

Article

Candidate Fault Scenario Generation in Substations by Integrating Multi-Source Time-Series Data

Qifei Lv ¹, Yuhao Liu ², Jinhua Huang ², Huiying Lu ², Li Li ² and Xianbo Wang ^{1,3,*}¹ College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China; lvqf@zju.edu.cn² Electric Power Research Institute Guangdong Power Grid Co., Ltd., Guangzhou 510080, China; m18729204813@163.com (Y.L.); arichuang@126.com (J.H.); luhuiying_dky@163.com (H.L.); liligz@163.com (L.L.)³ Hainan Institute, Zhejiang University, Sanya 572025, China

* Correspondence: xbwang@zju.edu.cn

Abstract

Following the occurrence of a substation fault, alarm systems frequently exhibit issues such as high-density redundancy, spurious alerts, undetected alarms, and temporal inconsistencies. In addition, the effective integration of heterogeneous multi-source data—including fault waveform records and phasor measurement unit (PMU) streams—remains a nontrivial challenge, further impeding the accuracy of fault diagnosis. To address these limitations, this paper proposes an enhanced framework for multi-source fault event modeling and candidate scenario generation based on an extended temporal constraint network. First, protection operations, circuit breaker status transitions, transient waveform features, and PMU-derived dynamic parameters are abstracted as fault event nodes. On this basis, an event representation model is constructed, incorporating event type, associated equipment, temporal attributes, data source identifiers, and reliability coefficients. Second, the conventional temporal constraint network is refined by jointly embedding temporal constraints, distance constraints, and event confidence metrics, thereby enabling the characterization of uncertain causal-temporal dependencies between fault origins and multi-source observations. Third, a dynamic time-window mechanism is introduced to perform online event clustering, followed by reverse temporal reasoning to generate a set of candidate fault scenarios. These scenarios are subsequently verified, merged, and ranked with reference to waveform and PMU data, facilitating the identification of false alarms, missing alarms, and timestamp anomalies. Finally, case studies conducted on a 220 kV substation and its interconnected transmission network demonstrate that the proposed method achieves a correct scenario coverage rate of 96.8%, a candidate scenario compression rate of 68.7%, and an abnormal alarm recognition accuracy of 89.5%, while sustaining an average scenario generation time of 0.28 s.



Academic Editor: José Matas

Received: 26 May 2026

Revised: 26 June 2026

Accepted: 1 July 2026

Published: 8 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: substation fault scenarios; extended temporal constraint network; multi-source information fusion; alarm processing; Phasor Measurement Unit (PMU)

1. Introduction

With the extensive implementation of the IEC 61850 standard, a unified framework has been established for information modeling, communication interaction, and interoperability among intelligent electronic devices in power automation systems [1,2]. This facilitates the transformation of substation secondary systems from traditional point-to-point data acquisition methods to high-speed, large-scale, and standardized information transmission

modes [3,4]. In the event of a substation failure, dispatch centers or remote monitoring systems can generally acquire multi-source data from remote substations, including protection actions of internal equipment, circuit breaker positions, fault waveforms, and synchronous phasors. The COMTRADE standard offers a universal format for the exchange of transient data associated with power system faults, tests, and simulations, enabling the consistent utilization of fault waveform data for fault process analysis [5,6]. The synchronous phasor measurement standard stipulates the performance requirements for Phasor Measurement Units (PMUs), allowing dynamic parameters such as voltage, current phasors, frequency, and the rate of change of frequency (ROCOF) to be utilized in power system operation monitoring and dynamic analysis [7,8]. These datasets comprehensively depict substation fault evolution processes from diverse perspectives, including discrete operational events, localized transient waveforms, and wide-area dynamic responses, providing a crucial data foundation for substation fault classification.

Upon the occurrence of a substation fault, protection devices, circuit breakers, and automatic devices generate a substantial volume of alarms within a short period, which are then transmitted to the dispatch center. Nevertheless, these fault timing records are frequently incomplete and exhibit excessive redundancy, thereby significantly augmenting the complexity of manual fault diagnosis [9,10]. Factors such as communication congestion, device malfunctions, failures in the data acquisition link, and time synchronization deviations can result in false alarms, missing alarms, or temporal discrepancies in the alarm information. Although fault waveform recordings capture the transient changes in electrical parameters during the fault occurrence and clearance processes, their data structure, sampling resolution, and acquisition mechanisms differ significantly from discrete alarm data [11,12]. PMUs offer wide-area synchronous dynamic information; however, they mainly focus on observing system dynamics and necessitate temporal correlation and semantic alignment with discrete events such as protection actions and circuit breaker positions [13]. Under conditions of information distortion and multi-source heterogeneity, reconstructing potential fault evolution processes from chaotic observational data remains of critical importance for enhancing the reliability of subsequent fault diagnosis.

Extensive research has been carried out both domestically and internationally on power system fault diagnosis [14,15]. In the early stage, the methodologies mainly relied on expert systems, rule-based reasoning, and protection action logic, using predefined rules to depict the causal relationships among faulty equipment, protective devices, and circuit breakers [16,17]. Although these approaches possess clear physical significance and strong interpretability, they are still susceptible to the quality of original alarms when confronted with concurrent problems such as redundancy, omissions, and false alarms [18]. To enhance the utilization of alarm temporal information, online alarm processing models based on temporal constraint networks have been utilized to characterize the temporal constraints between alarms, offering a structured approach for fault process analysis [19]. Other studies have adopted temporal traceability reasoning methods to model grid fault diagnosis and alarm interpretation, thereby enhancing the fault analysis capabilities in complex alarm environments [20,21]. These studies indicate that temporal information plays a crucial role in differentiating genuine fault sequences from false alarms.

Apart from temporal analysis models, model-driven approaches, including Petri nets, fuzzy Petri nets, and Bayesian networks, are extensively utilized in power system fault diagnosis [22,23]. The time-series-based fuzzy Petri net methods can effectively manage uncertain information during the operations of protection and circuit breakers [24]. Fuzzy Petri nets incorporating temporal constraints further integrate event timing constraints into the diagnostic reasoning process, thereby enhancing fault tolerance in complex failure situations [25,26]. Recent advancements, such as fuzzy temporal Petri nets, aim to

directly integrate alarm temporal relationships into hierarchical Petri net architectures, improving the efficiency and adaptability of fault diagnosis under conditions of information distortion [27,28]. Nevertheless, these methods generally employ predefined alarm sets or switch states as inputs, mainly concentrating on fault component identification and protection action evaluation, while paying inadequate attention to multi-source event organization, temporal consistency verification, and candidate fault scenario generation at the diagnostic front-end.

With the utilization of dynamic information sources, such as PMUs and fault recording systems, diagnostic methods relying solely on alarms and switch signals are no longer capable of fully exploiting the existing substation data resources. The IEEE C37.118.2 standard delineates the message formats and communication protocols for the real-time exchange of synchronous phasor measurement data among PMUs, phasor data concentrators, and other applications, laying the groundwork for integrating PMU data into online analysis systems [29,30]. Nevertheless, in practical applications, PMU data, fault recording data, and alarm data demonstrate substantial disparities in sampling frequency, temporal resolution, spatial coverage, and data representation formats [31,32]. In the absence of a unified event abstraction and temporal correlation mechanism, these multi-source datasets cannot be directly employed by diagnostic models. Consequently, it is imperative to establish a candidate fault scenario generation stage prior to the final fault diagnosis, converting the raw multi-source observations into interpretable sets of current events that serve as the foundation for fault identification models.

As summarized in Table 1, existing fault diagnosis studies have made substantial progress in alarm reasoning, temporal modeling, probabilistic inference, Petri-net-based diagnosis, PMU-based monitoring, and multi-source data fusion [33,34]. However, most existing methods still focus on final fault component identification or alarm interpretation. They generally lack a unified front-end mechanism for converting heterogeneous alarm, fault-recording, and PMU measurements into temporally constrained event nodes, and they seldom generate candidate fault evolution scenarios before final diagnosis. In contrast, the proposed method emphasizes candidate scenario generation under information distortion [35,36]. By integrating unified event modeling, extended temporal constraints, reverse temporal reasoning, and multi-source evidence validation, the proposed framework can not only infer possible fault causes but also identify false alarms, missing alarms, and timestamp anomalies.

Table 1. Taxonomy and comparison of existing fault diagnosis methods.

Method Category	Typical Methods	Main Focus	Limitation Compared with This Work
Rule-based and expert reasoning	Expert systems and protection logic [16–18]	Alarm-based fault inference	Sensitive to false or missing alarms
Alarm-sequence and temporal reasoning	Alarm timing and traceability reasoning [12,15,19–21,28,37]	Alarm sequence interpretation	Limited Digital Fault Recorder (DFR)/PMU fusion
Petri-net and probabilistic diagnosis	Petri nets, fuzzy Petri nets, and Bayesian models [22–27,32–35,38,39]	Uncertain action reasoning	Mainly final fault identification
PMU- and waveform-based analysis	PMU events and DFR features [5,6,11,29–31,36]	Physical disturbance verification	Weak alarm-chain reconstruction
General multi-source fusion	Alarm, PMU, and DFR fusion [11,31,39]	Robust evidence fusion	Lacks unified temporal events
Proposed method	Extended temporal constraint network	Candidate scenario generation	Unified events, reverse reasoning, and anomaly detection

To address the aforementioned challenges, this paper proposes a method for modeling multi-source fault events and generating candidate fault scenarios based on an extended temporal constraint network. The term “candidate fault scenarios” refers to one or more

fault evolution hypotheses capable of explaining current multi-source observation events. First, protection actions, circuit breaker positions, waveform-recording transient characteristics, and PMU dynamic characteristics are treated as fault event nodes to establish an event representation model encompassing event types, associated devices, temporal information, data source labels, and reliability coefficients. Second, by incorporating external constraints such as temporal point restrictions, distance constraints, and event confidence levels, the traditional temporal constraint network is enhanced to describe the uncertain causal temporal relationships between fault causes and multi-source observation events. Subsequently, a dynamic time window is employed for online event grouping, and candidate fault scenarios for substations are generated through reverse temporal reasoning. These candidate scenarios are then validated, merged, and prioritized using waveform recordings and PMU data to identify false alarms, missing alarms, and timestamp anomalies. Finally, the effectiveness of this approach is demonstrated through case studies involving a 220 kV substation and its associated transmission grid.

