IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

IEEE POWER & ENERGY SOCIETY SECTION

Received October 15, 2020, accepted October 26, 2020, date of publication November 2, 2020, date of current version November 11, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3035215

Event Sequence Model Application for Prioritization and Detection of Pre-Fault Waveforms on Power Distribution Lines

SANG KEUN MOON, (Member, IEEE), AND BYUNG-SUNG LEE
Korea Electric Power Corporation Research Institute, Daejeon 34056, South Korea
Corresponding author: Sang Keun Moon (sk.moon@kepco.co.kr)

ABSTRACT In this article, a conceptual approach is presented for pre-fault detection regarding the waveform analytic aspects of distribution monitoring and measuring devices. Also included are event patterns specifically arranged by feature and classification methods. A waveform class patterning algorithm on time-series is applied experimentally to field waveforms that were obtained for several years. The waveforms are processed with consideration of waveform classification and event sequence processing because these detect fault-related phenomena. This approach demonstrates that conspicuous patterns in fault-related sequences can be discovered by the data-driven structure described in the paper by designing the event structure depending on the network configuration. The application is applied in a pattern-learning and pattern-detection process so that integrating both approaches provides a meaningful consolidation for detecting abnormal conditions on distribution lines. This event-based fault prevention is employed using actual acquisition data from a domestic-scale distribution system and a unique sequence model is constructed to determine normal and abnormal conditions. Event index manipulation analysis on different risk levels defines the pattern and its impact on monitoring results. The proposed model guides recognition of event patterns and waveforms that can be pre-emptively detected in advance of distribution line failures.

INDEX TERMS Fault sequence detection, distribution system monitoring, pre-fault detection, power waveform classification, condition monitoring.

I. INTRODUCTION

Distribution network monitoring and control infrastructure has grid operation capabilities with regard to system automation and observation perspectives. However, their performance is limited because they lose event signals without managing recorded event logs on the system. Device enhancement with advanced technologies provides numerous possibilities for better system-condition monitoring. However, there have been no specific solutions yet when it comes to preventive maintenance and operation.

The fault prevention approaches that have been proposed before now for the distribution network have two parts: monitoring and maintenance. First, there is system monitoring and electrical measurement by means of sensor signals, which are not for predictive maintenance. This is because the monitoring devices adopted in distribution networks are fundamentally implemented to measure fault and status using real-time detection for conventional post-processing operation. Although the current monitoring infrastructure has provided major advantages and greatly enhanced system security, protection, and even economic benefits, this monitoring does not have the primary aim of fault prevention. The correspondence of pre-emptive actions and prevention of accidents and failures is highly questionable in this stage of infrastructure and studies. Second, preventive maintenance is conducted with asset management and reliability analysis on the distribution network. Maintenance and replacement of distribution components are regularly conducted by noticing the current status of equipment, historical data related to impending faults, or estimating degradation. Approaches for condition analytic prevention show definite limitations due to deficiencies of monitoring and of the data recorded from distribution components, as well as that the data are not verified sufficiently for fault prevention. The data recorded on distribution systems is repeatedly raising questions about accuracy and consistency [1]. To address issues mentioned above, research in preventive maintenance [2], [3] has suggested the

The associate editor coordinating the review of this manuscript and approving it for publication was Suman Maiti.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

use of event-prevention models by fault association analysis. Based on numerous historical fault data, this approach has emerged as a promising method for solving the condition problem. This method suggests the potential of revealing inner connections among faults and shows great potential to promote fault prevention techniques. Fortunately, the growing necessity for distribution network monitoring, data management, and data manipulation, means that new approaches are becoming possible to overcome these limitations using broad data domains, greater accuracy, and better techniques. Therefore, in this article, time-series data patterns are considered for use in fault prevention. To this end, a new integrated method is proposed by which to transform monitoring signals into event log structures. Numerous types of measurement data on domestic-level operation systems have been recorded over the last several years. However, the data have not been managed with practical regard to taking preventive action other than real-time operation and protection of the network. Here, a procedure is proposed for analyzing the measured signals that are already being used in operation. These signals are converted to designed event patterns to assess network health conditions, maintenance requirements before a failure, and inherent problems. Pre-fault analysis and predictive maintenance are specific interests in the field of electrical measurement because existing monitoring infrastructure is predominantly based on waveform data.

Electrical waveforms in the distribution network are considered essential measured signals in this article. Their measurement is recorded by feeder remote terminal units, digital relays, and power quality (PQ) meters because these devices are already installed. In waveform analytics, post-fault measures are commonly conducted in the field. The use of cause identification of faults has been proposed [4], [5] to assist restoration in case of momentary or temporary faults. This would guide operators in deciding on mitigation actions, and afterwards, in achieving optimized preparation and dispatching field engineers. Patterning the waveform itself is also provided in waveform analytic research to identify the system status at a given moment [6]–[8]. On the subject of detection and classification of electrical line faults, studies have conducted for effective fault protection algorithm using feature extraction and classification based on signal processing approaches [9]–[11]. Operation and monitoring based on the waveform signal provides status-event logs on distribution lines (DLs). If specific patterns can be recognized from a recorded dataset, advanced methodologies such as pattern learning and feature extraction can be applied to predict upcoming events. From the perspective of online monitoring [12], research has described the investigation of events by having a large number of measuring devices to understand better the fault-degradation relationships from electrical waveforms [13].

Fault events are a representative detection signal provided through several electrical measurement devices. Therefore, if the event sequences are not important anymore after detecting faults in the operation, then the event log is generally neglected. However, the event-data-driven method can be applied selectively to evolve knowledge-based fault diagnosis into predictive systems [14], [15]. The dynamic fault prediction is able to provide decision-making basis for practical condition-based operation and maintenances [16]. The effectiveness of dynamic early warning and incipient fault prediction in sub-health status of in-service power transformers has been proved. Incipient faults are usually detected on distribution system as a disturbance with a comparatively less current and shorter time period of less than 1/4th cycle to 4 cycles. These current changes are not detectable by common protection devices due to its instant time and less amplitude contrary to fault current detection scheme [17]. Field results of incipient fault examples on distribution cable systems are described in [18] to explain the incipient fault is common indicators of component degradations.

Condition monitoring is also counted as an event-prediction approach based upon the assumption that an individually measured signals can form event patterns. In terms of the condition of distribution equipment, electrical waveform events taken as power quality measures, can be considered for a predictive maintenance strategy to be implemented in advance before complete failure occurs [19]. Due to the nature of steady-state data, measuring devices on DLs are to indicate long-term variations of various power system conditions. Although, fairly little of the steady-state data have been used efficiently to track the power system performance. Extensive research on PQ disturbance analysis has shown that electrical signatures in voltage and current waveforms are suitable as non-invasive parameters for online-condition monitoring [20].

