

Article

Confidence-Aware Topology Identification in Low-Voltage Distribution Networks: A Multi-Source Fusion Method Based on Weakly Supervised Learning

Siliang Liu , Can Deng *, Zenan Zheng, Ying Zhu, Hongxin Lu and Wenze Liu

Research Center of Smart Energy Technology, School of Electric Power, South China University of Technology, Guangzhou 510640, China; 201910102409@mail.scut.edu.cn (S.L.); 202421015376@mail.scut.edu.cn (Z.Z.); epzhuying@mail.scut.edu.cn (Y.Z.); 202420114212@mail.scut.edu.cn (H.L.); liuwzh@scut.edu.cn (W.L.)

* Correspondence: 202410183049@mail.scut.edu.cn

Abstract

The topology identification (TI) of low-voltage distribution networks (LVDNs) is the foundation for their intelligent operation and lean management. However, the existing identification methods may produce inconsistent results under measurement noise, missing data, and heterogeneous load behaviors. Without principled multiple method fusion and meter-level confidence quantification, the reliability of the identification results is questionable in the absence of ground-truth topology. To address these challenges, a confidence-aware TI (Ca-TI) method for the LVDN based on weakly supervised learning (WSL) and Dempster–Shafer (D-S) evidence theory is proposed, aiming to infer each meter’s latent topology connectivity label and quantify the meter-level confidence without ground truth by fusing different identification methods. Specifically, within the framework of data programming (DP) in WSL, different TI methods were modeled as labeling functions (LFs), and a weakly supervised label model (WSLM) was adopted to learn each method’s error pattern and each meter’s posterior responsibility; within the framework of D-S evidence theory, an uncertainty-aware basic probability assignment (BPA) was constructed from each meter’s posterior responsibility, with posterior uncertainty allocated to ignorance, and was further discounted according to the missing data rate; subsequently, a consensus-calibrated conflict-gated (CCCG)-enhanced D-S fusion rule was proposed to aggregate the TI results of multiple methods, producing the final TI decisions with meter-level confidence. Finally, the test was carried out in both simulated and actual low-voltage distribution transformer areas (LVDATAs), and the robustness of the proposed method under various measurement noise and missing data was tested. The results indicate that the proposed method can effectively integrate the performances of various TI methods, is not adversely affected by extreme bias from any single method, and provides the meter-level confidence for targeted on-site verification. Further, an engineering deployment scheme with cloud–edge collaboration is further discussed to support scalable implementation in utility environments.



Academic Editor: José Matas

Received: 28 January 2026

Revised: 23 February 2026

Accepted: 1 March 2026

Published: 18 March 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: low-voltage distribution network; topology identification; confidence-aware; weakly supervised label model; Dempster–Shafer evidence theory; cloud–edge collaboration

1. Introduction

The LVDN is a key link for power grid companies to provide high-quality customer service and enable intelligent electricity consumption. TI forms the basis of the intelligent operation and lean management of LVDNs, and it provides important support for

LVDTA planning, loss management, outage management, power quality mitigation, and value-added customer services [1,2]. In recent years, LVDN TI has attracted increasing attention from researchers.

LVDN TI mainly addresses the problem of identifying node connectivity, which generally includes three levels: consumer–transformer connectivity identification (C-TCI), consumer–phase connectivity identification (C-PCI), and consumer–branch connectivity identification (C-BCI). Traditional identification methods include manual inspection and momentary outage-based identification. These methods are time-consuming and labor-intensive, have poor operability, and their accuracy is difficult to guarantee. With the development of the Power Internet of Things (PIoT), digital and information technologies have provided new approaches for LVDN TI [3,4], mainly including the signal injection-based method [5–8], broadband power line carrier-based (BPLC-based) method [9–11], and data driven-based method [12–20]: (1) The signal injection-based method refers to installing online monitoring terminals/modules with signal transmitting and receiving capabilities on the transformer side and/or the consumer side and automatically sorting out the LVDN topology through master–slave signal identification. The identification signals mainly include pulse currents [5], carrier communication signals [6], characteristic currents [7], and fundamental-frequency zero-crossing distortion signals [8]. (2) The BPLC-based method refers to TI based on the new interoperable BPLC technology for low-voltage power lines [9]. By leveraging the high-precision synchronized timer in the module, the BPLC-based method accurately captures the zero-crossing offset of the voltage sinusoidal waveform to identify the LVDN topology [10,11]. (3) The data driven-based method refers to mining intrinsic features from massive electricity consumption data under advanced metering infrastructure (AMI) using big data analytics and artificial intelligence and then inferring the LVDN topology. Depending on the type of metering data used, the data driven-based method can be further categorized into energy/current/power fitting [12–14], voltage time-series correlation analysis [15–17], and joint voltage correlation analysis with current/power fitting [18–20]. Beyond TI, data-driven learning on high-dimensional time-series data has also been widely adopted in power applications; for example, deep feature selection for LSTM-based PV power forecasting highlights the importance of identifying informative variables to improve robustness and generalization under complex and noisy operating conditions [21].

Although LVDN TI has achieved substantial research progress, most studies concentrate on improving a single class of algorithms and enhancing the identification accuracy. However, different types of identification methods differ in their characteristics and application scenarios and using only one type of method can hardly satisfy practical application requirements. Enabling complementary integration and coordinated interaction among different methods across scenarios is key to practical LVDN TI. From a system perspective, the cloud–edge collaborative architecture of the PIoT provides a framework for massive power data applications [4]. Similar engineering considerations have been discussed in other grid-facing power data applications, such as comparative studies of grid-tied PV maximum power point tracking (MPPT) methods, where reliability-oriented performance trade-offs under practical operating constraints are explicitly analyzed [22]. Nevertheless, application architectures and data interaction mechanisms tailored to LVDN TI still require further investigation.

From a methodological perspective, multi-source fusion for LVDN TI has been explored but is still limited in terms of reliability-aware fusion and confidence quantification. For example, Reference [23] proposed two fusion mechanisms, KMeans labeling and voting decision making, and then developed a consumer–branch connectivity identification method by fusing seven clustering algorithms. However, the fusion mechanism is

essentially the frequency counting of identification results, which is easily affected by individual methods and has difficulty handling contradictions and conflicts among multiple identification results.

Under bad-data conditions, such as missing data, measurement noise, and blurred features, single-method outputs can become unstable, and naive fusion may still suffer from uncertainty and conflict if source reliability is unknown. This motivates the need for Ca-TI, jointly producing TI decisions and quantifying their reliability. Several studies have attempted confidence evaluation. Reference [24] used data quality scores and the corresponding TI accuracy as training samples to build a mapping model between the data quality scores and the accuracy range of data driven-based identification and then output the confidence of the TI results based on the data quality score of a sample. Reference [25] constructed a latent tree model for LVDN TI and inferred the most probable topology using a Bayesian network. While probabilistic graphical models such as Bayesian networks provide a principled way to perform posterior inference, their practical use in brownfield LVDNs may be affected by the availability of reliable priors and the fidelity of model assumptions. In particular, topology-related priors such as structure priors and parameter priors may be unavailable or uncertain; meanwhile, the lack of ground-truth topology makes it difficult to systematically validate or calibrate these priors and structural assumptions. Reference [14] proposed a consumer-branch connectivity identification method based on hierarchical stepwise iterative regression and evaluated the confidence of results using the significance level of a *t*-test and performed correction. However, the existing confidence measures are often provided at an aggregate LVDTA level or are tightly coupled with specific identification methods, which limits their ability to deliver unified, meter-level confidence that can directly guide on-site topology verification. Moreover, ground-truth topology labels are frequently missing and expensive to obtain in practical engineering, making meter-level confidence computation particularly challenging without supervision. In recent years, WSL has emerged, whose core idea is to train models using incomplete, inexact, or inaccurate labeled data to reduce dependence on fully labeled data [26]. And it has been widely applied in AI fields such as image classification and semantic segmentation. WSL can be divided into three categories: incomplete supervision, inexact supervision, and inaccurate supervision. Among them, inaccurate supervision aims to correct models by identifying potentially misclassified samples. Data programming is a typical approach in inaccurate supervision; its core idea is to generate noisy labels through multiple labeling functions, train a label model, and denoise/aggregate high-quality labels without ground truth. By analogy to LVDN TI, noisy labels correspond to single-source identification results, and the denoised high-quality labels correspond to the latent true labels under multi-source consensus. Therefore, data-programming methods in WSL can be used to obtain probabilistic estimates of the latent ground-truth topology and to express the confidence of identification results in a probabilistic manner.

Based on the above analysis, this paper proposes a Ca-TI method for LVDNs based on WSL and multi-source fusion. The main contributions are as follows:

- (1) We develop a WSLM that models different TI methods as LFs under the inaccurate-supervision setting of DP, and we infer each method's error pattern and each meter's posterior responsibility over TI hypotheses without ground-truth topology.
- (2) Based on each meter's posterior responsibility, we develop an uncertainty-aware BPA construction method that jointly incorporates posterior uncertainty and missing data rates to allocate an adaptive uncertainty mass, yielding single-source confidence prior to D-S fusion.
- (3) We propose a CCCG-enhanced D-S fusion rule to combine multi-source evidence under contradictions, yielding the final TI decisions with meter-level confidence.

The method employs a conflict-gating mechanism to refine the classical D-S fusion rule, and it utilizes multi-source identification methods to optimize the fusion rule hyperparameters based on consensus consistency, which is robust to inter-method conflicts and data quality degradation.

- (4) Computational analysis is conducted using data samples from simulated and actual LVDTAs, demonstrating that the proposed method can track the best-performing source and avoid performance collapse caused by any single source failure under measurement noise and missing data. And we further discuss a cloud–edge collaborative engineering deployment scheme to support scalable utility implementation.

The rest of this paper is organized as follows: Section 2 offers the problem definition and preliminaries, including problem definitions of the Ca-TI, WSL and DP preliminaries and applicability and D-S evidence theory preliminaries and applicability. Section 3 proposes the Ca-TI method in the LVDN, including the unified modeling for meter-level TI posterior responsibility inference based on the WSLM and the Ca-TI based on CCCG-enhanced D-S evidence fusion. Section 4 conducts various case studies to validate the effectiveness of the proposed method. In Section 5, the engineering deployment scheme with cloud–edge collaboration is further discussed. Finally, Section 6 provides the conclusions and future work.

2. Problem Definition and Preliminaries

2.1. Problem Definition of Ca-TI

TI aims to determine each meter's topology connectivity label in the LVDN, such as the C-TC, C-PC and C-BC. In practice, the ground-truth topology is often missing or unreliable. Moreover, measurement noise, missing data, and heterogeneous load behaviors may cause different identification methods to output inconsistent TI results. Therefore, the task of Ca-TI is not only to produce a TI decision but also to quantify how reliable this decision is. Let the TI hypotheses set be $\Theta = \{h_1, h_2, \dots, h_j, \dots, h_N\}$. For meter $i = \{1, 2, \dots, M\}$, its unknown true topology connectivity label is denoted by $Z_i = \{1, 2, \dots, j, \dots, N\}$, satisfying $Z_i = j \Leftrightarrow h_j$ is true, and N is the number of TI hypotheses. The formulation of Ca-TI can be described as follows:

$$\hat{y}_i = \underset{j}{\operatorname{argmax}}(P_{i,j}), j = \{1, 2, \dots, N\} \quad (1)$$

$$\hat{r}_i = \operatorname{conf}_i, \hat{r}_i \in [0, 1] \quad (2)$$

where $\hat{y}_i$ is the TI decision of meter i ; $P_{i,j}$ is the probability of TI as hypothesis j for meter i , and it is described by the pignistic probability in the paper; $\hat{r}_i$ is the reliability of the TI decision for meter i , and the meter-level confidence conf_i is proposed to quantify it. The meter-level confidence can be used to rank meters for on-site verification operationally: meters with lower confidence should be prioritized for field inspection or further data collection.

2.2. WSL and DP Preliminaries and Applicability

We consider S TI methods as information sources. For source s , the TI result for meter i is written as $y_i^{(s)} \in \{1, 2, \dots, N\} \cup \{\perp\}$, where $\perp$ denotes abstention (no output). In the Ca-TI, the application of WSL arises in two complementary forms:

- (1) Inaccurate supervision (noisy labels): $y_i^{(s)} \in \{1, 2, \dots, N\}$ may be incorrect due to measurement noise, missing data, heterogeneous load behaviors or algorithmic bias.
- (2) Incomplete supervision (missing labels): $y_i^{(s)} \in \{\perp\}$ occurs when outputs are unavailable because of communication loss, missing measurements, or fusion infeasibility.

DP, as a representative WSL approach under inaccurate supervision, is well suited for Ca-TI because: (i) ground-truth topology labels are often missing, unreliable or expensive to obtain in LVDNs; (ii) different TI methods provide complementary but imperfect outputs whose agreement/disagreement patterns enable learning source reliability without labels; (iii) incomplete supervision is naturally represented by abstention $\{\perp\}$, so missing outputs are not forced into incorrect labels. Therefore, DP is adopted to learn source reliability from agreement/disagreement patterns without ground truth while explicitly handling incomplete supervision by allowing abstentions.

