

Bus Protection Using a Wavelet Based ANN

Ahmad Abdullah
Texas A&M University
Electrical and Computer Engineering Department
College Station, TX, 77840
Email: ahmad@tamu.edu

Abstract— This paper presents a new application of wavelet based artificial neural networks to the field of power system bus protection. In this paper it is shown that high frequency content of the bus voltage and any feeder current can be used not only to detect internal and external faults but also identify the faulted line in case of external fault. A scheme will be presented that uses only the bus voltage and any current from any of the lines connecting to the bus to detect internal and external bus faults, classify transients on adjacent lines and pinpoint the line that is causing the transient disturbance. Modal transformation is used to transform phase quantities to modal quantities. Wavelets are used to extract high frequency components of modal voltages and currents. A vector consisting of level 3 details of the two aerial modes of voltages and currents stacked in one column is used to train a neural network. The method is shown to be very robust against bus capacitance and surge arrester operation. Results show that very accurate classification can be made with one eighth of a cycle of post event data.

Index Terms-- ART neural networks, Classification algorithms, Discrete wavelet transforms, Wavelet coefficients, Modal analysis, Power system faults, Power system measurements, Current transformers, lightning.

I. INTRODUCTION

Power system bus protection has been traditionally been accomplished with voltage or current differential schemes [1]. In current differential scheme current transformers are installed on ongoing and outgoing feeders. Currents from CTs are then summed for a fault criterion. If a fault is outside the bus (external fault) then current summation equals to zero, if the currents don't add to zero then the fault is on the bus (internal fault). The major disadvantage of current differential schemes is CT saturation which impairs the method greatly [2]. High impedance current differential scheme have been proposed to counteract CT saturation [2], however they introduce other problems. The problem of CT saturation is eliminated completely at the source by the use of linear couplers with voltage differential principles. Voltages are used instead of current for fault discrimination. Currents in the feeders are converted to voltages using linear couplers [3] which produce linear input output voltage relation over a very large range of input currents.

Traditional bus protection schemes doesn't provide backup for incoming and outgoing feeders nor can they discriminate between various transient events occurring on adjacent lines.

We seek to provide; fault protection for the bus, back up for adjacent lines and classification between transients occurring on adjacent lines identifying the line causing the event in our approach.

The application of ANN for classification of faults has been given in [4]. Wavelets applications for power systems have been introduced in [5]. Wavelets have proven to be a very useful tool for transmission line fault detection and classification. Application of wavelets has received much attention for classification of power quality disturbance in [6], [7], [8] and [9]. In these applications and others energy [10] based methods seem to dominate the application of Wavelets in power systems. In energy based methods, the energy of travelling waves can be quantized in time because of the localization property of wavelets which makes it very useful in fault location application [11]. The application of ANN to bus protection is given in [12] where currents are fed to ANN for training and validation. In [13] a wavelet based scheme is presented for bus protection where currents going in and out of bus zone are analyzed using Discrete Wavelet Transform (DWT). DWT produces a series of details coefficients which are used to infer the polarity and the direction of currents thus issuing a trip signal when a fault occurs.

Our paper uses the local bus voltage information and one of the currents on the feeders connected to the bus for bus protection and backup protection for lines connected to the bus. We show that local data carries information enough for bus protection and providing backup for lines connecting to that bus. We also show that the local information can be used to classify transient events on lines connected to the bus and identify the line causing the transient disturbance. We use modal transformation [11] to transform phase signals to modal signals. DWT is used to extract one eights of a cycle of post fault data. The feature vector used to train ANN is a vector consisting of level 3 details coefficients of the two aerial modes [11] of voltages and currents stacked on top of each other. The transients studied in this paper are lightning, faults and line switching. The paper is organized as follows. A condensed material on ANN and wavelets is given in the background in section II. Methodology is given in section III. EMTP model and ANN training is given in section IV. Results are shown in sections V while conclusions are presented in the last section.