This article describes pre-fault phenomena that can be detected prior to the system failure using event patterns. For examining massive measurement waveform data, sequence-oriented signal feature extraction was conducted in a prior stage, as was learning the data feature patterns needed to obtain event classes. The method employs zero sequence current signals to discover transient waveforms by machine learning-based pattern classification models [21]. Distinct patterns of fault-related phenomena were discovered by the data-driven model. The proposed model has sequence structure and it constructs an event sequence model (ESM) by analyzing field-measured waveforms. In practice, applying generalized patterns to compare with distribution conditions, the ESM time-series analysis is used to detect pre-event patterns. The subjects addressed in this article cover condition monitoring and the detection of predictive patterns on the distribution network as a fault-anticipation method. The pre-fault duration and event characterization are also represented for timing the sequence of pattern detection. Moreover, structural analysis was conducted to reflect the qualities of various types of measurement devices, and their locational dependencies, on connected distribution lines.

In consideration of power-system characteristics, a method of structural pattern analysis is proposed here for a power distribution system. The remainder of the paper is organized as follows:

1) Concept and methodology of the event structure
2) Data manipulation of extracting fault signals by the ESM
3) Pre-fault pattern detection with risk value determination

II. DISTRIBUTION SYSTEM EVENT STRUCTURE

A. CLASSIFICATION OF MV WAVEFORM FEATURES

Waveform events are an essential indicator for recognizing the status of DLs. In terms of a distribution disturbance and the fault circuit status, waveform analysis is conventionally applied to classify the event occurrence. The classification for event waveforms has been applied to propose a practical disturbance classifier for empirical distribution monitoring devices [21]. The previous study correspondingly shows that features of the waveforms and of the classifier could potentially be implemented for waveform shape identification, to be transformed to event classes.

Distribution system events are defined as waveform classes and triggered conditions that contain a class of fault transactions. Each event is a subset of combined classification and triggered data as a consequence of determining measured waveforms. With this approach, functionalities of DL monitoring devices are implemented by updating the essential parameters for feeder-device classification models. First, after a device obtains event waveforms based on trigger conditions, previously trained models classify the waveforms into event classes. Next, because the class is clearly independent of time due to constrained recording cycles, the event classes are arranged to construct the event structure as the ESM.

B. EVENT AND CLASS SEQUENCE MODEL

In this section, subsequent to obtaining event classes from the measurement devices installed in the field, the use of time-sequence event processing is explained to employ additional pattern-detection work. An event pattern is referred to a fault-triggered database consisting of the network structure and hierarchy. The application describes event classes not only having the condition of the electrical-system status at the moment an event occurs, but also identifying any discernible sequence pattern using a timely arranged model. The fault associated analysis in the model is designed as an online detection model initially; however, manipulating intermittent events to reveal a recognizable pattern is considered predictive action so that the event model detects sequence patterns in the event model. For example, the fault-associated approach returns results in the form of waveform disturbance classes that have already been classified in the monitoring phase. In addition, the voltage and current classes are separately marked on the ESM along with triggered information of binary values. Therefore, all events in the sequence construct event patterns in chronological order, as shown in Fig. 1. This illustrates a conceptual schematic diagram of the proposed ESM upon which the classes with event signals and risk values on the DL are based. The measured classes of $\mathrm{C}_t^{\mathrm{V}}$ and $\mathrm{C}_t^{\mathrm{I}}$ are the classified voltage and current-disturbance vector, respectively. The event vector E_t^2 of the fault and trigger values, defines types of fault current as a lozenge shape in Fig. 1, so that the model determines pre-event sequence patterns prior to a fault. Accordingly, specified ranges of the pattern interval can be stressed according to the DL pattern importance based on the time period of the fault events then provides the pattern duration for a preventive action. The ESM is then implemented for preventive pattern-detection of distribution faults, with revision of a fault pattern after a fault occurs, and vice versa.

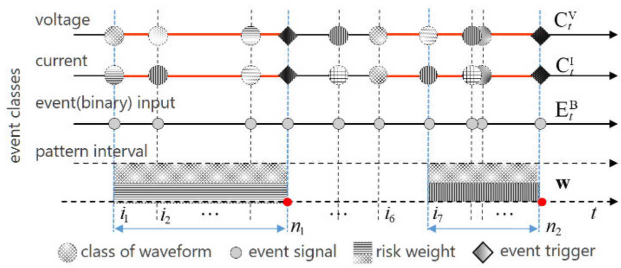


FIGURE 1. Schematic diagram of the distribution-waveform-event process structure.

III. EVENT SEQUENCE MODELING PROCESS

A. STRUCTURE-BASED EVENT SEQUENCE MODEL

Distribution system structures in practical applications are based on network design and planning criteria. In typical radial feeder networks, the configuration is operated through DLs connected to a certain supplying substation. Occasional DL interconnections by switch gears are designed to allow interacting electrical flows between DLs. The measurement devices are installed depending on the location of sectional operations. The event waveforms obtained from the devices represent mutual relations between the configuration of DLs and substations. In this article, to recognize the DL events more precisely, the ESM was selectively adopted based on the hierarchical structure of conventional DLs and on the time-series sequences of the waveform classes.

B. WAVEFORM EVENT-STRUCTURE CATEGORIZATION

Primarily, class labels such as normal and abnormal are contextual according to what events are required to be classified for a distribution system. For a normal event, the classification model ignores the event and is triggered only sporadically, although the event may become abnormal afterward. On the matter of event-state changing, the proposed method focuses on time interval and configurational analysis by analyzing the event sequence occurring on the DLs. Finding apparent symptoms before a fault and patterning for next events is the purpose of this approach. Moreover, because the waveform data generally provide a more complete picture than logical determination based on the field devices, the method uses waveform-feature classification for proposed conditions of a distribution network.

As illustrated in Fig. 2, the distribution system structures are depicted as a network hierarchy containing three concept levels. The structure represents a substation with several DLs, a DL including monitoring devices, and a single-feeder device on the DL separately. All of the components are included at the substation level in the proposed event-structure model according to their hierarchical relationship and these are divided into structure levels. The monitoring of integrated distribution according to the proposed disturbance classes is designed to detect disturbance events in accordance with the network configuration.