Under the DP framework, each meter is treated as an instance, and its unknown true topology connectivity (Z_i) is modeled as a latent label. Each TI method is treated as an LF that provides a noisy label ($y_i^{(s)}$), which is the output of the TI method s for meter i . When measurements are missing or a method fails to output, the LF abstains, which directly corresponds to incomplete supervision caused by communication loss or missing readings. A generative label model parameterizes the reliability of each method via source-specific error patterns based on confusion matrices and estimates these parameters from unlabeled data through the expectation maximization (EM) algorithm by exploiting agreement/disagreement statistics among TI methods. The label model outputs meter-level posterior responsibilities, which serve as denoised probabilistic labels and form the basis for the subsequent D-S evidence fusion. The mapping between the DP components and Ca-TI problem is summarized in Table 1.

Table 1. Mapping between DP components and Ca-TI.

DP Component	Mapping in Ca-TI
Instance	Meter
Latent label	Meter i 's unknown true topology connectivity label (Z_i)
LFs	Different TI methods are treated as LFs.
Out of LFs	Result of TI method s for meter i $y_i^{(s)}$
Noisy labels	Incorrect TI results caused by meter errors, measurement noise, heterogeneous loads, or algorithmic bias under inaccurate supervision, including abstention (no output) because of communication loss, missing measurements, or fusion infeasibility.
Generative label model	WSLM that parameterizes each TI method's reliability based on confusion matrix.
Source reliability parameters	Source-specific error patterns by confusion matrices capturing how likely a TI method outputs each label given the latent topology connectivity.
Unsupervised training	EM-based estimation of source reliability parameters using agreement/disagreement statistics among TI methods without ground-truth topology.
Output (denoised labels)	Meter-level posterior responsibilities used for subsequent D-S evidence fusion.

2.3. D-S Evidence Theory Preliminaries and Applicability

Due to electromagnetic interference, communication anomalies and other factors, LVDNs widely suffer from data quality problems, such as missing data and measurement noise. Using only one single type of method for identification is uncertain and inaccurate, and the reliability of identification results is difficult to guarantee. Multi-source information fusion aims to integrate results from multiple sources to obtain more objective and accurate estimates [27]. Common fusion methods include voting systems, Bayesian inference, D-S evidence theory and so on. Among them, D-S evidence theory is suitable for fusion

scenarios where multi-source information contains uncertainty, and it does not rely on prior data [28], which matches the problem considered in this paper.

In D-S evidence theory, hypotheses are firstly defined via a framework of discernment, then the BPA is introduced to quantify the belief mass for each hypothesis proposition, and fusion rules are finally used to fuse the belief masses provided by multiple evidence sources. The fusion framework based on D-S evidence theory for confidence-aware LVDN TI is shown as follows:

$$\Theta_{\text{TI}}^i = \{h_{\text{TI}}^y, h_{\text{TI}}^n\} \quad (3)$$

$$\Theta_{\text{PI}}^i = \{h_{\text{PI}}^A, h_{\text{PI}}^B, h_{\text{PI}}^C\} \quad (4)$$

$$\Theta_{\text{BI}}^i = \{h_{\text{BI}}^1, h_{\text{BI}}^2, \dots, h_{\text{BI}}^l, \dots, h_{\text{BI}}^L\} \quad (5)$$

where Θ_{TI}^i , Θ_{PI}^i and Θ_{BI}^i denote the identification frameworks of C-TC, C-PC, and C-BC for meter i , respectively; M is the total number of meters in the LVDTA; h_{TI}^y and h_{TI}^n represent the hypotheses for identifying the C-TC of meters belonging to the target LVDTA and not belonging to the target LVDTA, respectively; h_{PI}^A , h_{PI}^B and h_{PI}^C represent the hypotheses for identifying the C-PC of meters, which are phases A, B, and C, respectively; h_{BI}^l represents the hypothesis for identifying the C-BC of meters belonging to the LV branch (l); L is the total number of LV branches. It is noted that the method proposed in this paper performs evidence fusion analysis on the TI results for each meter node, and:

$$\begin{aligned} \Theta_{\text{TI}}^1 &= \Theta_{\text{TI}}^2 = \dots = \Theta_{\text{TI}}^i = \dots = \Theta_{\text{TI}}^M \\ \Theta_{\text{PI}}^1 &= \Theta_{\text{PI}}^2 = \dots = \Theta_{\text{PI}}^i = \dots = \Theta_{\text{PI}}^M \\ \Theta_{\text{BI}}^1 &= \Theta_{\text{BI}}^2 = \dots = \Theta_{\text{BI}}^i = \dots = \Theta_{\text{BI}}^M \end{aligned} \quad (6)$$

To facilitate subsequent analysis, Equations (3)–(5) are uniformly represented as:

$$\Theta = \{h_1, h_2, \dots, h_j, \dots, h_N\} \quad (7)$$

where h_j is the j th hypothesis of Θ , and the power set of Θ can be defined as:

$$2^\Theta = \{\emptyset, \{h_1\}, \{h_2\}, \dots, \{h_N\}, \{h_1, h_2\}, \dots, \Theta\} \quad (8)$$

where $\emptyset$ denotes the empty set.

BPA is a function mapping from 2^Θ to $[0, 1]$, satisfying:

$$m : 2^\Theta \rightarrow [0, 1] \quad (9)$$

$$m(\emptyset) = 0 \text{ and } \sum_{A \subseteq \Theta} m(A) = 1 \quad (10)$$

It can be seen that in the power set 2^Θ , each hypothesis proposition should be assigned a BPA within $[0, 1]$, and the sum of the BPAs over all propositions equals 1.

When there are S evidence sources that provide S pieces of BPA for the same hypothesis proposition, namely, $m_1, m_2, \dots, m_S$, and they are mutually independent, the classical D-S fusion rule can be used to compute the fused BPA, as shown in Equations (11) and (12):

$$m(A) = \begin{cases} 0, & A = \emptyset \\ \frac{1}{1-k} \sum_{A_1 \cap A_2 \cap \dots \cap A_S = A} m_1(A_1) m_2(A_2) \dots m_S(A_S), & A \neq \emptyset \end{cases} \quad (11)$$

$$k = \sum_{A_1 \cap A_2 \cap \dots \cap A_S = \emptyset} m_1(A_1) m_2(A_2) \dots m_S(A_S) \quad (12)$$

where k reflects the degree of conflict among evidence sources, and it ranges from 0 to 1, with a larger k value indicating a higher degree of conflict. Because data quality problems in LVDNs are significant, evidence sources may exhibit extreme bias on TI hypotheses, which can easily induce evidence conflict. A CCG-enhanced D-S fusion method will be introduced to mitigate evidence source conflicts in Sections 3.2.2 and 3.2.3.

Based on the fused BPA from multiple evidence sources, the final TI result of each meter can be decided, and the confidence score of the fused result can be provided; this will be discussed in Section 3.2.4.

3. Ca-TI Method in LVDN

3.1. Unified Modeling for Meter-Level TI Posterior Responsibility Inference Based on WSLM

Accurate BPA computation is key for D-S evidence fusion. Typical methods include frequency calibration [29], similarity-based computation [30], and self-learning approaches [31,32]. Frequency calibration constructs a mass function through sample frequency statistics and calculates the BPA. Similarity calculation determines the similarity between samples and hypotheses to compute the BPA. Both methods are computationally straightforward and offer strong physical interpretations, but they are highly sensitive to the quality of the sampled data. Self-learning incorporates machine learning and deep learning methods to construct training models from samples to BPA, enabling the deep exploration of complex trust patterns toward hypothetical propositions across diverse heterogeneous evidence sources. However, this approach involves high computational complexity and relies heavily on training with large volumes of labeled samples. In practical applications of LVDN TI, the samples of ground-truth topology are often scarce and costly to acquire. The WSLM can infer posterior probability distributions by analyzing consensus error patterns across multiple weakly supervised sources when true labels are scarce or absent, making it well suited for LVDN TI requirements [33,34].

The original dataset (D) is split into T time windows as subsets (D') using a sliding-window strategy. S TI methods are selected as evidence sources. Evidence source s performs TI based on the data of time window t and outputs the TI result of meter i as:

$$y_{i,t}^{(s)} \in \{1, 2, \dots, N\} \cup \{\perp\} \quad (13)$$

where $\perp$ denotes missing output; in other words, the evidence source cannot produce an identification result based on the data samples within the current time window.

Define the latent variables of the true topology connectivity, and construct a confusion matrix to describe the overall error patterns in each evidence source, as shown in the following equations:

$$Z_i \in \{1, 2, \dots, N\}, Z_i = j \Leftrightarrow h_j \text{ is true} \quad (14)$$

$$\Pi_{j,k}^{(s)} = \mathbb{P}(Y^{(s)} = k | Z = j), j, k \in \{1, \dots, N\} \quad (15)$$

where Z_i is the latent variable representing the true topology connectivity of meter i ; $\Pi_{j,k}^{(s)}$ is the element in row j , column k of the confusion matrix for evidence source s . The confusion matrix represents the probability that evidence source s outputs a topology label (k) when the true topology connectivity is j . In practical applications, the true topology connectivity is unknown. Therefore, the confusion matrix is solved using the EM algorithm, as shown in the following equations:

E-step:

$$\gamma_{i,j}^{(x)} \triangleq \mathbb{P}(Z_i = j \mid \{y_{i,t}^{(s)}\}) \propto \pi_{i,j} \prod_{s=1}^S \prod_{t \in \mathcal{O}_i^{(s)}} \left(\Pi_{j,y_{i,t}^{(s)}}^{(s)(x-1)} \right)^{w_{i,t}^{(s)}}, \gamma_{i,j}^{(x)} \leftarrow \gamma_{i,j}^{(x)} / \sum_{k=1}^N \gamma_{i,k}^{(x)} \quad (16)$$

M-step:

$$\Pi_{j,k}^{(s)(x)} \leftarrow \sum_i \gamma_{i,j}^{(x)} \sum_{t \in \mathcal{O}_i^{(s)}} w_{i,t}^{(s)} \mathbf{1}[y_{i,t}^{(s)} = k] / \sum_i \gamma_{i,j}^{(x)} \sum_{t \in \mathcal{O}_i^{(s)}} w_{i,t}^{(s)} \quad (17)$$

where $\pi_{i,j} = \mathbb{P}(Z_i = j)$ is the prior distribution of Z_i ; $\mathcal{O}_i^{(s)} = \{t : y_{i,t}^{(s)} \neq \perp\}$ denotes the set of identification labels for evidence source s under non-outcome-missing time windows for meter i ; $w_{i,t}^{(s)} \in [0, 1]$ represents the time window weight, used to apply differential weighting to time windows with missing data; $w_{i,t}^{(s)} = (1 - d_{i,t}) / W_{sw}$, where $d_{i,t}$ is the number of missing data points for meter i at time window t , and W_{sw} is the width of the time window; $\Pi_{j,k}^{(s)}$ is the element in row j , column k of the confusion matrix for evidence source s at iteration x ; $\gamma_{i,j}^{(x)}$ is the posterior responsibility at iteration x ; $\mathbf{1}[\bullet]$ denotes the indicator function—if the condition holds, take 1; otherwise, take 0.

In the absence of reliable topology priors, the confusion matrix is initialized as a diagonal-dominant matrix:

$$\Pi_{j,k}^{(s)(0)} = \begin{cases} \eta, & j = k \\ (1 - \eta) / (N - 1), & j \neq k \end{cases} \quad (18)$$

where η is the accuracy initialization of the evidence source. Similarly, we initialize the prior distribution as a uniform distribution ($\pi_{i,j} = 1/N$). Given these initial parameters, the initial posterior responsibilities can be obtained by the E-step. And the EM algorithm achieves iterative convergence by setting a threshold for the relative gain in log-likelihood and a maximum iteration limit, as shown in the following equation:

$$|\mathcal{L}^{(x)} - \mathcal{L}^{(x-1)}| / |\mathcal{L}^{(x-1)}| + \varepsilon < \tau \text{ and } x \geq \text{iter}_{\max} \quad (19)$$

where $\mathcal{L}^{(x)}$ is the log-likelihood of the observed data at iteration x ; τ is the relative gain threshold; $\text{iter}_{\max}$ is the maximum number of iterations; ε is the numerical safeguard term.

According to the confusion matrix of each evidence source and the posterior responsibility based on cross-source consensus, the posterior responsibilities of each evidence source can be computed as follows:

$$\gamma_{i,j}^{(s)} \propto \pi_{i,j} \prod_{t \in \mathcal{O}_i^{(s)}} \left(\Pi_{j,y_{i,t}^{(s)}}^{(s)} \right)^{w_{i,t}^{(s)}}, \gamma_{i,j}^{(s)} \leftarrow \gamma_{i,j}^{(s)} / \sum_{k=1}^N \gamma_{i,k}^{(s)} \quad (20)$$

where $\gamma_{i,j}^{(s)}$ is the posterior responsibility of evidence source s .

