Fault diagnosis of power grids based on grey relational analysis

Huayong Lu¹,², Hongzuo Guo³, Zhe Liu²,³, Xiao Yang² and Bing Leng²

¹College of Energy and Electrical Engineering, Hohai University, Nanjing, China; ²Electric Power Dispatch Control Center, State Grid Henan Electric Power Company, Zhengzhou, China; ³E-mail: lhy01@163.com

Abstract. A fault diagnosis method of power grids was presented based on grey relational analysis in this paper. For each suspected fault line in the power cut area, the grey relational of original characteristics sequence, amplitude and energy are calculated by using the current sampling value before and after the fault time. Then an active reference line was selected, followed by the calculation of the horizontal grey correlation between reference line and suspected fault line according to current sampling after the fault. Finally, the longitudinal grey relational degree and the horizontal grey relational degree of each line are weighted and fused to obtain their comprehensive grey relational degree, and the fault criterion is given to identify the faulty component accurately. The simulation results of the IEEE 14-node system show that the fault elements can be identified under different fault locations, various fault types and different transition resistances accurately.

1. Introduction

Although fault diagnosis method in power systems are widely used based on switching quantity diagnostic, such as expert systems [1], optimization techniques [2], etc. However, it is easy to produce erroneous diagnosis results if the multi-position protection and the circuit breaker action information are incorrect or incomplete, and the anti-interference ability of switch information is poor. Considering that the information of electrical quantities has a great amount of information, the problem of single information resource and less information quantity can be solved when the information of electrical quantities are applied to the power grid fault diagnosis. Meanwhile, the electrical quantity information has the advantage that the switch quantity cannot be compared in accuracy and anti-interference. When the power grid fails, the relevant protection and circuit breaker action, although the protection and circuit breaker mis-operation and refusal, but the fault components must be in the blackout area. Fault diagnosis of components in blackout area has been studied based on electrical quantity information in relevant literatures [3-5].

The grey relational analysis [6] has been applied to transformer fault identification, transmission line fault phase judgment and power system station domain protection [7-11]. When the grey relational analysis is used in transformer fault identification and transmission line fault phase identification, it is necessary to establish a standard fault pattern sequence. By calculating the grey relational degree between the actual sequence and the standard fault pattern sequence, the fault types of transformer and transmission line are analyzed. Using the fault current transient component, a station protection algorithm is proposed based on gray correlation calculation, which can quickly
identify the fault by using the fault transient current within 10ms after the fault. However, these existing algorithms take the geometric similarity of time series of electrical quantities to judge the degree of correlation, and need to establish a standard fault pattern sequence library.

The grey relational degree analysis can not only describe the similarity of geometric shapes of time series, but also describe the numerical differences between the two sequences. When the power grid fails, the current flowing through the fault line changes greatly before and after the fault time, and the current flowing through the fault line and the current flowing through the normal line after the fault is also quite different. In this paper, a new fault diagnosis method based on grey relational analysis is proposed, which uses electrical quantity information. Considering that when multiple faults occur in the power grid, the fixed resolution coefficient may not be able to correctly identify all fault elements. Therefore, this paper does not use fixed resolution coefficient, but uses dynamic resolution coefficient to calculate the grey relational degree.

2. Construct the grey relational degree for power grid fault diagnosis

2.1. The basic principle of grey relational analysis

Let $X_i$ be a time series, $x_i(k) (k = 1, 2, L, n)$ are the data at time $k$, where $n$ is the maximum sampling point. Then the characteristic behavior sequence of agent $i$ is

$$X_i = (x_i(1), x_i(2), L, x_i(n))$$

Suppose that there are $m$ agents in the system, the characteristic behavior sequences of each agent are $X_0, X_1, L, X_m$. Take characteristic behavior sequences $X_0$ as reference sequence, $X_1, X_2, L, X_m$ as compare sequences.