The contributions of this paper are as follows:

(1) This paper provides a unified event abstraction mechanism for heterogeneous substation fault information. Existing fault diagnosis studies mainly rely on alarm sequences or switching states, while fault recording data and PMU measurements differ from alarms in time scale, sampling frequency, and semantic representation. The proposed representation converts protection alarms, circuit breaker positions, fault-recording transient features, and PMU dynamic responses into temporally comparable multi-source fault events, thereby providing a common semantic and temporal basis for jointly using discrete operational events and continuous electrical measurements in fault scenario reconstruction.

(2) This paper extends the conventional temporal constraint network from single-source alarm reasoning to multi-source fault-event reasoning. Conventional temporal constraint networks mainly process alarm sequences and have limited ability to represent the reliability differences of multi-source evidence or the heterogeneous dynamic responses reflected in fault recording data and PMU measurements. By incorporating time-point constraints, time–distance constraints, data source identifiers, and event reliability coefficients, the proposed network describes both the causal temporal relationships among fault causes, protection actions, and circuit breaker operations and the confidence differences among alarms, fault recordings, and PMU measurements.

(3) This paper extends the diagnostic front-end from directly producing a single fault diagnosis result to generating and cleaning candidate fault scenarios. Through dynamic time-window grouping and reverse temporal reasoning, the method constructs multiple candidate fault evolution hypotheses that can explain the current observations. These hypotheses are then validated, merged, and ranked using fault recording data and PMU evidence, enabling the identification of false alarms, missing alarms, and timestamp anomalies before final fault diagnosis.

The novelty of this work does not lie in using temporal reasoning or dynamic time windows alone, but in coupling unified multi-source event modeling, extended temporal constraints, reverse temporal reasoning, and DFR/PMU evidence verification into a front-end candidate scenario generation framework. This integration enables candidate fault evolution hypotheses to be generated, validated, ranked, and cleaned before final fault diagnosis under false alarms, missing alarms, and timestamp anomalies.

The remainder of this paper is organized as follows: Section 2 introduces a unified modeling method for multi-source fault events; Section 3 constructs an extended temporal constraint network and presents temporal reasoning rules; Section 4 proposes a method for generating candidate fault scenarios based on dynamic time windows and backward temporal reasoning; Section 5 presents anomaly detection methods and result analysis;

Section 6 discusses practical implications and limitations; Section 7 summarizes the overall research findings.

2. Feature Representation of Multi-Source Data for Fault Events

Upon the occurrence of a substation fault, the dispatch center and station-side monitoring systems generally acquire multi-source information, including protection actions, circuit breaker positions, fault waveforms, and synchronous phasors [38,39]. This information reflects the fault evolution from aspects such as relay protection operation procedures, transient physical fault mechanisms, and system-wide dynamic responses. Nevertheless, there are substantial differences in data sampling frequency, data structure, time scale, spatial coverage, and reliability. Consequently, prior to constructing fault temporal constraint models, it is imperative to standardize state and switch action data from various sources and formats into unified representation information for fault events, which encompasses temporal attributes, equipment correlations, and physical semantics.

2.1. Feature Types of Multi-Source Fault Data

An accurate comprehension of the data characteristics from diverse information sources is a prerequisite for the utilization of multi-source information in substation fault diagnosis. This paper undertakes an analysis from three dimensions: alarm signals, fault recording data, and PMU data.

Alarm signals mirror the operational processes of protective devices and circuit breakers, featuring high real-time performance and well-defined engineering semantics. They function as the core input for fault timing modeling. Nevertheless, they are prone to communication congestion, device malfunctions, and acquisition channel failures, which result in problems such as false alarms, missing alarms, and time scale irregularities [40].

Fault recording data captures the transient variations in voltage, current, and switching quantities before and after a fault, facilitating the extraction of crucial features such as fault initiation, current surges, voltage sags, and fault clearance. This provides a physical foundation for screening potential fault scenarios in substations. However, data retrieval and analysis are accompanied by time delays, and the coverage is restricted by the installation locations of the devices.

PMU data provides a unified time scale and wide-area synchronous measurement capabilities, enabling the extraction of characteristics such as voltage sags, phase-angle abrupt changes, frequency exceedances, and dynamic recovery. This facilitates system-level verification of fault locations and accident impact ranges. Nevertheless, its point density and sampling frequency are still limited.

The three aforementioned data categories demonstrate significant complementarity in terms of temporal scale, spatial coverage, and physical meaning. Through unified event-based modeling, they offer a coherent data foundation for constructing extended temporal constraint networks, generating candidate fault scenarios, and identifying abnormal information. Figure 1 summarizes the characteristics and complementary relationships among the three primary information sources employed in substation fault scenario generation.

2.2. The Temporal Representation of Multi-Source Data

To accomplish the temporal modeling of substation alarms, waveform recordings, and PMU data, this paper abstracts the raw data from diverse sources into unified fault event nodes. A fault event is defined as an information unit that possesses distinct temporal attributes, equipment associations, and physical significance during fault occurrence, protection actions, fault clearance, and system dynamic responses. By transforming discrete alarms, transient waveform characteristics, and synchronous phasor characteristics into a

unified data structure, this method offers a standardized data source for constructing an extended temporal constraint network model.

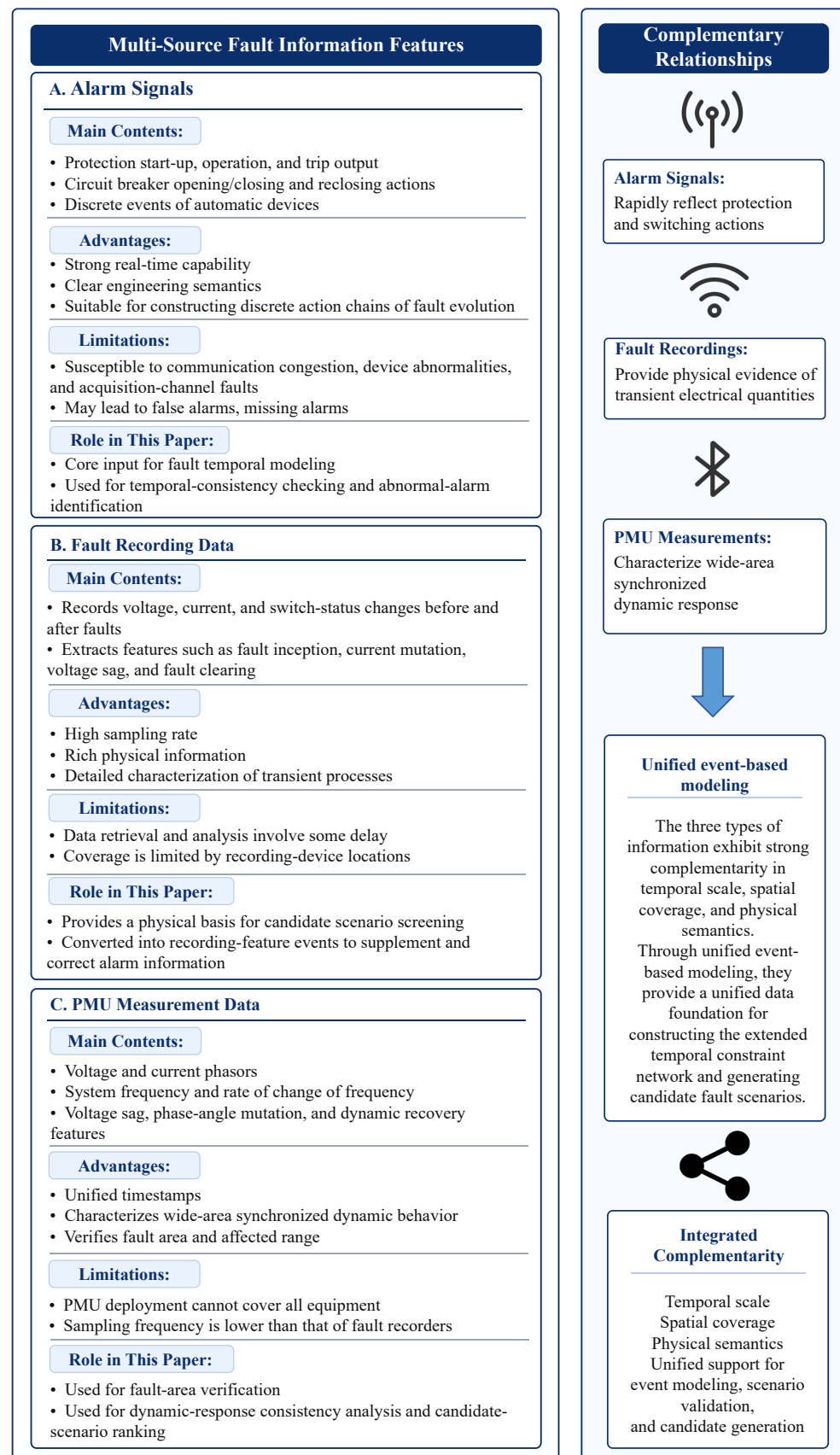


Figure 1. Multiple fault characteristics and their role in generating candidate fault scenarios.

