A Hybrid Method Based on Optimized Neuro-Fuzzy System and Effective Features for Fault Location in VSC-HVDC Systems

REZA ROHANI AND AMANGALDI KOOCHAKI
Department of Electrical Engineering, Aliabad Katoul Branch, Islamic Azad University, Aliabad Katoul 4941793451, Iran
Corresponding author: Amangaldi Koochaki (koochaki@aliabadiau.ac.ir)

ABSTRACT
One of the most important issues with HVDC systems are the occurrence of various faults that can lead to considerable electrical power losses, serious damage to expensive equipment and huge financial losses. Hence, it is highly required to design an accurate and automatic fault location method in HVDC systems for maintaining uninterrupted supply of energy and protecting sensitive equipment such as rectifiers and inverters. Accordingly, this paper proposes a new hybrid system based on adaptive neuro-fuzzy inference system (ANFIS) with optimal parameters and Hilbert-Huang (HH) transform for fault location in voltage sourced converter-HVDC (VSC-HVDC) systems. The proposed fault location method consists of three major sections. In the first section, HH transform is applied to extract new features from current signal. In the second part, ANFIS uses the extracted features to estimate the fault location in transmission lines. Learning algorithm determines the accuracy and efficiency of each machine-learning algorithm. In the third section of the developed system, enhanced version of particle swarm optimization (PSO) algorithm named chaotic dynamic weight PSO (CDWPSO) algorithm is implemented as learning algorithm to train the ANFIS. The developed fault detection and location system was tested on a VSC-HVDC system with 250 km length and the obtained results using MATLAB simulations have shown that combination of new features, and CDWPSO-based ANFIS has high accuracy in fault detection and location in VSC-HVDC systems. High fault location accuracy, robust performance of neuro-fuzzy system, optimal training of ANFIS, extraction of novel effective features from current signal and fault location only with six features are the main contribution of the developed system.

INDEX TERMS
HVDC transmission, feature extraction, fault location, fuzzy neural networks, Particle swarm optimization.

I. INTRODUCTION
High-voltage direct current transmission lines or HVDC systems are an effective option for transmission of huge amounts of electrical energy from power plants that are located in remote areas such as offshore wind farms [1], [2]. These systems have crucial function in modern power grids due to their potential of long distance transmission and large capacity [3]–[5]. Over the past few decades, the appearance of semiconductor-based power devices, especially insulated-gate bipolar transistors (IGBT), have led to the development of voltage source converters (VSC) for transmission of electrical power.

The VSC-based HVDC system (VSC-HVDC) has couple of advantages compared to common HVDC systems such as self-reliant control of reactive and active power, reversal of direction of power with no need for changing the polarity that is very beneficial in multi-terminal DC systems, voltage support at the converter node that is very advantageous for stability improvement, and no demands or fast communication between converter terminals [6].

The VSC-HVDC systems have widespread applications in the fields of modern power networks such as offshore wind farms, renewal of urban power networks, and island electric power supply and many other applications [7]. Power loss, serious damage to expensive and sensitive equipment such as converters and rectifiers, and huge financial losses are some destructive consequences of an undetected fault within
the VSC-HVDC systems. Due to long transmission lines, varied topography, and harsh environmental condition, fault location in VSC-HVDC systems become an extremely tough task for electrical engineers and experts. Thus, it is of utmost importance to design and develop an accurate and automatics fault location schemes for VSC-HVDC systems [8].

Given the importance of the issue, several studies have been conducted in recent years by researchers. Schemes based on differential protection are frequently utilized for fault location and HDVC systems protection. Generally, differential protection based approaches offer a swift response time and a precise operation. Furthermore, there is no need for any particular hardware for signal sampling with high sampling rate. However, needing for communication link between sending/receiving terminals makes these methods more expensive and less dependable as their accuracy rely on the communication media [9]–[12].

