High impedance fault location in 11 kV underground distribution systems using wavelet transforms

A.H.A. Bakar a,*, M.S. Ali b, ChiaKwang Tan a, H. Mokhlis a,b, H. Arof b, H.A. Illias a,b

a UM Power Energy Dedicated Advanced Centre (UMPEDAC), Level 4, Wisma R&D UM, Jalan Pantai Bahara, University of Malaya, 59990 Kuala Lumpur, Malaysia
b Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia

ARTICLE INFO
Article history:
Received 17 August 2012
Received in revised form 3 October 2013
Accepted 9 October 2013

Keywords:
Discrete Wavelet Transform-based MRA
Faulted section
High impedance fault
AAD score

ABSTRACT
Detecting and locating high impedance fault (HIF) in distribution system is a crucial task. HIF causes insulation degradation and over time will lead to supply interruptions. However, it is quite challenging to locate HIF since changes of voltage or current signal during the fault is insignificant and not able to trigger protection system. Besides, the complexity of the distribution system such as non-homogeneous line and lateral with branches increase the difficulty in detecting and locating HIF. Considering these issues, this paper presents a method to detect and locate HIF based on Discrete Wavelet Transform (DWT) Multi-Resolution Analysis. In the method, fault features from voltages measured at primary substation are extracted using DWT and matched with the pre-developed database from simulation. Due to single measurement and multiple branches, the matching will produce multiple possible faulted sections. This problem is then solved by applying ranking analysis to rank the possible faulted sections from the most likely to the least likely faulted section. The proposed method has been tested with all types of faults on a 38-node distribution network system in Malaysia using the PSCAD/EMTDC software. The test results revealed the effectiveness of the method. Since only single measurement of voltage signal is needed, the method is considered economical for practical implementation.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

A common fault in a distribution network is the high impedance fault (HIF) whereby no substantial increase in current since the high impedance restricts the flow of fault current rendering difficulty in detection. For an underground cable, HIF is normally caused by insulation defects that exposes the conductor to contact with non-conducting elements. Different HIF detection schemes have been proposed in the past [1–12]. The process of HIF detection comprises two basic steps; feature extraction and pattern matching or classification. In the feature extraction step, a number of features are extracted from voltage and current signals using various feature extractors or digital signal processing tools including wavelet transform [1–3,8,10,13], Fourier transform [14], Prony analysis [15], S-transform [16], TT-transform [17] and phase space reconstruction [18]. Irregularities in the voltage and current waveforms caused by HIF at a certain section produce unique signatures embedded in the extracted features. In the classification step, the features are fed to a classifier that is trained to identify the unique features and the associated faulted section where the HIF occurs. The classifier is also trained to discriminate HIF from other similar fault phenomena, such as inrush current, load switching, line switching, insulator leakage current and harmonic load [2].

Previously various research studies on locating low impedance fault (LIF) have been attempted. In [19–22], voltage sags pattern was employed to evaluate the fault location. The impedance-based technique have been used in [23–27]. Also, there are few intelligent techniques such as expert system [28,29], genetic algorithm [30], artificial neural network [31,32], and fuzzy system [33–35] have been utilized to locate the fault during a low impedance fault. Unlike LIF, comparatively less research studies have been conducted to identify the HIF location. There are several approaches to determine the HIF location such as network topology technique [5–7,36–39], travelling wave technique [40–43] and knowledge-based technique [44–48].

It has been found that utilizing topology technique promising an accurate high impedance fault localization. However, this technique requires a sensor device or measurement unit to be installed on each node. Thus, it is not a cost effective technique due to the need of high cost installation and maintenance. Besides, it is compulsory to ensure that all the sensor devices or measurement units are in a good condition and functioning well to measure and transmit voltage and current signals to main substation. Failure to do so will affect the overall performance of locating the fault. It will cause fault location to be wrongly identified thus slow down the system restoration.

* Corresponding author.
E-mail addresses: halim5389@gmail.com, a.halim@um.edu.my (A.H.A. Bakar).

0142-0615/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ijepes.2013.10.003
Travelling wave technique is an interesting technique because it has the capability to identify the fault location regardless of the fault type. Also, this technique gives an excellent reliability and high accuracy in identifying fault location which leads to a faster restoration process. However, this technique requires a Global Positioning System (GPS) to transmit a measured data from both ends to be analyzed. Therefore, a reliable communication link is necessary to send the data successfully to the main substation. Besides that, it was noticed that the travelling wave technique is commonly used to locate faults in the transmission system. Till now, there are comparatively less research being conducted to investigate the fault location for distribution system during the occurrence of HIF. This is due to the topology of the distribution system such as radial and branches which requires a high speed data acquisition device to be installed on each node. Thus, this technique is costly and only suitable to be implemented on transmission system which has long distance between two nodes.