II. BACKGROUND

In this Section an overview of both wavelets and neural networks as used in this paper is provided.

A. Wavelets

In this paper the dyadic wavelet transform is used [14]. The transform takes the signal and applies low and high pass filters to it. The transform convert the original signal into an independent set of signals spanning certain frequency bands. The frequency bands are determined by the number of levels we want to analyze the signal. The word independent means that we can't derive one level from another, i.e., level 2 can't be derived from level 1. The frequency band of each level depends on the sampling frequency. Nyquist theory still holds here. The highest frequency we see in the signal will be at most equal to half of the sampling frequency. Given a certain sampling frequency the dyadic transform will first apply a high and low pass filter 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 so you can't derive anyone from the other. We can stop at this step or we can apply another low and high pass filter to the first approximation to get level 2 and approximation 2. Again 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 half of the frequency band. Continuing this manner by applying successive low and high pass filters we obtain a set of levels also called details and one last approximation. In theory the last approximation should correspond to a pure sine wave assuming the high frequency components imposed on the power frequency have high frequency not to be included in the frequency band of the last approximation.

Taking numerical example, if we use a 100 kHz sampling frequency then we have: level 1 will occupy the frequency band from 50 kHz to 25 kHz, level 2 will occupy the frequency band from 25 kHz to 12.5 kHz, level 3 will occupy the frequency band from 12.5 kHz to 7.25 kHz and approximation 3 will correspond to 7.25 kHz to zero hertz.

The equation used for the transform is given below in (1) where $\phi(t)$ is the wavelet is used, the parameter a causes scaling (which is level determination) while the parameter b causes shift in time. In practice it is not needed to apply the transform all over the infinite line. Since wavelets have a strong localization property then it is only needed to apply the transform to the time period under study. The equation for the details coefficients are given below in equation (1) [14]

$$Wf(a, b) = a^{-1/2} \int_{-\infty}^{\infty} f(t) \overline{\phi\left(\frac{t-b}{a}\right)} dt \quad (1)$$

In this paper, we use Daubechies type 4 wavelets (db4). Any family of continuous wavelets can be used but we found that db2 gives the same results. Other types of wavelets would make the neural network used bigger which is against computational efficiency sough

B. Classification with neural networks

Neural networks are very well known tool for

classification. The classification done in this paper is not probabilistic but rather deterministic. The inputs are vectors in the n -dimensional space, the size of an input vector, that need to be mapped to another vector space the size of which is determined by the number of output classes we want to map to. In simple terms this classification can be shown to be carried out by a neural network [15].

All neural networks in this paper consist of three layers: an input layer, a hidden layer and 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 is greatly affected by the number of neurons in the hidden layer. The output layer consists of neurons that are activated by a function the input of which comes from the hidden layer. Classes or more specifically output classes are the patterns we want to map the input to.

The main idea behind classification in the n -dimensional space is that if the inputs belong to subspaces of the n -dimensional space, then using neural networks planes can be drawn in that space so that the space between those planes contain only one subspace and each subspace 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 synapsis are adjusted during the training phase such that the vector space is divided to subspaces each of which corresponds to certain class. Any output neuron receives the input vector through synaptic weights. The input is then applied to the function of the neuron which has the form $g(x)=0$ which is an equation of the plane in the n -dimensional space. The neuron is activated only if the output is positive.

In this paper we use the widely known and used backward propagation algorithm for the calculation of weights. Knowledge is stored in the network through those weights. We only do supervised learning in this paper. The algorithm of the steepest gradient is used throughout the paper and is given in [15].