The proposed meter-level TI posterior responsibility inference based on WSLM is summarized in Algorithm 1.

Algorithm 1: Meter-level TI Posterior Responsibility Inference Based on WSLM

Input: number of meters M ; number of hypotheses N ; width of time window W_{sw} ; number of time window T ; TI results of S evidence sources $y_{i,t}^{(s)}$; accuracy initialization of evidence source η ; maximum number of iterations $iter_{max}$; relative gain threshold τ ; numerical safeguard term ε .

Output: posterior responsibility of evidence source $\gamma_{i,j}^{(s)}$.

```

1 Initialize  $\Pi_{j,k}^{(s)(0)}$  and  $\pi_{i,j}$ .
2 for  $x = 1$  to  $iter_{max}$  do
3   for  $i = 1$  to  $M$  do
4      $\gamma_{i,j}^{(x)} \leftarrow \pi_{i,j} \prod_{s=1}^S \prod_{t \in \mathcal{O}_i^{(s)}} \left( \Pi_{j,y_{i,t}^{(s)}}^{(s)(x-1)} \right)^{w_{i,t}^{(s)}}$ ,  $\gamma_{i,j}^{(x)} \leftarrow \gamma_{i,j}^{(x)} / \sum_{k=1}^N \gamma_{i,k}^{(x)}$ 
5   for  $s = 1$  to  $S$  do
6     for  $j = 1$  to  $N$  do
7       for  $k = 1$  to  $N$  do
8          $\Pi_{j,k}^{(s)(x)} \leftarrow \sum_i \gamma_{i,j}^{(x)} \sum_{t \in \mathcal{O}_i^{(s)}} w_{i,t}^{(s)} \mathbf{1}[y_{i,t}^{(s)} = k] / \sum_i \gamma_{i,j}^{(x)} \sum_{t \in \mathcal{O}_i^{(s)}} w_{i,t}^{(s)}$ 
9       if  $|\mathcal{L}^{(x)} - \mathcal{L}^{(x-1)}| / |\mathcal{L}^{(x-1)}| + \varepsilon < \tau$ 
10        break
11   for  $s = 1$  to  $S$  do
12     for  $i = 1$  to  $M$  do
13        $\gamma_{i,j}^{(s)} \leftarrow \pi_{i,j} \prod_{t \in \mathcal{O}_i^{(s)}} \left( \Pi_{j,y_{i,t}^{(s)}}^{(s)} \right)^{w_{i,t}^{(s)}}$ ,  $\gamma_{i,j}^{(s)} \leftarrow \gamma_{i,j}^{(s)} / \sum_{k=1}^N \gamma_{i,k}^{(s)}$ 

```

3.2. Ca-TI Method Based on CCCG Enhanced D-S Evidence Fusion

During evidence fusion, conflicts among evidence sources are unavoidable. The classical Dempster rule filters the conflict mass through normalization and reallocates the remaining mass proportionally to competing hypotheses. As a result, when high conflict occurs, the fused outcome tends to be aggressive and may yield counter-intuitive results. By contrast, the classical Yager rule directly assigns the conflict mass to the uncertainty focal element, thereby avoiding forced decisions under high conflict. Consequently, the fusion outcome becomes more conservative and robust [35,36]. To balance information gain under low conflict and robustness under high conflict, this section exploits the complementarity between the Dempster and Yager rules and proposes an improved D-S fusion rule based on a conflict-gating mechanism.

3.2.1. Uncertainty-Aware BPA Construction

For D-S fusion, the meter-level posterior responsibilities are mapped to BPAs via uncertainty-aware mass assignment, and the resulting BPAs are further discounted considering the missing data rate.

In the uncertainty-aware mass assignment, the meter-level posterior responsibility is mapped to the BPA, with the posterior uncertainty controlling the reliability factor that trades off hypothesis masses against ignorance mass. The posterior uncertainty of evidence source s can be computed as follows:

$$\text{conc}_i^{(s)} = 1 - H(\gamma_i^{(s)}) / \log N, \quad H(\gamma_i^{(s)}) = - \sum_{j=1}^N \gamma_{i,j}^{(s)} \log(\gamma_{i,j}^{(s)}) \quad (21)$$

where $\text{conc}_i^{(s)}$ and $\gamma_i^{(s)}$ represent the posterior uncertainty and the posterior responsibility of the TI result for meter i from evidence source s ; $H(\cdot)$ is the Shannon entropy function, and it is used to calculate the uncertainty of the posterior responsibility. The smaller the entropy value, the more concentrated the posterior responsibility becomes, indicating that the evidence source is more certain about the TI results for the meters. Conversely, it indicates that the evidence source finds it difficult to determine the TI results for the meters.

Due to missing data issues in LVDNs, time window data samples may be incomplete, potentially preventing evidence sources from performing TI. This may result in identification outcomes failing to meet subsequent evidence fusion requirements. Therefore, when calculating the BPA, the impact of missing data should be incorporated, as shown in the following equation:

$$\delta_i^{(s)} = 1 - T_i^{(s)} / T, \quad T_i^{(s)} = \sum_{t=1}^T w_{i,t}^{(s)} \mathbf{1}[y_{i,t}^{(s)} \neq \perp] \quad (22)$$

where $\delta_i^{(s)}$ is the missing data rate in time windows of the TI for meter i by evidence source s ; $T_i^{(s)}$ is the number of effective time windows for evidence source s to perform TI on meter i . According to the posterior uncertainty of the evidence source and the impact of the data quality, the BPA can be calculated as follows:

$$m_i^{(s)}(\{h_j\}) = r_i^{(s)} \gamma_{ij}^{(s)}, \quad m_i^{(s)}(\Theta_i) = 1 - r_i^{(s)} \quad (23)$$

$$r_i^{(s)} = G\left((1 - \delta_i^{(s)}) \cdot \text{conc}_i^{(s)}, 0, 1\right) \quad (24)$$

where $m_i^{(s)}(\{h_j\})$ denotes the BPA assigned by evidence source s to hypothesis $\{h_j\}$ for meter i ; $m_i^{(s)}(\Theta_i)$ denotes the BPA assigned by evidence source s to uncertainty for meter i , which is used to avoid overconfident identification by the evidence source when the posterior responsibility is not concentrated or when missing data are severe. $r_i^{(s)}$ is the reliability factor of evidence source s for the TI result of meter i ; $G(e, 0, 1)$ denotes the truncation function, which can restrict e to the interval $[0, 1]$.

3.2.2. Improved D-S Evidence Fusion Rule Based on Conflict Gating

According to the BPA computation in Section 3.2.1, the conflict mass (C_i) among evidence sources for the TI result of meter i can be computed as follows:

$$C_i = 1 - \prod_{s=1}^S m_i^{(s)}(\Theta_i) - \sum_{j=1}^N \left[\prod_{s=1}^S \left(m_i^{(s)}(\{h_j\}) + m_i^{(s)}(\Theta_i) \right) - \prod_{s=1}^S m_i^{(s)}(\Theta_i) \right] \quad (25)$$

Then, the Dempster rule and the Yager rule are respectively used to obtain preliminary fusion BPAs for the TI result of meter i , as follows:

$$m_{D,i}(\{h_j\}) = \left[\prod_{s=1}^S \left(m_i^{(s)}(\{h_j\}) + m_i^{(s)}(\Theta_i) \right) - \prod_{s=1}^S m_i^{(s)}(\Theta_i) \right] / \max(1 - C_i, \varepsilon) \quad (26)$$

$$m_{D,i}(\Theta_i) = \prod_{s=1}^S m_i^{(s)}(\Theta_i) / \max(1 - C_i, \varepsilon)$$

$$m_{Y,i}(\{h_j\}) = \prod_{s=1}^S \left(m_i^{(s)}(\{h_j\}) + m_i^{(s)}(\Theta_i) \right) - \prod_{s=1}^S m_i^{(s)}(\Theta_i) \quad (27)$$

$$m_{Y,i}(\Theta_i) = \prod_{s=1}^S m_i^{(s)}(\Theta_i) + C_i$$

where $m_{D,i}(\bullet)$ and $m_{Y,i}(\bullet)$ represent the preliminary fusion BPAs for the TI result of meter i based on the Dempster rule and Yager rule, respectively.

Further, a dynamic gating function is defined to construct the improved D-S fusion rule as follows:

$$g_i = \sigma(a(C_i - C_0)) \quad (28)$$

$$m_i(A) = (1 - g_i)m_{D,i}(A) + g_im_{Y,i}(A), \quad A \in \{\{h_1\}, \dots, \{h_N\}, \Theta_i\} \quad (29)$$

where $m_i(A)$ represents the final fusion BPA for the TI of meter i based on the improved D-S fusion rule; g_i denotes the dynamic gating coefficient for fusing the TI result of meter i ; $\sigma(\bullet)$ is a sigmoid function satisfying $\sigma(x) = 1/(1 + e^{-x})$; a and C_0 represent the gating slope and gating threshold, respectively, both of which are hyperparameters of the dynamic gating function.

3.2.3. Computation of Dynamic Gating Hyperparameters

As shown in Section 3.2.2, the hyperparameters a and C_0 must be specified in advance, which introduces subjectivity. Therefore, this section establishes a hyperparameter calibration model based on consensus consistency among evidence sources. The model uses the Jousselme distance to measure the similarity between two sets of BPAs [36]. Specifically, the average Jousselme distance between the fused BPA and the BPAs of each evidence source is minimized as the objective function, with the lower and upper bounds of the hyperparameters as constraints. The mathematical formulation of the model is given as follows:

$$\begin{aligned} \min f &= \frac{1}{M} \sum_{i=1}^M \frac{1}{S} \sum_{s=1}^S d_J(\mathbf{M}_i(a, C_0), \mathbf{M}_i^{(s)}) \\ \text{s.t.} &\begin{cases} a_{\min} \leq a \leq a_{\max} \\ 0 \leq C_0 \leq 1 \end{cases} \end{aligned} \quad (30)$$

where $d_J(\bullet, \bullet)$ is the Jousselme distance metric function; $\mathbf{M}_i(a, C_0)$ is the fusion BPA vector for the TI results of meter i satisfying $\mathbf{M}_i(a, C_0) = [m_i(\{h_1\}), \dots, m_i(\{h_N\}), m_i(\Theta_i)]$, and it is a function of the hyperparameters a and C_0 . $\mathbf{M}_i^{(s)}$ is the BPA vector of evidence source s for the TI result of meter i , satisfying $\mathbf{M}_i^{(s)} = [m_i^{(s)}(\{h_1\}), \dots, m_i^{(s)}(\{h_N\}), m_i^{(s)}(\Theta_i)]$. $a_{\max}$ and $a_{\min}$ represent the upper and lower limits of the gating slope (a), respectively.

3.2.4. Fusion Result Decision and Meter-Level Confidence Computation

According to the fusion BPA computed in Section 3.2.3, the pignistic function is used to compute the probability of each hypothesis. The final TI result of each meter is decided according to the maximum pignistic probability, as shown in Equation (31). The meter-level confidence is defined as the product of the informativeness and consistency, as shown in Equation (32). The fused uncertainty mass ($m_i(\Theta_i)$) quantifies residual uncertainty due to insufficient evidence, while the conflict mass (C_i) quantifies inter-source contradiction. Hence, the confidence decreases when the evidence is either insufficient (high $m_i(\Theta_i)$) or contradictory (high C_i), and it becomes high only when the evidence is both informative and consistent:

$$\hat{y}_i = \arg \max_j [m_i(\{h_j\}) + m_i(\Theta_i)/N] \quad (31)$$

$$\text{conf}_i = 1 - [\lambda m_i(\Theta_i) + (1 - \lambda)C_i] \quad (32)$$

where $\hat{y}_i$ represents the final TI decision for meter i ; conf_i represents the meter-level confidence of the final TI result for meter i . λ is the weight coefficient.

The proposed Ca-TI method based on CCCG-DS evidence fusion is summarized in Algorithm 2.

Algorithm 2: Ca-TI Method Based on CCCG-DS Evidence Fusion

Input: number of hypotheses N ; posterior responsibility of evidence source $\gamma_{i,j}^{(s)}$; number of time window T ; time window weight $w_{i,t}^{(s)}$; TI results of S evidence sources $y_{i,t}^{(s)}$; TI results of S evidence sources $y_{i,t}^{(s)}$; numerical safeguard term ε ; upper and lower limits of the gating slope $a_{\max}$ and $a_{\min}$; weight coefficient λ .