Fault detection and location schemes which exploit the travelling wave are the most popular approach that has been investigated by researchers [13]–[20]. Travelling wave-based methods, by knowing the wave velocity and wave-head arrival time seen from the contrary terminal, the fault location can be assessed with high accuracy. The most important benefits of travelling wave-based methods are their high accuracy and swift response time. In travelling wave-based methods, the discriminative characteristics of the these waves could be extracted from the current signal (I) and voltage signal (V) by using different feature extraction techniques such as singular value decomposition (SVD), wavelet decomposition transform (WDT), fast Fourier transform (FFT), and autoregressive method (ARM). The extracted features can be used as the input of machine learning algorithms such as support vector machine (SVM) with different kernel functions, multilayer Perceptron neural networks (MLPNN), support vector regression (SVR), fuzzy logic systems (FLS), radial basis function neural network (RBFNN), random forest (RF), and other machine learning algorithms. In general, these features can be grouped into frequency domain and time domain features [21], [22].

According to published research works regarding the fault location in VSC-HVDC systems, effective feature extraction method and proper type of machine learning algorithm have huge impact on fault location accuracy. In this article, application of adaptive neuro-fuzzy inference system (ANFIS) with optimal parameters and effective time-frequency-energy domain features achieved by Hilbert-Huang (HH) transform is proposed for fault location purpose in VSC-HVDC systems. The HH transform is a novel non-stationary signal analysis method, which employs the concept of instantaneous frequency and can capture informative and effective features from frequency and time domain during fault occurrence. ANFIS employ a combination of linguistic knowledge and numeric information to analyze a process. The ANFIS has several advantages such as robust performance, ability to capture the nonlinear structure of a process, adaptation capability, and fast learning capacity. The ANFIS has been successfully implemented for fault detection, function approximation, time series forecasting, control, and nonlinear processes modeling [23]–[27].

The efficient design of ANFIS-based schemes for fault location in VSC-HVDC systems need for accurate parameter training for further accuracy and strengthen performance. Combination of least squares (LS) method and back propagation (BP) is the commonly utilized learning algorithm for ANFIS. Both of LS and BP are derivative-based methods, which have weak performance in ANFIS training [28], [29]. In the developed method, we used enhanced version of particle swarm optimization (PSO) algorithm named chaotic dynamic weight PSO (CDWPSO) algorithm as learning algorithm. The CDWPSO is a new metaheuristic algorithm that its high accuracy in solving complicated and nonlinear optimization problems has been proved [30].

II. BASCI CONCEPTS

In the proposed method, combination of HH transform, ANFIS and CDWPSO algorithm are utilized. In this section, these algorithms are introduced.

A. ANFIS

ANFIS is a powerful machine-learning algorithm that combines artificial neural networks concept with fuzzy logic. ANFIS is consisted of five main layers. The second layer parameters known as Antecedent parameters and parameters between third and fourth layer known as Conclusion parameters determine the performance of ANFIS. Standard ANFIS uses combination of LS and BP learning algorithm for tuning these parameters [33]. ANFIS utilizes K-Means clustering algorithm for fuzzy rules extraction. According to computed cluster centers for input data and Euclidean distance of input samples from cluster centers, fuzzy rules are obtained automatically.

B. OPTIMIZATION ALGORITHM

PSO is a nature-based stochastic optimization algorithm that is motivated from social behavior of fishes and birds in a big group. The PSO algorithm is very popular because of its simplicity, easy implementation and acceptable accuracy. However, this algorithm has some basic weaknesses especially in local search phase. In order to overcome this drawback, CDWPSO algorithm is presented in [30]. In the CDWPSO algorithm, a dynamic weight and a chaotic map are utilized to modify the global and local search capabilities. More detail regrading CDWPSO algorithm can be found in [30].

