A Modified Neural Network Relaying for High Efficiency Biomedical Power Transformer Protection

Balan K.[1]*, Sudha Mohanram[2] and Karpagam M.[3]

1. Department of Electrical and Electronics Engineering, Sriguru Institute of Technology Coimbatore, INDIA
2. Sri Eshwar College of Engineering, Coimbatore, INDIA
3. Department of Electrical and Electronics Engineering, Hindusthan College of Engineering and Technology Coimbatore, INDIA

*josephjennu@gmail.com

Abstract

The main element in a power system is the power transformer. A transformer set-up forms the central element in transmitting and distributing power. There is heavy transmission loss if the power system is defective. The transformers should be protected from internal faults and magnetic inrush currents. The power demands for a biomedical research center or hospital are higher and need to be consistent for the working of all biomedical equipments used. This paper proposes transient detection techniques using Wavelet transform (WT). The advantage of WT is that, it can represent current and voltage signals. Empathy of transients is very fast and precise so this work tends to elaborate on a new wavelet method to classify the inrush currents from power system disturbances. The classification problems in Artificial Intelligence (AI) are solved using neural networks. The neural networks derive its inspiration from neurons present in human body. The performance of NN is assessed based on the design of the hidden layers and by evaluating the weights that connect the nodes. There are no modifications made to the design of the hidden layers as the weight calculated is more important than the hidden structure. An optimum solution can be obtained by calculating the weights of NN by decreasing the function cost or error. This work aims to propose a modified neural network. The neural networks are trained using Firefly Algorithm (FA). Since this protocol is a simple and an efficient metaheuristic optimization technique which draws inspiration from fireflies which tend to move towards brighter light. The experimental results prove the computational efficiency of training process using firefly optimization technique.

Keywords: Power Transformer Protection, Artificial Neural Network (ANN), Wavelet Transform (WT), Firefly Algorithm (FA).

Introduction

Reliable and efficient power supply is required for hospitals, biomedical centers to operate the wide-ranging medical equipment that has to be powered round the clock. The biomedical equipments require steady power. The most significant component of power system is the transformer. Transformers must be protected to have a stabilized power system. In the recent years, the number of industries has increased at an alarming rate which in turn has led to installation of equipments that are bigger in size with high ratings. The present protection system consists of minimal circuit breakers, relays and other protective devices. However, there is no protection against abnormal or fault conditions. Such conditions can be overcome by using digital relays. These relays use software protocols with advanced logics ensuring better protection.

Protecting the biomedical systems is of at most importance. The digital relays have advantages like high operating speed, reliability, stability, flexibility and it is very economical. These aspects have fascinated the researchers to use digital relays in protection of power systems. These relays can also be integrated with computer system within substation and the grid. Artificial intelligence is applied to relays to improve its decision-making capability.

Another most popular protection system is differential protection. The current that enter the system is matched with the current that leave the system. If the sum of total currents is equal to zero, then there is no fault in the system. But if the sum of the incoming and outgoing currents is not 0 then it implies there is an error within the system. Consequently, the differential system identifies and isolates the faulty zone from the system. The functioning of the traditional differential protection system is hindered by three major setbacks. The differential relay is prompted to generate a false trip signal even during normal conditions. [1]

The primary part of the transformer experiences a transient magnetizing inrush, every time the transformer is turned on and the rate of instantaneous voltage has not been 90. During this under the steady state condition, the peak flux is higher the top leading flux wave. This results in internal fault thereby leading to differential currents. The magnetizing current's peak value is several times higher than the peak value of full load current. The differential relay trips the transformer if there are inrush currents taking it as faults in the transformer. The working principle of differential relay is that it matches the two currents from either side of the transformer as discussed above. In the case of transformers there is inrush on only one side and it has a significant impact on the transformer. So, the differential relays must be designed in such a way it recognises inrush current as a normal phenomenon rather than tripping the circuit.

In the previous methods, the transients were treated by desensitizing or delaying the relay. These methods proved less efficient as the transformers were exposed to long unprotected times. To overcome this problem and to guarantee improved security and reliability, second harmonic content was used instead of the fundamental one and this process is called as harmonic restraint differential protection. On the other hand, it was observed that the second harmonics were significantly present in winding faults. In addition, amorphous material with lower loss is used in the hub of modern transformers, to generate in dash currents with low harmonics and high magnitude. Under these circumstances the relation between the second harmonic and the basic restraining condition is modified by means of relations described at high frequency. There are many other proposed methods like Hidden Markov's Model (HMM), fuzzy-logic-based methods, wave shaped recognition method as well as ANN based learning pattern approach to achieve quick response, more accurate results with low computational burden. [2]

The Wavelet transformer transient phenomenon is considered to be more efficient as it can recover data from transitory signals that are present in time and frequency domains. The WTs are becoming more popular and it is applied in various fields like power system transients, power quality, to detect problems and fault locations. It is also used to analyse the transient currents caused due to faults in power transformer and the magnetizing inrush currents. The experimental results illustrate that certain components of WTs can be used in differentiating internal faults and magnetizing currents.

The Artificial neural networks derive its inspiration from neurons present in human body and provide wide range of advantages. The important features of ANN are its flexible nature, proficiency, reliability and the ability to analyse and to resolve problems related to in patterns classification, functions approximation, patterns matching as well as associative memories. Amongst the available ANN models, the Multi-Layer Feed Forward Neural Network (MLFF) is more efficient and it is popular for its universal approximation capabilities.

