Statistical Feature Extraction for Fault Locations in Nonintrusive Fault Detection of Low Voltage Distribution Systems

Authors:

Hsueh-Hsien Chang, Nguyen Viet Linh

Abstract:

This paper proposes statistical feature extraction methods combined with artificial intelligence (AI) approaches for fault locations in non-intrusive single-line-to-ground fault (SLGF) detection of low voltage distribution systems. The input features of the AI algorithms are extracted using statistical moment transformation for reducing the dimensions of the power signature inputs measured by using non-intrusive fault monitoring (NIFM) techniques. The data required to develop the network are generated by simulating SLGF using the Electromagnetic Transient Program (EMTP) in a test system. To enhance the identification accuracy, these features after normalization are given to AI algorithms for presenting and evaluating in this paper. Different AI techniques are then utilized to compare which identification algorithms are suitable to diagnose the SLGF for various power signatures in a NIFM system. The simulation results show that the proposed method is effective and can identify the fault locations by using non-intrusive monitoring techniques for low voltage distribution systems.

Record Type: Published Article

Statistical Feature Extraction for Fault Locations in Nonintrusive Fault Detection of Low Voltage Distribution Systems

Hsueh-Hsien Chang [1],* and Nguyen Viet Linh [2]

[1] Department of Electronic Engineering, JinWen University of Science and Technology,
    New Taipei 23154, Taiwan
[2] Department of Electrical, Electronic, Computers and Systems Engineering, University of Oviedo,
    Oviedo 33003, Spain; nguyenvietlinh.eee@gmail.com
*   Correspondence: sschang@just.edu.tw; Tel.: +886-2-8212-2000 (ext. 6813)

Abstract: This paper proposes statistical feature extraction methods combined with artificial intelligence (AI) approaches for fault locations in non-intrusive single-line-to-ground fault (SLGF) detection of low voltage distribution systems. The input features of the AI algorithms are extracted using statistical moment transformation for reducing the dimensions of the power signature inputs measured by using non-intrusive fault monitoring (NIFM) techniques. The data required to develop the network are generated by simulating SLGF using the Electromagnetic Transient Program (EMTP) in a test system. To enhance the identification accuracy, these features after normalization are given to AI algorithms for presenting and evaluating in this paper. Different AI techniques are then utilized to compare which identification algorithms are suitable to diagnose the SLGF for various power signatures in a NIFM system. The simulation results show that the proposed method is effective and can identify the fault locations by using non-intrusive monitoring techniques for low voltage distribution systems.

Keywords: artificial intelligence (AI); wavelet transform; non-intrusive fault monitoring (NIFM); distribution systems; feature extraction

1. Introduction

1.1. Background

Supervisory control and data acquisition (SCADA) has been the traditional load monitoring method for several years. In this method, sensors are installed at each load point to detect the actions of switches or breakers. Then, sensors deliver messages to load recorders and the data center. However, as power systems have become more complex, this approach incurs in significant time delays and cost for installation and maintenance. Also, the increase in the number of sensors will complicate the system and reduce the reliability.

To alleviate these problems, the initial conception of a non-intrusive load monitoring (NILM) system was mentioned by Hart [1]. Compared with the SCADA-based system, NILM only places a set of voltage and current sensors in the entrance of utility (electrical service entry, ESE) instead of the load points. At this ESE, load monitoring, communication and control schedules are carried out. This future load-monitoring system tends to minimize the number of instruments and thus, reduce the hardware and costs [2]. Therefore, NILM schemes have long-term potential for the coming generation of smart grids and power energy management systems.

In general, a NILM platform includes three stages: data acquisition, feature analysis, and pattern recognition. Figure 1 shows a simple schematic diagram of a NILM system. Firstly, the power signatures are acquired from the power inlets. The features of voltages and currents are then analyzed based on suitable feature extraction or selection techniques. The purpose of feature analysis is to reduce the dimensionality of the original raw waveforms, and obtain the subsets of principal variables. Machine learning algorithms finally compare those captured variables with various patterns corresponding to each appliance in a database.
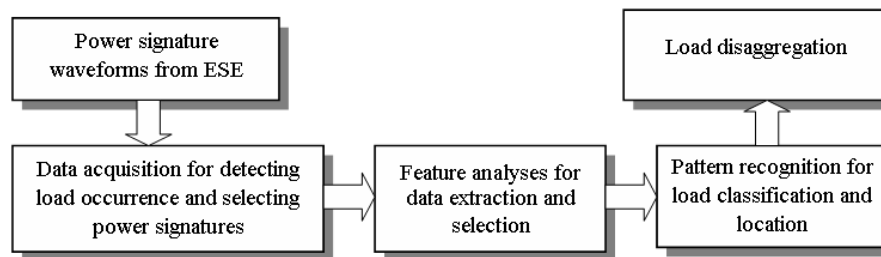

Figure 1. Schematic diagram of a non-intrusive load monitoring (NILM) system.

For the past two decades, a number of studies have been conducted for load identification in NILM systems. Laughman et al. [3] as well as Zeifman and Roth [4] reviewed the background and content of the NILM literature. In term of feature analysis, the commonly used signatures can be grouped into steady-state features and transient features. The steady-state features involve active power (*P*), reactive power (*Q*), harmonics, voltage-current (*V-I*) trajectory and admittance waveform. In Hart [1], the loads were identified based on the real and reactive powers. Cole and Albicki [5] also evaluated data extraction for NILM using power-based methods. Nevertheless, this has the limitation that various loads have the same power consumption. Also, the power-based methods may not recognize the appliances where the consumed powers change non-discretely [6]. Roos et al. [7] analyzed the details of load under steady-state operation for identification. This analysis, however, puts computers in a computationally complicated situation for accurate power signature data. Akbar and Khan [8] deployed the current harmonics as the steady-state features. Nevertheless, several highly resistive loads may not be detected by this approach due to the low level of the current harmonics. The applications of *V-I* trajectory in NILM are discussed in [9,10]. In this scheme, the instantaneous voltage and current are respectively measured and normalized based on their peaks. Then, the *V-I* plot traces different shapes corresponding to different appliances. This *V-I* trajectory-based extraction requires the intensive computation due to the two-dimensional matrix of *V-I* and complicated wave-shapes. The main drawback of using steady-state features is the insufficient information of the load behavior. Different appliances may have similar patterns in the steady-state operations, which may lead to the mistaken identification.

Various research papers studying NILM using transient features have been published in order to overcome the shortcomings of the steady-state feature [11–15]. Starting current or switching transient waveform are typical transient features. Chang et al. [11] proposed a turn-on transient energy approach combined with the discrete wavelet transform (DWT). Since the turn-on transient energy can represent the unique characteristics of appliances, this approach is reliable for load identification. Similarly, Gillis et al. [15] also applied the DWT principle and re-constructed the wavelet design for extracting the transient features. In [13], the energizing and de-energizing transient feature are adaptively adjusted by an artificial immune algorithm (AIA) and Fisher criterion. The energizing events are also processed by the S-transform to obtain the feature vectors in the complex domain [14]. Generally, the use of transient features requires the implementation of signal processing such as DWT or S-transform at the high sampling rate to capture the transient effects. In electrical transient analyses, these wavelet multi-resolution analysis (WMRA) techniques are useful to monitor power system small signal oscillations, power quality (PQ), and electric power disturbance [16–19]. For example,

the authors [19] employed harmonic voltages and wavelet coefficients as PQ features for placement of power quality measurement facilities to identify PQ problems. The main disadvantage of using transient features with those transforms is that the sampling frequency should be higher than 1 kHz to extract the distinction [20]. This is somewhat expensive in terms of the hardware cost for load-monitoring purposes.

Apart from the steady state and transient behaviors, several hybrid techniques of feature extraction are also developed. Guedes et al. [21] used the second and fourth-order cumulants of high-order statistics (HOS) to extract the information from the current waveform. Genetic algorithms (GAs) are then applied for feature selection and dimension reduction. Those HOS features can retrieve the important information of electrical signals, so they possess the capability of discrimination in the feature space. Bouhouras et al. [22] described the concept of spectra distribution analysis for individual load activation. In that work, the raw signals are transformed to the corresponding fast Fourier Transform (FFT) spectra. Afterwards, the complete spectrum is divided to a number of non-overlapping frequency bands. The spectral band energy and entropies are finally calculated for feature extraction. Both [21] and [22] reported the high performance in load disaggregation. Nevertheless, those outcomes were achieved at high sampling frequency, 15.4 kHz and 10.24 kHz, respectively.

In the pattern recognition stage, machine learning tools play a crucial role in classification. Several papers, e.g., Roos et al. [7] and Srinivasan et al. [23], have mentioned the use of artificial neural networks (ANNs) to facilitate the load recognition and harmonic sources for NILM systems. In [23], the authors used ANN to deal with harmonic issues, but the results do not involve the various operational modes of loads under different voltage sources. In NILM systems, most studies mainly focus on ANN with back-propagation (BP-ANN) for load recognition algorithms [2,6,7,11,13,21,23,24]. With simple implementation, *k*-Nearest Neighbors (*k*-NN) is also a commonly used recognizer. Saitoh et al. [24] implemented *k*-NN and support vector machine (SVM) to recognize household appliances. In that work, 1-NN was the most effective classifier compared with SVM. Tsai et al. [13] also drew a comparison between BP-ANN and *k*-NN. Both methods gave almost the same performance, and *k*-NN is preferable for its simplicity.