III. STATEMENT OF THE PROBLEM AND METHODOLOGY

In this paper we show that the information present in the transient signal available at the protected bus caused by sudden network topology changes contains sufficient information for classification not only between internal and external bus faults but also between transients on adjacent lines which effectively provides back up for lines terminating on the bus. Any change of configuration on the lines terminating on the protected bus causes a traveling wave to be generated travelling from the point of change towards the ends of the line. In the simplest case this wave will be just a step but in reality will have a lot of oscillating components. Fourier analysis is not suitable for such waveforms because the oscillations will be distorted and attenuated as they arrive at the line terminals. However, applying discrete wavelet transform will enable us to see both the spectral and temporal

variations of the traveling wave.

We extract the information from the traveling waves using the DWT at level 3. Fourier transform is not suitable for this purpose since the signals' frequency content of the signal changes with time. We choose level 3 because of computational efficiency. Using one eighth of a cycle of post fault data corresponds to only 45 coefficients per mode, so the feature vector will be of length 180. At such lengths, A PC can be used for ANN training. Higher levels will have more coefficients – basically doubling the length of feature vector for each level we go higher (higher levels means levels 1 and 2). Before we apply DWT we decouple phase currents and voltages from each other using modal analysis [16]. Since current transformer has a bandwidth of 100 kHz [17] we choose a sampling rate of 200 kHz for our ATP simulations. After decoupling phase quantities, DWT is applied to the voltages and currents 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 coefficients is detected. Once the event is detected, a window of one eighth of a cycle of post fault information is used for neural network training. The outcome of the DWT is a series of coefficients. We stack the coefficients of the two aerial modes of currents and voltages on top of each other to build one vector that will be used to train the network. A neural network of an appropriate size is selected for classification. The neural network is then trained using various scenarios for transients on lines. If we run internal and external fault simulations on the bus and its adjacent lines then using the feature vector built above, we can tell whether the bus is faulted or which line is being faulted. If we only run line switching cases the lines terminating on the bus then using local data of protected bus, we can tell which line is being switched. If we run only lightning cases, then using local data of protected bus we can identify the adjacent line being hit by a strike. Now the aggregate of all transient event types can be distinguished from each other: this means if we train ANN using all transient events, then we can have three output classes each of which corresponds to fault, line switching and lightning. In summary, the procedure is as follows:

1. Currents and voltage are decoupled using modal analysis.
2. Apply DWT to the two aerial modes of one eighth of a cycle of level 3 coefficients of post event information.
3. Build feature vector consisting of details coefficients of the 4 aerial modes on top of each other to train the neural network. The 4 modes are two aerial modes for bus voltage and two aerial modes for any feeder current.
4. A neural network of suitable size is selected for training. Training is accomplished using transient scenarios. Those scenarios include: faults, line switching and lightning.

IV. SIMULATION PLATFORM

This section describes the EMTP model used and the cases used to train the neural network and how the data was prepared as input for ANN.

A. ATP/EMTP Model Description

We present the topology of the system under study in fig. 1. This topology is part of IEEE 118 bus case [18]. Voltage is

