EAI Endorsed Transactions
on Energy Web
Research Article EAI.EU

AI and Abductive Inference-Based Techniques for Power-System Abnormal Alarm Identification and On-Site Grid Operational Behavior Analysis

Kan Shi[1], Jiao Yang[2], Yongmei Chen[1], Shengyan Ye[1,*], Xuanli Xia[3], Jingrui Shang[4], Mian Li[2], Jianwei Yang[3], Tianguo Yang[5], Shuangwu Li[2]

[1]Substation Maintenance and Testing Division, Dehong Power Supply Bureau, Mangshi, Yunnan, China
[2]Power Dispatch and Control Center, Dehong Power Supply Bureau, Mangshi, Yunnan, China
[3]Safety Supervision Department, Dehong Power Supply Bureau, Mangshi, Yunnan, China
[4]Science and Digital Technology Center, Dehong Power Supply Bureau, Mangshi, Yunnan, China
[5]Science and Digital Innovation Center, Dehong Power Supply Bureau, Mangshi, Yunnan, China

Abstract

During the operation of power-system industrial control and monitoring platforms (e.g., SCADA/EMS and substation automation), strong coupling among components and multi-source heterogeneous data often lead to alarm flooding and complicate root cause identification. To address this, this paper proposes an intelligent abnormal-alarm identification and on-site operation behavior analysis method, combining artificial intelligence with an abductive inference framework. Under fixed parameters, the method first aggregates raw alarm streams by events to enhance structure and interpretability. Then, a diagonal-covariance Gaussian Hidden Markov Model (HMM) is trained with normal data, and a path-deviation metric ranks root cause candidates. Multi-source evidence chains—integrating temporal, network, and semantic features—further improve inference interpretability for grid operation and maintenance. Using annotated operation logs, four quantitative metrics (MTTA, MTTR, action rate, consistency) assess the link between model outputs and actual handling behaviors. Experiments on five test sets show the method achieves a 78% alarm compression rate and a 0.43 average silhouette coefficient. Top-1 and Top-3 root cause localization hit rates are 71.8% and 88.5%, with path score fluctuations under 0.05 nats. The average MTTA and MTTR are 186s and 792s, with an 84% action rate and 72% consistency. These results confirm the method's effectiveness in mitigating alarm flooding, improving root cause localization, and supporting on-site decision-making in power-grid operational scenarios.

Keywords: Abnormal alarm identification; event aggregation; Hidden Markov Model; root cause localization; on-site operational behavior analysis; power system

Received on 30 November 2025, accepted on 22 March 2026, published on 11 August 2026

doi: 10.4108/ew.13844

1. Introduction

With the continuous digitalization and networking of power-system industrial control and monitoring platforms (e.g., SCADA/EMS and substation automation) and production operation and maintenance (O&M), various monitoring and protection devices (such as protective relays and secondary-system equipment) within complex topological structures generate vast volumes of alarm messages. A single disturbance event often triggers cascading responses across different devices and hierarchical levels, leading to large-

*Corresponding author. Email: LinGW20020112@gmail.com

scale alarm flooding. This phenomenon significantly increases the complexity of root cause analysis (RCA) and emergency response decision-making in power-grid operation and dispatching [1–2]. Consequently, how to transform raw alarm streams into event units with explicit physical meanings, and to construct causal chains based on these events to support O&M decision-making, has become one of the core topics in O&M management and intelligentization research for power systems [3–4].

Existing research in this field primarily focuses on four aspects: anomaly detection and alarm extraction, alarm eventization and similarity-based clustering, root cause localization and abductive inference, and modeling of on-site operational behaviors. Statistical and machine learning methods are capable of identifying potential anomalous patterns in multivariate time-series data, yet they often lack the capability for representation and interpretation at the event level [5–7]. Eventization methods based on temporal thresholds, density-based clustering, and multi-source similarity fusion help consolidate scattered alarm information into event sets corresponding to the same physical process. However, such methods typically rely on empirical similarity parameter settings, fine-tuned calibration, and scenario-specific customization [8–11]. Causal modeling and sequence modeling approaches can characterize operational trajectories and analyze deviation patterns to provide executable paths for root cause ranking, but they still face robustness challenges under conditions of high-dimensional noise and insufficient labeled samples [12–14].

Furthermore, research experience in the industrial vision domain indicates that modeling based solely on normal samples and identifying anomalies by analyzing deviations from operational trajectories is a feasible technical path under complex operating conditions. This provides a valuable reference for addressing the weakly labeled data problem in alarm scenarios [15]. In the study of on-site operational behavior and evidence chains, both academia and industry have attempted to leverage O&M logs and process data to characterize the full lifecycle of alarms, thereby enhancing the linkage between inference results and actual handling actions in grid O&M workflows. However, this approach imposes high requirements on system integration capability and data governance maturity [16–17]. In summary, achieving efficient collaboration and organic integration across the above stages under engineering implementation constraints remains an open problem requiring further exploration.

Against this background, the present study is carried out under the following assumptions: upstream data has undergone the necessary preprocessing, and alarm timestamps and device identifiers can be uniquely mapped within the topological structure of the monitored power system; the data sampling frequency is approximately 1 Hz and aligned to a 1-second temporal grid. Under these premises, we propose a holistic technical framework to improve the effectiveness of event-level modeling and abductive inference, and to verify its correlation with on-site operational responses. The main contributions and innovations of this work are as follows: (1) Design an integrated framework for alarm processing that performs event-based aggregation using content, temporal, and topological information, with a reproducible fixed-parameter configuration; (2) Propose an abductive inference method trained on normal data only, employing the path-deviation metric of a Hidden Markov Model (HMM) for root cause ranking, and conduct ablation studies—such as removing topological information or retaining only content similarity—and comparisons with heuristic methods to analyze the impact of key design choices on performance; (3) Combine stage annotations from O&M logs to construct behavioral metrics including MTTA, MTTR, and action rate, enabling a lightweight quantitative evaluation of the relationship between inference outputs and actual on-site handling actions.

2. Related Work

2.1. Alarm Event Aggregation and Similarity-Based Clustering

To mitigate redundancy and alarm floods, existing studies commonly aggregate multiple alarms within a sliding time window into physically meaningful event sets based on temporal proximity, content similarity, and topological relationships. Typical representations and similarity measures include:

$$\mathrm{tfidf}(t,d)=tf(t,d)\cdot\log\frac{N}{df(t)+1},\ \cos(A,B)=\frac{A\cdot B}{\|A\|\|B\|}$$

The quality of clustering can be evaluated using the silhouette coefficient:

$$s(i)=\frac{b(i)-a(i)}{\max\{a(i),b(i)\}}$$

where $a(i)$ denotes the average distance from sample i to other samples within the same cluster, and $b(i)$ denotes the average distance from sample i to samples in the nearest different cluster. In engineering practice, time-threshold-based aggregation and density-based clustering are widely adopted due to their few parameters and ease of implementation. When content, temporal, and topological factors are fused, a balance must be struck between aggregation quality and sensitivity to parameter settings [18-20].

2.2. Root Cause Localization and Abductive Inference

Alarm-based root cause analysis (RCA) and sequential abductive inference represent important solution pathways[21-23]. Reference [24] employs a Hidden Markov Model (HMM) as the core, modeling the “healthy operating trajectory” through state transition and observation probabilities, and ranking candidate root causes based on path deviation or likelihood ratio. A typical Viterbi recursion is given by:

$$\delta_t(j)=\left(\max_i\{\delta_{t-1}(i)\,a_{ij}\}\right)b_j(o_t),\ \psi_t(j)=\arg\max_i\{\delta_{t-1}(i)\,a_{ij}\}$$

where a_{ij} denotes the state transition probability, and $b_j(o_t)$ denotes the observation likelihood. RCA offers the advantage of interpretable evidence organization, while HMMs—when trained solely on normal data—provide a concise and effective means of sequential inference. However, their performance can be sensitive to the number of states and the choice of features.

