

Ultrafast Transmission Line Fault Detection Using a DWT-Based ANN

Ahmad Abdullah, *Member, IEEE*

Abstract—Digital impedance protection of transmission lines suffers from known shortcomings not only as a principle but also as an application. This necessitates developing a new relaying principle that overcomes those shortcomings. Such a principle is offered in this paper and is currently being validated using field data. The principle is a new application of wavelet-based artificial neural networks (ANNs). The application uses high-frequency content of a subset of local currents of one end of a protected line to classify transients on the line protected and its adjacent lines. The scheme can classify transients—including faults—occurring on a protected line, categorize transients on adjacent lines, and pinpoint the line causing the transient event. It is shown that the feature vector of the event can be determined from a subset of local currents without using any voltages altogether. The subset of local currents consists of the two aerial modes of the local current. Modal transformation is used to transform phase currents to modal quantities. Discrete wavelet transform (DWT) is used to extract high-frequency components of the two aerial modal currents. A feature vector is built using the wavelets details coefficients of one level of the aerial modes and is used to train an ANN. Results show that the classes corresponding to each transient event type on the protected line and its adjacent lines are almost linearly separable from each other. Results demonstrate that very accurate classification using one-eighth of a cycle of postevent data is possible.

Index Terms—Artificial neural networks (ANNs), current transformers, line switching, lightning strike, modal analysis, power system faults, transmission line relaying, wavelets transform.

I. INTRODUCTION

PROTECTION of a transmission line involves installing relays at both ends of the line that constantly monitor voltages and currents and act when a fault occurs on a line. Traditional protection uses phasor estimation to estimate the fundamental component of voltages and currents and takes a decision when certain criteria corresponding to certain fault conditions are met. Distance protection is the most widely used method for transmission line protection. Modern numerical distance relays make use

of low-pass and anti-aliasing filters—to reject high-frequency content and apply Nyquist sampling theorem—and special digital signal processing (DSP) hardware to perform sophisticated math functions [1]. DWTs are generally used to measure the voltages making them available to the relay. It is known that due to the interaction between the capacitive voltage divider and the transformer inductance, oscillations are imposed on the fundamental frequency measured [2]. This puts more stringent requirement on the filters used in the relay. Additionally, distance relay can only protect up to 85% of the line instantaneously [3]. This necessitates the use of a communication link between the relays at the two ends of the line to achieve fast tripping from both ends. With relays connected to substations, a communication link between relays is a cyber-threat. With all limitations mentioned above, a new relaying principle is needed. This principle has to be using currents only, fast, and finally does not need any communication link for its operation. Such a principle is provided and theoretically tested in this paper.

Except for traveling wave fault location techniques that are still not very popular with relay manufacturers [4], little is done with the high-frequency components of the transients even though it has been known that high-frequency components of faults and transient events contain rich information [5]. This is mainly because the traditional mathematical method at hand was fast Fourier transform (FFT) [6]. The FFT requires the signal be stationary in the wide sense for the calculation of coefficients to be accurate, i.e., the signal cannot have any temporal variations for the calculation of coefficients to be correct and accurate. Short window Fourier transform [7] can solve some of the problems with the FFT but it introduces other issues. However, most signals encountered in power systems are not stationary but have their characteristics change with time.

Multiresolution analysis (MRA) is a signal processing tool that has been introduced in the nineties [8] to solve some of the problems inherent in Fourier transform analysis methods. Wavelets [9] are usually used along with MRA to solve this very exact problem. Wavelets have a strong localization property that enables studying changes not only in time but also in frequency feasible. It is known from the uncertainty principle [10] that if the signal spans a small portion of the time domain, then its Fourier transform will span a large portion of frequency domain. The discrete wavelet transform (DWT) on the other hand does not suffer from such a limitation. On the contrary, the signal is approximated at various levels—frequency bands—and changes in time manifest themselves at the coefficients of the wavelets only around the time in which the event occurred. Such property

Manuscript received May 5, 2016; revised October 25, 2016 and May 13, 2017; accepted November 9, 2017. Paper 2016-PSPC-0365.R2, presented at the 2016 IEEE Power and Energy Conference at Illinois, Champaign, IL, USA, Feb. 19–20, and approved for publication in the IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS by the Power System Protection Committee of the IEEE Industry Applications Society.

The author is with the Department of Electrical Power and Machines, Faculty of Engineering, Cairo University, Giza 12613, Egypt, and also with Electric Power Engineers, Inc., Austin, TX 78738 USA (e-mail: ahmad.abdullah@cu.edu.eg).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIA.2017.2774202

makes it ideal to detect disturbances and study power system transients [11].

With the advent of wavelets, the power system community has seen a surge in application of wavelets-based methods for various power system problems notably in the area of power system protection—mainly in transient-based protection schemes, transients' classification, and fault location. The localization property of wavelets makes it very convenient in locating faults [12].

The use of artificial neural network (ANN) for fault classification and detection has been given in [13]–[15] where voltage and current samples are fed directly to the ANN for fault detection and classification on the protected line. In [14] and [13], samples of voltages and currents are used as a feature vector to train ANN. With the advent of wavelets, special wavelets transforms are applied to voltage and current waveforms before they are fed to the ANN for training. In [16] and [17], the DWT is applied to voltages and currents but instead of feeding the details coefficients to the ANN, entropy (energy) of the signal is captured and fed instead to the ANN to make fault-type classification between faults on the same transmission line. In [18] and [19], the DWT is applied to voltage and current signals at a relay location resulting in series of details coefficients that are fed to the ANN for fault detection and type classification. In [20], the energy of certain current levels and approximations is used to train a probabilistic classifier, and using this energy feature, a decision is made whether a transient signal is due to a fault or nonfault condition on a line. A detailed look at the literature reveals that very few attempts have been carried out to harness the power of the wavelets details coefficients alone and understand their underlying structure. Moreover, existing classification techniques as the ones in [19] do not take transients on adjacent lines into account, which could mislead these schemes. Additionally, symmetrical line configuration has been universally assumed, which is a configuration nearly nonexistent in real power systems.

In this paper, the wavelets details coefficients are directly used to train the ANN without using any entropy-based method. This paper proposes a novel application for wavelet-based ANN. It is shown that using the details coefficients of a subset local currents only, distinction can be made between fault and nonfault conditions on a protected transmission line without using any voltage signals at all. The proposed algorithm can also distinguish between faults on the protected line and faults on adjacent lines. The algorithm provided can not only tell the difference between forward and reverse faults but also can determine which adjacent line is faulted. It is also shown that the same applies for lightning strikes and line switching cases, i.e., the algorithm can tell which line is causing which transient event. Just as the FFT produces coefficients that correspond to certain frequencies, wavelets details also give information regarding the oscillatory components of the signal localizing them in time. In this regard, the DWT is used to extract useful oscillatory information about the signal. Oscillatory information is manifested in the wavelets details coefficients at various levels. It is shown that any transient event on a specific transmission line causes currents to oscillate in a unique way and these oscillations can be detected in the wavelets details coefficients themselves at

various levels. The use of these wavelets details coefficients shows the possibility of building ultrahigh speed relays.

The paper is organized as follows. A quick overview of both the DWT and ANN is offered in Section II. The rationale behind the proposed principle is presented in Section III. The solution methodology is given in Section IV. The simulation platform, which consists of a description of MATLAB model and the creation of transient cases for ANN training, is described in Section V. The structure of the feature vector is given in Section VI. Signal analysis is provided in Section VII. Simulation results are presented in Section VIII. Error analysis is provided in Section IX. Comparison between the method proposed in this paper and existing methods in the literature is presented in Section X. Conclusion and future research are provided in Section XI.

II. BACKGROUND

In this section, a brief overview of both the DWT and ANN as used in this paper is provided.

A. Discrete Wavelet Transform

In this paper, the dyadic wavelet transform is used [10]. The transform takes the signal and applies low- and high-pass filters to it. The transform converts the original signal into independent signals each of which spanning a certain frequency band. The frequency bands are determined by the number of levels the original signal is sought to be analyzed and each level corresponds to a unique frequency band. The word independent means that one level cannot be derived from another, i.e., level 2 cannot be derived from level 1. The frequency band of each level depends on the sampling frequency and Nyquist theory. The highest frequency that can be seen in the decomposed signal will be at most equal to half of the sampling frequency. Given a certain sampling frequency, the dyadic transform will first apply high- and low-pass filters to the signal resulting in two signals. The signal corresponding to the output of the high-pass filter is called level 1 and the other signal is called approximation 1. Both of those signals are independent. One can stop at this step or can apply high- and low-pass filters to the first approximation to get level 2 and approximation 2. Level 2 will span a frequency band corresponding to the upper-half of the frequency band of approximation 1, while approximation 2 will occupy the lower frequency band. Continuing this manner, by applying successive low- and high-pass filters, one obtains a set of levels—also called details—and one last approximation. In theory, the last approximation should correspond to a pure sine wave assuming that the high-frequency components imposed on the power frequency have frequencies that are not included in the frequency band of the last approximation. Taking a numerical example, if one uses 200 kHz sampling frequency, then the resulting DWT decomposition is shown in Fig. 1. The details coefficients of the DWT transform are provided in (1) [10], where $\varphi(t)$ is the mother wavelet used, $f(t)$ is the signal to be analyzed, parameter “ a ” causes scaling (which determines the level) and “ b ” causes shift in time. In practice, it is not needed to apply the transform all over the infinite time line. Since wavelets have a strong localization property, it is only needed to apply the transform to

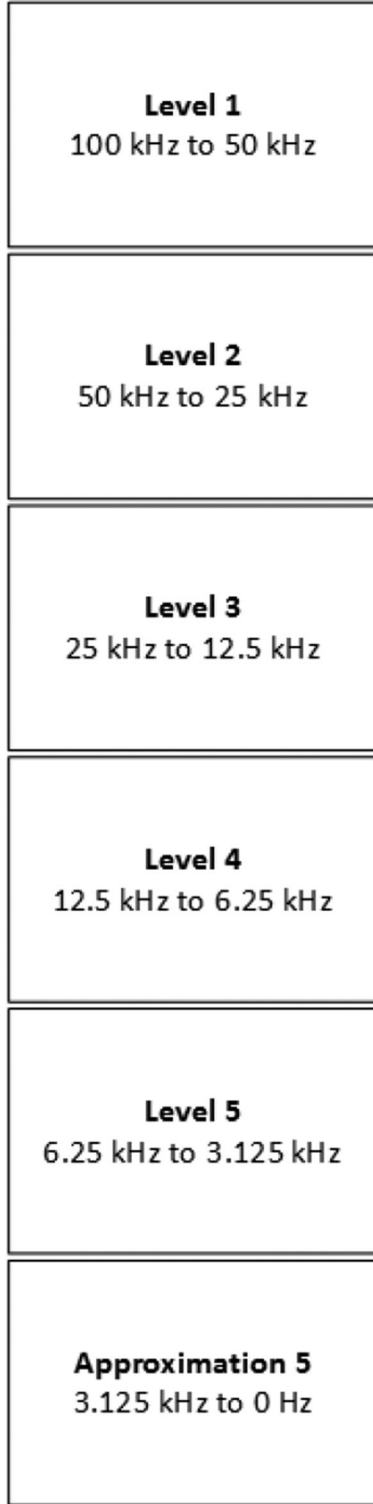


Fig. 1. DWT at 100 kHz sampling rate for five levels.

197 the time period under study

$$Wf(a, b) = \frac{1}{\sqrt{a}} \times \int_{-\infty}^{\infty} f(t) \times \overline{\varphi\left(\frac{t-b}{a}\right)} dt. \quad (1)$$

B. Classification With Artificial Neural Networks

198

ANN is a very well-established tool for classification. The classification done in this paper is not probabilistic but rather deterministic. The inputs to ANN are vectors in the n -dimensional space, where n is the size of each input vector that needs to be mapped to another output vector space. The size of the output vector space is determined by the number of output classes one wants to map the input vectors to. In simple terms, this classification can be shown to be carried out by ANN [21]. All ANNs in this paper consist of three layers: an input layer, a hidden layer, and an output layer. Each layer consists of a certain number of neurons. The connections between those neurons are called synapses. The input layer consists of junctions that represent the input. The number of junctions must equal to “ n ,” which is the dimension of the input vector. The hidden layer can consist of any number of neurons. The classification accuracy is greatly affected by the number of neurons in the hidden layer. The output layer consists of neurons each of which is activated by a function. The inputs to this function come from the hidden layer. Classes or more specifically output classes are the patterns one wants to map the input to. The main idea behind ANN classification is that if the inputs belong to linearly separable regions of the n -dimensional space, then using ANN, one can draw hyperplanes in that space so that the space between those hyperplanes contain only one region and each region is then mapped to one of the output classes using the activation function in the output neuron. The weights or more specifically synaptic weights of the synapses are adjusted during the training phase such that the vector space is divided into regions each of which corresponds to a certain class. Any output neuron receives the input vector through synaptic weights. The resultant is then applied to the function of the neuron that has the form $g(x) = 0$, which is an equation of the hyperplanes in the n -dimensional space. The function $g(x) = 0$ is realized in the ANN by the neurons and the synaptic weights. The output of the neuron is activated only if $g(x)$ is positive for the corresponding input vector. In this paper, the author uses the widely known backward propagation algorithm for the calculation of weights. Knowledge is stored in the ANN through those weights. Only supervised learning is done in this paper. The algorithm of the steepest gradient is used throughout the paper [21].

III. RATIONALE BEHIND THE PROPOSED METHOD

239

When a fault or any transient event occurs, it launches a traveling wave as well as high-frequency oscillatory components. Traveling waves can be easily quantified as they arrive at the line terminals but the high-frequency oscillatory components cannot. The reason for the oscillatory components not being easily quantified before in the literature will be explained below.