Output: Final TI decision $\hat{y}_i$ and meter-level confidence $conf_i$.

```

1  for  $s = 1$  to  $S$  do
2    for  $i = 1$  to  $M$  do
3       $\text{conc}_i^{(s)} \leftarrow 1 - H(\gamma_i^{(s)}) / \log N$ ,  $H(\gamma_i^{(s)}) \leftarrow -\sum_{j=1}^N \gamma_{i,j}^{(s)} \log(\gamma_{i,j}^{(s)})$ 
4       $\delta_i^{(s)} \leftarrow 1 - T_i^{(s)} / T$ ,  $T_i^{(s)} \leftarrow \sum_{t=1}^T w_{i,t}^{(s)} \mathbb{1}[y_{i,t}^{(s)} \neq \perp]$ 
5       $r_i^{(s)} \leftarrow G((1 - \delta_i^{(s)}) \cdot \text{conc}_i^{(s)}, 0, 1)$ 
6      for  $j = 1$  to  $N$  do
7         $m_i^{(s)}(\{h_j\}) \leftarrow r_i^{(s)} \gamma_{i,j}^{(s)}$ ,  $m_i^{(s)}(\Theta_i) \leftarrow 1 - r_i^{(s)}$ 
8    for  $i = 1$  to  $M$  do
9       $C_i \leftarrow 1 - \prod_{s=1}^S m_i^{(s)}(\Theta_i) - \sum_{j=1}^N \left[ \prod_{s=1}^S (m_i^{(s)}(\{h_j\}) + m_i^{(s)}(\Theta_i)) - \prod_{s=1}^S m_i^{(s)}(\Theta_i) \right]$ 
10      $m_{D,i}(\Theta_i) \leftarrow \prod_{s=1}^S m_i^{(s)}(\Theta_i) / \max(1 - C_i, \varepsilon)$ ,  $m_{Y,i}(\Theta_i) \leftarrow \prod_{s=1}^S m_i^{(s)}(\Theta_i) + C_i$ 
11     for  $j = 1$  to  $N$  do
12        $m_{D,i}(\{h_j\}) \leftarrow \left[ \prod_{s=1}^S (m_i^{(s)}(\{h_j\}) + m_i^{(s)}(\Theta_i)) - \prod_{s=1}^S m_i^{(s)}(\Theta_i) \right] / \max(1 - C_i, \varepsilon)$ 
13        $m_{Y,i}(\{h_j\}) \leftarrow \prod_{s=1}^S (m_i^{(s)}(\{h_j\}) + m_i^{(s)}(\Theta_i)) - \prod_{s=1}^S m_i^{(s)}(\Theta_i)$ 
14   Initial  $f^* \leftarrow +\infty$ 
15   for each  $a \in [a_{\min}, a_{\max}]$  do
16     for each  $C_0 \in [0, 1]$  do
17       for  $i \leftarrow 1$  to  $M$  do
18          $g_i \leftarrow \sigma(a(C_i - C_0))$ ,  $m_i(A) \leftarrow (1 - g_i)m_{D,i}(A) + g_i m_{Y,i}(A)$ ,  $A \in \{\{h_1\}, \dots, \{h_N\}, \Theta_i\}$ 
19          $f(a, C_0) \leftarrow \frac{1}{M} \sum_{i=1}^M \frac{1}{S} \sum_{s=1}^S d_j(M_i(a, C_0), M_i^{(s)})$ 
20         if  $f(a, C_0) < f^*$  then
21           Update  $f^* \leftarrow f(a, C_0)$ ,  $(a^*, C_0^*) \leftarrow (a, C_0)$ 
22   for  $I \leftarrow 1$  to  $M$  do
23      $g_i \leftarrow \sigma(a^*(C_i - C_0^*))$ ,  $m_i(A) \leftarrow (1 - g_i)m_{D,i}(A) + g_i m_{Y,i}(A)$ ,  $A \in \{\{h_1\}, \dots, \{h_N\}, \Theta_i\}$ 
24      $\hat{y}_i = \arg \max_j [m_i(\{h_j\}) + m_i(\Theta_i) / N]$ ,  $conf_i = 1 - [\lambda m_i(\Theta_i) + (1 - \lambda)C_i]$ 

```

3.3. Overall Algorithm Procedure

In summary, the overall procedure of the proposed method is shown in Figure 1.

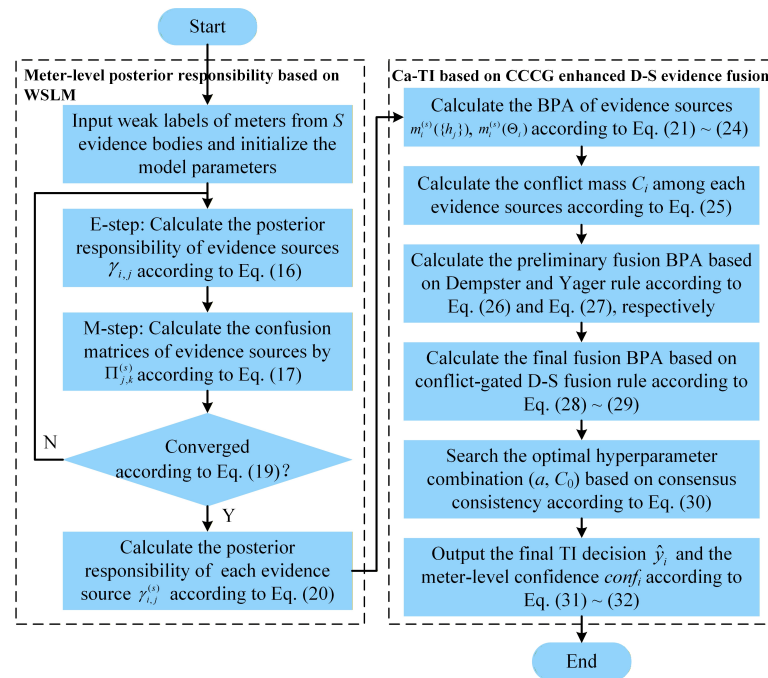


Figure 1. Flow chart of proposed Ca-TI method for LVDNs.

4. Case Study

4.1. Case Descriptions

Case 1 (simulated LVDTA): A simulated LVDTA case is established, as shown in Figure 2. The LVDTA contains three low-voltage branches and 65 smart meters. The line models are set as follows: the main trunk line uses BLVV-150 × 4, and the meter service drop uses BLVV-50 × 2; line lengths are marked in the figure. The required voltage, current, and power data are obtained by three-phase four-wire power flow simulation for LVDNs [37]. The data sampling interval is 15 min, generating a total of 6 days of data, i.e., a data length of 576 points. To better match practical conditions, Gaussian noise with a mean of 0 and 3σ of 5% is added to the voltage and current curves obtained from the power flow calculation.

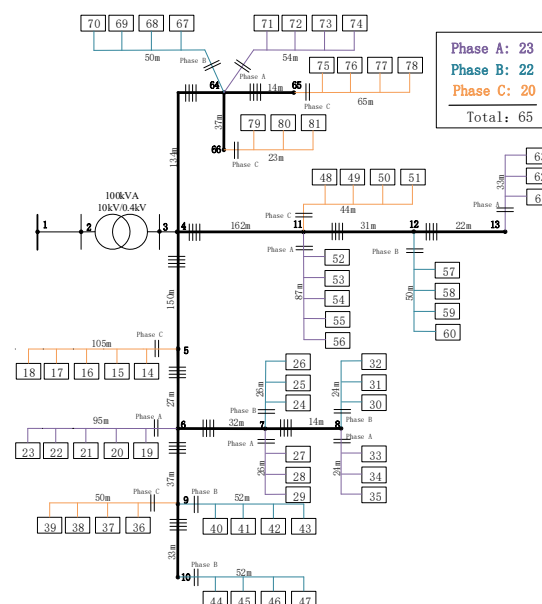


Figure 2. Topology structure of simulation substation in Case 1.

Case 2 (actual LVDTA): An actual LVDTA in Guangdong Province is selected for computational analysis, with its topology structure shown in Figure 3. This area is a typical urban residential service area, containing four low-voltage branches and 127 smart meters. For each meter, voltage, current, and power data are collected for 6 days, with a sampling interval of 15 min, i.e., a data length of 576 points.



Figure 3. Schematic diagram of actual substation in Case 2.

This paper uses C-PCI as an example for illustration. The identification methods proposed in References [13,16,18] are adopted as three evidence sources for fusion analysis and are defined as method 1, method 2, and method 3. The parameters of the algorithm are set in Table 2.

Table 2. Parameter setting of algorithm.

Parameter	Initial Value	Parameter	Initial Value
Number of hypotheses, N	3	Relative gain threshold of EM, τ	1×10^{-7}
Width of time window, W_{sw}	288	Numerical safeguard term, ϵ	1×10^{-7}
Sliding-window interval	1	Upper limit of gating slope, a_{max}	50
Number of time windows, T	289	Lower limit of gating slope, a_{min}	5
Accuracy initialization of evidence source, η	0.7	Weight coefficient, λ	0.5
Maximum number of iterations of EM, $iter_{max}$	100	Confidence threshold, ϕ	0.8

To comprehensively evaluate both the overall identification performance and the decision reliability induced by confidence scores, three metrics are proposed, the overall accuracy (A_{oa}), high-confidence region accuracy (A_{hc}), and low-confidence error recall (R_{lc}), as shown in the following equations:

$$A_{oa} = \sum_{i=1}^M \mathbf{1}[\hat{y}_i = y_i] / M \times 100\% \quad (33)$$

$$A_{hc} = \sum_{i \in \Omega} \mathbf{1}[\hat{y}_i = y_i] / |\Omega| \times 100\% \quad (34)$$

$$R_{lc} = \sum_{i \in \Psi} \mathbf{1}[\hat{y}_i \neq y_i] / |\Delta| \times 100\% \quad (35)$$

where y_i is the ground-truth topology of meter i , and in the case study of this paper, it refers to the true C-PC label; Ω , Ψ are the high-confidence set and low-confidence set, respectively, satisfying $\Omega = \{i | \text{conf}_i \geq \phi\}$, $\Psi = \{i | \text{conf}_i < \phi\}$; Δ is the misidentification set, satisfying $\Delta = \{i | \hat{y}_i \neq y_i\}$.

4.2. Identification Procedure

According to the WSLM proposed in this paper, the confusion matrices for method 1~method 3 can be calculated via the EM algorithm, as shown in Figure 4. In the confusion matrix, if the diagonal elements are large and close to 1, and the off-diagonal elements are small and close to 0, it indicates that the evidence source identification result has high stability; otherwise, it indicates randomness in the evidence source identification result. Specifically, the diagonal elements of the confusion matrices for method 1 (identifying Phase A and Phase B) and method 2 (for three-phase identification) are all greater than 0.85, indicating high stability, whereas method 3 corresponds to a confusion matrix where the diagonal elements are less than 0.8 (the diagonal elements of identifying Phase B and Phase C especially are significantly low, at only 0.634 and 0.337, respectively), indicating lower stability. According to the spatiotemporal characteristics of the consumer voltage analyzed in Reference [18], since the simulated LVDTA has small three-phase voltage unbalance, short supply lines, and light load levels, the voltage differences among meters are small. Method 3 is essentially based on voltage correlation for C-PC, which easily leads to significant randomness in identification results. It is evident that the reliability of identification results from a single method alone is difficult to guarantee. Through the proposed WSLM, fusion evaluation among multiple methods can be achieved, and the qualitative observation of the reliability of each method's identification results can be realized without ground-truth topology labels.

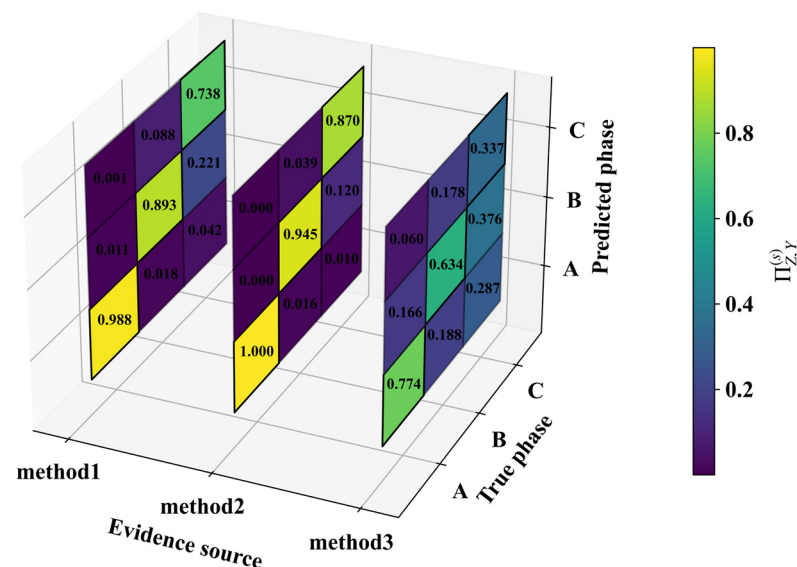


Figure 4. Confusion matrices of three evidence sources.

