This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2728015, IEEE Access

Research on Distribution Network Fault Recognition Method based on Time-frequency Characteristics of Fault Waveforms

Xue QIN [1], Peng WANG [2], Yadong LIU [1], Linhui GUO[1], Gehao SHENG[1], Xiuchen JIANG[1]

Abstract— **Accurate recognition of distribution line fault types can provide directional guidance for line operation and maintenance personnel. Based on the analysis of time-frequency features of fault waveform, a recongnized method of distribution line fault type was proposed in this paper. Through modeling and theoretical analysis of waveforms of different fault types, characteristic parameters which could characterize waveforms of different fault types from three aspects, time domain, frequency domain, and electric arc were put forward. Calculation formula for extracting characteristic parameters according to fault waveform data was proposed, recognition logic was established by taking multi-parameter fusion as a basis, and then automatic recognition of distribution line fault types caused by different factors was realized through detection and classification of characteristic parameters of input waveform data. Finally, 136 groups of field fault waveform data provided by EPRI (Electric Power Research Institute) were used to do closed-loop control and verification of the algorithm, and results indicated that recognition success rate reached 90%, which verified the feasibility of using time-frequency characteristics of fault waveform to realize recognition of distribution line fault types.**

Index Terms—**Distribution Line; Fault Cause; Time-frequency Characteristics; Electric Arc Model; Recognition Logic**

I. Introduction

It is well known that 10kV overhead distribution line is of large scale, wide range, long wire, low insulation level and persistent high tripping fault rate. As an important constituent part of power distribution network, accurate judgment of line fault type of distribution network through fault field waveform is of great significance to improving the operation and maintenance level thereby reducing the fault rate of power distribution line.

A significant amount of work has been carried out in the areas of fault diagnosis and automatic classification of power quality disturbance events. For example, in [1] the distribution network fault is classified by logistic regression and artificial neural network. A method to solve the problem of PQ data uncertainty by using the rough cloud theory is proposed in [2]. In addition, a time-domain simulation model is used to identify and locate the cable fault in the literature [3].

However, not much importance has been given to the waveform characterization of fault events, which is the first step to root-cause analysis. Literature [4] presents an identification method of lightning strikes based on traveling waves' time-domain parameters. The author of [5] analyzed characteristics of faults caused by animals and proposed corresponding preventive measures, but didn't mention line faults caused by other factors. In literature [6], faults caused by animals, lightning and trees were analyzed, three-phase voltage and current waveforms of the fault were proposed as judgment basis. However, it only discussed about the feasibility of the judgment basis but didn't establish complete and reliable recognition logic system.

In this paper, for faults caused by animals, trees, vehicles, lightning and equipment, characteristic parameters which could characterize waveform features of these faults from three aspects, time domain, frequency domain and electric arc features were proposed. The calculation formula of extracting characteristic parameters was put forward according to fault waveform data. And complete recognition logic was established by taking multi-parameter fusion as a basis. Then automatic recognition of distribution line fault types caused by different factors was realized through detection and classification of characteristic parameters of input waveform data. Finally, 136 groups of field fault waveform data provided by EPRI (Electric Power Research Institute) were used to do closed-loop control and verification of the algorithm, and results indicated that recognition success rate reached 90%.

II. Analysis of Fault Waveform Characteristics

There are many factors causing distribution line faults with interference on the quality of electric energy, which lead to five fault types, tree swinging cross fault, faults caused by animals, vehicles, lightning and equipment.

Firstly, waveform characteristics of different fault types were analyzed in the time domain [7]. Typical voltage and current waveforms of distribution line fault are as shown in Fig. 1. Number of faulted phases, fault duration, fault current component and current attenuation degree under open line fault

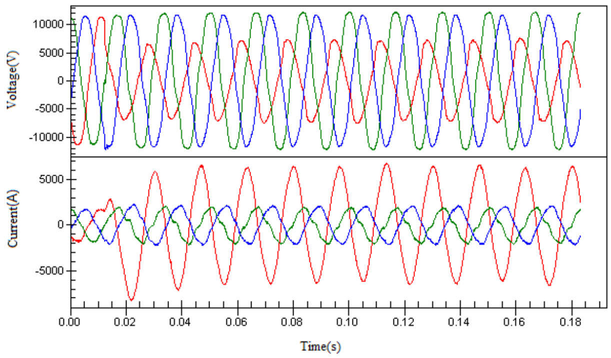
1. School of Electronic Information and Electrical Engineering, Shanghai Jiaotong Unversity, Shanghai 200240; 2. Electric Power Research Institute of Electric Power of Henan, State Grid, Zhengzhou, Henan Province 450052;

Corresponding author: Yadong LIU (lyd@sjtu.edu.cn)

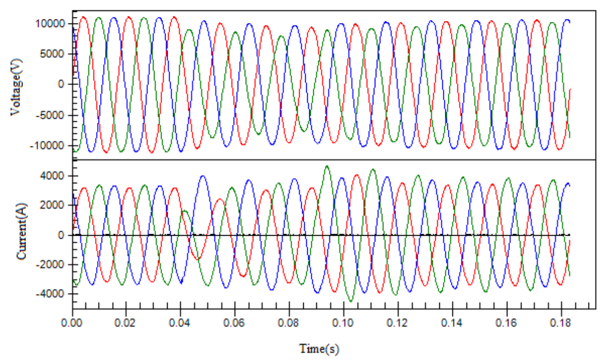
This work was supported in part by the National Natural Science Foundation of China under Grant 51307109 and in part by the State Grid Corporation of China.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2728015, IEEE Access

was extracted as characteristic parameters for follow-up analysis.