Represent the i -th failure event as:

$$e_i = \{id_i, type_i, dev_i, t_i, a_i, s_i, q_i\} \quad (1)$$

where id_i denotes the event number; $type_i$ represents the event type; dev_i indicates the associated primary or secondary device; t_i specifies the event occurrence time or time interval; a_i is the event characteristic parameter; s_i is the data source tag; and q_i is the event reliability coefficient. The event type $type_i$ distinguishes between fault cause events, alarm events, waveform-recording feature events, and PMU dynamic events. The associated device dev_i describes the relationship between the event and lines, busbars, transformers, circuit breakers, protection devices, or measurement points. Time information t_i can represent either a precise moment or an uncertain time interval. When event timing accuracy is high, t_i may be expressed as a single time point; when timing errors or uncertainties exist, t_i should be represented as an interval. The characteristic parameter a_i records key numerical values of the event, such as voltage drop magnitude, current surge magnitude, frequency offset, and phase-angle variation. The data source tag s_i identifies whether the event originates from an alarm, fault waveform recording, or PMU measurement. The reliability coefficient q_i measures the event's credibility, with a value range of $[0, 1]$.

The reliability coefficient q_i is not assigned arbitrarily, but is determined according to source reliability, timestamp reliability, data quality, and semantic consistency:

$$q_i = \omega_s q_i^s + \omega_t q_i^t + \omega_d q_i^d + \omega_c q_i^c, \quad \omega_s + \omega_t + \omega_d + \omega_c = 1 \quad (2)$$

where q_i^s denotes the source reliability, q_i^t denotes the timestamp reliability, q_i^d denotes the data quality reliability, and q_i^c denotes the semantic consistency with topology and protection logic. All sub-scores are within $[0, 1]$. For example, the timestamp reliability can be expressed as:

$$q_i^t = \exp\left(-\frac{|\Delta t_i|}{\tau_i}\right) \quad (3)$$

where Δt_i denotes the timestamp deviation and τ_i denotes the allowable time-deviation scale of the corresponding event type. The source reliability can be assigned according to device status, communication quality, and historical reliability of SOE, DFR, or PMU data. The data quality term reflects missing values, noise, or abnormal measurements, while the semantic consistency term evaluates whether the event is consistent with topology, protection configuration, and adjacent events.

Multi-source data can be summarized as a set of events:

$$E = \{e_1, e_2, \dots, e_n\} \quad (4)$$

where n represents the number of events within the current time window. This event set may include both actually received observation events and expected events inferred based on topological relationships, protection configurations, and temporal constraints.

2.3. Classification of Fault Event Types in Substations

As shown in Table 2, based on the event origin and physical implications, this paper categorizes fault events into four types: cause events, alarm events, waveform-recording characteristic events, and PMU dynamic events.

A causal event pertains to the fundamental cause that has the potential to initiate protective actions, circuit breaker tripping, or electrical disturbances. It typically corresponds to malfunctions in substation primary equipment or irregularities in critical devices. Common causal events encompass line faults, busbar faults, main transformer

malfunctions, circuit breaker non-tripping, and protection device malfunctions. Generally, causal events cannot be directly observed by the dispatching system. Instead, they are deduced through reverse reasoning based on observational data such as alarms, waveform recordings, and PMU measurements.

Table 2. Classification of Fault Event Types.

Event Type	Typical Events	Data Source	Functions & Usage
Causal Events	Line Faults, Busbar Faults, Transformer Faults	Topology & Protection Logic Inference	Starting point for candidate fault scenarios
Alarm Events	Protection Actuation, Protection Trip, Breaker Status Change	SCADA/Fault Information System	Construct discrete action chains
Recorded Characteristic Events	Current Transients, Voltage Sags, Fault Clearance Phase-Angle Jumps, Frequency Excursions, Dynamic Recovery Inflection Points	DFR	Provide transient physical verification
PMU Dynamic Events		Wide-Area Measurement System (WAMS)/PMU	Verify fault location and dynamic response

An alarm event is discrete action information uploaded by protective devices, circuit breakers, or automatic devices. Typical alarm events include protection activation, protection tripping, protection discharge, circuit breaker opening, reclosing actions, and failure protection trips [41,42]. Alarm events carry significant practical physical implications and directly reflect the operational status of the substation's secondary systems and switching equipment.

Recorded characteristic events are transient data extracted from fault recording data, which possess distinct physical significance. Typical recorded characteristic events include fault initiation events, current mutation events, voltage drop events, zero-sequence component anomalies, negative-sequence component anomalies, and fault clearance events. These events reflect the physical essence of the fault process itself and serve as crucial evidence for verifying the authenticity of alarm events and the rationality of candidate scenarios.

PMU dynamic events represent the system's dynamic response characteristics extracted from synchronized phasor data. Typical PMU dynamic events include voltage amplitude drop events, voltage phase-angle abrupt change events, branch current phasor abrupt change events, frequency limit exceedance events, abnormal frequency rate-of-change events, and dynamic recovery inflection point events. These events comprehensively reflect the impacts of faults on the system's operational status from a wide-area perspective.

3. Construction of an Extended Temporal Constraint Network Model

In the fault evolution process of substations, there are temporal correlations among fault occurrence, protection actions, circuit breaker tripping, abrupt changes in electrical parameters, and system dynamic responses. Temporal constraint networks are frequently employed to depict the uncertain temporal relationships between events. Nevertheless, traditional temporal constraint networks can only deal with alarm events and are unable to process cause events, waveform-recording characteristic events, or PMU dynamic events. To tackle this issue, this paper presents data source identifiers and event reliability coefficients to construct an extended temporal constraint network tailored for multi-source fault events.

3.1. Time-Point Constraints and Time-Distance Constraints

Assuming the occurrence time point of event e_i is T_i , when there is uncertainty about the event's time point, its temporal constraint is represented by a time interval:

$$T_i = [t_i^-, t_i^+] \quad (5)$$

where, t_i^- and t_i^+ represent the lower and upper bounds of the event e_i 's occurrence time, respectively. When $t_i^- = t_i^+$, the event time becomes a deterministic instant. Assuming a causal relationship exists between events e_i and e_j , their temporal distance constraint is expressed as:

$$D_{ij} = [d_{ij}^-, d_{ij}^+] \quad (6)$$

where d_{ij}^- and d_{ij}^+ represent the lower and upper bounds of the time interval between event e_i and event e_j , respectively. If the temporal constraints of the preceding event e_i and the time interval constraints between events are known, the temporal constraint for the subsequent event e_j can be derived through forward reasoning:

$$T_j = T_i \oplus D_{ij} = [t_i^- + d_{ij}^-, t_i^+ + d_{ij}^+] \quad (7)$$

If the temporal constraints of the subsequent event e_j and the time interval constraints between events are known, the temporal constraint of the preceding event e_i can be derived through backward reasoning:

$$T_i = T_j \ominus D_{ij} = [t_j^- - d_{ij}^+, t_j^+ - d_{ij}^-] \quad (8)$$

For causal paths composed of multiple events, the time distance can be obtained by interval superposition. If an event chain $e_i \rightarrow e_j \rightarrow e_k$ exists, then:

$$D_{ik} = D_{ij} \oplus D_{jk} = [d_{ij}^- + d_{jk}^-, d_{ij}^+ + d_{jk}^+] \quad (9)$$

Here, $\oplus$ denotes interval addition for forward temporal propagation, $\ominus$ denotes interval subtraction for reverse temporal reasoning, and interval intersection is used to retain only mutually consistent candidate time intervals. These temporal constraints and distance constraints provide the computational foundation for establishing temporal correlations among multi-source events, deducing candidate causes, and identifying timestamp anomalies.

3.2. Extended Temporal Constraint Network

This paper defines the extended temporal constraint network as a seven-element set:

$$G = (V, A, T, C_1, C_2, S, Q) \quad (10)$$

where $V = (v_1, v_2, \dots, v_n)$ represents the set of event nodes, with each node corresponding to a fault event; $A \subseteq V \times V$ denotes the directed edge set, representing the causal triggering relationships between events; $T = (t_1, t_2, \dots, t_n)$ is the set of event time constraints; C_1 denotes the set of single-time constraints; C_2 denotes the set of binary time-distance constraints; $S = (s_1, s_2, \dots, s_n)$ is the set of event data source identifiers; and $Q = (q_1, q_2, \dots, q_n)$ is the set of event reliability coefficients. Compared with traditional temporal constraint networks, Equation (10) introduces the data source identifier S and the event reliability coefficient Q .

Compared with traditional temporal constraint networks, the proposed extended temporal constraint network retains the basic elements of event nodes, directed causal edges, time-point constraints, and time-distance constraints. The main extension lies in the introduction of multi-source event nodes, data source identifiers, and event reliability coefficients. Therefore, alarm events, fault-recording characteristic events, and PMU dynamic events can be represented within the same temporal reasoning framework. Traditional temporal constraint networks mainly describe temporal relationships among alarm sequences, whereas the proposed network further supports multi-source evidence validation and candidate scenario generation under false alarms, missing alarms, and timestamp anomalies.

The data source identifier distinguishes alarms, waveform recordings, and PMU events; the event reliability coefficient quantifies the credibility of different observation events. The extended temporal constraint network not only captures temporal dependencies between events but also integrates quality variations across multi-source data, establishing a temporal modeling framework for generating candidate scenarios under information distortion conditions.