C. HH TRANSFORM

The study and decomposition of nonstationary and nonlinear signals into orthogonal constituent is a problem regarding distinct fields of the electrical engineering. The wavelet and Fourier analysis represent traditional decomposition approaches which have been applied to nonstationary and nonlinear signals processing with success. Recently, a new adaptive method for time-frequency analysis known as HH
transform has been introduced. The HH transform is able to compute the instantaneous amplitude and frequency of nonstationary and nonlinear signals like current and voltage signals of HVDC systems. Unlike WDT and FFT, the HH transform does not involve the concept of frequency resolution or time resolution, but introduces the concept of instantaneous frequency. More details regarding HH transform can be found in [32], [33].

III. PROPOSED METHOD

This study puts forward a hybrid method for fault locating in VSC-HVDC transmission line using one terminal current signal and optimized neuro-fuzzy system. The proposed fault location system consists of three major modules as shown by Figure 1. The proposed fault location system makes full exploit of the time, frequency and energy information to grab the fault’s main attributes. In the second module, ANFIS uses the extracted features to estimate the fault location in transmission lines. Learning algorithm determines the accuracy and efficiency of each machine-learning algorithm. In the third module of the proposed method, CDWPSO algorithm is used as learning algorithm to train the ANFIS with highest accuracy.

In Figure 2, a VSC-HVDC system and its main parts including converter station, HVDC system pole, AC and DC filter, smoothing reactor, DC capacitor, converter reactor, reactive power source, converter transformers, DC transmission lines etc. are illustrated. The faults that happen in h-j section are external faults, and faults that occur in j-k section are internal faults. In the AC section, decoupling of signals and computing zero-mode and aerial-mode are required for fault location purpose. The different modes of current signal, including 0-mode and 1-mode, could be computed in term of \( I_p \) and \( I_n \) using Eq. (1):

\[
\begin{pmatrix}
I_0 \\
I_1
\end{pmatrix} =
\begin{pmatrix}
1 & 1 \\
\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}
\end{pmatrix} 
\begin{pmatrix}
I_p \\
I_n
\end{pmatrix}
\]

(1)

In Eq. (1), \( I_0 \), \( I_1 \), \( I_p \) and \( I_n \) indicate the 0-mode current, 1-mode current, positive pole current and negative pole current respectively.

In the simulations, the propagation coefficient \( \gamma_i \) and surge impedance \( Z_i \) are expressed by Eq. (2) and (3):

\[
Z_i = \sqrt{R_i + j\omega L_i - G_i + j\omega C_i}
\]

(2)

\[
\gamma_i = \alpha_i + j\beta_i
\]

(3)

In Eq. (2), \( i \) indicates i-th mode, \( R_i \) shows resistance, \( L_i \) indicates inductance, \( C_i \) is capacitance and \( G_i \) indicates the conductance. In Eq. (3), \( \alpha_i \) and \( \beta_i \) are attenuation and distortion constants respectively. In addition, traveling wave velocity can be calculated using following equation:

\[
v_i = \frac{\omega}{\beta_i}
\]

(4)

For our VSC-HVDC test system with 250 km length, \( \beta_0 \) is 0.0431 and \( \beta_1 \) is 0.0143. When we compare \( \beta_0 \) and \( \beta_1 \), it can be seen that \( v_1 \) is substantially higher than \( v_0 \). In order to enhance the vividness of the stated formulas, Figure 3 shows the extension of each mode (\( I_0 \) and \( I_1 \)) component during a fault occurrence.

In VSC-HVDC transmission lines, the amount of natural frequency is only connected to velocity of traveling wave and distance of fault. In real VSC-HVDC systems, huge shunt capacitors are installed on both terminals of the DC lines. As a result, the fault distance can be stated as follow:

\[
l = \frac{v}{2f_1}
\]

(5)

In Eq. (5), \( l \) represents the fault distance and \( f_1 \) indicates the primary natural frequency constituent. Precisely extraction of natural frequency and velocity calculation of traveling wave determine the accurate fault location in VSC-HVDC.
systems. Hence, it is extremely challenging task to estimate the distance of fault simply by utilizing natural frequency.