At this point, the author needs to introduce new terminology. Speaking of a certain transmission line, the author calls a fault on that line a fault case. A fault case has its parameters. Those parameters are incipient angle, fault resistance, fault type, and fault location. Thus, a certain fault case on a specific transmission line causes voltages and currents to oscillate in a manner

in accordance with the parameters of that fault case. The formal solution of the currents and voltages of a single phase line is given by (called telegrapher equations) [22]

$$\frac{\partial V}{\partial x} = R \times I + L \times \frac{\partial I}{\partial t} \quad (2)$$

$$\frac{\partial I}{\partial x} = G \times V + C \times \frac{\partial V}{\partial t} \quad (3)$$

where I and V are the current and the voltage anywhere on the line; R and G are the resistance and conductance per unit length of the line; L and C are inductance and capacitance of the line per unit length; x is the distance from a zero reference frame generally taken to be at either end of the line.

A general closed-form solution of (2) and (3) is hard and generally impossible as it depends on the boundary and initial conditions of the case involved. Perhaps the most straightforward solution method is to obtain the solution in terms of an infinite time series expansion. Using infinite time series expansion would mystify the solution, because the author's intention is to analyze the frequency content of the signal as this content changes over time. Application of Laplace and Fourier transforms to those equations produces integrals that are yet to be solved formally. Moreover, when a three-phase line is studied, the above-mentioned equations become six equations (two equations for each phase: 1) a voltage; and 2) current equation). Thus, decoupling them is hard undertaking if not impossible in most cases. If we have mutual coupling between two lines sharing the same tower, then we have three more equations making them a total of nine equations. Bearing in mind that those nine equations are for one line only, those equations then need to be solved simultaneously along with all other equations in the system. For today's power systems, which consist of thousands of lines, solving all equations analytically in a closed form is behind human ingenuity.

The main aim is to detect the fault once it is initiated. It is very apparent that the high-frequency fault-generated transients are different for each fault case on a specific line. This is because each fault case on a specific line has its set of equations as given in (2) and (3) with unique boundary (lines connected to it) and initial conditions. Moreover, the same fault case on different lines will cause different fault-generated transients because each case has its own boundary and initial conditions, which is different for all neighboring lines. For this reason, no attempt is made to solve the equations formally but solve them numerically by running EMTP simulations, extracting the high-frequency content and then analyzing them.

Since the author argues that those frequencies—which change over time—are different for each line, the author only needs to see the aggregate of all those frequency components for each line. If intuition is true, then the aggregate of those fault-generated high frequencies will be different for each line. What remains is to investigate how the feature space can be defined such that this difference can be attained and this is done in Section VI. If it is possible to map those frequency components to an n -dimensional space, then faults corresponding to a certain line should be separable from faults on adjacent lines. The

author takes this idea further and investigates whether the aggregate of all faults from each line is actually linearly separable [21] from each other. This is the main argument behind the use of ANN as a pattern recognition approach.

For any fault occurring on a transmission line, the transient-generated frequency components depend on the initial and the boundary conditions (the lines connected to the faulted line). Since by definition all adjacent lines are different, the frequency-generated transients for faults on each transmission line should be unique for each line. The question is whether those frequency components can be combined in a certain way to form a pattern to be recognized.

The same argument not only applies to faults but also applies to switching events and lightning strikes on transmission lines. The transient event types that are studied in this paper are faults, lightning strikes, and line switching. Evidently, the frequency content of faults—on all lines—should be different from switching, which in turn is different from lightning because the equations governing the solutions are always different as different transient event types have different boundary and initial conditions. Those event types can be first recognized and then the event can be traced back to the line originating it as given in Section VIII-D. This basically means that there is a mapping that can divide a certain subspace to linearly separable regions that correspond to those event types. If one zooms in that event subspace, one can see which line is causing that event type.

IV. SOLUTION METHODOLOGY

In this paper, it is argued and shown that the oscillatory information present in the transient signal captured at a single relay location caused by sudden network topology changes contains sufficient information for classification not only between transients on the same line but also between transients on adjacent lines [23]. Any change of configuration on the line causes a traveling wave to be generated traveling from the point of change toward the ends of the line. In the simplest case, this wave will be just a pulse—a step—but in reality will be accompanied by high-frequency oscillating components. Fourier analysis is not suitable for analyzing such waveforms because those oscillations will be distorted and attenuated as they arrive at the line terminals. However, applying the DWT will enable us to see both the spectral and temporal variations of those high-frequency oscillations.

Historical treatment of traveling waves—as used in the ladder diagram, for example—treats them as if they were pulses traveling down the line with no regard to the oscillatory components they carry. A certain mathematical entity—which is the feature vector to be defined in Section VI—is derived from the oscillatory components and called the transient signature of the event. The transients that are studied in this paper are lightning, line switching, and faults. At a certain terminal of the line, which is typically a relay location, it is argued and shown that those signatures are unique to the event that originated them and to the line that initiated those events. This means that us-

ing this oscillatory information of the event, the line causing it can be determined. Those signatures are a function of the line parameters, the network topology and the instance of the event, and, of course, the type of the event. In essence, the signature of the fault occurring on a certain line will be different from the signature of the fault on adjacent lines. Also, line switching will cause the currents to oscillate in a manner that is different from faults and lightning striking the same line and those oscillatory components are different for different lines as well.

It is argued that the aggregate of all transient signatures caused by a certain transient event—switching or faults, for example—originating from a certain line occupies a specific subspace in the n -dimensional space. This subspace is almost linearly separable [21] from all other subspaces generated by other transient events by adjacent lines. The n -dimensional space is the space spanned by all n -dimensional feature vectors used to train the ANN. A relay that is programmed to use these wavelets details not only can detect and classify transients on a protected line but also can detect faults and classify transients on adjacent lines. ANN classification is used to show that this is indeed the case. Static feedforward ANN is a tool used to map a vector from the n -dimensional space to another m -dimensional space. In the case at hand, the feature vector is mapped, which is n -dimensional space to another output space. These output spaces will be described in Section VIII, but for now if this mapping is successful, then the original n -dimensional subspaces are linearly separable as explained in [21].

Information is extracted from the high-frequency components using the DWT at various levels. Before the DWT is applied, phase currents are decoupled from each other using a modal analysis [24], [25]. Modal transformation will decouple current waveforms, thus eliminating mutual coupling and untransposed line effects. The modal matrix will be explained in Section VI. It is known that CVT introduces transients to the measured voltage signal. It is shown in [26] that a typical CVT with active ferroresonance suppression circuit is a low-pass filter with a cut-off frequency of around 1 kHz, which drops to almost 200 Hz when passive ferroresonance suppression circuit is employed. On the other hand, classification with currents is preferable because the cut-off frequency of current transformers is much larger than the potential transformer or CVT. The useful passband of current transformers is typically 100 kHz [27], and in some cases, can go up to 400 kHz [28] or 500 kHz [29]. Since a conservative stance is chosen, the sampling rate is selected such that the maximum frequency available in the signal is 100 kHz to attain the 100 kHz passband of current transformers. The time step used in the EMTP simulations is 1 μ s, which should theoretically give a maximum frequency of 500 kHz in sampled output according to Nyquist theory. However, it is given in [30] that the maximum frequency in the EMTP simulations is only one-fifth of that, so the maximum frequency is 100 kHz.

After decoupling phase quantities, the DWT is applied to the currents modes to convert the signals to a series of coefficients that will be used for training of the neural networks. The event is detected once a change of level 3—or any level depending on which level is used—coefficients is detected. This detection


is simply because the pre-event data are assumed steady state so no high-frequency components exist in the pre-event data. Once the event is detected, a window of one-eighth of a cycle of postevent information is used for neural network training. Simulations show that at least one-eighth of a cycle of postevent data has to be used for correct classification. Numerous trials have shown that at least one-eighth of a cycle of postevent information has to be present for correct classification or classification fails completely.

The outcome of the DWT is a series of coefficients for each mode and for each level. The coefficients of one level of the two aerial modes of currents are stacked on top of each other to build one feature vector that will be used to train the network. Building the training vector will be explained in Section VI. A neural network of an appropriate size is then selected for classification. The choice of the size of network is reached at by trial and error and no universal size has been applied to all classification problems. That is, the size of the network is changed till the required classification accuracy is achieved. Once the size of ANN is selected, the ANN is then trained using various scenarios for transients on lines. It is shown that using only two modes of local current information at one end of the line—corresponding to a relay, distinction can be made between various transient events on different lines. It is argued and shown that the feature vector built has sufficient information to distinguish between a fault and nonfault condition on a protected transmission line. It is also shown that the adjacent line causing certain transient events can be determined. This means using local information we can see what is happening on adjacent lines.

In summary, the procedure is as follows.

- 1) Decouple currents using modal analysis.
- 2) Apply the DWT to the aerial modes using one-eighth of a cycle of postevent currents.
- 3) Stack the series of coefficients of the two aerial modes of a certain level on top of each other to create a vector used to train ANN.
- 4) A neural network of appropriate size is used for training. Training is done using transient scenarios. These scenarios include faults, line switching, and lightning.

V. SIMULATION PLATFORM

In this section, the /EMTP model is presented in Section V-A, while in Section V-B, creation of the transient cases for the training, validation, and testing of ANN are fully described.

A. EMTP Model

The area under study is shown in Fig. 2. The IEEE 118 bus is used [31] as a test system. All data are taken from [31] except for machine data and the lines in the area under study. The dynamic data of IEEE 118 bus system are taken from [32]. The line under study is the line connecting buses 23 and 25 (line 23-25). The selection of this is arbitrary but this specific line was attractive to the simulations for the reasons that follow. Line 23-25 is bordered by three lines connected to bus 23, namely: line 23-22, line 23-32, and line 23-24. Line 23-25 is also connected

have been created, i.e., AG, BG, CG, AB, BC, CA, ABG, CBG, ACG, ABC, and ABCG. Incipient angles are from 0 to 350° in 10° increments. Fault resistance assumes the values: 0, 20, 100, and 1000 Ω. The distance takes the following values: 5%, 15%, 35%, 50%, 65%, 80%, and 95%, which are all percentages of total line length. Simulations are created for line 23-25 and it adjacent lines.

A total of 8066 cases per batch have been created, which gives a total of 48 396 cases because line 25-27 is double circuit line, i.e., two batches of faults for line 25-27 have been created: one for each circuit. Fault arc although simulated was not taken into account during training. This is because fault arc causes, if any, very little distortion to phase currents, which is consistent with the results in [20] and [13]. The previous statement has been validated using the fault arc model given in [40].

2) *Creation of Line Switching Cases*: For line switching cases, a batch for each line has been created giving a total of six batches. Each batch consists of smaller batches; each of the smaller batches corresponds to switching by one of the circuits breakers at each terminal. For example, line 23-22 would have a smaller batch for switching using the breaker installed on terminal 23 and another smaller batch for switching using the breaker on terminal 22. The variable in switching cases is the moment of switching. Switching times ranges from 0 to 360° in 2° increment. This way 360 cases per batch have been created, giving a total of 2160 cases.

3) *Creation of Lightning Strike Cases*: For lightning strike cases, a batch of each line has been again created. An ATP Heidler type lightning source with rising time equal to 4 μs and τ equal to 10 μs has been used. The rising and tailing times are kept constant during simulations but amplitudes have been varied, which were set to 5, 10, 15, 20, and 30 kA. Striking distances were the same as the ones used for fault batches: 5%, 15%, 35%, 50%, 65%, 80%, and 95%, which are all percentages of total line length. The incipient angle was 0, 90°, 180°, 270°, and 330°. This amounts to 630 cases per batch giving a total of 3780 cases.

VI. BUILDING OF FEATURE VECTOR OF TRANSIENT SIGNATURE

After batches have been created, a special subroutine is used to translate the phase quantities to modal quantities. The phase quantities are decoupled using the modal matrix calculated at 10 kHz even though the matrix is frequency dependent. Different modal frequencies for the modal matrix have been tried but this did not affect the classification results at all. The two aerial modes are the only ones used for training. The modal matrix with real entries is given in (4). The signs of the entries are only shown to emphasize their physical meaning

$$\begin{pmatrix} \text{Mode1} \\ \text{Mode2} \\ \text{Mode3} \end{pmatrix} = \begin{pmatrix} + & + & + \\ - & + & + \\ + & + & - \end{pmatrix} \times \begin{pmatrix} I_a \\ I_b \\ I_c \end{pmatrix}. \quad (4)$$

It should be apparent that Mode1 is the weighted sum of all phase currents, which is nothing other than the ground mode current. This ground mode current is known to be very well

dependent on the frequency and ground resistance on the contrary to the aerial modes that are more or less frequency independent [41]. This is the main reason that the ground current mode has not been used for building the feature vector for training.

Once the modal quantities are available, we run the DWT with db4 for four levels, starting from level 1 to level 4. Db4 has been used since it has produced good results in the previous literature. The outcome of the DWT is a series of coefficients for each level and for each mode. Since the first one-eighth of a cycle following the transient event is only sought, the details coefficients corresponding to that period are only obtained. Two modes out of the three modes are only used. The details coefficients of one level of Mode2 and Mode3 are stacked on top of each other. It should be emphasized that building the n -dimensional vector this way does not change the temporal content of the feature vector, as long as all other n -dimensional vectors are built consistently this way, i.e., the order of the modes in the feature vector has to be preserved and the vector has to be built using one level only. If the order of modes in the feature vector is changed, this amounts to a rotation of n -dimensional space but should not change the results of classification, again as long as all vectors are built consistently.

At this point, one should note that the application uses only two thirds of modes of the currents as contrasted to other applications that use all phase voltage and current waveforms for ANN training [18], which amounts to drastic reduction in computational power needed for ANN.

Having created n -dimensional vectors, the input for the ANN is ready. Training of the ANN is then started. The cases are randomly divided into the following categories: 70% of all cases for training, 15% for validation, and another 15% for testing.

