Risk Element Identification of Grid Communication System Based on Improved Bayesian Network under SG and UPIOT

Yunli Yue1*, Xuefeng Jia2 and Cunbin Li2

1Economic and technological research institute, State Grid Jibei Electric Power Corporation, 21 East Yangfangdian Rd., Haidian District, Beijing, 100038, China
2School of Economic and Management, North China Electric Power University, Beijing, 102206, China
*Yunli Yue’s e-mail: 39213654@qq.com

Abstract. Now, Smart Grid and Ubiquitous Power Internet of Things are booming in China. The degree of coupling between the grid and the power communication system determines the degree of development of Smart Grid and Ubiquitous Electric Internet of Things. The rapid growth of Smart Grid has resulted in the risk transfer of Smart Grid Communication Systems with many levels and a strong correlation of indicators. This paper sorts out the risk indicators and provides a three-layer risk transfer network considering equipment, environment, business and operational dimension. Moreover, the triangular fuzzy number and DS evidence theory are used to assign the Bayesian network root node probability. At last, this paper identifies the risk element of SGCS and proposes preventive measures.

1. Introduction
With the development of smart grid (SG) and Internet of things (UPIOT), power grid has realized informatization, networking and optimization, ensuring its safety, reliability, flexibility, efficiency and economy.[1]. The power communication network can process real-time data in the smart grid, which is a combination of the power network and the information network. The Smart Grid Communication System (SGCS) should meet the needs of power generation, transmission, substation, distribution, sales and dispatching. It should also include a centralized analysis of data in each link to provide data support for grid management.

Many scholars have researched risk assessment, and there are also many risk assessments for smart grids. Yan Xingli et al. used a fuzzy analytic hierarchy process to assess the quality of power quality [2]. Li Peng and other researches on SG were dispatching with information risk element uncertainty, using TFN-AHP model to analyze the risk transfer mechanism of the power grid [3]. Z Chen proposed a new wind power assessment method, which combines wind farm and superconducting magnetic energy storage (SMES) equipment by sequential Monte Carlo simulation to investigate the impact of SMES capacity and maximum charging/discharging power on power system risk [4]. Gatzert Nadine provided a comprehensive overview of current risk and risk management solutions for renewable energy projects and identifies vital gaps [5]. He Yongxiu analyzed risk identification and analyzed the planning process of renewable energy generation [6]. Jiang Xinzhen started from the TOPSIS intuitionistic fuzzy multi-attribute decision-making method and studies the problem based on the intuitionistic fuzzy interval multi-attribute decision-making model [7]. Besides, increasingly risk transfer studies adopt fuzzy sets such as interval fuzzy numbers, fuzzy triangular numbers, and random fuzzy numbers [8]–[11].
However, at present, with the increasing degree of physical information fusion of power grids, the traditional research methods of communication network risk transfer have not been applied to SGCS research. This paper will adopt the Bayesian network model to conduct risk transfer research on SGCS. Relevant investigations include Li Cunbin and Huang Min studied the power supply risk with learning Bayesian [12]. Liao Yuanxi used Bayesian network theory to construct real-time network topology and calculate fault probability [13]. Nie Ruihua used Bayesian network method and support vector machine to establish a risk early warning model [14]. Islam proposed a new fuzzy Bayesian network model to study the risks of power plant projects [15]. Boudali delved into the Bayesian network of separated time series and found a suitable reliability framework for dynamic systems [16].

In this paper, using the theoretical method proposed in the reference[11], the risk indicators are sorted out, and a three-layer risk transfer network considering communication equipment, external environment, business, and maintenance management is proposed, and the triangular fuzzy number and DS evidence theory are used to determine the Bayesian network root node probability. Finally, this paper identifies the critical risk nodes in the communication network.

2. Based on DS theory and Bayesian network SGCS risk transfer model

2.1. Bayesian Network
Suppose $\mathcal{B} = (\mathcal{G}, \Phi)$ is a Bayesian Network. And $\mathcal{G}$ represents a directed acyclic graph (DAG). In this graph, each node represents an attribute variable. If there is a dependency between variables, a directed line segment is used to point to the dependent variable from the dependent variable. $\Phi$ quantitatively depicts all the dependencies in the graph, including the conditional probability table for each variable.

$(x_1, x_2, ..., x_n)$ is the value of all variables in the DAG, and $\text{par}(x_i)$ is the collection representing the parent node of variable $x_i$. Its joint probability can be expressed as $p(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} p(x_i | \text{par}(x_i))$.