The posterior responsibilities of each evidence source are iteratively inferred using the EM algorithm until convergence. According to Equations (21)–(24), the BPA of each hypothesis for the C-PCI of each meter can be calculated, as shown in Figure 5. It can be seen that: (1) Method 1 and method 2 have high identification certainty. The identification result of each meter is assigned to a specific phase with a high probability value. (2) In contrast, method 3 shows randomness in the identification results, and the phase assignment for some meters is uncertain. Moreover, it can be intuitively observed that although various

methods yield highly confident identification results for certain meters, the C-PCI labels differ and conflict with one another. Consequently, it is impossible to determine which result should ultimately be trusted.

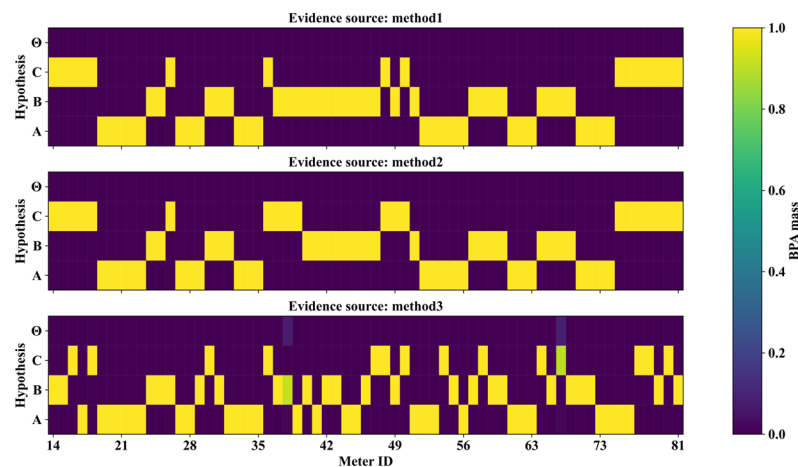


Figure 5. BPA of C-PCI hypotheses for each meter based on three evidence sources.

Furthermore, the proposed CCCG-DS evidence fusion method was used to fuse the identification results of the three evidence sources, and the fusion results are shown in Figure 6. Compared with the true phase labels, the relationship curve between the accuracy of meters with coverage and the meter coverage rate can be output, as shown in Figure 7. The x -axis represents the coverage statistics obtained by ranking meters in descending order of confidence scores. And the overall accuracy (A_{oa}), high-confidence region accuracy (A_{hc}), and low-confidence error recall (R_{lc}) are calculated as shown in Table 3. According to Figures 6 and 7 and Table 3, it can be seen that the fusion method achieved results consistent with method 1, with an overall accuracy (A_{oa}) of 90.77%, and it improved by -6.35% and 51.28% compared with method 2 and method 3, respectively; the fusion method can effectively integrate the performances of various identification methods, and it is not adversely affected by extreme bias from any single method (such as method 3). Moreover, the fusion method can provide meter-level confidence. And when the confidence threshold (ϕ) is set as 0.8, the high-confidence region accuracy (A_{hc}) and the low-confidence error recall (R_{lc}) can achieve 100%. This means that the fusion result for the meters in the high-confidence set (Ω) can be trusted, and only the meters in the low-confidence set (Υ) should be on-site-verified, which can significantly reduce the on-site verification work.

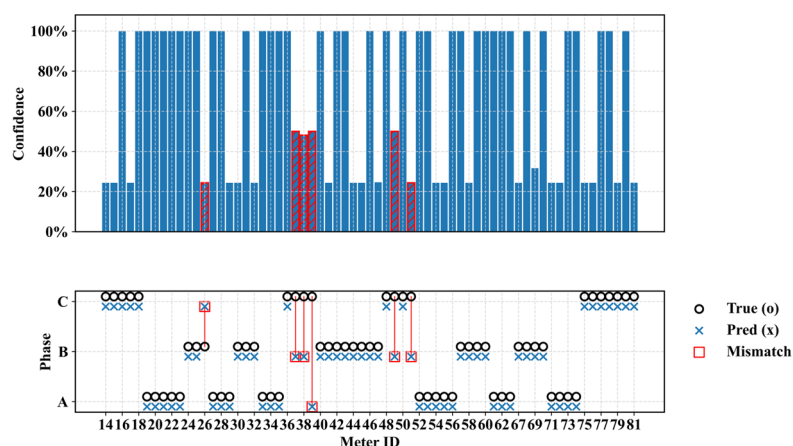


Figure 6. Results of C-PCI and confidence based on proposed fusion method in Case 1.

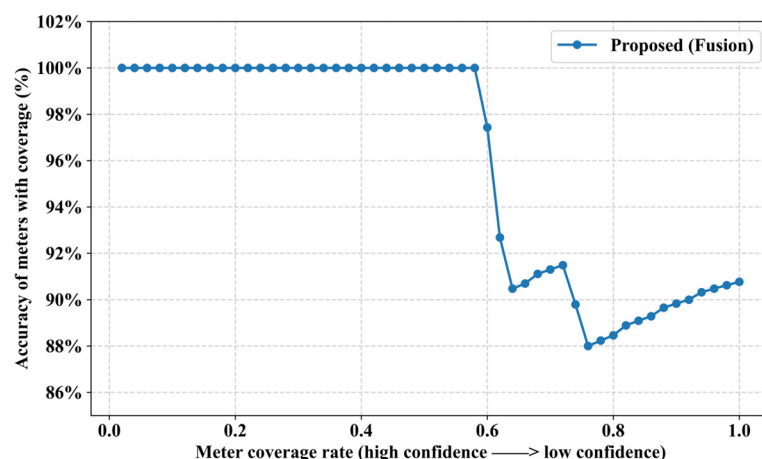


Figure 7. Curve of meters with coverage to accuracy based on proposed fusion method.

Table 3. Performance comparison of methods 1/2/3 and proposed method in Case 1.

Method	A_{oa}	A_{hc}	R_{lc}
Method 1	90.77%	/	/
Method 2	96.92%	/	/
Method 3	60.00%	/	/
Proposed method	90.77%	100%	100%

To further explain the high accuracy performance of the fusion method, Figure 8 provides the relationship curves for the dynamic gating coefficient, confidence score and conflict mass. It can be intuitively observed that as the confidence score decreases, the conflict mass increases, and the dynamic gating coefficient gradually increases. For identification results with high conflict, the proposed conflict-gating mechanism can bias them towards the Yager rule, thereby avoiding forced judgment and decision bias under high conflict, making the fusion result more robust. For meters with high conflict, the proposed method can effectively assign them to the low-confidence region. This can effectively reduce the deviation caused by the overconfidence of a certain single method, and it can effectively distinguish between high-confidence and low-confidence meter regions, helping to guide targeted and efficient on-site operation and maintenance.

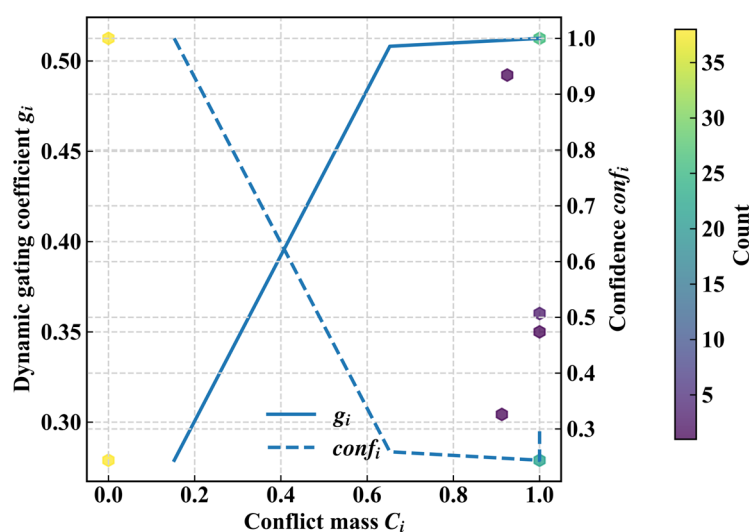


Figure 8. Curve of conflict mass to dynamic gating coefficient and confidence.

4.3. Effectiveness Verification of Proposed Algorithm

4.3.1. Verification of Proposed WLSM

To verify the effectiveness of the WSLM in learning source error patterns without ground-truth labels, and to quantify its contribution to downstream fusion performance, two comparison scenarios are set as follows (to isolate the effect of the CCCG-DS fusion method, the fusion rule is fixed to Dempster-only, and only the presence/absence of the WSLM is compared):

Scenario 1: WSLM + Dempster-only fusion: The WSLM is trained via EM using multi-source outputs to estimate source confusion matrices. And BPAs are constructed, followed by Dempster's combination rule to obtain the fused prediction.

Scenario 2: No-WSLM + Dempster-only fusion: The WSLM is disabled, and no confusion matrix is learned. For each source, $m_i^{(s)}(\{h_j\})$ is obtained by frequency counting across windows [18], and $m_i^{(s)}(\Theta_i)$ is set as zero, followed by the same Dempster-only fusion.

Since the proposed CCCG-DS fusion method was not adopted, the meter-level confidence could not be calculated. So, only the overall accuracy (A_{oa}) was used for the performance comparison of two scenarios. To mitigate the random effects of Gaussian noise, we repeated the experiment 100 times and calculated the expectation (exp) and standard deviation (std) of the evaluation metrics. The results are shown in Table 4. Compared with Scenario 2, enabling the WSLM increased the exp of A_{oa} from 91.29% to 98.71% and decreased the std of A_{oa} from 5.12% to 1.71%. This demonstrates that the WSLM can effectively learn each source's error pattern from multi-source consensus without ground-truth labels, thereby improving the evidence construction and mitigating the adverse impact of unreliable sources in fusion. Consequently, the WSLM constitutes a critical component of the proposed framework rather than a classical frequency-counting strategy.

Table 4. Performance comparison of Scenario 1 and Scenario 2.

Scenario	A_{oa}	
	exp	std
Scenario 1	98.71%	1.71%
Scenario 2	91.29%	5.12%

4.3.2. Verification of Proposed CCCG-DS Fusion

To verify the effectiveness of the CCCG-enhanced D-S evidence fusion method proposed in Section 3.2, five comparison scenarios are set, as described in Table 5. It should be noted that the voting-based fusion method [25] and the WSLM without fusion cannot provide meter-level confidence. To better compare the performances of the methods in various scenarios, the maximum value of frequency statistics was used to approximate the confidence for the voting-based fusion method, and the maximum posterior responsibility was used to approximate the confidence for the WSLM without fusion. Similarly, we repeated the experiments 100 times and calculated the exp and std of the evaluation metrics. The evaluate metrics for each scenario are shown in Table 6.

Table 5. Comparison scenario setting.

Scenario	Description
Scenario 3	Voting-based fusion [25]
Scenario 4	WSLM only (without fusion)
Scenario 5	WSLM + Dempster-only fusion
Scenario 6	WSLM + Yager-only fusion
Scenario 7	Proposed method (WSLM + CCCG-DS fusion)

Table 6. Performance comparison of Scenario 3–Scenario 7.

Scenario	A_{oa}		A_{hc}		R_{lc}	
	exp	std	exp	std	exp	std
Scenario 3	94.77%	3.46%	95.41%	3.31%	12.31%	19.25%
Scenario 4	98.71%	1.71%	98.71%	1.71%	0%	0%
Scenario 5	98.71%	1.71%	99.73%	0.74%	84.52%	33.00%
Scenario 6	98.49%	1.72%	99.72%	0.75%	87.63%	28.21%
Scenario 7	98.58%	1.70%	99.73%	0.74%	86.02%	31.65%

According to Table 5, the results show that:

- (1) The WSLM only in Scenario 4 achieved a comparable overall accuracy to that of the fusion-informed methods in Scenario 5–Scenario 7. However, R_{lc} of the WSLM-only scenario is 0, which indicates that erroneous meters were not flagged as low-confidence for inspection, and A_{hc} degenerated to A_{oa} ($98.71\% \pm 1.71\%$). In contrast, the fusion-informed methods preserved high A_{oa} while substantially improving error triage. For example, WSLM + CCCG-DS fusion achieved high A_{oa} ($98.58\% \pm 1.70\%$) and high A_{hc} ($99.73\% \pm 0.74\%$) and significantly higher error recall in the low-confidence region ($86.02\% \pm 31.65\%$). Therefore, the advantage and necessity of informed fusion are not tied to a monotonic improvement in A_{oa} but rather to providing an interpretable conflict-aware confidence that enables targeted verification for low-confidence meters, thereby reducing unnecessary field checks.
- (2) Comparing CCCG-DS fusion against Dempster-only fusion and Yager-only fusion can further validate the effectiveness of the CCCG mechanism. Their A_{oa} and A_{hc} are all high and close. Yet, CCCG exhibits a more balanced risk behavior: its R_{lc} lies between the other two rules, indicating that sample-wise gating adaptively trades off between Dempster's more aggressive normalization and Yager's more conservative conflict absorption, yielding a more even trade-off between high-confidence meters without verification and low-confidence meters with error recall.
- (3) The voting-based fusion method achieved relatively high A_{oa} and high A_{hc} . However, its R_{lc} performed significantly poorer than the proposed method, indicating that the voting mechanism based on superficial agreement does not reliably flag misidentified meters. The proposed method in this paper can mine the identification consensus among multiple sources and assign high and low probabilities to decisions for identification consensus and difference, respectively, which better supports a practical deployment for “trust and verification”.