Lastly, knowledge-based technique is an easy and fast technique to determine the fault location based on the learning vector. However, this technique requires a large amount of data to be trained and tested to produce a reliable net file in order to identify the fault location accurately. Besides, the accuracy of this technique relies on the reliability and validity of the input data.

In this paper, a new method is proposed in which the faulted section is located based on the matching technique. In this technique, it calculates the difference between wavelet coefficients extracted from the voltage signal and samples of wavelet coefficients stored in databases representing cases of HIF for various locations. Then, it modulus the calculated value to give an absolute value. Since there are more than one absolute values, the average of all the values are calculated which refers as the Average of Absolute Difference (AAD). In this technique, only measured three-phase voltage signal at the main substation will be considered and analyzed using the Discrete Wavelet Transform-based Multi-Resolution Analysis. The proposed method utilizes the first, second and third level resolutions of detail coefficients obtained from the wavelet multi-resolution decomposition. The proposed fault location method was tested on a typical 11 kV distribution network system in Malaysia.

2. Discrete Wavelet Transform-based Multi-Resolution Analysis

Wavelet is a wave-like oscillation with an amplitude that starts out at zero, increases and then decreases back to zero within a limited duration that has an average value of zero. Wavelet also behaves as a mathematical function that satisfies certain mathematical requirement to represent the signal in time domain. Basically, wavelet is utilized to extract distinctive information from different kind of data such as signal and image.

In wavelet analysis, the original signal will be decomposed into low and high frequency components equally by two complementary filters (low-pass and high-pass filters). The low frequency component is a high-scaled decomposition also known as approximation coefficient. The high frequency component on the other hand is a low-scaled decomposition known as detail coefficient.

Discrete Wavelet Transform (DWT)-based Multi-Resolution Analysis (MRA) is the extension of the Discrete Wavelet Transform, in which the process of decomposition is iterated with successive approximation components. It can split the analyzed signal into many lower resolution levels until the individual detail component consists of a single sample. The advantage of the DWT-based MRA is that it gives better signal representation because its resolution is balanced at any time and frequency.

![Fig. 1. Discrete Wavelet Transform-based Multi-Resolution Analysis of a faulted signal, S.](image)

3. Proposed method for the high impedance fault location

In this paper, the identification of high impedance fault location algorithm is based on the Average of Absolute Difference between the extracted features of the voltage signal and stored features from several databases. The process to determine the faulted section is depicted in Fig. 2. It involves two major steps which are HIF detection and HIF location.

3.1. Feature extraction & classification

The three-phase voltage signal measured at main substation is analyzed using the Daubechies fourth order wavelet, Daub4. Once the HIF is detected, approximation and detail coefficients from the first and second cycles relative to the position of the HIF in the voltage signal are analyzed to determine the type of high impedance fault (HIF) and its location. An example of an anomaly caused by an HIF is shown in Fig. 3 where it occurs in the middle cycle of the signal. The cycle that contains the anomaly is considered as the first and the one that follows it is considered as the second of the post disturbance.

Firstly, in every cycle of the voltage signal, 128 samples are taken and transformed by the Daub4 wavelet into 64 approximation coefficients and 64 detail coefficients. The values of the 64 detail coefficients are analyzed and compared to a threshold value. The...
approximation coefficient obtained from the normal cycle given Fig. 6c. This approach provides an easy means to identify an anomaly in the voltage signal. If any detail coefficient values exceed the set threshold value, HIF is considered normal and HIF conditions. If it is observed that any of the detail coefficients fluctuation of the detail components. If any detail coefficient value exceeds the threshold value, HIF is considered detected. This approach provides an easy means to identify an anomaly in the voltage signal as shown in Fig. 6c.

After detecting an HIF, it is necessary to classify its type and only the sum of first level approximation coefficients from the second cycle is needed. This sum is divided by the sum of first level approximation coefficients obtained from the normal cycle given by the following equation:

\[
AR = \frac{\sum a(HIF)}{\sum a(normal)} \tag{1}
\]

Based on the value of this ratio, the HIF is classified into one of the four fault types and thus the database to use as follows:

(i) Single line to ground fault (SLGF) – faulted phase will have an AR magnitude lower than 1.0 and the other two higher than 1.0.