The grey relational of reference sequence and compare sequences can be written as

$$r(X_0, X_i) = \frac{1}{n} \sum_{k=1}^{n} r(x_0(k), x_i(k))$$

where

$$r(x_0(k), x_i(k)) = \frac{\min_k \min_i |x_0(k) - x_i(k)| + \xi \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \xi \max_i \max_k |x_0(k) - x_i(k)|}$$

In (3), $\xi \in (0, 1)$ is the resolution coefficient of grey relation, which is used to improve the difference between gray correlations. $|x_0(k) - x_i(k)|$ is the absolute difference of $X_0$ and $X_i$ at time $k$. $\min_k \min_i |x_0(k) - x_i(k)|$ is the two-level minimum difference of $X_0$ and $X_i$ at time $k$. $\max_i \max_k |x_0(k) - x_i(k)|$ is the two-level maximum difference of $X_0$ and $X_i$ at time $k$.

For simplification, denote

$$\Delta_i(k) = |x_0(k) - x_i(k)|, i = 1, 2, L, m$$
$$\Delta = (\Delta_1), (\Delta_2), L, (\Delta_m), M = \max_i \Delta_i(k), N = \min_i \Delta_i(k)$$

Then, (3) can be rewritten as

$$r_{0i}(k) = \frac{N + \xi M}{\Delta(k) + \xi M}$$

The grey relational of reference sequence and compare sequences (2) can be rewritten as

$$r_{0i} = \frac{1}{n} \sum_{k=1}^{n} r_{0i}(k)$$
2.2. Extraction of current sequence
As shown in Figure 1, let \( I_a, I_b, I_c \) denote the instantaneous current sampling value of \( A, B, C \) phase of the line side respectively. Let \( S \) denote the transmission line transmitter side, and \( I_s \) denote the current of transmitter side. Let \( R \) denote the transmission line receiving terminal, and \( I_R \) denote the current of receiving terminal.

Suppose \( I_s = I_a + 2I_b - 3I_c \), then \( I_s \) always contains fault current no matter what type of fault occurs, and \( I_s \) is small when the line is normal. Similarly, let \( I_R = I_a + 2I_b - 3I_c \).

When a fault occurs in the \( k_1 \) zone, the integrated current \( I_{z_1} = I_s + I_R > I_{k_1} \), where, \( I_{k_1} \) is short-circuit current. When an out-of-zone failure occurs, the integrated current \( I_{z_2} = I_s + I_R \approx 0 \).

![Figure 1. Transmission line model.](image)

For the faulty line, the integrated current of the line before fault is smaller, and the integrated current of the line after fault is larger. For normal lines, the integrated current is always small. Therefore, the integrated current of faulty lines before and after faults varies greatly. Then, the fault element can be judged vertically in time sequence by comparing the gray correlation of the integrated current before and after the fault of the same line.

Similarly, when the power grid fails, the integrated current of the fault line is larger, the integrated current of the normal line is smaller, and the integrated current between the fault line and the normal line is quite different. A normal transmission line outside the power outage area is selected as a reference line, and then the gray correlation degree between each suspected fault line and the reference line is calculated separately, then the faulty component can be determined in the lateral direction of the space.

2.3. Extraction of feature sequences
When a fault occurs in the power grid, the network connection analysis is performed to find out the power outage area, and the lines \((l_1, l_2, \ldots, l_n)\) in the power outage area are regarded as suspected fault lines.

Considering that the current at both ends of the normal line will change after the fault occurs in the power grid, but the current at both ends of the normal line is equal in magnitude and opposite in direction, so the integrated current of the normal line is small. For the suspected fault line \( l_i \), the integrated current of the line \( l_i \) with three cycles before and after faults is taken as the characteristic sequence \( X_i \) of the line \( l_i \). Through the network topology analysis, the active line with the smallest electrical distance from all suspected fault lines is selected as the reference line \( l_0 \). For the reference lines \( l_0 \) and the suspected fault line \( l_i \), take the three-cycle line integrated current after the fault as the characteristic sequence \( X_0 \) and \( X_i \) of line \( l_0 \) and \( l_i \) respectively.