The ANN has become more successful due to its design, training algorithm and its configurations employed in training. It is applicable in various optimization and mathematical issues are classification, object and image recognition and in several domains like Signal processing, seismic event predictions, temperature as well as weather forecasting, bankruptcy, tsunami intensities, earthquakes, as well as sea levels and so on. Back Propagation Neural Network (BPNN) protocol was utilized to train ANNs to improve it network performance. Conversely, it has two main demerits like lower convergence rate and unsteadiness due to the high risks of getting trapped in local minima and overshooting the minimum error plane.[3]

Generally, in engineering the majority of optimisation problems are nonlinear with many constraints. Proficient optimisation algorithms are used in solving such nonlinear problems. These optimisation algorithms are classified into two as: deterministic and stochastic. Hill Climbing is a type of deterministic algorithm that will produce the same solutions for a given set of initial values. Unlike deterministic algorithm, stochastic protocols can produce different solution for the same set of initial values. The required accuracy is achieved, as the results though not identical, converges at the same optimal point.

All the local search protocols belong to deterministic algorithms and are quite capable of finding the optimum solution. Despite this, it is possible with the protocol to get forced in local minima as the global minima are not within reach. So, the general practise is to make these protocols stochastic by introducing certain stochastic component which enables the algorithm to get away from such locality. As a result, stochastic algorithms have both the deterministic component as well as random component. The components of stochastic algorithm are capable of taking different varieties like easy randomization by means of random sampling the search area or with arbitrary walks. Many of the stochastic protocols may be treated as meta-heuristic. Suitable instances are Genetic Algorithms (GAs) and Particle Swarm Optimisation (PSO). The basis for these algorithms is the nature inspired swarm intelligence. The new protocols that are developed are becoming more competent and efficient. For instance, the FA is proven to be more superior to other conventional methods.

FA is presented in this work deals with optimization problem training of ANN. FA is a modern metaheuristic algorithm motivated from nature. The basis of FA is derived from the fireflies move i.e. the tendency of the firefly to move towards brighter light. Firefly back propagation (BP) algorithm is used in analysing the performance and behaviour of ANN training using FA. Following this, Section 2 presents an overview of the related work in literature. Section 3 details the various methods used and Section 4 details the outcomes of experiments. Section 5 presents the conclusion.

Related Works

Zellagui and Chaghi[4] presented distance communication settings performance based on analytic (AM) and ANN scheme with 400 kV higher voltage of transmission line presented in Eastern Algerian transmission networks of Sonelgaz Group. It was balanced by sequence of Flexible AC Transmission System (FACTS) that is, Thyristor Controlled Series Reactor (TCSR) had been attached intermediate at electrical transmission line. The TCSR insertion effects on the overall impedance defended by the transmit distance of a transmission line and the customized setting region defending in capacitive as well as inductive boost up form for 3 zones. Investigation was carried out for two different methods to prevent the circuit breaker from tripping for unwanted reasons as well to enhance the performance of distance relay.

Bigdeli et al[5] proposed a modern transfer function analysis method to identify the transformer winding faults. This analysis is carried out by using vector fitting and probabilistic NN. The NN is trained using the results from transfer functions estimation through vector fitting and the faults can be classified with probabilistic NN. To identify the type of fault the required information is acquired from two types of transformers namely a typical 20 kV transformer as well as a latest transformer subjected to normal and abnormal circumstances like axial displacement, radial deformation, disc space variation, as well as short circuit of windings. The outputs are compared with other well established methods and it is understood that the proposed method is superior to others.

Gholizadeh[6] proposed two strategies for computation which is an effective methodology for Performance-Based Optimum Seismic Design (PBOSD) of steel moment frames. The Modified Firefly Algorithm (MFA) is the initial strategy to effectively identify the PBOSD at the performance levels. The structural responses can be worked out on performance level by means conducting analysis on nonlinear static pushover, but in turn it increases the processing time. The second strategy was presented to minimise the computational burden by an innovative NN model NN namely Wavelet Cascade-Forward Back-Propagation (WCFBP) to accurately forecast the outputs for non-linear push over analyses in optimization method. The outputs show that the presented soft computing method for PBOSD steel constructions is inexpensive in terms of computation.

Tripathy et al[7] presented the optimal PNN to differentiate among the magnetizing inrush as well as the interior error of a power transformer. PSO was utilized to acquire an optimum smoothing aspect of PNN as a significant constraint for PNN. To protect transformer a protocol is developed under the theme, conventional differential protection. The operating condition is determined by finding the ratio between voltage-to-frequency as well as the amplitude of dissimilar current. The traditional homoscedastic-type PNN,FeedForward Back Propagation (FFBP) NN, as well as the conventional harmonic restraint technique are used in evaluating the performance of the suggested heteroscedastic-type PNN.

Nayak et al[8] presented a higher level NN based on Firefly to classify data and to maintain the learning speed by avoiding the exponential increase of processing units. Many researchers were conducted to analyse the previously developed methods. The results from several standard data sets from UCI machine was utilized to assess the working of this suggested scheme. The outcomes obtained were evaluated with the performance of other models. The simulation effects showed that the suggested method was steady, quick, and reliable with exact classification.