VII. SIGNAL ANALYSIS

Even though ANN training is done offline, yet the principle proposed in this paper can be used for online applications. It should be noted that training has been done using hundreds of thousands of cases. For example, 8066 fault cases per line and a total of 54 336 transient cases have been used to train ANN as given in Section V-B. These cases encompass all variations of any transient that can occur on a transmission line. ATP/EMTP is mature enough that it can capture the actual system response. A real event oscillography would look like the simulated case in EMTP as long as the EMTP model is an accurate representation of the system. It is not possible to exhaust all foreseen transient parameters, for example, it is not computationally feasible to include all values of fault resistance into consideration in the training stage. However, it is evident that the oscillations present in the transient signal corresponding to a fault case with a fault resistance of 20 Ω should have oscillations that fall in between the same fault case but with a fault resistance of 0 and 100 Ω. This notion of “in-betweenness” is illustrated in Fig. 4. As explained in Section IV, the transient signal is eventually translated into a vector that belongs to the n -dimensional space. Thus, if the fault cases corresponding to 5%, 20%, 80%, and 95% line length can be linearly separated from all other transients, then the fault cases corresponding to 50% line length should fall in the

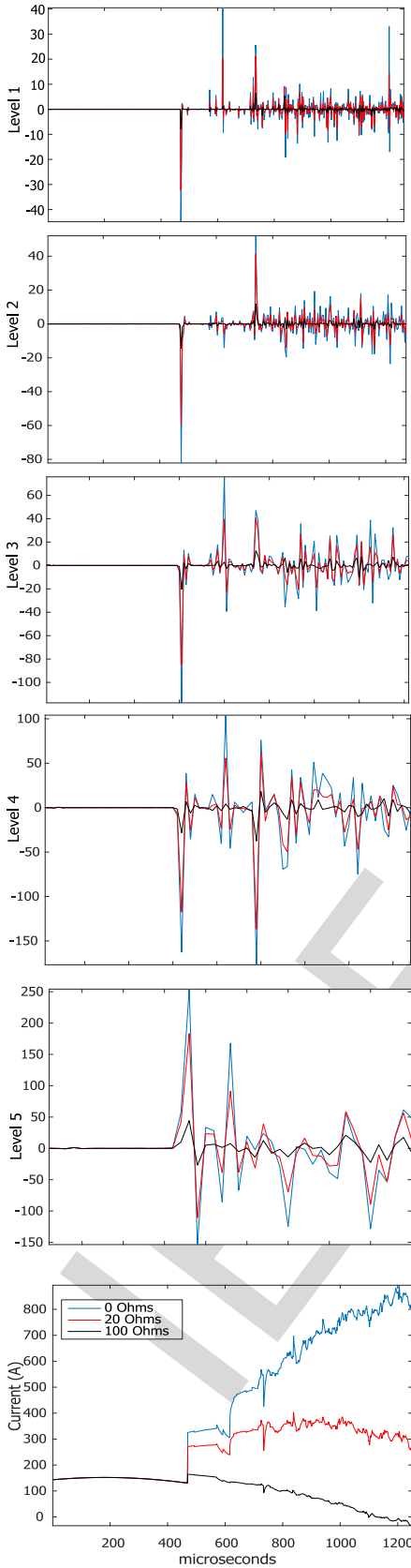


Fig. 4. Mode2 current of a certain fault case at different fault resistances. The same fault type, fault location, and incipient angle are evaluated but at different fault resistances.

TABLE I
FAULTS CLASSIFICATION ACCURACY

Transmission line	Accuracy (%)
Line under study	100
Line 23-32	100
Line 23-24	100
Line 23-22	100
Circuit (1) line 25-27	46
Circuit (2) line 25-27	54

region of those fault cases. The same argument applies for all event types. This idea of “in-betweenness” makes this approach applicable to online applications as long as the extreme cases are included in the offline training.

VIII. RESULTS

This section presents the results of the classification. In Section VIII-A, the results of classification in case all events were faults are shown. In Section VIII-B, results in case events were all line switching are presented. In Section VIII-C, classification output in case of all events were lightning strikes is shown. Finally, event-type classification is provided in Section VIII-D.

In Section IV, the mapping from n -dimensional space to the m -dimensional space has not been discussed. The dimension of the n -dimensional space is fixed by the length of the feature vector, while the dimension of the m -dimensional space varies according to which space we want to map the n -dimensional space to. In Sections VIII-A–VIII-C, the events are being mapped to the lines causing them, which amounts to a mapping from the n -dimensional space to a 6-dimensional space (one dimension per line, line 25-27 counted twice for the double circuit). In Section VIII-D, all events are being mapped to an event-type space, which is a mapping from n -dimensional space to a 3-dimensional space (faults, lightning, and switching). It should be apparent that the dimension of the feature space (n -dimensional space) depends on the DWT level used. It should be obvious as well that this dimension is almost halved as we go one level higher, i.e., the dimension of the n -dimensional space using level 1 is double the dimension of level 2 n -dimensional.

A. Line Identification in Case of Faults

This section provides the results for classification in case all transients were of faults.

First, six batches of faults are created as described in Section V-B1. Creating the feature vector using currents from terminal 23 as described before using any levels from 1 to 4 to train ANN gives the classification accuracy presented in Table I. One immediately sees that classification accuracy for the double circuits of line 25-27 fails. This is to be expected if the reader consults Fig. 3 for the tower configuration, which shows that from the Thevenin’s point of view at the terminal 23 or 25 of the line under study, the same fault case on any circuit will produce the same response due to the symmetry of the tower

TABLE II
SWITCHING CLASSIFICATION ACCURACY

Transmission line	Accuracy (%)
Line under study	100
Line 23-32	100
Line 23-24	100
Line 23-22	100
Circuit (1) line 25-27	48
Circuit (2) line 25-27	52

TABLE III
LIGHTNING CLASSIFICATION ACCURACY

Transmission line	Accuracy (%)
Line under study	100
Line 23-32	100
Line 23-24	100
Line 23-22	100
Circuit (1) line 25-27	49
Circuit (2) line 25-27	51

TABLE IV
TRANSIENT EVENT TYPE CLASSIFICATION ACCURACY

Event type	Accuracy (%)
Faults	100
Lightning	97.2
Switching	99.8

TABLE V
TESTING ANN IN TABLE I WITH NOISY AND DENOISED DATA
T IN CASE OF FAULTS

SNR	Transmission line	Classification accuracy	
		Noisy data (%)	Denoised data (%)
60	Line under study	99	99.3
	Line 23-32	98.2	98.7
	Line 23-24	98.4	99.1
	Line 23-22	98	98.6
	Circuit (1) line 25-27	48.1	48.9
	Circuit (2) line 25-27	49	49.4
40	Line under study	96	98.2
	Line 23-32	95.2	97.2
	Line 23-24	94	98
	Line 23-22	93.1	97.1
	Circuit (1) line 25-27	46	47
	Circuit (2) line 25-27	47	48.3
30	Line under study	90	97.3
	Line 23-32	89	96.5
	Line 23-24	89	96.8
	Line 23-22	88.7	97
	Circuit (1) line 25-27	42	46.5
	Circuit (2) line 25-27	44	47.9

25-27. It should be pointed out that the faults on one of the circuits of line 25-27 are confused for faults on the other circuit and vice versa but not with other lines. The percentage accuracy does not change if the currents at terminal 25 are used instead. Training with levels 1, 2, 3, and 4 gives the same classification accuracy but the difference is mainly in the ANN size needed. For reproducibility purposes, an ANN of size 40 in the hidden layer is needed to achieve the results presented in Table I for level 3. Lower levels (levels corresponding to higher frequencies) needed less neurons but the feature vector becomes so long (corresponding to more detailed time resolution). This in turn demands for a high super computer, which has been undertaken for level 1 (frequency band from 100 to 50 kHz) and level 2 (50 to 25 kHz). Level 3 (25–12.5 kHz) and level 4 (12.5–6.25 kHz) can be done on a PC such as the one the author used with Core i7, 3.2 GHz speed, 4 cores, and 8 GB of RAM. If the double circuit tower 25-27 is considered to be one line, a 100% correct classification is achieved for faults.

B. Line Identification in Case of Switching

This section shows the results of classification if all transients are for line switching. Creating the switching events as described in Section V-B2 and using either end currents, the accuracy presented in Table II is achieved. Training with any level from 1 to 4 gives the same classification result, but the only difference is being the size of the hidden layer. Level 3 required 30 neurons in the hidden layer with lower levels requiring less neurons and higher levels requiring more neurons. Once again switching events on one of the circuits of line 25-27 are confused for the other circuit and vice versa.

C. Line Identification in Case of Lightning Strikes

If all transient events are lightning only as described in Section V-B3, then using any end data gives the results presented in Table III irrespective of the terminal used or the level used (levels 1–4). Level 3 required 30 neurons in the hidden layer with lower levels requiring less neurons and higher levels requiring more neurons. One has to note that in Tables II and III, the algorithm cannot still distinguish between the events on any of the circuits of line 25-27 due to the symmetry with respect to the line being studied.

D. Transient Event Type Classification

Finally, the algorithm performance is shown when it comes to classification between different transients types. The types of studies in this paper are faults, line switching, and lightning. ANN is trained with all transient cases given in Sections V-B1–V-B3. The output is shown in Table IV. Once again levels 1–4 give the same classification accuracy whether bus 23 or 25 is used. The ANN used for Table IV has a size of 30 when trained with level 3 currents. It should be pointed out that in Table IV, most lightning strikes with amplitude 5000 A on line 23-22 have been misclassified for faults. A 5000 A lightning strike is unlikely in real-world lightning strike cases. If we remove all 5000 A strikes from training, a 100% accurate classification is achieved for event-type classification.

TABLE VI
TESTING ANN IN TABLE II WITH NOISY AND DENOISED DATA IN CASE OF SWITCHING

SNR	Transmission line	Classification accuracy	
		Noisy data (%)	Denoised data (%)
60	Line under study	98	99
	Line 23-32	97	98
	Line 23-24	97.1	98
	Line 23-22	97.5	98.5
	Circuit (1) line 25-27	47.5	49.2
40	Circuit (2) line 25-27	48.8	49.5
	Line under study	95	98
	Line 23-32	94.6	97.5
	Line 23-24	93.5	97.2
	Line 23-22	92.5	98
30	Circuit (1) line 25-27	45.5	48.7
	Circuit (2) line 25-27	46.1	49
	Line under study	87	96.5
	Line 23-32	86	95.2
	Line 23-24	86.2	95.6
	Line 23-22	87.4	96
	Circuit (1) line 25-27	42.2	47.6
	Circuit (2) line 25-27	43.5	48.5

TABLE VII
TESTING ANN IN TABLE III WITH NOISY AND DENOISED DATA IN CASE OF LIGHTNING

SNR	Transmission line	Classification accuracy	
		Noisy data (%)	Denoised data (%)
60	Line under study	100	100
	Line 23-32	99	99.3
	Line 23-24	99.1	99.5
	Line 23-22	99.7	99.7
	Circuit (1) line 25-27	49.3	49.3
40	Circuit (2) line 25-27	49.5	49.5
	Line under study	98	99
	Line 23-32	97	97.8
	Line 23-24	97.2	98
	Line 23-22	97.9	98.5
30	Circuit (1) line 25-27	46	47.3
	Circuit (2) line 25-27	47	48.5
	Line under study	95	98
	Line 23-32	94	96.7
	Line 23-24	94.6	96.2
	Line 23-22	94.8	97.6
	Circuit (1) line 25-27	42	46.6
	Circuit (2) line 25-27	41.5	48.1

It is now straightforward to construct a two-layer feedforward network that classifies any event to fault, lightning, or switching in the first layer, then identifies the line causing the event in the second layer. Alternatively, a relay can be programmed to classify the events into fault, switching, or lightning as per Table IV. Once this is done, the same event can be applied to the ANN corresponding to that event type.

IX. ERROR ANALYSIS

Since power systems are subject to harsh environmental conditions, error always exists in the measurements regardless of the bandwidth of the measuring devices used. The author

TABLE VIII
TESTING ANN IN TABLE IV WITH NOISY AND DENOISED DATA IN CASE OF ALL TRANSIENT EVENTS

SNR	Event type	Classification accuracy	
		Noisy data (%)	Denoised data (%)
60	Faults	99	99.2
	Lightning	96	96.8
	Switching	98.1	98.5
40	Faults	97	98.1
	Lightning	95.2	95.7
	Switching	96.5	97
30	Faults	95	97.5
	Lightning	94.4	95
	Switching	94.1	96.5

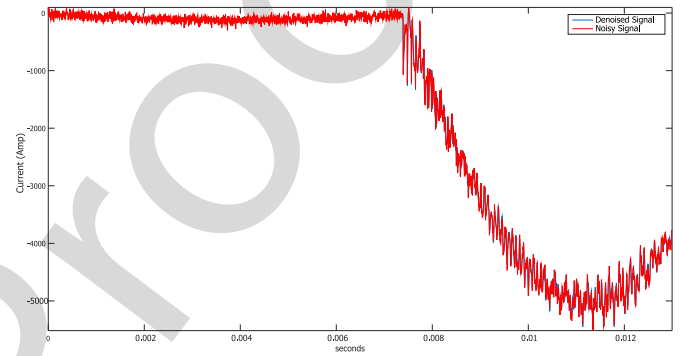


Fig. 5. Mode2 of a fault case with noise and after denoising.

analyzed the performance of the proposed algorithm when noise exists in the current measurements. Noise has been added to the data that have been generated in Section V-B. The noise that has been added is Gaussian white noise at different signal-to-noise ratios. The noisy data were then applied to the ANNs trained in Sections VIII-A–VIII-D to see how noise affects the classification accuracy. As can be seen from Tables V–VIII, classification accuracy decreases as SNR decreases. Even in the best case with high SNR, classification accuracy is not as good as it was in Sections VIII-A–VIII-D. This may be because ANN has bad generalization capability, which could signal that a better classification approach is needed.