(a) Non-metallic single-phase ground fault



(b) Disconnection fault

Fig. 1. Typical distribution line fault voltage and current waveforms

Secondly, in consideration, that fault waveform also contains shore pulse or radio-frequency components like a high-frequency oscillation, which is caused by external damages from trees and lightning may contain more energy. Therefore, three-scale orthogonal wavelet transform was used in this paper to analyze fault phase voltage in the frequency domain, wavelet transform results are as shown in Fig. 2, and then frequency spectrum and energy distribution of fault waveform in different frequency bands were studied.

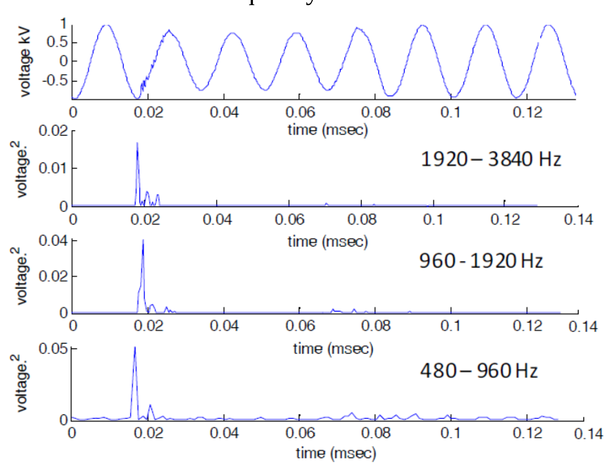
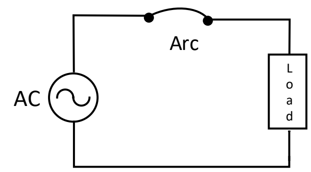
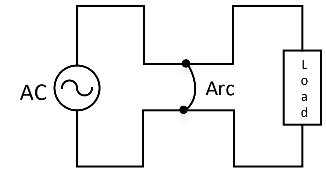


Fig. 2. Wavelet transform results of fault phase voltage

In addition, external faults like tree pressing and internal faults like insulator breakdown will generate electric arcs in the medium, and different fault types can be equivalent to series arc or parallel arc model for research as shown in Fig. 3. Hence, besides analysis in the time domain and frequency domain, this paper will analyze electric arc characteristics of different fault types.



(a) Series arc fault



(b) Parallel arc fault

Fig.3. Different types of arc fault model

A. Time-domain features of different fault types

1. Number of faulted phases

When a tree swinging cross fault occurs, adverse weather like gale will easily occur in summer and autumn, so under wind effect, branches will easily contact wire or trees so as to press the wire, which is generally single-phase ground fault [8]. In addition, branches falling on the wire, the interphase short circuit may be caused or trees collapse to press the wire and directly cause a three-phase open line fault.

Animal activities in spring and summer are quite frequent, animals like snakes will easily climb onto transformer top or pole tower top during the predation process, as snakes are long, they may contact high-voltage binding post or cause insulator short circuit which gives rise to a network fault. Moreover, in the spring, birds build nest and incubate on pole tower of distribution line, when birds on pole tower are too close to the wire, the birds nest will easily be blown by gale onto wire or insulator, which then causes insulator short circuit and forms single-phase grounded short circuit.

When an automobile runs into cable protection pile on a telegraph pole, the inside cable will be damaged and even the telegraph pole will collapse, mixed wire connection will be caused when telegraph pole is hung on the wire so that interphase short circuit fault will occur [9]. In addition, line fault caused by crane arm pressing will probably happen. When crane construction is being implemented beside distribution line, if the crane is too close to the wire during lifting or arm swiveling process, phase-earth short circuit will happen.

When multi-phase simultaneous flashover is caused by induced over-voltage generated by lightning, arcs will be built on the channel of lightning-stricken flashover so as to form power-frequency follow current, and current amplitude of power-frequency short circuit is high (several kA) and time is short (0.1s level). As electric arc roots generated by lightning fault can't slide along the insulating wire but can only be fixed at lightning-stricken flashover to do centralized discharging [10], power-frequency short-circuit current can burn out the insulated wire. So most faults caused by lightning are multi-phase disconnection faults.

Humid air in spring and summer may dampen electrical equipment so that the insulating property of the equipment will degrade, which will seriously affect normal and safe system actuation. Equipment faults (like branch fuse insulation breakdown on the line, high-voltage lead disconnection of distribution transformer, insulation breakdown of fuse or lightning arrester on distribution transformer, or insulator

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2728015, IEEE Access

breakdown, high-voltage winding single-phase insulation and grounding or breakdown of distribution transformer, etc.) can result in wire disconnection of one phase, or the wire may fall on cross arm to form ground fault. Hence, faults caused by insulation aging inside the equipment are generally single-phase faults.

To sum up, trees, animals, vehicles and equipment fault will cause single-phase ground fault, both vehicles and trees will cause an interphase fault, both trees and lightning will cause a three-phase open-line fault, number of faulted phases can be used to differentiate fault types.

2. Fault duration

Many overhead line faults can by self-restored and even they have already been removed before actuation of protection device like swinging cross caused by gale or branches, etc., these faults have been solved before actuation of the protection device, so their durations are commonly short [11]. However, when ground short-circuit of the distribution line is caused by animals climbing onto the line or transformer, large current flowing into their bodies will instantaneously electrocute them and cause persistent ground fault until actuation of the protection device[12].

Generally speaking, zero-sequence current or zero-sequence power protection can be adopted for single-phase grounded short circuit of the 10kV distribution network. Actuation time of relay set is usually about 20ms, actuation time of breaker =protection actuation time + inherent breaker opening time, protective actuation time is usually about 30ms, and breaker opening time is generally 40-60ms[13]. Hence, it takes at least 90ms from fault occurrence to fault disappearance, namely 5-6 cyclic waves, which can be differentiated according to the long duration of single-phase ground fault generated by animals.