3.3. Event Causal Relationship Modeling

The causal relationships between substation fault events are primarily determined by the power grid topology, protection configurations, and equipment operation logic. For instance, a line fault may trigger main protection actions on both sides of the line; these protection actions may cause circuit breakers to trip; a failed breaker operation may activate failure protection mechanisms; while the fault occurrence and clearance process induces abrupt changes in recorded electrical parameters and dynamic responses from PMUs. Let $\mathbb{R}$ denote the set of event causal rules, with the r -th rule expressed as:

$$r : e_i \xrightarrow{D_{ij}, q_r} e_j \quad (11)$$

where e_i denotes the preceding event, e_j the subsequent event, D_{ij} represents the temporal distance constraint between them, and q_r indicates the credibility of the causal rule. Based on protection setpoints, circuit breaker operating times, and measurement data response characteristics, temporal distance constraints between typical events can be established, as shown in Table 3.

Table 3. Example of time–distance constraints for typical fault events.

Predecessor Event	Successor Event	Time Interval Constraint (ms)	Description
Line Fault	Primary Protection Actuation	[15, 25]	Fast operation of primary protection
Primary Protection Actuation	Protection Trip	[0, 10]	From protection actuation to trip output
Protection Trip	Breaker Opening	[40, 60]	Breaker mechanical opening time
Breaker Failure	Breaker Failure Protection Actuation	[190, 210]	Breaker failure protection time delay
Fault Occurrence	Current Transient	[0, 5]	Transient waveform characteristics
Fault Occurrence	PMU Voltage Sag	[0, 40]	PMU dynamic response

The time–distance constraints in this table are not arbitrary empirical numbers. They are determined according to relay protection operating characteristics, circuit breaker mechanical operation time, breaker failure protection delay settings, and the time resolution of DFR and PMU event extraction. For example, the main protection action interval corresponds to the fast operating time of pilot differential protection, the breaker opening interval reflects the mechanical tripping time of the circuit breaker, and the breaker failure protection interval is consistent with the configured failure protection delay. The DFR and PMU response intervals are determined according to their sampling and event extraction characteristics. In practical applications, these intervals should be updated according to the actual protection settings, device manuals, and historical SOE/DFR/PMU records of the target substation.

3.4. Construction of Multi-Scale Time Windows

Given the significant temporal variations among fault events, this paper employs a multi-scale dynamic time window to perform real-time grouping of continuously arriving event streams. The current time window is defined as follows:

$$W_k = [\tau_k^-, \tau_k^+] \quad (12)$$

When a new event e_i arrives, its characteristic time window can be represented as:

$$W(e_i) = [t_i - \Delta_i^-, t_i + \Delta_i^+] \quad (13)$$

where t_i denotes the event time stamp, while Δ_i^- and Δ_i^+ represent the temporal margins before and after the event, respectively. If W_i overlaps with the current time window W_k , the current time window is extended accordingly:

$$W_k \leftarrow [\min(\tau_k^-, t_i - \Delta_i^-), \max(\tau_k^+, t_i + \Delta_i^+)] \quad (14)$$

If the new event does not overlap with the current time window and the current window meets the closing conditions, close the current window and start a new analysis window. The window closing conditions can be expressed as:

$$t_{now} - t_{last} > \Delta_{silent} \quad (15)$$

where t_{now} is the current time point, t_{last} is the arrival time point of the last correlated event in the current window, and Δ_{silent} is the waiting duration.

Based on the actual evolution process of substation faults, this paper divides the time window into three categories: millisecond-level operation window, second-level dynamic window, and fault evolution window. The millisecond-level operation window is used to handle protection operations, circuit breaker status changes, and transient recording events. The second-level dynamic window is used to process PMU dynamic response events. The fault evolution window is used to deal with delayed events such as backup protection, failure protection, and reclosing. The introduction of these multi-scale windows helps avoid incorrect splitting of related events within the same fault process and reduces misassociation between different fault sequences.

4. Method for Generating Candidate Fault Scenarios and Evaluating Their Matching

The objective of generating candidate fault scenarios is to construct hypotheses explaining the evolution of current multi-source observational data, given a set of observed events, topological relationships, and temporal constraints. This paper employs a workflow comprising event grouping, backward reasoning, multi-source verification, merging and sorting, and anomaly identification to generate candidate fault scenarios. The process for generating candidate fault scenarios is shown in Figure 2.

4.1. Candidate Fault Scenario Definition

The candidate fault scenarios referred to in this paper can be defined as:

$$H_m = \{c_m, T_m, E_m, \Gamma_m, S_m\} \quad (16)$$

where c_m denotes the candidate causal event; T_m represents the time interval for occurrence of the candidate causal event; E_m is the set of observational events most strongly supporting this candidate scenario; Γ_m indicates the corresponding set of causal paths for the scenario; and S_m signifies the comprehensive support degree of the candidate scenario.

Each candidate scenario encompasses at least the fault cause, the time interval of occurrence, a set of interpretable events, and their causal paths. It is noteworthy that candidate fault scenarios are not equivalent to final fault diagnosis conclusions; instead, they serve as

potential interpretations of the current observational data. The fault event diagnosis model facilitates reliability assessment, protection action evaluation, and scheduling support decision-making based on the collection of candidate scenarios.

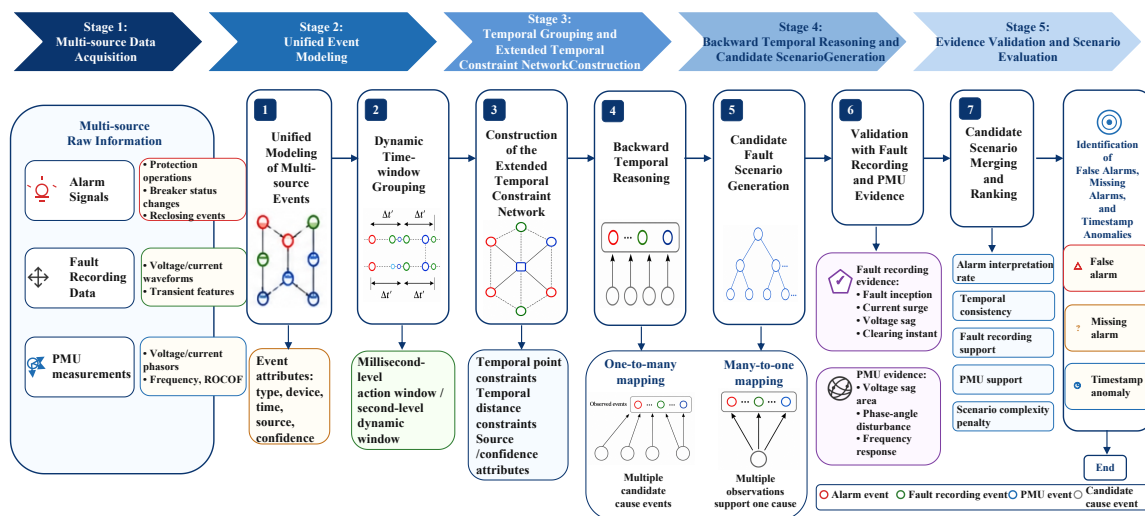


Figure 2. Flowchart for generating candidate fault scenarios.

4.2. Candidate Cause Generation Based on Reverse Temporal Reasoning

For each observation event e_j within the current time window, perform a backward search along the directed edges in the extended temporal constraint network to identify potential causal events that may have led to this observation. If a path exists:

$$p: c_m \rightarrow e_{p1} \rightarrow e_{p2} \rightarrow \dots \rightarrow e_j \quad (17)$$

Observing event e_j allows reverse reasoning to derive the temporal constraint of cause event c_m . Let the cumulative time distance from c_m to e_j along path p be defined as:

$$D_{mj}^p = \oplus_{(u,v) \in p} D_{uv} \quad (18)$$

Therefore, the candidate time intervals for event c_m can be expressed as:

$$T_m^p = T_j \ominus D_{mj}^p \quad (19)$$

For the same cause event c_m , multiple candidate time intervals can be derived through backward reasoning from various observational events, which can be represented as $\{T_m^1, T_m^2, \dots, T_m^r\}$.

If these intervals have a non-empty intersection, merge them to obtain the final time constraint for the cause event:

$$T_m = \bigcap_{k=1}^r T_m^k \quad (20)$$

If the interval intersection is empty, it indicates that this cause event cannot simultaneously explain multiple observed events; its priority should be reduced or it should be eliminated.

4.3. Validation of Candidate Fault Scenarios Based on Waveform-Recording Data and PMU Data

The support degree S_m^w of the recording data for candidate scenario H_m can be defined as:

$$S_m^w = \alpha_1 \eta_{time} + \alpha_2 \eta_{type} + \alpha_3 \eta_{clear} \quad (21)$$

where η_{time} represents the consistency at the fault initiation moment, η_{type} denotes the consistency of fault type characteristics, and η_{clear} indicates the consistency at the fault clearance moment; α_1 , α_2 , and α_3 are weight coefficients satisfying $\alpha_1 + \alpha_2 + \alpha_3 = 1$.

The weighting coefficients in Equation (21) reflect the relative diagnostic importance of different types of fault recording evidence in candidate scenario validation. The fault inception time consistency η_{time} is assigned a relatively high priority because it determines whether the candidate cause can temporally explain the observed transient disturbance. The fault type characteristic consistency η_{type} is also important because it verifies whether the extracted electrical features, such as current mutation and voltage sag, are consistent with the assumed fault type. The fault clearing time consistency η_{clear} is used as supplementary evidence to check whether the protection trip and circuit breaker operation agree with the physical clearing process. Since no fixed universal values are imposed in this study, α_1 , α_2 , and α_3 are determined according to engineering knowledge of fault-recording interpretation and adjusted using historical or simulated fault cases, while satisfying $\alpha_1 + \alpha_2 + \alpha_3 = 1$.