2.2. Root node probability determination based on the triangular fuzzy number
The intervals number can eliminate the fuzziness of the value, representing a number with the closed interval $[x^-, x^+]$. $x^-$ and $x^+$ represents the upper and lower limits of the number. The triangular fuzzy number $[x^-, x^*, x^+]$ is an extension of the general interval number and the maximum possible value $x^*$ of the number is included in the interval in the way of the interval center of gravity, satisfying $x^- \leq x^* \leq x^+$. If $f(x)$ is used to indicate the probability of the number in the interval, that is, its membership function is

$$
\mu(x) = \begin{cases}
0, x & x < x^- \\
\frac{x-x^-}{x^*-x^-}, x^- \leq x \leq x^* \\
\frac{x^*-x}{x^*-x^*}, x^* \leq x \leq x^+ \\
1, x^+ < x
\end{cases}
$$

For any two triangular fuzzy numbers $X = [x^-, x^*, x^+]$ and $Y = [y^-, y^*, y^+]$, there are the following operation rules:

$$
X \Theta Y = [x^- + y^-, x^* + y^*, x^+ + y^+]
$$

To improve the accuracy of expert evaluation, five scales are used to judge, as shown in Table 1, the expert opinions are converted into five fuzzy probability values by five semantic values. Suppose there are $l$ experts who assign the linguistic variable to the probability value of multi-state root node $X_i$ in state $j$, and the triangular fuzzy number of the $k^{th}$ expert assignment is expressed as

$$
p_k^i = [(x^-)^{k}_{i,j}, (x^*)^{k}_{i,j}, (x^+)^{k}_{i,j}]
$$

Using the arithmetic average method to synthesize expert opinions, the fuzzy probability value of the comprehensive evaluation of $X_i$ in state $j$ by $l$ expert nodes is

$$
P_{i,j} = \frac{p_1^i \oplus p_2^i \oplus ... \oplus p_l^i}{l} = [(x^-)^{l}_{i,j}, (x^*)^{l}_{i,j}, (x^+)^{l}_{i,j}]
$$
Use the “Average Area Method” to convert the fuzzy probability values into exact values, as (5):

\[ p_{i-j}'' = \frac{(x^-)_{i-j} + 2(x^=)_{i-j} + (x^+)_{i-j}}{4} \]  

To ensure that the sum of all probability values is 1, the probability value of the root node should be treated as “unitized”, as (6):

\[ p_{i-j} = \frac{p_{i-j}'}{\sum p_{i-j}'} \]  

| NO. | Semantic value | Triangular fuzzy number |
|-----|---------------|------------------------|
| 1   | Highest       | (0.7, 1.0, 1.0)        |
| 2   | Higher        | (0.5, 0.7, 0.9)        |
| 3   | Medium        | (0.3, 0.5, 0.7)        |
| 4   | Lower         | (0.1, 0.3, 0.5)        |
| 5   | Lowest        | (0, 0, 0.3)            |

2.3. Determination of conditional probability based on DS credit theory

Let \( \Theta \) be the identification frame composed of the mutually exclusive values of the variable \( X \), and the power set constitutes the set \( 2^\Theta \). Suppose \( \forall A \subset \Theta \), for the set function \( m \) to satisfy \( 2^\Theta \rightarrow [0,1] \), which is

\[ \sum_{A \in \Theta} m(A) = 1, m(\emptyset) = 0 \]  

The \( m \) is called the Basic Reliability Allocation (BPA) on the recognition framework \( \Theta \), which characterizes the extent to which evidence supports the likelihood of an event. The time trust function and the likelihood function are determined according to the basic reliability allocation. In the same way, this paper constructs the \([ Bel(A), P(A) ]\) confidence interval. According to the Dempster synthesis rule, two mass functions \( m_1, m_2 \) are synthesized:

\[ m_1 \oplus m_2 (A) = \frac{1}{1-K} \sum_{B \cap C} m_1(B) \cdot m_2(C) \]  

In the (8), \( K = \sum_{B \cap C} m_1(B) \cdot m_2(C) \) is the inter-evidence conflict coefficient.

| Suitability value          | value |
|----------------------------|-------|
| Extremely suitable         | 6     |
| Strongly to extremely      | 5     |
| Strongly suitable          | 4     |
| Generally to strongly      | 3     |
| Generally suitable         | 2     |

Considering the impact of the scoring experts' own research fields, business level, and academic titles, the decision-making criteria of Table 2 are constructed, indicating that the experts judge the conditional probability in the Bayesian network model. Suppose \( t \) experts \( (e_1, e_2, ..., e_t) \) make an important judgment based on the recognition frame \( \Theta \) from the \( n \) dimensions \( (c_1, c_2, ..., c_n) \) to the combined object \( x_1, x_2, ..., x_n \).