3. Fault current component

When the circuit is under normal operation, current is decided by load power, and according to code standards for urban electric power planning, the 10kV distribution network is generally smaller than 15,000kW, so phase current under normal operation is smaller than $\frac{15000}{\sqrt{3}\times 10}=866A$. Tree swinging cross and vehicle running into pole may cause an interphase fault. When interphase short circuit happens because of tree swinging cross, the broken branches are hung on two wires to build up a “bridge”, as persistent current heating will give rise to carbonization of branches, reduced resistance facilitates aggravation of carbonization process, finally a carbonization channel with extremely low resistance and large fault current are formed.

At the moment, composite sequence network is as shown in Fig. 4, and fault-phase current amplitude is $I=\frac{\sqrt{3}E_a}{Z_1+Z_2+3Z_k}$ where $Z_1=Z_2$ is positive sequence impedance of the circuit, Z_k is the resistance of carbonization channel, being about 1Ω, overhead insulated wire is usually used for 10kV distribution network and its resistance is about 0.5Ω [9]. Through calculation, fault-phase current can reach 4,330A and fault current component is $4330-866=3465A$.

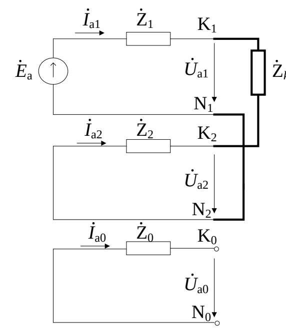


Fig. 4. Phase short circuit composite sequence network

When a mixed wire connection is caused by a vehicle which hits a pole, line impedance is large, so fault current amplitude is much smaller than the interphase fault current amplitude caused by tree swinging cross, and fault current component is the basis for differentiating interphase fault.

4. Attenuation degree of open circuit fault current

Both lightning and tree pressing off the wire can cause an open circuit fault. The horizontal distance between neighboring pole towers of 10kV distribution line is usually 100m, so stress analysis is conducted when trees at two sides of the line collapse to press the line, as the acting force perpendicular to the wire is too small to cause three-phase simultaneous disconnection. As the point of three-phase open circuit caused by tree is generally close to middle part of the line, and broken line-phase current is half of sound phase current; however, as 10kV overhead distribution line itself is not of strong lightning attraction effect, its tail end is close to user side, there are many towers and keen-edged buildings in the surroundings, some electromagnetic interference facilities may result in enhanced activities which may cause ground lightning, so open circuit point is generally close to tail end of the line, and broken line-phase current can only be slightly smaller than that under normal operation. The ratio of broken line-phase current to phase current under normal operation is defined in this paper as current attenuation degree in broken line fault, namely:

$$\alpha_{ATT}=\frac{I_{pkMin}}{I_{0pk}} \quad (1)$$

Where I_{pkMin} is minimum value of phase current amplitude in fault duration, I_{0pk} is phase current amplitude before fault occurrence, and α_{ATT} can be used as characteristic parameters of differentiating different fault types.

B. Frequency-domain characteristics of different fault types

When a current shock component appears in the system, it may trigger a series of shock components, short pulses or high-frequency components like high-frequency oscillation. Wavelet transform can be used to analyze frequency spectrum, and three-scale orthogonal wavelet transform is used in this paper to analyze fault-phase voltage. Fig. 5 and Fig. 6 are a wavelet transform results of fault-phase voltage when faults are respectively caused by lightning and tree to the lines of the same voltage class, it can be seen that amplitudes and energy distribution of frequency bands are different in the line faults caused by different factors. The fault caused by lightning

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2728015, IEEE Access

results from traveling waves generated during pulse propagation, and consequently, the energy of high-frequency component in the 1,920-3,840Hz frequency band is 10 times of that of the fault caused by tree swinging cross.

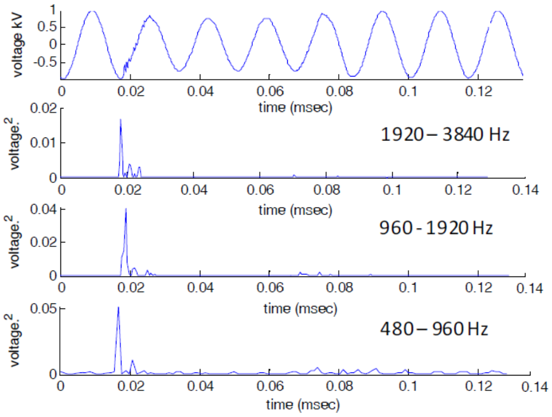


Fig. 5. Wavelet transform of fault phase voltage caused by lightning

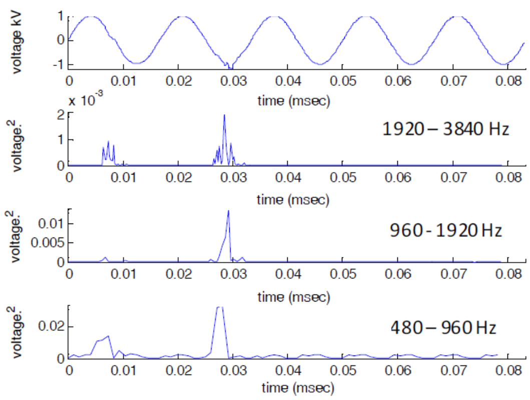


Fig. 6. Wavelet transform of fault phase voltage caused by tree

C. Electric arc characteristics of different fault types

An electric arc is a continuously glowing and discharging process after the insulating medium is penetrated. Fault arc can be divided into series type and parallel type by occurrence method. For electric arcs in series, within the single charged conductor, if poor contact happens to the connection point, then voltage difference will be generated to break down the gap and form the electric arc, and at the moment, electric arcs will develop along a conductive surface of insulator part as shown in Fig. 3 (a). For electric arcs in parallel, if insulator aging happens, a carbonization channel will be formed after a while. When the insulation level of the line is reduced to a certain critical value, carbonization channel will be broken down by the voltage between conductors to form arcing short circuit as shown in Fig. 3 (b). When an external fault happens, namely tree or vehicle swinging cross will cause high-voltage discharging, high-voltage electricity is grounded through tree trunks so that high-voltage single-phase ground fault is caused, so it can be regarded as parallel arc fault; but equipment fault faults caused by insulation aging and poor contact can be regarded as a series arc fault.