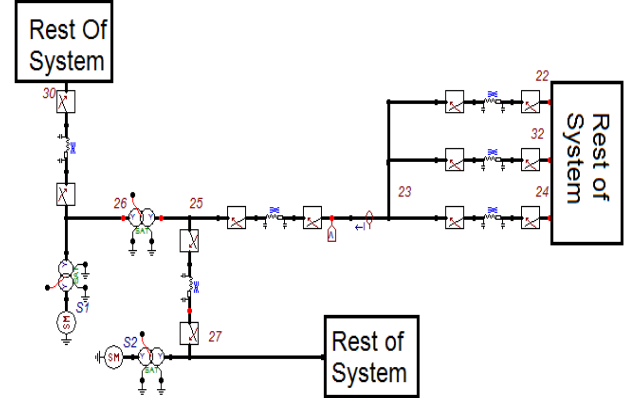


Fig. 1 Area under study

being measured at bus 23 and the current being measured from bus 23 to bus 25. Any current can be used but we only choose to report on this current due to paper length requirement. Our selection for bus 23 is for study because it has almost all power system components around it. Bus 23 has three incoming feeders into it, namely line connecting bus 23 to bus 22 (line 23-22), line 23-32 and line 23-24. The outgoing feeder is line 23-25. Bus 25 has a generator connected to it through a step up transformer and power transformer connecting it to bus 26. Bus 26 has another generator connected to it via a step up transformer. All lines are modeled as frequency dependent models with no ground return. Complete transposition has been assumed. We apply transient cases only to lines 23-25, 23-22, 23-24, 23-32. We don't aim to study transients beyond those lines. All generators in the system are assumed unregulated machines with no exciter or governors. This done mainly because we only interested in one eighth of a cycle during which exciters and governors time constants are large compared to the period under study. Power transformers are modeled using lumped parameters which is justified for our level 3 frequency band [19]. Bus capacitance and transformer stray capacitance is accounted for in our simulation platform due to their high importance for switching studies. Surge arresters have been added at all buses surrounds the study area. These arresters are selected according to the material given in [20].

Creation of transient cases has been automated using the toolbox released in [21]. The code builds on another code we used for a previous paper and the reader should consult the paper in [22] for inner working of the toolbox. For this paper we need to create cases of faults, line switching and lightning. For fault, we create all types of faults at varying fault resistances and different incipient angles. A total of 8066 fault cases per line have been generated. Those faults span all types of faults at all distances from 5% to 95% of line length and with fault resistances up to 1000 Ohms. A total of 1120 of fault cases have been also generated for bus faults. These include all faults types with varying incipient angles and various fault impedances including high resistance faults.

In the next batch of simulations, we stop creating fault cases and generate lightning cases. We strike the adjacent lines by lightning strikes with different amplitudes and at different instants of times fixing the waveform rise and tail times. We use an ATP Heidler type lightning with rising time equal to 4 μ s and a τ equal to 10 μ s. Amplitudes ranging from 5000 A to 30000 A at distances up to 95% have been simulated. It is very rare that a lightning strikes a bus in a substation since buses are almost always inside housing so no lightning strike cases has been generated for the bus under study. A total of 630 cases per line have been generated and used for neural network training. This gives a total of 3150 cases for lightning strikes.

Finally we stop creating fault cases and lightning cases, we energize the lines terminating on the bus at different instants of time. The switch at each end of the line is switched every two degrees making the total number of all cases 360 per line giving a total of 1800 cases of line switching.

B. ANN Training and Preparation of Data

Once transient cases are available from simulations, we have a subroutine that transforms pl4 files to mat files using Matlab. After that Matlab is invoked to import simulations into workspace, another subroutine is used to transform phase signals to modal signals. At this stage we apply DWT to the modal data. We extract the details coefficients corresponding to level 3 aerial modes. Once we extract these coefficients, we build the feature vector consisting of two aerial modes coefficients of the currents stacked on top of the two aerial modes coefficients. One should note that we can build the same feature vector by stacking the modes in any order but we have to be consistent in building the vector, i.e., we must choose one method throughout our training or we will get contradicting results. It should be clear at this point that using only two modes instead of three phase quantities make our approach faster and requires less computational power for training as opposed to other methods that use all voltages and currents for training [23].

After we build the feature vector we randomly divide the data set. We choose to have seventy percent of all cases have been used for training while fifteen percent has been used for validation and the last fifteen percent has been used for testing.

V. RESULTS

This section provides classification results for faults, line switching and lightning.

A. Line Identification in Case of Switching

In case of line switching, we build the feature vector as described in section III with level 3 details of the two aerial modes of currents stacked on top of level 3 details of the two aerial modes of voltages. After we train the ANN with a hidden layer of 40 neurons, we find that a 100% classification was possible. The user has to keep in mind that 70% of all cases have been used for training while 15% of all cases have been used for validation and another 15% for testing. This means that the feature vector can tell which line is being switched if all transient events were just switching.