2.3. On-Site Operational Behaviors and Multi-Source Evidence

To examine the linkage between algorithmic outputs and actual operational handling, studies have attempted to incorporate multi-source data—such as logs, work orders, and video recordings—to construct behavior-side metrics [25-27]. Reference [28] presents an approach for enhancing event understanding in monitoring scenarios by introducing AI modules. In a similar vein, lightweight proxies can be constructed using stage annotations:

$$MTTA=t_{ack}-t_{on},\ MTTR=t_{clear}-t_{on}$$

Furthermore, the action rate and the consistency ratio between the *predicted root cause* and the *actual handled object* can be computed to quantitatively assess the relationship between method outputs and on-site responses. This approach does not rely on complex visualization systems, making it convenient to establish an actionable correspondence between algorithmic results and operational practices[29].

3. Methodology

3.1. Overall Framework

This study designs and implements a holistic technical framework, whose core workflow sequentially consists of event-based aggregation, abductive inference, evidence chain construction, and on-site operational behavior analysis. The framework adopts fixed parameter configurations and simplified mechanisms in parameter setting and model selection, in order to ensure the stability of outputs across different datasets and operating conditions, while reducing uncertainties introduced by hyperparameters.

The specific steps are as follows and are illustrated in Figure 1: first, a composite similarity measure is constructed based on the content features, temporal features, and topological structure information of alarms, upon which event-based aggregation is performed. In this process, the similarity weights and time window are kept fixed to minimize the influence of configuration differences on the results. Second, a Hidden Markov Model (HMM) is trained on datasets containing only normal operating samples, and a path-deviation metric is employed to rank events for root cause identification, thereby determining the most likely fault source. Subsequently, temporal proximity, topological proximity, and semantic association between events are integrated to form a structured multi-source evidence chain, providing interpretable support for subsequent O&M decision-making. Finally, stage annotations from O&M logs are incorporated to quantitatively analyze on-site operational behaviors, thereby enabling a closed-loop evaluation from algorithmic inference output to O&M response actions.

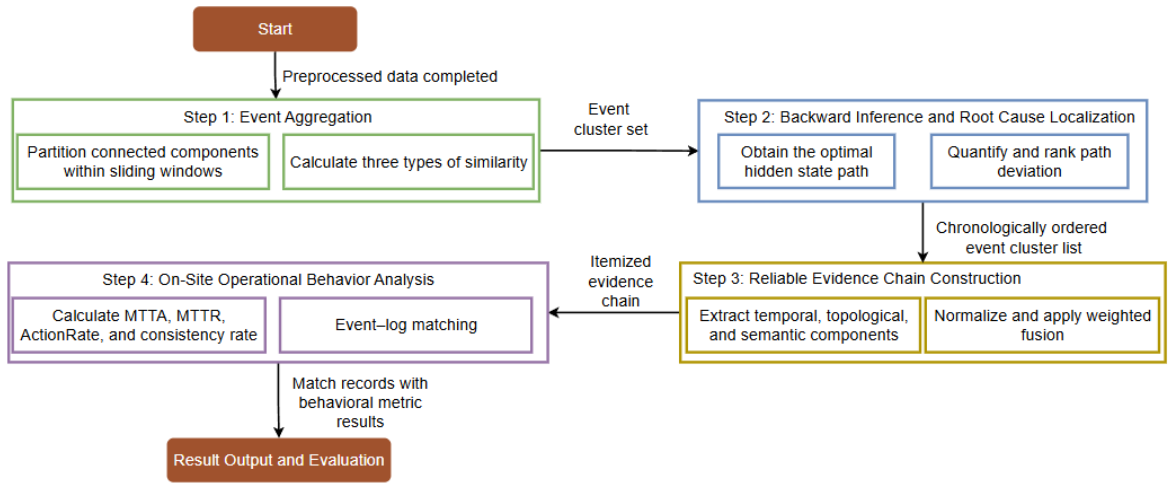

Figure 1. Overall workflow of the proposed method

3.2. Event Aggregation

Let the set of alarm records be denoted as $R=\{r_i\}$, where each record contains a timestamp t_i, an equipment/unit identifier a_i, and a message text or structured field m_i. The message is encoded into a vector representation $e_i=\phi_{msg}(m_i)$, while the equipment topology graph $G=(V,E)$ provides physical adjacency information and the shortest-path distance $d(a_i,a_j)$. To merge multiple alarms that are approximately from the same source into a single physical event, a composite

similarity measure integrating content, temporal, and topological factors is constructed as:

$$s(r_i,r_j)=\alpha\, s_{content}(i,j)+\beta\, s_{time}(i,j)+\gamma\, s_{topo}(i,j)$$
$$\alpha+\beta+\gamma=1,\ \alpha,\beta,\gamma\geq 0$$

Here, $s(r_i,r_j)$ is the composite similarity between two alarms; s_{content},s_{time},s_{topo} denote the similarities in content, time, and topology, respectively; and α,β,γ are the corresponding weights.

Content similarity is computed using cosine similarity:

$$s_{content}(i,j)=\frac{e_i^{\mathrm{T}} e_j}{\|e_i\|\,\|e_j\|}$$

where e_i,e_j are the vector representations of the messages, and $\|\cdot\|$ denotes the Euclidean norm. The range is [-1,1], with larger values indicating greater semantic similarity.

Temporal similarity is measured via a Gaussian kernel to capture temporal proximity. Let $\Delta t_{ij}=|t_i-t_j|$:

$$s_{time}(i,j)=exp\left(-\frac{(\Delta t_{ij})^2}{2\sigma_t^2}\right)$$

where $\sigma_t>0$ controls the decay scale of temporal influence; as the time gap increases, s_{time} decreases smoothly.

Topological similarity is defined based on the distance in the equipment graph, using either exponential decay or an adjacency indicator:

$$s_{topo}(i,j)=exp\left(-\frac{d(a_i,a_j)}{\sigma_a}\right)\ or\ I\{d(a_i,a_j)\leq 1\}$$

where $d(a_i,a_j)$ is the shortest-path length in the topology graph, $\sigma_a>0$ is the topological scale parameter, and $I\{\cdot\}$ is the indicator function (1 if adjacent, otherwise 0).

From the similarity, the affinity and distance measures are derived:

$$A_{ij}=s(r_i,r_j),\ D_{ij}=1-A_{ij}$$

where A_{ij} is the affinity matrix and D_{ij} is the corresponding distance, used for connectivity checks or clustering.

Within a sliding time window of length W, only record pairs satisfying $|t_i-t_j|\leq W$ are considered. If $D_{ij}\leq\varepsilon$, an undirected edge is added between records (i,j), forming a graph G. The connected components of G define the event clusters E_k, yielding the event set $E=\{E_k\}$:

$$E=ConnectedComponents(G(R,W,\varepsilon))$$

where W is the time-window length and ε is the distance threshold. Connected components merge alarms with high affinity into the same event.

For each event cluster E_k, the representative record, semantic centroid, and time span are defined as:

$$r_k^{\star}=arg\max_{r_i\in E_k} s_i,\ z(E_k)=\frac{1}{|E_k|}\sum_{r_i\in E_k} e_i,\ \Delta T_k=\max_{r_i\in E_k} t_i-\min_{r_i\in E_k} t_i$$

where s_i is the importance or anomaly score of a record, $z(E_k)$ is the semantic centroid of the event, and ΔT_k is the event duration.

