Arcing Fault Detection with Interpretable Learning Model Under the Integration of Renewable Energy

Yousaf Hashmy, Qiushi Cui, Zhihao Ma and Yang Weng
School of Electrical, Computer and Energy Engineering
Arizona State University, Tempe, Arizona, USA
Email: {shashmy, qiushi.cui, zhihaoma, yang.weng}@asu.edu

***Abstract*—Under the trend of deeper renewable energy integration, active distribution networks are facing increasing uncertainty and security issues, among which the arcing fault detection (AFD) has baffled researchers for years. Existing machine learning based AFD methods are deficient in feature extraction and model interpretability. To overcome these limitations in learning algorithms, we have designed a way to translate the non-transparent machine learning prediction model into an implementable logic for AFD. Moreover, the AFD logic is tested under different fault scenarios and realistic renewable generation data, with the help of our self-developed AFD software. The performance from various tests shows that the interpretable prediction model has high accuracy, dependability, security and speed under the integration of renewable energy.**

***Index Terms*—Arcing Fault Detection, Distribution Networks, Power System Protection, and Renewable Energy.**

I. Introduction

RESEARCHERS and engineers have kept on exploring new ways for AFD since the 1970s. Arcing fault (AF) is usually associated with an undowned or downed conductor. The undowned conductor scenario involves the contacts between overhead lines and tree limbs that have large impedance. Similarly, if a downed conductor falls on a poorly conductive surface such as sand, asphalt, grass, soil, and concrete, the fault current might be too low to reach the pickups of traditional ground overcurrent relays. Additionally, the past two decades have witnessed a rapid growth in distributed energy resources (DERs) worldwide. This trend has added more uncertainties on top of AF's irregular, non-linear and asymmetric attributes.

At the early stage, enhancements of conventional relays are proposed, leading to a proportional relaying algorithm [1], impedance-based method [2], and PC-based fault locating and diagnosis algorithm [3]. However, these methods are ineffective in detecting AFs with a low fault current. For this problem, harmonics patterns are utilized to capture AF characteristics, such as magnitudes and angles of 3^{rd} and 5^{th} harmonics [4], even order harmonic power, and interharnomic currents. Besides, a Kalman-filter-based method is proposed to monitor harmonics in AFDs. This type of method actively injects higher than fundamental frequency signals like positive/zero voltage signals into the grid for AFD. Unfortunately, most of these attempts at addressing AFD issues rely on simple thresholds and logic, which lack a systematical procedure that determines the most effective features for various distribution systems and scenarios during AFs. Therefore, it is getting necessary to introduce a systematic design for a learning framework so that information gain in high-dimensional correlation can be quantified for better AFDs.

With the advancement of artificial intelligence, methods on expert systems, neural networks, and machine learning are gaining popularity in recent years. Wavelet transform is adopted to extract data and Bayes classifier to differentiate fault cases. Wavelet transform is also used and the classification method used here is the nearest neighborhood rule. Originating from the analysis and processing of geometrical structures, mathematical morphology (MM) method is presently gaining popularity in data extracting upon the inception of fault, [5] proposed a method based on MM alone to detect AFs, while some papers combined MM with neural network, and support vector machine (SVM). Moreover, principal component analysis and SVM are utilized. The aforementioned machine learning methods share a common drawback in model interpretability – their prediction models are non-transparent and therefore not explainable. From the implementation perspective, this issue is vital since a non-transparent model causes tremendous problems to application engineers, even though we assume future application engineers have acquired necessary training on neural networks, Naive Bayes, nearest neighborhood, etc.