B. Line Identification in Case of Lightning

If all transient events were just lightning strike cases, then the feature vector built the way described in the methodology section again was able to determine the line being hit by a strike by 100% accuracy. An ANN with only 40 neurons in the hidden layer was enough for 100% accurate classification. Once again that 70% of all cases have been used for training while 15% of all cases have been used for validation and another 15% for testing

C. External and Internal Bus Faults

In this case, we assume that all transient events are faults; we build the feature vector as described before in the methodology; use trial and error to get an appropriate size of ANN that gives good classification results. If we are only interested in knowing whether a fault is an internal bus or external bus fault, then an ANN with 20 neurons in the hidden layer gives 100% classification accuracy. If we need to see if the neural network can recognize the adjacent line being faulted, then an ANN with 80 neurons is necessary for 100% accurate classification. The size of the data used for training, testing and validation is the same as subsection A and B.

D. Transient Event Type Classification

In this section we investigate whether the feature vector proposed can identify the transient event type. For this, we train the ANN with all transient cases given in subsections A, B and C such that its output layer has three layers: one for switching, another one for lightning and a third one for faults. Building the ANN this way and using 50 neurons in the hidden layer gives the classification output in table (1). It has to be kept in mind that the feature vector has to be built consistently the same way for all cases. For table (1) we used

Table 1
Classification Accuracy for All Transient Events with Level 3
Feature Vector

Event	Classification Accuracy
Faults	100%
Switching	99.5%
Lightning	97%
Total Accuracy	98.5%

currents on top of voltages.

A relay that is programmed such that it processes the local current and voltage using modal analysis building the feature vector as we proposed can first identify the event type as given in table (1). Once the event is identified, the feature vector is then applied to corresponding ANN of the corresponding type to determine the line causing such transient event.

VI. CONCLUSIONS

A new application of wavelet based ANN for bus protection has been presented in this paper. The contributions of the paper are:

- The scheme provides ultra-high speed protection with only one eighth of a cycle of post event data.

- The approach doesn't suffer from CT saturation which is considered an improvement over modern differential schemes.
- It has been shown that the backup protection can be provided to adjacent lines as well as identifying internal and external faults.
- The algorithm can not only detect faults, but also can classify transient disturbances on adjacent lines identifying the line causing the transient event.

Results show that surge arrester has minimal effect on classification even though it chops travelling waves. We showed that feature vector must contain both voltages and currents for classification.

A drawback of this approach is the use of voltages for the feature vector. It is known that the bandwidth of CVT is only around 200 Hz [24]. The level 3 voltage used in this paper would be then highly attenuated and distorted. Modern microprocessor relays however can compensate for the CVT response by programming. Optical voltage transformers can solve this problem as well. Another disadvantage of this approach is that the ANN architecture used in this paper cannot detect faults that are caused by lightning strikes. Further research is still needed to rectify those issues. The effect of line transposition and mutual coupling has been ignored in this research and is a subject for further investigations.

