Journal of Electronic Research and Application, 2026, Volume 10, Issue 9
https://bbwpublisher.com/journal/JERA
Online ISSN: 2208-3510
Print ISSN: 2208-3502

Dynamic Human–AI Authority Allocation for Distribution Network Fault Management under Diagnostic Uncertainty

Biwei Li[1], Caiyu Zhang[2*], Weijie Chen[1], Hai Yang[1]

[1]Zhaoqing Power Supply Bureau of Guangdong Power Grid Co., Ltd., Zhaoqing 526000, Guangdong, China
[2]Guangdong Technology College, Zhaoqing 526100, Guangdong, China

Author to whom correspondence should be addressed.

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: To address the mismatch between diagnostic credibility and machine authority in distribution network fault management, a dynamic human–AI authority allocation model under diagnostic uncertainty is proposed. Based on candidate fault hypotheses and posterior probabilities, comprehensive diagnostic uncertainty is quantified from data evidence, model outputs, and state information. Decision risk is calculated by considering safety, load-loss, operating-limit violation, switching-operation, and time-related losses of candidate actions under different fault states. Allowable authority is determined according to risk thresholds, safety constraints, action reversibility, and operator availability. Three authority modes are configured: machine-led operation, artificial intelligence (AI) recommendation with human confirmation, and human-led operation with AI assistance. A closed loop is established through human takeover, evidence supplementation, rediagnosis, and offline parameter calibration. Illustrative scenarios are constructed using the IEEE 33-bus distribution network as the network context and compared with fixed authority-allocation strategies. The comparison indicates that machine-led operation is selected under low uncertainty and low risk; when a single factor constrains machine authority, AI recommendation with human confirmation is adopted; and when risk thresholds or safety constraints further restrict authority, human-led operation with AI assistance is applied. Authority allocation based only on diagnostic uncertainty may retain excessive machine authority in low-uncertainty, high-risk scenarios. Diagnostic credibility should therefore not be the sole basis for authorizing autonomous machine actions; diagnostic uncertainty, decision risk, and safety constraints should be jointly considered in dynamic human–AI authority allocation.

Keywords: Distribution network fault management; Diagnostic uncertainty; Decision risk; Human–AI collaboration; Dynamic authority allocation

Online publication: October 30, 2026

1. Introduction

Fault handling in distribution networks is characterized by high time criticality, complex operational

constraints, and sensitive safety consequences. Once a fault occurs, fault diagnosis, isolation, load transfer and restoration, and result confirmation must be completed rapidly; diagnostic errors or operational mistakes may expand the outage scope, cause loss of supply to critical loads, and create equipment operation risks. The development of distribution automation, intelligent diagnosis, and power supply restoration technologies has enhanced capabilities in fault isolation, network reconfiguration, and coordination of distributed resources [1–4]. Intelligent diagnostic modules can use multi-source information—such as terminal measurements, protection operations, network topology, and equipment status—to form candidate fault determinations. Consequently, a problem that needs further resolution is how to allocate handling decision-making authority between machines and operators when diagnostic credibility and the potential consequences of candidate operations differ.

Existing research mainly falls into two categories: distribution network fault diagnosis and supply restoration, and human-machine collaborative decision-making. The former improves diagnostic reliability and scheme feasibility through information fusion, fault location, network reconfiguration, and restoration optimization [5–7]; the latter focuses primarily on levels of automation, trust calibration, decision authorization, shared situational awareness, and human oversight [8–18]. However, both categories still lack a human-machine authority allocation mechanism that links diagnostic uncertainty with handling consequences. Diagnostic uncertainty reflects the credibility of fault determination, whereas whether a machine can autonomously execute a scheme also depends on the potential handling consequences under diagnostic errors. Authorizing solely based on diagnostic confidence or uncertainty may grant excessive authority to operations that are diagnostically clear but have high potential handling consequences; fixed manual confirmation, in turn, makes it difficult to leverage the rapid handling advantages of low-uncertainty, low-risk scenarios.