To remove noise from data, the author used a wavelet denoising approach. The wavelets toolbox in MATLAB has been used to perform wavelet denoising. The denoised data are used for testing the ANNs trained in Sections VIII-A–VIII-D. The waveform of Mode2 current of a certain fault case with noise (SNR = 40) and after denoising is shown in Fig. 5. The results of this testing is presented in Tables V–VIII. As can be seen from the tables, denoising improves the performance of ANN. This is more noticeable when classification is done using data that have lower SNR. A detailed error analysis is beyond the scope of this paper and will be undertaken along with field validation.

X. COMPARISON TO EXISTING METHODS IN THE LITERATURE

The major drawbacks of existing methods have been summarized in Section I. In this section, the method proposed will

TABLE IX
PNN CLASSIFICATION ACCURACY

Event type	Accuracy (%)
Faults	98
Lightning	96
Switching	100

be compared to the existing methods in the literature that use single end data to reach a decision regarding the fault.

Even though the subject of transmission line relaying based on high-frequency components has and continues to receive attention from academia and the industry, a handful of papers study transients other than faults and their effect on the proposed methods.

In [18], the details coefficients of all phase voltages and currents are fed to ANN for training. However, for the method to perform well, the high-frequency transients imposed on the fundamental voltage have to be measured, which is not feasible if CVTs are used, as the bandwidth of CVTs does not exceed 1 kHz [26]. Additionally, the method ignores all transients types other than faults, which can mislead the scheme most of the times.

In [16], the details coefficients of all phase voltages and currents are still used to train the ANN in addition to using an energy quantity as a trip restraint. However, transients on the line being protected are being taken into consideration while training the ANN. A drawback of this method is the need to perform many offline studies to determine the threshold for the trip restraint. Moreover, the method suffers from the limited bandwidth of the CVTs.

In [20], the energy of the three-phase currents is used to train a probabilistic classifier. Voltages have been abandoned for the purpose of classification due to the aforementioned shortcomings of CVTs. However, the simulated system is small and the methods proposed cannot identify the line causing the transient event. Additionally, the use of energy as a feature requires a detailed study on the proximity of the transient to the measuring CT. The results in [20] are a special case of the results in this paper and correspond to the results in Section VIII-D.

For the purpose of statistical comparison, the author has reproduced the PNN classifier given in [20] and applied it to the system shown in Fig. 2. The training dataset that had been developed in Section V-B is then applied to the PNN classifier. The results of classification are presented in Table IX. As can be seen from both Tables IV and IX, not all lightning strikes are fully recognized. This occurs because some lightning strikes are confused for faults. The ones that are confused for faults are the ones that have 5 kA amplitude, which is consistent with the observation in Section VIII-D. Moreover, some faults are not well recognized using the PNN due to high resistance faults, which is consistent with the observation in [20]. The main advantage of the method given in this paper is that fewer features are needed to obtain satisfactory classification accuracy.

In all the methods mentioned above, at least one-quarter of a cycle is needed to reach a secure decision. On the other hand, the approach proposed in this paper uses only one-half of that

to reach a secure decision about whether the transient belongs to the line being protected and whether the transient is a fault. Additionally, actual physical tower configurations have not been taken into consideration in the literature.

XI. CONCLUSION AND FUTURE RESEARCH

This paper presented an argument that high-frequency signals can be used for high-speed power system fault detection via transient classification identifying the line causing the transient event.

The contributions of the paper are as follows.

- 1) It has been shown that currents alone can be used for transient signature of the event.
- 2) Only two modes are necessary for classification.
- 3) Only one-eighth of a cycle of postevent data is necessary for classification.

Although the results in this paper have been shown only to a specific tower, the author has tried different tower configurations and confirmed that the algorithm works for all tower configurations considered. It should be clear that the current method fails if a lightning strike evolves to a fault. This drawback has been addressed separately in a different publication [42]. Additionally, detection of faults on mutually coupled lines has not been addressed in this paper and will be treated in a different publication.

One issue that has not been investigated in this paper is insulation breakdown. The author has not assumed any failure of insulation in the simulations. The cutting of the signal associated with insulation breakdown can potentially mislead the scheme proposed. Further investigations are needed to completely quantify its effect on the proposed relaying scheme. It should be noted, however, that lightning strikes are being chopped by the surge arrester but still being correctly recognized.

Field validation is being performed and results will be shared once all investigations are done.

ACKNOWLEDGMENT

This paper could not have seen the light without the resources at Texas A&M Supercomputing Facility. The author would like to thank all at the facility for their help and support. The author also would like to express his sincere thanks for J. Woodworth of ArresterWorks for his guidance regarding arrester selection. Finally, the author would like to thank G. Gurralla from the Indian Institute of Science for his constructive review of the manuscript.

REFERENCES

- [1] A. Phadke and J. Thorp, *Computer Relaying for Power Systems*. Hoboken, NJ, USA: Wiley, 2009.
- [2] L. Kojovic and M. Kezunovic, "A new method for the CCVT performance analysis using field measurements, signal processing and EMTP modeling," *IEEE Trans. Power Del.*, vol. 9, no. 4, pp. 1907–1915, Oct. 1994.
- [3] J. Thorp, A. Phadke, S. Horowitz, and J. Beehler, "Limits to impedance relaying," *IEEE Trans. Power App. Syst.*, vol. PAS-98, no. 1, pp. 246–260, Jan. 1979.






- [4] E. O. Schweitzer, A. Guzmán, M. V. Mynam, V. Skendzic, B. Kasztenny, and S. Marx, "Locating faults by the traveling waves they launch," in *Proc. 2014 67th Annu. Conf. Protective Relay Eng.*, 2014, pp. 95–110.
- [5] S. Santoso, E. J. Powers, W. M. Grady, and P. Hofmann, "Power quality assessment via wavelet transform analysis," *IEEE Trans. Power Del.*, vol. 11, no. 2, pp. 924–930, Apr. 1996.
- [6] R. G. Lyons, *Understanding Digital Signal Processing*. Reading, MA, USA: Addison-Wesley, 1997.
- [7] A. V. Oppenheim *et al.*, *Discrete-time Signal Processing*, vol. 2. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989.
- [8] S. G. Mallat, "A theory for multiresolution signal decomposition - The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.
- [9] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, Sep. 1990.
- [10] A. Boggess and F. J. Narcowich, *A First Course in Wavelets With Fourier Analysis*. Hoboken, NJ, USA: Wiley, 2009.
- [11] W. A. Wilkinson and M. Cox, "Discrete wavelet analysis of power system transients," *IEEE Trans. Power Syst.*, vol. 11, no. 4, pp. 2038–2044, Nov. 1996.
- [12] F. H. Magnago and A. Abur, "Fault location using wavelets," *IEEE Trans. Power Del.*, vol. 13, no. 4, pp. 1475–1480, Oct. 1998.
- [13] T. Dalstein and B. Kulicke, "Neural-network approach to fault classification for high-speed protective relaying," *IEEE Trans. Power Del.*, vol. 10, no. 2, pp. 1002–1011, Apr. 1995.
- [14] M. Kezunovic, I. Rikalo, and D. J. Sobajic, "High-speed fault-detection and classification with neural nets," *Elect. Power Syst. Res.*, vol. 34, no. 2, pp. 109–116, 1995.
- [15] M. Oleskovicz, D. Coury, and R. Aggarwal, "A complete scheme for fault detection, classification and location in transmission lines using neural networks," in *Proc. 2001 7th Int. Conf. Develop. Power Syst. Protection*, 2001, pp. 335–338.
- [16] K. M. Silva, B. A. Souza, and N. S. D. Brito, "Fault detection and classification in transmission lines-based on wavelet transform and ann," *IEEE Trans. Power Del.*, vol. 21, no. 4, pp. 2058–2063, Oct. 2006.
- [17] P. L. L. Mao and R. K. Aggarwal, "A novel approach to the classification of the transient phenomena in power transformers using combined wavelet transform and neural network," *IEEE Trans. Power Del.*, vol. 16, no. 4, pp. 654–660, Oct. 2001.
- [18] F. Martin and J. A. Aguado, "Wavelet-based ANN approach for transmission line protection," *IEEE Trans. Power Del.*, vol. 18, no. 4, pp. 1572–1574, Oct. 2003.
- [19] G. Zwe-Lee, "Wavelet-based neural network for power disturbance recognition and classification," *IEEE Trans. Power Del.*, vol. 19, no. 4, pp. 1560–1568, Oct. 2004.
- [20] N. Perera and A. Rajapakse, "Recognition of fault transients using a probabilistic neural-network classifier," *IEEE Trans. Power Del.*, vol. 26, no. 1, pp. 410–419, Jan. 2011.
- [21] J. M. Zurada, *Introduction to Artificial Neural Systems*. St. Paul, MN, USA: West, 1992.
- [22] J. J. Grainger and W. D. Stevenson, *Power System Analysis*, vol. 31. New York, NY, USA: McGraw-Hill, 1994.
- [23] A. Abdullah, "Towards a new paradigm for ultrafast transmission line relaying," in *Proc. 2016 IEEE Power Energy Conf. Illinois*, 2016, pp. 1–8.
- [24] D. Hedman, "Propagation on overhead transmission lines I—Theory of modal analysis," *IEEE Trans. Power App. Syst.*, vol. PAS-84, no. 3, pp. 200–205, Mar. 1965.
- [25] D. Hedman, "Propagation on overhead transmission lines II—Earth-conduction effects and practical results," *IEEE Trans. Power App. Syst.*, vol. PAS-84, no. 3, pp. 205–211, Mar. 1965.
- [26] D. Hou and J. Roberts, "Capacitive voltage transformer: Transient overreach concerns and solutions for distance relaying," in *Proc. 1996 Can. Conf. Elect. Comput. Eng.*, 1996, pp. 119–125.
- [27] H. Lee, "Development of an accurate transmission line fault locator using the global positioning system," in *Proc. 1994 Nat. Tech. Meeting Inst. Navig.*, 1994, pp. 111–116.
- [28] D. A. Douglass, "Current transformer accuracy with asymmetric and high-frequency fault currents," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 3, pp. 1006–1012, Mar. 1981.
- [29] M. Redfern, S. Terry, F. Robinson, and Z. Bo, "A laboratory investigation into the use of MV current transformers for transient based protection," in *Proc. 2003 Int. Conf. Power Syst. Transients*, New Orleans, LA, 2003, pp. 1–5.
- [30] J. C. G. de Siqueira, B. D. Bonatto, J. R. Marti, J. A. Hollman, and H. W. Dommel, "A discussion about optimum time step size and maximum simulation time in EMTP-based programs," *Int. J. Elect. Power Energy Syst.*, vol. 72, pp. 24–32, 2015.
- [31] Washington State University, "118 bus power flow test case." [Online]. Available: https://www.ee.washington.edu/research/pstca/pf118/pg_tca118bus.htm
- [32] P. Demetriou, M. Asgrou, J. Quiros-Tortos, and E. Kyriakides, "Dynamic IEEE test system transient analysis," *IEEE Syst. J.*, 2017, to be published.
- [33] S. D. Cho, "Parameter estimation for transformer modeling," Ph.D. dissertation, Michigan Technol. Univ., Houghton, MI, USA, 2002.
- [34] N. Perera, A. Rajapakse, and T. Buchholzer, "Isolation of faults in distribution networks with distributed generators," *IEEE Trans. Power Del.*, vol. 23, no. 4, pp. 2347–2355, Oct. 2008.
- [35] A. M. Gashimov, A. Kabayeva, and A. Nayir, "Transmission line transposition," in *Proc. 2009 Int. Conf. Elect. Electron. Eng.*, 2009, pp. 364–367.
- [36] ALSTOM, *Network Protection & Automation Guide*. Stafford, U.K.: ALSTOM Grid, 2002.
- [37] J. Woodworth, "Arrester selection guide." [Online]. Available: http://www.arresterworks.com/arresterfacts/pdf_files/Arrester_Characteristics_for_ATPDraw_Users.xls
- [38] A. Greenwood, *Electrical Transients in Power Systems*, 2nd ed. New York, NY, USA: Wiley, 1991.
- [39] A. Abdullah, "Atpmat: An open source toolbox for systematic creation of EMTP cases in ATP using MATLAB," in *Proc. 2015 North Amer. Power Symp.*, Oct. 2015, pp. 1–6.
- [40] V. V. Terzija and H. F. Koglin, "On the modeling of long arc in still air and arc resistance calculation," *IEEE Trans. Power Del.*, vol. 19, no. 3, pp. 1012–1017, Jul. 2004.
- [41] A. Ametani, N. Nagaoka, Y. Baba, and T. Ohno, *Power System Transients: Theory and Applications*. Boca Raton, FL, USA: CRC Press, 2013.
- [42] A. Abdullah, "A wavelet entropy approach for detecting lightning faults on transmission lines," in *Proc. 2016 IEEE Power Eng. Soc. Transm. Distrib. Conf. Expo.*, May 2016, pp. 1–5.



Ahmad Abdullah (M'xx) received the B.Sc. and M.Sc. degrees from the Department of Electrical Power and Machines, Faculty of Engineering, Cairo University, Giza, Egypt, in 2009 and 2012, respectively, and the Ph.D. degree from the Electrical and Computer Engineering Department, Texas A&M University, College Station, TX, USA, in 2017.

Since 2015, he has been with Electric Power Engineers, Inc., Austin, TX, USA, as a Power Systems Engineer responsible for conducting power system design studies.

QUERIES

- Q1. Author: Please provide the expansion of acronyms “CVT,” “EMTP,” “ATP,” and “PNN,” at first occurrence. 
- Q2. Author: Please provide the expanded form of “LANs” 
- Q3. Author: Please complete and update Ref. [32]. 
- Q4. Author: Please provide the areas of study in which Ahmad Abdullah received the B.Sc., M.Sc., and Ph.D. degree 
- Q5. Author: Please provide the year in which Ahmad Abdullah became a member of the IEEE. 