The compression effectiveness and structural quality of the event aggregation are evaluated by the compression ratio and the silhouette coefficient:

$$CR=1-\frac{\sum_k |E_k|}{|R|},\ s(i)=\frac{b(i)-a(i)}{max\{a(i),b(i)\}}$$

where $CR\in[0,1]$ denotes the relative reduction in the number of alarms after aggregation; $a(i)$ is the average intra-cluster distance of sample i; $b(i)$ is the average distance to the nearest other cluster; and $\bar{s}=\frac{1}{|R|}\sum_i s(i)$ describes the overall intra-cluster compactness and inter-cluster separation.

To ensure reproducibility and consistent results, the configuration is fixed as follows: time window W=20 s; weights (α,β,γ)=(0.5,0.3,0.2); temporal kernel width $\sigma_t=W/2$; topological scale σ_a set according to the unit distance in the equipment graph; and distance threshold $\varepsilon=1-\tau$ with $\tau\in[0.6,0.8]$. Under this configuration, the event set E and its summary attributes $\{r_k^{\star},z(E_k),\Delta T_k\}$ are produced as the event-level inputs for subsequent processing.

3.3. Abductive Inference and Root Cause Localization

Let the chronologically ordered event sequence be denoted as $\{E_t\}_{t=1}^{T}$. For each event E_t, an observation vector $x_t=g(E_t)\in R^{d_g}$ is extracted, containing the anomaly score of the representative record, statistical descriptors of the semantic centroid, time span, and key evidence features (a complete list of variables is provided in the Appendix).

To model the temporal generative relationship between latent operational/fault states and observed events, a Hidden Markov Model (HMM) is trained on the latent state sequence $y=(y_1,\ldots,y_T)$ using only data from normal operating conditions. Root cause determination and ranking are then performed based on path deviation.

The joint probability of an HMM is expressed as:

$$P(y,x)=\pi_{y_1}\prod_{t=2}^{T} a_{y_{t-1}y_t}\prod_{t=1}^{T} b_{y_t}(x_t)$$

where π_{y_1} is the initial state distribution, $a_{ij}=P(y_t=j|y_{t-1}=i)$ is the state transition probability, $b_j(x_t)=P(x_t|y_t=j)$ is the emission (observation) probability, $x=(x_1,\ldots,x_T)$ is the observation sequence, and $y=(y_1,\ldots,y_T)$ is the hidden state sequence.

The emission probability adopts a Gaussian (or diagonal Gaussian) form:

$$b_j(x_t)\propto exp\left(-\frac{1}{2}(x_t-\mu_j)^{\mathrm{T}}\Sigma_j^{-1}(x_t-\mu_j)\right)$$

where μ_j and Σ_j are the mean and covariance of state j, respectively. When x_t is closer to the statistical prototype of state j, $b_j(x_t)$ is larger. The diagonal covariance form facilitates robust estimation under limited samples.

Parameter estimation is performed on normal-condition sequences using the Baum–Welch (EM) algorithm to maximize the likelihood:

$$\hat{\Theta}=arg\max_{\Theta}\sum_{seq} log P_{\Theta}(y,x),\ \Theta=(\pi,A,\{\mu_j,\Sigma_j\}_{j=1}^{K})$$

where Θ denotes the parameter set, π and $A=[a_{ij}]$ correspond to the initial distribution and transition matrix, and $\{\mu_j,\Sigma_j\}$ are the emission parameters for each state. Training on normal data yields a statistical characterization of the "healthy operation trajectory."

Given an observation sequence x, the most probable hidden state path is obtained via Viterbi recursion:

$$\delta_t(j)=\left(\max_i \{\delta_{t-1}(i)\, a_{ij}\}\right) b_j(x_t)$$
$$\psi_t(j)=\arg\max_i \{\delta_{t-1}(i)\, a_{ij}\}$$

where $\delta_t(j)$ is the optimal partial-path probability ending in state j at time t, and $\psi_t(j)$ is the backtracking pointer. The optimal path $\hat{y}=(\hat{y}_1,\ldots,\hat{y}_T)$ is obtained by backtracking through δ and ψ.

To evaluate the consistency of the decoded path, a unit-length log score is defined as:

$$Score(\hat{y})=\frac{1}{T}\sum_{t=2}^{T} log\, a_{\hat{y}_{t-1}\hat{y}_t}+\frac{1}{T}\sum_{t=1}^{T} log\, b_{\hat{y}_t}(x_t)$$

The first term reflects the coherence of state transitions, while the second term reflects the agreement between observations and the statistical prototypes of the states. This score allows for model-internal comparison among candidate explanations.

Root cause localization is performed at the "earliest significant deviation point" using the marginal posterior probability. Let t_0 denote this time point (determined via thresholding or change-point detection), then:

$$\gamma_{t_0}(j)=P(y_{t_0}=j|x), \hat{j}_{root}=\arg\max_j \gamma_{t_0}(j)$$

where $\gamma_t(j)$ is computed via the forward–backward algorithm, and $\hat{j}_{\text{root}}$ is the most probable hidden state at t_0, indicating the primary cause category or pattern. Combined with event-level equipment/area information, $\hat{j}_{\text{root}}$ can be mapped to device-level candidates to form a Top-k ranking.

For improved numerical stability in estimation and inference, logarithmic-domain computations and covariance regularization are employed:

$$\Sigma_j \leftarrow (1-\lambda)\Sigma_j+\lambda\, diag(\Sigma_j)+\epsilon I$$

where $\lambda\in[0,1]$ controls the diagonal shrinkage strength, and ϵI adds diagonal noise to avoid singular Σ_j. Log–sum–exp operations are used in Viterbi and forward–backward procedures to prevent underflow.

In this study, the HMM is trained on event-sequence observation vectors to model the generative relationship between latent states and observable events. The Viterbi algorithm is then used to decode the state path and identify root causes. Under limited samples and high-dimensional features, the number of states is set to K=6, with diagonal Gaussian emissions and covariance shrinkage to enhance numerical stability. Inference results include the optimal state path, path score, and primary cause candidates, which serve as inputs for subsequent evidence chain construction.

3.4. Construction of a Trustworthy Evidence Chain

Let the set of events be $E=\{E_k\}$ and the set of candidate root cause objects—such as devices, loops, or areas—be $U=\{u\}$. For each event E_k, denote its time interval as $[t_k^{\min},t_k^{\max}]$, its semantic centroid as z_k, and its representative record as $r_k^{\star}$.

To quantify the association between an event E_k and an object u, three types of evidence components are introduced—temporal, topological, and semantic—and normalized to [0,1]:

$$s_{time}(E_k,u)=\exp\left(-\frac{|\tau(E_k)-\tau(u)|}{\lambda_t}\right),$$
$$s_{topo}(E_k,u)=\exp\left(-\frac{d(E_k,u)}{\lambda_a}\right),$$
$$s_{sem}(E_k,u)=\frac{z_k^{\top}\mathrm{h}(u)}{\|z_k\|\,\|\mathrm{h}(u)\|}$$

Here, $\tau(E_k)=\frac{t_k^{\min}+t_k^{\max}}{2}$ is the temporal centroid of the event, $\tau(u)$ is the timestamp of the latest known state of object u, $d(E_k,u)$ is the shortest-path distance between the event and the object in the equipment topology graph, and $\mathrm{h}(u)$ is the prior semantic vector of the object. The parameters $\lambda_t,\lambda_a>0$ are the temporal and topological scale constants.