Suja and Jerome[9] introduced the purpose of wavelet based NN in determining the harmonics. The DWT when combined with PNN model can construct the classifier of harmonics. Further, the effectual significances of three phase active as well as reactive power can be calculated using the DWT. Similarly, the classifier of harmonics can be built by combining DWT with PNN model. Harmonics can be identified in four stages; in the first stage an inverter circuit with three phase was simulated. In the next stage the 3-phase current as well as the voltage is converted into alpha-beta reference frame using Power Quality (PQ) theory. In the third stage multi resolution analysis is carried on these alpha-beta vectors. The signals are later disintegrated by DWT at six levels. In the last stage, the values of active power, reactive power and energy are calculated using the generated coefficients.

Abdullah and Al-Zyoud[10] proposed a simultaneous formulation in which ANN optimizes reactive power control problem. This method mainly aims at minimizing loss in active power as well as increasing the voltage profile of the provided system. The voltage as well as reactive power controlled by usage of the power transformer transformation ratios as well as the inserted reactive power to be optimized system concert on the basis of Feed-forward ANN with BP training protocol was utilized while the training data was got through optimum purpose transformation parts and for multiple operating modes reactive power injected were performed on Jordanian Electrical Power System by means of ANN as well as contrasted with another regression model. The BP ANN model was validated for predicting the minimization for active power loss as well as for voltage control.

Nandy et al[11] employed a metaheuristic algorithm with BP techniques for training a Feed-Forward Neural Network (FFNN). To obtain a fast and enhanced convergence rate in training FFNN, FA a meta-heuristic algorithm was integrated with the BP protocol. The suggested method was experimented with a data set of and the results revealed that this method was superior to other methods as it required only little iteration to produce enhanced convergence rate. The result was positive with improved convergence rate when the proposed method was tested against Genetic Algorithm (GA) based BP.

Shah and Bhalja[12] proposed a modern differential protection scheme to discriminate bwteen the faults. The method derives its basis from support vector machine (SVM). This method could effectively differentiate between internal faults with other abnormal conditions like inrush currents and over excitation. The input to SVM classifier was the output from feature extraction using WT. The software package PSCAD/EMTDC was used to simulate the existent power transformer of Gujarat Energy Transmission Corporation Ltd. (GETCO). The tests included simulation of internal errors as well as other disturbances using varying mistake and system constraints. This proposed algorithm was experimented using a simulated on 5442 cases of data set. From the results, it can be deduced that an accuracy of more than 99% is achievable in overall fault discrimination.

Rasoulpoor and Banejad[13] suggested a new method to differentiate among inrush currents and internal errors using differential defending method in power transformer. The basis for this technique is WT and correlation coefficient. This work uses DWT in describing the existing signal for varying time and frequency modules. The signal energy with this module was also used in this algorithm. Additionally, an arithmetical parameter correlation coefficient established a criterion for this discriminative algorithm. The experimental results inferred from analysis were that it can accurately differentiate inrush currents in a power transformer protection from error currents in a frame time below quarter power frequency cycle.

Ramesh and Sushama[14] proposed a modern method for protecting power transformer. This method was based upon current signals that are extracted from various operating conditions like energized transformers and several fault circumstances. Fast Wavelet Transform (FWT) methods are inappropriate for data having oscillating features. An efficient solution involves a simultaneous signal analysis of both in time and frequency domains to distinguish between the inrush current from internal error cases. The wavelet packet transform is an extension of FWT that finely characterizes the signal information together in time and frequency. This work uses the Matlab simulation environment to simulate A 138/13.8 kV transformer. From the outputs it is clear that the suggested method of protection was stable, dependable and fast in differentiating among magnetizing inrush currents and internal faults.

Methodology

In this work, the ANN, wavelet based ANN and FA-ANN methods are elaborated.

Artificial Neural Network (ANN): A network or a mathematical model with several nonlinear artificial neurons running parallel to each other is known as Artificial Neural Network (ANN). ANN can be generated as either single or double layered. In 1943, McCulloch & Pitts developed the idea of artificial neurons. Though it was developed long ago, only after the development of BP method, there has been a major increase in the application of ANN. With drastic improvement, ANN is able to solve complex problems which are otherwise not possible with conventional methods as the quality and the quantity of the data is restricted. The ANN models have specific features that are significantly applicable for dynamic nonlinear system modelling and are referred as “black box” models. This method when compared to other conventional models has the biggest advantage of not requiring mathematical expression of the process that is under consideration. The application of ANN in hydrology varies from actual-time to occasion-based modelling.

The Multi-Layer Perceptron (MLP) is trained by using BP protocol and is a well-known ANN architecture in hydrological modelling. The input layer, one or other hidden layers with computation nodes as well as an output layer form a MLP network. The actual input and output variables nature determine input as well as output nodes number. Based on the problem nature, the modeller uses trial and error scheme for determining the hidden nodes number. The given input signal travels in a forward direction through the network. The signal is processed by the hidden and output node. The process includes multiplying the signal with a weight and then the values are added up to pass them by a non-linear transfer function to generate the output. The NN uses gradient descent technique to alter the chosen weights of random nodes used for the training process. The errors among actual output and target values are used as basis to modify these weights. This process is termed as learning or training. The process is stopped when there are minimal errors or when another terminating criterion is encountered.