992

993

994

995

996

997

Ultrafast Transmission Line Fault Detection Using a DWT-Based ANN

Ahmad Abdullah, *Member, IEEE*

Abstract—Digital impedance protection of transmission lines suffers from known shortcomings not only as a principle but also as an application. This necessitates developing a new relaying principle that overcomes those shortcomings. Such a principle is offered in this paper and is currently being validated using field data. The principle is a new application of wavelet-based artificial neural networks (ANNs). The application uses high-frequency content of a subset of local currents of one end of a protected line to classify transients on the line protected and its adjacent lines. The scheme can classify transients—including faults—occurring on a protected line, categorize transients on adjacent lines, and pinpoint the line causing the transient event. It is shown that the feature vector of the event can be determined from a subset of local currents without using any voltages altogether. The subset of local currents consists of the two aerial modes of the local current. Modal transformation is used to transform phase currents to modal quantities. Discrete wavelet transform (DWT) is used to extract high-frequency components of the two aerial modal currents. A feature vector is built using the wavelets details coefficients of one level of the aerial modes and is used to train an ANN. Results show that the classes corresponding to each transient event type on the protected line and its adjacent lines are almost linearly separable from each other. Results demonstrate that very accurate classification using one-eighth of a cycle of postevent data is possible.

Index Terms—Artificial neural networks (ANNs), current transformers, line switching, lightning strike, modal analysis, power system faults, transmission line relaying, wavelets transform.

I. INTRODUCTION

PROTECTION of a transmission line involves installing relays at both ends of the line that constantly monitor voltages and currents and act when a fault occurs on a line. Traditional protection uses phasor estimation to estimate the fundamental component of voltages and currents and takes a decision when certain criteria corresponding to certain fault conditions are met. Distance protection is the most widely used method for transmission line protection. Modern numerical distance relays make use

of low-pass and anti-aliasing filters—to reject high-frequency content and apply Nyquist sampling theorem—and special digital signal processing (DSP) hardware to perform sophisticated math functions [1]. CVTs are generally used to measure the voltages making them available to the relay. It is known that due to the interaction between the capacitive voltage divider and the transformer inductance, oscillations are imposed on the fundamental frequency measured [2]. This puts more stringent requirement on the filters used in the relay. Additionally, distance relay can only protect up to 85% of the line instantaneously [3]. This necessitates the use of a communication link between the relays at the two ends of the line to achieve fast tripping from both ends. With relays connected to substation LANs, a communication link between relays is a cyber-threat. With all limitations mentioned above, a new relaying principle is needed. This principle has to be using currents only, fast, and finally does not need any communication link for its operation. Such a principle is provided and theoretically tested in this paper.

Except for traveling wave fault location techniques that are still not very popular with relay manufacturers [4], little is done with the high-frequency components of the transients even though it has been known that high-frequency components of faults and transient events contain rich information [5]. This is mainly because the traditional mathematical method at hand was fast Fourier transform (FFT) [6]. The FFT requires the signal be stationary in the wide sense for the calculation of coefficients to be accurate, i.e., the signal cannot have any temporal variations for the calculation of coefficients to be correct and accurate. Short window Fourier transform [7] can solve some of the problems with the FFT but it introduces other issues. However, most signals encountered in power systems are not stationary but have their characteristics change with time.

Multiresolution analysis (MRA) is a signal processing tool that has been introduced in the nineties [8] to solve some of the problems inherent in Fourier transform analysis methods. Wavelets [9] are usually used along with MRA to solve this very exact problem. Wavelets have a strong localization property that enables studying changes not only in time but also in frequency feasible. It is known from the uncertainty principle [10] that if the signal spans a small portion of the time domain, then its Fourier transform will span a large portion of frequency domain. The discrete wavelet transform (DWT) on the other hand does not suffer from such a limitation. On the contrary, the signal is approximated at various levels—frequency bands—and changes in time manifest themselves at the coefficients of the wavelets only around the time in which the event occurred. Such property

Manuscript received May 5, 2016; revised October 25, 2016 and May 13, 2017; accepted November 9, 2017. Paper 2016-PSPC-0365.R2, presented at the 2016 IEEE Power and Energy Conference at Illinois, Champaign, IL, USA, Feb. 19–20, and approved for publication in the IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS by the Power System Protection Committee of the IEEE Industry Applications Society.

The author is with the Department of Electrical Power and Machines, Faculty of Engineering, Cairo University, Giza 12613, Egypt, and also with Electric Power Engineers, Inc., Austin, TX 78738 USA (e-mail: ahmad.abdullah@cu.edu.eg).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIA.2017.2774202

makes it ideal to detect disturbances and study power system transients [11].

With the advent of wavelets, the power system community has seen a surge in application of wavelets-based methods for various power system problems notably in the area of power system protection—mainly in transient-based protection schemes, transients' classification, and fault location. The localization property of wavelets makes it very convenient in locating faults [12].

The use of artificial neural network (ANN) for fault classification and detection has been given in [13]–[15] where voltage and current samples are fed directly to the ANN for fault detection and classification on the protected line. In [14] and [13], samples of voltages and currents are used as a feature vector to train ANN. With the advent of wavelets, special wavelets transforms are applied to voltage and current waveforms before they are fed to the ANN for training. In [16] and [17], the DWT is applied to voltages and currents but instead of feeding the details coefficients to the ANN, entropy (energy) of the signal is captured and fed instead to the ANN to make fault-type classification between faults on the same transmission line. In [18] and [19], the DWT is applied to voltage and current signals at a relay location resulting in series of details coefficients that are fed to the ANN for fault detection and type classification. In [20], the energy of certain current levels and approximations is used to train a probabilistic classifier, and using this energy feature, a decision is made whether a transient signal is due to a fault or nonfault condition on a line. A detailed look at the literature reveals that very few attempts have been carried out to harness the power of the wavelets details coefficients alone and understand their underlying structure. Moreover, existing classification techniques as the ones in [19] do not take transients on adjacent lines into account, which could mislead these schemes. Additionally, symmetrical line configuration has been universally assumed, which is a configuration nearly nonexistent in real power systems.

In this paper, the wavelets details coefficients are directly used to train the ANN without using any entropy-based method. This paper proposes a novel application for wavelet-based ANN. It is shown that using the details coefficients of a subset local currents only, distinction can be made between fault and nonfault conditions on a protected transmission line without using any voltage signals at all. The proposed algorithm can also distinguish between faults on the protected line and faults on adjacent lines. The algorithm provided can not only tell the difference between forward and reverse faults but also can determine which adjacent line is faulted. It is also shown that the same applies for lightning strikes and line switching cases, i.e., the algorithm can tell which line is causing which transient event. Just as the FFT produces coefficients that correspond to certain frequencies, wavelets details also give information regarding the oscillatory components of the signal localizing them in time. In this regard, the DWT is used to extract useful oscillatory information about the signal. Oscillatory information is manifested in the wavelets details coefficients at various levels. It is shown that any transient event on a specific transmission line causes currents to oscillate in a unique way and these oscillations can be detected in the wavelets details coefficients themselves at

various levels. The use of these wavelets details coefficients shows the possibility of building ultrahigh speed relays.

The paper is organized as follows. A quick overview of both the DWT and ANN is offered in Section II. The rationale behind the proposed principle is presented in Section III. The solution methodology is given in Section IV. The simulation platform, which consists of a description of EMTP model and the creation of transient cases for ANN training, is described in Section V. The structure of the feature vector is given in Section VI. Signal analysis is provided in Section VII. Simulation results are presented in Section VIII. Error analysis is provided in Section IX. Comparison between the method proposed in this paper and existing methods in the literature is presented in Section X. Conclusion and future research are provided in Section XI.

II. BACKGROUND

In this section, a brief overview of both the DWT and ANN as used in this paper is provided.

A. Discrete Wavelet Transform

In this paper, the dyadic wavelet transform is used [10]. The transform takes the signal and applies low- and high-pass filters to it. The transform converts the original signal into independent signals each of which spanning a certain frequency band. The frequency bands are determined by the number of levels the original signal is sought to be analyzed and each level corresponds to a unique frequency band. The word independent means that one level cannot be derived from another, i.e., level 2 cannot be derived from level 1. The frequency band of each level depends on the sampling frequency and Nyquist theory. The highest frequency that can be seen in the decomposed signal will be at most equal to half of the sampling frequency. Given a certain sampling frequency, the dyadic transform will first apply high- and low-pass filters to the signal resulting in two signals. The signal corresponding to the output of the high-pass filter is called level 1 and the other signal is called approximation 1. Both of those signals are independent. One can stop at this step or can apply high- and low-pass filters to the first approximation to get level 2 and approximation 2. Level 2 will span a frequency band corresponding to the upper-half of the frequency band of approximation 1, while approximation 2 will occupy the lower frequency band. Continuing this manner, by applying successive low- and high-pass filters, one obtains a set of levels—also called details—and one last approximation. In theory, the last approximation should correspond to a pure sine wave assuming that the high-frequency components imposed on the power frequency have frequencies that are not included in the frequency band of the last approximation. Taking a numerical example, if one uses 200 kHz sampling frequency, then the resulting DWT decomposition is shown in Fig. 1. The details coefficients of the DWT transform are provided in (1) [10], where $\varphi(t)$ is the mother wavelet used, $f(t)$ is the signal to be analyzed, parameter “ a ” causes scaling (which determines the level) and “ b ” causes shift in time. In practice, it is not needed to apply the transform all over the infinite time line. Since wavelets have a strong localization property, it is only needed to apply the transform to

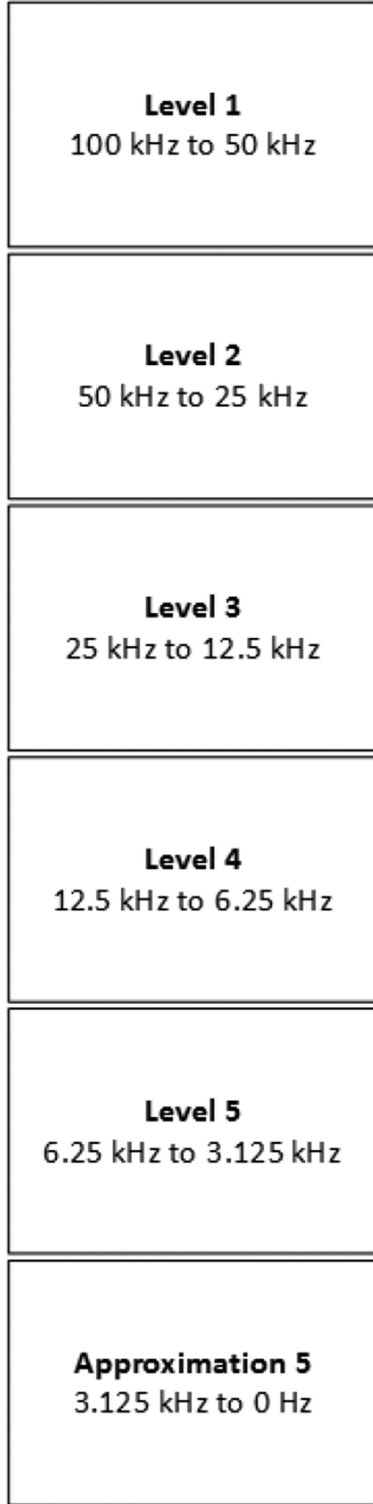


Fig. 1. DWT at 100 kHz sampling rate for five levels.

B. Classification With Artificial Neural Networks

ANN is a very well-established tool for classification. The classification done in this paper is not probabilistic but rather deterministic. The inputs to ANN are vectors in the n -dimensional space, where n is the size of each input vector that needs to be mapped to another output vector space. The size of the output vector space is determined by the number of output classes one wants to map the input vectors to. In simple terms, this classification can be shown to be carried out by ANN [21]. All ANNs in this paper consist of three layers: an input layer, a hidden layer, and an output layer. Each layer consists of a certain number of neurons. The connections between those neurons are called synapses. The input layer consists of junctions that represent the input. The number of junctions must equal to " n ," which is the dimension of the input vector. The hidden layer can consist of any number of neurons. The classification accuracy is greatly affected by the number of neurons in the hidden layer. The output layer consists of neurons each of which is activated by a function. The inputs to this function come from the hidden layer. Classes or more specifically output classes are the patterns one wants to map the input to. The main idea behind ANN classification is that if the inputs belong to linearly separable regions of the n -dimensional space, then using ANN, one can draw hyperplanes in that space so that the space between those hyperplanes contain only one region and each region is then mapped to one of the output classes using the activation function in the output neuron. The weights or more specifically synaptic weights of the synapses are adjusted during the training phase such that the vector space is divided into regions each of which corresponds to a certain class. Any output neuron receives the input vector through synaptic weights. The resultant is then applied to the function of the neuron that has the form $g(x) = 0$, which is an equation of the hyperplanes in the n -dimensional space. The function $g(x) = 0$ is realized in the ANN by the neurons and the synaptic weights. The output of the neuron is activated only if $g(x)$ is positive for the corresponding input vector. In this paper, the author uses the widely known backward propagation algorithm for the calculation of weights. Knowledge is stored in the ANN through those weights. Only supervised learning is done in this paper. The algorithm of the steepest gradient is used throughout the paper [21].

III. RATIONALE BEHIND THE PROPOSED METHOD

When a fault or any transient event occurs, it launches a traveling wave as well as high-frequency oscillatory components. Traveling waves can be easily quantified as they arrive at the line terminals but the high-frequency oscillatory components cannot. The reason for the oscillatory components not being easily quantified before in the literature will be explained below.