A hybrid recognizer that combines a supervised Self-Organizing Feature Map with Bayesian classification is also described in [25]. This method, however, encounters the difficulty of multi-state transitions in home appliances. In [26], Multiple conditional factorial hidden Markov model (MCFHMM) algorithm is adopted for classification. This innovation is a modified version of the hidden Markov model for unsupervised learning. Although the identification accuracy of the unknown appliances exceeds 80%, MCFHMM may not correctly predict unknown linear loads. In NILM, it requires further investigation on unsupervised learning to detect the new appliances without prior information.

1.2. Problems and Contributions

In power system protection analyses, power system faults are one of the most important research topics. These faults which are mainly due to short-circuit phenomena, can drastically affect the operations of power systems. Faults in power systems are divided into three-phase balanced and un-balanced faults. Different types of un-balanced faults are single-line-to-ground fault (SLGF), line-to-line fault (LLF), and double-line-to-ground fault (DLGF). Among the faults of power systems, SLGF is the most frequent type. Due to the serious consequences of fault damage in power systems, many researchers have devoted their attention to this research topic. Chunju et al. [27] employed a wavelet fuzzy neural network (WFNN) to extract the fault characteristics from the SLGF signals in an industrial distribution power system. However, this method does not contain the high frequency information of pre-fault currents and only considers SLGF on phase A of Line 1. In fact, the equivalent capacitance and mutual inductance will change when the fault location changes, the charging and discharging currents will be also changed. Hence the high frequency currents, related with the distance from the relay point to the fault location, are different for different fault locations. Ekici et al. [28]

proposed a SLGF location method simulated in a single 380kV and 360 km long power transmission line. In that paper, the wavelet energy and entropy criterion of the wavelet packet transform (WPT) coefficients are deployed for every faulty current and voltage signal to extract features. Those features are later used for training and testing of ANNs. However, the technique will be unsuitable for a more complicated power system. Reddy et al. [29] used the wavelet multi-resolution analysis (WMRA) technique to diagnose the faulty conditions. An adaptive neuro-fuzzy inference system (ANFIS) and ANN in conjunction with GPS are applied to locate the faults when faults occur randomly during the real-time smart grid operation of transmission lines. This paper focuses on the amplitudes of the second- and third-order harmonics generated during fault current occurrence to track the fault location. Among different coefficients pertaining to different decomposition levels, only the summation of the fifth-level detailed coefficients (d5) is considered for the sampling rate of 6 kHz. Borghetti et al. [30] built specific mother wavelets inferred from the recorded fault-originated transient waveforms to improve the wavelet analysis for distribution networks. In that paper, the network topology and the traveling wave speeds of the various propagation modes are assumed to be known. However, as concluded in [30], the orthogonalization of transient-inferred mother wavelet is expected to improve the algorithm accuracy by means of proper integration of time-domain fault location approaches. Bezerra Costa [31] has presented a wavelet-based methodology for real-time detection of fault-induced transients in transmission lines, where the wavelet coefficient energy takes into account the border effects of the sliding windows. As a result, the performance of the proposed energy analysis is not affected by the choice of the mother wavelet and presents no time delay in real-time fault detection.

In the trend of energy disaggregation, fault identification, localization and classification have become the new challenge in non-intrusive monitoring system. Even with only one set of voltage and current sensors in ESE, the system must still ensure the effectiveness of protection. Faults should be detected and cleared as fast as possible. As far as the authors are aware, fault identification using non-intrusive monitoring mechanisms in power systems is still in the primary development stage. To enhance this recognition accuracy for a more complicated low voltage distribution system by using non-intrusive fault monitoring (NIFM) techniques, a new fault identification approach in load buses and distribution lines for SLGF is the topic of this paper. Figure 2 illustrates the block diagram for the SLGF diagnoses in NIFM systems.
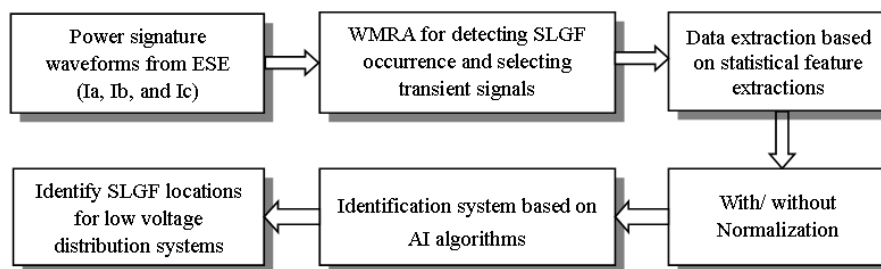

Figure 2. Block diagram of a non-intrusive power fault identification system. ESE: electrical service entry; SLGF: single-line-to-ground fault; and WMRA: wavelet multi-resolution analysis.

First, the WMRA technique is utilized to detect SLGF occurrence and select the effective transient signals which are the post-fault events. Second, the statistical moment transformation and normalization are employed for feature extraction. In general, the characteristics of any random data can be described by the statistical moments (SMs) [32,33]. These moments are widely deployed in signal processing, ex. blind decomposition [34], noise and signal detection [35,36]. In particular, the use of the statistical moments is popular in terms of condition monitoring and bearing defect detection [37–39]. Since these statistical moments can represent the uniqueness of the raw power signatures, the features extracted in this stage could be the good candidates to solve the problem of the fault disaggregation.

In the final step for comparisons, different artificial intelligence (AI) algorithms—BP-ANN, *k*-NN, and SVM—are used to diagnose the SLGF of distribution lines or load buses for various power signatures in a NIFM system. Among those classifiers, SVM is proved to be the most effective learning tool for these case studies of this work. By using the proposed methods (SK and SVM), the number of transient signal features can be reduced without losing its fidelity. The accuracy rate of the proposed methods is tested in a model system and simulated by using the Electromagnetic Transient Program (EMTP) software. These results show that the proposed methods can analyze the signals efficiently and effectively, thus enhancing the performance of the fault identification in the NIFM system.

The remainder of the paper is organized as follows: the data preparations including the data acquisition, experimental datasets and fault detection are described in Section 2. The feature analysis of transient fault currents using the statistical moments (SMs), Z-scores, Skewness and Kurtosis (SK), as well as SVMs for the fault location are addressed in Section 3. Based on the feature analysis techniques and the intelligent classifiers, a series of experiments covering the distribution-line and load-bus faults is analyzed in Section 4 to prove the feasibility of the proposed methods. This approach can also distinguish between the inrush current caused by the motor starting and overcurrent caused by the SLGF, which verifies its superiority. An experiment is discussed in Section 5.

2. Data Preparation

2.1. Data Acquisition

Figure 3 shows a typical NIFM system for the experimental case studies in this paper. Based on the statistical feature extraction, the identification algorithm in the NIFM system recognizes the distribution-line faults and load-bus faults of five 220-V loads measured on a three-phase 480-V common bus These loads include a 45-hp induction motor, a 55-hp induction motor, a 30-hp induction motor, a 20-hp induction motor driven by line frequency variable voltage drives, and a load bank supplied by a six-pulse thyristor rectifier for AC power.
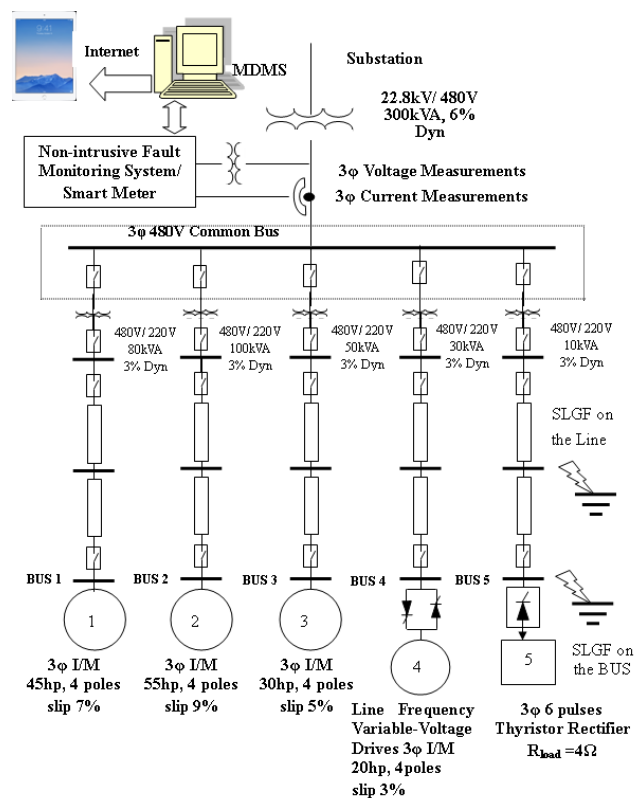

Figure 3. Power fault identification system for a NIFM system. MDMS: meter data management system.