4.4. Robustness Testing

To further verify the robustness of the proposed method, various scenarios involving measurement noise and missing data were set, and the methods set in Table 4 were tested. Specifically, for the analysis of the measurement noise impact, Gaussian noise with a mean of 0 and 3σ of 0%, 5%, 10%, 15% and 20% was added to the voltage and current curves without missing data; for the analysis of the missing data impact, missing data rates of 0%, 5%, 10%, 15% and 20% were simulated with 5% measurement noise. The evaluation metrics for each method are shown in Tables 7 and 8.

Table 7. Performances of methods under different measurement noise levels.

Scenario	Measurement Noise	A_{oa}		A_{hc}		R_{lc}	
		exp	std	exp	std	exp	std
Scenario 3	0%	100%	0	100%	0	/	/
Scenario 4		100%	0	100%	0	/	/
Scenario 5		100%	0	100%	0	/	/
Scenario 6		100%	0	100%	0	/	/
Scenario 7		100%	0	100%	0	/	/
Scenario 3	5%	94.77%	3.46%	95.41%	3.31%	12.31%	19.25%
Scenario 4		98.71%	1.71%	98.71%	1.71%	0%	0%
Scenario 5		98.71%	1.71%	99.73%	0.74%	84.52%	33.00%
Scenario 6		98.49%	1.72%	99.72%	0.75%	87.63%	28.21%
Scenario 7		98.58%	1.70%	99.73%	0.74%	86.02%	31.65%
Scenario 3	10%	81.29%	5.19%	82.36%	5.05%	8.93%	8.33%
Scenario 4		94.92%	3.39%	94.92%	3.39%	0%	0%
Scenario 5		92.34%	3.37%	97.61%	3.05%	86.56%	19.15%
Scenario 6		87.85%	4.85%	97.58%	3.09%	90.67%	14.12%
Scenario 7		89.94%	4.02%	97.60%	3.08%	89.30%	17.09%
Scenario 3	15%	73.82%	6.03%	74.78%	6.10%	7.73%	7.08%
Scenario 4		90.89%	4.36%	90.89%	4.36%	0%	0%
Scenario 5		86.31%	3.93%	94.55%	4.40%	86.97%	11.49%
Scenario 6		77.97%	4.99%	94.83%	4.65%	92.17%	7.36%
Scenario 7		82.37%	4.15%	94.62%	4.35%	90.02%	8.79%
Scenario 3	20%	70.77%	5.00%	71.85%	4.89%	7.37%	5.79%
Scenario 4		86.58%	8.73%	86.58%	8.73%	0%	0%
Scenario 5		82.71%	9.15%	91.81%	10.13%	85.25%	12.33%
Scenario 6		73.66%	7.79%	91.76%	10.34%	90.29%	8.78%
Scenario 7		77.94%	8.79%	91.84%	10.15%	88.28%	10.49%

Table 8. Performances of methods under different missing data rates.

Scenario	Missing Data	A_{oa}		A_{hc}		R_{lc}	
		exp	std	exp	std	exp	std
Scenario 3	0%	94.77%	3.46%	95.41%	3.31%	12.31%	19.25%
Scenario 4		98.71%	1.71%	98.71%	1.71%	0%	0%
Scenario 5		98.71%	1.71%	99.73%	0.74%	84.52%	33.00%
Scenario 6		98.49%	1.72%	99.72%	0.75%	87.63%	28.21%
Scenario 7		98.58%	1.70%	99.73%	0.74%	86.02%	31.65%
Scenario 3	5%	86.24%	3.97%	87.58%	4.02%	9.86%	11.37%
Scenario 4		96.85%	2.68%	96.85%	2.68%	0%	0%
Scenario 5		94.52%	2.84%	97.93%	2.79%	84.76%	23.05%
Scenario 6		91.21%	2.69%	97.88%	2.65%	86.55%	19.63%
Scenario 7		92.78%	2.75%	97.90%	2.72%	85.22%	21.54%
Scenario 3	10%	80.86%	4.83%	82.02%	4.78%	8.52%	10.01%
Scenario 4		94.69%	3.42%	94.69%	3.42%	0%	0%
Scenario 5		90.78%	3.71%	95.66%	3.79%	86.42%	19.90%
Scenario 6		87.40%	3.65%	95.89%	3.68%	88.76%	17.28%
Scenario 7		88.89%	3.70%	95.85%	3.66%	87.59%	18.79%

Table 8. Cont.

Scenario	Missing Data	A_{oa}		A_{hc}		R_{lc}	
		exp	std	exp	std	exp	std
Scenario 3	15%	75.24%	5.69%	76.83%	5.80%	7.22%	7.54%
Scenario 4		92.36%	4.67%	92.36%	4.67%	0%	0%
Scenario 5		86.63%	4.92%	93.26%	4.88%	87.30%	17.75%
Scenario 6		83.22%	4.80%	92.92%	4.83%	89.84%	15.17%
Scenario 7		84.51%	4.86%	93.03%	4.79%	88.38%	16.56%
Scenario 3	20%	71.35%	5.88%	72.75%	5.69%	7.01%	6.85%
Scenario 4		89.83%	5.52%	89.83%	5.52%	0%	0%
Scenario 5		82.45%	7.92%	90.94%	8.11%	88.46%	15.37%
Scenario 6		78.81%	7.10%	90.44%	8.36%	90.79%	12.99%
Scenario 7		81.06%	7.55%	90.75%	8.20%	89.54%	14.24%

According to Table 7, as the noise level increases from 5% to 20%, all methods exhibit a monotonic drop in each evaluation metric, accompanied by noticeably larger standard deviations. However, the proposed method (WSLM + CCCG-DS fusion) remains advantageous because its R_{lc} stays consistently high and relatively stable across all noise levels. This indicates that even under stronger noise, the proposed method can still concentrate most misidentified meters into the low-confidence region, enabling targeted on-site verification. In contrast, the voting-based method suffers a sharper accuracy decline (A_{oa} and A_{hc}) while keeping a persistently low and further decreasing R_{lc} , implying that superficial agreement alone is insufficient for robust identification.

Moreover, comparing the other fusion methods in Scenario 5 and Scenario 6, it can be observed that WSLM + Dempster-only fusion performs better on A_{oa} and A_{hc} , whereas WSLM + Yager-only fusion performs better on R_{lc} . This further indicates that Dempster's rule is more decisive, while Yager's rule is more conservative, and that the proposed method can achieve an effective balance, ensuring robustness against the measurement noise.

According to Table 8, the overall impact pattern of missing data is similar to that of measurement noise. The proposed method remains advantageous under heavier missingness: by learning source-specific error patterns in a weakly supervised manner and applying conflict-gated fusion, it explicitly characterizes the insufficient evidence and conflicting evidence, enabling the low-confidence region more stably. This yields a more robust and verifiable decision behavior under adverse data conditions, highlighting its robustness for practical deployment.

4.5. Sensitivity Analysis of Preset Thresholds

The proposed method involves several preset thresholds, and a sensitivity analysis of these parameters is required. The main preset thresholds are listed in Table 2. Except for the weight coefficient (λ) of the meter-level confidence calculation and the confidence threshold (ϕ) of the evaluation metrics, the other thresholds are not decision-critical. They mainly ensure numerical stability and do not change the predicted labels, such as τ , ε , η , and $iter_{max}$. And the hyperparameters a and C_0 of CCCG-DS fusion are adaptively learned rather than manually tuned. Thus, the sensitivity analysis of λ and ϕ was carried out. Gaussian noise with a mean of 0 and 3σ of 5% was added to the voltage and current curves without missing data. The experiment was repeated 100 times, and the exp and std of A_{hc} and R_{lc} were calculated. The results are shown in Figures 9 and 10.

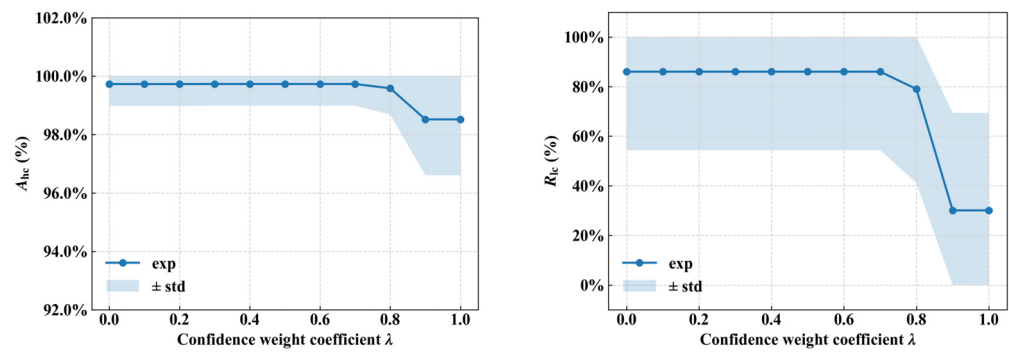


Figure 9. Exp and std of A_{hc} and R_{ic} under different weight coefficients (λ).

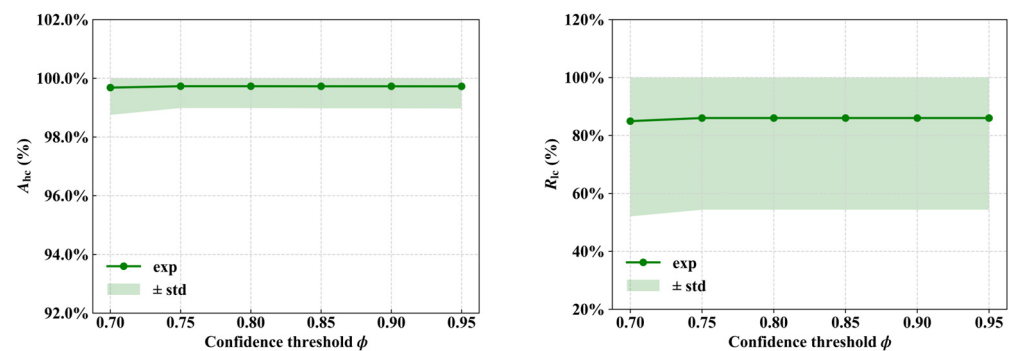


Figure 10. Exp and std of A_{hc} and R_{ic} under different confidence thresholds (ϕ).

According to Figure 9, the result shows that A_{hc} and R_{ic} both remain stable when λ is in the low-to-medium range (e.g., $0 \leq \lambda \leq 0.7$) but degrades sharply when λ becomes large. This degradation is expected because overly weighting the uncertainty term effectively downplays the conflict component in the confidence computation, making source disagreement less likely to be reflected as low-confidence risk. Therefore, we recommend choosing λ in a moderate-to-low range to obtain robust confidence partitioning and reliable error control.

According to Figure 10, the result shows that the variation in λ has a minor impact on A_{hc} and R_{ic} , with both remaining largely stable, indicating that the proposed method is insensitive to λ and provides consistent high-confidence decisions and error recall.

4.6. Validation Based on Actual LVDTA

To further verify the applicability of the proposed method in practical scenarios, a performance comparison analysis between single methods and the fusion method was carried out based on the actual LVDTA in Case 2. The identification results of the fusion method proposed are shown in Figure 11, and the performance comparison of the three single methods and the proposed fusion method is shown in Table 9.

Table 9. Performance comparison of methods 1/2/3 and proposed method in Case 2.