At this point, the author needs to introduce new terminology. Speaking of a certain transmission line, the author calls a fault on that line a fault case. A fault case has its parameters. Those parameters are incipient angle, fault resistance, fault type, and fault location. Thus, a certain fault case on a specific transmission line causes voltages and currents to oscillate in a manner

the time period under study

$$Wf(a, b) = \frac{1}{\sqrt{a}} \times \int_{-\infty}^{\infty} f(t) \times \varphi\left(\frac{t-b}{a}\right) dt. \quad (1)$$

in accordance with the parameters of that fault case. The formal solution of the currents and voltages of a single phase line is given by (called telegrapher equations) [22]

$$\frac{\partial V}{\partial x} = R \times I + L \times \frac{\partial I}{\partial t} \quad (2)$$

$$\frac{\partial I}{\partial x} = G \times V + C \times \frac{\partial V}{\partial t} \quad (3)$$

where I and V are the current and the voltage anywhere on the line; R and G are the resistance and conductance per unit length of the line; L and C are inductance and capacitance of the line per unit length; x is the distance from a zero reference frame generally taken to be at either end of the line.

A general closed-form solution of (2) and (3) is hard and generally impossible as it depends on the boundary and initial conditions of the case involved. Perhaps the most straightforward solution method is to obtain the solution in terms of an infinite time series expansion. Using infinite time series expansion would mystify the solution, because the author's intention is to analyze the frequency content of the signal as this content changes over time. Application of Laplace and Fourier transforms to those equations produces integrals that are yet to be solved formally. Moreover, when a three-phase line is studied, the above-mentioned equations become six equations (two equations for each phase: 1) a voltage; and 2) current equation). Thus, decoupling them is hard undertaking if not impossible in most cases. If we have mutual coupling between two lines sharing the same tower, then we have three more equations making them a total of nine equations. Bearing in mind that those nine equations are for one line only, those equations then need to be solved simultaneously along with all other equations in the system. For today's power systems, which consist of thousands of lines, solving all equations analytically in a closed form is behind human ingenuity.

The main aim is to detect the fault once it is initiated. It is very apparent that the high-frequency fault-generated transients are different for each fault case on a specific line. This is because each fault case on a specific line has its set of equations as given in (2) and (3) with unique boundary (lines connected to it) and initial conditions. Moreover, the same fault case on different lines will cause different fault-generated transients because each case has its own boundary and initial conditions, which is different for all neighboring lines. For this reason, no attempt is made to solve the equations formally but solve them numerically by running EMTP simulations, extracting the high-frequency content and then analyzing them.

Since the author argues that those frequencies—which change over time—are different for each line, the author only needs to see the aggregate of all those frequency components for each line. If intuition is true, then the aggregate of those fault-generated high frequencies will be different for each line. What remains is to investigate how the feature space can be defined such that this difference can be attained and this is done in Section VI. If it is possible to map those frequency components to an n -dimensional space, then faults corresponding to a certain line should be separable from faults on adjacent lines. The

author takes this idea further and investigates whether the aggregate of all faults from each line is actually linearly separable [21] from each other. This is the main argument behind the use of ANN as a pattern recognition approach.

For any fault occurring on a transmission line, the transient-generated frequency components depend on the initial and the boundary conditions (the lines connected to the faulted line). Since by definition all adjacent lines are different, the frequency-generated transients for faults on each transmission line should be unique for each line. The question is whether those frequency components can be combined in a certain way to form a pattern to be recognized.

The same argument not only applies to faults but also applies to switching events and lightning strikes on transmission lines. The transient event types that are studied in this paper are faults, lightning strikes, and line switching. Evidently, the frequency content of faults—on all lines—should be different from switching, which in turn is different from lightning because the equations governing the solutions are always different as different transient event types have different boundary and initial conditions. Those event types can be first recognized and then the event can be traced back to the line originating it as given in Section VIII-D. This basically means that there is a mapping that can divide a certain subspace to linearly separable regions that correspond to those event types. If one zooms in that event subspace, one can see which line is causing that event type.

IV. SOLUTION METHODOLOGY

In this paper, it is argued and shown that the oscillatory information present in the transient signal captured at a single relay location caused by sudden network topology changes contains sufficient information for classification not only between transients on the same line but also between transients on adjacent lines [23]. Any change of configuration on the line causes a traveling wave to be generated traveling from the point of change toward the ends of the line. In the simplest case, this wave will be just a pulse—a step—but in reality will be accompanied by high-frequency oscillating components. Fourier analysis is not suitable for analyzing such waveforms because those oscillations will be distorted and attenuated as they arrive at the line terminals. However, applying the DWT will enable us to see both the spectral and temporal variations of those high-frequency oscillations.

Historical treatment of traveling waves—as used in the ladder diagram, for example—treats them as if they were pulses traveling down the line with no regard to the oscillatory components they carry. A certain mathematical entity—which is the feature vector to be defined in Section VI—is derived from the oscillatory components and called the transient signature of the event. The transients that are studied in this paper are lightning, line switching, and faults. At a certain terminal of the line, which is typically a relay location, it is argued and shown that those signatures are unique to the event that originated them and to the line that initiated those events. This means that us-

ing this oscillatory information of the event, the line causing it can be determined. Those signatures are a function of the line parameters, the network topology and the instance of the event, and, of course, the type of the event. In essence, the signature of the fault occurring on a certain line will be different from the signature of the fault on adjacent lines. Also, line switching will cause the currents to oscillate in a manner that is different from faults and lightning striking the same line and those oscillatory components are different for different lines as well.

It is argued that the aggregate of all transient signatures caused by a certain transient event—switching or faults, for example—originating from a certain line occupies a specific subspace in the n -dimensional space. This subspace is almost linearly separable [21] from all other subspaces generated by other transient events by adjacent lines. The n -dimensional space is the space spanned by all n -dimensional feature vectors used to train the ANN. A relay that is programmed to use these wavelets details not only can detect and classify transients on a protected line but also can detect faults and classify transients on adjacent lines. ANN classification is used to show that this is indeed the case. Static feedforward ANN is a tool used to map a vector from the n -dimensional space to another m -dimensional space. In the case at hand, the feature vector is mapped, which is n -dimensional space to another output space. These output spaces will be described in Section VIII, but for now if this mapping is successful, then the original n -dimensional subspaces are linearly separable as explained in [21].

Information is extracted from the high-frequency components using the DWT at various levels. Before the DWT is applied, phase currents are decoupled from each other using a modal analysis [24], [25]. Modal transformation will decouple current waveforms, thus eliminating mutual coupling and untransposed line effects. The modal matrix will be explained in Section VI. It is known that CVT introduces transients to the measured voltage signal. It is shown in [26] that a typical CVT with active ferroresonance suppression circuit is a low-pass filter with a cut-off frequency of around 1 kHz, which drops to almost 200 Hz when passive ferroresonance suppression circuit is employed. On the other hand, classification with currents is preferable because the cut-off frequency of current transformers is much larger than the potential transformer or CVT. The useful passband of current transformers is typically 100 kHz [27], and in some cases, can go up to 400 kHz [28] or 500 kHz [29]. Since a conservative stance is chosen, the sampling rate is selected such that the maximum frequency available in the signal is 100 kHz to attain the 100 kHz passband of current transformers. The time step used in the EMTP simulations is 1 μ s, which should theoretically give a maximum frequency of 500 kHz in sampled output according to Nyquist theory. However, it is given in [30] that the maximum frequency in the EMTP simulations is only one-fifth of that, so the maximum frequency is 100 kHz.

After decoupling phase quantities, the DWT is applied to the currents modes to convert the signals to a series of coefficients that will be used for training of the neural networks. The event is detected once a change of level 3—or any level depending on which level is used—coefficients is detected. This detection

is simply because the pre-event data are assumed steady state so no high-frequency components exist in the pre-event data. Once the event is detected, a window of one-eighth of a cycle of postevent information is used for neural network training. Simulations show that at least one-eighth of a cycle of postevent data has to be used for correct classification. Numerous trials have shown that at least one-eighth of a cycle of postevent information has to be present for correct classification or classification fails completely.

The outcome of the DWT is a series of coefficients for each mode and for each level. The coefficients of one level of the two aerial modes of currents are stacked on top of each other to build one feature vector that will be used to train the network. Building the training vector will be explained in Section VI. A neural network of an appropriate size is then selected for classification. The choice of the size of network is reached at by trial and error and no universal size has been applied to all classification problems. That is, the size of the network is changed till the required classification accuracy is achieved. Once the size of ANN is selected, the ANN is then trained using various scenarios for transients on lines. It is shown that using only two modes of local current information at one end of the line—corresponding to a relay, distinction can be made between various transient events on different lines. It is argued and shown that the feature vector built has sufficient information to distinguish between a fault and nonfault condition on a protected transmission line. It is also shown that the adjacent line causing certain transient events can be determined. This means using local information we can see what is happening on adjacent lines.

In summary, the procedure is as follows.

- 1) Decouple currents using modal analysis.
- 2) Apply the DWT to the aerial modes using one-eighth of a cycle of postevent currents.
- 3) Stack the series of coefficients of the two aerial modes of a certain level on top of each other to create a vector used to train ANN.
- 4) A neural network of appropriate size is used for training. Training is done using transient scenarios. These scenarios include faults, line switching, and lightning.

V. SIMULATION PLATFORM

In this section, the ATP/EMTP model is presented in Section V-A, while in Section V-B, creation of the transient cases for the training, validation, and testing of ANN are fully described.

A. EMTP Model

The area under study is shown in Fig. 2. The IEEE 118 bus is used [31] as a test system. All data are taken from [31] except for machine data and the lines in the area under study. The dynamic data of IEEE 118 bus system are taken from [32]. The line under study is the line connecting buses 23 and 25 (line 23-25). The selection of this is arbitrary but this specific line was attractive to the simulations for the reasons that follow. Line 23-25 is bordered by three lines connected to bus 23, namely: line 23-22, line 23-32, and line 23-24. Line 23-25 is also connected

have been created, i.e., AG, BG, CG, AB, BC, CA, ABG, CBG, ACG, ABC, and ABCG. Incipient angles are from 0 to 350° in 10° increments. Fault resistance assumes the values: 0, 20, 100, and 1000 Ω. The distance takes the following values: 5%, 15%, 35%, 50%, 65%, 80%, and 95%, which are all percentages of total line length. Simulations are created for line 23-25 and it adjacent lines.

A total of 8066 cases per batch have been created, which gives a total of 48 396 cases because line 25-27 is double circuit line, i.e., two batches of faults for line 25-27 have been created: one for each circuit. Fault arc although simulated was not taken into account during training. This is because fault arc causes, if any, very little distortion to phase currents, which is consistent with the results in [20] and [13]. The previous statement has been validated using the fault arc model given in [40].

2) *Creation of Line Switching Cases*: For line switching cases, a batch for each line has been created giving a total of six batches. Each batch consists of smaller batches; each of the smaller batches corresponds to switching by one of the circuits breakers at each terminal. For example, line 23-22 would have a smaller batch for switching using the breaker installed on terminal 23 and another smaller batch for switching using the breaker on terminal 22. The variable in switching cases is the moment of switching. Switching times ranges from 0 to 360° in 2° increment. This way 360 cases per batch have been created, giving a total of 2160 cases.

3) *Creation of Lightning Strike Cases*: For lightning strike cases, a batch of each line has been again created. An ATP Heidler type lightning source with rising time equal to 4 μs and τ equal to 10 μs has been used. The rising and tailing times are kept constant during simulations but amplitudes have been varied, which were set to 5, 10, 15, 20, and 30 kA. Striking distances were the same as the ones used for fault batches: 5%, 15%, 35%, 50%, 65%, 80%, and 95%, which are all percentages of total line length. The incipient angle was 0, 90°, 180°, 270°, and 330°. This amounts to 630 cases per batch giving a total of 3780 cases.

VI. BUILDING OF FEATURE VECTOR OF TRANSIENT SIGNATURE

After batches have been created, a special subroutine is used to translate the phase quantities to modal quantities. The phase quantities are decoupled using the modal matrix calculated at 10 kHz even though the matrix is frequency dependent. Different modal frequencies for the modal matrix have been tried but this did not affect the classification results at all. The two aerial modes are the only ones used for training. The modal matrix with real entries is given in (4). The signs of the entries are only shown to emphasize their physical meaning

$$\begin{pmatrix} \text{Mode1} \\ \text{Mode2} \\ \text{Mode3} \end{pmatrix} = \begin{pmatrix} + & + & + \\ - & + & + \\ + & + & - \end{pmatrix} \times \begin{pmatrix} I_a \\ I_b \\ I_c \end{pmatrix}. \quad (4)$$

It should be apparent that Mode1 is the weighted sum of all phase currents, which is nothing other than the ground mode current. This ground mode current is known to be very well

dependent on the frequency and ground resistance on the contrary to the aerial modes that are more or less frequency independent [41]. This is the main reason that the ground current mode has not been used for building the feature vector for training.

Once the modal quantities are available, we run the DWT with db4 for four levels, starting from level 1 to level 4. Db4 has been used since it has produced good results in the previous literature. The outcome of the DWT is a series of coefficients for each level and for each mode. Since the first one-eighth of a cycle following the transient event is only sought, the details coefficients corresponding to that period are only obtained. Two modes out of the three modes are only used. The details coefficients of one level of Mode2 and Mode3 are stacked on top of each other. It should be emphasized that building the n -dimensional vector this way does not change the temporal content of the feature vector, as long as all other n -dimensional vectors are built consistently this way, i.e., the order of the modes in the feature vector has to be preserved and the vector has to be built using one level only. If the order of modes in the feature vector is changed, this amounts to a rotation of n -dimensional space but should not change the results of classification, again as long as all vectors are built consistently.

At this point, one should note that the application uses only two thirds of modes of the currents as contrasted to other applications that use all phase voltage and current waveforms for ANN training [18], which amounts to drastic reduction in computational power needed for ANN.

Having created n -dimensional vectors, the input for the ANN is ready. Training of the ANN is then started. The cases are randomly divided into the following categories: 70% of all cases for training, 15% for validation, and another 15% for testing.