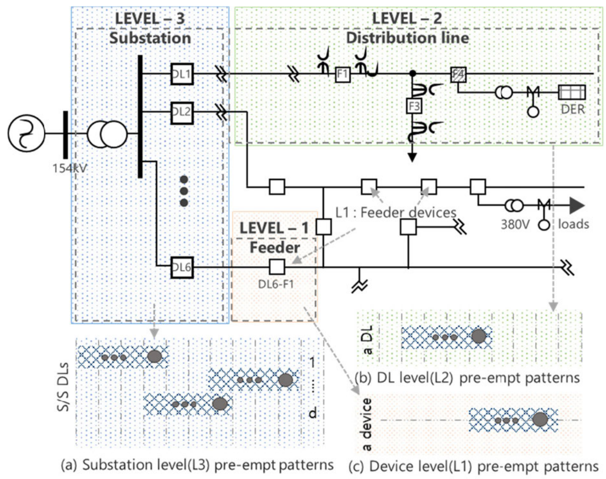


FIGURE 2. Explanatory depictions of network categorization levels with monitoring devices and the ESM class arrangement.

1) LEVEL-1: A FEEDER DEVICE

The ESM on level-1 has event structures consisting of classified labels of individual measurement devices on a time scale. Because the label has classes of voltage and current, the events on the ESM describe combined information at the time of each event. Therefore, the sequence is depicted subsequent to events, following the disturbance by several moments. A device at the feeder can be drawn as a single line that consists of classified multiple labels, as shown in Fig. 2(c).

2) LEVEL-2: A DISTRIBUTION LINE

Monitoring devices are illustrated as being connected with a DL on the same substation as a merged line with feeder devices, as shown in Fig. 2(b) and Fig 3(b). Thus, all feeder device events on the DL are integrated. For detailed event analysis, expanding the merged classes to a class representation is applied to identify specific events on the DL (depicted in Fig. 3c). Because most events are generally combined in the model, the class-separation analysis assists in distinguishing the ESM pattern and recognizing the distribution condition.

3) LEVEL-3: SUBSTATION

To broaden the viewpoint of the monitoring, the condition structure integrates ESM level-2 with the substation structure of ESM level-3, as illustrated in Fig. 3(a). Apart from all connection and independent event cases, every event on the network is observed by the ESM process. This includes interconnected events identified by loop operations and switching between DLs.

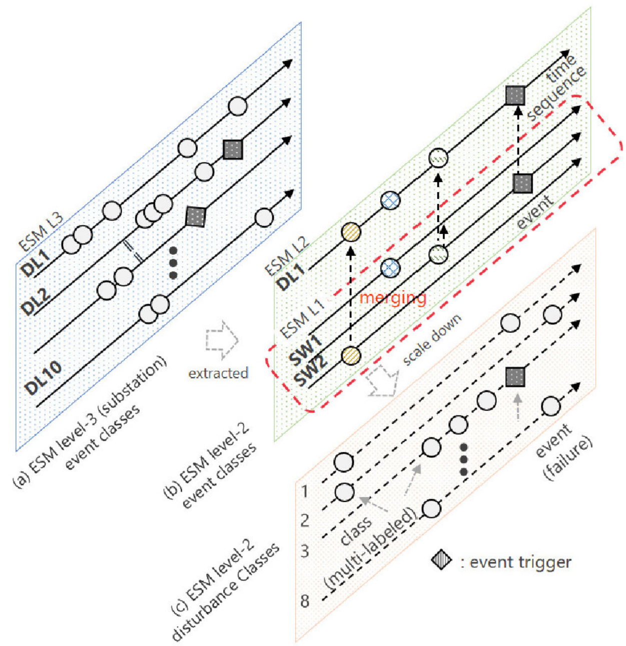


FIGURE 3. ESM hierarchy illustration of DL integrated level-3 to level-2 of detailed classes on a DL.

C. PRE-EVENT EXTRACTION ON ESM STRUCTURES

The event classes in the ESM have classified waveform values along with time-stamp and event-types—disturbance, fault, et cetera. Regarding the distribution power supply and operation, structural identification elements such as DL, switch number, and line-connection information are dependently applied to the ESM analysis. The arranged event value using the ESM structure is described in this section to construct a time-sequential matrix of the classified events. Each obtained and classified waveform is defined as a class C having its own classification number based on the machine-learning classifier [21]. Therefore, C is annotated with time and DL indices to comprise a sequence matrix:

$$C_{i,j}^{\mathrm{L}} = \text{classification}(\,) \tag{1}$$

where the function ''classification'' represents the classified labels of waveform disturbances; the index i is the time-based sequence in order of occurrence ($\forall i \in \{1, \ldots, t\}$ where t is the overall time of the provided dataset). The DL index j represents and $\forall j \in \{1, \ldots, d\}$ from the interconnected substation. The identifier L indicates that $\mathrm{L} \in \{\mathrm{V}, \mathrm{I}\}$ where V and I represent the voltage and current, respectively. Equation (1) is obtained by the class set in which $C_{i,j}^{\mathrm{L}} \in \{1, 2, 3, 4, 5, 6, 7, 8\}$ and the classes are defined as index numbers corresponding to waveform labels (1: sag, 2: swell, 3: interruption, 4: flicker, 5: oscillation and impulse, 6: notch or transient, 7: spike, 8: harmonics). The class $C_{i,j}^{\mathrm{L}}$ does not represent exact distribution system phenomena, but instead indicates waveform-shape patterns within the classification boundaries that are determined by the classification model. For example, the class notch covered shapes of most transients and very short disturbances whereas the oscillation class is the shape of short-time changes in magnitude following the nominal frequency. The bipolar oscillating and impulse waveforms are classified as representative types of the class.