(ii) Three line fault (LLLF) – all of the three phases have almost identical AR magnitudes.

(iii) Double line to ground fault (LLGF) – faulted phases will have an AR magnitude lower than 1.0 and the other phase vice versa.

(iv) Double line fault (LLF) – healthy phase will have an AR magnitude the same as before the fault (AR = 1), whereas the faulted phases change (one of the phases has AR magnitude lower than 1.0 and the other phase vice versa).

Then the final step is to find the section where the fault occurred. For this purpose, detail coefficients from the first and second cycles of the three voltage phases are calculated. For each cycle, the detail coefficients at the first, second and third levels of wavelet expansion are added. So, altogether there are 18 wavelet features to be compared against sets of stored features from the selected database. For each set, the Average of Absolute Difference (AAD) between the extracted wavelet features and the stored features is calculated as a measure of similarity between them. Out of all possible cases, only three candidates with the lowest AAD score are retained. Test results show that all of the simulated faults can be detected in the three candidates with the lowest AAD score.

3.2. Database development

For each fault type, five impedance values of 60 Ω, 70 Ω, 80 Ω, 90 Ω and 100 Ω are used to simulate faults in the middle of each line section. Since the network has 34 sections, each impedance value contributes 34 cases to the databases. Altogether, there are 170 fault cases in each of the four databases where each case is represented by 18 wavelet features. The process of developing the databases is summarized by the following steps:

1. A fault in the middle of each line section, starting from section 1 to section \(N\) is generated (where \(N\) is the total number of line sections).
2. The first and second cycles of post-disturbance voltage waveform are analyzed using the DWT-based MRA. The first cycle of post-disturbance is determined based on the sharp fluctuation of the detail components. If any detail coefficient value exceeds the threshold value, HIF is considered detected.
3. The sum of the first, second and third levels of detail coefficients for each voltage phase are extracted.
4. The calculated sum of detail coefficients is stored in the database for that particular line section.
5. The value of \(j\) is checked. If \(j\) is not equal to \(N\), steps 1 and 4 are repeated.
6. If \(j\) is equal to \(N\), the process is terminated.

3.3. Faulted section prediction based ranking analysis

After constructing a reference database that contains samples of wavelet coefficients of HIF cases from different sections in the network, the effectiveness of the proposed method in locating the faulted section is tested. Given a signal of an HIF case where fault is applied in the middle of a line section with defined fault impedance value, a set of wavelet features is extracted. Then the Average of Absolute Difference (AAD) between the extracted wavelet features from the signal and each set from the database is calculated. Fig. 4 shows the first six wavelet features from the measured signal matched against the same number of features from 4 samples of the reference database that most closely resemble them. It should be noted that only the first six coefficients out of 18 are shown in the figure due to the need to show the differences in their values with sufficient magnification. The Average of Absolute Difference (AAD) between the features of the test signal and each case of the reference database can be calculated as follows:

\[
AADm = \frac{\sum_{i=1}^{n} \sum d_{i}(measured) - r_{mi}}{n} \tag{2}
\]
where \( m = 1,2, \ldots, 34 \) (number of line section), \( n \) is number of data, i.e., \( n = 18 \), \( \sum d_{(\text{measured})} \) is a sum of detail coefficients of the test voltage signal, \( r_{(r)} \) is reference database.

Once the AADs for all sets in the database have been calculated, they are arranged from the smallest to highest value where the smallest value is ranked as the first. Finally, the proposed method will check the suspected faulted sections associated with the lowest few AAD values that give a high probability of occurrence.

It is necessary to use a few low AAD values to detect the faulted section because in some cases the lowest AAD value does not point to the correct fault section. This is due to the complexity of the distribution network such as branches, non-homogenous lines and high impedance fault that results in variation of fault location. Thus, if the section related to the lowest AAD value is not faulted, the HIF is searched at sections related to the second and third lowest AADs. This is done by physically inspecting the suspected locations. In practice, when any fault occurs, engineers have to go to the location where the fault occurs in order to clear the fault thus expedite the power system restoration.
A schematic diagram of a typical 11 kV distribution network system in Malaysia is shown in Fig. 5. It consists of 38 nodes that represent 34 sections of the network. The test system frequency is 50 Hz and the sampling frequency is 6.4 kHz, resulting in 128 recorded data for each cycle. The data is sampled at a measurement node connected to the 132/11 kV radial distribution network as shown in the figure.