The basic structure of ANN is given in figure 1 as below. There are two major categories for the learning methods of adaptive NNs:

Supervised learning is carried out with an external teacher. Here the output unit is directed to produce the desired response for a. given input signal. Global information is sought in the learning course. Error-correction learning, reinforcement learning as well as stochastic learning are the theory behind supervised learning.

The major issue is error convergence in supervised learning, which means, the error among desired and computed values must be minimum. The main purpose is to minimise the error values by identifying the appropriate set of weights. Least Mean Square (LMS) convergence is a commonly used popular method to many learning concepts.

Unsupervised learning employs no exterior teacher and requires only local facts. This learning is termed as self-organization. It can organize its data that is to be in the network and it can detect their developing collective properties. Hebbian learning and competitive learning are included in unsupervised learning.

To Artificial Neurons from Human Neurons previous learning aspect distress the difference of separate phase or not in network training and operation phase. When the two phases namely learning and operation is different than NN learns off-line. While learning online the NN tends to learn and operate at the similar time. Generally, supervised learning is an off-line process, where unsupervised learning is an on-line process.

The following are the four stages of training algorithm of BP[15]:

- Initialization of weights
- Feed forward

- BP of errors
- Weights as well as biases are updated.

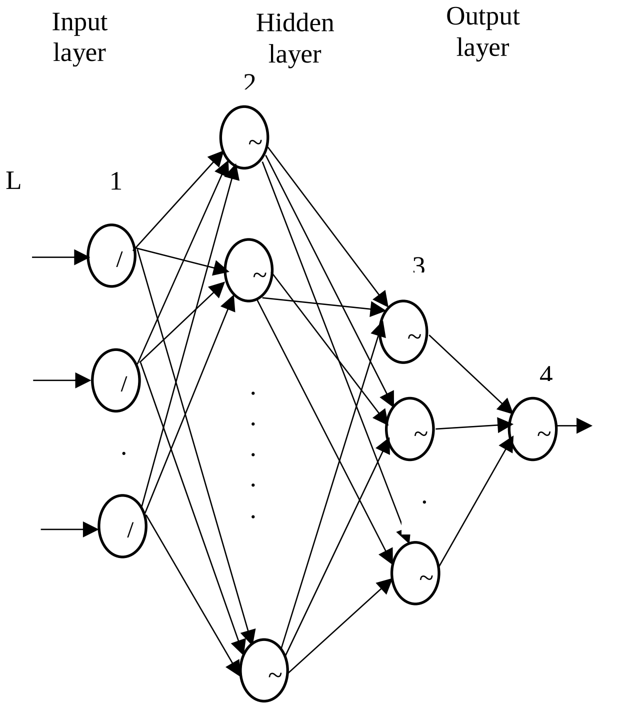

Figure 1: Basic Structure of ANN

In the first stage the random variables are designated to the initialized weights. In the feed forward phase every input unit (X_i) obtains an input signal as well as sends that signal to all hidden units $z_1......z_p$. Then hidden unit will evaluate the activation function as well as transmits the signal z_j to all output units. The output unit assesses the activation function to create the response of the network for the provided input pattern. In BP of errors, every output unit makes a comparison of compound activation functions y_k with target values t_k for determining the related fault for the pattern with the unit. On the basis of the error, the factor $\delta_k (k = 1,..., m)$ is calculated as well as utilized for distributing errors at output unit y_k back to every unit in the earlier layer. Likewise, the factor $\delta_j (j = 1,..., p)$ is calculated for all hidden units z_j. In the final phase, the δ factor as well as activation are used to update the weight and biases. The step by step protocol is given thus.

Initialization of weights
Step 1: Weight is set to small arbitrary value between arbitrary values between 0 and 1.

Step 2: If termination criterion is false, steps 3-10 are iterated.
Step 3: For every training pair steps 4-9 are carried out.

Feed forward
Step 4: All input nodes receive the input signal xi as well as transmit it to every node in the previous layer, that is, the hidden units.

Step 5: All hidden units $(z_j, j = 1,..., p)$ sum the weighted input signals in equation (1):

$$z_{in_j} = V_{o_j} + \sum_{i=1}^{n} (x_i . V_{i_j}) \tag{1}$$

Utilizing the activation function in (2):

$$Z_j = f(z_{in_j}) \tag{2}$$

And this signal is sent to every unit in the previous layer, that is, to output units.

Step 6: Each output unit $y_k (k = 1,..., m)$ sums its weighted input signals in equation (3):

$$y_{in_k} = W_{o_k} + \sum_{j=1}^{p} (z_j W_{j_k}) \tag{3}$$

And activation function is applied for computing the output signal in equation (4):

$$Y_k = f(y_{in_k}) \tag{4}$$

BP of errors

Step 7: All output units $y_k (k = 1,..., m)$ obtains a target pattern related to the input pattern. Error information term is computed in equation (5):

$$\Delta_k = (t_k - y_k) f^1 (y_{in_k})$$

(5)

Wherein $f^1 (y_{in_k}) = f(y_{in_k})(1 - f(y_{in_k}))$

Step 8: All hidden units $(z_j, j = 1,..., p)$ sum its delta inputs in equation (6) from units in the previous layer.