Waveform data acquired from several distribution monitoring systems were used to build a substation scale event matrix that has time-series sequences proposed as the ESM level-3 matrix as follows:

$$\mathbf{M}_{t,d}^{\mathrm{L3}} = \left[\begin{array}{ccccc} C_{1,1}^{\mathrm{V}} & C_{1,2}^{\mathrm{V}} & \cdots & C_{1,d-1}^{\mathrm{V}} & C_{1,d}^{\mathrm{V}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{t,1}^{\mathrm{V}} & C_{t,2}^{\mathrm{V}} & \cdots & C_{t,d-1}^{\mathrm{V}} & C_{t,d}^{\mathrm{V}} \\ \hline C_{1,1}^{\mathrm{I}} & C_{1,2}^{\mathrm{I}} & \cdots & C_{1,d-1}^{\mathrm{I}} & C_{1,d}^{\mathrm{I}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{t,1}^{\mathrm{I}} & C_{t,2}^{\mathrm{I}} & \cdots & C_{t,d-1}^{\mathrm{I}} & C_{t,d}^{\mathrm{I}} \end{array}\right] \quad (2)$$

where, L3 denotes ESM level-3. Equation (2) shows DL divided columns having synchronized time by class sequences on the ESM; therefore, every event on the DL is simultaneously analyzed at the moment of fault and when the disturbance occurs. Subsequently, the matrix is concatenated to provide corresponding voltage and current sequences. The ESM level-2 matrix is sequentially derived to apply the DL select vector $\mathrm{E}_d^{\mathrm{A}}$ where $\mathrm{E}_d^{\mathrm{A}} = [E_1^{\mathrm{A}}, E_2^{\mathrm{A}}, \cdots E_d^{\mathrm{A}}]^{\mathrm{T}}$ with T transpose. The vector $\mathrm{E}_d^{\mathrm{A}}$ denotes selected binary fault values in the DLs. Consequently, the extracted ESM level-2 is derived (as follows) as the voltage and current matrix:

$$\mathbf{M}_{t,d}^{\mathrm{L2}} = \mathbf{M}_{t,d}^{\mathrm{L3,V}} \cdot \mathrm{E}_d^{\mathrm{A}} \,|\mathbf{M}_{t,d}^{\mathrm{L3,I}} \cdot \mathrm{E}_d^{\mathrm{A}} \quad (3)$$

where L2 represents the ESM level-2 and (3) extracts events on the DL level. The selected event is the instantaneous and permanent fault in the time sequence from the dataset, which is included in the **M** matrix. Accordingly, the event moment is represented as the event vector $\mathrm{E}_t^{\mathrm{B}}$, which contains multiple fault events in sequence ($\mathrm{E}_t^{\mathrm{B}} = [E_1^{\mathrm{B}}, E_2^{\mathrm{B}}, \cdots E_t^{\mathrm{B}}]$). Hence, the vector $\mathrm{E}_t^{\mathrm{B}}$ indicates DL event occurrences where more than a single fault event is detected on the DL. For the purpose of identifying each event as a candidate of the pre-fault pattern, the vector $\mathrm{E}_t^{\mathrm{B}}$ of the triggered time t is transformed to the sequence vector $\mathrm{E}_q^{\mathrm{B}}$ ($\mathrm{E}_q^{\mathrm{B}} = [E_1^{\mathrm{B}}, E_2^{\mathrm{B}}, \cdots E_q^{\mathrm{B}}]$). The sequence q has indices of the event occurrence time; thus, the calculation process can specify the pre-fault interval through the calculation process. For instance, the fault index $\mathrm{E}_q^{\mathrm{B}} = [1, 3, \cdots 9]$ indicates that the first, third, and ninth sequences have fault events. Furthermore, based on the $\mathrm{E}_q^{\mathrm{B}}$ of the event sequence, the pre-fault pattern extractor $\mathbf{E}_{t,s}^{\mathrm{C}}$ of the occurrence time t and the pre-fault interval s matrix for each $q$event can be derived as follows:

$$\mathbf{E}_{t,s}^{\mathrm{C}} = \left[\lambda_{i,k}\right]_{t \times s} \quad (4)$$

where, $\mathbf{E}_{t,s}^{\mathrm{C}}$ has the time t by the sequence s matrix of the event extraction, which has the s diagonal expansion for the pre-patterns—a reversed diagonal with q of $\mathrm{E}_q^{\mathrm{B}}$ by k order columns where $k = [1, \ldots, s]$. With respect to the observing interval s, developing precursor events are determined by analyzing the dataset and discovering how many pre-patterns are considered essential. Around 4-10 consecutive events are patterned in this model. The pre-event indicator $\lambda_{i,k}$ is the s shift tensor, which is obtained as:

$$\lambda_{i,k} = \begin{cases} 1, & i = E_q^{\mathrm{B}} - s + k - 1 \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

The sequence index i is derived with $\mathrm{E}_q^{\mathrm{B}}$ by subtracting the previous s interval and shifting the k steps until the interval s is reached. The indicator $\lambda_{i,k}$ is obtained from every q event pre-pattern of a sequence prior to the fault moment; thus, each individual pre-pattern is extracted by the $\mathbf{E}_{t,s}^{\mathrm{C}}$ extractor. Consequently, the extracted pre-fault pattern can be obtained by multiplying the $\mathbf{E}_{t,s}^{\mathrm{C}}$ and the ESM matrix.

$$\mathbf{M}_{t,s}^{\mathrm{L2}} = \mathbf{M}_{t,d}^{\mathrm{L2}} \cdot \mathbf{E}_{t,s}^{\mathrm{C}} \quad (6)$$

$$\mathbf{M}_{t,s}^{\mathrm{L1}} = \mathbf{M}_{t,d}^{\mathrm{L1}} \cdot \mathbf{E}_{t,s}^{\mathrm{C}} \quad (7)$$

where, $\mathbf{M}_{t,s}^{\mathrm{L2}}$ and $\mathbf{M}_{t,s}^{\mathrm{L1}}$ represent the pre-event extracted matrices for a DL level event and a switch end-device level event, respectively. The extraction process is iterated for the q event set.

D. INDEX MANIPULATION FOR PRE-EMPT EVENTS

The proposed ESM has index values to emphasize the pre-event importance in advance of actual failures of power lines. The importance index, which can be interpreted as the pre-empt value, is represented to indicate DL vulnerabilities by means of converging event information regarding a system fault. Thus, the index is determined to inform of anticipated disturbances, associated with degrees of importance. The proposed method evaluates two types of values from the systematically obtained data patterns. The first is the class characteristic index of the voltage and current values, which counts events and collects specified classes in the DL. Hence, the index reveals more credible fault relevant signals by examining the ESM. Second is the pre-fault duration and the span index, which are derived from the interval of the time sequence on the assumption that an event indicated in the interval has a considerable pattern regarding how closely the pre-fault event is located by the fault.