As the most important characteristic parameter of the electric arc, arc voltage contains key information causing power quality disturbance. High-frequency component manifested by nonlinear vibration of electric arc approximates arc voltage waveform to a rectangular wave (actual arc voltage and current waveform as shown in Fig. 7).

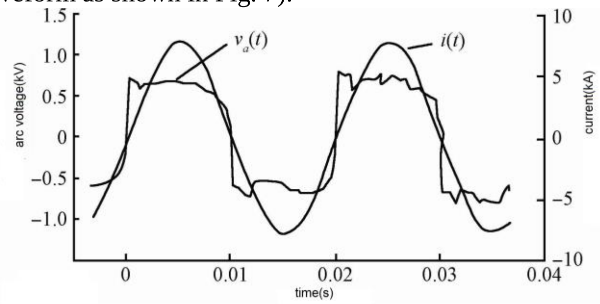


Fig.7. The actual arc voltage and current waveform

It's assumed that arc voltage is a square wave with the phase position being equal to arc fault current, then the arc voltage can be expressed in a simple mathematical expression:

$$v_a(t) = V_a \operatorname{sgn}[i(t)] + \xi(t) \tag{2}$$

Where $v_a(t)$ and $i(t)$ are respectively arc voltage and current, V_a is amplitude of square wave, $sign$ is sign function and $\xi(t)$ is zero-mean white noise. Based on the arc voltage model of the above expression, the computer simulation results of arc voltage are as shown in Fig. 8.

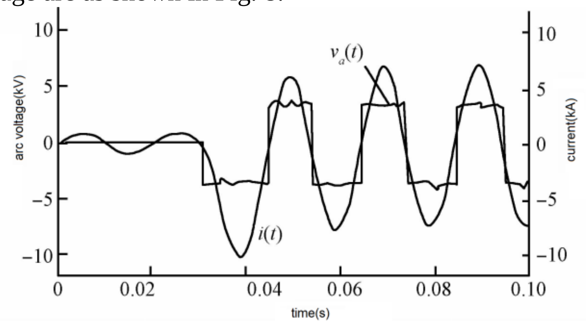


Fig. 8. Simulation of arc voltage and current waveform

Hence, according to the formula proposed in [14-15], the amplitude of single-phase fault arc voltage can be calculated according to the model as shown in Fig. 9.

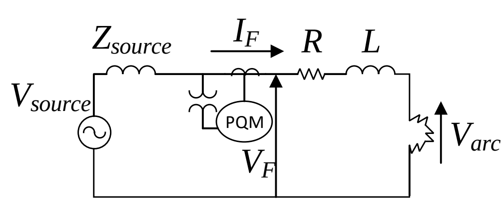


Fig. 9. Arc voltage magnitude estimation model of single-phase fault [16]

So voltage V_f monitored at fault side can be expressed by equation (3), least square method can be used to solve the solution of the mathematical expression decided by multiple factors, and then the arc voltage V_{arc} is obtained.

$$V_f = R \times I_f + L\frac{dI_f}{dt} + V_{arc} \times sign(I_f) \tag{3}$$

V_f : Fault phase voltage L : Line inductance

R : Line resistance V_{arc} : Arc voltage

$sign(I_f) = 1$, if $I_f > 0$ and -1 if $I_f \le 0$

When an external fault happens, for example, high-voltage electricity is grounded through trunks because of a tree or vehicle swinging cross, as there is no metallic hard connection between the tree and high-voltage power line, and loosened electrical connection will lead to arc generation. When there is loosened contact at connection points of the electrical circuit, when inter-point voltage is enough to break down air gap, air conductivity will be formed. If air gap of connection points is large and coincides with a peak value of voltage waveform, electric arcs will be pulled up in the air; if air gap of connection points is large, then even though the voltage is not large, air can also be broken down so as to generate an electric arc. In this way, continuous air ionization is caused, and as contacts are continuously pulled open, the electric arc will become longer. However, electric arc caused by a vehicle is of long discharging distance, so voltage amplitude is usually larger. Under general conditions, discharging in air gap mostly happens on the shortest path between grounding body and electric conductor, but the channel formed by soot and heat dispersed in the fault process is of low insulation strength, and as a result, discharging usually happens not along the gap between pole tower-wire or insulator string but along firework channel from relatively long distance, so electric arc is long.

Comparatively speaking, for internal faults like equipment fault, the insulator is heated for a long term, leakage current or spark discharge happens between wires and the generated heat will give rise to insulator decomposition, a conductive carbonization channel is then formed between wires so as to cause electric arc, so it can be regarded as series arc fault. As there is no contact with the external environment, the electric arc is short.

Field intensity of fault arc nearly keeps constant along the arc and mean value is 1.3kV/m [17]. Hence, fault arc voltage is in direct proportion to arc length, and arc voltage can be used to differentiate external fault from internal fault.

To sum up, number of faulted phases, fault duration, fault current component and open circuit fault current attenuation degree can be used to differentiate fault types in the time domain. Fault duration can be used to differentiate a single-phase fault caused by animal from a single-phase fault caused by others, and threshold value $n_{CycleSet}$ is initially set as 6 cyclic waves; fault current component can differentiate interphase short circuit caused by trees from other interphase fault types, and threshold value $I_{faultSet}$ is initially set as 3,465A; when open circuit fault happens, attenuation degree of fault current can be used for fault differentiation, and initial threshold value α_{ATTset} is set as 0.5. In the frequency domain, line faults caused by different factors are different in their amplitude and energy distribution in different frequency bands, so amplitude and energy distribution can be used to differentiate different fault types. In addition, arc voltage can be used to differentiate internal fault from external fault, when the internal fault of the line happens, arc length is usually smaller than 40cm [17], so initial threshold value V_{arcSet} is set as $1300 \times 0.4 = 520V$.