The rate of variance contribution could be proper index for measuring the relative importance of the main constituent. Accordingly, in order to distinguish the external and internal area, high-frequency variance contribution rate (HVCR) of intrinsic mode function (IMF) constituents are utilized in the developed method.

$$D_i = \frac{1}{N} \sum_{k=1}^{N} |c_i(k \Delta t)|^2 - \left[ \frac{1}{N} \sum_{k=1}^{N} c_i(k \Delta t) \right]^2$$  \hspace{1cm} (6)

$$M_i = \frac{D_i}{\sum_{i=1}^{n} D_i}$$  \hspace{1cm} (7)

$$HVCR = 100 \times M_i$$  \hspace{1cm} (8)

In these equations, $D_i$ and $c_i$ represent the variance and coefficients of i-th IMF respectively. In addition, $\Delta t$ represents the sampling rate and time interval between two sampling measure.

For an intelligent fault location scheme, it is an extremely essential issue that the developed model selects the informative and educational features to learn the process and make an accurate model between input-output pairs. According to conducted researches in the field of signal processing, HH transform can capture the information of frequency and time domain. The developed method uses the sampled current signal at single-ended DC bus to locate unipolar short-circuit faults. It is a well-known fact that $I_0$ and $I_1$ have different transmission velocities, because of the faster attenuation of $I_0$. Therefore, using instantaneous frequency analysis of the first IMF component, the time delay can be achieved. Subsequently, the boundary spectrum of $I_1$ is as the characteristic frequency connected to the distance of fault. Figure 4 illustrate the wave shape of extracting temporal data. In addition, Figure 5 shows the boundary spectrum of HH transform.

The developed model is simulated by MATLAB in order demonstrate the recognizing capability of the HVCR in different fault distances. The obtained result shown by Figure 6 confirms that fault area is connected to HVCR value. In this figure, the horizontal axis shows the fault distances in Km and vertical axis shows the value of HVDC in percent. A proper criterion for separating the external and internal sections could be defined via simulation as faults take place in the vicinity of capacitors. Accordingly, 0.004 with additional margin is selected for HVCR$\_set$.

High-frequency energy ($\omega_1$) of $I_1$ and $I_0$, and energy attenuation coefficient ($\lambda$) of $I_1$ and $I_0$ are utilized for validating the fault distances as well as characteristic frequency of $I_1$ and time difference between $I_1$ and $I_0$. The $\lambda$ and $\omega_1$ are defined by Eq. (9) and Eq. (10):

$$\omega_1 = \frac{1}{N} \sum_{k=1}^{N} |c_i(k \Delta t)|^2$$  \hspace{1cm} (9)

$$\lambda = \frac{\omega_{\max}}{\omega_{\min}}$$  \hspace{1cm} (10)

In these equations, for $I_0$, $\omega_{\max}$ is $\omega_2$, for $I_1$, $\omega_{\max}$ is $\omega_4$ and $\omega_{\min}$ is the high-frequency energy $\omega_1$.

For ANFIS training, the developed method take the negative pole grounded fault instances that take place on the transmission line for each five km. Each fault sample only use current signal to achieve required features. According
FIGURE 6. The HVCR value of $I_1$ at side M in various distances for internal fault occurrence at negative pole.

FIGURE 7. Extracted feature (time difference between $I_0$ and $I_1$) for occurred faults in various locations.

to obtained results from MATALB simulations, it is more appropriate to normalize the inputs in the range of [0, 1], in order to improve the generalization ability and remove the differences of multiple input cases. Eq. (11) is utilized for normalization purpose:

$$f_{\text{new}} = \frac{f - f_{\text{min}}}{f_{\text{max}} - f_{\text{min}}}$$  \hspace{1cm} (11)

For the purpose of better explaining the validity of the proposed features, Figures 8 to 13 show the relationship between the fault distance and extracted feature. In these figures, the horizontal axis shows the fault distances in Km and vertical axis shows the value of feature. As illustrated by Figure 7, the time differences between $I_0$ and $I_1$ (feature 1) are approximately linear to the fault distance and characteristic frequency of $I_1$ (feature 2) decrease when the fault distance increases (Figure 8).