In this scheme, the smart meter/NIFM manages and measures the operations and power demands of each load. Then, this information of each fault event is sent to the meter data management system (MDMS) via wireless sensor communication. Through the Internet or Web systems, the client terminal can access these load data and fault events.

2.2. Experimental Datasets

Experimental data sets were generated by simulating the data on the current waveforms at the ESE. Each final sample consists of ($T \times 60 \times 256$) data samples obtained from 256 sampling points for one cycle over a period of T in a 60 Hz power system. Each power signature example includes a voltage variation from -3% to $+3\%$ at 0.5% intervals. This yields 13 samples of each power signature for each bus and $N \times 13 \times 3$ raw data for all power signatures (I_a, I_b, and I_c) where N is the number of buses in a three-phase power system network. The full input data set comprises a ($N \times 13 \times 3$) $\times$ ($T \times 60 \times 256$) matrix as the experimental data sets. Basically, the fuse or circuit breaker must interrupt the fault quickly (generally less than 4 ms) in order to provide the maximum protection for equipment and personnel [40]. Underwriters Laboratories Inc. (UL, Northbrook, IL, USA) defines breaker current limitation as a breaker that interrupts and isolates a fault in less than 1/2 of an AC cycle. 1/2 a cycle is completed in 8.3 ms [41]. The data samples are calculated from the statistical feature extraction methods of the current signal within one cycle ahead from the fault inception. In the case studies of this paper, T and N are 1/60 seconds and 5, respectively. Figure 4 shows the current and voltage waveforms simulated on the ESE when the line 2 of the case study in Figure 3 has SLGF on the phase A.
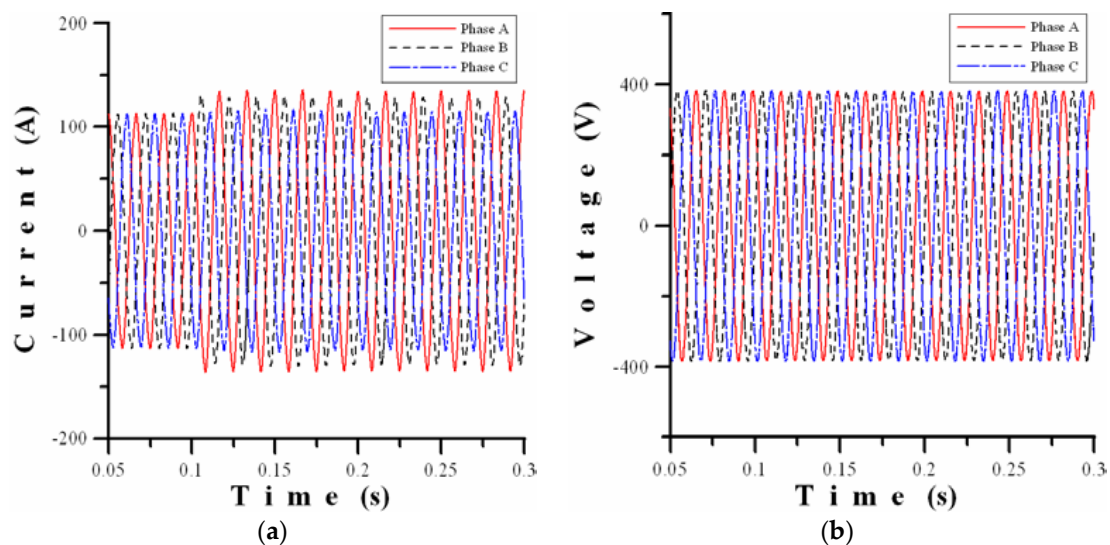

Figure 4. Current and voltage waveforms of SLGF on the phase A for the distribution line 2 of the case study. (**a**) Current waveforms; and (**b**) voltage waveforms.

2.3. Fault Detection

First, the wavelet transform (WT) technique of multi-resolution analyses is utilized to detect the time of SLGF occurrence in power systems. The post-fault signals are preprocessed before inputting the data into AI identification algorithms. The WT is a powerful tool for studying transients by characterizing band-pass filters [42,43]. Therefore, it is suitable to detect the occurrence of faults in power systems. WT decomposes power signatures into different scales of WT coefficients (WTCs). WT can be grouped into continuous WT (CWT), discrete WT (DWT), stationary WT (SWT) and wavelet packet analysis. The CWT is expressed as:

$$C_{a,b}(t) = \int\limits_{-\infty}^{\infty} \psi_{a,b}(t)x(t)dt \tag{1}$$

where $x(t)$ is the original signal and $\psi_{a,b}(t)$ is the daughter wavelet. The daughter wavelet is defined by Morlet and Grossman [43] as:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right); \ a \in R^+, \ b \in R \tag{2}$$

where $\psi(t)$ is the chosen mother wavelet, a is the scaling factor, and b is the shift factor.

In the discrete case, the scaling and shift factors are represented as follows:

$$\begin{aligned} a &= a_0^m \\ b &= nb_0 a_0^m \end{aligned} \tag{3}$$

Then, the DWT is obtained as:

$$D_{m,n} = a_0^{-\frac{m}{2}} \sum_k x[k]\psi[a_0^{-m}k - nb_0]; \ a_0 > 1; \ b_0 \neq 0 \tag{4}$$

where m is a scaling index; n is a sampling time point, for $n = 1, 2, ..., N$; and N is the number of sampling points; $m, n, b_0 \in Z$, Z is a set of integers, and k is an operating index; a_0 is selected as a spacing factor.

Set:

$$a_0 = 2; \ b_0 = 1 \tag{5}$$

Then the following results of the DWT for dyadic (octave) grid:

$$D_{m,n} = 2^{-\frac{m}{2}} \sum_k x[k]\psi[2^{-m}k - n] \tag{6}$$

The dyadic grid is perhaps the simplest and most efficient discretization for practical purposes and lends itself to the construction of an orthonormal wavelet basis.

The WMRA technique is an existing and well known method widely utilized in fault detection applications [27–29,31,44–46]. The time–frequency location feature of the WT makes it suitable to be used in fault detection method based on high and fundamental frequency components. This research uses the MRA technique to decompose the current signals into some approximation and detail levels of resolution. By analyzing certain levels of approximation and details, the stages of fault detection are processed. Based on the [45], the occurrence (T_{fault}) of SLGF is detected.

Figure 5 shows the fault-detection procedure, where IF is a counter that expresses the sample number under SLGF. SUM_d1 is the sum value of the detailed output (d1 coefficients) for a one-cycle period and is represented as an absolute value. Fault criterion (FC) is the signal magnitude threshold as the lower limit of SUM_d1, while Ns is the sample number that signifies the duration time for which a transient event has to persist continuously. When SUM_d1 is greater than or equal to FC, the value of IF is incremented and as soon as it attains the level Ns. This indicates an internal fault, and a trip signal is initiated. As shown in Figure 5, the sampling rate is 15.360 kHz; i.e., 256 samples/cycle at 60 Hz. The summated values associated the three phases are compared with a preset threshold level FC. The whole process is based on a moving window approach where the one-cycle window is moved continuously by one sample. In these cases of the paper, T_{fault} is 0.1 s. The optimal settings for FC and Ns are 0.085 and 128, respectively.
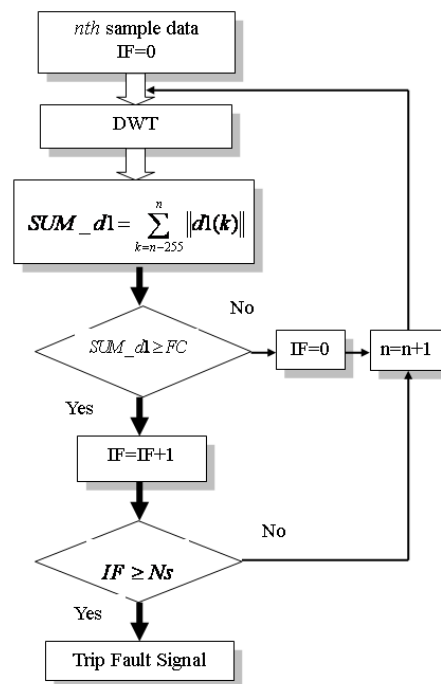
Figure 5. Block diagram of fault detection.

3. Proposed Methods

3.1. Statistical Moments

It is generally not possible to directly identify SLGF in power systems based upon these complicated raw waveforms. Thus, some transformations of the recorded SLGF time-series are implemented to extract the time-invariant features. These are statistical moments (SMs). Those transformations can also reduce the vulnerability and the variation of the datasets [38]. In the faults, the energy (mean square value) in the fault signals is expected to increase. The features of *n*th-order SMs are generated from the raw fault data, which is defined by Equation (7). The SMs used here are the four (first to fourth order) moments, stored in a matrix of size (65 × 4) for each power signature. Specially, the first and second order moments are the mean and the variance of the fault signals, respectively:

$$f^{(n)} = \frac{1}{N} \sum_{i=1}^{N} x_i^n \tag{7}$$

for *n* = 1, 2, 3, and 4.