III. Recognition Method of Distribution Line Fault Types Based on Time-Frequency Characteristics

A. Extraction and calculation of characteristic parameters

Based on the analysis of section II, number of faulted phases, fault duration, fault current component, open circuit fault current attenuation degree and frequency-domain energy of fault voltage and arc voltage can be characteristic parameters to realize differentiation and recognition of distribution network fault types. Hence, firstly processing method of input waveform data will be expounded in this section and then calculation formulas of these characteristic parameters will be put forward.

It's assumed that input waveform data are monitored and sampled data on the fault site including sampling time, three-phase current and three-phase voltage. N represents total a number of sampling points, then:

$$N_0 = \frac{N}{60 \times [t(N) - t(1)]} \tag{4}$$

$$n_{cycle} = \frac{N}{N_0} \tag{5}$$

In the above formulas, N_0 represents number of sampling points of one cycle and n_{cycle} represents number of sampling cycles.

1. Number of faulted phases

Number of faulted phases is judged according to mutation of fault phase current, and method is as below:

Firstly maximum value of sampling sequence of three-phase current, namely

$$\begin{aligned} I_{a\max} &= \max\left|I_{a(n)}\right|, \ n \in [1, N] \\ I_{b\max} &= \max\left|I_{b(n)}\right|, \ n \in [1, N] \\ I_{c\max} &= \max\left|I_{c(n)}\right|, \ n \in [1, N] \end{aligned} \tag{6}$$

In the above formula, $I_{a(n)}$, $I_{b(n)}$ and $I_{c(n)}$ are respectively sampling sequences of three-phase current, and their ratio relationships are calculated as below:

$$R1_{I_f} = \frac{\max(I_{a\max}, I_{b\max}, I_{c\max})}{median(I_{a\max}, I_{b\max}, I_{c\max})} \tag{7}$$

$$R2_{I_f} = \frac{median(I_{a\max}, I_{b\max}, I_{c\max})}{\min(I_{a\max}, I_{b\max}, I_{c\max})} \tag{8}$$

In the above formulas, $R1_{I_f}$ and $R2_{I_f}$ represents the ratios between every two phases after maximum values of sampling points of three phases of the three-phase current are sorted in a descending order. The judgment formula of fault phase $n_{faultPhase}$ proposed in this paper is shown as below:

$$n_{faultPhase} = \begin{cases} 3, (R1_{I_f} \le k_1) \& (R2_{I_f} \le k_2) \\ 2, (R1_{I_f} \le k_3) \& (R2_{I_f} \ge k_4) \\ 1, (R1_{I_f} \ge k_5) \& (R2_{I_f} \le k_6) \end{cases} \tag{9}$$

2. Fault duration

For the research convenience, firstly current of one phase is firstly selected and stipulated as fault phase current I_{fault}, and its judgment formula is as below:

$$I_{pkMax} = \max(I_{a\max}, I_{b\max}, I_{c\max}) \tag{10}$$

Then

$$I_{fault} = \begin{cases} I_a, I_{pkMax} = I_{a\max} \\ I_b, I_{pkMax} = I_{b\max} \\ I_c, I_{pkMax} = I_{c\max} \end{cases} \tag{11}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2728015, IEEE Access

$$I_{pk(j)} = \max\left|I_{fault(n)}\right| \quad (12)$$

Where $j \in [1, n_{cycle}]$ and $n \in [(j-1)N_0, jN_0]$

In the above formulas, $I_{pk(j)}$ represents the peak current value of the j[th] cycle of fault phase current, when it is greater than one threshold value, then it can be considered that this cycle is under fault status so that the number of cycles that the fault lasts can be judged and its formula is as below:

$$n_{faultCycle} = \sum_{j=1}^{n_{cycle}} \mathrm{sgn}\left\{\left[\frac{I_{pk(j)}}{I_{pk0}}\right]\right\} \quad (13)$$

3. Fault current component

Calculation formula of fault current component $I_{faultAmp}$ is as below:

$$I_{faultAmp} = I_{pkMax} - I_{0pk} \quad (14)$$

In the above formula, I_{0pk} is peak value of three-phase current under normal operation before the fault and it is obtained through the following formula:

$$I_{0pk} = \min(I_{a(n)}, I_{b(n)}, I_{c(n)}) \ , \quad n \in [1, N_0] \quad (15)$$

4. Open circuit fault current attenuation degree

For open circuit fault, fault phase current attenuates, then peak current value of the j[th] cycle of the three-phase current is as below:

$$\begin{cases} I_{apk(j)} = \max\left|I_{a(n)}\right| \\ I_{bpk(j)} = \max\left|I_{b(n)}\right| \\ I_{cpk(j)} = \max\left|I_{c(n)}\right| \end{cases} \quad (16)$$

Where $j \in [1, n_{cycle}]$ and $n \in [(j-1)N_0, jN_0]$

Minimum value of peak three-phase current value is taken as below

$$I_{pkMin} = \min(I_{apk}, I_{bpk}, I_{cpk}) \quad (17)$$

Then fault current attenuation degree α_{ATT} is defined and its calculation formula is as below:

$$\alpha_{ATT} = \frac{I_{pkMin}}{I_{0pk}} \quad (18)$$

5. Frequency-domain energy percentage

For frequency-domain analysis of fault phase voltage energy distribution, normalization processing is firstly conducted:

$$V_{fault} = \begin{cases} V_a, I_{pkMax} = I_{a\max} \\ V_b, I_{pkMax} = I_{b\max} \\ V_c, I_{pkMax} = I_{c\max} \end{cases} \quad (19)$$

$$V_{norm} = \frac{V_{fault}}{V_{fault\max}} \quad (20)$$

The wavelet transform is used to calculate energy percentages of the fundamental frequency and high-frequency bands:

$$(\mathrm{E}_{BF}, \mathrm{E}_{HF}) = \mathrm{wenergy}(\mathrm{wave}dec(V_{norm})) \quad (21)$$