This paper proposes a framework that bridges the gap between the non-transparent machine learning algorithm and a real detection logic. Specifically, the AFD features are firstly engineered to reduce their dimension through feature extraction method. Then, the induction method of decision tree is deployed for model interpretability. Subsequently, the AFD logic is designed by considering a variety of realistic issues in its application. It provides practical insights on utilizing machine learning based prediction model. Furthermore, we have developed an AFD software that automates this process. The numerical validation conducts a comprehensive fault analysis and compares the performance of the proposed method with five other methods over six evaluation metrics. This paper contributes to use variable-importance-based feature selection method to identify an effective feature set out from a large feature pool. Specifically, we conduct a systematic design of AF feature pool by looking into when the fault happens, how long it lasts, and what the magnitude of the fault is. For when,

978-1-7281-0407-2/19/$31.00 ©2019 IEEE

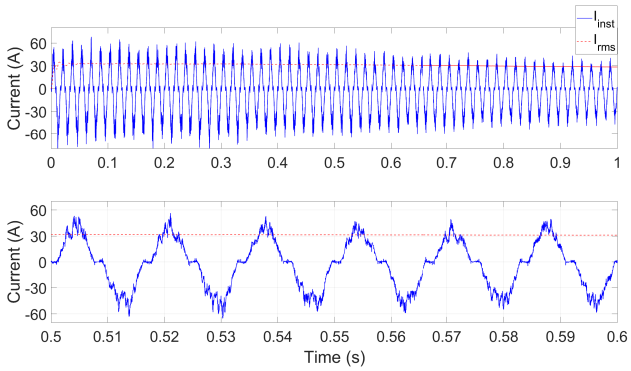


Fig. 1. The current waveforms upon AF. The upper waveform shows the instantaneous and RMS AF currents during 1 sec. The lower waveform is zoomed in from the upper waveform from 0.5 to 0.6 sec.

we first calculate different quantities such as active power and reactive power based on the voltage and current time series. Then, we use the derivative of these quantities to tell when there is a potential change due to AF.

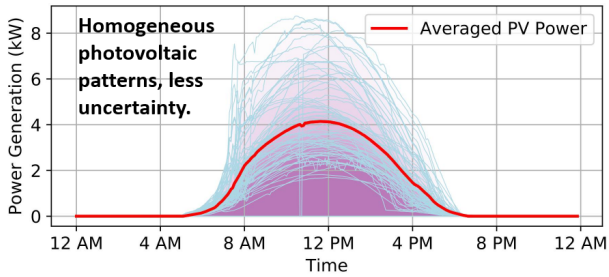
This paper has the following structure: Section II introduces the AF modeling techniques. Section III explains the formulation of the interpretable AFD model. The details of the AFD logic is elaborated in Section IV, followed by the performance analysis in Section V. Section VI presents the conclusions.

II. Modeling of Arcing Faults

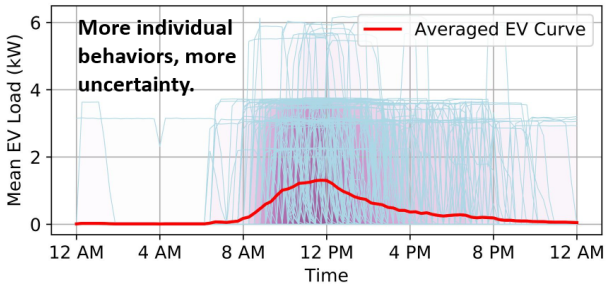
Although AF phenomena are difficult to model in general, there are mainly three ways to model AFs including both downed and undowned types for analysis. Each way provides acceptable similarity with real AFs from its own perspective. In the following, we briefly explain each of them and the motivation behind the chosen model.

- The first one is called the transient analysis of control systems (TACS) controlled switch. This model emulates arc conduction, re-ignition, and extinction. The advantage of this model is the adjustable phase difference between the applied voltage and fault current.
- The second way originates from the Kizilcay model which utilizes a dynamic arc model derived from the viewpoint of control theory based on the energy balance in the arc column.
- The third way of modeling AF is the employment of two anti-parallel DC-sources connected via two diodes, plus two variable resistors. The nonlinear impedance was included to add the non-linearity of fault current [4]. Later on, the model is extended with two anti-parallel DC-sources connected via two diodes, which modeled the asymmetric nature, as well as the intermediate arc extinction around current zero. This kind of model is able to model the effective impedance and thus the randomness of the resulting fault current.

Due to the ease of implementation in Matlab Simulink for multiple simulations to realize the proposed machine learning-based method, we employ the third model. In addition, we further improve this model by replacing the two variable resistors with two controlled resistors. Each controlled resistor has an integrator to represent the moisture changing process in the vicinity of the point of contact of the conductor with the ground, a randomizer to introduce more randomness during AF and a first-order transfer function to tune the response to the introduced randomness.