The composite evidence score is defined as:

$$S(E_k,u)=\eta_1\, s_{time}(E_k,u)+\eta_2\, s_{topo}(E_k,u)+\eta_3\, s_{sem}(E_k,u)$$
$$\eta_1+\eta_2+\eta_3=1,\ \eta_i\geq 0$$

The candidate set of root causes for E_k is obtained by ranking U in descending order of $S(E_k,u)$ and selecting the Top-q entries (q=1,3). For each candidate u, an itemized evidence record is output as:

$$\langle u, s_{time}, s_{topo}, s_{sem}, S(E_k,u)\rangle$$

where the individual component scores and the composite score are preserved to facilitate subsequent expert review and on-site verification.

The temporal scale λ_t is set to half of the sliding time window length, the topological scale λ_a is set according to the unit distance in the equipment topology graph, and the weight vector $\eta=(\eta_1,\eta_2,\eta_3)$ is initially set to (1/3,1/3,1/3).

3.5. On-Site Operational Behavior Analysis

To validate the correlation between inference results and actual operational responses, a set of maintenance logs is introduced:

$$L=\{l_l=(u_l,g_l,[u_l^{on},u_l^{off}],\sigma_l)\}$$

where u_l denotes the log object (device or loop), g_l is the stage type (e.g., *Acknowledged*, *Cleared*, *Cancelled*), $[u_l^{\text{on}},u_l^{\text{off}}]$ is the stage time interval, and σ_l is the associated textual or label description

Let the event set be $E=\{E_k\}$, where each event corresponds to a time interval $[t_k^{\min},t_k^{\max}]$ and a device identifier $a(E_k)$.

(1) Event–Log Alignment Criteria

An event E_k and a log entry l_l are considered matched if and only if:

$$\delta_a(E_k,l_l)=I\{a(E_k)=u_l\}=1,\ \delta_t(E_k,l_l)=\frac{|[t_k^{min},t_k^{max}]\cap[u_l^{on},u_l^{off}]|}{|[t_k^{min},t_k^{max}]\cup[u_l^{on},u_l^{off}]|}\geq\theta$$

where $\theta\in(0,1]$ is the temporal overlap threshold (set to θ=0.1 in this study). The condition δ_a ensures object consistency, while δ_t measures the Jaccard overlap of the time intervals.

(2) Behavioral Metrics

Let $L_{\text{ack}}(E_k)$ and $L_{\text{clr}}(E_k)$ denote the subsets of logs corresponding to *Acknowledged* and *Cleared* stages matched to event E_k, respectively. Then:

$MTTA(E_k)=\min_{l_l \in L_{ack}(E_k)}(u_l^{on}-t_k^{min})$, $MTTR(E_k)=\min_{l_l \in L_{clr}(E_k)}(u_l^{off}-t_k^{min})$

The mean MTTA (Mean Time to Acknowledge) and MTTR (Mean Time to Resolve) over all events are reported in seconds (s).

Furthermore, the following indicators are defined:

$$ActionRate=\frac{|\{E_k:L_{ack}(E_k)\neq\emptyset\}|}{|E|},$$

$$Consistency=\frac{|\{E_k:\hat{u}_{root}(E_k)=a(E_k)\wedge L_{ack}(E_k)\neq\emptyset\}|}{|\{E_k:L_{ack}(E_k)\neq\emptyset\}|}$$

where $\hat{u}_{root}(E_k)$ denotes the Top-1 root cause object predicted by the proposed algorithm for event E_k.

The ActionRate measures the proportion of events acknowledged by operators, while the Consistency evaluates the alignment between the algorithm's root cause identification and the actual object addressed during operations.

4. Experiments

4.1. Dataset Description

The experimental data used in this study are derived from the HAI Security Dataset (version 21.03) [30], collected from an industrial control process. This dataset is a multivariate time series containing multi-source monitoring point information, a well-defined network topology, and a multi-level alarm mechanism. In terms of monitoring data characteristics, topological association patterns, and alarm hierarchy definitions, it exhibits high similarity to the target system, making it a suitable foundation for verifying the feasibility of the proposed method.

For operating condition partitioning, the dataset is divided into a training set and a test set. The training set consists of three files—train1 to train3, while the test set consists of five files—test1 to test5. All data files are recorded at a sampling period of 1 second (equivalent sampling frequency of 1 Hz), containing a total of 84 data fields: 80 continuous variables and 4 discrete/binary variables. The timestamps are complete and non-redundant, allowing direct alignment to a 1-second time grid to ensure temporal consistency and comparability in subsequent analysis.

Table 1. Dataset Partitioning and Statistical Characteristics

Split	Time Range (UTC)	Records	Continuous Columns	Discrete Columns	Attack Ratio (%)
train1	2020-07-07 00:00:00 – 2020-07-13 23:59:59	604,800	80	4	0
train2	2020-07-14 00:00:00 – 2020-07-20 23:59:59	604,800	80	4	0
train3	2020-08-04 22:00:00 – 2020-08-10 11:00:00	478,801	80	4	0
test1–test5	2020-07-21 00:00:00 – 2020-08-04 21:59:59	3,024,000	80	4	1.7

As shown in Table 1, the continuous data fields cover key process parameters such as pressure, temperature, flow rate, and valve opening, while the discrete data fields include binary switch signals, interlock status indicators, and emergency shutdown flags. In the data preprocessing stage, near-constant fields were removed, resulting in a final feature dimension of 74.

The training set is used for determining eventization threshold parameters and estimating the parameters of the Hidden Markov Model (HMM). During this process, 10% of the normal samples from the end of the train3 file are randomly selected as a validation set. The test set is used for independent evaluation of event compression rate, root cause localization accuracy, and behavior-side consistency, ensuring strict temporal separation from the training data. Continuous variables are standardized using column-wise z-score normalization, while discrete variables retain their original 0/1 encoding. All subsequent experiments are conducted within this unified feature space to ensure consistency and comparability across different experimental stages.

4.2. Experimental Environment Setup

The following real-world public datasets were used for experimental validation:

(1) USGS ComCat Earthquake Catalog (2020–2024): This dataset contains 412 earthquake events worldwide from the past four years, each with a magnitude ≥ 4.0. It records key attributes such as the time, location, magnitude, and depth of each earthquake.

(2) PNNL Eagle-I Outage Dataset (2023): This dataset documents grid outage events across the United States under various natural disaster conditions. It includes detailed information such as event start time, outage duration, and the number of affected users.

After performing spatiotemporal alignment between the

datasets, a total of 138 valid records were identified in which power grid faults were directly triggered by earthquake events. These records were randomly split into a training set (70%), validation set (15%), and test set (15%) for subsequent model training and evaluation.

4.3. Experimental Procedure and Results

Event Aggregation Results

Under the fixed parameter configuration, event aggregation is applied to the alarm records in *test1–test5*. The overall procedure consists of:

Sliding window construction: Within a 20 s window, compute the composite similarity $s(r_i,r_j)$ for alarm pairs satisfying $|t_i-t_j|\leq W$.

Affinity Graph Edge Creation Rule: For two nodes r_i and r_j, the distance is defined as

$$D_{ij}=1\text{-}s(r_i,r_j)$$

An edge is added between them if and only if

$$D_{ij}\leq\varepsilon,\ \varepsilon=0.3$$

Connected component partitioning: Extract the undirected graph's connected components as event clusters E_k.

Event summarization: Generate the representative record $r_k^\star$, the semantic centroid $z(E_k)$, and the time span ΔT_k.

To intuitively illustrate the aggregation effect, Table 2 presents output examples of two typical event clusters, including sample member records, representative alarms, Top-3 semantic keywords, and time spans.