$$\delta_{in_j} = \sum_{k=1}^{m} (\delta_j) W_{j_k} \tag{6}$$

The error term is computed as in equation (7):

$$\delta_j = \delta_{in_j} f^1 (z_{in_j}) \tag{7}$$

Where $f^1 (z_{in_j}) = f(z_{in_j})(1 - f(z_{in_j}))$

Weight and biases are updated

Step 9: All output units $y_k (k = 1,..., m)$ update their biases as well as weights (j=0,……..,p)
The weight correction term is provided in in equation (8):

$$\Delta W_{j_k} = n \delta_k z_j \tag{8}$$

and the bias correction term is provided by equation (9-11):

$$\Delta W_{o_k} = n \delta_k \tag{9}$$

Hence

$$W_{j_k}(new) = W_{j_k}(old) + \Delta W_{j_k} + m[W_{j_k}(old) - W_{j_k}(old-1)] \quad (10)$$

and

$$W_{o_k}(new) = W_{o_k}(old) + \Delta W_{o_k} \quad (11)$$

All hidden units $(z_j,\ j = 1,...,p)$ update their biases as well as weights(i=0,….,n)The weight correction term in equation (12):

$$\Delta V_{i_j} = \Delta \delta_j x_i \quad (12)$$

As well as the bias correction term in equation (13-15):

$$\Delta V_{0_j} = \Delta \delta_j \quad (13)$$

Hence

$$V_{i_j}(new) = V_{i_j}(old) + \Delta V_{i_j} + m[V_{i_j}(old) - V_{i_j}(old-1)]$$

(14)

and

$$V_{o_k}(new) = V_o(old) + \Delta V_{o_j} \quad (15)$$

Step 10: The terminating criterion is verified (minimizing errors).

Wavelets based Artificial Neural Network (ANN): In faults and inrush currents of a power transformer, there is occurrence of current signals related to high-speed electromagnetic transients. The current signals are generally nonperiodic as well as have high-frequency oscillations as well as localised impulses overlaid on the power frequency as well as its harmonics. The above features lead to problem with conventional Fourier analysis as this presumes a periodic signal. The disturbances that occur in power system are subjected to transient as well as nonperiodic elements. So Discrete Fourier Transform (DFT) alone is not enough for signal analysis. This drawback can be overcome by applying WT for analysing the transient phenomena in power system. WT has replaced Fourier analysis in many applications as it has the capability to obtain data from transient signals that occur in time as well as frequency domains. The WT functions by expanding the signal by wavelets instead of trigonometric polynomial, generating it using translation time and dilation time (shift in time) as well as dilation (compression in time) of the fixed wavelet function known as the mother wavelet. This wavelet function is restricted to time as well as frequency yield wavelet coefficients at multiple levels. Signals with restricted transient components the is compactly analysed by WT.

The process of extracting features based on wavelet transform is gaining popularity and many investigations are performed to check its applicability in medical images processing, military radar application, power system error protection, and so on. The 2 main features that help wavelet to gain importance in fault detection are time localization capability as well as multi-resolution analysis.

As discussed, WT is used to analyze the transients in power transformer. Discrete Wavelet Transform (DWT) coefficients for a given signal are calculated by giving the resulting data from MAT LAB simulations to MATLAB WT tool. The various mother wavelets are Harr, Daubichies, Coiflet and Symmlet wavelets. An important role is played by the choice of mother wavelet as this detects and localizes the distinct type of fault transients. Further depending on a specific application, the required choice is made. This work deals with methods that are more attention to detect and to analyze lower amplitude, shorter duration, rapid decay as well as oscillating kind of higher frequency current signals. Daubichies wavelet is considered as one among the most well-known mother wavelets that is best suited for wide range of applications.