The support degree S_m^p of PMU data for candidate scenario H_m can be defined as:

$$S_m^p = \beta_1 \eta_v + \beta_2 \eta_\theta + \beta_3 \eta_f \quad (22)$$

where η_v represents the consistency of voltage drop range, η_θ denotes the consistency of phase-angle disturbance, and η_f indicates the consistency of frequency response; β_1 , β_2 , and β_3 are weight coefficients satisfying $\beta_1 + \beta_2 + \beta_3 = 1$. When partial missing values occur in waveform-recording data or PMU data, the corresponding evidence support values S_m^w and S_m^p are excluded from calculation, preventing erroneous exclusion of candidate scenarios due to unavailable single data sources.

The weighting coefficients in Equation (22) distinguish the relative contributions of different PMU dynamic features. The voltage sag range consistency η_v is emphasized because it is directly related to fault area verification. The phase-angle disturbance consistency η_θ is treated as important supporting evidence because it reflects disturbance propagation and power-flow redistribution after the fault. The frequency response consistency η_f is regarded as system-level auxiliary evidence, because frequency dynamics are generally less location-specific than voltage and phase-angle features in short-term substation fault scenario verification. Accordingly, β_1 , β_2 , and β_3 are selected according to the diagnostic relevance of the corresponding PMU features, with $\beta_1 + \beta_2 + \beta_3 = 1$.

4.4. Merging and Sorting Candidate Fault Scenarios

Candidate fault scenarios derived through backward inference from different observation events may describe the same fault process. If two candidate fault scenarios H_a and H_b satisfy the following conditions, they are merged into a new candidate fault scenario H_{ab} :

$$dev(H_a) = dev(H_b), T_a \cap T_b \neq \emptyset \quad (23)$$

The new candidate failure scenario H_{ab} encompasses all observed events explainable by both failure scenarios H_a and H_b , with their temporal constraints and comprehensive support values recalculated. The comprehensive support value for candidate scenarios defined in this paper is as follows:

$$S_m = \lambda_1 S_m^a + \lambda_2 S_m^t + \lambda_3 S_m^w + \lambda_4 S_m^p - \lambda_5 C_m \quad (24)$$

where S_m^a denotes the alarm interpretation rate; S_m^t represents the consistency of the time series; S_m^w indicates the evidence support level of waveform recording; S_m^p signifies the

PMU evidence support level; C_m is the scenario complexity penalty term; and λ_1 to λ_5 are the weight coefficients.

Here, c_m denotes the candidate causal event, whereas C_m in Equation (24) denotes the scenario complexity penalty term.

The alarm interpretation rate indicates the proportion of alarms that a candidate scenario can explain relative to the total number of alarms in the current window; temporal consistency reflects the degree of alignment between actual event timestamps and the expected time interval; the scenario complexity penalty term serves to suppress complex hypotheses containing excessive cause events or excessively long causal chains. Candidate scenarios are ranked in descending order of comprehensive support, with the top K scenarios selected as subsequent diagnostic inputs.

4.5. Evaluation of the Match Between Candidate Fault Scenarios and Observed Events

Based on the matching relationship between candidate fault scenarios and observed events, this paper employs three sets of metrics: false alarm, missing alarm, and timestamp anomaly detection.

(1) False alarm: If an alarm event e_j cannot be explained by any candidate scenario, i.e.,: $e_j \notin \bigcup_{m=1}^K E_m$. If no corresponding physical response is detected in the waveform recording or PMU data, the alarm is classified as a suspected false alarm.

(2) Missing alarm inference: If event e_j^{exp} exists in the expected event set of candidate scenario H_m , but the corresponding event does not appear in the actual observed event set, while the recorded waveform data or PMU data indicate that the relevant equipment status or electrical quantities have changed, then it is inferred that this event is missing.

(3) Timestamp anomaly detection: If an alarm event can be explained by a candidate scenario, but its actual timestamp t_j falls outside the expected time interval t_j^-, t_j^+ , then the timestamp deviation is defined as:

$$\Delta t_j = \begin{cases} t_j^- - t_j, & t_j < t_j^- \\ 0, & t_j^- \leq t_j \leq t_j^+ \\ t_j - t_j^+, & t_j^+ < t_j \end{cases} \quad (25)$$

When Δt_j exceeds the specified threshold Δt_{thr} , it is determined that an anomaly in the alarm's timing exists.

4.6. Handling of Delayed, Missing, and Out-of-Order Data

The dynamic time-window mechanism provides limited tolerance for delayed events, because related events can still be accepted before the current window is closed by the silent waiting duration Δ_{silent} . Out-of-order data are processed according to event timestamps and temporal intervals rather than their arrival order. When fault recording data or PMU data are missing, the corresponding support term is excluded or down-weighted instead of directly eliminating the candidate fault scenario. However, if multiple sources are simultaneously missing or severely delayed, the method may generate more candidate scenarios, reduce ranking confidence, and degrade abnormal information identification accuracy.

4.7. Computational Complexity and Scalability Analysis

The computational complexity of the proposed method is mainly determined by dynamic time-window grouping, reverse temporal reasoning, multi-source evidence validation, candidate scenario merging, and scenario ranking. Let n denote the number of observed events in the current dynamic time window, b denote the average number of predecessor branches searched during reverse temporal reasoning, L denote the maximum

backward search depth, and M denote the number of initially generated candidate fault scenarios. In the worst case, the complexity of reverse temporal reasoning is approximately $O(nb^L)$, because each observed event may trace multiple causal paths in the extended temporal constraint network. Candidate scenario merging has a straightforward pairwise complexity of $O(M^2)$, while scenario ranking requires $O(M \log M)$. Multi-source evidence validation using fault recording data and PMU data is mainly related to candidate-event matching and is approximately linear with respect to the number of candidate scenarios and associated evidence events.

In practical applications, the actual computational burden is lower than the theoretical worst case. First, the dynamic time-window mechanism limits the number of events processed in each analysis window and prevents unrelated fault sequences from being searched together. Second, the power grid topology and protection-zone constraints restrict the search range, while temporal distance constraints remove infeasible causal paths at an early stage. Third, interval intersection during reverse temporal reasoning eliminates candidate causes that cannot simultaneously explain multiple observed events. Finally, candidate scenario merging reduces redundant hypotheses before final ranking, thereby decreasing the number of scenarios passed to subsequent diagnostic modules.

From an engineering scalability perspective, the proposed method can be extended to larger substations or regional grids because event grouping is performed within local time windows, reverse reasoning is constrained by topology and protection configurations, and fault-recording/PMU evidence is used only to validate relevant candidate scenarios rather than searching the whole network. In addition, different candidate causes or candidate scenarios can be evaluated independently, which makes the method suitable for parallel implementation in online fault analysis systems.

5. Simulation Analysis and Case Studies

The case study is constructed based on the topology and protection configuration of a 220 kV substation, while the fault scenarios and information distortion cases are simulated to evaluate the proposed method. The simulated data include SOE alarms, DFR transient features, and PMU dynamic events.

The selected 220 kV substation and its corresponding transmission grid are employed as the test system, as shown in Figure 3. The substation employs a double-busbar configuration, which comprises three 220 kV transmission lines, two main transformers, and twelve sets of circuit breakers. The lines are equipped with longitudinal differential main protection and distance protection serving as near-backup protection, while zero-sequence over-current protection acts as far-backup protection; each circuit breaker is also equipped with a failure protection device. The busbars are protected by busbar differential protection.

The case study integrates three types of information sources:

(1) Alarm timing data, which encompasses Sequence of Events (SOE) such as protection actions, protection outputs, circuit breaker positions, and failure protection actions. Under normal circumstances, the alarm timing error does not exceed ± 5 ms; in scenarios with information distortion, false alarms, missing alarms, and timestamp anomalies are simulated.

(2) Fault recording data, including voltage and current waveforms before and after the fault, from which characteristic events such as fault initiation, current surges, voltage drops, and fault clearance are extracted.

(3) PMU data, including bus voltage phasors near the fault area, line current phasors, system frequency, and frequency change rate, are utilized to extract PMU events such as voltage sags, phase-angle abrupt changes, and dynamic recovery.

The temporal constraints of various events are illustrated in Table 3. It is evident that after a line fault occurs, the surge in recording current and the PMU voltage drop must be detected within 5 ms and 40 ms, respectively. This underscores the critical need for ultra-high real-time performance in sensing and measuring the initial fault state within power systems. From fault occurrence to complete physical isolation, the primary protection path exhibits a distinct stepwise rapid-response characteristic: the primary protection operates within 15–25 ms, with protection discharge completed within 10 ms. In contrast, the circuit breaker's tripping action experiences a delay of 40–60 ms, as circuit breakers are mechanical devices whose tripping actions are constrained by mechanical inertia and cannot be executed instantaneously like electronic signals.