In Table 3, 1 is the comparison with the elements' own, and 0 is the comparison without none. \( s_k \) is the \( k^{th} \) element under the attribute \( c_j \); \( a_k \) is the comparison coefficient between \( s_k \) and recognition frame \( \Theta \). \( p_{ij} \) is the weight of the expert in \( e_i \) under the attribute \( c_j \).

Solve the knowledge matrix based on expert information and determine the BPA of each focal element. Its maximum eigenvalue is \( \lambda_{d+1} = 1 + \sqrt{d} \), and eigenvector is \((x_1, x_2, ..., x_d, x_{d+1})\). They satisfy (9) and (10).

\[ x_j = \frac{a_{j}p_{ij}}{\sum_{i=1}^{d} a_{i}p_{ij} + \sqrt{d}}, j = 1,2, ..., d \]  

\[ x_{d+1} = \frac{\sqrt{d}}{\sum_{i=1}^{d} a_{i}p_{ij} + \sqrt{d}} \]
Table 3 Expert $e_j$ knowledge matrix at $c_j$ latitude

| $c_j$ | $s_1$ | $s_2$ | …… | $s_r$ | $\Theta$ |
|-------|-------|-------|-----|-------|---------|
| $s_1$ | 1     | 0     | …… | 0     | $(a_1 p_{ij})$ |
| $s_2$ | 0     | 1     | …… | 0     | $(a_2 p_{ij})$ |
| ……   | ……   | ……   | …… | ……   | ……      |
| $s_r$ | 0     | 0     | …… | 1     | $(a_3 p_{ij})$ |
| $\Theta$ | 1/$(a_1 p_{ij})$ | 1/$(a_2 p_{ij})$ | …… | 1/$(a_3 p_{ij})$ | 1 |

3. SGCS system risk assessment analysis

The risk transfer research process of SGCS is shown in Fig 1.

3.1. Identify the risk factors of SGCS

This paper divides the SGCS risk into three layers. The first layer is $R_0$ SGCS failure risk, which indicates the overall goal of risk transfer. Because of the emphasis, it is not appropriate to divide too many nodes in the second layer. It is divided into four nodes: $R_1$ Equipment, $R_2$ Environment, $R_3$ Business and $R_4$ Operational; the third layer is the specific factor layer, which details the incentives for triggering the node risk in the second layer, which consists of 15 nodes in total. The particular risk element transfer indicator system is shown in Fig 2.

3.2. Determine the Bayesian network root node probability value

Table 4 Fuzzy probability statistics table of root node $R_{31}$ wind level

| Expert serial number | High risk  | Status    |
|----------------------|------------|-----------|
| 1                    | (0.1,0.3,0.5) Lower | (0.5,0.7,0.9) Higher |
| 2                    | (0,0,0.3) Lowest   | (0.7,1,0,1) Highest |
| 3                    | (0,0,0.3) Lowest   | (0.7,1,0,1) Highest |
| 4                    | (0.5,0.7,0.9) Higher| (0.1,0.3,0.5) Lower |

This paper selects the most empirical research object of this method in a certain SGCS in Tianjin. Due to a large number of nodes, the root node probability is determined only by taking $R_1$ and $R_32$ as examples. $R_{31}$ is set as an example to control the possibility of service interruption and $R_{32}$ is to manage the service terminal possibility as a 2-state node (high-intensity, low-intensity), and $R_3$ is a 3-state node (strong risk, general risk, and weak risk). Consult the four authoritative experts in the relevant fields to
judge the possible occurrence value of the risk in the SGCS. Refer to Table 2 to establish the fuzzy probability statistics table of the R_{31} wind level, as shown in Table 4.

According to (4), the fuzzy average of the R_{31} wind level state can be obtained:

$$p'_{31-1} = (0.6, 1.0, 2.0), \ p'_{41-2} = (2.0, 3.0, 3.4)$$

According to (5), the R_{31} wind level state defuzzification can be obtained, and the unified (6) can be used to obtain the R_{31} wind level probability value:

$$p_{31-1}'' = 0.29, \ p_{41-2}'' = 0.71$$

The probability value of other root nodes can be solved by the method of solving the probability value of the R_{31} wind level.