3.2. Z-Scores

Normalization plays a crucial rule in preprocessing, as proven by Bishop [47]. This normalization works well for populations that are normally distributed. To normalize, the input variable of statistical moments f_{old} is converted into zero mean and unit variance, the new input variable of statistical moments f_{new} is called Z-scores or standard score normalization. The formula is illustrated by Equation (8):

$$f_{new}^{(n)} = \frac{f_{old}^{(n)} - \mu^{(n)}}{\sigma^{(n)}} \tag{8}$$

for *n* = 1, 2, 3, and 4, where μ and σ are the mean value and the standard deviation for the feature vector of the statistical moments f_{old}, respectively.

3.3. Skewness and Kurtosis

Another method of normalization is skewness and kurtosis (SK) or standardized central moments, using standard deviation as a measure scale. The *n*th-order standardized central moments are defined by Equation (9):

$$f_{new}^{(n)} = \frac{1}{N\sigma^{(n)}} \sum_{i=1}^{N} \left(f_{old}^{(n)} - \mu^{(n)} \right)^n \tag{9}$$

for *n* = 1, 2, 3, and 4.

Equation (9) shows that the 1st and 2nd standardized central moments are constants, with values of 0 and 1, respectively. This implies that these constants are not necessary for inputs of AI algorithms. In other words, skewness (normalized 3rd central moment) and kurtosis (normalized 4th central moment) are the sufficient features for fault identification. These moments are reduced into a matrix of size (65 × 2) for each power signature. To examine the effects of this normalization, the results of SLGF identification were compared between normalized and un-normalized statistical moments.

By examining the current distributions of energy from orders for SLGF on different five lines for different proposed methods, i.e., SMs, Z-scores, and SK, in Figures 6–8, respectively, some characteristics can be found. Firstly, the distributions of energy among Line 1, Line 2 and Line 3 are always in similar phases. There are only some minor differences in term of their magnitudes. This is understandable as Load 1, Load 2, and Load 3 are the same type of loads which are induction machines (IMs). Load 4, however, is the IM with the accompaniment of the line frequency variable-voltage drives; while Load 5 is the load bank supplied by a six-pulse thyristor rectifier for ac power. For this reason, Line 4 or Line 5 forms significantly different patterns compared with the remaining four lines, and can be easily discriminated.

To distinguish the faults which have the similar patterns, it is necessary to combine with the powerful tool of classification algorithms even though these statistical methods can effectively extract features and reduce the size of datasets.
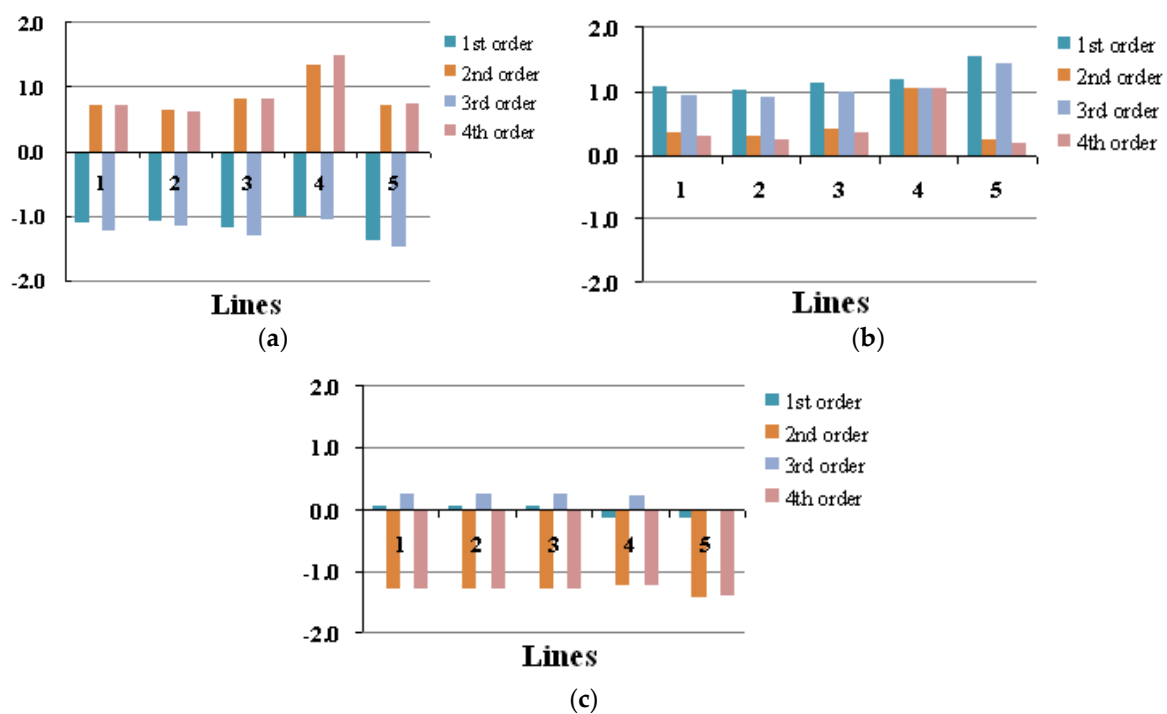

Figure 6. Current distributions of each order for the SLGF on different five lines by using the SMs. (**a**) Phase A; (**b**) Phase B; and (**c**) Phase C.
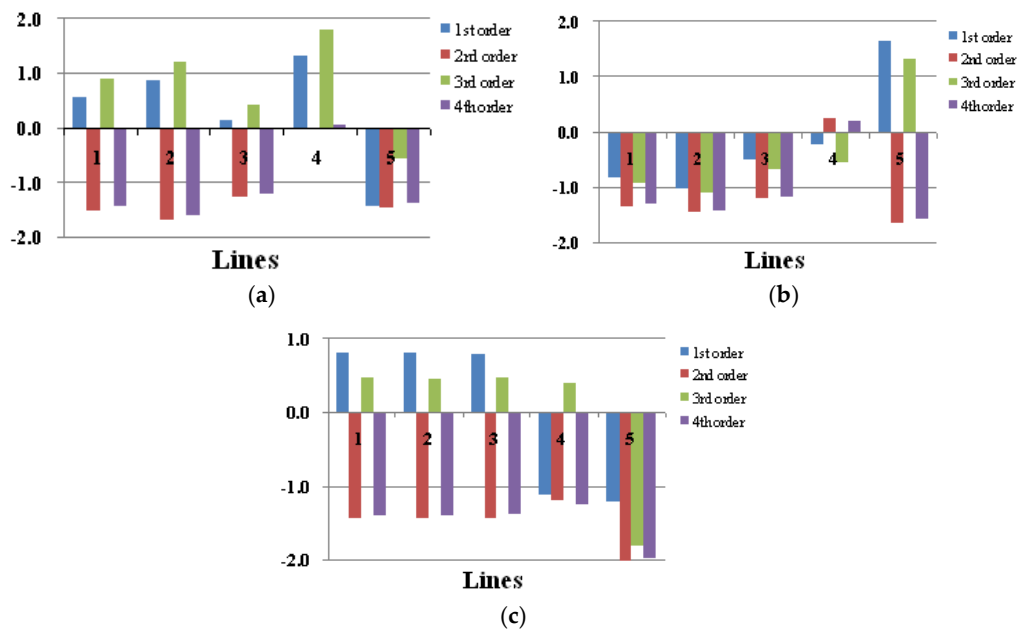
Figure 7. Current distributions of each order for the SLGF on different five lines by using the Z-scores. (**a**) Phase A; (**b**) Phase B; and (**c**) Phase C.
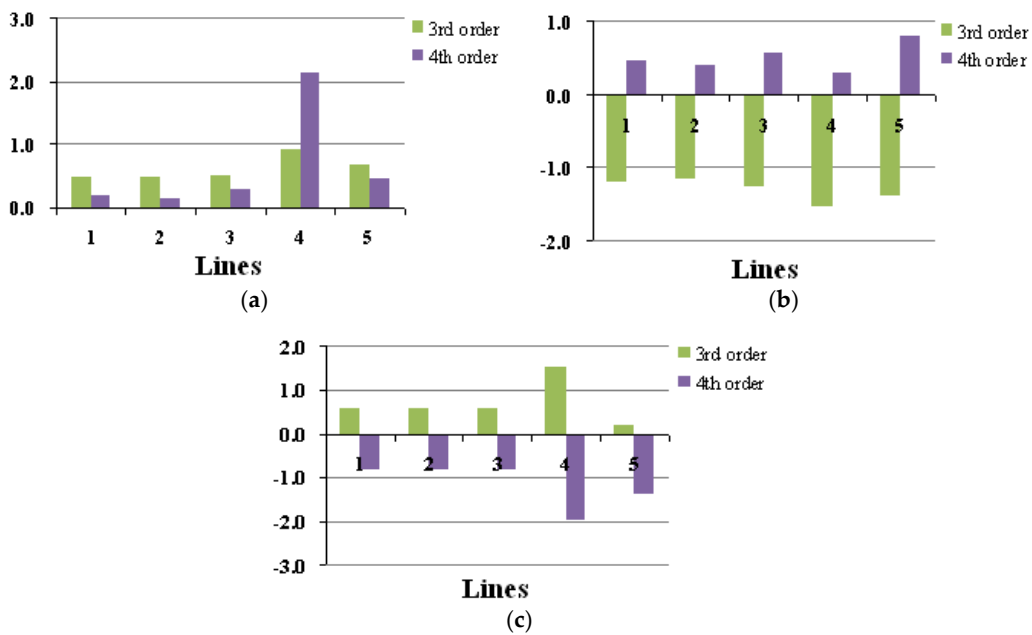
Figure 8. Current distributions of each order for the SLGF on different five lines by using the SK. (**a**) Phase A; (**b**) Phase B; and (**c**) Phase C.