VII. REFERENCES

- [1] A.R. Warrington, *Protective Relays Their Theory and Practice Power Systems*, vol. I. Springer, 3rd edition.
- [2] P.M. Anderson, *Power System Protection*, Wiley-IEEE Press, 1st edition.
- [3] C. Mason, *The Art and Science of Protective Relaying*, Wiley, 1 edition.
- [4] Kezunovic, M.; Rikalo, I., "Detect and classify faults using neural nets," *Computer Applications in Power, IEEE*, vol.9, no.4, pp.42,47, Oct 1996
- [5] Robertson, D.C.; Camps, O.I.; Mayer, J.S.; Gish, William B., "Wavelets and electromagnetic power system transients," *Power Delivery, IEEE Transactions on*, vol.11, no.2, pp.1050, 1058, Apr 1996
- [6] Gaouda, A.M.; Kanoun, S.H.; Salama, M.M.A.; Chikhani, A.Y., "Wavelet-based signal processing for disturbance classification and measurement," *Generation, Transmission and Distribution, IEE Proceedings-*, vol.149, no.3, pp.310, 318, May 2002
- [7] Gaouda, A.M.; Salama, M.M.A.; Sultan, M.R.; Chikhani, A.Y., "Application of multiresolution signal decomposition for monitoring short-duration variations in distribution systems," *Power Delivery, IEEE Transactions on*, vol.15, no.2, pp.478, 485, Apr 2000
- [8] Gaouda, A.M.; Salama, M.M.A.; Sultan, M.R.; Chikhani, A.Y., "Power quality detection and classification using wavelet-multiresolution signal decomposition," *Power Delivery, IEEE Transactions on*, vol.14, no.4, pp.1469, 1476, Oct 1999
- [9] Santoso, S.; Powers, E.J.; Grady, W.M.; Hofmann, P., "Power quality assessment via wavelet transform analysis," *Power Delivery, IEEE Transactions on*, vol.11, no.2, pp.924, 930, Apr 1996
- [10] Ibrahim, W.R.A.; Morcos, M.M., "Artificial intelligence and advanced mathematical tools for power quality applications: a survey," *Power Delivery, IEEE Transactions on*, vol.17, no.2, pp.668,673, Apr 2002
- [11] Magnago, F.H.; Abur, A., "Fault location using wavelets," *Power Delivery, IEEE Transactions on*, vol.13, no.4, pp.1475,1480, Oct 1998
- [12] Xiucheng Dong; Han Han; Mei Shu, "Research of bus protection based on artificial neural network," *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*, vol., no., pp.4261,4264, 25-27 June 2008
- [13] N. Perera, A.D.; Rajapakse, D. Muthumuni, "Wavelet Based Transient Directional Method for Busbar Protection", International Conference on Power Systems Transients (IPST2011) in Delft, the Netherlands June 14-17, 2011
- [14] Albert Boggess, Francis J. Narcowich, *A First Course in Wavelets with Fourier Analysis*, College Station, TX: Wiley 2009
- [15] Jacek M. Zurada, *Introduction to Artificial Neural Systems*, Louisville, KY: West Group, 1992
- [16] Hedman, D.E., "Propagation on Overhead Transmission Lines I-Theory of Modal Analysis," *Power Apparatus and Systems, IEEE Transactions on*, vol.84, no.3, pp.200,205, March 1965
- [17] Douglass, D.A., "Current Transformer Accuracy with Asymmetric and High Frequency Fault Currents," *Power Apparatus and Systems, IEEE Transactions on*, vol.PAS-100, no.3, pp.1006,1012, March 1981
- [18] Power Systems Test Case Archive, [online] https://www.ee.washington.edu/research/pstca/pf118/pg_tca118bus.htm
- [19] N. Chiesa, "Power Transformer Modelling Advanced Core Model", MSc thesis, Politecnico di Milano, Italy.
- [20] Arrester Selection Guide, [online] http://www.arresterworks.com/resources/selection_guide.php
- [21] [online] <https://bitbucket.org/ahmadmabdullah/atpmat>
- [22] A. Abdullah, A. Esmailian, G. Gurralla, P. Dutta, T. Popovic, M. Kezunovic, "Test Bed for Cascading Failure Scenarios Evaluation", International Conference on Power System Transients-IPST 2013, Vancouver, Canada, Jul 2013
- [23] Martin, F.; Aguado, J.A., "Wavelet-based ANN approach for transmission line protection," *Power Delivery, IEEE Transactions on*, vol.18, no.4, pp.1572,1574, Oct. 2003
- [24] Daqing Hou; Roberts, J., "Capacitive voltage transformer: transient overreach concerns and solutions for distance relaying," *Electrical and Computer Engineering, 1996. Canadian Conference on*, vol.1, no., pp.119, 125 vol.1, 26-29 May 1996