Determine the conditional probability value of the intermediate node. Taking R_3 as an example, the risk level of the R_3 Business factor is directly determined by R_{31} for controlling business interruption possibility and R_{32} for managing business interruption possibility. In the case of equal treatment of the root nodes R_{31} and R_{32}, the conditional probability values are determined by four combinations: \{high-high\}, \{high-low\}, \{low-high\}, and \{low-low\}. Now take the case of \{high-high\} as an example, consult three experts in related fields, assuming that the weights of the three experts are 0.2, 0.7 and 0.1 and establish the knowledge matrix of each expert, as shown in Table 5, Table 6 and Table 7.

Solve the maximum eigenvalues of \(\det \left( C_i - \lambda I \right) (i = 1, 2, 3)\) and the corresponding eigenvectors, synthesize the reliability function according to the principle of evidence theory synthesis in (8), and obtain the conditional probability BPA value, such as TABLE 8.

### Table 5 Establish knowledge matrix A_1 with expert 1 preference

| \{High-High\} | Strong | General | Weak | \(\Theta\) |
|----------------|--------|---------|------|-----------|
| Strong         | 1      |         |      | 0.8       |
| General        |        | 1       |      | 0.6       |
| Weak           |        |         | 1    | 0.4       |
| \(\Theta\)     | 2.5    | 1.7     | 2.5  | 1         |

### Table 6 Establish knowledge matrix A_2 with expert 2 preference

| \{High-High\} | \{Strong, General\} | Weak | \(\Theta\) |
|----------------|---------------------|------|-----------|
| \{Strong, General\}| 1                   |      | 3.5       |
| Weak           |                     | 1    | 1.4       |
| \(\Theta\)     | 0.3                 | 0.7  | 1         |

### Table 7 Establish knowledge matrix A_3 with expert 3 preference

| \{High-High\} | Strong | \{General, Weak\} | \(\Theta\) |
|----------------|--------|-------------------|-----------|
| Strong         | 1      |                   | 0.4       |
| \{General, Weak\}| 1      |                   | 0.2       |
| \(\Theta\)     | 2.5    | 5                 | 1         |

### Table 8 Conditional probability BPA distribution table

| Events          | \{Strong risk\} | \{General risk\} | \{Low risk\} |
|-----------------|-----------------|-----------------|-------------|
| BPA             | 0.678           | 0.254           | 0.068       |

### 3.3. SGCS risk transfer structure

The triangular fuzzy number theory is used to synthesize expert experience and historical data to determine the probability of each root node. For the 3-state intermediate node, the conditional probability distribution can be determined according to the evidence relationship between it and the root node.

The risk transmission network of the SGCS is shown in Fig 3. The indicators such as R_2 environmental factors and R_4 operation and maintenance factors can be determined based on historical...
data, and the risk indicators can be determined by changing the corresponding inputs when performing a risk assessment. After inputting the similar indicator values, the Bayesian network will make a reasoning diagnosis based on the conditional relationship between the nodes to determine the vital nodes where the risk situation may occur. The critical nodes that will trigger the SGCS failure risk are the R_{11} cable loss level indicator in the R_{1} device factor and the R_{41} device maintenance time in the R_{4} operation and maintenance indicator. Therefore, in the SGCS, the reliability of the communication equipment must be ensured. If equipment failure occurs, advanced fault location technology should be used to locate the fault and repair the defect in time. Pay attention to the operation and maintenance of the equipment, and regularly check the equipment to reduce loss.

Fig 3 Bayesian network structure diagram

4. Conclusion
SGCS involves multiple links, and there are many potential risks and the correlation is extremely high. In this paper, the risk transmission of the smart grid communication network is combed into a three-tier structure, which is divided into four categories: equipment, environment, service and operation and maintenance, and the conditional probability of each node is determined. The conclusions are as follows:

(1) This paper combines the risk index of SGCS integrating equipment factors, environmental factors, business factors and operation and maintenance factors to determine multiple sub-risk elements. And this paper constructs a risk assessment model of the Bayesian network.

(2) This paper applies the triangular fuzzy number to the 2-state root node sub-risk element lacking valid historical data relying on the expert experience information, and solves the fuzzy probability value.

(3) For the determination of the probability value of the risk element condition of the intermediate state of the 3-state, the DS evidence theory is adopted to construct the knowledge matrix, and the conditional probability distribution of the 3-state corresponding to the root node is determined. This method eliminates the comparison between the analytic hierarchy process and the two. The tedious process of consistency checking can effectively reduce the uncertainty.