To address these issues, this paper proposes a dynamic human-machine decentralized decision-making model for distribution network fault handling under diagnostic uncertainty. Unlike methods that allocate machine authority solely based on diagnostic credibility, this paper quantifies comprehensive diagnostic uncertainty from data evidence, model outputs, and state information; by combining the posterior probabilities of candidate faults with the handling losses of schemes under different true fault states, it constructs scheme-level decision risk, distinguishing “whether the diagnosis is credible” from “whether the handling is suitable for autonomous execution.” Furthermore, by integrating risk thresholds, safety constraints, operation reversibility, and conditions for human intervention, it establishes a dynamic switching and human takeover mechanism comprising machine-led, Artificial Intelligence (AI) recommendation–human confirmation, and human-led–AI assistance modes, and illustrates the authority allocation characteristics under different combinations of uncertainty and risk through typical scenarios.

2. Analysis of distribution network fault handling and diagnostic uncertainty

2.1. Fault handling business process

Distribution network fault handling includes alarm ingestion, fault diagnosis, fault isolation, load transfer and restoration, and result confirmation. The system gathers information such as automated terminal data, protection operations, network topology, equipment ledgers, and work orders to form candidate fault hypotheses; the AI module can perform information fusion, candidate fault ranking, handling scheme generation, and safety constraint verification, while operators confirm, correct, or take over high-risk operations. After handling is completed, the diagnostic conclusions and handling schemes are verified against

actual results. This paper only discusses the rapid post-fault handling process [19–22].

2.2. Multi-source fault diagnosis information

The observation information is denoted as:

$$x_t = \left\{ x_{mea,t}, x_{prot,t}, x_{topo,t}, x_{asset,t}, x_{work,t}, x_{field,t} \right\}$$

Where $x_{mea,t}$、$x_{prot,t}$ $x_{topo,t}$ $x_{asset,t}$ $x_{work,t}$ and $x_{field,t}$ represent terminal measurements, protection and fault indications, topology and switching states, equipment asset and defect records, dispatch work orders and operation records, and field inspection and user repair request information, respectively. Measurement and protection information reflect the electrical response to a fault; topology and switching states define the fault propagation paths and the operable scope; the remaining information is used for fault attribution and verification of handling measures. The timeliness, completeness, and consistency of the information affect the quality of diagnostic evidence [23–25].

2.3 Sources of diagnostic uncertainty

Diagnostic uncertainty refers to the degree to which, under current observation conditions, the system finds it difficult to form a stable and credible judgment about the true fault state. Incomplete multi-source information, weakened fault features, model discrepancies, and uncertain evidence fusion methods may all affect the identification of candidate fault states [26–28]. According to the role that uncertainty plays in the diagnostic process, this paper classifies it into data evidence uncertainty, model output uncertainty, and state information uncertainty:

$$U_{D,t} = f(U_{data,t}, U_{model,t}, U_{state,t})$$

where $U_{data,t}$ denotes the evidence uncertainty caused by measurement errors, missing data, communication delays, and abnormal protection information; $U_{model,t}$ denotes the output uncertainty arising from the AI diagnostic module's lack of concentration and discriminability in determining candidate faults; and Ustate, $U_{state,t}$ denotes the uncertainty resulting from difficulties in consistently confirming topology, switch positions, and key operating states. Renewable energy output and load fluctuations are incorporated into $U_{D,t}$ only when they affect fault diagnosis; if they affect only the transfer capacity or the feasibility of the scheme, they are considered in subsequent constraints, and no separate parallel uncertainty indicator is established [1,4–6,25,29].

2.4. Limitations of the fixed human–machine authority mode

The fixed authority mode struggles to adapt to changes in diagnostic conditions and handling risks. Full manual decision-making throughout the process may increase handling delays in low-risk scenarios; fixed AI recommendations make it difficult to unleash machine execution capabilities for low-risk operations; and authorization based solely on diagnostic confidence or uncertainty may neglect the consequences of plan handling, thereby granting excessive authority to operations with relatively clear diagnoses but high potential consequences. Therefore, it is necessary to incorporate diagnostic uncertainty, plan decision-making risk, and safety constraints into a dynamic authority allocation mechanism [8–16].