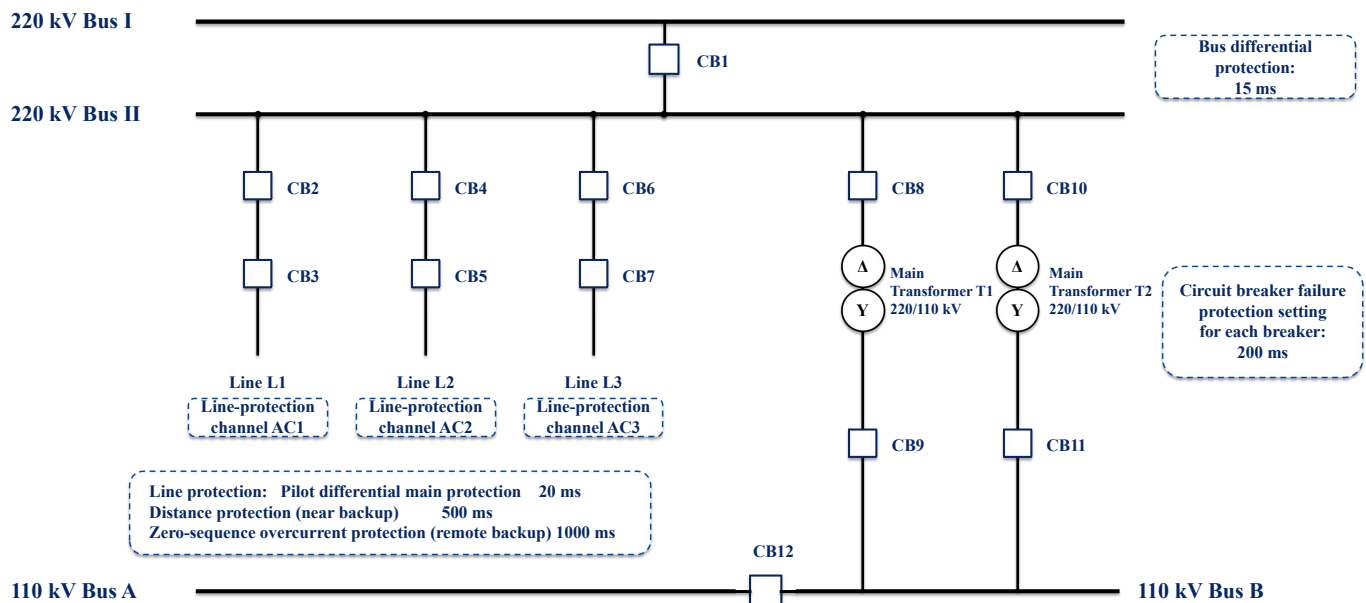


Figure 3. The main electrical wiring diagram of the example system.

The case-study time–distance constraints follow the same principle as Table 3 and are configured according to the protection settings, breaker operation characteristics, and DFR/PMU event extraction settings of the tested 220 kV substation system. The comprehensive support weights in Equation (24) can be set as: $\lambda_1 = \lambda_2 = 0.3$, $\lambda_3 = 0.2$, $\lambda_4 = 0.15$, and $\lambda_5 = 0.05$. The weights are selected according to the diagnostic contribution of each evidence component. The alarm interpretation rate S_m^a and temporal consistency S_m^t are assigned the largest weights because they directly determine whether a candidate fault scenario can explain the observed discrete event chain and satisfy the expected temporal constraints. The waveform-recording support S_m^w is assigned a medium weight because fault recording data provide direct physical evidence of fault inception, transient behavior, and clearing. The PMU support S_m^p is assigned a slightly lower but still significant weight because PMU data provide wide-area dynamic verification, while their spatial coverage in substations may be limited. The scenario complexity penalty term C_m is assigned the smallest weight to suppress unnecessarily complex hypotheses without eliminating plausible multi-event scenarios. In this study, the weights were first initialized according to engineering knowledge and the diagnostic roles of different evidence sources, and then checked or adjusted using historical fault records or simulated cases to obtain stable scenario ranking and anomaly identification performance for false alarms, missing alarms, and timestamp anomalies.

To evaluate the adaptability of the proposed method under various information distortion conditions, this paper presents test cases for four typical fault scenarios, as shown

in Table 4. Case 1 represents a single-phase ground fault on line L1, assessing the model's scenario generation capability under normal, complete data conditions. Case 2 involves a fault on line L1, evaluating the model's false-negative inference performance. Case 3 concerns a fault on busbar I, testing the model's false-positive identification capability. Case 4 describes a fault on line L3 accompanied by CB6 failure, verifying the model's anomaly detection performance.

Table 4. Settings of Four Test Cases.

Cases	Fault Settings	Information Distortion Situation	Verification Target
Case 1	L1 Line Single-Phase-to-Ground Fault	Intact Alarms	Scenario Generation Capability
Case 2	L1 Line Fault	Missing Alarm for Primary Protection Actuation (One Side)	Missing Data Inference Capability
Case 3	Bus I Fault	False Alarm from L2 Backup Protection	False Alarm Identification Capability
Case 4	L3 Line Fault with CB6 Failure	Timestamp Anomaly in CB6 Trip Signal	Anomaly Detection Capability

5.1. Evaluating Indicators

This paper introduces three metrics: correct scenario coverage R_c , candidate scenario compression rate R_s , and abnormal alarm recognition accuracy R_a to evaluate the effectiveness of candidate fault scenario generation.

(1) Correct scenario coverage rate R_c :

$$R_c = N_c / N \quad (26)$$

where N_c denotes the number of occurrences in the candidate set for the correct fault scenario, while N represents the total number of tests.

(2) Candidate scenario compression rate R_s :

$$R_s = 1 - N_{out} / N_{raw} \quad (27)$$

where N_{raw} represents the initial number of candidate scenarios, while N_{out} denotes the number of merged and sorted candidate scenarios.

(3) Average anomaly alarm recognition accuracy R_a :

$$R_a = N_{right} / N_{abn} \quad (28)$$

where N_{right} represents the number of correctly detected abnormal alarms, and N_{abn} denotes the actual number of abnormal alarms.

5.2. Experimental Analysis

5.2.1. Results of Generated Candidate Fault Scenarios

As depicted in Table 4, the generated candidate fault scenarios for the test cases of the aforementioned four typical fault scenarios are presented in Table 5. Evidently, when all alarms are present, the proposed method precisely generates candidate scenarios that are consistent with the actual fault progression. In instances of missing main protection action alarms, the method infers missing events by utilizing current transients in waveform recordings, voltage drops detected by PMUs, and temporal correlations between adjacent alarms, thus maintaining high confidence levels for correct fault scenarios. Regarding false alarms, which are challenging to characterize solely through bus fault causality chains or waveform/PMU data, the method identifies them as suspected false alarms. Finally, under timestamp anomalies in circuit breaker trip alarms, the method calculates timestamp

deviations based on temporal constraints between protection output signals and circuit breaker opening events, facilitating effective anomaly detection.

Table 5. Results of generated candidate fault scenarios.

Cases	Top Candidate Fault Scenario	Support Degree	Correct Ranking	Anomaly Identification Result
Case 1	L1 Line Fault, Correct Primary Protection Operation	0.94	1	No Anomaly
Case 2	L1 Line Fault, Missing One-Side Primary Protection Alarm	0.88	1	Identified Primary Protection Alarm Loss
Case 3	Bus I Fault, False Alarm from L2 Backup Protection	0.85	1	Identified False Alarm from L2 Backup Protection
Case 4	L3 Line Fault, Timestamp Anomaly in CB6 Alarms	0.82	1	Identified CB6 Timestamp Anomaly

5.2.2. Comparative Analysis of the Generated Results Using Different Comparison Methods

To further validate the advanced nature of the proposed method, this paper also introduces a candidate scenario generation approach based solely on alarm rules (Method I) in [37] and the temporal constraint network (Method II) in [24].

Method I is implemented as an alarm-rule-based candidate scenario generation method. Its inputs include protection alarms and circuit breaker status records, and its reasoning logic depends on predefined protection action rules and alarm matching. Therefore, it is sensitive to false alarms and missing alarms because no additional physical evidence from fault recording data or PMU data is used. Method II is implemented as a temporal constraint network method. Its inputs include alarm events and circuit breaker events, and its reasoning logic uses temporal constraints among protection actions and breaker operations to eliminate temporally inconsistent scenarios. However, Method II mainly focuses on the temporal consistency of alarm-event sequences and does not fully incorporate fault-recording characteristic events or PMU dynamic events.

As presented in Table 6, Method I exclusively depends on alarm rules and encounters difficulties in resolving the problems of missing alarms and false alarms. This leads to a lower coverage rate of correct scenarios and a lower accuracy in identifying abnormal information. By incorporating temporal constraints, Method II utilizes the temporal relationship between protective actions and circuit breaker positions to eliminate certain unreasonable candidate scenarios, thus enhancing its performance. The proposed method further integrates the unified modeling of multi-source events, waveform-recording evidence, and PMU dynamic data, in addition to traditional temporal constraints. It achieves a correct scenario coverage rate of 96.8%, an abnormal alarm recognition accuracy of 89.5%, and a candidate scenario compression rate of 68.7%. Although the average generation time of the proposed method is marginally longer than that of the previous two methods, it still remains within seconds, fulfilling the requirement for rapid post-fault scenario reconstruction. The slightly increased average generation time is consistent with the additional computational steps analyzed in Section 4, including multi-source evidence validation and candidate scenario merging; nevertheless, the average generation time remains 0.28 s in the tested 220 kV substation case, indicating that the proposed method can still support rapid post-fault scenario reconstruction.

Table 6. Performance comparison of candidate fault scenario generation under different methods.

Methods	R_c (%)	R_s (%)	R_a (%)	Average Generation Time (s)
Method I in [37]	83.5	42.1	61.8	0.16
Method II in [24]	91.2	55.4	76.3	0.21
Proposed method	96.8	68.7	89.5	0.28

Figure 4 visualizes the comparison results reported in Table 6.