(a) PV generation curves.



(b) Electric vehicle charging loads.

Fig. 2. Realistic renewable energy data from industrial partners. Red line: averaged PV generation and EV load curve. Grey line: individual curves.

The adopted model is, therefore, more accurate, since the moisture change and system dynamic response are incorporated. The obtained AF current waveforms are presented in Fig. 1, which clearly displays the irregular and decreasing current waveforms upon an AF. Test results of this AF model reveal a good modeling performance and are validated in the simulation and field test results.

III. Interpretable Arcing Fault Detection Method

To generate an interpretable detection model, we propose an interpretable AFD method. The expert knowledge and the system information are firstly needed before collecting all necessary data. After data preparation, we rely on feature extraction to reduce the dimension of the input signals. After that, different induction algorithms are examined. We finally choose the decision tree classifier to translate the prediction model into an AFD logic, which is elaborated in Section IV.

A. Data Preparation: System-level Events

Data preparation is assisted with the expert knowledge of useful system-level events. The adopted technique is transferable on different feeders because the event category and event type in Table I are suitable for most of the distribution feeders during the training. In our study, we include realistic generation data from PV and load data from electric vehicle charging stations, the purpose of which is to avoid false tripping of the proposed pattern recognition method under deep penetration levels of renewables. Fig. 2 depicts the typical

curves from our industrial partner. We run simulations at spot load by integrating realistic distributed generation and load over a 24-hour horizon.

TABLE I
EVENT CATEGORY OF SYSTEM UNDER STUDY

Event Category	Event Type	Number of Events
System Operating Condition	Loading Condition (30%-100%)	8
	DER Tech. (SG, inverter, hybrid)	3
Fault Event	Type 1: SLG, LLG, LL, LLLG	10
	Type 2: Downed conductor	3
	Fault impedance	6
	Inception Angle (0°, 30°, 60°, 90°)	4
	Fault location	3
Non-fault Event	Normal State	1
	Load Switching	6
	Capacitor Switching	2

Moreover, the event category in Table I is flexible and can be tailored for other special systems by adding or deleting some of them. In this case study, comprehensive scenarios are considered in the event category. A loading condition ranging from 30% to 100%, in a step of 10%, is simulated. Furthermore, eight loading conditions and three DG technologies are examined respectively on top of the base case scenario.

B. Feature Selection for Arcing Fault

If we have to use hundreds of features for AFD, its applicability is compromised. Therefore, it is necessary to identify key feature set and reduce the data amount. We elaborate on the way of selecting the key features in this section. The variable-importance in feature evaluation is firstly explained, followed by the selected feature set.

1) Variable-importance in Feature Evaluation: An effective and unbiased feature evaluator is required to calculate the merit of each tested feature before the classification between AF event and non-AF event. Here, we take advantages of the information gain and minimum description length (MDL)-based discretization algorithm to select important features during AF. For the convenience of power background readers, we call MDL score the variable of importance in this paper.

The score of variable-importance is one type of selection measures in machine learning. The problem of selecting the best attribute can be stated as the problem of selecting the most compressive attribute. Assuming that all features are discrete, the objective is to find the best features that maximize the selection measure. Let C, A and V denote the number of classes, the number of features, and the number of values of the given feature. With this notation, we have the entropy of the classes $H_C = -\sum_i p_{i.} \log p_{i.}$, the values of the given feature $H_A - \sum_j p_{.j} \log p_{.j}$, the joint events class-feature value $H_{CA} = -\sum_i \sum_j p_{ij} \log p_{ij}$, and the classes given the value of the attribute $H_{C|A} = H_{CA} - H_A$, where $p_{ij} = n_{ij}/n_{..}$, $p_{i.} = n_{i.}/n_{..}$, $p_{.j} = n_{.j}/n_{..}$ and $p_{i|j} = n_{ij}/n_{.j}$. "$n_{..}$" denotes the number of training instances and "$n_{i.}$" is the number of training instances from class C_i, $n_{.j}$ is the number of instances with the j-th value of the given attribute, and n_{ij} is the number of instances from class C_i with the j-th value of the given attribute. The approximation of the total number of bits that are needed to encode the classes of $n_{..}$ is:

$$\text{Prior MDL}' = n_{..} H_C + \log \binom{n_{..}+C-1}{C-1}, \tag{1}$$

and the approximation of the number of bits to encode the classes of examples in all subsets corresponding to all values of the selected attribute is:

$$\text{Post MDL}' = \sum_j n_{.j} H_{C|j} + \sum_j \log \binom{n_{.j}+C-1}{C-1} + \log A.$$

The last term ($\log A$) is needed to encode the selection of an attribute among A attributes. However, this term is constant for a given selection problem and can be ignored. The first term equals $n_{..} H_{C|A}$. Therefore, the MDL′ measure evaluates the average compression (per instance) of the message by an attribute. The measure is defined by the following difference, Prior MDL′ − Post MDL′, normalized with $n_{..}$:

$$\text{MDL}' = \text{Gain} + \frac{1}{n_{..}} \Big(\log \binom{n_{..}+C-1}{C-1} - \sum_j \log \binom{n_{.j}+C-1}{C-1} \Big). \tag{2}$$

However, entropy H_C can be used to derive MDL′ if the messages are of arbitrary length. If the length of the message is known, the more optimal coding uses the logarithm of all possible combinations of class labels for a given probability distribution:

$$\text{Prior MDL} = \binom{n_{..}}{n_{1.},\ldots,n_{C.}} + \log \binom{n_{..}+C-1}{C-1}. \tag{3}$$

Similarly, if we use the priori minus the posterior of the MDL, we have

$$\text{MDL} = \frac{1}{n_{..}} \Big(\binom{n_{..}}{n_{1.},\ldots,n_{C.}} - \sum_j \log \binom{n_{.j}}{n_{1j},\ldots,n_{Cj}} + \log \binom{n_{..}+C-1}{C-1} - \sum_j \log \binom{n_{.j}+C-1}{C-1} \Big). \tag{4}$$

2) Effective Feature Set (EFS): To get an EFS using the score of variable-importance, we systematically design a pool [6] with many practical and reliable features widely used in microprocessor-based relays. Currently, we have 246 features in the feature pool, as indicated in [7]. Some of them are discrete Fourier transform (DFT) based, while others are Kalman Filter (KF) based. But many other features in the literature can be added to the feature pool. The feature pool is flexible and expandable to maximize the AFD accuracy. Table II summarizes the obtained EFS for unbalanced AFD.

TABLE II
REFERENCE TO THE UNBALANCED FAULT DETECTION FEATURES IN EFS

Features in EFS	Reference
V_2, I_2	[4], [8]
$\theta_{V_2} - \theta_{V_0}$, $\theta_{I_2} - \theta_{I_0}$	[4], [9]
$KF\_V_a\_cos\_H3$	[4] (3rd harmonic), [10] (KF and low-order odd harmonic), (KF harmonic decomposition)
$KF\_V_a\_sin\_H3$	

Remark 1. *These features are extracted mainly through two techniques: DFT and KF. Both techniques are simple, reliable and implementable in engineering fields. The DFT is used to capture the majority of physical quantities in fault detection as is widely used in microprocessor-based relays. On the other hand, the utilization of the KF-based algorithm is motivated by the fact that it can accurately track the harmonics and inter-harmonics coefficients at given frequency components.*

IV. From Interpretable Decision-Tree Model to Arcing Fault Detection Logic

Inspired by the tree structure of the machine learning classifier, we further explore the possibility of relating the EFS and the detection logic using simple thresholds as most of the commercial products and patents do. Statistically, since three-phase faults take up only $2\% - 3\%$ of the fault occurrences, an AFD logic is designed in this regard for unbalanced AFs.

The AFD logic is targeted to be implemented in a microprocessor-based digital relay. Before the explanation of the AFD logic, the logic circuit is presented first in Fig. 3. Generally, the proposed AFD scheme updates its comparison and decision logic according to the obtained tree structure.