In the simulation, a database was created and the proposed algorithm was tested with different values of high impedance fault simulating SLGF, LLGF, LLF and LLLF cases. The simulation was carried out using the PSCAD software to obtain the fault voltage signal during HIF event. The measured voltage signal was analyzed using the Discrete Wavelet Transform program in MATLAB. In each voltage cycle, 128 data were recorded and analyzed to extract the coefficients. It consists of 18 features which are the sums of detail coefficients obtained from the wavelet transform with three levels of phase A, B and C for two cycles of post-disturbance voltage signal. Compare these features to the stored ones, the absolute difference of each is calculated and summed.

4.1. Wavelet waveform analysis

To investigate the effectiveness of the proposed method, different fault impedance values were applied to the middle point of the line section. The tested high impedance fault values were 75 Ω, 85 Ω and 95 Ω. The voltage signal is recorded before and after the fault was applied. A sample of voltage signal containing the fault occurrence is shown in Fig. 6a. There was a very small fluctuation in one of the cycle that is hardly noticeable to the naked eye. However it can be observed clearly after the cycle is magnified and shown in Fig. 6b. It is the first cycle after the fault was applied to line section 6 with 85 Ω fault impedance.

Generally, this small deviation cannot be detected by the common impedance-based methods because it behaves like a normal signal. In order to distinguish this type of anomaly, a digital signal processing technique is required. In the proposed method, Discrete Wavelet Transform is applied to decompose the first and second post-disturbance cycles into the approximation and detail coefficients. After the decomposition of the two cycles, sharp fluctuations can be seen in its detail coefficients as shown in Fig. 6c–e. They are the detail coefficients of the first, second and third level of the DWT-based MRA respectively.

(a) Instantaneous voltage signal.
(b) 1st Cycle of post-fault.
(c) 1st Level of detailed coefficients.
(d) 2nd Level of detailed coefficients.
(e) 3rd Level of detailed coefficients.

In this work, the HIF is detected by observing the first level of detailed coefficient. The threshold values are set to 0.02 and –0.02 within which the detail coefficients are regarded as normal. If the detail coefficients exceed 0.02 or fall below –0.02, then HIF conditions have occurred as shown in Fig. 6c. To identify the faulted section, the sum of detail coefficients measured above is compared to the ones in the selected database. The overall process for estimating the faulted section using the proposed method is depicted in Fig. 7.

4.2. Analysis of test results

To observe the efficiency of the proposed method, different fault impedance values were applied at four different line sections one by one. Line sections 2 and 9 were used to represent faults at feeder 1 while line sections 29 and 34 were selected to represent faults at feeder 2. Table 1 summarizes the tested locations and their respective fault impedance values while Table 2 shows the test results for the respective tested section for SLGF, LLLF, LLGF and LLF.
As shown in Table 2, the first and second columns are the tested sections and fault impedance values. The third, fourth, fifth and sixth column contains the ranking numbers of the tested sections when they are subjected to SLGF, LLLF, LLGF and LLF faults respectively. Then, the same test was performed on the remaining 30 sections of the network and the overall results are presented in Section 4.3.

It is observed that sections 2 and 9 have been correctly identified as the faulted sections when they are assigned AAD rank 1 for all fault conditions. For section 29 ad 34, AAD rank 2 has been assigned for some faulted cases. This is not bad considering that they are just the second choice after the section with AAD rank 1 out of 34 sections.

Table 3 shows the result of simulating an SLGF using 75 Ω fault impedance at the middle of line section 29. The features are extracted from the input signal, and the AAD value for each case is calculated using the proposed method. Then the AAD values are arranged from the lowest to the highest and the respective ranking is assigned accordingly. The five lowest AAD values are shown in Table 3 along with their respective sections but only the lowest three are selected for the faulted section. The first column shows the AAD value measured by comparing the input features to those from the database, the second column provides the faulted section associated with the AAD values and the third column contains the rank number.