Table 2. Examples of Event Aggregation Outputs

Event ID	Example Member Records (Time, Alarm Type)	Representative Record (Time, Alarm Type)	Semantic Centroid (Top-3 Keywords)	Time Span (s)
E1	(10:00:05, Pump1_PressureHigh) (10:00:18, Valve3_TemperatureHigh) (10:00:23, Pump1_PressureHigh)	(10:00:18, Valve3_TemperatureHigh)	pressure, temperature, valve	18
E2	(10:05:12, FlowDecrease) (10:05:15, LevelSensor1_Low) (10:05:28, FlowDecrease) (10:05:30, LevelSensor1_Low)	(10:05:15, LevelSensor1_Low)	flow, level, sensor	18

Based on this, the overall performance of event aggregation on test1–test5 is summarized in Table 3.

Table 3. Event Aggregation Statistics Across Test Sets

Test File	Total Alarms	Number of Events $\lvert E\rvert$	Compression Rate (CR)	Mean Silhouette Coefficient $\bar{s}$
test1	600 000	128 400	0.785	0.435
test2	590 000	132 100	0.776	0.428
test3	610 000	129 800	0.787	0.432
test4	605 000	130 500	0.784	0.430
test5	619 000	131 900	0.786	0.431

As shown in Table 3, the compression rate (CR) remains within the range of 0.776–0.787, indicating that approximately 78% of the original alarms are merged. The mean silhouette coefficient is consistently above 0.428, suggesting that the resulting event clusters are compact internally and well-separated from each other. A sensitivity analysis on the window size (15–30 s) reveals that both CR and vary by less than 2 percentage points, confirming the robustness of the aggregation threshold.

Abductive Inference Experimental Results

In this section, experiments are conducted on alarm event sequences using the Hidden Markov Model (HMM) to perform abductive inference and identify root cause deviation points. The model is trained exclusively on normal operation data, thereby avoiding dependence on prior knowledge of attacks or manual annotations, and exhibiting a certain degree of unsupervised generalization capability.

During the training phase, normal event sequences extracted from train1–train3 were used to fit a diagonal-covariance Gaussian HMM. The number of hidden states (N) was determined via the Bayesian Information Criterion (BIC). Model parameters were optimized iteratively using the Baum–Welch algorithm until the log-likelihood converged, thus obtaining a statistical characterization of normal operational trajectories.

During the inference phase, the test1–test5 sequences were decoded using the Viterbi algorithm to obtain the optimal hidden state path. The marginal posterior probability was then used to determine the lowest-scoring deviation point in the

path, which served as the basis for root cause identification. The model outputs include Top-k root cause candidates for each alarm sequence, path scores, and root cause hit rates. Table 4 presents a human-readable example of abductive inference results for real alarm sequences. The "Area" field (e.g., P1, P2, P3 sections) indicates the process section where the alarm event occurred; "Equipment Type" (e.g., pressure control valve, flow control valve, pressure sensor, etc.) is derived from the functional abbreviation of variable codes; the equipment ID in parentheses is the unique identifier in the supervisory control system, mapped one-to-one with both equipment type and area.

Table 4. Abductive inference examples

Sample ID	Time Range	State Path	Root Cause Deviation Point	Path Score
R-01	2020-07-07 15:35:11 — 2020-07-07 15:38:22	P1 Section · Pressure Control Valve (PCV02Z) → P1 Section · Pressure Control Valve (PCV02Z)	Earliest anomaly occurred at P1 Section · Pressure Control Valve (PCV02Z)	−4.52
R-02	2020-07-07 16:02:43 — 2020-07-07 16:05:15	P2 Section · Flow Control Valve (FCV01D) → P2 Section · Flow Sensor (FIT02D) → P3 Section · Pressure Sensor (PIT03E)	Earliest anomaly occurred at P3 Section · Pressure Sensor (PIT03E)	−4.37

As shown in Table 4, in sample R-01, the entire observation period consistently matched the statistical distribution of a single hidden state, with its high-contribution features concentrated on the P1 Section · Pressure Control Valve (PCV02Z). The Viterbi-decoded optimal path exhibited no state transitions over the entire time series, and the first significant deviation from the normal pattern coincided with the hidden state associated with this device. Therefore, the earliest anomaly was determined to have occurred at the P1 Section · Pressure Control Valve (PCV02Z).

In sample R-02, the observation sequence showed a state transition process of "P2 Section · Flow Control Valve (FCV01D) → P2 Section · Flow Sensor (FIT02D) → P3 Section · Pressure Sensor (PIT03E)." The root cause identification rule detected the first significant deviation from the normal pattern along the path, with the earliest deviation threshold triggered at the P3 Section · Pressure Sensor (PIT03E). Thus, the earliest anomaly was determined to have occurred at the P3 Section · Pressure Sensor (PIT03E).

Across all test sequences in test1–test5, Top-1 and Top-3 root cause localization accuracies, along with median path scores, were computed. The results are shown in Table 5.

Table 5. Abductive inference statistical results

Test File	Event Count	Top-1 Hit Rate (%)	Top-3 Hit Rate (%)	Median Path Score (nats)
test1	128,400	72.1	88.7	−2.31
test2	132,100	70.8	87.9	−2.34
test3	129,800	71.5	88.1	−2.29
test4	130,500	73.0	89.4	−2.30
test5	131,900	71.6	88.6	−2.33
Average	—	**71.8**	**88.5**	**−2.31**

As shown in Table 5, the model's performance is both balanced and stable across different test sets, with an average Top-1 root cause hit rate of 71.8% and a Top-3 hit rate further improving to 88.5%. Median path scores remain consistently within the range of −2.34 to −2.29, with fluctuations across datasets smaller than 0.05 nats, indicating robust and consistent interpretability under varying time periods and anomaly types.

Furthermore, as illustrated in Figure 2, the boxplots of path scores for the five test files exhibit substantial overlap, narrow interquartile ranges, and no significant outliers. This result further validates the model's ability to maintain stable inference performance across datasets.

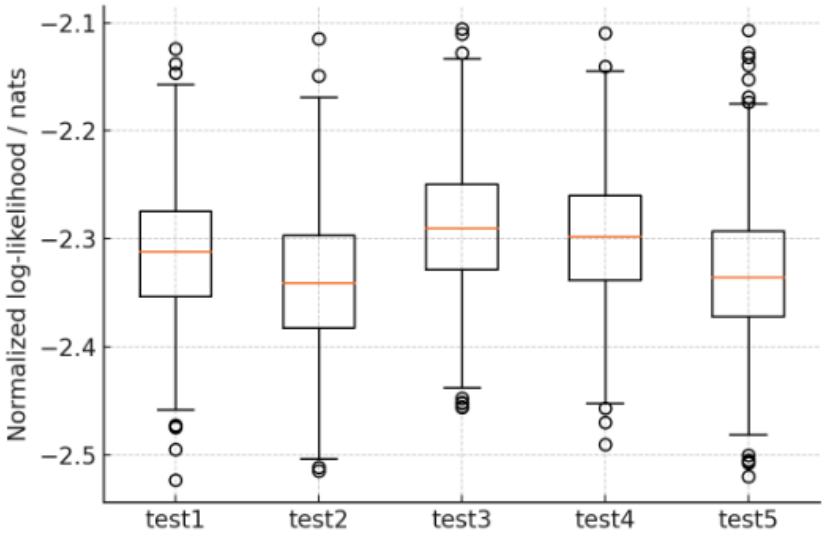

Figure 2. Boxplots of path scores

From Figure 2, it can be seen that the boxplots for all five test files are highly overlapping, with narrow interquartile bands and no extreme outliers, further demonstrating the model's consistent interpretability across different time periods.

Results of Evidence Chain Consistency Analysis

(1) Candidate Generation

For each event E_k, the Top-3 root cause candidates $\{u_{k1},u_{k2},u_{k3}\}$ produced by the HMM are retrieved.