1) CLASS CHARACTERISTIC INDEX

With this approach, the class characteristic index is proposed to model the event-count value. Accordingly, voltage and current indices are considered with respect to class importance to obtain the weight value. From the DL event manipulation, the class index $\mathrm{W}_{q,\mathrm{VI}}^{\mathrm{A}}$ is derived to explain class counts and importance by the time sequence interval. After the pre-fault-pattern extracting procedure, pattern events are generated where $n = [1, \ldots, q]$ because the sequence q represents the total number of events in the set. The types of certain repetitive events, classification classes, and occurring frequencies are indexed as:

$$\mathrm{w}_{q,\mathrm{VI}}^{\mathrm{A}} = (\textstyle\sum_{k=1}^{s} w_k^{\mathrm{c}})^{\mathrm{T}} \cdot \mathrm{w}_{\mathrm{VI}}^{\mathrm{B}}, \quad \forall n \tag{8}$$

where, w_k^c is the specified class importance within the pattern interval k where $k = [1, \ldots, s]$. For example, the class importance w_k^c has the weights of $\mathrm{w}_s^{\mathrm{c}} = [1, 2, 1, 2, 2, 3, 3, 1]$, which correspond to $C_{i,j}^{\mathrm{L}}$. The class importance determined by the classification label is empirically determined in advance. On the other hand, the characteristic index $\mathrm{W}_{\mathrm{VI}}^{\mathrm{B}}$ has signal importance between the voltage and current measurement from the local device. In this sense, the voltage and current characteristics are applied in this article as $\mathrm{w}_{q,\mathrm{VI}}^{\mathrm{A}} = [1,\ 1.5]$ for the reason that a large amount of fault events are identified experientially by the current measurement of short-circuit currents. The class index $\mathrm{w}_{q,\mathrm{VI}}^{\mathrm{A}}$ is obtained as an s-sequence index by the individual n fault event of multiplying the row vectors by each other.

The class index implies that occurring frequencies of certain events indicate network instability and a potential system fault in connection with past events. The combined proportion typically has higher current values than voltage values. The reason is that impulse waveforms are frequently trigged by discharge-current phenomena in the ESM extraction procedure. In contrast with conventional PQ detection, the class characteristic also shows that the current class is more essential, for example, that a pre-fault has considerably more impulse transient patterns than other waveforms, which fact is applied to the class weight.

2) PRE-FAULT INTERVAL INDEX

On the other hand, the time-interval index is derived from the sequence time calculation. The interval is exclusively designed for a fault-event pattern to express the importance of the time interval between the event and failure. Therefore, the interval is emphasized by increase in the importance of approaching the fault moment. This viewpoint is based on the notion that fault-adjacent events have high potential for failure to explain the fault condition; thus, certain event classes have more importance than others with regard to fault associations and exclusion of unrelated events. The considered duration weight $\mathrm{w}_s^{\mathrm{C}}$ of the pre-fault duration is obtained as follows:

$$\mathrm{w}_s^{\mathrm{C}} = \begin{cases} (\mathrm{r}_s)^{1+\alpha}, & (\mathrm{r}_s)^{1+\alpha} \geq 1 \\ 1, & \text{otherwise} \end{cases} \tag{9}$$

where, r_s is the step index vector according to the sequence of the s interval. The additivity weight α is the interval gradient parameter for the purpose of emphasizing the event-class adjacent to a fault moment. Selecting a feasible α is highly dependent on finding an event class related to the fault event, and 0.5 was applied in this study after trial calculations using the obtained dataset. In addition to the step index, the term $\mathrm{w}_s^{\mathrm{C}}$ has a pending constant "1", which adjusts all values to positive when applying integrations of several indices.

Despite the fact that the interval s is the variable that is constantly changing, the index $\mathrm{w}_s^{\mathrm{C}}$ should consider the index magnitude that is in a fixed range to maintain a similar scale of importance. This is because the importance is exponentially weighted in the patterning process. Therefore, the step index r_s is proposed by the step variable k starting with 1 to the interval s. The target constant σ of the ceiling value at the fault moment where $\mathrm{r}_s = \{k \cdot (\sigma \cdot s)^{-1}\}_{k=1}^{s}$. The constant value ($\sigma = 2$) is determined in accordance with the pre-empt pattern interval of 4-7 sequences practically applied in this study. Meanwhile, because the interval s has a different time interval according to the difference in time of occurrence of events in each s, another index is required to reflect variation in time. The span index calculates time between initial signal and failure in order to quantify the time duration. The index $\mathbf{w}_{s,q}^{\mathrm{D}}$ is presented to the s as the same value, but different indices are placed on each n fault event.

$$\mathbf{w}_{s,q}^{\mathrm{D}} = \left[\Delta\tau_{E_n^{\mathrm{B}}}\right]_{s\times q}^{\beta}, \quad \forall n \tag{10}$$

$$\Delta\tau_s = \left\{\frac{1}{\tau_{E_n^{\mathrm{B}}} - \tau_{E_n^{\mathrm{B}}-k}}\right\}_{k=1}^{s} \tag{11}$$

where, τ indicates the time stamp of the n of every event; The event index E_n^{B} specifies sequence indices at the τ where $n = [1, \ldots, q]$. The time interval $\Delta\tau_s$ under the interval s is acquired by extracting the time between each event and the moment of fault. The interval $\Delta\tau_s$ has a minimum starting time interval necessarily longer than 50 ms where $\forall\Delta\tau \geq 50$ ms, which is 3 cycles of the circuit-breaker-operation time limit. Because the events occur repetitively and simultaneously, the event labels tend to overlap for the same event. For example, a phase disconnection or failure might have a following ground-fault event that is not individually considered, especially in conventional monitoring.

The graphical depiction in Fig. 4(a) illustrates that the values chosen at the duration index $\mathrm{w}_s^{\mathrm{C}}$ do not indicate a fixed value. The time importance is gradually increased depending on the weight parameter α as the pre-empt pattern sequence proceeds. Contrary to the constant sequence in Fig. 4(a), the span index in Fig. 4(b) regards the amount of time. In accordance with time intervals expanded from the moment of the fault, the longer the interval is, the less importance is applied in the event of E_n^{B}. The weight β adjusts the time index regarding the time-span differences. The time relationship between events in the pattern interval is varying and inconstant; thus, it is impossible to model the index linearly. Moreover, with respect to events adjacent to faults, their importance is related to an exponent and subordinate importance is considered for distance from the event.

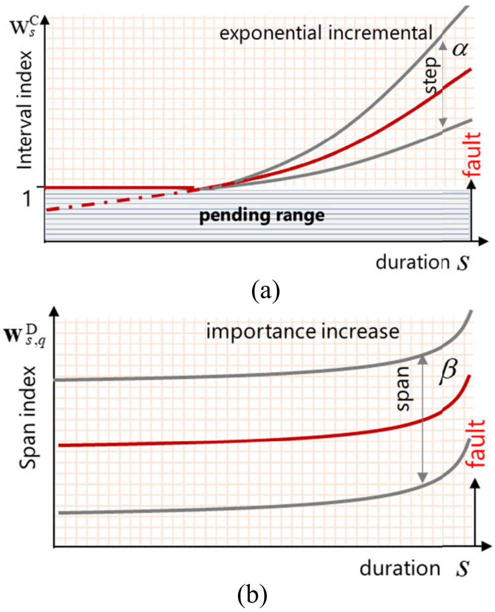


FIGURE 4. Graphical depiction of event indices through the importance value transfers: (a) Duration index and (b) Span weight.