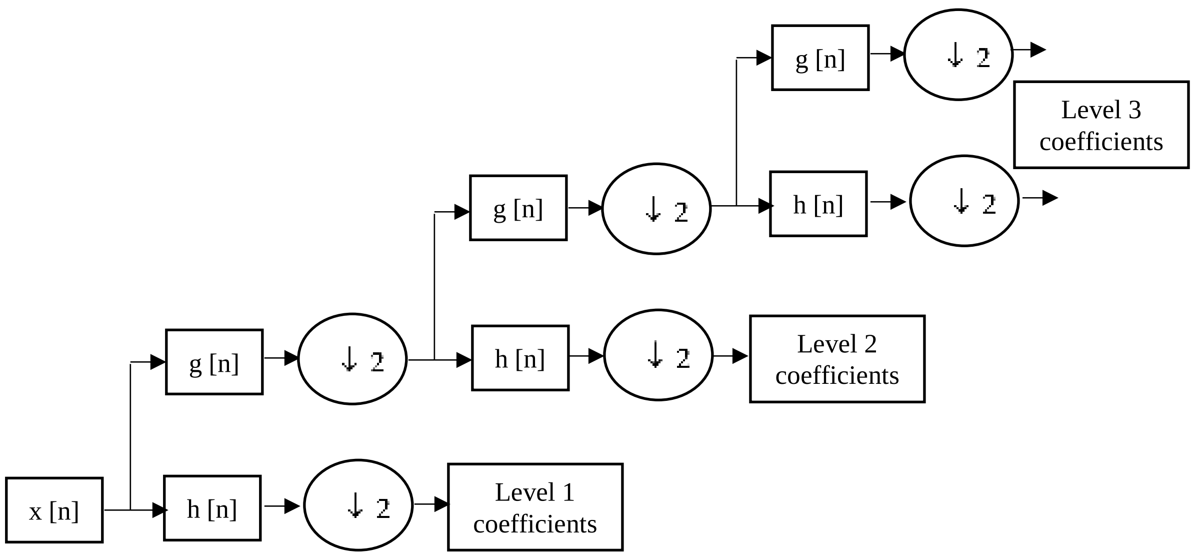

Figure 2: Filter Analysis

Figure 2 demonstrates the execution procedure of a DWT, which represent parameters like x[n] the original signal, h[n] as well as g[n] low- as well as high-pass filters, correspondingly. During the initial phrase the original signal is separated into 2 halves of the frequency bandwidth and it is passed through both high- as well as low-pass filter. The resulting output from low pass filter is then split into half based on frequency bandwidth as well as again sent to the second phase. The process is iterated till the signal is disintegrated to a fixed level.

The signals thus obtained are related the same input signal however have varying frequency bands. The original sampling rate of the signal and the frequency band of each DWT are directly proportional to each other. According to Nyquist's theorem the greatest frequency that a signal can reach is $F_s/2$ Hz if the original signal was sampled at F_s Hz. The first detail obtained would be the frequency from high frequency filter. From this the band of frequencies between as well as would be is obtained in detail 1; likewise, the band of frequencies between as well as would be is obtained in detail 2, etc. If the wavelet decomposition is executed on signal then DWT coefficient of the Level 1 signals are obtained.[16]

Detailed coefficients at level1 in equation (16):

$$d^1[n] = \sum_{k=0}^{N-1} x[k]\, h[n-k] \qquad (16)$$

Approximation coefficients at level 1 in equation (17):

$$a^1[n] = \sum_{k=0}^{N-1} x[k]\, g[n-k] \qquad (17)$$

The maximum, normalised and average values of detailed and approximation coefficients are used to train the NN in equation (18-19):

$$\begin{aligned} d^1_{\max} &= \max imum\ value\ of\ d^1[n] = \max[d^1[n]] \\ d^1_{norm} &= normalisation\ value\ of\ d^1[n] = norm[d^1[n]] \\ d^1_{avg} &= average\ value\ of\ d^1[n] = avg[d^1[n]] \end{aligned} \qquad (18)$$

$$\begin{aligned} a^1_{\max} &= \max imum\ value\ of\ a^1[n] = \max[a^1[n]] \\ a^1_{norm} &= normalisation\ value\ of\ a^1[n] = norm[a^1[n]] \\ a^1_{avg} &= average\ value\ of\ a^1[n] = avg[a^1[n]] \end{aligned} \qquad (19)$$

For each phase, 6 input values so for three phase 18 input values are used to train the NN.

The ANN that is developed is one of the ANNs that derive its inspiration from biological structure of human body. That input map sets into output sets by means of NN. The individual neurons present in every layer generate one normalize output function for every normalized inputs. The considered network structure is multi-layer comprising 1 input, 1 hidden, as well as 1 output layer. The layers belong to feed forward type and are completely linked to one. The results belong to non-linear functions of the input. They are restricted using weights which are calculated in learning procedure. Supervised type of learning process is considered while BP is the learning paradigm.

Firefly Algorithm based Artificial Neural Network (FA-ANN): This work employs FA such that the weights of ANN model are optimized represented as FA-ANN in order to get the most favourable parameter settings to train the network of ANN as well as also reduce the error rate.

The recently developed FA belongs to swarm intelligence methods and it was proposed by Yang in 2008. It is a sort of stochastic algorithm, nature-inspired, metaheuristic protocol which is capable of resolving toughest optimization problems and also NP-hard problems. It is stochastic which a sort of randomization is to search for a solution set. It draws inspiration from fireflies flashing lights in nature. The word heuristic refers to the finding of resolutions by trial and error'. By this there is no assurance for attaining optimal solution within a considerable time.

The unique feature of fireflies namely the flashing properties may be briefed into three rules as below:

- The fireflies are attracted to other fireflies regardless of its sex.
- Attractiveness as well as brightness are directly proportional to one another. For instance, for a given pair of fireflies, the fly with lower brightness will move toward the brighter one. As the distance increases between the two fireflies, its brightness decreases. The brightest firefly will move arbitrarily since there is no other fly to attract it
- The purpose function is to be optimized the affects or determines the brightness of a firefly.

The two main problems in FA are: light intensity variation and formulating attractiveness. To simplify this, it could be stated as the attractiveness of the firefly relies upon the brightness that related to programmed objective function.

In the easy case for most of the optimization problems, the brightness I of a firefly at a specific location x may be selected as $I(x) \alpha f(x)$. Additionally, the attractiveness β is relative; sit depends on the judgement by some other fireflies. Hence, it will diverge the distance r_{ij} among firefly i as well as j. As well, light intensity reduces with increase in distance from its source, while light is absorbed in the media, so the attractiveness must be allowed to differ with the degree of absorption. In the most basic format, the light intensity I(r) differs as per the inverse square law $I(r) = \frac{I_s}{r^2}$ wherein I_s the source intensity. In a medium of fixed absorption light coefficient γ , I differ with r in equation (20). It could be written as:

$$I = I_0 e^{-\gamma r} \tag{20}$$

Here I_0 refers to original light intensity.