6. Arc voltage

Calculation formula of arc voltage is obtained according to formula (3) as below:

$$V_{fault(n)} = R_{(j)} \times I_{fault(n)} + L_{(j)} \frac{dI_{fault(n)}}{dt_{(n)}} + V_{arc(j)} \times sign(I_{fault(n)}) \quad (22)$$

Where $\mathrm{j} \in [1,\ \mathrm{n}_{cycle} \times 2 - 1]$ and $n \in \left[\frac{N_0 j}{2} - \frac{N_0}{2} + 1, \frac{N_0 j}{2}\right]$

The least square method is used to solve the variance:

$$\sigma^2(\mathrm{j}) = \frac{\sum_{i=j}^{j+2}(V_{\mathrm{arc(i)}} - \frac{1}{3}(V_{\mathrm{arc(j)}} + V_{\mathrm{arc(j+1)}} + V_{\mathrm{arc(j+2)}}))^2}{3} \quad (23)$$

Where $\mathrm{j} \in [1,\ \mathrm{n}_{cycle} \times 2 - 3]$

When the minimum value of $\sigma^2(\mathrm{j})$ is taken, it's deemed that arc voltage amplitude is stable, and mean value of three calculated values within the interval is taken as arc voltage V_{arc} of this model.

B. Recognition logic

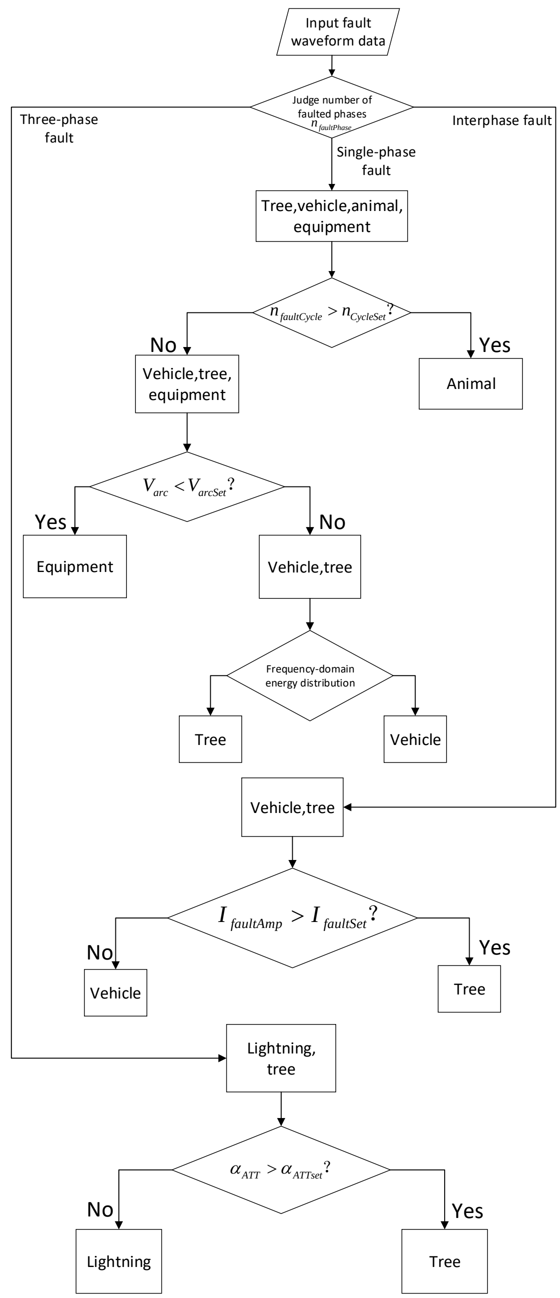


Fig. 10. Distribution Line Fault Type Identification Logic

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2728015, IEEE Access

Based on the analysis in section II and calculation formulas put above, the classifier is built based on multi-parameter fusion in this section, and automatic recognition of distribution line fault types caused by different factors is realized through detection and classification of input waveform data. The overall logic is shown in Figure 10.

1) Input fault waveform data and judge number of faulted phases $n_{faultPhase}$.

2) Re-differentiation of interphase fault. Fault current component is used to differentiate interphase fault caused by tree swinging cross from that caused by vehicle colliding pole. When $I_{faultAmp} > I_{faultSet}$, it's judged that tree swinging cross fault happens.

3) Re-differentiation of three-phase fault. The fault current attenuation degree is used to differentiate three-phase open circuit fault caused by lightning from that caused by tree pressing off the wire. When $\alpha_{ATT} > \alpha_{ATTset}$, it's judged that broken line fault caused by tree happens.

4) Re-differentiation of single-phase fault. Firstly fault duration is used to differentiate fault caused by animal from other fault types, when $n_{faultCycle} > n_{CycleSet}$, the fault is judged as fault caused by animal; secondly arc voltage amplitude is used to differentiate internal fault from external fault, and when $V_{arc} < V_{arcSet}$, it's judged that an internal fault happens; finally frequency-domain energy distribution is used to differentiate single-phase ground fault caused by tree from that caused by a vehicle.