From Figures 9 to 12, it can be seen that the energy attenuation coefficient $\lambda$ (features 3 and 4) and high-frequency energy $\omega_1$ (features 5 and 6) have normal fluctuations with various fault distances. It can also be seen that the time differences between $I_0$ and $I_1$ was used for coarse global refinement, and characteristic frequency of $I_1$, the energy attenuation coefficient $\lambda$ as well as high-frequency energy $\omega_1$ were used for fine local refinement.

In the developed method, six extracted features by HH transform are implemented as the inputs of optimized ANFIS. The extracted features by HH transform are as follow:

F1: Time difference between $I_0$ and $I_1$
F2: Characteristic frequency of $I_1$
F3: Energy attenuation coefficient ($\lambda$) of $I_0$
F4: Energy attenuation coefficient ($\lambda$) of $I_1$
In the second module, ANFIS is used as intelligent estimator. The issue of learning algorithm type and its convergence speed is an extremely important topic in building ANFIS model. In ANFIS training, antecedent parameters (found in the second layer) and conclusion parameters (found between third and fourth layers) are selected using learning algorithm. In this study, the mentioned parameters are selected by using CDWPSO algorithm. In the proposed method, we used mean square error (MSE) as fitness function. The mathematical representation of MSE is as follow:

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} (d_i - O_i)
\]  

(12)

In Eq. (12), \(O_i\) is the true output value of i-th training samples or target \(d_i\) is the expected value, and N is the number of samples.

F5: High-frequency energy (\(\omega_1\)) of \(I_0\)
F6: High-frequency energy (\(\omega_1\)) of \(I_1\)

In the second module, ANFIS is used as intelligent estimator. The issue of learning algorithm type and its convergence speed is an extremely important topic in building ANFIS model. In ANFIS training, antecedent parameters (found in the second layer) and conclusion parameters (found between third and fourth layers) are selected using learning algorithm. In this study, the mentioned parameters are selected by using CDWPSO algorithm. In the proposed method, we used mean square error (MSE) as fitness function. The mathematical representation of MSE is as follow:

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} (d_i - O_i)
\]  

(12)

In Eq. (12), \(O_i\) is the true output value of i-th training samples or target \(d_i\) is the expected value, and N is the number of samples.

In the second module, ANFIS is used as intelligent estimator. The issue of learning algorithm type and its convergence speed is an extremely important topic in building ANFIS model. In ANFIS training, antecedent parameters (found in the second layer) and conclusion parameters (found between third and fourth layers) are selected using learning algorithm. In this study, the mentioned parameters are selected by using CDWPSO algorithm. In the proposed method, we used mean square error (MSE) as fitness function. The mathematical representation of MSE is as follow:

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} (d_i - O_i)
\]  

(12)

In Eq. (12), \(O_i\) is the true output value of i-th training samples or target \(d_i\) is the expected value, and N is the number of samples.

In the second module, ANFIS is used as intelligent estimator. The issue of learning algorithm type and its convergence speed is an extremely important topic in building ANFIS model. In ANFIS training, antecedent parameters (found in the second layer) and conclusion parameters (found between third and fourth layers) are selected using learning algorithm. In this study, the mentioned parameters are selected by using CDWPSO algorithm. In the proposed method, we used mean square error (MSE) as fitness function. The mathematical representation of MSE is as follow:

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} (d_i - O_i)
\]  

(12)

In Eq. (12), \(O_i\) is the true output value of i-th training samples or target \(d_i\) is the expected value, and N is the number of samples.
three membership functions are utilized for each input and eight fuzzy rules are obtained. Each Gaussian MF has two parameters including sigma (σ) and center (C). Therefore, we have $6 \times 3 = 18$ antecedent parameters and $8 \times 7 = 56$ conclusion parameters. Thus, $(6 \times 3) + (8 \times 7) = 74$ unknown parameters are optimized using CDWPSO to build an ANFIS with the highest accuracy.