As the attractiveness of firefly is proportional to the light intensity viewed by adjacent fireflies, then the attractiveness β of a firefly is defined by equation (21):

$$\beta = \beta_0 e^{-\gamma r^2} \tag{21}$$

Wherein β_0 is the attractiveness at r = 0.

The pseudocode of FA-ANN is given below. In the first step, For a given problem, the initial population of required candidate solutions is generated in this case it is the weights of ANN. The most attractive firefly is identified by calculating the light intensity for each fire fly. Later the attractiveness and the distance for each fly are evaluated. Using this, the flies move toward the most attractive fly in the search space. In the last step, the attractive fly moves arbitrarily in the search space. The procedure continues until it creates maximum generation number and considered as stopping criterion. [17]

Pseudo code for FA-ANN

Begin
Generate the initial solution randomly
Evaluate every individual in the population $f(x)$ *onthebasisof error rate*
Find the best solution from the population
While (stopping criterion satisfied)
For $i = 1$ *to n do*
For $j = 1$ *to n do*
If $(f(x_j) < f(x_i))$
Calculate attractive fireflies by Eq.22
Calculate the distance between each fireflies i and j by Eq.23
Move all firefly (x_i) *to the best solution* (x_j) *by Eq.24*
End if
End for j
End for i
Moves best sloution randomly by Eq.25
Discover the best solution from the new population
End while
Return best (TP), (TN), (FP) and (FN)
End of the algorithm.

Generally the error rate is used in calculating time series value classification. The error rate is measured based upon True Positive (TP), True Negative (TN), False Positive (FP) as well as False Negative (FN). The process begins with an initial population of erratically created individuals. Using Eq (1) all individual value is evaluated and the most suited solution is chosen. In FA, the attractiveness function format of a firefly is illustrated by the equation (22):

$$\beta(r) = \beta_0 \exp(-\gamma r^2) \tag{22}$$

here

r = distance between any 2 fireflies
β_0 = initial attractiveness at r = 0 & initialized to 1
γ = absorption coefficient that controls the reduction of light intensity and is set to 1.

The distance between any 2 fireflies i as well as j, at positions xi & xj, correspondingly, may be expressed Cartesian or Euclidean distance as in equation (23):

$$r_{ij} = \| x_i - x_j \| = \sqrt{\sum_{k=1}^{d} (x_{i,k} - x_{j,k})^2} \tag{23}$$

Wherein, d represents the dimensionality of the particular issue.
The movement of a firefly i that is attracted by a brighter firefly j is expressed as (24-25):

$$x_i = x_i + \beta_0 * \exp(-\gamma r_{ij}^2) * (x_j - x_i) + \alpha * \left(rand - \frac{1}{2} \right) \tag{24}$$

$$x_i = x_i + \alpha * \left(rand - \frac{1}{2} \right) \tag{25}$$

In Eq. (24), the first term represents the existing firefly position, the 2nd term gives the attractive firefly, in light intensity by means of adjacent fireflies and the third corresponds to the arbitrary part, as it requires the brighter ones. The α coefficient is a random parameter defined for the problem of interest, whereas rand is an arbitrary number generator reliable distributed in the space (0, 1). In Eq. (25), best candidate movement is done randomly.

Results and Discussion

In this section, ANN, FF-ANN, wavelet ANN and wavelet FF-ANN methods are evaluated. A generic 750 MVA, 27/420KV, Δ/Y power transformer linked between a 27KV source at the sender end as well as a 420KV transmission line linked to an infinite bus power system at the receiver end. The generated data was utilized by MATLAB to validate the suggested method. The neural networks are trained for several training patterns of fault as well as inrush conditions. The target for internal fault currents are trained to be one while those for inrush are trained to be zero. The figures 3 & 4 show the fitness and mean square error.

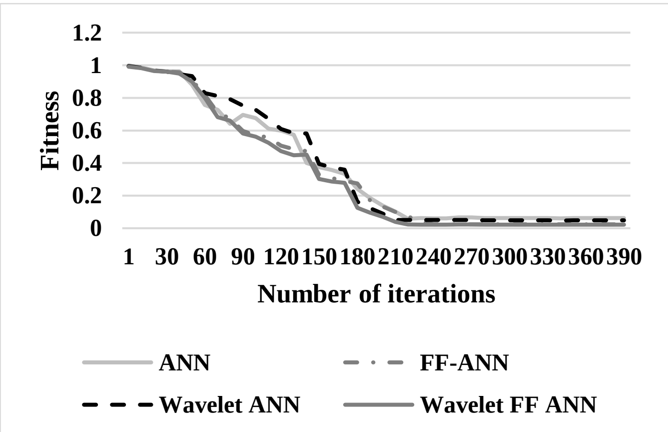

Figure 3: Fitness

From the figure 3, it can be observed that the ANN has higher average fitness by 7.5% for FF-ANN and by 13.27% for Wavelet FF ANN. The ANN has lower by average fitness by 1.93% for Wavelet ANN when compared with various number of iterations respectively.