3. Human–machine dynamic authority allocation architecture oriented to diagnostic uncertainty

3.1. Overall architecture

The human–machine dynamic authority allocation architecture for distribution network fault handling oriented to diagnostic uncertainty is shown in Figure 1. The system integrates multi-source observational information to form candidate fault hypotheses and posterior probabilities, and quantifies comprehensive diagnostic uncertainty. Based on candidate fault states and operating conditions, candidate handling plans are generated, and handling losses and decision-making risks under different true fault states are calculated. Combined with diagnostic uncertainty, safety constraints, and operational characteristics, the allowable authority set and actual authority mode are determined. Subsequently, executable plans are screened under the corresponding authority and personnel intervention conditions, and the optimal final plan is selected. Operators can trigger re-diagnosis through confirmation, takeover, and supplementary evidence, and the handling results are used for diagnostic evaluation and authority parameter calibration. (**Figure 1**)

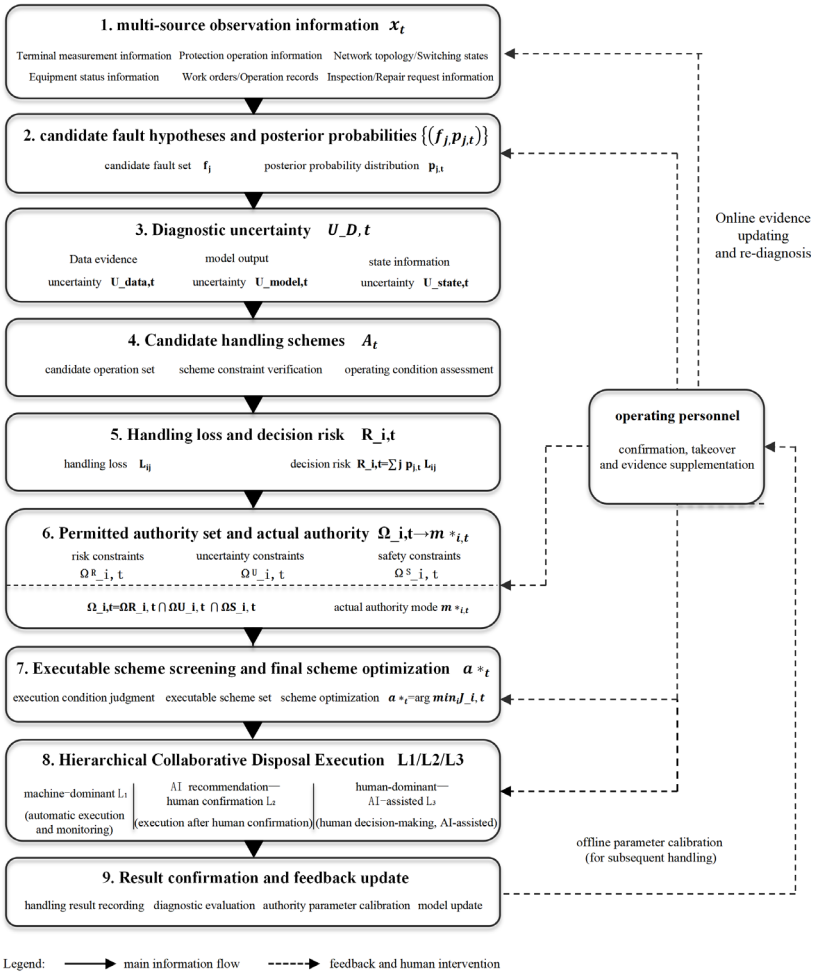


Figure 1. Dynamic human–AI authority allocation framework for distribution network fault management under diagnostic uncertainty.

3.2. Human–machine functional boundary

AI and operators form a complementary division of labor. AI can undertake information fusion, calculation of candidate fault probabilities, generation or recommendation of candidate schemes, risk alerts, and verification of operational constraints; this paper does not evaluate the recognition accuracy or scheme generation performance of specific AI algorithms. Operators are responsible for confirming high-risk operations, overriding abnormal results, supplementing evidence, and providing final authorization for critical operations. Human–machine authority is dynamically adjusted according to diagnostic uncertainty, handling risk, and safety conditions; operations involving major safety consequences always retain a human authorization boundary [11–16].

3.3. Decision authority levels

To distinguish the degree of decision control in fault handling, three levels of authority are established:

$$L=\{L_1,L_2,L_3\} \tag{1}$$

Where: L1 denotes machine-dominant operation, in which AI may generate, verify, and execute schemes within pre-authorized safety constraints, while operators are responsible for monitoring and retain the rights of suspension and override; L2 denotes AI recommendation–human confirmation, in which AI generates schemes and indicates risks, and operators execute after confirmation and may revise or reject the recommendations; L3 denotes human-dominant operation with AI assistance, in which operators are responsible for scheme judgment, final decision-making, and operational authorization, while AI only provides information analysis, candidate schemes, and risk alerts. At all levels, operators retain the rights of abnormal override and emergency takeover. AI may participate in diagnosis and decision support at all levels, but the qualification for autonomous execution is determined by the dynamic authority allocation mechanism [8–10,13].

