IOP Conference Series: Earth and Environmental Science

PAPER • **OPEN ACCESS**

Internal Over-Voltage Identification Method of Distribution Network Based on AD-SVM Algorithm

To cite this article: Yufei Liao *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **223** 012013

Internal Over-Voltage Identification Method of Distribution Network Based on AD-SVM Algorithm

Yufei Liao[1], Gengjie Yang[1,*], Wei Gao[1] and Moufa Guo[1]

[1]College of Electrical Engineering and Automation, Fuzhou University, Fuzhou, 350116, China;

*ygj23802@126.com; Tel.: +86-138-0506-3892

Abstract. An internal over-voltage identification method based on atomic decomposition(AD) and Support Vector Machine(SVM) is proposed. Firstly, the three-phase voltages of bus are decomposed by AD algorithm to get the optimal atoms. According to the atom frequency, the optimal atoms are constructed as characteristic atom matrix, which can reflect the time-frequency characteristics of the over-voltage signal. Features are obtained from the characteristic atom matrix by singular value decomposition. Finally, the features are input into the SVM to achieve the identification of seven types of internal over-voltages. The proposed method is verified on the simulation platform. The results show that the proposed method has high accuracy.

1. Introduction

Operational experience and research have shown that insulation of equipment and lines is easily broken down by over-voltage, it will affect the normal operation of the power system, and ultimately bring huge economic losses and personal hazards. Accurately identifying the type of over-voltage is helpful for engineers and technicians to find out the cause of the accident in a timely manner, which provides a basis of the improvement in the method of over-voltage suppression and the insulation coordination of the distribution network. It is of great significance for improving the self-healing capacity of the distribution network and ensuring the safe and stable operation of the power grid.

Over-voltage identification contains two main parts: selecting appropriate mathematical method to extract features and choosing appropriate classifier. The traditional Fourier transform lacks the ability to analyse the time domain of the signal and is unable to process non-stationary signal; Wavelet transform[1], which is a local transform of the frequency and time of the signal, can extract valid information on the signal, yet the selection of the wavelet base has great influence on the transform result; S-transform[2] has better time-frequency characteristics, but its performance is susceptible to sampling frequency, band and rate; Hilbert-Huang transform (HHT) can describe the different frequency components of the signal, nevertheless there exists the end effect and mode mixing in the process of EMD[3], which will bring great errors to the analysis results; atomic decomposition(AD) algorithm expresses the signal based on adaptive selection of basis function, it has good time-frequency characteristics, whose result is intuitive and unsusceptible to noise.

At present, the classifier based on machine learning, such as Support Vector Machine(SVM), Extreme Learning Machine(ELM)[4], BP neural network[5], is generally adopted. Among them, SVM has advantages in solving the problems of small samples, nonlinearities, local minimum points and so on. Therefore, SVM is widely used in the field of fault diagnosis[6].

In this paper, an internal over-voltage identification method based on AD algorithm and SVM is proposed. The AD algorithm and singular value decomposition(SVD) are combined to form feature quantities. And these feature quantities are input to SVM classifier to complete the identification of seven typical internal over-voltages in distribution networks.

2. Atomic decomposition

The AD algorithm is derived from the idea proposed by Mallat and Zhang that the signals are decomposed on the over-complete atom library, in which the redundant time-frequency atom replaces the traditional orthogonal basis function. It overcomes the limitation of the signal expressed by a fixed basis function.

Matching pursuits (MP) is an iterative algorithm for signal adaptive atomic decomposition. After *m*-step iterations by MP algorithm, the signal can be expressed as

$$x = \sum_{n=1}^{m} \left\langle R^n x, g_{\gamma_n} \right\rangle g_{\gamma_n} + R^{m+1} x \tag{1}$$

Where g_{γ_n} is the optimal atoms obtained at the *n*th iteration; $R^n x$ is the current signal; $\left\langle R^n x, g_{\gamma_n} \right\rangle$ is denoted as inner product operation whose absolute value is defined as fitness G; $R^{m+1}x$ is the residual signal.

The MP algorithm is a greedy algorithm. In order to reduce the computational complexity, Imperial Colony Competitive Algorithm (ICA) is used to optimize the MP algorithm. The principle of the ICA algorithm can be referred to [7].