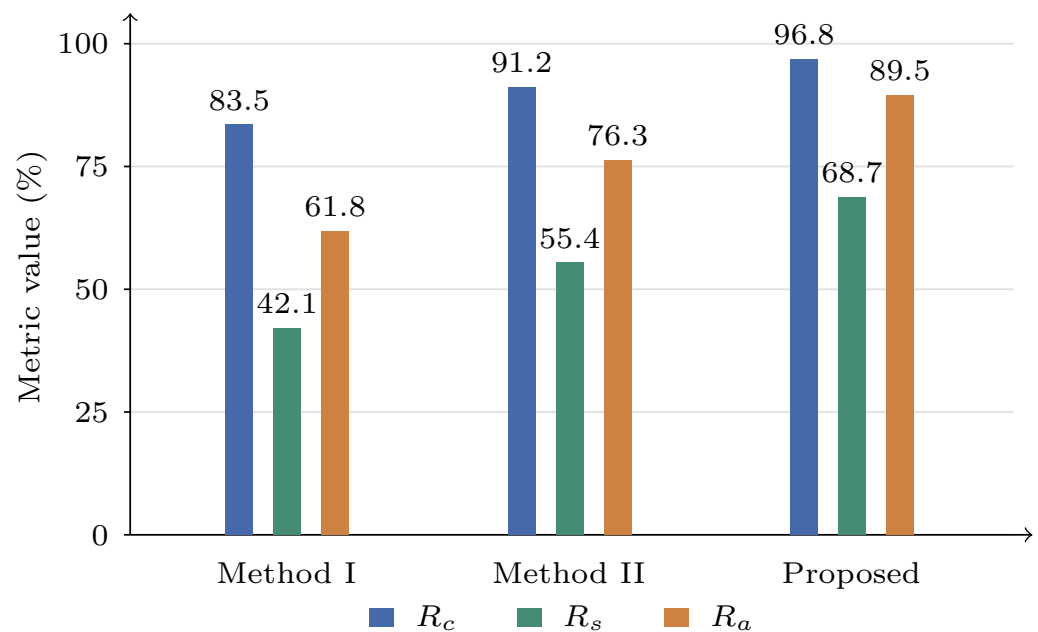


Figure 4. Performance comparison of candidate fault scenario generation methods based on the metrics reported in Table 6.

6. Discussion

Although the proposed method improves candidate scenario generation by integrating alarm events, DFR evidence, and PMU dynamic responses, its performance still depends on the quality of multi-source data and the correctness of predefined temporal constraints. From the computational perspective, the main cost comes from reverse temporal reasoning, multi-source evidence validation, and candidate scenario merging. As analyzed in the computational complexity subsection, the search space is reduced by dynamic time-window grouping, topology constraints, protection-zone constraints, and interval intersection. The experimental results in Table 6 further show that the average generation time remains 0.28 s in the tested 220 kV substation case, which satisfies rapid post-fault reconstruction requirements.

In practical applications, SOE, DFR, and PMU data may arrive through different communication channels and protocols, resulting in delayed, missing, or out-of-order events. The proposed method handles these issues by using event timestamps and temporal intervals rather than only relying on data arrival order. Missing DFR or PMU evidence does not directly eliminate a candidate scenario; instead, the corresponding evidence term is excluded or down-weighted during scenario evaluation. However, when multiple data sources are simultaneously missing or when synchronization errors exceed the predefined temporal tolerance, the number of candidate scenarios may increase and the ranking confidence may decrease.

The proposed method is not restricted to a single network topology because the temporal constraint network is constructed from topology, protection configuration, and event causal rules. It can be applied to different network structures, such as busbar, radial, or ring networks, if the corresponding causal rules, protection zones, and temporal constraints are available. However, different topologies and protection schemes require updating the rule library, protection action paths, time–distance constraints, and candidate causal event set. In addition, limited PMU/DFR coverage, incomplete protection configuration information, or severe synchronization errors may reduce scenario ranking confidence and abnormal information identification accuracy. Systematic tests under more topologies and protection configurations will be considered in future work.

7. Conclusions

To address the issues of false alarms and missing alarms caused by heterogeneous multi-source data and dense redundant alarms following substation faults, this paper proposes a method for generating candidate fault scenarios by integrating multi-source time-series data. First, recognizing the significant differences in sampling frequency, temporal scale, and physical semantics among protection alarms, fault waveforms, and PMU data, we developed a unified multi-source fault event representation model. This model abstracts discrete action signals, transient electrical characteristics, and wide-area dynamic responses into standardized event nodes containing event types, associated devices, time intervals, data source labels, and reliability coefficients, thereby eliminating integration barriers arising from data heterogeneity. Building on this foundation, we extended traditional temporal constraint networks by incorporating dimensions such as temporal point constraints, distance constraints, data source identifiers, and event confidence levels. This approach precisely captures the uncertain causal temporal relationships between fault causes, protection actions, circuit breaker positions, and electrical parameter anomalies. A multi-scale dynamic time-window mechanism enables accurate event stream grouping for millisecond-level actions, second-level dynamics, and accident evolution processes. Subsequently, a reverse temporal reasoning strategy is employed to trace potential fault origins backward from observed events, generating multiple candidate fault evolution hypotheses that explain current observations. These hypotheses are validated using waveform transient features and PMU dynamic response evidence, followed by similarity-based merging and priority ranking based on comprehensive support measures, effectively eliminating redundant hypotheses. The simulation case was constructed based on a 220 kV substation system. Compared with methods relying solely on alarm rules or traditional temporal constraint networks, the proposed method achieves a correct scenario coverage rate of 96.8%, a candidate scenario compression rate of 68.7%, an abnormal alarm recognition accuracy of 89.5%, and an average generation time of 0.28 s in the tested 220 kV substation case. By improving scenario coverage, compressing redundant candidate scenarios, enhancing abnormal information identification, and maintaining sub-second generation time, this method satisfies the requirement for rapid post-fault scenario reconstruction and provides highly reliable decision-making support for remote power system monitoring.

Author Contributions: Conceptualization, Q.L. and Y.L.; methodology, Q.L.; software, Q.L.; validation, Y.L. and J.H.; formal analysis, H.L.; investigation, L.L.; resources, L.L.; data curation, Q.L.; writing—original draft preparation, Q.L.; supervision, X.W.; project administration, Y.L. and X.W.; funding acquisition, X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the Science and Technology Project of China Southern Power Grid (Grant No. GDKJXM20240116 (030000KC24010016)), by the Hainan Provincial Sanya Yazhou Bay Science and Technology Innovation Joint Project (Grant No. ZDYF2025GXJS142), and by the Project of Sanya Yazhou Bay Science and Technology City (Grant No: SKJC-JYRC-2025-53).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: Author Yuhao Liu, Jinhua Huang, Huiying Lu and Li Li were employed by the company Electric Power Research Institute, China Southern Power Grid (China). The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Alolabi, Y.; Nweke, L.O. Risk–Observability Mismatch in an IEC 61850 Digital Substation: A Structured Cyber-Physical Assessment. *Appl. Sci.* **2026**, *16*, 5237. [\[CrossRef\]](#)
- Gaede, T.C.; Glende, E.; Wolter, H.M. Integrated Communication Protocols for Substation Automation: HiL Validation of IEC 61850, IEEE C37.118 and IEC 60870-5-104. In Proceedings of the 2025 57th North American Power Symposium, NAPS, Storrs, CT, USA, 26–28 October 2025; IEEE: New York, NY, USA, 2025; 6p. [\[CrossRef\]](#)
- Saadi, K.; Abbou, R. On IEC 61850 Communication Networks in Smart Grids System: Methodology of Implementation and Performances Analysis on an Experimental Platform. *Int. J. Energy Res.* **2022**, *46*, 89–103. [\[CrossRef\]](#)
- Guo, J.; Dou, C.; Yue, D.; Zhang, Z.; Zhang, Z.; Zhang, B. Cyber-Physical Integration Design for Frequency Regulation in Virtual Power Plants Facing Diverse Communication Issues. *IEEE Trans. Power Syst.* **2025**, *40*, 5237–5249. [\[CrossRef\]](#)
- Jin, W.; Xinran, L.; Sheng, S. Research on the Application of the Recorded Fault Data Analysis based on COMTRADE Standard in Electric Power Load Modeling. In Proceedings of the 2005 International Power Engineering Conference, Singapore, 29 November–2 December 2005; Volume 2, pp. 1058–1063. [\[CrossRef\]](#)
- Lin, Z.; Zhang, J.; Chen, Y.; Tian, Q.; Lin, Z.; Huang, G. A New Approach to Power System Fault Diagnosis based on Fuzzy Temporal Order Petri Nets. *Energy Rep.* **2022**, *8*, 969–978. [\[CrossRef\]](#)
- Wang, H.; Tang, Z.; Tan, D.; Zhou, F. A Short-Term Load Forecasting Method for Traction Substations Based on Physical Information Collaboration and Spatiotemporal Correlation. *Energies* **2026**, *19*, 2514. [\[CrossRef\]](#)
- Pandey, S.; Srivastava, A.K.; Dubey, A.; Rahmatian, F. A Novel Architecture and Estimation Algorithm for Adaptive Distribution Phasor Measurement Unit in Renewable-Rich Electrical Distribution System. *IEEE Trans. Ind. Appl.* **2025**, *61*, 2148–2157. [\[CrossRef\]](#)
- Lin, L.; Qiu, Y. Fault Detection of Relay Protection Secondary Circuit in the Digital Substation. In Proceedings of the 2023 3rd International Conference on New Energy and Power Engineering (ICNEPE), Huzhou, China, 24–26 November 2023; pp. 752–756. [\[CrossRef\]](#)
- Filho, H.V.H.S.; De Oliveira, J.M.A.M.; Barboza, A.G.; Silva, L.M.; De Melo, M.T.; Barboza, D.C.P.; Novo, L.R.G.S.L.; Dos Santos, R.G.M.; Tarrago, V.L. Bandpass Filter Using Open Loop Microstrip Resonators for a Power Substation Eyebolt Fault Detection System. In Proceedings of the 2023 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC), Castelldefels, Spain, 5–9 November 2023; pp. 178–180. [\[CrossRef\]](#)
- Patel, V.; Kapoor, A.; Sharma, A.; Chakrabarti, S. Taxonomy of Outlier Detection Methods for Power System Measurements. *Energy Convers. Econ.* **2023**, *4*, 73–88. [\[CrossRef\]](#)
- Li, J.; Zheng, C.; Wu, Z.; Chen, H.; Wang, X.B.; Yang, Z.X. Review of Fault Diagnosis for Rotating Machinery: Prior Knowledge Integration in Data-driven Methods Benefits Model Interpretability and Generalizability. *Mech. Syst. Signal Process.* **2026**, *242*, 113623. [\[CrossRef\]](#)
- Liu, S.; Zhao, Y.; Lin, Z.; Liu, Y.; Ding, Y.; Yang, L.; Yi, S. Data-Driven Event Detection of Power Systems Based on Unequal-Interval Reduction of PMU Data and Local Outlier Factor. *IEEE Trans. Smart Grid* **2020**, *11*, 1630–1643. [\[CrossRef\]](#)
- Mosayyebi, S.R.; Shahalami, S.H.; Mojallali, H. Fault Ride-Through Capability Improvement in a DFIG-Based Wind Turbine using Modified ADRC. *Prot. Control Mod. Power Syst.* **2022**, *7*, 1–37. [\[CrossRef\]](#)
- Xue, P.; Shi, L.; Zhou, Z.; Liu, J.; Chen, X. An Online Fault Detection and Diagnosis Method of Sensors in District Heating Substations based on Long Short-term Memory Network and Adaptive Threshold Selection Algorithm. *Energy Build.* **2024**, *308*, 114009. [\[CrossRef\]](#)
- Raja, V.K.; Ideris, A.I. Enhancement for Circuit Breaker Failure Protection in Transmission Substations. In Proceedings of the 2022 IEEE International Conference on Power Systems Technology (POWERCON), Kuala Lumpur, Malaysia, 12–14 September 2022; pp. 1–6. [\[CrossRef\]](#)
- Mircea, P.M.; Ciontu, M.; Mircea, I.; Buzatu, C.; Rusinaru, D.; Ursu, D.; Marin, I.; Butoarca, E. Analyzing the Efficiency of Neutral Treatment with Resistor and Shunt Circuit Breaker in a MV Distribution Substation. In Proceedings of the 2018 International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO), Prague, Czech Republic, 19–21 May 2018; pp. 223–228. [\[CrossRef\]](#)
- Xu, Z.; Ma, R.; Cheng, Y.; Xu, J.; Li, T.; Zhang, L. Modeling and Fault Mechanism Analysis for AC Intrusion into DC System in Power Substations. In Proceedings of the 2025 5th IEEE International Conference on Energy Engineering and Power Systems (EEPS), Shenzhen, China, 31 October–2 November 2025; pp. 380–384. [\[CrossRef\]](#)
- Xu, P.; Zhang, S.; Luo, P.; Wang, X.; Yan, Y.; Qi, D. Data-feature Hybrid Optimization for Ultra-short-term Wind Power Forecasting. *Int. J. Electr. Power Energy Syst.* **2026**, *176*, 111742. [\[CrossRef\]](#)
- Xin, J.; Liao, Z.; Wen, F. Intelligent Alarm Processing and Fault Diagnosis in Digital Substations. In Proceedings of the 2010 International Conference on Power System Technology, Hangzhou, China, 24–28 October 2010; pp. 1–5. [\[CrossRef\]](#)
- Chen, H.; Li, J.M.; Wang, X.B.; Yu, L.Q.; Yang, Z.X. Review of Intelligent Fault Diagnosis for Rotating Machinery under Imperfect Data Conditions. *Expert Syst. Appl.* **2025**, *285*, 127726. [\[CrossRef\]](#)