It can be observed that the lowest AAD value belongs to line section 30 although linesection 29 is the faulted one. According to our approach, we start by inspecting line section with rank 1. Since the fault is not found in section 30, we go to the line section with rank 2 and we find that section 29 is faulted. Although the actual faulted section is successfully located after the second attempt, in practice, when any fault occurs, engineers have to do a physical inspection by visiting the suspected locations. Hence,

**Table 1**

<table>
<thead>
<tr>
<th>Test section</th>
<th>Fault impedance value (Ω)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 9 (main at feeder 1)</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>95</td>
</tr>
<tr>
<td>Section 34 (main at feeder 2)</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>95</td>
</tr>
<tr>
<td>Section 2 (branch at feeder 1)</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>95</td>
</tr>
<tr>
<td>Section 29 (branch at feeder 2)</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>95</td>
</tr>
</tbody>
</table>

**Table 2**

<table>
<thead>
<tr>
<th>Test section</th>
<th>Impedance (Ω)</th>
<th>SLGF</th>
<th>LLLF</th>
<th>LLGF</th>
<th>LLF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 9 (main at feeder 1)</td>
<td>75</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Section 34 (main at feeder 2)</td>
<td>75</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Section 2 (branch at feeder 1)</td>
<td>75</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Section 29 (branch at feeder 2)</td>
<td>75</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 3**

<table>
<thead>
<tr>
<th>AAD value</th>
<th>Faulty section candidate</th>
<th>Ranking number</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00010</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>0.00018</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td>0.00034</td>
<td>31</td>
<td>3</td>
</tr>
<tr>
<td>0.00055</td>
<td>28</td>
<td>4</td>
</tr>
<tr>
<td>0.00103</td>
<td>27</td>
<td>5</td>
</tr>
</tbody>
</table>
the proposed method reduces the high number of possibilities to just three.

Table 4 shows the result when two fault values are applied at several locations along the line of section 12. In this case, the two different fault impedance values are 65 Ω and 75 Ω. The fault locations are at 10%, 25%, 50%, 75% and 90% of the line section length.

From the table, it shows that the proposed method can still identify the faulted section although the location of the fault varies along the line section. Also, it is observed that the accuracy increases (1st rank) as the fault is shifted near to the end of the line section.

4.3. Overall performance

The overall performance of the proposed method in identifying the actual faulted section is presented in Figs. 8a–c. A fault is applied at the middle line of each section of the network and then the capability of the proposed method to locate the fault is analyzed. It is found that most of the faulted sections can be traced to the sections associated with the lowest AAD. The remaining faulted sections are mainly located in the sections of the second lowest AAD and a few can be traced with sections of the third lowest AAD. The effectiveness of the proposed method is evaluated based on the number of times the actual faulted sections are found in the sections associated with the three lowest AAD values. The x-axis represents the type of faults and the y-axis represents the total number of correct faulted sections associated with the first, second and third rank AAD values.

The results show that most of the faulted sections can be located in the first rank AAD values, thereby they could be found in the first sections checked. For the LLGF and LLF faults, 97% and 100% accuracies are obtained if sections of the first rank AAD are searched. However, there are a few faulted sections found in the second or third ranked sections. From the theoretical point of view, it is acceptable to inspect a maximum of three sections for a fault because locating the faulted section in a distribution system is quite difficult due to the branch topology of the network. Narrowing the choices to just three possible choices reduces a lot of time taken to identify the actual faulted section. Instead of patrolling the feeder or checking section by section, the faulted section can be discovered more efficiently by the proposed method. Also, the AAD value of 75 Ω, 85 Ω and 95 Ω are different from the values used to construct the databases which are 70 Ω, 80 Ω, 90 Ω and 100 Ω. It is assumed that if the same fault impedances were used, a higher accuracy is achievable. This proves that the proposed method is effective in locating a faulted section in an underground distribution network for SLGF, LLLF, LLGF and LLF.

5. Conclusion

This paper presents an approach to detect and locate an HIF using a Discrete Wavelet Transform (DWT) Multi-Resolution Analysis (MRA) and a pre-defined faulted section database as a reference. In this approach, the Average of Absolute Difference between the wavelet coefficients of the network voltage signal and those of the reference database is used as an indicator of the location of the faulted section. Since the proposed method utilized pre-stored databases, the whole process of estimating the faulted section is efficient and fast.

The method was simulated using the real time simulation environment of PSCAD/EMTDC software for a typical 11 kV underground distribution network. The test results of SLGF, LLF, LLGF and LLLF show that most of the tested faulted sections are associated with the sections with the lowest AAD score. The remaining few cases are accounted for by sections with the second or third lowest scores. In summary, the proposed method is suitable to be adopted for detecting underground cable fault location.

Acknowledgments

This work is supported by the Ministry of Higher Education under High Impact Research Grant (HIR-MOHE D000004-16001) and the University of Malaya.
References