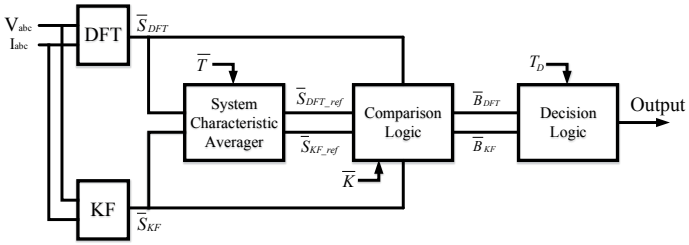


Fig. 3. The proposed AFD logic scheme.

As indicated in the detection logic, three-phase voltage and current signals are sent to DFT and KF for feature extraction. This section takes the obtained EFS in Section III-B as an example. (5) and (6) show the extracted instantaneous signals after the DFT and KF blocks:

$$\bar{S}_{DFT} = \{s_1, s_2\, s_3, s_4\} = \{V_2, I_2, \theta_{V_2} - \theta_{V_0}, \theta_{I_2} - \theta_{I_0}\}, \quad (5)$$

$$\begin{aligned}\bar{S}_{KF} = \{&s_5, s_6, s_7, s_8, s_9, s_{10}\} = \{KF\_V_a\_cos\_H3,\\ &KF\_V_b\_cos\_H3, KF\_V_c\_cos\_H3, KF\_V_a\_sin\_H3,\\ &KF\_V_b\_sin\_H3, KF\_V_c\_sin\_H3\}.\end{aligned} \quad (6)$$

A. System Characteristic Averager

The inputs of the System Characteristic Averager are the extracted instantaneous signals after the DFT and KF blocks. Meanwhile, the time duration $\bar{T}$ needs to be provided to this averager. Specifically, the System Characteristic Averager has a memory that stores the signals for a predefined duration of $\bar{T} = \{t_1, t_2, t_3, t_4, t_5, t_6,\ t_7, t_8, t_9, t_{10}\}$. In other words, $\bar{T}$ is the time constant that is a vector of ten elements associated with $\bar{S}_{DFT}$ and $\bar{S}_{KF}$. The input signals are stored and calculated at every $18,000$ cycles (5 minutes) [4]. The five minutes interval is subject to change depending on the case-specific analysis.

B. Comparison Logic

The block of Comparison Logic is depicted in Fig. 4. Based on the feature extraction technique discussed in Section III-B, the extracted instantaneous signal s_i can be understood as the system background signal superimposed by the extra signal contributed from the AF behavior. The comparison is, therefore, made between the extracted instantaneous signal s_i and its reference value s_{i_ref} [4].

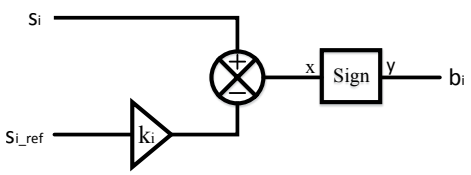


Fig. 4. Comparison Logic in the proposed AFD logic.

The sensitivity gain of k_i is incorporated in order to 1) set the margin of detection and 2) add a handle to the detection sensitivity. The undefined parameter of $\bar{K}$ stands for:

$$\bar{K} = \{k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8, k_9, k_{10}\}. \quad (7)$$

The sensitivity gain $\bar{K}$ is set at 1.2 (adjustable for each element). The 20% above and below margin is adjustable and is taken as a typical blackout region where the AF tripping is not required [4]. This k_i value can be set to close to 1.0 after getting more confidence in AFD scheme. After the summation block in Fig. 4, a Sign function is employed to output "1" when input is larger than zero. The output of the comparison logic is the comparison assertion bit of b_i $(i = 1, 2, \cdots, 10)$ and $\bar{B}$ is the input to the decision logic.