VII. SIGNAL ANALYSIS

Even though ANN training is done offline, yet the principle proposed in this paper can be used for online applications. It should be noted that training has been done using hundreds of thousands of cases. For example, 8066 fault cases per line and a total of 54 336 transient cases have been used to train ANN as given in Section V-B. These cases encompass all variations of any transient that can occur on a transmission line. ATP/EMTP is mature enough that it can capture the actual system response. A real event oscillography would look like the simulated case in EMTP as long as the EMTP model is an accurate representation of the system. It is not possible to exhaust all foreseen transient parameters, for example, it is not computationally feasible to include all values of fault resistance into consideration in the training stage. However, it is evident that the oscillations present in the transient signal corresponding to a fault case with a fault resistance of 20 Ω should have oscillations that fall in between the same fault case but with a fault resistance of 0 and 100 Ω. This notion of “in-betweenness” is illustrated in Fig. 4. As explained in Section IV, the transient signal is eventually translated into a vector that belongs to the n -dimensional space. Thus, if the fault cases corresponding to 5%, 20%, 80%, and 95% line length can be linearly separated from all other transients, then the fault cases corresponding to 50% line length should fall in the

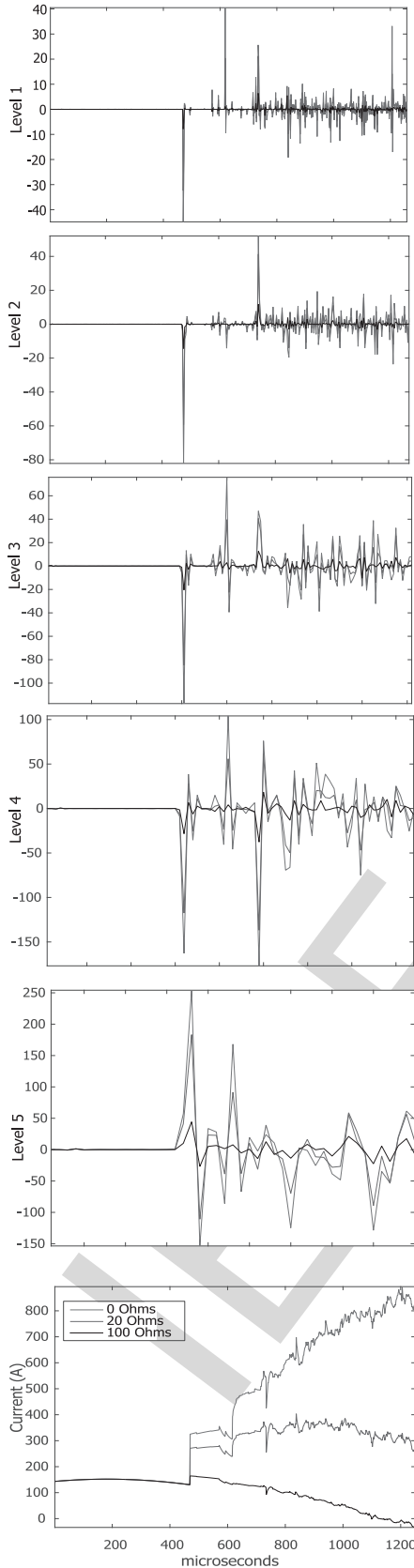


Fig. 4. Mode2 current of a certain fault case at different fault resistances. The same fault type, fault location, and incipient angle are evaluated but at different fault resistances.

TABLE I
FAULTS CLASSIFICATION ACCURACY

Transmission line	Accuracy (%)
Line under study	100
Line 23-32	100
Line 23-24	100
Line 23-22	100
Circuit (1) line 25-27	46
Circuit (2) line 25-27	54

region of those fault cases. The same argument applies for all event types. This idea of “in-betweenness” makes this approach applicable to online applications as long as the extreme cases are included in the offline training.

VIII. RESULTS

This section presents the results of the classification. In Section VIII-A, the results of classification in case all events were faults are shown. In Section VIII-B, results in case events were all line switching are presented. In Section VIII-C, classification output in case of all events were lightning strikes is shown. Finally, event-type classification is provided in Section VIII-D.

In Section IV, the mapping from n -dimensional space to the m -dimensional space has not been discussed. The dimension of the n -dimensional space is fixed by the length of the feature vector, while the dimension of the m -dimensional space varies according to which space we want to map the n -dimensional space to. In Sections VIII-A–VIII-C, the events are being mapped to the lines causing them, which amounts to a mapping from the n -dimensional space to a 6-dimensional space (one dimension per line, line 25-27 counted twice for the double circuit). In Section VIII-D, all events are being mapped to an event-type space, which is a mapping from n -dimensional space to a 3-dimensional space (faults, lightning, and switching). It should be apparent that the dimension of the feature space (n -dimensional space) depends on the DWT level used. It should be obvious as well that this dimension is almost halved as we go one level higher, i.e., the dimension of the n -dimensional space using level 1 is double the dimension of level 2 n -dimensional.

A. Line Identification in Case of Faults

This section provides the results for classification in case all transients were of faults.

First, six batches of faults are created as described in Section V-B1. Creating the feature vector using currents from terminal 23 as described before using any levels from 1 to 4 to train ANN gives the classification accuracy presented in Table I. One immediately sees that classification accuracy for the double circuits of line 25-27 fails. This is to be expected if the reader consults Fig. 3 for the tower configuration, which shows that from the Thevenin’s point of view at the terminal 23 or 25 of the line under study, the same fault case on any circuit will produce the same response due to the symmetry of the tower

TABLE II
SWITCHING CLASSIFICATION ACCURACY

Transmission line	Accuracy (%)
Line under study	100
Line 23-32	100
Line 23-24	100
Line 23-22	100
Circuit (1) line 25-27	48
Circuit (2) line 25-27	52

TABLE III
LIGHTNING CLASSIFICATION ACCURACY

Transmission line	Accuracy (%)
Line under study	100
Line 23-32	100
Line 23-24	100
Line 23-22	100
Circuit (1) line 25-27	49
Circuit (2) line 25-27	51

TABLE IV
TRANSIENT EVENT TYPE CLASSIFICATION ACCURACY

Event type	Accuracy (%)
Faults	100
Lightning	97.2
Switching	99.8

TABLE V
TESTING ANN IN TABLE I WITH NOISY AND DENOISED DATA
T IN CASE OF FAULTS

SNR	Transmission line	Classification accuracy	
		Noisy data (%)	Denoised data (%)
60	Line under study	99	99.3
	Line 23-32	98.2	98.7
	Line 23-24	98.4	99.1
	Line 23-22	98	98.6
	Circuit (1) line 25-27	48.1	48.9
	Circuit (2) line 25-27	49	49.4
40	Line under study	96	98.2
	Line 23-32	95.2	97.2
	Line 23-24	94	98
	Line 23-22	93.1	97.1
	Circuit (1) line 25-27	46	47
	Circuit (2) line 25-27	47	48.3
30	Line under study	90	97.3
	Line 23-32	89	96.5
	Line 23-24	89	96.8
	Line 23-22	88.7	97
	Circuit (1) line 25-27	42	46.5
	Circuit (2) line 25-27	44	47.9

25-27. It should be pointed out that the faults on one of the circuits of line 25-27 are confused for faults on the other circuit and vice versa but not with other lines. The percentage accuracy does not change if the currents at terminal 25 are used instead. Training with levels 1, 2, 3, and 4 gives the same classification accuracy but the difference is mainly in the ANN size needed. For reproducibility purposes, an ANN of size 40 in the hidden layer is needed to achieve the results presented in Table I for level 3. Lower levels (levels corresponding to higher frequencies) needed less neurons but the feature vector becomes so long (corresponding to more detailed time resolution). This in turn demands for a high super computer, which has been undertaken for level 1 (frequency band from 100 to 50 kHz) and level 2 (50 to 25 kHz). Level 3 (25–12.5 kHz) and level 4 (12.5–6.25 kHz) can be done on a PC such as the one the author used with Core i7, 3.2 GHz speed, 4 cores, and 8 GB of RAM. If the double circuit tower 25-27 is considered to be one line, a 100% correct classification is achieved for faults.

B. Line Identification in Case of Switching

This section shows the results of classification if all transients are for line switching. Creating the switching events as described in Section V-B2 and using either end currents, the accuracy presented in Table II is achieved. Training with any level from 1 to 4 gives the same classification result, but the only difference is being the size of the hidden layer. Level 3 required 30 neurons in the hidden layer with lower levels requiring less neurons and higher levels requiring more neurons. Once again switching events on one of the circuits of line 25-27 are confused for the other circuit and vice versa.

C. Line Identification in Case of Lightning Strikes

If all transient events are lightning only as described in Section V-B3, then using any end data gives the results presented in Table III irrespective of the terminal used or the level used (levels 1–4). Level 3 required 30 neurons in the hidden layer with lower levels requiring less neurons and higher levels requiring more neurons. One has to note that in Tables II and III, the algorithm cannot still distinguish between the events on any of the circuits of line 25-27 due to the symmetry with respect to the line being studied.

D. Transient Event Type Classification

Finally, the algorithm performance is shown when it comes to classification between different transients types. The types of studies in this paper are faults, line switching, and lightning. ANN is trained with all transient cases given in Sections V-B1–V-B3. The output is shown in Table IV. Once again levels 1–4 give the same classification accuracy whether bus 23 or 25 is used. The ANN used for Table IV has a size of 30 when trained with level 3 currents. It should be pointed out that in Table IV, most lightning strikes with amplitude 5000 A on line 23-22 have been misclassified for faults. A 5000 A lightning strike is unlikely in real-world lightning strike cases. If we remove all 5000 A strikes from training, a 100% accurate classification is achieved for event-type classification.

TABLE VI
TESTING ANN IN TABLE II WITH NOISY AND DENOISED DATA IN CASE OF SWITCHING

SNR	Transmission line	Classification accuracy	
		Noisy data (%)	Denoised data (%)
60	Line under study	98	99
	Line 23-32	97	98
	Line 23-24	97.1	98
	Line 23-22	97.5	98.5
	Circuit (1) line 25-27	47.5	49.2
40	Circuit (2) line 25-27	48.8	49.5
	Line under study	95	98
	Line 23-32	94.6	97.5
	Line 23-24	93.5	97.2
	Line 23-22	92.5	98
30	Circuit (1) line 25-27	45.5	48.7
	Circuit (2) line 25-27	46.1	49
	Line under study	87	96.5
	Line 23-32	86	95.2
	Line 23-24	86.2	95.6
	Line 23-22	87.4	96
	Circuit (1) line 25-27	42.2	47.6
	Circuit (2) line 25-27	43.5	48.5

TABLE VII
TESTING ANN IN TABLE III WITH NOISY AND DENOISED DATA IN CASE OF LIGHTNING

SNR	Transmission line	Classification accuracy	
		Noisy data (%)	Denoised data (%)
60	Line under study	100	100
	Line 23-32	99	99.3
	Line 23-24	99.1	99.5
	Line 23-22	99.7	99.7
	Circuit (1) line 25-27	49.3	49.3
40	Circuit (2) line 25-27	49.5	49.5
	Line under study	98	99
	Line 23-32	97	97.8
	Line 23-24	97.2	98
	Line 23-22	97.9	98.5
30	Circuit (1) line 25-27	46	47.3
	Circuit (2) line 25-27	47	48.5
	Line under study	95	98
	Line 23-32	94	96.7
	Line 23-24	94.6	96.2
	Line 23-22	94.8	97.6
	Circuit (1) line 25-27	42	46.6
	Circuit (2) line 25-27	41.5	48.1

It is now straightforward to construct a two-layer feedforward network that classifies any event to fault, lightning, or switching in the first layer, then identifies the line causing the event in the second layer. Alternatively, a relay can be programmed to classify the events into fault, switching, or lightning as per Table IV. Once this is done, the same event can be applied to the ANN corresponding to that event type.

IX. ERROR ANALYSIS

Since power systems are subject to harsh environmental conditions, error always exists in the measurements regardless of the bandwidth of the measuring devices used. The author

TABLE VIII
TESTING ANN IN TABLE IV WITH NOISY AND DENOISED DATA IN CASE OF ALL TRANSIENT EVENTS

SNR	Event type	Classification accuracy	
		Noisy data (%)	Denoised data (%)
60	Faults	99	99.2
	Lightning	96	96.8
	Switching	98.1	98.5
40	Faults	97	98.1
	Lightning	95.2	95.7
	Switching	96.5	97
30	Faults	95	97.5
	Lightning	94.4	95
	Switching	94.1	96.5

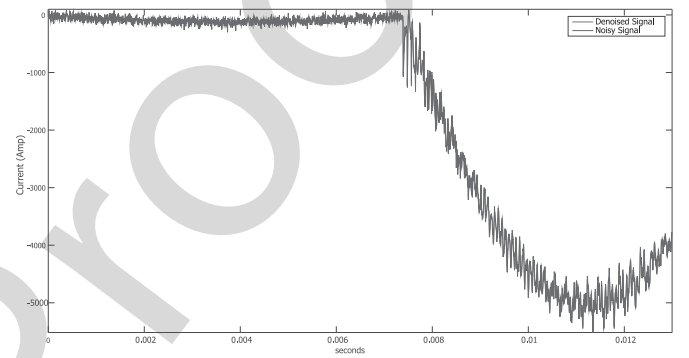


Fig. 5. Mode2 of a fault case with noise and after denoising.

analyzed the performance of the proposed algorithm when noise exists in the current measurements. Noise has been added to the data that have been generated in Section V-B. The noise that has been added is Gaussian white noise at different signal-to-noise ratios. The noisy data were then applied to the ANNs trained in Sections VIII-A–VIII-D to see how noise affects the classification accuracy. As can be seen from Tables V–VIII, classification accuracy decreases as SNR decreases. Even in the best case with high SNR, classification accuracy is not as good as it was in Sections VIII-A–VIII-D. This may be because ANN has bad generalization capability, which could signal that a better classification approach is needed.