22. Xu, B.; Yin, X.; Yin, X.; Wang, Y.; Pang, S. Fault Diagnosis of Power Systems Based on Temporal Constrained Fuzzy Petri Nets. *IEEE Access* **2019**, *7*, 101895–101904. [[CrossRef](#)]
23. Xu, C.; Cheng, X.; Zhuang, X.; Li, J. Fault diagnosis of distribution network based on time constraint intuition fuzzy Petri nets. *Meas. Energy* **2025**, *5*, 100034. [[CrossRef](#)]
24. Guo, W.; Wen, F.; Liao, Z.; Wei, L.; Xin, J. An Analytic Model-Based Approach for Power System Alarm Processing Employing Temporal Constraint Network. *IEEE Trans. Power Deliv.* **2010**, *25*, 2435–2447. [[CrossRef](#)]
25. Kang, T.; Wu, W.; Zhang, B. Temporal Abductive Reasoning based Intelligent Alarm for Power System. In Proceedings of the 2010 Asia-Pacific Power and Energy Engineering Conference, Hangzhou, China, 24–28 October 2010; pp. 1–5. [[CrossRef](#)]
26. Feng, Y.; Meng, Y. Event-triggered Control Design of DC Microgrids Based on Dual-Channel Network Switching System under Denial-of-Service Attacks. In Proceedings of the 2023 3rd International Conference on New Energy and Power Engineering (ICNEPE), Huzhou, China, 24–26 November 2023; pp. 815–822. [[CrossRef](#)]
27. Yuan, Z.; Zhao, M.; Lu, M. A Fault Diagnosis Method With Temporal Constraints For Power System Based On Intuitionistic Fuzzy Colored Petri nets. In Proceedings of the 2026 IEEE PES International Meeting (PES IM), Hong Kong, China, 18–21 January 2026; pp. 1–5. [[CrossRef](#)]
28. Zhang, Y.; Chung, C.Y.; Wen, F.; Zhong, J. An Analytic Model for Fault Diagnosis in Power Systems Utilizing Redundancy and Temporal Information of Alarm Messages. *IEEE Trans. Power Syst.* **2016**, *31*, 4877–4886. [[CrossRef](#)]
29. Li, H.; Yu, H.; Liu, Z.; Li, F.; Wu, X.; Cao, B.; Zhang, C.; Liu, D. Long-term Scenario Generation of Renewable Energy Generation using Attention-based Conditional Generative Adversarial Networks. *Energy Convers. Econ.* **2024**, *5*, 15–27. [[CrossRef](#)]
30. IEEE Std C37.118.2-2024 (Revision of IEEE Std C37.118.2-2011); IEEE Standard for Synchrophasor Data Transfer for Power Systems. IEEE: Piscataway, NJ, USA, 2024; pp. 1–79. [[CrossRef](#)]
31. Liang, X.; Wallace, S.A.; Nguyen, D. Rule-Based Data-Driven Analytics for Wide-Area Fault Detection Using Synchrophasor Data. *IEEE Trans. Ind. Appl.* **2017**, *53*, 1789–1798. [[CrossRef](#)]
32. Li, J.; Wang, X.B.; Chen, H.; Yang, Z.X. Physical-Knowledge-Guided and Interpretable Deep Neural Networks for Gear Fault Severity Level Diagnosis. *IEEE Trans. Ind. Informat.* **2025**, *21*, 4892–4903. [[CrossRef](#)]
33. Huang, H.; Lv, Y.; Xue, L.; Zhang, X.; Shen, Z.; Jin, B. Power Grid Fault Diagnosis based on Support Vector Machine. In Proceedings of the 2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2), Taiyuan, China, 28 October 2021; pp. 2301–2307. [[CrossRef](#)]
34. Jiang, Z.; Wang, H.; Wang, W. A Petri Net Strategy for Fault Diagnosis and Location in Power Distribution Systems to Prevent Local Power Shortages. *IEEE Access* **2024**, *12*, 161038–161053. [[CrossRef](#)]
35. Kejie, L.; Nianwen, X.; Jianben, L. A Multi-sensor Fusion Approach for Susceptibility Analysis in Smart Substation based on Deep Bayesian Network. In Proceedings of the 2022 IEEE International Conference on High Voltage Engineering and Applications (ICHVE), Chongqing, China, 25–29 September 2022; pp. 1–4. [[CrossRef](#)]
36. Mallinath, S.; Dutta, S.; Jayalakshimi, N.S.; Jadoun, V.K. Shifting of Research Trends in Fault Detection and Estimation of Location in Power System. *IEEE Access* **2025**, *13*, 70721–70740. [[CrossRef](#)]
37. Park, Y.M.; Kim, G.W.; Sohn, J.M. A Logic based Expert System (LBES) for Fault Diagnosis of Power System. *IEEE Trans. Power Syst.* **1997**, *12*, 363–369. [[CrossRef](#)]
38. Ma, C.; Li, J.; Zhang, N.; Bu, F.; Yang, Z. Open-circuit Radial Stray Magnetic Flux Density based Noninvasive Diagnosis for Mixed Eccentricity Parameters of Interior Permanent Magnet Synchronous Motors in Electric Vehicles. *IEEE Trans. Ind. Electron.* **2022**, *70*, 1983–1992. [[CrossRef](#)]
39. Ye, K.; Li, Q.; Yan, Y.; Wang, X.; Qi, D. DTA-Head: Dynamic Task Alignment Head for Regression and Classification in Small Object Detection. *Appl. Sci.* **2025**, *15*, 9060. [[CrossRef](#)]
40. Xiao, X.; Zhou, Y.; Zhang, W.; Wang, Y.; Zheng, Z.; Hu, W. Power Disturbance Waveform Analysis and Proactive Application in Power Systems. *Energy Convers. Econ.* **2023**, *4*, 123–136. [[CrossRef](#)]
41. Chen, R.; Li, X.; Chen, Y. Optimal Layout Model of Feeder Automation Equipment Oriented to the Type of Fault Detection and Local Action. *Prot. Control Mod. Power Syst.* **2023**, *8*, 2. [[CrossRef](#)]
42. Wang, X.; Du, H.; Gao, J.; Wei, X.; Liang, Z.; Guo, L.; Liu, W. Grounding Fault Location Method of Overhead Line Based on Dual-Axis Magnetic Field Trajectory. *Prot. Control Mod. Power Syst.* **2023**, *8*, 4 [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.