3.4. SVMs

Support vector machines (SVMs) are used as intelligent tools for identifying faulty lines and buses that is finding the location with respect to the ESE. They were first introduced by Vapnik [48] on the basis of structural risk minimization principle. The aim of SVM is to create a line of hyper-plane among different datasets for identification. The hyper-plane is defined by subsets of the training data called support vectors. Support vectors can create complex boundaries and maximize the margin separation through the quadratic minimizations [49]. The SVM is able to solve both linear and

nonlinear identification problems. Based on the principle of structural risk minimization, SVM could avoid the local-minimum issue which is the fundamental challenge of BP-ANN.

The implementation of SVM multi-class classification is probably the one-against-all (one-vs-rest) method. It constructs k SVM models where k is the number of classes. The mth SVM is trained with all of the examples in the mth class with positive labels, and all other examples with negative labels. Thus, given l training data $(x_1, y_1), \ldots, (x_l, y_l)$, where $x_i \in R^n$, $i = 1, \ldots, l$, $y_i \in \{1, \ldots, k\}$ is the class of x_i. The hyper-plane of the mth SVM is optimized to solve the following problem [49]. The first term in the objective function represents the model complexity and the second term represents the model accuracy (i.e., classification error in the training data):

$$
\begin{aligned}
&\min_{w^m, b^m, \xi^m} \tfrac{1}{2} (w^m)^T \omega^m + C \sum_{i=1}^{l} \xi_i^m \\
&(w^m)^T \phi(x_i) + b^m \geq 1 - \xi_i^m, \text{ if } y_i = m, \\
&(w^m)^T \phi(x_i) + b^m \leq -1 + \xi_i^m, \text{ if } y_i \neq m, \\
&\xi_i^m \geq 0, \ i = 1, \ldots, l
\end{aligned}
\tag{10}
$$

where x_i is the ith training data which are mapped to a higher dimensional space by the function ϕ and C is the penalty parameter. y_i is the class label value which is either +1 or −1.

Then the largest value of the decision function is:

$$
\arg \max_{m=1,\ldots,k} ((w^m)^T \phi(x) + b^m)
\tag{11}
$$

The basic concept of SVM is to search for a balance between the regularization term $\tfrac{1}{2}(w^m)^T \omega^m$ and the training errors. Practically the Lagrangian dual problem of (10) is solved. Hence k l-variable quadratic programming problems are solved as follows:

$$
\begin{aligned}
&\max f(\alpha_1 \ldots \alpha_n) = \sum_i \alpha_i - \tfrac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j k(x_i x_j) \\
&\text{s.t. } o \leq \alpha_i \leq C \forall i, \sum_i \alpha_i y_i = 0
\end{aligned}
\tag{12}
$$

The merit of SVM which is the inner product in the feature space by using a kernel function $k(x_i, x_j)$ is that it ties to make the training data linear-separable in the high dimension feature space, thus achieve nonlinear-separable in the input space. Typical choices of kernel function include the follows:

(1)　Polynomial kernel: $k(x_i, x_j) = (\gamma(x_i^T \cdot x_j) + r)^d$;
(2)　Gaussian kernel: $k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$;
(3)　Sigmoid kernel: $k(x_i, x_j) = \tanh(\gamma(x_i^T \cdot x_j) + r)$.

where d is the exponent of polynomial function; $\gamma > 0$, $d > 0$ and r are kernel parameters.

It is obvious that Gaussian kernel function is the simplest model with the fewest hyper-parameters. Hence, Gaussian kernel function is reasonably used in this paper. To obtain a good performance, some parameters in SVM have to be chosen carefully [50]. These parameters include:

- The regularisation parameter C, which determines the trade-off between minimizing the training error and minimizing model complexity. The higher the value of C is, the narrower the separating margin of hyperplane is. In contrast, the hyperplane could easily misclassify more feature points if C is too low.
- Kernel function $k(x_i, x_j)$ and
- Parameter γ of the kernel function that implicitly defines the nonlinear mapping from input space to some high-dimensional feature space. The higher value of γ increases the complexity of the classification model, easily leading to overfitting. On the other hand, the lower γ may cause the under fitting error of classification.

In this paper, those parameters are tuned and configured during experiments. The Gaussian kernel function has been implemented with varying values for $C \in [0, 500]$, $\gamma \in [0, 1]$. The optimized values are 200 and 0.4 for C and γ, respectively.

3.5. Performance Evaluation

In the fault location of this paper, the full input dataset can respectively form a matrix with a size of (65×4) and (65×2) for each fault current. From the same input datasets, a matrix with the same size is additionally created for all three fault currents in the case of faulty phase detection. The data in those matrices are randomly divided into two equal sets. One set of the data is for training and the other is for testing. To confirm the inferential power of SVM, the input dataset is created from the raw signature waveforms by different proposed methods with SMs, Z-scores, or SK for comparisons.

The proposed protection scheme is depicted as in Figure 9. The one-vs-rest algorithm is applied for this binary classification. To illustrate, one class is set to be positive while the rests are negative, and the line/phase suffered from SLGF can be distinguished from the other lines/phases. Hence, this problem is divided into five sub-problems of binary classification. For example, two bits "0" and "1" respectively represent the negative and positive class. In the case study for the fault location, a set of five distribution lines represents a five-category attribute in the dataset. Those five classes indicate which line has a SLGF. A SLGF at Line 1, 2, 3, 4, or 5 can be respectively assigned as (1, 0, 0, 0, 0), (0, 1, 0, 0, 0), (0, 0, 1, 0, 0), (0, 0, 0, 1, 0), or (0, 0, 0, 0, 1). In the case study for faulty phase detection, phase A, B, and C can be set as (1, 0, 0), (0, 1, 0), and (0, 0, 1), respectively.
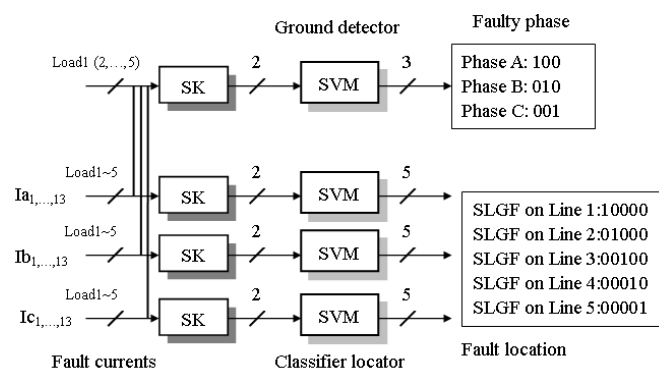
Figure 9. Protection scheme for proposed methods. SK: Skewness and Kurtosis; and SVM: support vector machine.

To evaluate the performance, the decision functions or classification rules should be applied. Firstly, the true positive (TP) and true negative (TN) are defined as the correctly classified positive cases and the correctly classified negative cases, respectively. The confusion matrix is shown in Table 1. Then, the classification accuracy is calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \qquad (13)$$

Table 1. A general confusion matrix.

Total Population		Predicted Condition	
		Prediction Positive (One Line)	Prediction Negative (Four Remaining Lines)
True Condition	Condition Positive	True Positive (*TP*)	False Negative (*FN*)
	Condition Negative	False Positive (*FP*)	True Negative (*TN*)

4. Experimental Results

The inputs of AI identification algorithms are three different power signatures, currents of three phases (I_a, I_b, I_c) pre-processed by the proposed extraction methods. Besides SVM, two other popular algorithms BP-ANN and *k*-NN (*k* = 1) are also implemented for comparison. Those identification algorithms are carried out by HeuristicLab (Ver. 3.3, Heuristic and Evolutionary Algorithms Laboratory (HEAL), Hagenberg, Austria) [51]. The programs were run to identify SLGF events for two case studies on a personal computer equipped with a 2.2-GHz Intel Core i3-2330M central processing unit. One is distribution-line faults; the other is load-bus faults. Each entry in the AI algorithms represents 100 different trials. The identification results are compared between the AI algorithms and power signatures for the proposed methods. A simulated NIFM system as shown in Figure 3 has five different loads on different buses. Those loads include a 45-hp induction motor, a 55-hp induction motor, a 30-hp induction motor, a 20-hp induction motor driven by variable-voltage drives, and a bank of loads supplied by a six-pulse thyristor rectifier for AC power.