To remove noise from data, the author used a wavelet denoising approach. The wavelets toolbox in MATLAB has been used to perform wavelet denoising. The denoised data are used for testing the ANNs trained in Sections VIII-A–VIII-D. The waveform of Mode2 current of a certain fault case with noise (SNR = 40) and after denoising is shown in Fig. 5. The results of this testing is presented in Tables V–VIII. As can be seen from the tables, denoising improves the performance of ANN. This is more noticeable when classification is done using data that have lower SNR. A detailed error analysis is beyond the scope of this paper and will be undertaken along with field validation.

X. COMPARISON TO EXISTING METHODS IN THE LITERATURE

The major drawbacks of existing methods have been summarized in Section I. In this section, the method proposed will

TABLE IX
PNN CLASSIFICATION ACCURACY

Event type	Accuracy (%)
Faults	98
Lightning	96
Switching	100

be compared to the existing methods in the literature that use single end data to reach a decision regarding the fault.

Even though the subject of transmission line relaying based on high-frequency components has and continues to receive attention from academia and the industry, a handful of papers study transients other than faults and their effect on the proposed methods.

In [18], the details coefficients of all phase voltages and currents are fed to ANN for training. However, for the method to perform well, the high-frequency transients imposed on the fundamental voltage have to be measured, which is not feasible if CVTs are used, as the bandwidth of CVTs does not exceed 1 kHz [26]. Additionally, the method ignores all transients types other than faults, which can mislead the scheme most of the times.

In [16], the details coefficients of all phase voltages and currents are still used to train the ANN in addition to using an energy quantity as a trip restraint. However, transients on the line being protected are being taken into consideration while training the ANN. A drawback of this method is the need to perform many offline studies to determine the threshold for the trip restraint. Moreover, the method suffers from the limited bandwidth of the CVTs.

In [20], the energy of the three-phase currents is used to train a probabilistic classifier. Voltages have been abandoned for the purpose of classification due to the aforementioned shortcomings of CVTs. However, the simulated system is small and the methods proposed cannot identify the line causing the transient event. Additionally, the use of energy as a feature requires a detailed study on the proximity of the transient to the measuring CT. The results in [20] are a special case of the results in this paper and correspond to the results in Section VIII-D.

For the purpose of statistical comparison, the author has reproduced the PNN classifier given in [20] and applied it to the system shown in Fig. 2. The training dataset that had been developed in Section V-B is then applied to the PNN classifier. The results of classification are presented in Table IX. As can be seen from both Tables IV and IX, not all lightning strikes are fully recognized. This occurs because some lightning strikes are confused for faults. The ones that are confused for faults are the ones that have 5 kA amplitude, which is consistent with the observation in Section VIII-D. Moreover, some faults are not well recognized using the PNN due to high resistance faults, which is consistent with the observation in [20]. The main advantage of the method given in this paper is that fewer features are needed to obtain satisfactory classification accuracy.

In all the methods mentioned above, at least one-quarter of a cycle is needed to reach a secure decision. On the other hand, the approach proposed in this paper uses only one-half of that

to reach a secure decision about whether the transient belongs to the line being protected and whether the transient is a fault. Additionally, actual physical tower configurations have not been taken into consideration in the literature.

XI. CONCLUSION AND FUTURE RESEARCH

This paper presented an argument that high-frequency signals can be used for high-speed power system fault detection via transient classification identifying the line causing the transient event.

The contributions of the paper are as follows.

- 1) It has been shown that currents alone can be used for transient signature of the event.
- 2) Only two modes are necessary for classification.
- 3) Only one-eighth of a cycle of postevent data is necessary for classification.

Although the results in this paper have been shown only to a specific tower, the author has tried different tower configurations and confirmed that the algorithm works for all tower configurations considered. It should be clear that the current method fails if a lightning strike evolves to a fault. This drawback has been addressed separately in a different publication [42]. Additionally, detection of faults on mutually coupled lines has not been addressed in this paper and will be treated in a different publication.

One issue that has not been investigated in this paper is insulation breakdown. The author has not assumed any failure of insulation in the simulations. The cutting of the signal associated with insulation breakdown can potentially mislead the scheme proposed. Further investigations are needed to completely quantify its effect on the proposed relaying scheme. It should be noted, however, that lightning strikes are being chopped by the surge arrester but still being correctly recognized.

Field validation is being performed and results will be shared once all investigations are done.

ACKNOWLEDGMENT

This paper could not have seen the light without the resources at Texas A&M Supercomputing Facility. The author would like to thank all at the facility for their help and support. The author also would like to express his sincere thanks for J. Woodworth of ArresterWorks for his guidance regarding arrester selection. Finally, the author would like to thank G. Gurralla from the Indian Institute of Science for his constructive review of the manuscript.

REFERENCES

- [1] A. Phadke and J. Thorp, *Computer Relaying for Power Systems*. Hoboken, NJ, USA: Wiley, 2009.
- [2] L. Kojovic and M. Kezunovic, "A new method for the CCVT performance analysis using field measurements, signal processing and EMTP modeling," *IEEE Trans. Power Del.*, vol. 9, no. 4, pp. 1907–1915, Oct. 1994.
- [3] J. Thorp, A. Phadke, S. Horowitz, and J. Beehler, "Limits to impedance relaying," *IEEE Trans. Power App. Syst.*, vol. PAS-98, no. 1, pp. 246–260, Jan. 1979.

- [4] E. O. Schweitzer, A. Guzmán, M. V. Mynam, V. Skendzic, B. Kasztenny, and S. Marx, "Locating faults by the traveling waves they launch," in *Proc. 2014 67th Annu. Conf. Protective Relay Eng.*, 2014, pp. 95–110.
- [5] S. Santoso, E. J. Powers, W. M. Grady, and P. Hofmann, "Power quality assessment via wavelet transform analysis," *IEEE Trans. Power Del.*, vol. 11, no. 2, pp. 924–930, Apr. 1996.
- [6] R. G. Lyons, *Understanding Digital Signal Processing*. Reading, MA, USA: Addison-Wesley, 1997.
- [7] A. V. Oppenheim *et al.*, *Discrete-time Signal Processing*, vol. 2. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989.
- [8] S. G. Mallat, "A theory for multiresolution signal decomposition - The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.
- [9] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, Sep. 1990.
- [10] A. Boggess and F. J. Narcowich, *A First Course in Wavelets With Fourier Analysis*. Hoboken, NJ, USA: Wiley, 2009.
- [11] W. A. Wilkinson and M. Cox, "Discrete wavelet analysis of power system transients," *IEEE Trans. Power Syst.*, vol. 11, no. 4, pp. 2038–2044, Nov. 1996.
- [12] F. H. Magnago and A. Abur, "Fault location using wavelets," *IEEE Trans. Power Del.*, vol. 13, no. 4, pp. 1475–1480, Oct. 1998.
- [13] T. Dalstein and B. Kulicke, "Neural-network approach to fault classification for high-speed protective relaying," *IEEE Trans. Power Del.*, vol. 10, no. 2, pp. 1002–1011, Apr. 1995.
- [14] M. Kezunovic, I. Rikalo, and D. J. Sobajic, "High-speed fault-detection and classification with neural nets," *Elect. Power Syst. Res.*, vol. 34, no. 2, pp. 109–116, 1995.
- [15] M. Oleskovicz, D. Coury, and R. Aggarwal, "A complete scheme for fault detection, classification and location in transmission lines using neural networks," in *Proc. 2001 7th Int. Conf. Develop. Power Syst. Protection*, 2001, pp. 335–338.
- [16] K. M. Silva, B. A. Souza, and N. S. D. Brito, "Fault detection and classification in transmission lines-based on wavelet transform and ann," *IEEE Trans. Power Del.*, vol. 21, no. 4, pp. 2058–2063, Oct. 2006.
- [17] P. L. L. Mao and R. K. Aggarwal, "A novel approach to the classification of the transient phenomena in power transformers using combined wavelet transform and neural network," *IEEE Trans. Power Del.*, vol. 16, no. 4, pp. 654–660, Oct. 2001.
- [18] F. Martin and J. A. Aguado, "Wavelet-based ANN approach for transmission line protection," *IEEE Trans. Power Del.*, vol. 18, no. 4, pp. 1572–1574, Oct. 2003.
- [19] G. Zwe-Lee, "Wavelet-based neural network for power disturbance recognition and classification," *IEEE Trans. Power Del.*, vol. 19, no. 4, pp. 1560–1568, Oct. 2004.
- [20] N. Perera and A. Rajapakse, "Recognition of fault transients using a probabilistic neural-network classifier," *IEEE Trans. Power Del.*, vol. 26, no. 1, pp. 410–419, Jan. 2011.
- [21] J. M. Zurada, *Introduction to Artificial Neural Systems*. St. Paul, MN, USA: West, 1992.
- [22] J. J. Grainger and W. D. Stevenson, *Power System Analysis*, vol. 31. New York, NY, USA: McGraw-Hill, 1994.
- [23] A. Abdullah, "Towards a new paradigm for ultrafast transmission line relaying," in *Proc. 2016 IEEE Power Energy Conf. Illinois*, 2016, pp. 1–8.
- [24] D. Hedman, "Propagation on overhead transmission lines I—Theory of modal analysis," *IEEE Trans. Power App. Syst.*, vol. PAS-84, no. 3, pp. 200–205, Mar. 1965.
- [25] D. Hedman, "Propagation on overhead transmission lines II—Earth-conduction effects and practical results," *IEEE Trans. Power App. Syst.*, vol. PAS-84, no. 3, pp. 205–211, Mar. 1965.
- [26] D. Hou and J. Roberts, "Capacitive voltage transformer: Transient overreach concerns and solutions for distance relaying," in *Proc. 1996 Can. Conf. Elect. Comput. Eng.*, 1996, pp. 119–125.
- [27] H. Lee, "Development of an accurate transmission line fault locator using the global positioning system," in *Proc. 1994 Nat. Tech. Meeting Inst. Navig.*, 1994, pp. 111–116.
- [28] D. A. Douglass, "Current transformer accuracy with asymmetric and high-frequency fault currents," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 3, pp. 1006–1012, Mar. 1981.
- [29] M. Redfern, S. Terry, F. Robinson, and Z. Bo, "A laboratory investigation into the use of MV current transformers for transient based protection," in *Proc. 2003 Int. Conf. Power Syst. Transients*, New Orleans, LA, 2003, pp. 1–5.
- [30] J. C. G. de Siqueira, B. D. Bonatto, J. R. Marti, J. A. Hollman, and H. W. Dommel, "A discussion about optimum time step size and maximum simulation time in EMTP-based programs," *Int. J. Elect. Power Energy Syst.*, vol. 72, pp. 24–32, 2015.
- [31] Washington State University, "118 bus power flow test case." [Online]. Available: https://www.ee.washington.edu/research/pstca/pf118/pg_tca118bus.htm
- [32] P. Demetriou, M. Asprou, J. Quiros-Tortos, and E. Kyriakides, "Dynamic IEEE test systems for transient analysis," *IEEE Syst. J.*, 2017, to be published.
- [33] S. D. Cho, "Parameter estimation for transformer modeling," Ph.D. dissertation, Michigan Technol. Univ., Houghton, MI, USA, 2002.
- [34] N. Perera, A. Rajapakse, and T. Buchholzer, "Isolation of faults in distribution networks with distributed generators," *IEEE Trans. Power Del.*, vol. 23, no. 4, pp. 2347–2355, Oct. 2008.
- [35] A. M. Gashimov, A. Kabayeva, and A. Nayir, "Transmission line transposition," in *Proc. 2009 Int. Conf. Elect. Electron. Eng.*, 2009, pp. 364–367.
- [36] ALSTOM, *Network Protection & Automation Guide*. Stafford, U.K.: ALSTOM Grid, 2002.
- [37] J. Woodworth, "Arrester selection guide." [Online]. Available: http://www.arresterworks.com/arresterfacts/pdf_files/Arrester_Characteristics_for_ATPDraw_Users.xls
- [38] A. Greenwood, *Electrical Transients in Power Systems*, 2nd ed. New York, NY, USA: Wiley, 1991.
- [39] A. Abdullah, "Atpmat: An open source toolbox for systematic creation of EMTP cases in ATP using MATLAB," in *Proc. 2015 North Amer. Power Symp.*, Oct. 2015, pp. 1–6.
- [40] V. V. Terzija and H. F. Koglin, "On the modeling of long arc in still air and arc resistance calculation," *IEEE Trans. Power Del.*, vol. 19, no. 3, pp. 1012–1017, Jul. 2004.
- [41] A. Ametani, N. Nagaoka, Y. Baba, and T. Ohno, *Power System Transients: Theory and Applications*. Boca Raton, FL, USA: CRC Press, 2013.
- [42] A. Abdullah, "A wavelet entropy approach for detecting lightning faults on transmission lines," in *Proc. 2016 IEEE Power Eng. Soc. Transm. Distrib. Conf. Expo.*, May 2016, pp. 1–5.



Ahmad Abdullah (M'xx) received the B.Sc. and M.Sc. degrees from the Department of Electrical Power and Machines, Faculty of Engineering, Cairo University, Giza, Egypt, in 2009 and 2012, respectively, and the Ph.D. degree from the Electrical and Computer Engineering Department, Texas A&M University, College Station, TX, USA, in 2017.

Since 2015, he has been with Electric Power Engineers, Inc., Austin, TX, USA, as a Power Systems Engineer responsible for conducting power system design studies.

QUERIES

992

Q1. Author: Please provide the expansion of acronyms “CVT,” “EMTP,” “ATP,” and “PNN,” at first occurrence.

993

Q2. Author: Please provide the expanded form of “LANs”

994

Q3. Author: Please complete and update Ref. [32].

995

Q4. Author: Please provide the areas of study in which Ahmad Abdullah received the B.Sc., M.Sc., and Ph.D. degrees.

996

Q5. Author: Please provide the year in which Ahmad Abdullah became a member of the IEEE.

997

IEEE Proof