3. Fault diagnosis of power network based on grey relational analysis
The principle of applying grey relational analysis model to grid fault diagnosis is as follows. Firstly, the grey correlation degree, amplitude grey correlation degree and energy value grey relational degree of the original characteristic sequence of each suspected fault line are calculated, and then the grey relational degree, amplitude grey relational degree and energy value grey relational degree of the
original characteristic sequence of each suspected fault line and the reference line are calculated. Weighted fusion of the six grey relational degree values, through normalization and difference treatment, get the component fault degree. Finally, the fault components are identified by comparing with threshold values.

3.1. Calculate the grey relational grade
Suppose that a fault takes place in power grid at time $t_l$, take the feature sequence $x_i(k), k = t_l - 3T, L, t_l - 1$ before fault as reference sequence for the $i$th suspected fault lines $l_i$, take the fault sequence $x_i(k), k = t_l, L, t_l + 3T$ after the fault as comparison sequence. The absolute difference of absolute difference and reference sequence is denoted as

$$\Delta_i(k) = |x_i(t_l + k) - x_i(k)|$$

According to (4), the longitudinal gray relational degree of each suspected line feature sequence can be obtained, which is denoted as $r_{tz1}$. After the characteristic sequence $X_i$ of the line $l_i$ is decomposed by the modified EMD [12], the first three IMF components are taken and the Hilbert transform is performed. If the instantaneous amplitude of the $j$th IMF component of line $l_i$ at time $k$ is $A_{ij}(k), j = 1, 2, 3$, then the instantaneous amplitude of line $l_i$ at time $k$ is

$$A_i(k) = \sqrt{A_{i1}^2(k) + A_{i2}^2(k) + A_{i3}^2(k)}$$

(7)

For the $i$th suspected fault lines $l_i$, the instantaneous amplitude sequence before the fault is taken as the reference sequence, and the instantaneous amplitude sequence after the fault is taken as the comparison sequence. The absolute difference of absolute difference and reference sequence is denoted as

$$\Delta_iA(k) = |a_i(t_l + k) - a_i(k)|$$

(8)

According to (4), the longitudinal amplitude gray correlation degree of each suspected line feature sequence can be obtained, which is denoted as $r_{Ad1}$. Next, the grey relational degree of longitudinal energy is calculated as follow. The HHT transform is applied to the characteristic sequence of the suspected fault line. The energy sequence of the line is formed by the marginal energy spectrum value of each integral interval. The energy sequence before the fault is taken as the reference sequence, and the energy sequence after the fault is taken as the comparison sequence. The absolute difference of absolute difference and reference sequence is denoted as

$$\Delta_iE(k) = |e_i(t_l + k) - e_i(k)|$$

(9)

According to (4), the grey relational degree of longitudinal energy for each suspected line feature sequence can be obtained, which is denoted as $r_{Ad2}$. Similarly, the horizontal grey correlation degree $r_{hz2}$, the horizontal amplitude grey correlation degree $r_{Ad2}$ and the horizontal energy value grey correlation degree $r_{Ed2}$ are calculated, respectively.

3.2. Calculation method of the dynamic resolution coefficient $\xi$
Considering that the dynamic resolution coefficient $\xi$ has a great influence on the calculation of grey correlation degree [13], a dynamic resolution coefficient calculation method is adopted in this paper.