C. Decision Logic

As mentioned in the previous subsection, the comparison assertion bit of b_i $(i = 1, 2, \cdots, 10)$ is the output of the comparison logic in Fig. 4. The decision logic in Fig. 5, is the execution part of the AFD logic. There are four groups of signal bits:

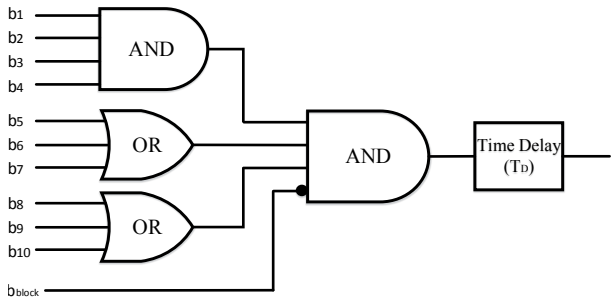


Fig. 5. Decision Logic in the proposed AFD logic.

1) DFT-based assertion bits. The four bits go through an AND gate. If any of the four signals are not asserted, the decision logic will not be set high.
2) KF-estimated in-phase components of third harmonic voltage. If none of the three-phase in-phase components

of third harmonic estimated from the KF gets asserted, the decision logic will not be set high.

3) KF-estimated in-quadrature components of harmonic voltage. If none of the three-phase in-quadrature components of third harmonic estimated from the KF gets asserted, the decision logic will not be set high.
4) The blocking bit b_{block}. If this bit is 1, the detection logic is blocked and none of AF events can be detected; if this bit is 0, AFD is enabled.

A time delay of T_D is implemented because an appropriate selection of T_D can effectively avoid the false operation resulting from normal switching, which sometimes contributes to third harmonics. The output of the AF logic is either alarming or tripping signal.

V. Performance Test of the Proposed AFD Logic

A. Testing Environment

The benchmark system utilized can be found in Fig. 6. It has a 9 MVA synchronous generator and a 1 MVA PV system. The proposed AFD logic is tested under 7884 new scenarios: 7776 unbalanced faults and 108 non-faults. To overcome the imbalance in the samples, the synthetic minority over-sampling technique (SMOTE) is employed to generate synthetic samples and shift the classifier learning bias towards minority class [11]. The fault locations under testing include faults near B-3, B-11, and B-16. Similar to the work in [4] and [5], the measurement point is at the substation. Its sampling frequency is $2,000$ Hz. The time delay in Fig. 5 is set to 100 ms. The average fault detection time is 0.126 sec using a real-time simulator. The signals measured are the voltage and current. The features used are derived from the measured signals and can be found in the EFS in Table II.

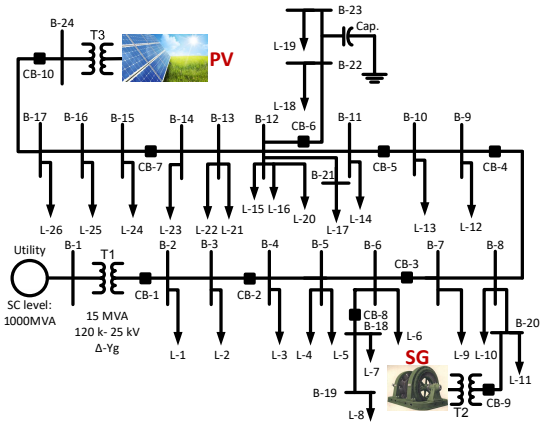


Fig. 6. Single line diagram of distribution feeder under study.

B. Performance of the AFD Logic

In order to compare the proposed technique with some existing ones in the field, we borrow the concepts of true positive (TP), true negative (TN), false positive (FP), and false negative (FN)[1] from statistical classification, and define them and their related evaluation criteria [12] as follows:

- Accuracy: $A = \frac{TP+TN}{TP+FP+TN+FN}\%$.
- Dependability: $D = \frac{TP}{TP+FN}\%$.
- Security: $S = \frac{TN}{TN+FP}\%$.
- Speed: $V = \frac{T_{one-cycle}}{T_{detection}}\%$, where $T_{one-cycle}$ and $T_{detection}$ are the time duration of one cycle and the detection time respectively.
- Objectivity (OBJ): the objectivity to fault type and network, indicating whether the technique is objective to the type of fault, and the network topology.
- Completeness (COM): the ability to hold important information, indicating the time window of the data that is needed for the method to make the crucial decision.