Although the two proposed indices have different scales of s and time, the model should treat the indices as having importance as integrated sequences in order to find pre-empt events in the ESM structure. Therefore, the $\mathrm{w}_s^{\mathrm{C}}$ and $\mathbf{w}_{s,q}^{\mathrm{D}}$ indices are merged with q sequences in s duration. Consequently, the combined pre-fault duration and span index is obtained using element-wise multiplication of these indices. The finalized integrated importance is derived as follows:

$$\mathbf{w}_{s,q} = \mathrm{w}_q^{\mathrm{A}} \circ \mathrm{w}_s^{\mathrm{C}} \circ \mathbf{w}_{s,q}^{\mathrm{D}} \tag{12}$$

The pre-pattern index is subject to discovery of features among the patterns, and it extracts considerable waveforms from the aspect of anticipating distribution line failures. The indices will systematically be adjusted by detecting and analyzing event occurrences in the operational environment. The preventive action is possibly implemented by the index with accordance of the pre-patterns. Defined priorities guide target patterns from the index then the pattern detection will alarm the system operator for preventive actions.

IV. RESULTS ON EMPIRICAL SYSTEM DATA

A. SYSTEM CONFIGURATIONS

The proposed system configuration was based on conventional measurement data from monitoring devices in the field. Among various types of measurement waveforms, the system obtained triggered events of 3-phase voltage and current waveforms from the installation to the present moment. The trigger condition already implemented at the field device was verified and the fundamental underlying IEEE standard 1159 and 181 [22], [23] of power quality and fault detection under which the field-measured waveforms were recorded, were based on the IEEE C37.111 standard [24]. Approximately 23,000 waveforms from 34 substations with 598 DLs were acquired and processed. These contained 1316 fault events in the interval from 2014-2018.

Furthermore, the event classification was performed in regard to the triggered waveforms, which have types of abnormal, fault, and voltage variation triggers, which monitoring devices captured. The classified voltage and current classes are shown in the results. Thus, a pair of classes between fault events and classified waveforms were concurrently treated to define the distribution system conditions. Once the events were combined, the pre-fault pattern analysis was provided for the detection process.

B. ESM PRE-PATTERN RECOGNITION CASE

After the classification of waveforms, the data developed were utilized to build the ESM structure proposed in this article. Because the obtained time-series event classes were not arranged spatially and temporally, the model first organized the data for structural analysis of the distribution system. Common waveform types were determined as classes in the classification process. In addition to these common types, event data such as electrical fault currents (instantaneous and permanent) were simultaneously applied to the dataset to extract the event duration in the ESM.

The empirical results were derived by the proposed ESM model using substation levels and distribution lines, respectively. Among the datasets, representative ESM level-3 of the $\mathbf{M}_{t,s}^{\mathrm{L3}}$ results are shown in Fig. 5. As illustrated in Fig. 5, the events of a substation "Geumchon" were arranged in horizontal time-sequences and vertical DL number segments. In the ESM level-3, y-axis lines appeared as divided DLs corresponding to synchronized time of class sequences. The fault pre-patterns $\mathbf{M}_{t,s}^{\mathrm{L2}}$ were displayed as colored squares in accordance with the class diversity. Extracted patterns based on the fault and trigger values were marked as colored squares followed by lozenge-shaped fault events at the pattern end. A pattern interval with a specified duration of 5 sequences was applied to extract the pattern. In the sequence map of Fig. 5, two separated sections are presented by a dotted line

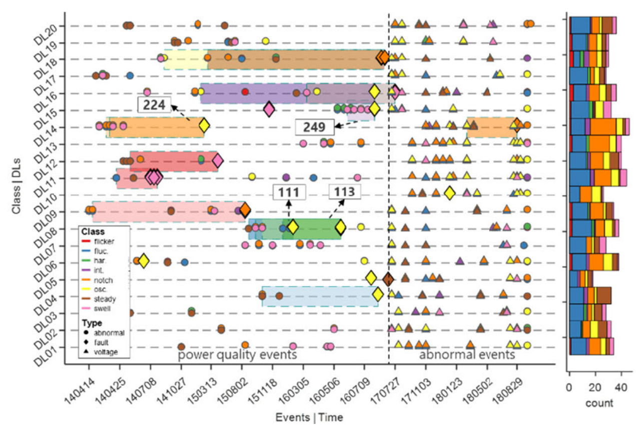
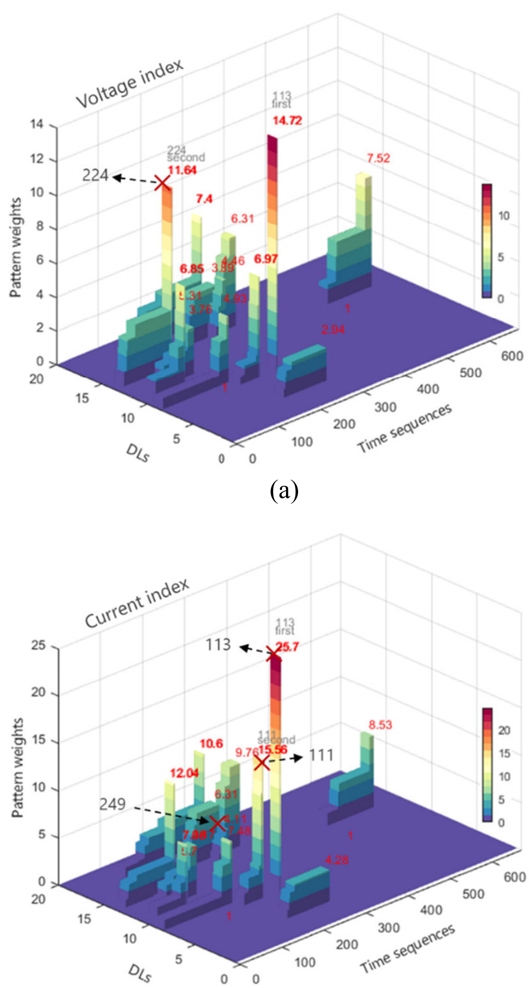


FIGURE 5. Classified time-sequence events and extracted patterns with the class histogram of ESM level-3.

in the middle. The left and right sides display power-quality-triggered classes and abnormal triggered classes, respectively. Because the sections indicate pre-patterns caused by the system fault, system malfunctions had more PQ-class patterns than the abnormal triggered ones did. The class histogram is also depicted on the right side in Fig. 5 to express the frequency of event occurrence. The event count showed that majority fluctuation classes were caused by small deviations such as load changing and switching. Thus, class selection was required to exclude unimportant classes and to consider the crucial classes directly dependent on the system failure according to the ESM pattern detected.