4. Quantification of uncertainty in distribution network fault diagnosis

4.1. Candidate fault hypotheses and probabilities

After acquiring multi-source observation information xtxt at time tt, the diagnosis module constructs a set of candidate fault hypotheses:

$$F_t=\{f_1,f_2,\ldots,f_n\} \tag{2}$$

Where fj*fj* denotes the j*j*-th candidate fault state consistent with the current observation, which can be characterized by the fault section, fault type, and equipment status. Its posterior probability is:

$$p_{j,t}=P\left(f_j|\ x_t\right),\quad p_{j,t}\geq 0,\quad \sum_{j=1}^{n}p_{j,t}=1 \tag{3}$$

$p_{j,t}$ represents the relative evidential support obtained by a candidate state, rather than a simple confidence level assigned to a single conclusion. This paper does not restrict itself to a specific diagnostic algorithm; instead, the diagnostic module that outputs candidate states and their probability distribution is treated as an interface. This distribution is simultaneously used for uncertainty quantification and decision risk calculation. When entering uncertainty evaluation, $n\geq 2$ is taken, and a single candidate state is handled as a deterministic boundary case [5,6,25].

4.2. Data evidence uncertainty

Data evidence uncertainty reflects the degree to which observational information cannot fully, promptly, and reliably characterize the true operating state. For the k type of evidence, the credibility $q_{k,t} \in [0,1]$ is calculated based on completeness, timeliness, consistency, and source reliability, and is weighted according to importance:

$$q_{data,t} = \sum_{k=1}^{K} \omega_k q_{k,t}, \qquad \sum_{k=1}^{K} \omega_k = 1, \qquad \omega_k \geq 0 \tag{4}$$

Accordingly, it is defined that:

$$U_{data,t} = 1 - q_{data,t} \tag{5}$$

A larger $U_{data,t}$ indicates lower evidence quality. A decline in evidence quality does not necessarily lead to dispersion of the output probabilities; therefore, data evidence uncertainty and model output uncertainty are quantified separately [5,6,25,30].

4.3. Model output uncertainty

Model output uncertainty represents the degree to which the diagnostic module lacks concentration and discriminability in its judgment of candidate faults. Let $p_t^{(1)}$ and $p_t^{(2)}$ be the largest and second largest posterior probabilities, respectively. The normalized probability entropy H_t and the class interval uncertainty D_t are defined as:'

$$H_t = -\frac{1}{\ln n} \sum_{j=1}^{n} p_{j,t} \ln p_{j,t} \tag{6}$$

$$D_t = 1 - \left(p_t^{(1)} - p_t^{(2)} \right) \tag{7}$$

H_t reflects the global dispersion across all candidate states, while D_t reflects the local competitive relationship between the largest and second-largest candidate states. The two are complementary but are not required to be statistically independent. They are defined as follows:

$$U_{model,t} = \lambda_H H_t + \lambda_D D_t, \quad \lambda_H + \lambda_D = 1, \quad \lambda_H, \lambda_D \geq 0 \tag{8}$$

$U_{model,t} \in [0,1]$A larger value indicates that the diagnostic output is less concentrated and that candidate states are more difficult to distinguish [30,31].

4.4. State information uncertainty

State information uncertainty refers to the extent to which, even after multi-source evidence has been made available, the key operating states remain difficult to confirm consistently. Let the set of key states be:

$$S_t = \left\{ s_{l,t} \mid l = 1,2,\ldots,N_s \right\} \tag{9}$$

Where the respective states denote the operating states of network topology, switch positions, relevant equipment, and distributed resources; only state variables that may alter fault diagnosis are included in this set. Let $r_{l,t} \in [0,1]$ be the degree of consistent confirmation for the l-th type of critical state, then:

Let $r_{l,t} \in [0,1]$ be the degree to which the l type of critical operating state can be consistently confirmed based on multi-source state evidence; then the state information uncertainty is defined as:

$$U_{state,t}=1-\sum_{l=1}^{4}\beta_l r_{l,t},\quad \sum_{l=1}^{4}\beta_l=1,\quad \beta_l \geq 0 \tag{10}$$

The larger $U_{state,t}$ is, the more difficult it becomes to confirm the critical state. $U_{data,t}$ focuses on evidence quality, whereas $U_{data,t}$ focuses on whether the available evidence can form a consistent state perception—for example, when multi-source information regarding the same switch position conflicts with one another [1,3,29,32]. See **Table 1**.