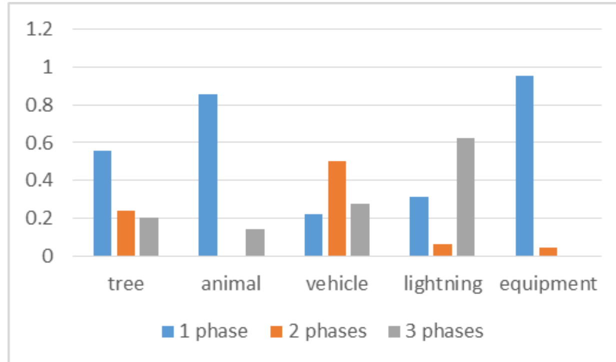
IV. CASE ANALYSIS AND RECOGNITION RESULTS

A. Data source

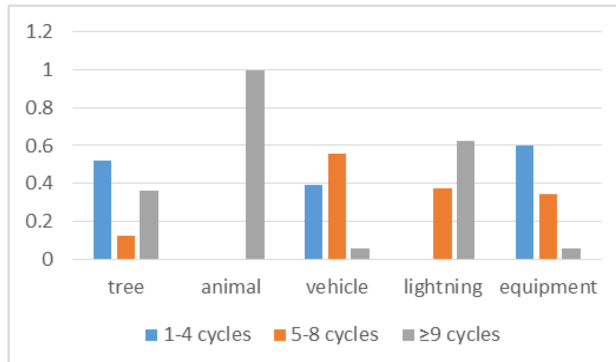
136 fault cases of distribution line provided by EPRI are used in this section to verify the logic algorithm proposed in section III, including 7 faults caused by animal climbing, 70 faults caused by equipment aging, 16 faults caused by lightning, 25 faults caused by tree pressing and 18 faults caused by vehicles. Monitored and recorded data include three-phase voltage, three-phase current and neutral current. Voltage class is 10-20kV, so it belongs to a high-voltage line. The sampling frequency is 60Hz and sampling numbers of different cycles are uncertain and usually are 16, 128 and 256.

B. Introduction of fault data characteristics

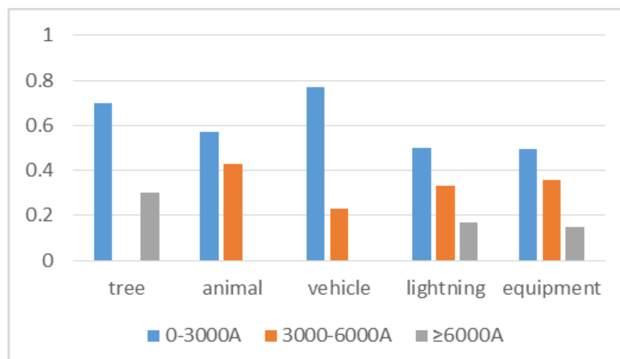
To guarantee reasonability and comprehensiveness of data verification, a statistic of time-frequency characteristic parameters of 136 groups of data is firstly made in this section, including number of faulted phases, fault duration, fault current component, arc voltage amplitude and fundamental frequency energy percentage distribution of different fault types and results are as shown in Fig. 11. It can be seen that the distribution results accord with the analysis made in section II so that comprehensiveness and reliability of data verification are guaranteed.



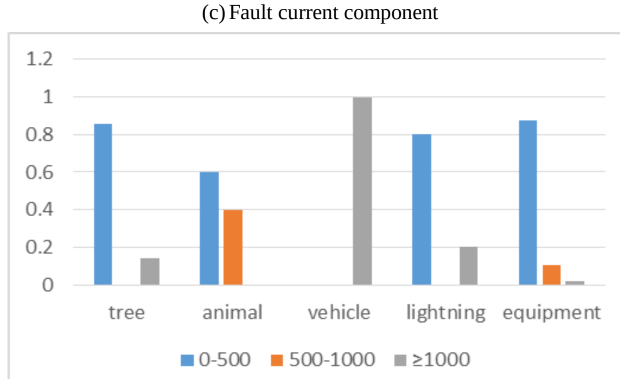
(a) Number of faulted phases



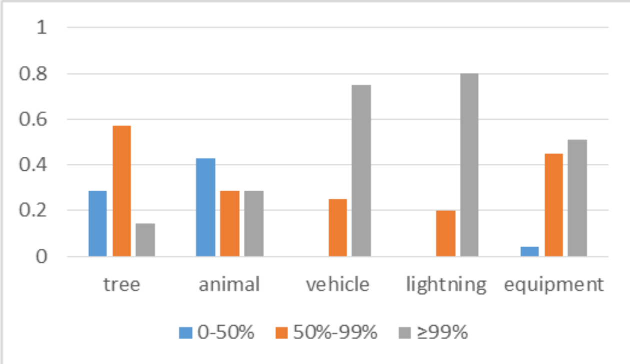
(b) Fault duration



(c) Fault current component



(d) Arc fault voltage amplitude



(e) Fundamental frequency energy percentage

Fig. 11. Time-frequency characteristics' distribution of test data

C. *Recognition process and results*

In consideration of the difference between American distribution network standard and domestic standard and to make sure that recognition results are more accurate, firstly, 30 groups of data are used in this section as training samples and input into the system to enter self-learning phase, and then the adjustment of threshold values and closed-loop control of the whole verification system are realized. These data include 2 faults caused by an animal, 15 ones caused by equipment aging, 4 ones caused by lightning, 5 ones caused by tree and 4 ones caused by a vehicle. Through learning of these training samples, initial threshold values set in section 2 are adjusted to a certain degree, and values before and after adjustment are seen in Table 1.

TABLE 1. Adjustment of the threshold values

Threshold value	Initial value	Adjusted value
Duration	6 cycles	9 cycles
Fault current component	3,465A	6,000A
Arc voltage	520V	500V

And then the rest 106 groups of data are used as test samples and input into the system in batches for verification, and faults types are output. Finally 95 groups are verified successfully and the average success rate reaches 90%, including 4 among 5 faults caused by animal climbing are verified successfully and the success rate is 80%; 50 among 55 faults caused by equipment aging are verified successfully and the success rate is 90%; 10 among 12 faults caused by lightning are verified successfully and the success rate is 81%; 18 among 20 faults caused by tree pressing are verified successfully and success rate is 88%; 13 among 14 faults caused by vehicle are verified successfully and the success rate is 95% as shown in Table 2.