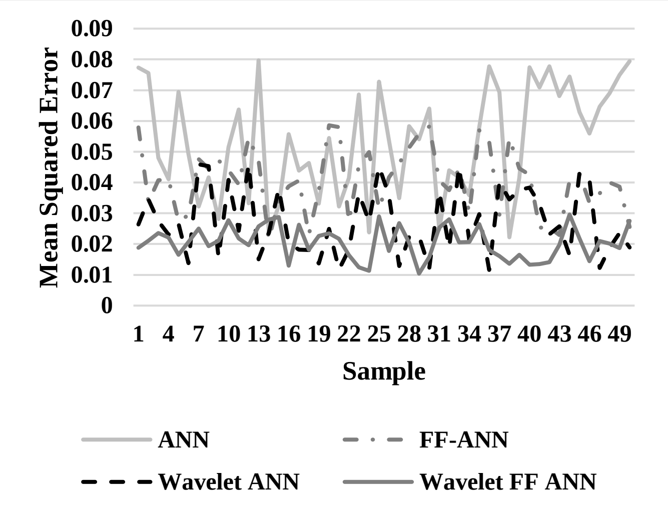

Figure 4: Mean Square Error

From the figure 4, it can be observed that the ANN has higher average mean square error by 26.94% for FF-ANN, by 65.39% Wavelet ANN and by 89.17% for Wavelet FF ANN when compared with various number of samples respectively.

Conclusion

In this work, the BPNN protocol is the most popular and widely accepted procedure to optimize the FFNN training. The traditional BPNN algorithm has certain setbacks like getting trapped in local minima and low rate of convergence. Optimization of complex problems is achieved by using meta-heuristic algorithms as it provides derivative-free solution. FA is the modern meta-heuristic search algorithm that is applicable to train BPNN such that is achieves quick convergence rate and minimises the training error. FA is executed to train the NN by increasing the objective function successfully. In order to achieve high accuracy, the number of training data, population of fireflies and iteration number must be significantly high. According to results the ANN has higher average fitness by 7.5% for FF-ANN and by 13.27% for Wavelet FF ANN. The ANN has lower by average fitness by 1.93% for Wavelet ANN when compared against various numbers of iterations respectively.

References

1. Sharma N.K. and Kumar Sahu N., Protection of Power Transformer By Advanced Differential Protection Scheme, *International Journal of Recent Scientific Research*, **6(11)**, 7427-7431 **(2015)**

2. Pavithra M.A. and Loganathan N., Artificial Neural Network Approach for Discriminating Various faults in Transformer Protection, *International Journal of Research and Engineering*, **2(4)**, 83-87 **(2015)**

3. Soundiraraju P. and Loganathan N., Wavelet Transforms Based Power Transformer Protection from Magnetic Inrush Current, *Indonesian Journal of Electrical Engineering and Computer Science*, **14(3)**, 381-387 **(2015)**

4. Zellagui M. and Chaghi A., Distance Protection Settings Based Artificial Neural Network in Presence of TCSR on Electrical Transmission Line, *International Journal of Intelligent Systems and Applications*, **4(12)**, 75 **(2012)**

5. Bigdeli M., Vakilian M. and Rahimpour E., A probabilistic neural network classifier-based method for transformer winding fault identification through its transfer function measurement, *International Transactions on Electrical Energy Systems*, **23(3)**, 392-404 **(2013)**

6. Gholizadeh S., Performance-based optimum seismic design of steel structures by a modified firefly algorithm and a new neural network, *Advances in Engineering Software*, **81**, 50-65 **(2015)**

7. Tripathy M., Maheshwari R.P. and Verma H.K., Power transformer differential protection based on optimal probabilistic neural network, *IEEE transactions on power Delivery*, **25(1)**, 102-112 **(2010)**

8. Nayak J., Naik B. and Behera H.S., A novel nature inspired firefly algorithm with higher order neural network: Performance analysis, *Engineering Science and Technology, an International Journal*, **19(1)**, 197-211 **(2016)**

9. Suja S. and Jerome J., Wavelet based signal analysis in three phase circuits using neural network, *Computer Applications in Engineering Education*, **18(3)**, 537-546 **(2010)**

10. Abdallah J.M. and Al-Zyoud A.R., Voltage and Reactive Power Control Simulations Using Neural Networks, *IJSSST*, **10(5)**, 64-72 **(2009)**

11. Nandy S., Sarkar P.P. and Das A., Analysis of a nature inspired firefly algorithm based back-propagation neural network training, *International Journal of Computer Applications*, **43(22)**, 8-16 **(2012)**

12. Shah A.M. and Bhalja B.R., Discrimination between internal faults and other disturbances in transformer using the support vector machine-based protection scheme, *IEEE Transactions on Power Delivery*, **28(3)**, 1508-1515 **(2013)**

13. Rasoulpoor M. and Banejad M., A correlation based method for discrimination between inrush and short circuit currents in differential protection of power transformer using discrete wavelet transform: theory, simulation and experimental validation, *International Journal of Electrical Power & Energy Systems*, **51**, 168-177 **(2013)**

14. Ramesh K. and Sushama M., Power transformer protection using wavelet packet transform, In Computer Communication and Informatics (ICCCI), 2014 International Conference, IEEE, 1-7 **(2014)**

15. Nanda S.K., Application of Neural Network for Transformer Protection (Doctoral dissertation) **(2013)**

16. Srinivasulu K., Vishwakarma D. and Sinha A., Wavelet and ann based differential protection of power transformer, 16th National Power Systems Conference, 15th-17th December **(2010)**

17. Alweshah M., Firefly algorithm with artificial neural network for time series problems, *Research Journal of Applied Sciences, Engineering and Technology*, **7(19)**, 3978-3982 **(2014)**.