Table 1. Evaluation metrics for data evidence quality and state confirmation degree

Indicator	Evaluation source	Normalization principle
Evidence completeness	Ratio of information that should be connected to actually connected information	The higher the completeness rate, the higher the credibility
Evidence timeliness	Difference between information timestamp and diagnosis time	The smaller the delay, the higher the credibility
Evidence consistency	Degree of logical consistency between homogeneous or heterogeneous evidence	The fewer the conflicts, the higher the credibility
Source reliability	Source levels such as supervisory control and data acquisition (SCADA), protection recording, on-site feedback, and user repair requests	Authoritative sources have higher credibility
Status confirmation degree $r_{l,t}$	Degree of consistency among remote signaling, operation records, on-site feedback, and equipment status	Approaches 1 when consistent, and approaches 0 when conflicting

4.5. Integrated diagnostic uncertainty

The weighted fusion of $U_{data,t}$, $U_{model,t}$, and $U_{state,t}$ is performed as follows:

$$U_{D,t}=\alpha_d U_{data,t}+\alpha_m U_{model,t}+\alpha_s U_{state,t},$$

$$\alpha_d+\alpha_m+\alpha_s=1,\quad \alpha_d,\alpha_m,\alpha_s \geq 0 \tag{11}$$

The weights can be determined from historical samples; when samples are insufficient, they can be set by combining operating procedures, expert experience, and sensitivity analysis. In a typical scenario, $\alpha_d=\alpha_m=\alpha_s=1/3$ which is used only to illustrate the logical response of the authority allocation mechanism and does not imply that the three types of factors have equal influence. When any key indicator reaches or exceeds the corresponding safety threshold, a mandatory downgrade is applied in the subsequent authority determination.

5. Diagnostic uncertainty—error consequence-driven human-machine dynamic authority allocation decision model

5.1. Generation of candidate fault handling schemes

After obtaining the candidate fault set F_t and the operating state, the AI decision-making module generates or recommends a candidate scheme set:

$$A_t=\{a_1,a_2,\ldots,a_M\} \tag{12}$$

Where: a_i denotes the i candidate scheme, and M is the number of schemes, which may include operations such as fault isolation, tie-line transfer supply, network reconfiguration, distributed generation or

energy storage support, and load restoration. Schemes entering subsequent evaluation should satisfy basic constraints including network connectivity, radial operation, power flow, voltage and current limits, equipment capacity, and operation sequence. For F_t this paper constructs a unified A_t to evaluate the handling loss $L(a_i,f_j)$ of the same scheme under different true fault states f_j [1–4,29,32].

5.2. Quantification of incorrect handling consequences

The handling loss of scheme a_i under the true fault state f_j is defined as:

$$L_{ij}=L(a_i,f_j)=\gamma_1 L_{ij}^{safe}+\gamma_2 L_{ij}^{load}+\gamma_3 L_{ij}^{viol}+\gamma_4 L_{ij}^{op}+\gamma_5 L_{ij}^{time},$$

$$\sum_{r=1}^{5}\gamma_r=1,\quad \gamma_r \geq 0 \tag{13}$$

The respective components represent personal and equipment safety, loss of supply to critical loads and outage propagation, operational limit violations, incorrect switching or isolation, and restoration delay losses, all normalized to [0,1], so that Lij ∈ [0,1]*Lij* ∈ [0,1]. Lij*Lij* is a comprehensive handling loss index and is not equivalent to precise economic loss. When a scheme does not match the true state, incorrect isolation, failure to clear the fault, or transferring supply with the fault present will increase the loss. Major safety consequences are separately constrained by the subsequent safety red line [1–3,32]. See **Table 2**.

