianjin and analyzed its risk level based on the science and technology project of State Grid Corporation. The sample risk data collected are shown in Table 2.

In this paper, the first 120 groups of data in the table are selected as network data and input training is carried out through the constructed model. The learning rate and error target were 0.001, and the last 30 groups were test samples. The training sample was adjusted for 2000 times to reach the termination condition. The discriminant results of the last 30 samples were obtained by using the trained network, as shown in Fig 4. The accuracy of MIC-FNN was 96.67%, higher than 80% of BP and 93% of FNN. It indicates that the MIC-FNN of training can be used for risk assessment.

![Fig 3 MIC value of each index and risk level](image)

![Fig 4 Result of FNN](image)
| No. | Operation | Management | Risk Level | MIC-FNN Membership |
|-----|-----------|------------|------------|--------------------|
| 1   | 0.40      | 0.37       | M          | (0.01,0.01,0.97,0.01) |
| 2   | 0.19      | 0.24       | VH         | (0.01,0.02,0.93,0.04) |
| 3   | 0.29      | 0.48       | M          | (0.01,0.01,0.97,0.01) |
| 4   | 0.09      | 0.11       | VH         | (0.03,0.03,0.91,0.03) |
| 5   | 0.77      | 0.85       | VL         | (0.58,0.38,0.03,0.01) |
| 6   | 0.75      | 0.83       | VL         | (0.58,0.38,0.03,0.01) |
| 7   | 0.19      | 0.24       | VH         | (0.01,0.02,0.95,0.02) |
| 8   | 0.09      | 0.12       | VH         | (0.02,0.02,0.93,0.02) |
| 9   | 0.75      | 0.83       | VL         | (0.58,0.38,0.03,0.01) |
| 10  | 0.13      | 0.17       | VH         | (0.02,0.03,0.92,0.03) |
| 11  | 0.70      | 0.77       | L          | (0.6,0.38,0.0,0.01) |
| 12  | 0.09      | 0.11       | VH         | (0.03,0.03,0.91,0.03) |
| 13  | 0.33      | 0.40       | M          | (0.01,0.01,0.98,0.01) |
| 14  | 0.72      | 0.80       | L          | (0.59,0.38,0.02,0.01) |
| 15  | 0.81      | 0.90       | L          | (0.58,0.38,0.04,0.01) |
| 16  | 0.77      | 0.85       | VL         | (0.58,0.38,0.04,0.01) |
| 17  | 0.72      | 0.79       | L          | (0.59,0.38,0.02,0.01) |
| 18  | 0.72      | 0.80       | L          | (0.59,0.38,0.02,0.01) |
| 19  | 0.18      | 0.23       | VH         | (0.02,0.03,0.93,0.03) |
| 20  | 0.09      | 0.11       | VH         | (0.03,0.03,0.91,0.03) |
| 21  | 0.75      | 0.83       | VL         | (0.58,0.38,0.03,0.01) |
| 22  | 0.70      | 0.77       | L          | (0.6,0.38,0.0,0.01) |
| 23  | 0.72      | 0.80       | L          | (0.59,0.38,0.02,0.01) |
| 24  | 0.72      | 0.80       | L          | (0.59,0.38,0.02,0.01) |
| 25  | 0.09      | 0.20       | H          | (0.02,0.01,0.01,0.96) |
| 26  | 0.07      | 0.09       | VH         | (0.02,0.02,0.94,0.02) |
| 27  | 0.73      | 0.81       | L          | (0.59,0.38,0.02,0.01) |
| 28  | 0.74      | 0.82       | VL         | (0.58,0.38,0.03,0.01) |
| 29  | 0.08      | 0.10       | VH         | (0.02,0.02,0.94,0.02) |
| 30  | 0.45      | 0.50       | M          | (0.01,0.98,0.0) |

5. Conclusion
This paper studies the risk of SGCS. Firstly, the index system of operational risk assessment is proposed. Based on the study of risk in historical literature, this paper combines FNN with MIC and proposes a new risk assessment model that considers subjective and objective factors. This model overcomes the shortage of traditional risk assessment methods that rely too much on expert experience. At the same time, proposed methods make prior knowledge well embedded in the network. Besides, through the collection and update of sample data, the model can quickly establish and adjust practical problems, which improves the adaptability of the evaluation model. The conclusion of the case analysis shows that this method can accurately evaluate the risk level of SGCS and can be used in practical application.