Method	A_{oa}	A_{hc}	R_{ic}
Method 1	99.21%	/	/
Method 2	91.34%	/	/
Method 3	54.33%	/	/
Proposed method	99.21%	99.21%	0%

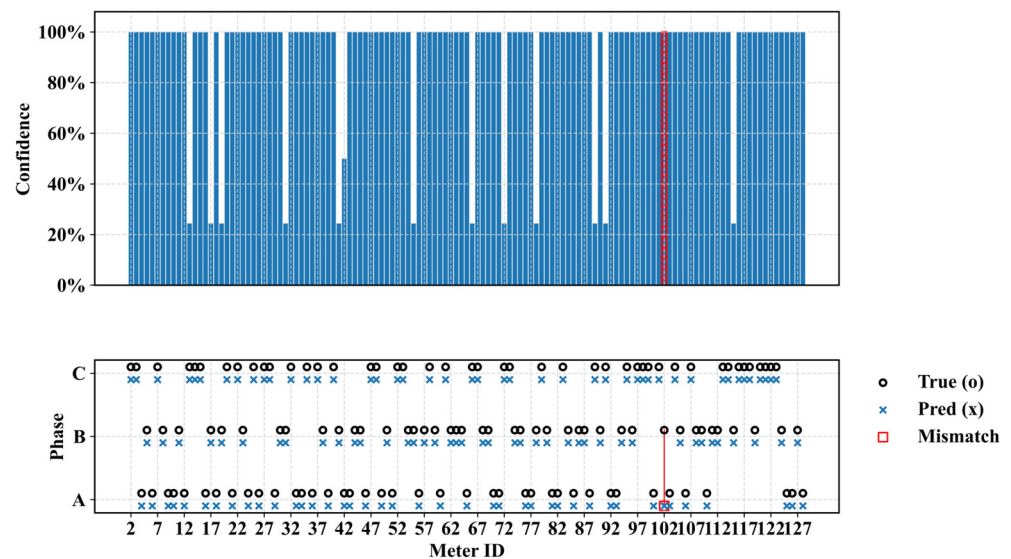


Figure 11. Results of C-PCI and confidence based on proposed fusion method in Case 2.

According to Table 8, the result shows that the A_{oa} of a single method varies in different cases. For example, method 2 performed the best in Case 1, whereas method 2 performed the best in Case 2. So, relying on a single method cannot guarantee stable accuracy. The proposed fusion method is not adversely affected by extreme bias from any single method and is able to maintain high identification accuracy. However, R_{lc} equals 0% mainly because the total number of misidentified meters is extremely small (only one error in this case) and the misidentified meter is assigned to the high-confidence set. This typically happens when sources exhibit consistent, correlated errors, leading to low uncertainty mass and low conflict, thus resulting in high confidence despite the incorrect prediction. Therefore, this reflects the challenge of detecting errors in correlated multi-source scenarios, which creates the need for future work.

5. Further Discussion on Engineering Deployment with Cloud–Edge Collaboration

5.1. Analysis of Mainstream LVDN TI Methods

The mainstream methods for LVDN TI can be broadly categorized into three types: the signal injection-based method, BPLC-based method and data driven-based method. The case study in Section 4 is analyzed based on the data driven-based method but is also applicable to the signal injection-based method and BPLC-based method. However, different types of identification methods differ in their characteristics and application scenarios and using only one type of method can hardly satisfy practical application requirements. The characteristics of these methods are summarized in Table 10. In the “Identification Ability” column, “√” indicates that the method only supports the identification of the meters registered in the metering archives. It cannot identify the meters that are actually supplied by the local LVDTA but not included in the metering archives. “√√” indicates that the method supports the identification of the meters registered in the metering archives. After data analysis in the cloud platform, the meters that are actually supplied by the local LVDTA but not listed in the archive can be identified.

Table 10. Summary of characteristics of mainstream LVDN TI methods.

Method	Application Requirement	Identification Ability	Deployment Location
Signal injection-based method	High configuration LVDAs with characteristic signal modules and BPLC communication module equipment	√ C-TCI √ C-PCI √ C-BCI	× Cloud √ Edge
BPLC-based method	Standard configuration LVDAs with BPLC communication module equipment	√ C-TCI √ C-PCI × C-BCI	× Cloud √ Edge
Data driven-based method	Other LVDAs without characteristic signal modules and BPLC communication module equipment	√√ C-TCI √√ C-PCI √√ C-BCI	√ Cloud √ Edge

Note: “√” indicates that the method only supports to identify the meters registered in the metering archives. It cannot identify the meters that are actually supplied by the local LVDTA but not included in the metering archives. “√√” indicates that the method supports to identify the meters registered in the metering archives. “×” indicates that the method is unable to identify.

According to Table 10: (1) In terms of application requirements, the signal injection-based method is the most demanding, requiring that devices such as transformer terminal units, branch terminal units and smart meters in the LVDTA be equipped with characteristic signal modules. The BPLC-based method is less demanding. With the continuous maturation of BPLC technologies, the local communication schemes in LVDAs are being progressively upgraded to BPLC; therefore, standard configuration LVDAs can generally support the BPLC-based method. The data driven-based method has the lowest requirements. Under the AMI, the measurement data in the LVDTA can be collected, and the data driven-based method can be carried out without characteristic signal modules and BPLC communication modules. (2) In terms of the identification ability and deployment location, because the signals monitored by the signal injection-based method and BPLC-based method are high-frequency and cannot be frequently reported to the cloud platform, these two types of methods can only be deployed on the edge side. Moreover, they can only identify meters included in the metering archives on the edge side. When there exist meters that are actually supplied by the target LVDTA but not included in the metering archives, these two methods do not support it. By contrast, the data driven-based method can be deployed either on the cloud side or on the edge side. In addition, measurement data from the LVDTA can be reported to the cloud platform to carry out the cross-LVDTA identification, achieving identification of the complete topology for the LVDTA.

5.2. Engineering Deployment Architecture

Based on the cloud–edge collaboration framework of the PloT, a system architecture for LVDN TI applications is designed, as shown in Figure 12.

Edge layer: This mainly refers to edge computing terminals and the various LVDN TI algorithms deployed on these terminals. Through communication technologies such as BPLC, the edge computing terminal can aggregate measurement data and signal information from smart meters and perform TI. As analyzed in Section 5.1, different types of TI methods have different application requirements; therefore, for high-configuration LVDAs, standard configuration LVDAs and other LVDAs, the types of algorithms deployed on the edge computing terminal should be different.

Cloud layer: This supports the rapid access of a large number of edge computing terminals and enables integration of large-scale data from LVDAs. In the form of microservice-based applications, the functional modules for LVDN TI can be rapidly deployed and put to use.

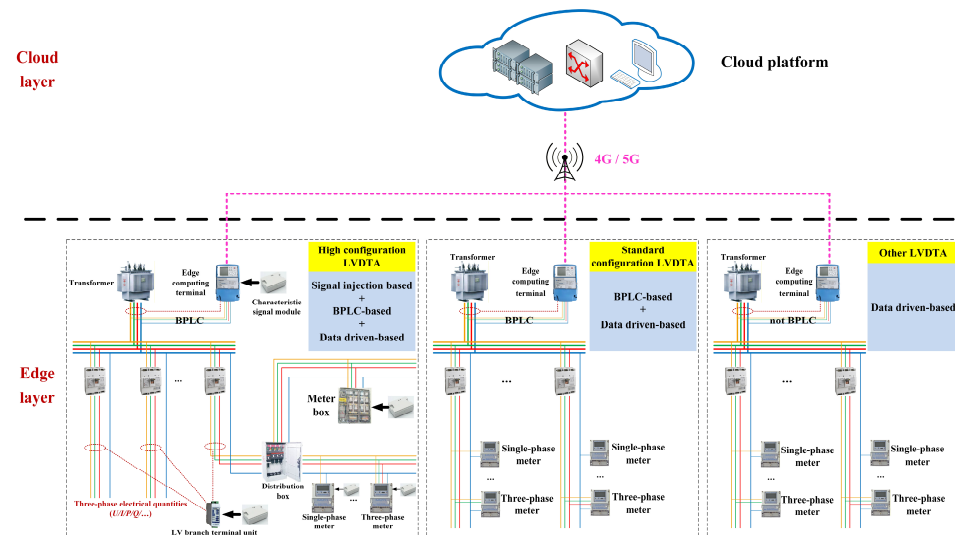


Figure 12. System architecture of topology identification for LVDN.

5.3. Function Realization and Data Interaction

Based on the aforementioned system architecture, this section further proposes the function realization and data interaction model for LVDN TI applications, as shown in Figure 13. “✓” indicates that the type of TI carried out in Figure 13.

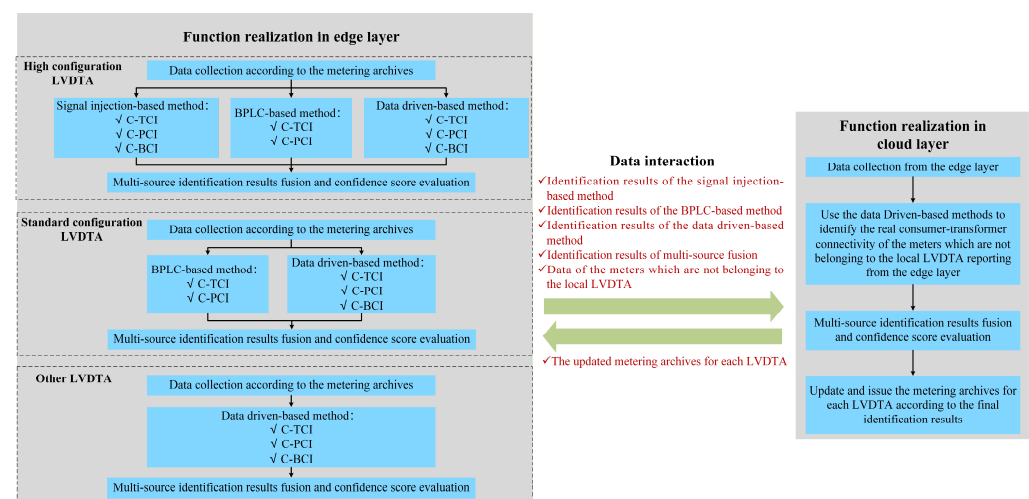


Figure 13. Diagram of function realization and data interaction of LVDN TI applications.

In terms of function realization, edge computing terminals are configured with different TI methods according to the application requirement of each method. TI on the edge side is only for meters that are in the metering archives of the target LVDTA, and the multi-source fusion identification results are reported to the cloud side. All the measurement data is collected on the cloud side, and the global re-attribution analysis can be conducted, thereby achieving the complete identification of the C-TC for each LVDTA. Further, the complete identification of the C-PC and C-BC can be realized. On this basis, the metering archives can be updated and issued back to each LVDTA, and the topology can be re-identified based on the updated metering archives. Through cloud–edge iterative optimization, the accurate and reliable identification of the LVDN topology can be achieved.

In terms of data interaction, the edge computing terminal primarily reports TI results and measurement data. The measurement data mainly includes the time-series data referring to 15 min interval monitoring data (energy, voltage, current, active power, power

factor, etc.) from various monitoring nodes in the area (including LV buses, LV branches, and smart meters). As the volume of this data is immense, uploading the full set would impose significant pressure on the uplink communication and cloud platform storage. Therefore, the system considers reporting only the time-series data of meters identified by the edge as “non-native” (not belonging to the target LVDTA) and meters with key consumption characteristics in the target LVDTA. This provides data for the cloud to perform re-allocation calculations for non-native meters. The cloud then issues the updated metering archive, enabling the edge to re-identify it.

5.4. Computational Complexity Analysis

The computational complexity mainly lies in posterior responsibility inference based on the EM algorithm, which can be analyzed as follows:

$$\vartheta_{\text{E-step}} = \vartheta(M \times N \times S \times \bar{T}) \quad (36)$$

$$\vartheta_{\text{M-step}} = \vartheta(S \times N^2 \times M \times \bar{T}) \quad (37)$$

$$\vartheta_{\text{EM}} = \vartheta(S \times N^2 \times M \times \bar{T}) \quad (38)$$

where $\bar{T} = \sum_{i,s} |\mathcal{O}_i^{(s)}| / MS$ is the average number of non-outcome-missing time windows per meter per evidence source.

In practical deployments, edge computing terminals commonly adopt ARM-based SoCs (e.g., Cortex-A9-class processors) with a typical configuration of a 1.0 GHz CPU, 2 cores, and 1.0 GB RAM, which is sufficient for non-real-time analytics tasks in PLoT environments. In a typical LVDTA with $M = 200$, $S = 3$, $T = 673$, and $N = 3$, the computation of each EM iteration is approximately 3.7×10^6 operations, and each EM iteration completes in approximately 300 ms with Python 3.13 coding. Calculated based on the maximum number of iterations, the total computation time is approximately 300 s. As topology identification is a quasi-static asset management task rather than a real-time protection/control function, EM can be executed at the edge for daily computation. Moreover, with the continuous advancement of embedded processor and accelerator technologies, the edge computing capability is expected to further improve, making EM-based training and confidence-aware TI increasingly practical for deployment on the edge side.

6. Conclusions

Aiming at the problems that the current LVDN TI research lacks the consideration of multiple method fusion and the confidence evaluation of identification results, and that the reliability and practicality of identification results are difficult to guarantee, this paper proposes a Ca-TI method for LVDNs based on the WSLM and CCCCCG-DS evidence fusion, obtaining the following conclusions:

- (1) The proposed framework delivers reliable TI results under practical bad-data conditions. Compared with naive agreement-based fusion, the proposed method mitigates performance collapse caused by a single unreliable method and remains robust under increasing measurement noise and missing data.
- (2) Beyond accuracy, the proposed method can provide the meter-level confidence to support deployment: high-confidence meters can be trusted with minimal verification, while low-confidence meters can be prioritized for verification, enabling a practical “trust-and-verify” workflow that reduces unnecessary on-site checks.
- (3) We further analyzed the preset threshold sensitivity for the proposed method. The confidence thresholds exhibit stable behavior over a broad range, indicating that the proposed confidence scoring supports consistent high-confidence decision quality

and error recall. The confidence weight coefficient remains stable in moderate-to-low ranges, whereas overly emphasizing the uncertainty term can weaken conflict-driven risk signaling; this provides clear guidance for practical parameter selection.