Table 2. Remedial action loss components and their quantification basis

Loss component	Engineering implication	Quantification principle
L_{ij}^{safe}	Safety consequences for personnel and equipment	Mapping of safety consequence levels
L_{ij}^{load}	Loss of supply to critical loads and outage expansion	Normalization of weighted unsupplied load
L_{ij}^{viol}	Voltage and current limit violations	Normalization of the maximum limit violation degree
L_{ij}^{op}	Incorrect switching or incorrect isolation	Mapping of incorrect operation severity
L_{ij}^{time}	Fault recovery delay	Normalization relative to reference recovery time

5.3. Decision risk calculation

The decision risk of the candidate scheme a_i at time t is defined as:

$$R_{i,t}=\sum_{j=1}^{n}p_{j,t}L_{ij}=\sum_{j=1}^{n}P(f_j|\ x_t)L(a_i,f_j) \tag{14}$$

$R_{i,t}$ represents the expected handling loss of executing a_i under the current posterior probability distribution, connecting diagnostic results with the physical consequences of the scheme. Unlike $U_{D,t}$ which describes diagnostic credibility, $R_{i,t}$ may still differ among different schemes under the same $U_{D,t}$. Low-probability major safety consequences are constrained by the safety red line in Section 4.4 [1–3,32].

5.4. Allowed authority set

To first exclude decision authorities that do not match the risk level of candidate schemes, this paper sets three-level authority risk tolerance thresholds based on Ri,t*Ri,t* to construct the basic allowed authority set for candidate scheme a_i :

$$\Omega_{i,t}^{R}=\{L_1 \in L \mid R_{i,t} \le \tau_1\}, \qquad 0 \le \tau_1<\tau_2<\tau_3=1 \tag{15}$$

Where L_1、L_2 and L_3 denote machine-dominated, AI-recommended–human-confirmed, and human-dominated–AI-assisted authority modes, respectively. It should be noted that L_1、L_2 and L_3 merely represent different human–machine authority modes and do not imply that the degree of machine autonomous execution increases with the numbering; in terms of the degree of machine autonomous execution, the three decrease successively. Correspondingly, as the scheme decision risk increases, higher autonomous authority such as machine dominance is progressively excluded, while human-dominated authority is always retained. Since $R_{i,t} \in [0,1]$ and $\tau_3=1$, $\Omega_{i,t}^{R}$ contains at least L_3.

Considering that comprehensive evaluation may weaken the impact of a single severe anomaly, for any case in which a key diagnostic uncertainty reaches or exceeds the corresponding safety threshold, the uncertainty safety constraint set is defined as:

$$\Omega_{i,t}^{U}=\begin{cases} L \setminus \{L_1\}, \exists k \in \{data, model, state\}, U_{k,t} \ge \theta_k \\ L, otherwise \end{cases} \tag{16}$$

Let the maximum safety consequence that a scheme may cause be:

$$S_{i,t}=\max_{j=1,?,n} L_{ij}^{safe}, \qquad \Omega_{i,t}^{S}=\begin{cases} \{L_3\}, S_{i,t} \ge \theta_{safe} \\ L, otherwise \end{cases} \tag{17}$$

The final permitted authority set is defined as:

$$\Omega_{i,t}=\Omega_{i,t}^{R} \cap \Omega_{i,t}^{U} \cap \Omega_{i,t}^{S} \tag{18}$$

θ_{safe} is determined according to operating procedures, equipment safety levels, and critical load assurance requirements. This paper only explains the constraining role of the safety red line and does not perform engineering statistical calibration. The final permitted authority set $\Omega_{i,t}$ represents the range of authority options permitted, rather than the final actual authority mode, and serves as a constraint for subsequent dynamic authority allocation decisions [8–16].

5.5. Criteria for dynamic authority allocation

Within the final permitted authority set $\Omega_{i,t}$, this paper determines the actual authority mode based on integrated diagnostic uncertainty, operational reversibility of candidate schemes, and operator intervention status:

$$m_{i,t}^{*}=g\left(U_{D,t}, v_i, h_t \mid \Omega_{i,t}\right), \qquad m_{i,t}^{*} \in \Omega_{i,t} \tag{19}$$

where $v_i \in [0,1]$ denotes the degree of operational reversibility; a larger value indicates that the system can more readily return to a safe operating state after execution; $h_t \in \{0,1\}$ indicates whether operating personnel can promptly confirm and take over. $h_t=0$ shall not serve as a basis for upgrading a scheme that originally requires manual confirmation or human leadership to machine leadership. For machine-led schemes that already satisfy Eq. (20) and fall within the pre-authorized safety boundary, the temporary inability of personnel to intervene in a timely manner does not alone constitute a condition for authority escalation. When a scheme requires manual confirmation or human leadership and $h_t=0$, the relevant operation should be suspended or the safe state maintained.