Using the amplitude sequence of each line, and let the amplitude distortion of each line be

$$q_i = \sum_{j=1}^{l-1} a_i(j) \sum_{j=1}^{l-1} a_i(j), i = 1, 2, L, m$$

(10)
Select the minimum number greater than 3 in $q_i$, the amplitude distortion corresponds to the $j$th line, which is used as the calculation line to calculate the resolution coefficients of six characteristic sequences, i.e. longitudinal characteristic sequence, longitudinal amplitude, longitudinal energy, horizontal characteristic sequence, horizontal amplitude and horizontal energy, respectively. The formula is as follows:

$$\xi = \frac{1}{2} \sum_{k=1}^{n} \Delta_j(k) \left( \max_{i=1}^{n} \max_{j=1}^{m} \Delta_{ij} \right)$$ (11)

### 3.3. Calculate the fault degree of lines

The comprehensive grey relational degree of the line is defined as follow.

$$r_i = \frac{2(r_{iTZ} + r_{iAD} + r_{iED}) + (r_{iTZ} + r_{iAD} + r_{iED})}{9}$$ (12)

Then, normalization and difference processing of the comprehensive grey correlation grade $r_i$ of each line are processed to obtain the fault degree $m_i$.

$$m_i = 1 - \frac{r_i - 0.5}{0.5}, \quad i = 1, 2, L , m$$ (13)

For each suspected fault line, if the fault degree $m_i$ is greater than the threshold value, it is judged as a fault line, otherwise it is a normal line. According to the above analysis, the fault diagnosis algorithm of power grid based on grey relational calculation is as follows.

(i) **Judge the suspected components.** According to the protection and circuit breaker action information received by the dispatcher, combined with the network topology structure, the power outage area is determined, and the components in the power outage area are regarded as suspected fault components.

(ii) **Detect the fault time.** Detection of fault time based on protection and circuit breaker action information and fault recording data.

(iii) **Calculate the grey relational of each suspected fault line.** According to (4), (6), (8), (9), $r_{iTZ}, r_{iAD}$ and $r_{iED}$ are calculated respectively. $r_{iTZ}, r_{iAD}$ and $r_{iED}$ are calculated similarly.

(iv) **Calculate the fault degree of each suspected fault element.** The obtained grey relational degrees are weighted to obtain the comprehensive grey relational degree of each suspected fault line by use of (12), and the failure degree of components is obtained by means of normalization using (13).

(v) **Determine the faulty component.** According to the fault degree of each suspect component, the faulty component is determined based on the fault threshold.

### 4. Simulation analysis

In this section, simulation analysis of different fault locations, different fault types and high resistance, double fault scenarios on IEEE 14 nodes have been done to verify the effectiveness of the proposed method. The typical structure of IEEE 14 node is shown in Figure 2.

According to the network connection analysis, the power outage area includes the lines L2, L3, L6, L15, L17, and the line in the power outage area is regarded as the suspected fault line, and one active line is selected as the reference line. The current sampling values of suspected fault lines and reference lines in each substation are collected, and the gray correlation degree calculation method described in this paper is used to identify the fault components of the power grid.

The electrical distance between the line L1 outside the power outage area and the suspected lines L2, L3, L6, L15, L17 is 0, 0, 1, 1, 2; the electrical distance between line L4 and each suspect line is 1, 0, 1, 1, 2; the electrical distance between line L5 and each suspect line is 1, 0, 0, 1, 2; the electrical distance between line L12 and each suspect line is 0, 0, 0, 1; line L16 and The electrical...
distance of the suspected fault line is 0, 0, 0, 0, 0. Therefore, L16 is selected as the reference line. The threshold value is set as 0.7.

![Figure 2. The typical structure of IEEE 14 node model.](image)

The amplitude distortion before and after each suspected line fault time are shown in Table 1. It can be seen that only the amplitude distortion of L15 is greater than 3, so L15 is selected as the calculation line. Then the dynamic resolution coefficients of each grey relational degree are calculated separately, and the dynamic resolution coefficients $\xi$ are shown in Table 2.