(4) Relying on the Bayesian network's diagnostic reasoning mechanism to determine the risk probability of SGCS, and to identify the key nodes in the SGCS, the model can be extended to the SGCS risk transfer research.

Acknowledgments
This study was 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 were grateful to the participants who help to improve the paper by many pertinent comments and suggestions.
References

[1] X. Li, W. Wei, Y. Wang, Z. Mu, and W. Gu, ‘Study on the development and technology of strong smart grid’, Power System Protection and Control, vol. 37, no. 17, pp. 1–7, 2009.

[2] X. Zhai, F. Wen, Z. Lin, and Y. Li, ‘Power quality comprehensive evaluation and sensitivity analysis based on fuzzy analytical hierarchy process’, Journal of North China Electric Power University (Natural Science Edition), vol. 40, no. 05, pp. 48–53, 2013.

[3] P. Li, C. Li, and L. Niu, ‘Smart Grid Dispatching Decision-making Risk Transmission Model Considering Information Risk Element Uncertainty’, Journal of North China Electric Power University (Natural Science Edition), vol. 42, no. 02, pp. 43–49, 2015.

[4] 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.

[5] N. Gatzer and T. Kosub, ‘Risks and risk management of renewable energy projects: The case of onshore and offshore wind parks’, Renew. Sust. Energ. Rev., vol. 60, pp. 982–998, Jul. 2016, doi: 10.1016/j.rser.2016.01.103.

[6] Y. He, Y. Zhou, B. Wang, T. Zhang, and W. Xiong, ‘Risk analysis and aversion of renewable energy supplies in China based on the Monte Carlo stochastic simulation method’, Appl. Math. Inf. Sci., vol. 6, no. 3, pp. 975–981, Nov. 2012.

[7] X. Jiang, Y. Hong, and Y. Zhang, New Energy Marketing Risk’s Assessment and Sorting in Power Supply Enterprise Based on Interval-valued Intuitionistic Fuzzy Multi-attribute Decision Making Model. New York: Ieee, 2012.

[8] J. Yuan and C. Li, ‘A New Method for Multi-attribute Decision Making with Intuitionistic Trapezoidal Fuzzy Random Variable’, Int. J. Fuzzy Syst., vol. 19, no. 1, pp. 15–26, Feb. 2017, doi: 10.1007/s40815-016-0184-y.

[9] C. Li, J. Yuan, D. Liu, W. Li, and S. Lin, ‘Risk group assessment based on the demands of important consumers with standby electric source’, Journal of Intelligent & Fuzzy Systems, vol. 33, no. 4, pp. 2177–2188, Sep. 2017, doi: 10.3233/JIFS-162420.

[10] J. Yuan and C. Li, ‘Intuitionistic Trapezoidal Fuzzy Group Decision-Making Based on Prospect Choquet Integral Operator and Grey Projection Pursuit Dynamic Cluster’, Mathematical Problems in Engineering, vol. 2017, pp. 1–13, 2017, doi: 10.1155/2017/2902506.

[11] C. Li and J. Yuan, ‘A New Multi-attribute Decision-Making Method with Three-Parameter Interval Grey Linguistic Variable’, Int. J. Fuzzy Syst., vol. 19, no. 2, pp. 292–300, Apr. 2017, doi: 10.1007/s40815-016-0241-6.

[12] C. Li and M. Huang, ‘Power supply risk transmission analysis based on learning Bayesian networks’, Journal of Central South University (Natural Science Edition), vol. 42, no. 08, pp. 2338–2344, 2011.

[13] Z. Li, T. Xu, J. Gu, and R. Wang, ‘Bayesian Network Conditional Probability Value Determination Method Based on DS/AHP’, Modern Defense Technology, vol. 45, no. 06, pp. 178–184, 2017.

[14] R. Nie and H. Shi, ‘Research on Risk Early Warning Model of Listed Securities Companies Based on Bayesian Network’, THE THEORY AND PRACTICE OF FINANCE AND ECONOMICS, vol. 39, no. 06, pp. 51–57, 2018.

[15] M. S. Islam and M. Nepal, ‘A Fuzzy-bayesian Model for Risk Assessment in Power Plant Projects’, Procedia Computer Science, vol. 100, pp. 963–970, 2016, doi: 10.1016/j.procs.2016.09.259.

[16] H. Boudali and J. B. Dugan, ‘A discrete-time Bayesian network reliability modeling and analysis framework’, Reliability Engineering & System Safety, vol. 87, no. 3, pp. 337–349, Mar. 2005, doi: 10.1016/j.ress.2004.06.004.