Here, the damped sinusoid atom library[8] is selected as over-complete library, the absolute value of the product of the atom and the signal or signal residual is used as the fitness function, and the parameters of damped sinusoid atom are used as the optimization variables.

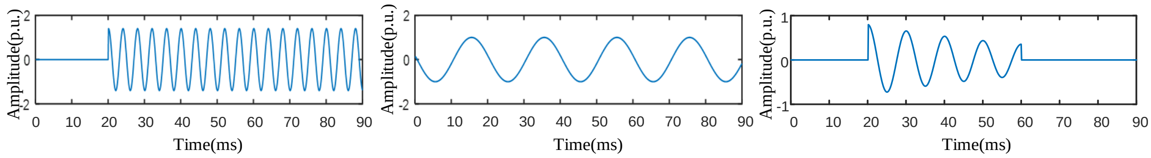
3. Feature extraction of over-voltage signal

3.1. Atomic decomposition of over-voltage signal

Constructing a test signal to verify the validity of AD algorithm. The fundamental frequency component, the 2nd decay harmonic component and the 5th harmonic component of different time distribution are superimposed as test signal, the expression is

$$x(t) = \begin{cases} 10\cos(100\pi t)\,, & 0 \le t \le 90\text{ms} \\ 12\cos(200\pi t)e^{-20t}\,, & 20 \le t \le 40\text{ms} \\ 14\cos(500\pi t)\,, & 20 \le t \le 90\text{ms} \end{cases} \tag{2}$$

At the same time, a Gaussian white noise(SNR=30dB) is added to the signal to prove anti-noise of the algorithm. Decomposing $x(t)$ three times to get three optimal atoms as shown in Figure 1, where U_1 is the first optimal atom, U_2 is the second optimal atom, U_3 is the third optimal atom. The parameters of atom are shown in Table 1, it can be seen that the parameters of atom are basically consistent with the original signal, indicating that the algorithm can accurately extract the internal characteristics of the signal.



(a)U_1 (a)U_2 (a)U_3

Figure 1. The first three optimal atoms extracted by atomic decomposition

Table 1. The parameters of the optimal atoms

Num.	f/Hz	ϕ /rad	A	t_s /ms	t_e /ms	ρ	G
1	250.02	0	14.03	19.98	90.00	0	26.24
2	50.02	1.39	10.01	0	90.00	0	21.23
3	100.04	0	11.90	0	90.00	20.28	7.95

3.2. Construction of characteristic atom matrix

For the purpose of better depicting the local feature of the over-voltage waveform in the time and frequency domain, the optimal atoms are reconstructed as certain frequency interval and divided into frequency bands.

Define the amplitude of the kth point in the band m as

$$E_m(k) = \left|\sum_{k=1}^{H} U_f(k)\right| \tag{3}$$

Where $U_f(k)$ is the amplitude of the atom belonging to the band m at the kth point; f is the atom frequency; H is the number of data points of the atom.

Each frequency band is N-divided in time period so as to make full use of the characteristics of the time domain. Define the amplitude of the atom in the frequency band m during the period n as

$$E_m^n = \sum_{k=(n-1)\frac{H}{N}+1}^{n\frac{H}{N}} E_m(k) \tag{4}$$

By calculating the amplitudes of all frequency bands and time periods, we can obtain the characteristic atom matrix $\boldsymbol{E}$:

$$\boldsymbol{E} = \begin{bmatrix} E_1^1 & E_1^2 & \cdots & E_1^n \\ E_2^1 & E_2^2 & \cdots & E_2^n \\ \cdots & \cdots & \cdots & \cdots \\ E_m^1 & E_m^2 & \cdots & E_m^n \end{bmatrix} \tag{5}$$

Based on frequency characteristics of over-voltage described in [2], it is found that the atom frequency is largely concentrated on 0~4 kHz, so the range is taken as the total bandwidth. The number of frequency band m is set to 15, the frequency band range is [0 10 20 30 40 80 200 300 400 500 600 700 1000 1500 2000 4000], the 1th frequency band is 0~10Hz, the 2th frequency band is 10~20Hz, and so on.