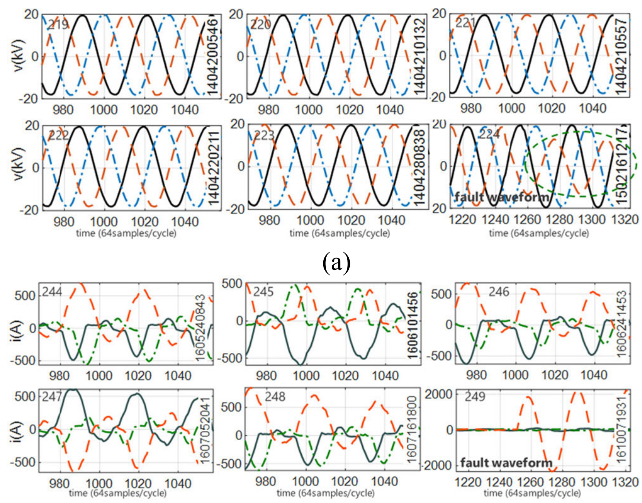
The pattern importance was predictable by means of event classes and time intervals. The patterns in Fig. 5 illustrate that long periods of time had less relation with the system fault. Therefore, the importance index specified the pre-patterns to signify considerable patterns, as shown in Fig. 6. Among the extracted pre-patterns for the substation, 2-3 particular sequences were distinguished from the ESM with importance indices. With the importance index, the order of priority through the pattern interval and the number of waveforms were found. In general, a higher index indicates higher order of precedence to explain the sequence of failure relevance. As shown in Fig. 6, cross marks on the bar graph give the order of the indices with event numbers followed by index values in red. From the waveform pattern, according to importance of the voltage and current in Fig. 6, representative pre-pattern waveforms are illustrated with respect to the primary and secondary current indices of 111 and 113 in Fig. 6(b). The indications show that specific waveform events due to current measurements have more expressive signals of system failures using the empirical approach. Although voltage indices 224 in Fig. 6(a) showed no particular abnormal conditions according to the followed waveform analysis, as illustrated in Fig. 7(a). In contrast, the pre-pattern waveforms of the current event are illustrated as the voltage and current waveforms in Fig. 8(a) and (b). The waveforms in the sequence contain different pattern signals even though the class patterns have been determined. The current pattern in Fig. 8(b) found two abnormal impulse fault currents on the day of the three-phase



(a)

(b)

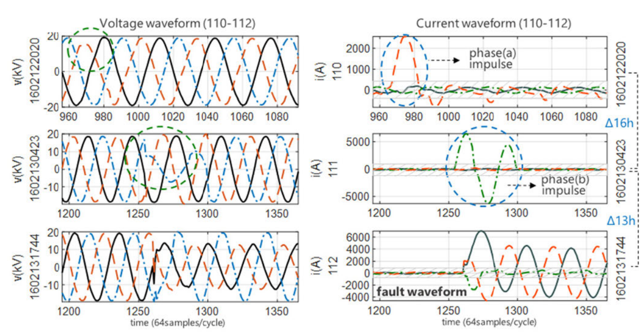
FIGURE 6. Pre-pattern index graphs of the pattern importance from the Geumchon Substation.



(a)

(b)

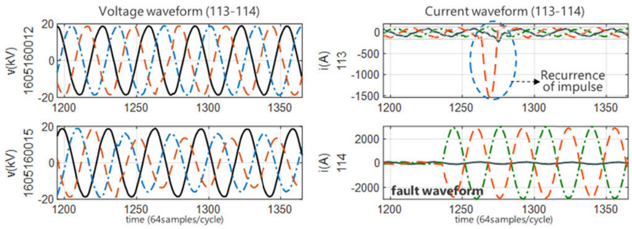
FIGURE 7. Additional waveform pattern representations with respect to the extracted patterns.



(a)

(b)

FIGURE 8. Extracted pattern waveform illustrations of representative pre-patterns (voltage, current).



(a) (b)

FIGURE 9. Post event pattern waveforms that result in system failures.

failure and displayed 13 and 16 hours of prior time intervals, which made the current index higher. The event was recorded on the fault report as a winding burning of the overhead transformer due to a lightning strike in stormy weather. Therefore, the pattern is eventually applied to the prediction for future events then provides more time margin for preventive actions before the fault.

The index 113 waveform is followed by 111 and showed the system fault two days later as 114 in Fig. 9(b). This represents recurring impulse and fault currents because the failure had not been cleared properly in spite of continuous pre-fault events. The cause of the three-fault event was recorded as unidentified. Despite this fact, the pattern importance index detected the pre-pattern as the highest index (25.7) from the event sequence indicated in Fig. 6(b). Additionally, event sequence 249 was discovered from the waveform patterns shown in Fig. 7(b). This represented the harmonic distortion that occurred repetitively a day prior to a phase fault. The analysis was applied to adjust the class index of a harmonic element having additional weight to identify the subsequent waveform and pattern detection.

V. DISCUSSION AND CONCLUSION

This research work includes methods and empirical applications from a proposed model and method found using actual field data from a domestic distribution operation and management infrastructure. Because the suggested pattern exploration model defined fundamental signal features of actual event patterns from the field data, the model was verified by presenting how the model was able to handle phase signals for condition recognition and pre-empt detection. The model demonstrated that voltage and current waveforms were appropriately classified into disturbance classes and that it also provided waveform sequence patterns reflecting what possibly happened in preceding system outages.

In this study, the proposed method was developed by classifying DL conditions of the disturbances in real networks and the signals could provide unique patterns for recognition of abnormal operations and fault conditions in the learning mechanism of the classification model. The learned model would necessarily replace the trigger and detection logic of the field measuring devices. This affirms the research goal of enabling the classification system to be updated by waveform modeling and feature-set reconfiguration using the learning process. The research also found that, in distribution system networks, conditional and monitoring structures of the ESM configurations are important for recognizing DL conditions. Despite the method is required to be examined extensively on the field, time-scale and hierarchical analysis are verified to this patterning approach. Even though the experimental results in this study showed limited examples of these pre-patterns, a great amount of detection and extracted event data is still in the process of discovery of particular features.