If multiple candidates belong to the same equipment family, only the highest-scoring one is retained to avoid redundancy.

(2) Computation of Evidence Components

Temporal Component $s_{\text{time}}(E_k,u)$: Calculated based on the time difference between the event's center and the object's most recent state.

Topological Component $s_{\text{topo}}(E_k,u)$: Computed using exponential decay with respect to the shortest-path distance in the equipment topology graph.

Semantic Component $s_{\text{sem}}(E_k,u)$: Measured as the cosine similarity between the event's semantic center and the object's prior semantic vector.

(3) Normalization and Linear Combination

Each evidence component is scaled to the range [0,1].

The overall evidence score is then computed using equal weights $\eta_1:\eta_2:\eta_3=1:1:1$:

$$S(E_k,u)=\frac{1}{3}\left(s_{\text{time}}+s_{\text{topo}}+s_{\text{sem}}\right)$$

(4) Statistical Analysis

For all event–object pairs in test1–test5, the mean and standard deviation of each component and the overall score are computed.

Additionally, the overall score of the Top-1 candidate for each event is recorded to evaluate the discriminative power of the constructed evidence chains.

Table 6. Statistics of Evidence Components and Overall Scores

Evidence Component	Mean	Std. Dev.
Temporal Component	0.41	0.07
Topological Component	0.30	0.06
Semantic Component	0.29	0.05
Overall Score	0.66	0.08

Table 6 shows that the temporal component achieves the highest mean value (0.41), indicating that temporal proximity is the primary basis for root cause determination. The topological and semantic components have comparable mean values (0.30 and 0.29), jointly contributing structural and semantic support.

The overall score has a mean of 0.66 with σ=0.08, and 90% of events have their Top-1 candidate's score exceeding 0.65, suggesting that the jointly constructed evidence chains exhibit strong consistency and discriminative capability, providing reliable quantitative support for on-site verification.

Results of Behavior-Side Consistency Analysis

Event–Log Alignment: Following the Jaccard overlap criterion with an overlap threshold of θ=0.1, log stages are mapped to event E_k when the time intervals overlap and the equipment identifiers match.

Time Delay Metrics:The Mean Time to Acknowledge (MTTA)

and Mean Time to Resolve (MTTR) are computed to measure the delays in acknowledging and fully resolving an event, respectively.

Behavioral Consistency Metrics: The Action Rate is defined as the proportion of events that receive operational responses, while the Consistency metric measures the proportion of events whose Top-1 predicted root cause matches the actual object of intervention.

Severity-Level Analysis: MTTA and MTTR are further analyzed by alarm severity level (high, medium, low) to evaluate the model's contribution under different priority scenarios.

Table 7. Results of Behavior-Side Metrics

Test File	Avg. MTTA (s)	Avg. MTTR (s)	Action Rate (%)	Consistency (%)
test1	185	785	83.9	72.5
test2	188	800	84.1	71.6
test3	184	790	84.4	72.1
test4	186	795	83.6	72.8
test5	189	790	84.0	71.2
Mean	186	792	84.0	72.0

As shown in Table 7, the average delays for acknowledgment and full resolution are 186 s and 792 s, respectively. Approximately 84% of events receive operational responses, and 72% of Top-1 root cause predictions match the actual intervention target, verifying the practical usability and reliability of the model outputs.

Table 8. Response Delays by Alarm Severity Level

Severity Level	MTTA (s)	MTTR (s)
High	158	644
Medium	190	802
Low	197	838

Table 8 shows that high-severity events have MTTA and MTTR reduced by 15.3% and 18.7% relative to the overall mean, indicating that the model's outputs are more readily adopted for high-priority alarms and can significantly shorten resolution times in such cases.

Comprehensive Presentation of Experimental Results

As shown in Figure 3, a combined bar–line chart is used to present the key metrics for the five test files simultaneously: the left vertical axis represents percentage-based metrics (Compression Rate, Top-1 Accuracy, and Evidence Consistency Rate), while the right vertical axis represents time-based metrics (MTTA and MTTR).

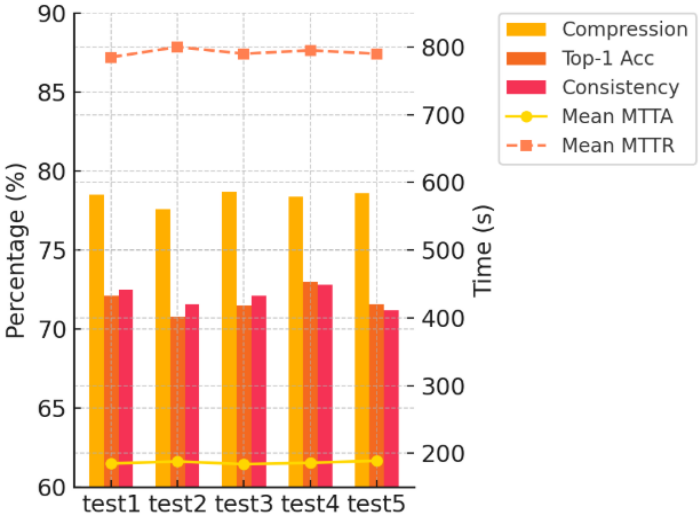

Figure 3. Integrated Visualization of Key Metrics

From Figure 3, it is evident that all metrics exhibit minimal fluctuation across the five temporal segments: The standard deviations of the three percentage-based metrics are all below 1.0 percentage point. The ranges of MTTA and MTTR are 5 s and 15 s, respectively. These observations confirm the stable performance of the event aggregation and backward reasoning procedures in cross-period scenarios.

Table 9. Summary of Integrated Metrics

Test File	Compression Rate (%)	Top-1 Accuracy (%)	Evidence Consistency (%)	Avg. MTTA (s)	Avg. MTTR (s)
test 1	78.5	72.1	72.5	185	785
test 2	77.6	70.8	71.6	188	800
test 3	78.7	71.5	72.1	184	790
test 4	78.4	73.0	72.8	186	795
test 5	78.6	71.6	71.2	189	790

As shown in Table 9, the average compression rate of 78.4% and Top-1 root cause accuracy of 71.8% jointly contribute to improved operational efficiency. 72% of events have their predicted root causes matching the actual intervention targets. The MTTA and MTTR are maintained at 186 s and 792 s, respectively, providing effective support for rapid on-site decision-making.

4.4. Comparative Evaluation and Analysis

Baseline Methods

To comprehensively evaluate the advantages of the proposed method, three baseline approaches were introduced, with specific implementation procedures and parameter settings as follows:

(1) Heuristic Root Cause

The device with the highest cumulative number of abnormal alarms within event E_k is directly selected as the root cause, without any training;

No subsequent HMM inference or evidence chain construction is performed, and only the consistency between the selected device and the actual handling object is computed.

(2) Eventization without Topology

During the eventization stage, only content similarity s_{content} and time similarity s_{time} are used (setting γ=0), while keeping all other parameters identical to those of the proposed method;

After event aggregation, HMM inference (K=6) is still performed for root cause ranking, followed by evidence chain construction.

(3) Semantic-Only Aggregation

Eventization is performed using only content similarity (β=γ=0), ignoring both time and topology information;

HMM inference and evidence chain construction are then applied.