On the basis of the above procedure, the test signal of equation (2) is converted into a characteristic atom matrix.The characteristic atom matrix of Figure 2(a) as shown in Figure 2(b), in which the energy is concentrated on the 5th, 6th and 7th frequency bands, representing the components of the original signals of 50Hz, 100Hz, and 250Hz. The depth of color represents the magnitude of the energy and also corresponds the amplitude of the original signal, thus the characteristic atom matrix can reflect the time-frequency characteristics of the waveform of Figure 2(a) very well. The time-frequency matrix constructed by HHT[9] is shown in Figure2(c).Compared with Figure2(b), there are the end effect and mode mixing which is occur at the frequency mutation, and the noise also appears. Therefore, the characteristic atom matrix can reflect the time-frequency characteristics of the original waveform better than the time-frequency matrix constructed by the HHT.

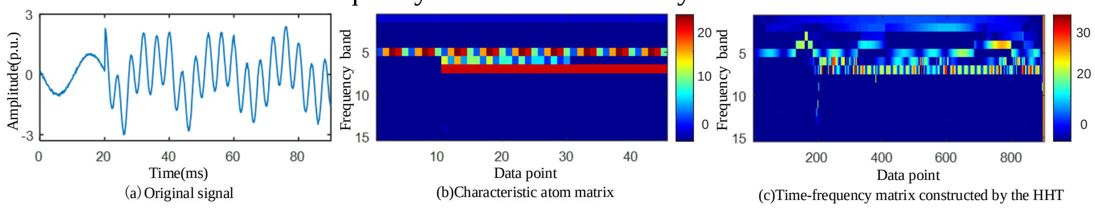


Figure 2. Analog signal waveform and its characteristic atom matrix

3.3. The SVD of characteristic atom matrix

From the previous section, the characteristic atom matrix reflect the characteristics of the waveform. However, the dimension of the characteristic atom matrix is too high to directly be used as the feature quantities. Therefore, according to the method of [10], we use SVD and calculate the statistical characteristics of singular values to reduce the dimension of the characteristic atom matrix.

SVD is equivalent to decomposing the matrix $\boldsymbol{A}$ with rank r into the weighted sum of r $m \times n$-order matrices $\boldsymbol{A}_h$ with rank 1. The weight λ_h is the singular value of the matrix, as shown in formula (6).

$$\boldsymbol{A} = \sum_{h=1}^{r} \lambda_h \boldsymbol{A}_h \tag{6}$$

In the light of singular value cumulative contribution rate[11], the first four singular values of each phase are taken as the main singular values. Due to the certain irregularities and dispersions of the same over-voltage waveform, the main singular spectrum maximum singular value λ_1, mean λ_{ave}, pulse factor I, entropy S_{sum}, and standard deviation S_{td}, which are based on the singular value of the characteristic atom matrix in each phase, are calculated as formula(7)-(10).

$$\lambda_{\text{ave}} = \frac{1}{4}\sum_{i=1}^{4} \lambda_i \tag{7}$$

$$I = \lambda_1 / \lambda_{ave} \tag{8}$$

$$S_{\text{sum}} = \sum_{i=1}^{4}(-p_i \lg p_i),\ p_i = \lambda_i / \sum_{i=1}^{4} \lambda_i \tag{9}$$

$$S_{\text{td}} = \sqrt{\frac{1}{4}\sum_{i=1}^{4}(\lambda_i - \lambda_{\text{ave}})} \tag{10}$$

4. Experiment And Simulation

4.1. Analysis of over-voltage data

The seven kinds of internal over-voltages commonly found in the distribution network include the over-voltage of single-phase to ground(T1), fundamental ferroresonance(T2), sub-frequency ferroresonance(T3), high-frequency ferroresonance(T4), intermittent arcing grounding(T5), capacitor switching (T6) and line switching(T7). The characteristic atom matrix can be constructed in according with the above construction method. The first 1 and last 3.5 periodical waveforms of the fault three voltages in the bus are displayed in Figure 3(a).