Acknowledgments
This study is supported by the science and technology project of State Grid Corporation of China “Research on key technologies of system construction and planning based on the ‘two-network fusion’ of Ubiquitous Power Internet of Things and Strong Smart Distribution Grid”. The authors are grateful to the participators who help to improve the paper by many pertinent comments and suggestions.
References

[1] Hui-sheng Gao, Jing-xue Ran, Hai-yang Xie, and C. Li, ‘Risk evaluation of communication network of electric power using logarithm least squares and restriction coefficient’, in 2008 7th World Congress on Intelligent Control and Automation, Chongqing, China, 2008, pp. 5801–5806, doi: 10.1109/WCICA.2008.4592815.

[2] H.-S. Gao, J.-X. Ran, and Y.-Q. Sun, ‘Risk Evaluation of Communication Network of Electric Power Based on Restriction Coefficient’, in 2007 International Conference on Wireless Communications, Networking and Mobile Computing, Shanghai, China, 2007, pp. 3127–3129, doi: 10.1109/WICOM.2007.776.

[3] M. H. Ansari, V. T. Vakili, B. Baharak, and P. Tavassoli, ‘Graph theoretical defense mechanisms against false data injection attacks in smart grids’, Journal of Modern Power Systems and Clean Energy, vol. 6, no. 5, pp. 860–871, Sep. 2018, doi: 10.1007/s40565-018-0432-2.

[4] G. Chaojun, P. Jirutitijaroen, and M. Motani, ‘Detecting False Data Injection Attacks in AC State Estimation’, IEEE Transactions on Smart Grid, vol. 6, no. 5, pp. 2476–2483, Sep. 2015, doi: 10.1109/TSG.2015.2388545.

[5] D. B. Rawat and C. Bajracharya, ‘Detection of False Data Injection Attacks in Smart Grid Communication Systems’, IEEE Signal Processing Letters, vol. 22, no. 10, pp. 1652–1656, Oct. 2015, doi: 10.1109/LSP.2015.2421935.

[6] W.-L. Chin, C.-H. Lee, and T. Jiang, ‘Blind False Data Attacks Against AC State Estimation Based on Geometric Approach in Smart Grid Communications’, IEEE Transactions on Smart Grid, vol. 9, no. 6, pp. 6298–6306, Nov. 2018, doi: 10.1109/TSG.2017.2708114.

[7] W.-H. Chen, Q.-Y. Jiang, Z.-Y. Wang, and Y.-J. Cao, ‘Risk assessment of cascading outages in power systems using fuzzy neural network’, in Advances in Neural Networks - Isnn 2006, Pt 2, Proceedings, vol. 3972, J. Wang, Z. Yi, J. M. Zurada, B. L. Lu, and H. Yin, Eds. Berlin: Springer-Verlag Berlin, 2006, pp. 1422–1427.

[8] Z. Chen, X. Y. Xiao, and C. S. Li, Risk Assessment of Power Systems Integrated with Wind Farm and SMES. New York: Ieee, 2015.

[9] X. Ji, H. Ji, Q. Yao, and M. Wang, ‘Risk Index Early-warning of Smart Grid Based on Neural Network’, in 2017 4th International Conference on Systems and Informatics (icsai), New York: Ieee, 2017, pp. 1723–1727.

[10] C. Liu, P. Fan, H. Wang, J. Guo, and R. Ke, ‘Modeling Forest Fire Risk Assessment Based on BP Neural Network of Transmission Line’, Power System Protection and Control, vol. 45, no. 17, pp. 100–105, 2017.

[11] X. He, W. Wang, X. Liu, and Y. Ji, ‘Risk Assessment of Communication Network of Power Company Based on Rough Set Theory and Multiclass SVM’, Physics Procedia, vol. 24, pp. 1226–1231, 2012, doi: 10.1016/j.phpro.2012.02.183.

[12] L. Cui, G. Li, T. Wang, J. Shi, and Y. Song, ‘Fuzzy Measurement of Information Communication System Risk Estimation in Smart Grid’, Electric Power Construction, vol. 38, no. 05, pp. 52–61, 2017.

[13] J. Liu, X. Wang, H. Sun, S. Niu, and Z. Zhang, ‘Feature Selection Analysis of Power System Transient Stability Prediction Using FJPD Mutual Information’, Smart Power, vol. 46, no. 10, pp. 78-85+95, 2018.

[14] A. R. Iyer, D. K. Prasad, and C. H. Quek, ‘PIE-RSPOP: A brain-inspired pseudo-incremental ensemble rough set pseudo-outer product fuzzy neural network’, Expert Systems with Applications, vol. 95, pp. 172–189, Apr. 2018, doi: 10.1016/j.eswa.2017.11.027.