TABLE 2. Recognition results of different fault types

Fault type	Number of test samples	Verification success	Verification failure	Success rate
Animal	5	4	1	80%
Equipment	55	50	5	90%
Lightning	12	10	2	81%
Tree	20	18	2	88%
Vehicle	14	13	1	95%
Total	106	95	11	90%

Through analysis of 11 cases of verification failure, it can be found that threshold values after closed-loop adjustment reaches high levels, but a small part of incident data still deviate from what is expected, so recognition error is caused. If the number of samples increases, system self-learning will be more perfect and accuracy will be greatly improved.

V. Conclusion

Accurate recognition of fault types of distribution line can help operation and maintenance unit to rapidly and accurately judge fault types and solve the faults, and in the meantime, specific precautionary measures can be formulated according to historical data. Characteristics of fault waveforms of 5 faults types in the time domain, frequency domain and arc voltage were concluded in this paper through modeling and theoretical analysis, and characteristic parameters which could characterize different fault types were proposed. Multi-parameter fusion was used to build recognition logic, and then automatic recognition of distribution line fault types caused by different factors was realized through detection and classification of input characteristic parameters of waveform data. 136 groups of waveform data on the fault site provided by EPRI were used to do verification and recognition and success rate reached 90%.

References

[1] L. Xu and Mo-Yuen Chow, "A classification approach for power distribution systems fault cause identification," in *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 53-60, Feb. 2006.

[2] Q. Sun, X. Liu, H. Zhang, "Power Distribution Fault Diagnosis Based on Rough-Cloud Sets," in *Asia-Pacific Power and Energy Engineering Conference*, March, 2009, pp. 1-4.

[3] W. Zhao, Y. H. Song, and Y. Min, "Wavelet Analysis Based Scheme for Fault Detection and Classification in Underground Power Cable Systems," *Elsevier Science Journal of Electric Power Systems Research*, vol. 53, no. 1, pp. 23-30, 2000.

[4] Y. Liu *et al.*, "Identification of Lightning Strike on 500-kV Transmission Line Based on the Time-Domain Parameters of a Traveling Wave," in *IEEE Access*, vol. 4, no.1 , pp. 7241-7250, 2016.

[5] Mo-yuen Chow and L. S. Taylor, "Analysis and prevention of animal-caused faults in power distribution systems," in *IEEE Transactions on Power Delivery*, vol. 10, no. 2, pp. 995-1001, Apr 1995.

[6] S. Kulkarni, D. Lee, A. J. Allen, S. Santoso and T. A. Short, "Waveform characterization of animal contact, tree contact, and lightning induced faults,"*IEEE PES General Meeting*, Minneapolis, MN, 2010, pp. 1-7.

[7] W. Fan and Y. Liao, "Automated analysis of voltage and current waveforms during faults," *Proceedings of the 2012 44th Southeastern Symposium on System Theory (SSST)*, Jacksonville, FL, 2012, pp. 214-219.

[8] L. Xu, M. y. Chow and L. S. Taylor, "Data Mining and Analysis of Tree-Caused Faults in Power Distribution Systems," *2006 IEEE PES Power Systems Conference and Exposition*, Atlanta, GA, 2006, pp. 1221-1227.

[9] Y. Liu *et al.*, "Distributed fault location based on comprehensive analysis of fault current," *2012 IEEE International Conference on Condition Monitoring and Diagnosis*, Bali, 2012, pp. 999-1002.

[10] H. Hou, G. Sheng and X. Jiang, "Localization Algorithm for the PD Source in Substation Based on L-Shaped Antenna Array Signal Processing," in *IEEE Transactions on Power Delivery*, vol. 30, no. 1, pp. 472-479, Feb. 2015.

[11] T. S. Sidhu and Z. Xu, "Detection of Incipient Faults in Distribution Underground Cables," in *IEEE Transactions on Power Delivery*, vol. 25, no. 3, pp. 1363-1371, July 2010.

[12] S. Kulkarni, S. Santoso and T. A. Short, "Incipient Fault Location Algorithm for Underground Cables," in *IEEE Transactions on Smart Grid*, vol. 5, no. 3, pp. 1165-1174, May 2014.

[13] Q. Yang *et al*. "Measurement and analysis of transient overvoltage distribution in transformer windings based on reduced-scale model," in *Electric Power Systems Research*, vol.140, pp. 70-77, 2016.

[14] Z. M. Radojevic and V. V. Terzija, "Fault Distance Calculation and Arcing Faults Detection on Overhead Lines Using Single End Data," *2008 IET 9th International Conference on Developments in Power System Protection (DPSP 2008)*, Glasgow, 2008, pp. 638-643.

[15] T. A. Short, D. D. Sabin and M. F. McGranaghan, "Using PQ Monitoring and Substation Relays for Fault Location on Distribution Systems," *2007 IEEE Rural Electric Power Conference*, Rapid City, SD, 2007, pp. B3-B3-7.

[16] V. Barrera Núñez, S. Kulkarni, S. Santoso and M. F. Joaquim, "Feature analysis and classification methodology for overhead distribution fault events," *IEEE PES General Meeting*, Minneapolis, MN, 2010, pp. 1-8.

[17] Z. M. Radojevic, V. V. Terzija and N. B. Djuric, "Numerical algorithm for overhead lines arcing faults detection and distance and directional protection," in *IEEE Transactions on Power Delivery*, vol. 15, no. 1, pp. 31-37, Jan 2000.