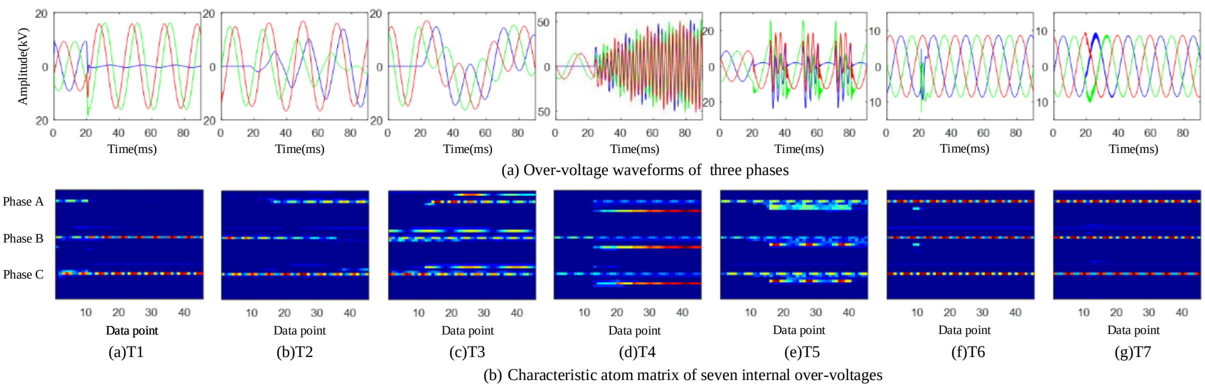


Figure 3. Three-phase waveforms and its characteristic atom matrix of seven internal over-voltages

Because the statistical characteristics of singular values obtained from (7)-(10) are hard to observe the difference between different over-voltages intuitively. Thence, we visualize the characteristic atom matrix to illustrate.

As can be observed in Figure 3(b) that the energy of T1 is mainly concentrated at 40~80Hz (5th frequency band), and the energy of the fault phase drops to zero after the faults occur. The difference of ferroresonance is that the frequency is distributed over different frequency bands. The energy of T2(Figure 3(b)) concentrates at 40~80Hz(5th frequency band); in addition to the 5th frequency band, T3(Figure 3(c)) also has energy distributed at 20~30Hz(4th frequency band); the energy of T4(Figure 3(d)) is distributed over 80~700Hz(6-11th frequency bands). The energy distribution of T5 is complex and distributed over 40~700Hz(5-11th frequency bands). The energy of both T6(Figure 3(f)) and T7(Figure 3(g)) are largely distributed over 40~80Hz(5th frequency band), but in the vicinity of data point 10, the over-voltage of T6, with respect to the T7, has more energy and shorter duration in the high frequency bands.

From the above analysis, the characteristics and differences of different over-voltages can be reflected by characteristic atom matrix, whose dimension is reduced by SVD. By splicing the five singular value of the three phases, the 1×15 feature quantities can be formed and sent to SVM to identify the type of over-voltage .

4.2. Identification test

A typical model of 10kV distribution system, which is based on actual primary wiring and basic data of a substation in Fujian Province, is established via ATP-EMTP, as shown in Figure 4. Among them, F10~F92 are the fault points; electromagnetic voltage transformer and capacitor bank are hung on the bus. The sampling rate is set to 10 kHz.

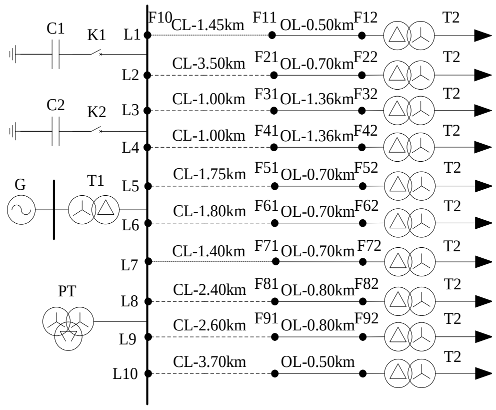


Figure 4. Simulation model of distribution network

To consider the impact of many factors on the identification accuracy, we change the parameters of fault location, fault resistance and fault phase in the model of Figure 4 to yield 2742 over-voltage waveform samples, half of which are used as training sets and half as test sets.

The identification results of the HHT combined with SVM and AD combined with ELM are given respectively to highlight the advantages of the proposed method. As shown in Table 2 that the accuracy of AD+SVM algorithm can reach 99.20%, which is higher than HHT+SVM. The main reason is that the mode mixing is more serious, resulting in that fundamental frequency ferroresonance is easily misidentified as sub-frequency ferroresonance, and line switching is easily misidentified as capacitor switching. Meanwhile, compared with the AD+SVM and AD+ELM, the SVM relative to ELM identifying seven types is multi-level identification, each layer is specific to the type, which is more targeted and has higher identification accuracy.