4.1. Case Study 1, Distribution-Line Faults

In case study 1, the AI algorithms in the NIFM system identify SLGF on phase A at five different 220-V distribution lines from different time based upon the proposed methods, as shown in Figure 3. The fault SLGF occurs on a specific single distribution line of the five distribution lines, the other distribution lines operate normally. There are essentially five 220-V distribution lines in the model system of Figure 3 because the distance between the 480 V/220 V distribution transformer and the load bus is long.

Each line is 3 km long and there are two sections per line each section being 1.5 km in length. This allows the user to apply faults at the section junctions. The lines are modelled using a constant parameter line model in EMTP program. The line conductor is a 1 AWG with an 8.4 millimeter diameter and a dc resistance of 0.5426 Ω/km at 25 °C. The line parameters are calculated at 60 Hz with an earth resistivity of 100 Ω-m.

In the case study of distribution-line faults where each distribution line has SLGF on phase A, the results for different AI algorithms and three extraction methods including SMs, Z-scores, and SKs are shown from Tables 2–4. In general, SVM obtains the most impressive results compared with the others, regardless whether any extraction method is used. On the other hand, *k*-NN (*k* = 1) has the unsatisfactory outcome where all average values of the test accuracy are below 72.22% in the case of SMs and Z-scores. Also, BP-ANN which does not exceed 84.35% in the average test accuracy could not well perform in SMs and Z-scores. Furthermore, the Z-scores normalization (SK) significantly improves the capability of disaggregation in 1-NN and SVM, but BP-ANN does not gain any benefit of this normalization. The approach of using the SK has taken a further result in the recognition accuracy, shown in Tables 2 and 3. Most of results achieve visible improvement. Those achievements show that almost all cases obtain 100% in recognition accuracy. There is an exception in the case of BP-ANN.

In term of phases, the inputs from phase A obviously achieve the best outcome than those of phase B and C in the case of SMs and Z-scores when distribution-line faults occur on phase A. Another remark is that BP-ANN requires much more execution time compared with SVM and 1-NN.

Table 2. Results of SLGF on the line faults with the SVM. SMs: statistical moments.

Phase	SMs				Z-Scores				SK			
	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.
Recognition Accuracy in Training (%)	100	97.14	88.57	95.24	100	97.14	82.86	93.33	100	100	100	100
Recognition Accuracy in Test (%)	100	86.67	66.67	84.45	100	93.33	73.33	88.89	100	100	100	100
Time (s) 100 trials	0.2813	0.3125	0.2969	0.2969	0.2857	0.2592	0.5238	0.3562	0.2344	0.1343	0.0782	0.1490

Table 3. Results of SLGF on the line faults with the *k*-nearest neighbors (*k*-NN) (*k* = 1).

Phase	SMs				Z-Scores				SK			
	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.
Recognition Accuracy in Training (%)	20	14.29	0	11.43	94.29	71.43	40	68.57	100	100	100	100
Recognition Accuracy in Test (%)	11.43	10	6.67	9.37	93.33	76.67	46.67	72.22	100	100	100	100
Time (s) 100 trials	0.1111	0.1071	0.1070	0.1084	0.1172	0.0993	0.1135	0.1100	0.0625	0.0781	0.0781	0.0729

Table 4. Results of SLGF on the line faults with the BP-ANN.

Phase	SMs				Z-Scores				SK			
	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.
Recognition Accuracy in Training (%)	100	97.49	65.71	87.73	99.89	99.51	68.43	89.28	60	100	79.40	79.8
Recognition Accuracy in Test (%)	100	96.43	56.63	84.35	99.83	96.23	53.33	83.13	60	100	79.40	79.8
Time (s) 100 trials	5.6908	5.3066	0.1200	3.7058	4.8694	5.7919	3.1896	4.6170	7.0159	8.2034	3.5001	6.2398

The results of using SK and SVM for faulty phase detection in distribution-line faults are shown as Table 5. It is obvious that the proposed methods lead to the satisfactory classification performance 100% for training and test, regardless of faulty or normal phases.

Table 5. Confusion matrix for faulty phase detection using SK and SVM in case study 1.

Training	Actual Class		
	Phase A (Faulty)	Phase B (Normal)	Phase C (Normal)
Predicted Phase A (faulty)	33	0	0
Predicted Phase B (normal)	0	33	0
Predicted Phase C (normal)	0	0	33
Accuracy = (33 + 33 + 33)/(33 + 0 + 0 + 0 + 33 + 0 + 0 + 0 + 33) = 100%			
Test	**Phase A**	**Phase B**	**Phase C**
Predicted Phase A (faulty)	32	0	0
Predicted Phase B (normal)	0	32	0
Predicted Phase C (normal)	0	0	32
Accuracy = (32 + 32 + 32)/(32 + 0 + 0 + 0 + 32 + 0 + 0 + 0 + 32) = 100%			

4.2. Case Study 2, Load-Bus Faults

In case study 2, the AI algorithms in the NIFM system identify SLGF on phase A at five different 220-V load buses based upon the proposed methods, as shown in Figure 3. The fault SLGF occurs on a specific one load bus of the five load buses, while the other load buses are operating normally. There are essentially five different loads on different 220-V load buses from five feeders in the model system of Figure 3. The loads include a 45-hp induction motor, a 55-hp induction motor, a 30-hp induction motor, a 20-hp induction motor driven by variable-voltage drives, and a bank of loads supplied by a six-pulse thyristor rectifier for AC power.

Tables 6–8 list the results of SLGF location for different AI algorithms and three statistical extraction methods are obtained when SLGF occurs on phase A of each load bus. Obviously, SVM remains the best classifier compared with the others, regardless of whether any extraction method is used. In k-NN ($k = 1$), the average values of test accuracy are below 50% in the case of SMs and Z-scores. On average, BP-ANN does not surpass 78.78% in the test recognition accuracy when the inputs are from SMs and Z-scores. The main drawback of BP-ANN is that it consumes more computer resources than SVM and 1-NN do, as shown in the time execution. As the good features, SK maximizes the capability of disaggregation in all algorithms.

In term of phases for this case study, the inputs of phase A do not achieve better outcome than those of phase B and C in the case of BP-ANN and 1-NN when load-bus faults occur on phase A, regardless whether any extraction method is used.

Table 9 shows the performance of using SK and SVM applied for faulty phase detection in the case of load-bus faults, which again verifies the high accuracy 95.95% in training and 90.625% in test. This means that SKs can extract the distinction among phases even when SLGF occurs.

To summarize case studies 1 and 2, the currents monitored in phase A gives better distinctive features than those measured in phase B and C when SVM is used and SLGF occurs on phase A, regardless whether any extraction method is used. Among classifiers for these case studies, SVM gives the best results. In this proposal, SK does not only reduce the number of inputs for AI classifiers, but it also provides the distinctive features for phase and line disaggregation.

Table 6. Results of SLGF on the bus faults with the SVM.

	SMs				Z-Scores				SK			
Phase	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.
Recognition Accuracy in Training (%)	100	100	100	100	100	94.4	100	98.13	100	100	100	100
Recognition Accuracy in Test (%)	100	96.97	100	98.99	100	93.33	100	97.78	100	100	100	100
Time (s) 100 trials	0.3281	0.1563	0.2656	0.2500	0.2969	0.2500	0.3438	0.2969	0.0625	0.1406	0.1406	0.1146

Table 7. Results of SLGF on the bus faults with the k-NN ($k = 1$).

	SMs				Z-Scores				SK			
Phase	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.
Recognition Accuracy in Training (%)	3.33	11.73	5.71	6.92	65.71	68.57	42.86	59.05	100	100	100	100
Recognition Accuracy in Test (%)	5.71	16.67	3.33	8.57	43.33	50	56.67	50	100	100	100	100
Time (s) 100 trials	0.2031	0.2344	0.1094	0.1823	0.0937	0.2188	0.1719	0.1615	0.0625	0.0313	0.0625	0.0521

Table 8. Results of SLGF on the bus faults with the BP-ANN.

	SMs				Z-Scores				SK			
Phase	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.	I_a	I_b	I_c	Aver.
Recognition Accuracy in Training (%)	67.97	62.86	92.49	74.44	68.1	65.7	88.94	74.25	96.2	100	99.06	98.42
Recognition Accuracy in Test (%)	71.37	73.33	91.63	78.78	70.9	63.3	81.9	72.03	96.2	100	99.03	98.41
Time (s) 100 trials	8.7192	4.5627	5.0471	6.1097	10.609	4.7033	4.4845	6.5989	5.5159	6.9378	4.1564	5.5367

Table 9. Confusion matrix for faulty phase detection using SK and SVM in case study 2.