There are no other systems that have been applied in the field because of insufficient data and validation difficulty. The preventive maintenance based on waveform pattern analysis and time sequences is necessarily be expended the field application to improve the recognition and restricted applications. However, despite that the proposed method is challenging, the approach is verified by this research and it now provides the potential to implement the methodology to create a pre-pattern detection system for distribution operation and maintenance.

REFERENCES

[1] J. H. Jürgensen, L. Nordström, and P. Hilber, "Estimation of individual failure rates for power system components based on risk functions," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1599–1607, Aug. 2019.
[2] Y. Liu, S. Gao, and L. Yu, "A novel fault prevention model for metro overhead contact system," *IEEE Access*, vol. 7, pp. 91850–91859, 2019.
[3] L. Zhang, H. Chen, Q. Wang, N. Nayak, Y. Gong, and A. Bose, "A novel on-line substation instrument transformer health monitoring system using synchrophasor data," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1451–1459, Aug. 2019.
[4] Y. Cai, M.-Y. Chow, W. Lu, and L. Li, "Statistical feature selection from massive data in distribution fault diagnosis," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 642–648, May 2010.
[5] M. A. Shuvra and A. Del Rosso, "Root cause identification of power system faults using waveform analytics," in *Proc. Clemson Univ. Power Syst. Conf. (PSC)*, Sep. 2018, pp. 1–8.
[6] T. Hong, D. Deswal, and F. de Leon, "An online data-driven technique for the detection of transformer winding deformations," *IEEE Trans. Power Del.*, vol. 33, no. 2, pp. 600–609, Apr. 2018.
[7] M. H. J. Bollen, "Algorithms for characterizing measured three-phase unbalanced voltage dips," *IEEE Trans. Power Del.*, vol. 18, no. 3, pp. 937–944, Jul. 2003.
[8] S. Kulkarni, S. Santoso, and T. A. Short, "Incipient fault location algorithm for underground cables," *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1165–1174, May 2014.
[9] S. R. Ola, A. Saraswat, S. K. Goyal, S. K. Jhajharia, B. Rathore, and O. P. Mahela, "Wigner distribution function and alienation coefficient-based transmission line protection scheme," *IET Gener., Transmiss. Distrib.*, vol. 14, no. 10, pp. 1842–1853, May 2020.
[10] G. S. Yogee, O. P. Mahela, K. D. Kansal, B. Khan, R. Mahla, H. H. Alhelou, and P. Siano, "An algorithm for recognition of fault conditions in the utility grid with renewable energy penetration," *Energies*, vol. 13, no. 9, p. 2383, May 2020.
[11] O. P. Mahela, J. Sharma, B. Kumar, B. Khan, and H. H. Alhelou, "An algorithm for the protection of distribution feeder using Stockwell and Hilbert transforms supported features," *CSEE J. Power Energy Syst.*, early access, Jul. 6, 2020, doi: 10.17775/CSEEJPES.2020.00170.
[12] A. van Deursen, P. Wouters, and F. Steennis, "Corrosion in low-voltage distribution networks and perspectives for online condition monitoring," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1423–1431, Aug. 2019.
[13] M. Jannati, B. Vahidi, and S. H. Hosseinian, "Incipient faults monitoring in underground medium voltage cables of distribution systems based on a two-step strategy," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1647–1655, Aug. 2019.
[14] P. Liu, Y. Liu, B. Cai, X. Wu, K. Wang, X. Wei, and C. Xin, "A dynamic Bayesian network based methodology for fault diagnosis of subsea christmas tree," *Appl. Ocean Res.*, vol. 94, Jan. 2020, Art. no. 101990.

[15] A. A. Razi-Kazemi and M. Shariatnasab, "A new approach on prioritization of the circuit breakers for installation of online monitoring systems," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1569–1577, Aug. 2019.

[16] J. Jiang, R. Chen, M. Chen, W. Wang, and C. Zhang, "Dynamic fault prediction of power transformers based on hidden Markov model of dissolved gases analysis," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1393–1400, Aug. 2019.

[17] B. Koch and Y. Carpentier, "Manhole explosions due to arcing faults on underground secondary distribution cables in ducts," *IEEE Trans. Power Del.*, vol. 7, no. 3, pp. 1425–1433, Jul. 1992.

[18] A. Hamel, A. Gaudreau, and M. Cote, "Intermittent arcing fault on underground low-voltage cables," *IEEE Trans. Power Del.*, vol. 19, no. 4, pp. 1862–1868, Oct. 2004.

[19] A. F. Bastos and S. Santoso, "Condition monitoring of circuit switchers for shunt capacitor banks through power quality data," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1499–1507, Aug. 2019.

[20] J. A. Wischkaemper, C. L. Benner, B. D. Russell, and K. Manivannan, "Application of waveform analytics for improved situational awareness of electric distribution feeders," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2041–2049, Jul. 2015.

[21] S.-K. Moon, J.-O. Kim, and C. Kim, "Multi-labeled recognition of distribution system conditions by a waveform feature learning model," *Energies*, vol. 12, no. 6, p. 1115, Mar. 2019.

[22] *Recommended Practice for Monitoring Electric Power Quality*, IEEE Standard 1159-2009, IEEE Standard Association, 2009.

[23] P. Ng, J. Antonellis, E. Balestrieri, J. Blair, J. Calvin, and D. Dallet, *IEEE Standard for Transitions, Pulses, and Related Waveforms*, IEEE Standard 181-2011 (Revision of IEEE Std. 181-2003), 2011.

[24] *Common Format for Transient Data Exchange (COMTRADE) for Power Systems*, IEEE Standard C37. 111-1999 (R2005) (Revision AV IEEE Standard C37. 111-1999), IEEE Standards Board, New York, NY, USA, 2012.

SANG KEUN MOON (Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from Hanyang University, Seoul, South Korea, in 2012 and 2019, respectively. He is currently a Senior Research Engineer with Korea Electric Power Corporation Research Institute (KEPCO), Daejeon, South Korea. His research interests include intelligent fault management and data processing in power distribution systems in the subject area of signal monitoring and network operation.

BYUNG-SUNG LEE received the Ph.D. degree in electrical engineering from Chungnam National University, Daejeon, South Korea, in 2005. He is currently a Principal Research Engineer with the Smart Distribution Laboratory, Korea Electric Power Corporation Research Institute (KEPCO), Daejeon. His research interests include facility reliability and evaluation, asset management, equipment diagnosis with smart sensors, and distribution system operation.