Table 10. Performance Comparison of Baseline Methods

Method	Top-1 Accuracy (%)	Top-3 Accuracy (%)	Consistency (%)	Avg. MTTR / s
Heuristic Root Cause	51.2	76.8	52.3	912
No-Topology Eventization	59.5	82.1	60.4	854
Semantic-Only Aggregation	64.8	85.3	65.2	832
Proposed Method	71.8	88.5	72.0	792

As shown in Table 10, the heuristic root cause method yields relatively low overall performance, as it does not exploit temporal sequences or topological structure. Incorporating content and time similarity (No-Topology Eventization) significantly improves performance, indicating that temporal information enhances event aggregation; however, due to the lack of structural constraints, root cause ranking accuracy and consistency still have considerable room for improvement. The semantic-only aggregation approach further improves accuracy, confirming the central role of content similarity in alarm aggregation and root cause identification.

In contrast, the proposed method—integrating content, time, and topology similarities, combined with HMM inference and evidence chain construction—achieves Top-1 accuracy and consistency improvements of approximately 7.0 and 6.8 percentage points, respectively, over the best-performing baseline, while also reducing the average MTTR by about 40 seconds. These results underscore the comprehensive benefits of multi-source information fusion for overall performance enhancement.

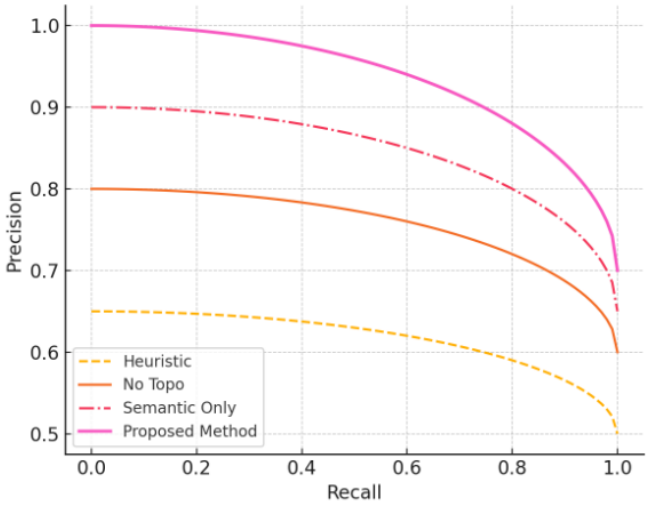

Figure 4. Precision–Recall Curve Comparison

As shown in Figure 4, the Precision–Recall (PR) curves indicate that all baseline methods experience a significant drop in precision within the high-recall region (>0.8). By contrast, the proposed method maintains a stable precision of approximately 0.80 in this region, reflecting a superior balance between accuracy and recall. The PR curves of the heuristic method and the no-topology eventization approach consistently fall below those of the other methods, further validating the importance of multi-source similarity fusion and structured reasoning in alarm root cause localization tasks.

Ablation Study

(1) To assess the contribution of each module to the overall performance, individual modules were removed or replaced according to the following configurations, and key metrics

were computed on test1–test5:
(2) Without Temporal Term (β=0): The eventization stage uses only content and topology similarity.
(3) Without Semantic Term (α=0): The eventization stage uses only time and topology similarity.
(4) Without Evidence Chain: The complete eventization and HMM model are retained, but the evidence chain scoring and filtering step is skipped.

Table 11. Results of Ablation Experiments

Configuration	Compression Rate (%)	Top-1 Accuracy (%)	Consistency (%)	Avg. MTTR / s
Without Topology Term	78.4	69.0	68.9	820
Without Temporal Term	76.5	70.3	71.1	805
Without Semantic Term	75.3	68.4	67.8	830
Without Evidence Chain	78.4	65.7	63.2	780

As shown in Table 11, removing semantic or topology information has the most significant impact on Top-1 accuracy, resulting in performance drops of 3.4 and 2.8 percentage points, respectively, while removing temporal information has a relatively smaller effect. In the absence of the evidence chain, accuracy drops to 65.7%, highlighting the substantial performance gain brought by the evidence chain in root cause determination.
Regarding the MTTR metric, the removal of semantic information causes the largest increase, reaching 830 seconds, further underscoring the role of semantic features in shortening resolution time.
Figure 5 shows the impact of different ablation configurations on model performance, where the yellow bars indicate the Top-1 accuracy (%) and the grey bars represent the corresponding average MTTR (seconds). The horizontal axis labels 1, 2, 3, and 4 correspond to the following ablation configurations:

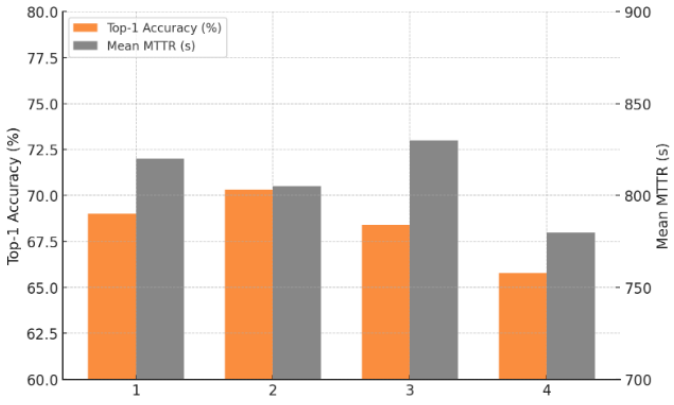

Figure 5. Comparison of Top-1 Accuracy and Average MTTR under Different Ablation Configurations

From Figure 5, it can be seen that removing semantic information or the evidence chain leads to a substantial drop in Top-1 accuracy, with the No Evidence Chain configuration producing the lowest accuracy (65%), despite its MTTR decreasing to approximately 775 seconds. In contrast, removing topology information has less impact on performance compared to removing temporal or semantic information.

Comparative Analysis

This section analyzes the experimental results presented in Tables 10 and 11, as well as Figures 4 and 5, as follows:
(1) Performance improvements over the baseline
From the data in Table 10, it can be observed that the proposed method improves the Top-1 root cause hit rate from 64.8% in the baseline model using only semantic aggregation to 71.8%, representing an increase of 7.0 percentage points; the Top-3 hit rate improves by 3.2 percentage points over the baseline. The average MTTR is reduced from 832 seconds to 792 seconds, a decrease of 40 seconds. Compared with the eventization scheme without topological information, the proposed method improves accuracy and consistency by 11.3 and 11.6 percentage points, respectively. This clearly demonstrates the critical role of incorporating topological information in enhancing root cause localization accuracy.
(2) Ablation study findings
The ablation results in Table 11 further validate the contribution of each functional module. When the content semantics module is removed (No Sem), the Top-1 hit rate drops by 3.4 percentage points, and MTTR increases significantly, indicating that content semantics play a decisive role in improving overall performance. Removing the structured evidence chain module (No Evidence Chain) results in a 6.1 percentage point decrease in the Top-1 hit rate, along with the largest fluctuation in MTTR, highlighting the central importance of the evidence chain in enhancing model reliability and interpretability. In contrast, the performance degradation caused by removing temporal information or topological information is relatively smaller, further confirming the higher design priority of the content semantics and evidence chain modules.
(3) Precision–recall performance
As shown in Figure 4, in the recall range above 0.8, the proposed method maintains a precision level of approximately 0.80, whereas the PR curves of various baseline models show a pronounced drop in this region. This indicates that the method can sustain stable performance under high-recall requirements while effectively reducing the false positive rate.

(4) Visualized impact of module removal
The ablation bar chart in Figure 5 visually illustrates the pronounced performance degradation under the No Sem and No Evidence Chain configurations, both showing the lowest accuracy and the highest MTTR. The extent of the performance drop is substantially greater than in cases where other modules are removed. The relatively moderate changes when temporal or topological information is removed are consistent with the quantitative results in Table 9.