Training	Actual Class		
	Phase A (Faulty)	Phase B (Normal)	Phase C (Normal)
Predicted Phase A (faulty)	29	0	0
Predicted Phase B (normal)	4	33	0
Predicted Phase C (normal)	0	0	33
Accuracy = (29 + 33 + 33)/(29 + 0 + 0 + 4 + 33 + 0 + 0 + 0 + 33) = 95.95%			
Test	**Phase A**	**Phase B**	**Phase C**
Predicted Phase A (faulty)	23	0	0
Predicted Phase B (normal)	9	32	0
Predicted Phase C (normal)	0	0	32
Accuracy = (23 + 32 + 32)/(23 + 0 + 0 + 9 + 32 + 0 + 0 + 0 + 32) = 90.625%			

5. Discussion

Magnetizing inrush currents in transformers result from abrupt changes in the magnetizing voltage. These currents in the transformer may be caused by energizing an unloaded transformer, occurrence of an external fault, voltage recovery after clearing an external fault and out of phase synchronization of a connected generator [52]. Also, the magnitude of the magnetizing inrush current is usually 10 to 15 times the rated current. Besides, a high inrush current occurs when an AC motor is energized. Typically, during the initial half cycle, the inrush current is often higher than 20 times the normal full load current. After the first half-cycle the motor begins to rotate and the starting current subsides to four to eight times the normal current for several seconds. As a motor reaches running speed, the current subsides to its normal running level [53]. Obviously, these inrush currents have many unfavourable effects, including differential protection maloperation, deterioration of the insulation material and mechanical support structure of windings, voltage sag and other power quality issues on the high voltage (HV), as well as on lower voltage terminals.

An overcurrent is either an overload current or a short-circuit current. Overload currents are most often between one and six times the normal current level. Usually, they are caused by harmless temporary surge currents that occur when motors start up or transformers are energized [54]. Such overload currents are normal occurrences. Since they are short duration, any temperature rise is trivial and has no harmful effect on the circuit components.

To verify the superiority of the proposed method (SK) in this paper, the proposed method can distinguish between the short-time overload current or inrush current caused by the motor starting and overcurrent caused by the SLGF. Figure 10 shows the current and voltage waveforms simulated on the ESE when the load 2 of the case study in Figure 3 has a motor starting. It is obvious by comparing with Figures 4 and 10 that the current waveforms of motor starting are different from those of SLGF. When the motor is starting and the SLGF occurs at a different time, Table 10 shows the results of identification for distinguishing between the short-time overload current caused by the motor starting and overcurrent caused by the SLGF on phase A. These currents which are monitored on the ESE are Ia, Ib, and Ic for phase A, B, and C, respectively. The total average value of the recognition accuracy is 100% for both training and test.

Table 10. Results of recognition accuracy in training and test between SLGF on the line 2 and motor starting current on the load 2 with the SK.

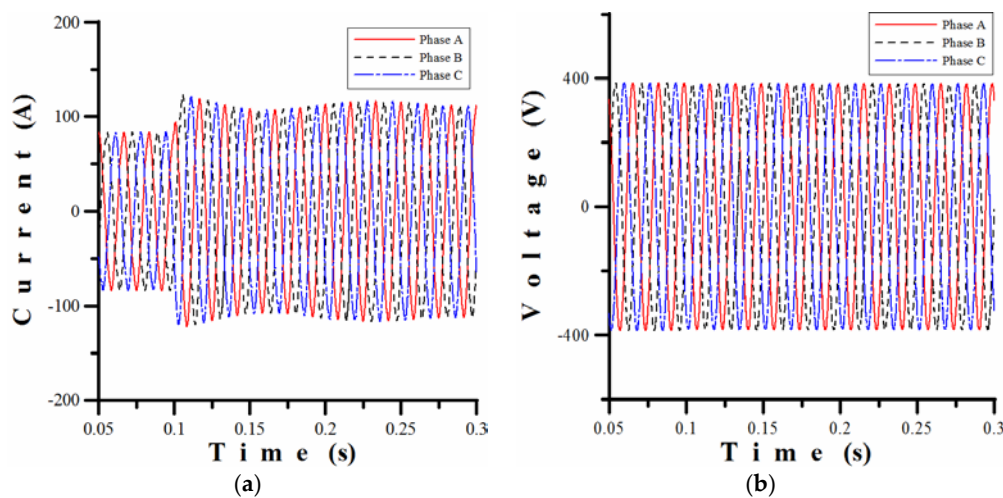| Recognition Accuracy (%)	SVM	BP-ANN	1-NN
I_a	100	100	100
I_b	100	100	100
I_c	100	100	100

Figure 10. Current and voltage waveforms of motor starting for the load bus 2 of the case study. (**a**) Current waveforms; and (**b**) voltage waveforms.

6. Conclusions

This paper has presented statistical feature extraction methods combined with SVM to enhance the recognition accuracy of fault locations and faulty phase detection in low voltage distribution systems using NIFM techniques. The statistical variables of the fault current data have been computed during one cycle since the SLGF occurrence is detected by DWT. In comparison with the previous work on NILM, one major advantage of this concept is that the extraction of transient feature is less computationally burdensome since the statistical variables only need one simple calculation within one cycle. This concept of feature extraction can effectively reduce the dimensions of data inputs for machine learning processing. To verify the validity of the proposed method, three popular AI algorithms used in NIFM are compared in this paper. If the proposed methods are followed, SVM is the best candidate for the disaggregation when SLGF occurs. The evolution of the outcomes in SVM from SMs to SK also clarifies that a proper feature extraction does not need a complicated model for disaggregation. In other words, the better the feature extraction is, the less the machine classifiers learn. Moreover, the effect of good feature extraction, i.e., SKs, for improving the disaggregation is also demonstrated in this paper.

Those statistical features characterize the distinction among the locations of the fault current signals as well as each phase. The experimental results also show that the monitored current of phase A is the best when SLGF occurs on the phase A by using the proposed methods.

The case studies include some of the most challenging scenarios for a NIFM system such as system voltage variations and SLGF on phase A. The proposed methods (SK and SVM) demonstrate the enhanced capability for fault location and faulty phase detection, compared with other existing statistical methods such as Z-scores and SMs. So far, the dataset just considers only the variations of transformer taps for the first attempt in the new approach of NIFM. In future works, the proposed techniques will be implemented for other fault types, ex., DLGF, DLF, and balanced faults. The variations of fault resistances and inception angles should be considered. Furthermore, more complicated variations of the dataset are also included in order to enhance the performance of the NIFM scheme in complicated power systems.

Author Contributions: Hsueh-Hsien Chang and Nguyen Viet Linh conceived and designed the experiments; Nguyen Viet Linh performed the experiments; Hsueh-Hsien Chang and Nguyen Viet Linh analyzed the data; Hsueh-Hsien Chang contributed reagents/materials/analysis tools; Hsueh-Hsien Chang and Nguyen Viet Linh wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Hart, G.W. Nonintrusive appliance load monitoring. *Proc. IEEE* **1992**, *80*, 1870–1891. [CrossRef]
2. Chang, H.H. Genetic algorithms and non-intrusive energy management system based economic dispatch for cogeneration units. *Energy* **2011**, *36*, 181–190. [CrossRef]
3. Laughman, C.; Lee, K.; Cox, R.; Show, S.; Leeb, S.B.; Norford, L.; Armstrong, P. Power signature analysis. *IEEE Power Energy Mag.* **2003**, *1*, 56–63. [CrossRef]
4. Zeifman, M.; Roth, K. Nonintrusive appliance loads monitoring: Review and outlook. *IEEE Consum. Electron.* **2011**, *57*, 76–84. [CrossRef]
5. Cole, A.I.; Albicki, A. Data extraction for effective non-intrusive identification of residential power loads. In Proceedings of the IEEE Instrumentation and Measurement Technology Conference (IMTE/98), St. Paul, MN, USA, 18–21 May 1998; pp. 812–815.
6. Chang, H.H.; Yang, H.T. Applying a non-intrusive energy management system to economic dispatch for a cogeneration system and power utility. *Appl. Energy* **2009**, *86*, 2335–2343. [CrossRef]
7. Roos, J.G.; Lane, I.E.; Lane, E.C.; Hanche, G.P. Using neural networks for non-intrusive monitoring of industrial electrical loads. In Proceedings of the IEEE Instrumentation and Measurement Technology Conference, Hamamatsu, Japan, 10–14 May 1994; pp. 1115–1118.
8. Akbar, M.; Khan, D. Modified nonintrusive appliance load monitoring for nonlinear devices. In Proceedings of the IEEE International Multi-topic Conference, Lahore, Pakistan, 28–30 December 2007; pp. 1–5.
9. Hassan, T.; Javed, F.; Arshad, N. An empirical investigation of VI trajectory based load signatures for non-intrusive load monitoring. *IEEE Trans. Smart Grid* **2014**, *5*, 870–878. [CrossRef]
10. Teshome, D.F.; Huang, T.D.; Lian, K.L. Distinctive load feature extraction based on fryze's time-domain power theory. *IEEE Power Energy Technol Syst. J.* **2016**, *3*, 60–70. [CrossRef]
11. Chang, H.H.; Chen, K.L.; Tsai, Y.P.; Lee, W.J. A new measurement method for power signatures of non-intrusive demand monitoring and load identification. *IEEE Trans. Ind. Appl.* **2012**, *48*, 764–771. [CrossRef]
12. Chang, H.H. Non-intrusive demand monitoring and load identification for energy management systems based on transient feature analyses. *Energies* **2012**, *5*, 4569–4589. [CrossRef]
13. Tsai, M.S.; Lin, Y.H. Modern development of an adaptive non-intrusive appliance load monitoring system in electricity energy conservation. *Appl. Energy* **2012**, *96*, 55–73. [CrossRef]
14. Lin, Y.H.; Tsai, M.S. Development of an improved time-frequency analysis-based non-intrusive load monitor for load demand identification. *IEEE Trans. Instrum. Meas.* **2013**, *63*, 1470–1483. [CrossRef]
15. Gillis, J.M.; Alshareef, S.M.; Morsi, W.G. Nonintrusive load monitoring using wavelet design and machine learning. *IEEE Trans. Smart Grid* **2016**, *7*, 320–328. [CrossRef]
16. Huang, S.J.; Jou, M.J. Application of arithmetic coding for electric power disturbance data compression with wavelet packet enhancement. *IEEE Trans. Power Syst.* **2004**, *19*, 1334–1341. [CrossRef]
17. Puliafito, V.; Vergura, S.; Carpentieri, M. Fourier, wavelet, and Hilbert-Huang transforms for studying electrical users in the time and frequency domain. *Energies* **2017**, *10*, 188. [CrossRef]
18. Hosseini, S.A.; Amjady, N.; Velayati, M.H. A Fourier based wavelet approach using Heisenberg's uncertainty principle and Shannon's entropy criterion to monitor power system small signal oscillations. *IEEE Trans. Power Syst.* **2015**, *30*, 3314–3326. [CrossRef]
19. Hong, Y.Y.; Chen, Y.Y. Placement of power quality monitoring using enhanced genetic algorithm and wavelet transform. *IET Gener. Transm. Distrib.* **2011**, *5*, 461–466. [CrossRef]
20. Zoha, A.; Gluhak, A.; Imran, M.A.; Rajasegarar, S. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors* **2012**, *12*, 16838–16866. [CrossRef] [PubMed]
21. Guedes, J.D.S.; Ferreira, D.D.; Barbosa, B.H.G. A non-intrusive approach to classify electrical appliances based on higher-order statistics and genetic algorithm: A smart grid perspective. *Electr. Power Syst. Res.* **2016**, *140*, 65–69. [CrossRef]
22. Bouhouras, A.S.; Milioudis, A.N.; Labridis, D.P. Development of distinct load signatures for higher efficiency of NILM algorithms. *Electr. Power Syst. Res.* **2014**, *117*, 163–171. [CrossRef]