Table 2. Identification results of the test sets

Type	Number of test sets	Accuracy of test sets		
		AD+ELM	HHT+SVM	AD+SVM
T1	432	96.06%	100%	100%
T2	165	90.91%	92.73%	99.13%
T3	165	95.76%	87.88%	97.58

T4	165	96.97%	100%	99.39
T5	105	98.10%	95.24%	100%
T6	167	89.82%	97.60%	99.40%
T7	172	89.53%	86.63%	97.09%
Total	1371	93.58 %	95.33%	**99.20%**

In order to simulate the noise interference from the electromagnetic environment, 30dB of Gaussian white noise is added to the test sets to verify anti-noise ability. The identification result is shown in Table 3. The accuracy is 97.22%. This is because the AD algorithm itself has a good filtering effect, so the influence of noise on the identification is very weak, indicating that the proposed method has good anti-noise performance.

Table 3. Identification results under noise interference

Type	T1	T2, T3, T4	T5	T6	T7	Total
Number of test sets	432	495	105	167	172	1371
Accuracy of test sets	99.07%	99.13%	97.14%	86.83%	95.93%	97.22%

5. Conclusion

An internal over-voltage identification method based on AD and SVM is suggested in this paper. The results show that the AD algorithm can effectively analyze the over-voltage signal and accurately extract the internal characteristics of the signal in the presence of noise interference. In addition, the characteristic atom matrix constructed by the atomic decomposition algorithm can restore the time-frequency characteristics of the original signal more accurately than the time-frequency matrix obtained by the HHT algorithm. The proposed method has high accuracy and strong adaptability, which lays a foundation for the next step of research on over-voltage suppression in the distribution network.

Reference

[1] Mokryani G, Haghifam MR, Esmaeilpoor J 2013 Identification of ferroresonance based on wavelet transform and artificial neural networks. *European Transactions on Electrical Power*, 19 474-486

[2] Wang J, Yang Q, Sima W, Yuan T, Zahn M 2011 A Smart Online Over-Voltage Monitoring and Identification System *Energies* 4 599-615

[3] Wang J, Yang Q, Sima W 2012 Classification and Identification of Over-voltage Based on HHT and SVM *High Voltage Engineering* 38 2068-2075

[4] Luo M, Li C, Zhang X 2016 Compound feature selection and parameter optimization of ELM for fault diagnosis of rolling element bearings *Isa Transactions* 65 556-566

[5] Zhang YY, Zhao JJ, Liang W 2014 The Fault Diagnosis of Electric Power Metering System Based on Momentum BP Neural Network *Applied Mechanics & Materials* 668-728

[6] Zhou S, Qian S, Chang W, Xiao Y, Cheng Y 2018 A Novel Bearing Multi-Fault Diagnosis Approach Based on Weighted Permutation Entropy and an Improved SVM Ensemble Classifier *Sensors* 18 6

[7] Zhu C, Jing C, Zhang C 2015 Applying Modified Colonial Competitive Algorithm to Solve Minimal Hitting Set Problems *China Mechanical Engineering* 26 917-923

[8] Lovisolo L, Silva EABD, Rodrigues MAM, Diniz PSR. 2005 Efficient coherent adaptive representations of monitored electric signals in power systems using damped sinusoids *IEEE Transactions on Signal Processing* 53 3831-3846

[9] Guo MF, Yang NC 2017 Features-clustering-based earth fault detection using singular-value decomposition and fuzzy c-means in resonant grounding distribution systems *International Journal of Electrical Power & Energy Systems* 93 97-108

[10] Yang Y, Li L, Du L 2012 Lightning and Switching Overvoltage Identification Based on Singular Value Decomposition of Time-Frequency Matrix and Multi-Level Support Vector Machine *Power System Technology* 36 31-37
[11] Guo MF, You LX, Hong C 2017 Identification Method of Distribution Network Faults Based on Singular Value of LCD-Hilbert Spectrums and Multilevel SVM *High Voltage Engineering* 43 1239-1247