In the standard ANFIS, combination of LS and BP (LS+BP) is used as the learning algorithm. Table 1 shows the antecedent parameters of second feature (Characteristic frequency of $I_1$) selected by the CDWPSO and LS+BP. These membership functions build according on selected parameters are shown in Figure 13. According to Figure 13, it can be observed that there is remarkable difference between membership functions built by the LS+BP and CDWPSO. These parameters have significant effect on ANFIS performance. Therefore, ANFIS will have different performance using different parameters.

In order to investigate the accuracy of the developed method in fault location at unknown distances, 14 randomly unipolar ground fault cases are considered. In order to assess the HVCR criterion, two external fault instances are considered in the simulation and experiments. The results achieved by the proposed method, optimally trained ANFIS (CDWPSO-based ANFIS) and six extracted features using HH transform, are listed in Table 2. In this table, NP means negative pole and PP means positive pole.

For comparison with standard ANFIS, the obtained results using LS+BP-based ANFIS are also listed in this table. In Figures 14 and 15, the accuracy of different methods are compared. It can be seen that input type, current signal or extracted features, and learning algorithm, LS+BP or CDWPSO, have significant impact on accuracy of ANFIS and fault location estimation. The main contribution of the proposed method is improving the performance of ANFIS by using new learning algorithm, improving the robustness of ANFIS, extracting new effective features from current signal and enhancing the fault location accuracy in VSC-HVDC systems.

### C. EFFECT OF NOISE

In a real VSC-HVDC system, there are many electromagnetic disturbances and noise resources that affect the measurements. Therefore, developed method should have ability to handle available noises and locate the fault with high accuracy. In this subsection, the performance of developed method is investigated under different noise levels and obtained results are listed in Tables 3 and 4. For this purpose, different signal-to-noise (SNR) are considered, and the value of MSE for 14 test samples is listed. In this experiment, CDWPSO-based ANFIS is used for fault location. The obtained results prove that the proposed method has good performance even when there are noise. This experiment shows that the extracted features improve the performance of ANFIS when there are noise. Moreover, the developed method will have a better anti-noise ability through collecting more training samples with different noise levels.

### D. COMPARISON AND DISCUSSION

Because of the importance of accurate and fast fault location in VSC-HVDC systems, number of surveys and studies have been done in over the past years and various schemes have been proposed to enhance the accuracy and speed of fault location.
been proposed by researchers. For example, in [34] authors have introduced a new method for fault location on VSC-HVDC transmission line using one terminal current data and the natural frequency of distributed parameter line model. In this method, Prony algorithm to obtain natural frequency. In [35] authors have proposed a one-end gap-based method for fault location in VSC-HVDC systems. In this method, current signal during fault occurrence is used as the initial input signal. In this method, frequency spectrum is produced using post-fault current time series for measuring the gaps between the contiguous peak frequencies. Table 4 and Figure 16 show and compare the performance of different method in term of MSE and the used inputs. The obtained results and comparisons show that the proposed method has much better performance and accuracy than other similar fault location methods.

V. CONCLUSION

In this study, an intelligent and accurate method based on ANFIS proposed for fault location in VSC-HVDC systems. In the proposed method, time and frequency domain features extracted in order to improve the ANFIS performance and fault location accuracy. Furthermore, CDWPSO algorithm used for ANFIS training instead of conventional approach, LS+BP. In first experiment, the performance of CDWPSO-based ANFIS and standard ANFIS tested using two kind of inputs: raw data (current signal) and proposed features (six extracted features from current signal using HH transform). The value of MSE for standard ANFIS using raw data and proposed features was 0.8774 and 0.1839 respectively. In addition, the value of MSE for CDWPSO-based ANFIS and standard ANFIS using raw data and proposed features was 0.5128 and 0.0592 respectively. The obtained results show the high impact of input type and learning algorithm. Moreover, the performance of the proposed method is much better than other similar methods in fault location in VSC-HVDC systems.