Future work will focus on the physical interpretability of TI confidence computation and relaxing the implicit independence assumption in D-S fusion because different TI methods may share common physical features (e.g., voltage/power variations), leading to correlated evidence and potentially overconfident decisions. We will investigate dependence-aware evidence fusion by explicitly modeling inter-source correlation or applying correlation-aware discounting to improve confidence calibration. In addition, to support scalable and online maintenance under evolving operating conditions, we will explore reinforcement-learning-driven sparse update strategies that selectively trigger re-identification or field verification only for persistently low-confidence/high-conflict meters [38]. Finally, we will study how confidence-aware TI can better support downstream grid operations under high EV integration, following recent system-level assessments of EV impacts on distribution networks [39].

Author Contributions: Conceptualization, S.L., C.D. and Z.Z.; experiments, Y.Z. and H.L.; data analysis, S.L., C.D. and Z.Z.; writing—original draft preparation, S.L., C.D. and Z.Z.; writing—review and editing, W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Natural Science Foundation of China (No. 52177085).

Data Availability Statement: The data presented in this study are available upon request from the corresponding author, while some of the data are not publicly available due to privacy restrictions.

Acknowledgments: All authors express gratitude for the support and cooperation provided by their respective institutions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kalinga, T.; Banfield, B.; Knott, J.C.; Robinson, D.A. K-Means Clustering and Linear Regression for User Phase Identification, Verification, and Topology Determination Under Varied Smart Meter Penetration. *Energies* **2026**, *19*, 183. [\[CrossRef\]](#)
2. Lei, Y.; Yang, F.; Feng, Y.; Hu, W.; Cheng, Y. A Two-Stage Topology Identification Strategy for Low-Voltage Distribution Grids Based on Contrastive Learning. *Energies* **2025**, *18*, 5886. [\[CrossRef\]](#)
3. Zhang, Y.; Liu, S.; Jiang, J.; Yi, Y.; Li, Q. Research Review on Low-Voltage Intelligent Distribution Network Technology. *Guangdong Electr. Power* **2019**, *32*, 1–12.
4. Wang, P.; Zhang, D.; Gan, L.; Zhang, Y. Key Technologies and Applications of Collaboration between Digital Power Grid and Internet of Things. *Digit. Twins Appl.* **2024**, *1*, 26–37. [\[CrossRef\]](#)
5. Ge, H.; Xu, B.; Chen, W.; Zhang, X.; Bi, Y. Topology identification of low voltage distribution network based on current injection method. *Arch. Electr. Eng.* **2021**, *70*, 297–306.
6. Fernandez, J.H.; Omri, A.; Di Pietro, R. Power grid surveillance: Topology change detection system using power line communications. *Int. J. Electr. Power Energy Syst.* **2023**, *145*, 108634. [\[CrossRef\]](#)
7. Lai, X.; Cao, M.; Liu, S.; Sun, C. Low-Voltage Distribution Network Topology Identification Method Based on Characteristic Current. In Proceedings of the 2021 6th Asia Conference on Power and Electrical Engineering (ACPEE), Chongqing, China, 8–11 April 2021; pp. 1233–1238. [\[CrossRef\]](#)
8. Zhou, C.; Dai, S.; Pan, Z.; Wu, H.; Chen, C. An Algorithm for Dynamic Identification of Low-Voltage Topology. *J. Phys. Conf. Ser.* **2023**, *2428*, 012036. [\[CrossRef\]](#)
9. *IEEE Std 1901–2020 (Revision of IEEE Std 1901–2010)*; IEEE Standard for Broadband over Power Line Networks: Medium Access Control and Physical Layer Specifications. IEEE: New York, NY, USA, 2021. [\[CrossRef\]](#)
10. Lavenu, C.; Aldebert, T.; Caqueux, M.; Bretand, D.; Chapoulie, A. Phase Detection in PLC-Based Advanced Metering Infrastructures. In Proceedings of the 25th International Conference on Electricity Distribution (CIRED 2019), Madrid, Spain, 3–6 June 2019; p. 136.
11. Xu, C.; Lei, Y.; Zou, Y. A Method of Low-Voltage Topology Identification. In Proceedings of the 2020 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS), Shenyang, China, 11–13 December 2020; pp. 318–323. [\[CrossRef\]](#)

12. Xu, M.; Li, R.; Li, F. Phase Identification with Incomplete Data. *IEEE Trans. Smart Grid* **2018**, *9*, 2777–2785. [\[CrossRef\]](#)
13. Yi, Y.; Liu, S.; Zhang, Y.; Xue, Y.; Deng, W.; Li, Q. Phase Identification of Low-Voltage Distribution Network Based on Stepwise Regression Method. *J. Mod. Power Syst. Clean Energy* **2023**, *11*, 1224–1234. [\[CrossRef\]](#)
14. Zhang, Y.; Yi, Y.; Deng, W.; Liu, S.; Zhou, L.; Lin, K.; Cai, Y. Consumer-Branch Connectivity Identification of Low-Voltage Distribution Networks Based on Data-Driven Approach. *Prot. Control Mod. Power Syst.* **2024**, *9*, 60–82. [\[CrossRef\]](#)
15. Bariya, M.; Deka, D.; von Meier, A. Guaranteed Phase & Topology Identification in Three-Phase Distribution Grids. *IEEE Trans. Smart Grid* **2021**, *12*, 3605–3612. [\[CrossRef\]](#)
16. Zhou, L.; Li, Q.; Zhang, Y.; Chen, J.; Yi, Y.; Liu, S. Consumer Phase Identification Under Incomplete Data Condition with Dimensional Calibration. *Int. J. Electr. Power Energy Syst.* **2021**, *129*, 106851. [\[CrossRef\]](#)
17. Wang, Z.; Yu, T.; Huang, Y.; Wang, K.; Pan, Z. A New Topology Identification Method for Low-Voltage Distribution Network Based on U-Shaplets Feature Subsequence. *Electr. Meas. Instrum.* **2025**, *62*, 68–79.
18. Zhou, L.; Zhang, Y.; Liu, S.; Li, K.; Li, C.; Yi, Y.; Tang, J. Consumer Phase Identification in Low-Voltage Distribution Network Considering Vacant Users. *Int. J. Electr. Power Energy Syst.* **2020**, *121*, 106079. [\[CrossRef\]](#)
19. Tang, J.; Cai, Y.; Zhou, L.; Yi, Y.; Chen, G.; Liang, S.; Zhang, Y. Data-Driven Based Identification Method of Feeder-Consumer Connectivity in Low-Voltage Distribution Network. *Autom. Electr. Power Syst.* **2020**, *44*, 127–134.
20. Li, H.; Liang, W.; Liang, Y.; Li, Z.; Wang, G. Topology Identification Method for Residential Areas in Low-Voltage Distribution Networks Based on Unsupervised Learning and Graph Theory. *Electr. Power Syst. Res.* **2023**, *215*, 108969. [\[CrossRef\]](#)
21. Mauladdawilah, H.; Balfaqih, M.; Balfagih, Z.; Pegalajar, M.d.C.; Gago, E.J. Deep Feature Selection of Meteorological Variables for LSTM-Based PV Power Forecasting in High-Dimensional Time-Series Data. *Algorithms* **2025**, *18*, 496. [\[CrossRef\]](#)
22. Elsafi, A.; Almohammed, A.A.; Balfaqih, M.; Balfagih, Z.; Sabri, S. Comparative Analysis of Maximum Power Point Tracking Methods for Power Optimization in Grid Tied Photovoltaic Solar Systems. *Discov. Appl. Sci.* **2025**, *7*, 976. [\[CrossRef\]](#)
23. Guo, X.; Xu, C.; Ming, Z.; Meng, B.; Yang, S.; Xu, L.; Zhu, Y. User-Feeder Topology Identification in Low-Voltage Residential Power Networks: A Clustering Fusion Approach. *Energies* **2025**, *18*, 4908. [\[CrossRef\]](#)
24. Li, K.; Zhou, L.; Zhang, Y.; Liu, X.; Li, G.; Huang, G. Credibility Evaluation of Topology Identification Results in Low-Voltage Distribution Network Based on Quality of Measured Data. *Autom. Electr. Power Syst.* **2021**, *45*, 99–107.
25. Zhang, H.; Zhao, J.; Wang, X.; Xuan, Y. Low-Voltage Distribution Grid Topology Identification With Latent Tree Model. *IEEE Trans. Smart Grid* **2022**, *13*, 2158–2169. [\[CrossRef\]](#)
26. Zhou, Z.-H. A brief introduction to weakly supervised learning. *Natl. Sci. Rev.* **2018**, *5*, 44–53. [\[CrossRef\]](#)
27. Khaleghi, B.; Khamis, A.; Karray, F.O.; Razavi, S.N. Multisensor data fusion: A review of the state-of-the-art. *Inf. Fusion* **2013**, *14*, 28–44. [\[CrossRef\]](#)
28. Zadeh, L.A. Review of Shafer's A Mathematical Theory of Evidence. *AI Mag.* **1984**, *5*, 81–83.
29. Denœux, T.; Li, S. Frequency-Calibrated Belief Functions: Review and New Insights. *Int. J. Approx. Reason.* **2018**, *92*, 232–254. [\[CrossRef\]](#)
30. Tang, Y.; Zhou, Y.; Ren, X.; Sun, Y.; Huang, Y.; Zhou, D. A New Basic Probability Assignment Generation and Combination Method for Conflict Data Fusion in the Evidence Theory. *Sci. Rep.* **2023**, *13*, 8443. [\[CrossRef\]](#)
31. Fu, W.; Yu, S.; Wang, X. A Novel Method to Determine Basic Probability Assignment Based on Adaboost and Its Application in Classification. *Entropy* **2021**, *23*, 812. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Fei, L.; Li, T.; Ding, W. A State-of-the-Art Survey on Neural Computation-Enhanced Dempster-Shafer Theory for Safety Accidents: Applications, Challenges, and Future Directions. *Neurocomputing* **2025**, *643*, 130419. [\[CrossRef\]](#)
33. Zhang, J.; Hsieh, C.-Y.; Yu, Y.; Zhang, C.; Ratner, A. A Survey on Programmatic Weak Supervision. *arXiv* **2022**, arXiv:2202.05433. [\[CrossRef\]](#)
34. Zhang, J.; Yu, Y.; Li, Y.; Wang, Y.; Yang, Y.; Yang, M.; Ratner, A. WRENCH: A Comprehensive Benchmark for Weak Supervision. In Proceedings of the Neural Information Processing Systems (NeurIPS) Datasets and Benchmarks, Online, 6–14 December 2021. [\[CrossRef\]](#)
35. Huang, L.; Ruan, S.; Decazes, P.; Denœux, T. Deep Evidential Fusion with Uncertainty Quantification and Reliability Learning for Multimodal Medical Image Segmentation. *Inf. Fusion* **2025**, *113*, 102648. [\[CrossRef\]](#)
36. Zhao, K.; Li, L.; Chen, Z.; Sun, R.; Yuan, G.; Li, J. A Survey: Optimization and Applications of Evidence Fusion Algorithm Based on Dempster-Shafer Theory. *Appl. Soft Comput.* **2022**, *124*, 109075. [\[CrossRef\]](#)
37. Gao, Y.; Wang, C.; Xin, J.; Liu, A.; Chen, Y. Calculation of Three-Phase Power Flow in Medium and Low-Voltage Distribution Network with Multiple Distributed Generators Taking Into Account Ground Impedance (1): Model. *Proc. CSEE* **2016**, *36*, 2619–2627.

38. Kaloev, M.; Krastev, G. Comprehensive Review of Benefits from the Use of Sparse Updates Techniques in Reinforcement Learning: Experimental Simulations in Complex Action Space Environments. In Proceedings of the 2023 4th International Conference on Communications, Information, Electronic and Energy Systems (CIEES 2023), Plovdiv, Bulgaria, 23–25 November 2023. [[CrossRef](#)]
39. Amann, G.; Escobedo Bermúdez, V.R.; Boskov-Kovacs, E.; Gallego Amores, S.; Giannelos, S.; Iliceto, A.; Ilo, A.; Chavarro, J.R.; Samovich, N.; Schmitt, L.; et al. *E-Mobility Deployment and Impact on Grids: Impact of EV and Charging Infrastructure on European T&D Grids: Innovation Needs*; Report No. MJ-09-22-246-EN-N; Publications Office of the European Union: Luxembourg, 2022. [[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.