The AFD logic performance is tested against realistic renewable energy data (refer to Fig. 2) from our industrial partners. We also compare the performance of the AFD logic with four representative AFD methods in [4], [12]–[14], as well as the combined conventional relay elements (including frequency, over/under voltage, over current) in Table III. The methods in the comparison group cover logic-gate based AFD, wavelet domain analysis, time-frequency domain analysis, and pattern recognition techniques.

TABLE III
AFD Logic Performance Tested Against Realistic Renewable Energy Data.

Solution under test	A (%)	D (%)	S (%)	V	OBJ	COM
The proposed EFS and AFD logic	98.3	98.3	98.2	0.13	No	Yes
The work in [4]	N/A	69.0	90.7	N/A	No	Yes
The work in [12]	93.6	100	81.5	1.00	Yes	Yes
The work in [13]	96	90	100	0.25	No	No
The work in [14]	94.9	90.0	90.9	0.11	No	Yes
Combined conventional relay elements	49.1	0.0	98.2	N/A	Yes	Yes

Note: N/A stands for "Not Available" in this table.

Comparing with the other five methods in Table III, it is indicated that the proposed method has a superior overall performance in terms of the six evaluation criteria. For example, the detection accuracy of the proposed method is the highest among the solutions under test; its detection speed $(1/60/0.126 = 0.13)$ is not the fastest but fits well in the AF detector requirements on response time. The detection time of less than 1 second, which means the minimum speed of 0.017 in a 60 Hz network is viewed as a conservative setting [5].

C. Fault Scenario Analysis

We have tested the EFS in various fault impedance and locations. The quantifier for evaluation is the variable of

[1]TP: the number of correctly detected fault events. TN: the number of correctly detected non-fault events. FP: the number of incorrectly detected fault events but they are actually non-fault events. FN: the number of incorrectly detected non-fault events but they are actually fault events.

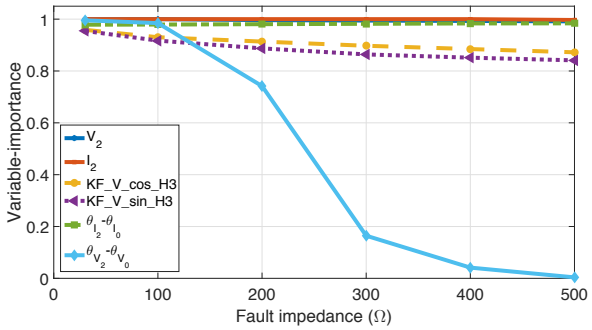


Fig. 7. Variable-importance of all features under faults in a grounded system.

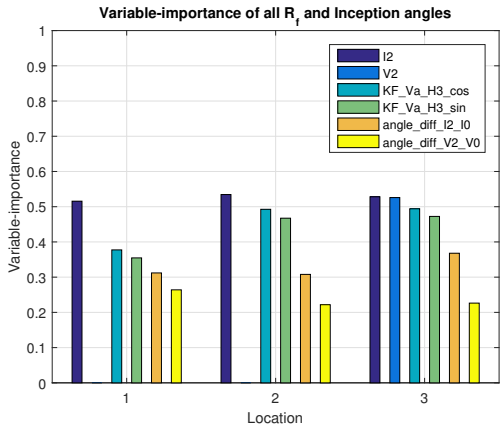


Fig. 8. Variable-importance at different fault locations.

importance explained in Section III-B1. In general, the EFS is robust in all fault scenarios. Specifically, we plot Fig. 7 to provide the impact of fault impedance on the variable-importance. To be practical, this paper investigates the fault impedance up to 500Ω to cover typical AFDs whose fault currents are as low as 10 amps. It is concluded that the negative sequence of voltage and current are the most reliable features that can keep unaffected during any unbalanced fault upon a varying fault impedance. We also notice that the feature of the angle difference between negative sequence voltage and zero sequence voltage is vulnerable to high fault impedance, and the third harmonic components estimated from KF gets slightly deteriorated when the fault impedance increases.