Table 1. The amplitude distortion before and after each suspected line fault time.

| Lines | L2  | L3  | L6  | L15 | L17 |
|-------|-----|-----|-----|-----|-----|
| The amplitude distortion | 0.9645 | 0.9828 | 1.1164 | 7.7743 | 1.4876 |

Table 2. The dynamic resolution coefficients of each grey relational degree.

| $r_{\text{Z1}}$ | $r_{\text{AO1}}$ | $r_{\text{EO1}}$ | $r_{\text{Z2}}$ | $r_{\text{AO2}}$ | $r_{\text{EO2}}$ |
|----------------|-----------------|-----------------|----------------|----------------|----------------|
| 0.4196         | 0.5719          | 0.6712          | 0.4206         | 0.5802         | 0.6662         |
Then, the gray relational degree $r_{TZ}$, $r_{AD}$, $r_{ED}$, $r_{TZ2}$, $r_{AD2}$ and $r_{ED2}$ are calculated respectively based on the dynamic resolution factor coefficients $\xi$, and the comprehensive grey relational degree $r$, the fault degree $m$ are obtained, which are shown in Table 3.

As can be seen from Table 3, the fault degree of L15 is greater than the threshold value, therefore, L15 is determined as a fault line. According to the diagnosis results, the following failure occurred: the circuit breaker CB9 refuse to move, malfunction of main protection on B4 side of L6, malfunction of backup protection of L2 and L3 and malfunction of circuit breaker CB14. The diagnosis is consistent with the actual failure.

Table 3. The grey relational degree and fault degree of each line.

| Lines | $r_{TZ}$ | $r_{AD}$ | $r_{ED}$ | $r_{TZ2}$ | $r_{AD2}$ | $r_{ED2}$ | $r$ | $m$ |
|-------|---------|---------|---------|---------|---------|---------|------|-----|
| L2    | 0.9979  | 0.9978  | 0.9996  | 0.5554  | 0.9871  |         |      |     |
| L3    | 0.9953  | 0.9997  | 0.9999  | 0.5036  | 0.9249  |         |      |     |
| L6    | 0.9994  | 0.9997  | 0.9999  | 0.5056  | 0.9653  |         |      |     |
| L15   | 0.9480  | 0.9705  | 0.9483  | 0.5543  | 0.9689  |         |      |     |
| L17   | 0.9557  | 0.9534  | 0.9228  | 0.5032  | 0.9487  |         |      |     |

In order to verify the effectiveness of the algorithm, the fault degree of each suspected fault line is calculated when the L15 failure under different locations, different types and high resistance is calculated, which is shown in Table 4. It can be seen that the fault elements can be correctly identified no matter what type of fault, fault location and transition resistance occurs in the event of a grid failure.

Furthermore, in the case of double faults in the power grid, the calculation results using fixed resolution coefficient and dynamic resolution coefficient are shown in Table 5.

Table 4. Fault probability of suspected faulty lines when different types of faults occur in different locations of L15.

| Fault type | Single phase grounding | Three-phase short circuit |
|------------|------------------------|--------------------------|
| Fault location | 10%  | 90%  | 10% high resistance | 90% high resistance | 10%  | 90%  |
| L2         | 0.0086 | 0.0272 | 0.0326 | 0.0412 | 0.0060 | 0.0132 |
| L3         | 0.0053 | 0.0212 | 0.0262 | 0.0192 | 0.0068 | 0.0091 |
| L6         | 0.1711 | 0.0342 | 0.0503 | 0.0446 | 0.0024 | 0.0067 |
| L15        | 0.9468 | 0.9576 | 0.9389 | 0.9523 | 0.9465 | 0.9498 |
| L17        | 0.308  | 0.0794 | 0.4540 | 0.0654 | 0.0101 | 0.0217 |

Table 5. Fault probability of suspected faulty lines when different types of faults occur in different locations of L15.