23. Srinivasan, D.; Ng, W.S.; Liew, A.C. Neural-network-based signature recognition for harmonic source identification. *IEEE Trans. Power Deliv.* **2006**, *21*, 398–405. [CrossRef]

24. Saitoh, T.; Osaki, T.; Konishi, R.; Sugahara, K. Current sensor based home appliances and state of appliance recognition. *SICE J. Contr. Meas. Syst. Integr.* **2010**, *3*, 86–93. [CrossRef]

25. Du, L.; Restrepo, J.A.; Yang, Y.; Harley, R.G.; Habetler, T.G. Nonintrusive, self-organizing, probabilistic classification and identification of plugged-in electric loads. *IEEE Trans. Smart Grid* **2013**, *4*, 1371–1380. [CrossRef]

26. Agyeman, K.A.; Han, S.; Han, S. Real-time recognition non-intrusive electrical appliance monitoring algorithm for a residential building energy management system. *Energies* **2015**, *8*, 9029–9048. [CrossRef]

27. Chunju, F.; Li, K.K.; Chan, W.L.; Weiyong, Y.; Zhaoning, Z. Application of wavelet fuzzy neural network in locating single line to ground fault (SLG) in distribution lines. *Electr. Power Energy Syst.* **2007**, *29*, 497–503. [CrossRef]

28. Ekici, S.; Yildirim, S.; Poyraz, M. Energy and entropy-based feature extraction for locating fault on transmission lines by using neural network and wavelet packet decomposition. *Expert Syst. Appl.* **2008**, *34*, 2937–2944. [CrossRef]

29. Reddy, M.J.B.; Rajesh, D.V.; Gopakumar, P.; Mohanta, D.K. Smart fault location for smart grid operation using RTUs and computational intelligence techniques. *IEEE Syst. J.* **2014**, *8*, 1260–1271. [CrossRef]

30. Borghetti, A.; Bosetti, M.; Silvestro, M.D.; Nucci, C.A.; Paolone, M. Continuous-wavelet transform for fault location in distribution power networks: Definition of mother wavelets inferred from fault originated transients. *IEEE Trans. Power Syst.* **2008**, *23*, 380–388. [CrossRef]

31. Bezerra Costa, F. Fault-induced transient detection based on real-time analysis of the wavelet coefficient energy. *IEEE Trans. Power Deliv.* **2014**, *29*, 140–153. [CrossRef]

32. Ehrenfeld, S.; Littauer, S.B. *Introduction to Statistical Methods*; McGraw-Hill: New York, NY, USA, 1964.

33. Mack, C. *Essentials of Statistics for Scientists and Technologists*; Helnemann Educational Books Ltd.: London, UK, 1966.

34. Cadzow, J.A.; Li, X.K. Blind deconvolution. *Digit. Signal Process.* **1995**, *5*, 3–20. [CrossRef]

35. Dwyer, R. Use of the kurtosis statistic in the frequency domain as an aid in detecting random signals. *IEEE J. Ocean. Eng.* **1984**, *9*, 85–92. [CrossRef]

36. Tesei, A.; Regazzoni, C.S. The asymmetric generalized Gaussian function: A new HOS-based model for generic noise pdfs. In Proceedings of the IEEE Signal Processing Workshop on Statistical Signal and Array Processing, Corfu, Greek, 24–26 June 1996; pp. 210–213.

37. Dyer, D.; Stewart, R.M. Detection of rolling element bearing damage by statistical vibration analysis. *ASME J. Mech. Des.* **1978**, *100*, 229–235. [CrossRef]

38. Martin, H.R.; Honarvar, F. Application of statistical moments to bearing failure detection. *Appl. Acoust.* **1995**, *44*, 67–77. [CrossRef]

39. Heng, R.B.W.; Nor, M.J.M. Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition. *Appl. Acoust.* **1998**, *53*, 211–226. [CrossRef]

40. OptiFuse. Rev A 01/2010. Available online: http://uk.trgcomponents.com/uploads/media/FuseSelectionGuide_RevA_01.pdf (accessed on 1 February 2017).

41. ABB Inc. *Miniature Circuit Breaker-Application Guide*; ABB Inc.: Wichita Falls, TX, USA, 2009.

42. Meyer, Y. *Wavelets: Algorithms and Applications*; SIAM: Philadelphia, PA, USA, 1993.

43. Robertson, D.C.; Camps, O.I.; Mayer, J.S.; Gish, W.B. Wavelets and electromagnetic power system transients. *IEEE Trans. Power Deliv.* **1996**, *11*, 1050–1056. [CrossRef]

44. Silva, M.D.; Oleskovicz, M.; Coury, D.V. A hybrid fault locator for three-terminal lines based on wavelet transforms. *Electr. Power Syst. Res.* **2008**, *78*, 1980–1988. [CrossRef]

45. Kim, C.H.; Kim, H.; Ko, Y.H.; Byun, S.H.; Aggarwal, R.K.; Johns, A.T. A novel fault-detection technique of high-impedance arcing faults in transmission lines using the wavelet transform. *IEEE Trans. Power Deliv.* **2002**, *17*, 921–929. [CrossRef]

46. Silva, K.M.; Souza, B.A.; Brito, N.S.D. Fault detection and classification in transmission lines based on wavelet transform and ANN. *IEEE Trans. Power Deliv.* **2006**, *21*, 2058–2063. [CrossRef]

47. Bishop, C.M. *Neural Networks for Pattern Recognition*; Clarendon Press: Oxford, UK, 1995.

48. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995.

49. Hsu, W.; Lin, C.J. A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Netw.* **2002**, *13*, 415–425. [PubMed]

50. Ravikumar, B.; Thukaram, D.; Khincha, H.P. Application of support vector machines for fault diagnosis in power transmission system. *IET* **2008**, *2*, 119–130. [CrossRef]

51. Wagner, S. Heuristic Optimization Software Systems—Modelling of Heuristic Optimization Algorithms in the HeuristicLab Software Environment. Ph.D. Thesis, Johannes Kepler University, Linz, Austria, 2009.

52. Jamali, M.; Mirzaie, M.; Asghar Gholamian, S. Calculation and analysis of transformer inrush current based on parameters of transformer and operating conditions. *Electron. Electr. Eng.* **2011**, *109*, 17–20. [CrossRef]

53. Cooper Bussmann. Motor Protection: Basic Explanation. 2005. Available online: http://www.cooperindustries.com/content/dam/public/bussmann/Electrical/Resources/solution-center/technical_library/BUS_Ele_Tech_Lib_Overload_Currents.pdf (accessed on 1 February 2017).

54. Cooper Bussmann. Fuseology: Overcurrents and Voltage Ratings. 2005. Available online: http://www.cooperindustries.com/content/dam/public/bussmann/Electrical/Resources/solution-center/technical_library/BUS_Ele_Tech_Lib_Overload_Currents.pdf (accessed on 1 February 2017).