The variable-importance of the features in EFS is presented at three fault locations (refer to Section V-A). The result is demonstrated in Fig. 8, including all fault impedance and all fault inception angles in Table I. The feature of negative sequence current keeps being unaffected at each location. However, the negative sequence of voltage is so low at location 1 and 2 that the variable of importance becomes almost zero. As the strong voltage source from the substation is ideally balanced, the negative sequence voltage deviation contributed from the AF is weak. Location 3 is selected, that is far from the substation, so the negative sequence voltage becomes a good AF indicator again. To a negligible extent, it is similar for the variable of importance performance of other features: the further the fault is, the less compromised the features are.

VI. Conclusions

This paper proposes a new framework for AF detection and classification. By introducing the MDL-based algorithm to rank a pool of systematically designed features, an effective feature set is generated. The detection capability of such a ranked feature set is evaluated through a comprehensive fault analysis. Furthermore, an interpretable AFD logic is successfully implemented based on the extensively used techniques of DFT and KF. It is shown that the proposed method achieves significantly enhanced performance in AFD with the effective feature set in different scenarios and that the proposed AFD logic exhibits satisfactory dependability, security, and detection time using the real-time simulator.

Acknowledgment

The authors would like to acknowledge the meaningful discussions with Amin Ghaderi from Schweitzer Engineering Laboratories, Inc.

References

[1] J. Carr, "Detection of high impedance faults on multi-grounded primary distribution systems," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 4, pp. 2008–2016, 1981.

[2] Myeon-Song Choi, Seung-Jae Lee, Duck-Su Lee, and Bo-Gun Jin, "A new fault location algorithm using direct circuit analysis for distribution systems," *IEEE Transactions on Power Delivery*, vol. 19, no. 1, pp. 35–41, 2004.

[3] Jun Zhu, D. L. Lubkeman, and A. A. Girgis, "Automated fault location and diagnosis on electric power distribution feeders," *IEEE Transactions on Power Delivery*, vol. 12, no. 2, pp. 801–809, 1997.

[4] D. C. Yu and S. H. Khan, "An adaptive high and low impedance fault detection method," *Power Delivery, IEEE Transactions on*, vol. 9, no. 4, pp. 1812–1821, 1994.

[5] S. Gautam and S. M. Brahma, "Detection of high impedance fault in power distribution systems using mathematical morphology," *Power Systems, IEEE Transactions on*, vol. 28, no. 2, pp. 1226–1234, 2013.

[6] Q. Cui, K. El-Arroudi, and G. Joos, "An effective feature extraction method in pattern recognition based high impedance fault detection," in *International Conference on Intelligent System Application to Power Systems*, 2017, pp. 1–6.

[7] Q. Cui, "Interconnection protection of distributed energy resources using intelligent schemes," Ph.D. dissertation, McGill University, 2018.

[8] H. Laaksonen and P. Hovila, "Method for high-impedance fault detection," *CIRED - Open Access Proceedings Journal*, vol. 2017, no. 1, pp. 1295–1299, 2017.

[9] S. Huang, L. Luo, and K. Cao, "A novel method of ground fault phase selection in weak-infeed side," *Power Delivery, IEEE Transactions on*, vol. 29, no. 5, pp. 2215–2222, 2014.

[10] A. A. Girgis, W. Chang, and E. B. Makram, "Analysis of high-impedance fault generated signals using a kalman filtering approach," *Power Delivery, IEEE Transactions on*, vol. 5, no. 4, pp. 1714–1724, 1990.

[11] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[12] A. Ghaderi, H. A. Mohammadpour, H. L. Ginn, and Y.-J. Shin, "High-impedance fault detection in the distribution network using the time-frequency-based algorithm," *Power Delivery, IEEE Transactions on*, vol. 30, no. 3, pp. 1260–1268, 2015.

[13] A. Sedighi, M. Haghifam, O. Malik, and M. Ghassemian, "High impedance fault detection based on wavelet transform and statistical pattern recognition," *Power Delivery, IEEE Transactions on*, vol. 20, no. 4, pp. 2414–2421, 2005.

[14] S. Sahoo and M. E. Baran, "A method to detect high impedance faults in distribution feeders," in *IEEE PES Transmission and Distribution Conference and Exposition*, 2014, pp. 1–6.