5. Conclusion

This paper proposes an intelligent alarm root cause localization and field operation behavior analysis method for power-system operation and maintenance (O&M) that fuses content, temporal, and topological similarity with Hidden Markov Model (HMM) reasoning, enabling automated alarm flood compression, causal path inference, and quantitative assessment of operational responses. The main conclusions are as follows: (1) Efficient event-based aggregation is achieved by integrating multi-dimensional similarities, effectively mitigating cross-device alarm flooding in power-grid monitoring and protection systems while maintaining physical consistency of event boundaries, thus enabling fine-grained alarm management. (2) Interpretable root cause ranking is accomplished using a diagonal Gaussian HMM trained on normal operation data, delivering stable performance across diverse testing scenarios. (3) Structured multi-source evidence chain construction, combined with operational log analysis for response efficiency and behavioral consistency evaluation, provides verifiable decision support for grid field operations, improving the visualization and traceability of fault localization and resolution, and promoting the shift toward automated and quantitative power-system O&M processes.
Despite the positive results achieved in several aspects, there is still room for improvement in adaptability and online deployment. Currently, fixed thresholds and state numbers are manually set, limiting adaptability to multi-condition scenarios; multi-modal data fusion has not yet been incorporated; and real-time on-site validation mechanisms are lacking. Future work will focus on developing online threshold adjustment and model adaptation strategies, exploring lightweight inference optimization methods, and advancing multi-modal data fusion as well as real-time closed-loop platform construction, to further enhance the practicality and robustness of the system for power-grid online operation.

Acknowledgements
This research was funded by the Key Technologies for AI-Based Auxiliary Decision-Making in Relay Protection Operation and Watchkeeping (Task 3: Digital and Intelligent Technologies for Protection Operation and Watchkeeping to Enhance Power-Supply Reliability in Regional Grids) under Grant YNKJXM20240572.

References

[1] Yang T, Zhang J Y, Huang Z Q, et al. A survey on industrial control system security[J]. Journal of Computer Research and Development, 2022, 59(05): 1035-1053.
[2] Shang S Y, Li H J, Song C, et al. A survey of KPI anomaly detection based on machine learning in Internet service scenarios[J]. Journal of Computer Research and Development, 2025, 62(01): 207-231.
[3] Kargaran A H, Neshastegaran A, Izadi I, et al. Analytical derivation and comparison of alarm similarity measures[J]. IFAC-PapersOnLine, 2021, 54(3): 360-365.
[4] Li D, Cheng X. A first-out alarm detection method via association rule mining and correlation analysis[J]. Entropy, 2023, 26(1): 30.
[5] Khaleghy H, Izadi I. Detection of correlated alarms using graph embedding[C]//2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS). IEEE, 2021: 1-7.
[6] Medianovskyi K, Pietarinen A V. On explainable AI and abductive inference[J]. Philosophies, 2022, 7(2): 35.
[7] Farahnakian F, Nicolas F, Farahnakian F, et al. A comprehensive study of clustering-based techniques for detecting abnormal vessel behavior[J]. Remote Sensing, 2023, 15(6): 1477.
[8] Panza M A, Pota M, Esposito M. Anomaly detection methods for industrial applications: A comparative study[J]. Electronics, 2023, 12(18): 3971.
[9] Rahaman M H, Alinezhad H S, Chen T. Real-time classification and early warning of industrial alarm floods using modified TF-IDF methods[J]. Control Engineering Practice, 2025, 164: 106485.
[10] Jadidi Z, Pal S, Hussain M, et al. Correlation-based anomaly detection in industrial control systems[J]. Sensors, 2023, 23(3): 1561.
[11] Wang J, Sun Q, Zhou C. Insider threat detection based on deep clustering of multi-source behavioral events[J]. Applied Sciences, 2023, 13(24): 13021.
[12] Cui J, Zhang G, Chen Z, et al. Multi-homed abnormal behavior detection algorithm based on fuzzy particle swarm cluster in user and entity behavior analytics[J]. Scientific Reports, 2022, 12(1): 22349.
[13] Manca G, Dix M, Fay A. Clustering of similar historical alarm subsequences using alarm series and characteristic coactivations[C]//Proc. 32nd Int. Workshop on Principles of Diagnosis. 2021: 2-9.
[14] Liu Y, Garg S, Nie J, et al. Deep anomaly detection for time-series data in industrial IoT: A communication-efficient on-device federated learning approach[J]. IEEE Internet of Things Journal, 2020, 8(8): 6348-6358.
[15] Zhang S, Liu M J. Industrial image anomaly detection based on deviation-reduced coupled hypersphere[J]. Progress in Laser and Optoelectronics, 2025, 62(12): 176-183.
[16] Birihanu E, Lendák I. Explainable correlation-based anomaly detection for industrial control systems[J]. Frontiers in Artificial Intelligence, 2025, 7: 1508821.
[17] Abdelaty M, Doriguzzi-Corin R, Siracusa D. DAICS: A deep learning solution for anomaly detection in industrial control systems[J]. IEEE Transactions on Emerging Topics in Computing, 2021, 10(2): 1117-1129.
[18] Song X, Liu Q, Dong M, et al. Chemical process alarm root cause diagnosis method based on the combination of data-

knowledge-driven method and time retrospective reasoning[J]. ACS omega, 2022, 7(24): 20886-20905.
[19] Rao H R M, Zhou B, Brown K, et al. Alarm correlation analysis with applications to industrial alarm management[J]. Control Engineering Practice, 2024, 143: 105812.
[20] Vogel-Heuser B, Fay A, Rupprecht B, et al. Exploring challenges of alarm root-cause analysis across varying production process types[J]. at-Automatisierungstechnik, 2024, 72(4): 369-386.
[21] Venkidasalapathy J A, Kravaris C. Hidden Markov model based fault diagnoser using binary alarm signals with an analysis on distinguishability[J]. Computers & Chemical Engineering, 2022, 160: 107689.
[22] Alinezhad H S, Roohi M H, Chen T. A review of alarm root cause analysis in process industries: Common methods, recent research status and challenges[J]. Chemical Engineering Research and Design, 2022, 188: 846-860.
[23] Rahaman M H, Alinezhad H S, Chen T. Identification of Most Critical Alarms for Alarm Flood Reduction[J]. IFAC-PapersOnLine, 2024, 58(14): 835-840.
[24] Ariamuthu Venkidasalapathy J, Kravaris C. Hidden Markov model based approach for diagnosing cause of alarm signals[J]. AIChE Journal, 2021, 67(10): e17297.
[25] Li Y, Liu Y, Zhang J, et al. Automated analysis and assignment of maintenance work orders using natural language processing[J]. Automation in Construction, 2024, 165: 105501.
[26] Sundaram S, Zeid A. Technical language processing for Prognostics and Health Management: applying text similarity and topic modeling to maintenance work orders[J]. Journal of Intelligent Manufacturing, 2025, 36(3): 1637-1657.
[27] Ahn J, Park J Y, Lee S S, et al. SafeFac: Video-based smart safety monitoring for preventing industrial work accidents[J]. Expert Systems with Applications, 2023, 215: 119397.
[28] Pazho A D, Neff C, Noghre G A, et al. Ancilia: Scalable intelligent video surveillance for the artificial intelligence of things[J]. IEEE Internet of Things Journal. 2023, 10(17): 14940-14951.
[29] Shaheen B W, Németh I. Integration of maintenance management system functions with industry 4.0 technologies and features—A review[J]. Processes, 2022, 10(11): 2173.
[30] Shin H K, Lee W, Yun J H, et al. Two ICS security datasets and anomaly detection contest on the HIL-based augmented ICS testbed[C]//Proceedings of the 14th USENIX Workshop on Cyber Security Experimentation and Test (CSET '21). ACM, 2021: 36-40.