| Fault type | Two phase grounding | Two phase short circuit |
|------------|---------------------|-------------------------|
| Fault location | 10%  | 90%  | 10% high resistance | 90% high resistance | 10%  | 90%  |
| L2         | 0.0086 | 0.0345 | 0.0220 | 0.0351 | 0.0154 | 0.0210 |
| L3         | 0.0057 | 0.0120 | 0.0162 | 0.0301 | 0.0096 | 0.0130 |
| L6         | 0.0056 | 0.0117 | 0.0216 | 0.0310 | 0.0110 | 0.0189 |
| L15        | 0.9273 | 0.9381 | 0.9328 | 0.9618 | 0.9429 | 0.9490 |
| L17        | 0.0154 | 0.0410 | 0.0406 | 0.0708 | 0.0171 | 0.0394 |
As shown in Table 5, when the resolution coefficient $\xi=0.5$, the L15 fault can be diagnosed, but the L16 fault cannot be diagnosed correctly. The fault components L15 and L16 can be correctly identified by the diagnosis method proposed in this paper. It shows that the dynamic resolution coefficient adopted in this paper improves the accuracy of the diagnosis results compared with the fixed resolution coefficient.

| Case1: High resistance grounding of phase A occurs at 90% of line L16 and phase A grounding fault occurs at 10% of line L15. | L2  | L3  | L6  | L15 | L16 | L17  |
|--------------------------------------------------|-----|-----|-----|-----|-----|-----|
| Fixed resolution coefficient                     | 0.0109 | 0.0101 | 0.0214 | 1.0000 | 0.14324 | 0.0287 |
| Dynamic resolution coefficient                   | 0.1112 | 0.1003 | 0.1735 | 1.0000 | 0.9237 | 0.2713 |

| Case2: High resistance grounding of phase A occurs at 90% of line L16 and phase A and phase B grounding fault occurs at 10% of line L15. | L2  | L3  | L6  | L15 | L16 | L17  |
|--------------------------------------------------|-----|-----|-----|-----|-----|-----|
| Fixed resolution coefficient                     | 0.0154 | 0.0153 | 0.0191 | 0.9972 | 0.1932 | 0.0264 |
| Dynamic resolution coefficient                   | 0.0901 | 0.0896 | 0.1152 | 1.0000 | 0.8963 | 0.1654 |

5. Conclusions
The method of power grid fault diagnosis based on grey relational degree is studied in this paper, due to the advantages of electrical quantity information in accuracy and reliability. The extraction method of feature sequence, the calculation steps of dynamic discriminant coefficient, the calculation method of gradient grey relational degree, the calculation method of transverse grey relational degree and the fault diagnosis criterion based on comprehensive grey relational degree are introduced. Finally, the validity of the diagnosis method is verified by a fault example.

References
[1] Xu, Feng X 2014 Applied Mechanics and Materials 539:659-663
[2] Ye W, Xinyan F, Zheng Y 2009 European Transactions on Electrical Power 19(7):969-977
[3] Zhong JY, Zhang Y, Wen FS, et al. 2015 Automation of Electric Power Systems 39(11):152-159
[4] Dong M, Zhang Y, Zhang Y, et al. 2013 Automation of Electric Power Systems 37(6):55-62
[5] Su GN, Zhang PC, Hu Y, et al. 2012 Automation of Electric Power Systems 36(1): 61-65
[6] Deng J L 1989 Journal of Grey Systems 1(1): 1-24
[7] Shi J, Ding Z, Lee W J, et al. 2014 IEEE Transactions on Smart Grid 5(1):521-526
[8] Jie ZY, Hong AN, Chen SY 2017 Equipment Environmental Engineering
[9] He Linghui, Deng Wei, Gong Jie, et al. 2009 Transformer 46(12):50-53
[10] Pan W, Lei C, Jia W, et al. 2018 Grey Relational Grade Based Quantitative Analysis of the Factors Influencing the Load Characteristics of A Power Grid. 3rd International Conference on Advances in Energy and Environment Research
[11] Wang Y, Huang Y, Zeng X, et al. 2017 IEEE Transactions on Power Delivery 32(1):55-61
[12] Li NY, Liang J 2014 Automation of Electric Power Systems 38(16):100-105
[13] Han YC, Tong XY 2017 Power System Technology 41(2):581-587