Let θ_D and θ_v be the upper limit of diagnostic uncertainty permitted for machine leadership and the minimum degree of operational reversibility, respectively. When the following expression is simultaneously satisfied, machine-led authority may be selected:

$$L_1 \in \Omega_{i,t}, \qquad U_{D,t} \le \theta_D, \qquad v_i \ge \theta_v \tag{20}$$

Equation (19) presents the overall mapping relationship for the actual authority mode, and Equation (20) further specifies the explicit executable conditions for the machine-dominated authority L1. For situations in which the machine-dominated conditions are not satisfied, the authority branches—including AI recommendation–human confirmation, human-dominated–AI-assisted, and deferred execution—are determined according to **Table 3**.

Table 3. Dynamic Human–AI Authority Allocation Rules

Condition	Actual authority mode
$\Omega_{i,t}=\{L_3\}$, or triggering the red line of major safety consequences	L_3, human-led with AI assistance
Satisfies Eq. (20)	L_1, machine-led
Does not satisfy Eq. (20), but $L_2 \in \Omega_{i,t}$ and $h_t=1$	L_2, AI recommendation with human confirmation
Requires human confirmation or human-led execution, but $h_t=0$	Suspend execution or maintain a safe state

Low UD,t*UD*,*t* does not necessarily permit machine dominance; high decision risk, low operational reversibility, or the safety red line can all restrict machine authority [8–16].

5.6. Scheme optimization for handling under constraints

Executable schemes are screened under the current authority and personnel conditions:

a_i meets the execution conditions under the authority mode $m^*_{i,t}$

$$A_t^{exe}=\left\{a_i \in A_t \mid a_i meets the execution conditions under the authority mode m^*_{i,t}\right\} \tag{21}$$

When manual confirmation or dominance is required and ht=0ht=0, the scheme does not enter Atexe*Atexe*. The final scheme is:

$$a_t^* = \arg\min_{a_i \in A_t^{exe}} J_{i,t} \tag{22}$$

where:

$$J_{i,t}=\eta_R R_{i,t}+\eta_{sw}\overline{N}_i^{sw}, \qquad \eta_R+\eta_{sw}=1, \quad \eta_R,\eta_{sw}\ge 0 \tag{23}$$

$$\overline{N}_i^{sw}=\frac{N_i^{sw}}{N_{max,t}^{sw}}, \qquad N_{max,t}^{sw}=\max_{a_i \in A_t^{exe}} N_i^{sw} \tag{24}$$

N_i^{sw} is the number of switching operations required for normal execution, which is distinct from the erroneous operation loss L_{ij}^{op}. $J_{i,t}$ is used only for secondary optimization after basic feasibility screening, risk assessment, and authority configuration, and does not replace power flow, topology, or operational safety verification [1–3,29,32].

6. Manual takeover and feedback update

6.1. Trigger conditions for manual takeover

When the permitted authority set retains only L_3, or when the dynamic authority allocation decision requires substantive judgment by operators, the system shifts to a human-led mode. Manual takeover can also be triggered by missing critical evidence, abnormal diagnostic module behavior, obvious conflict between human and machine judgments, major safety consequences, and low-reversibility operations. After takeover, the AI ceases autonomous execution and retains only functions such as information summarization, risk alerts, scheme comparison, and constraint verification, while the operators bear the final authorization responsibility [8–16].

6.2. Manual evidence supplementation and re-diagnosis

After manual takeover, operators supplement evidence concerning key uncertain links, including field inspections, user repair requests, protection recordings, and equipment and switching states, thereby forming updated observation information x_{t+1}. The diagnostic module recalculates:

$$p_{j,t+1}=P(f_j \mid x_{t+1}) \tag{25}$$

Based on this, the integrated diagnostic uncertainty, decision risk, and authority configuration are updated, and operators confirm or revise the diagnostic conclusion according to the updated results [5,16].

6.3. Feedback on handling results

After the handling is completed, information such as the actual fault, final scheme, execution and power supply restoration results, and reasons for manual override or takeover is recorded. This information is used to evaluate the consistency among diagnostic results, decision risks, and actual handling outcomes, and to identify misdiagnosis, risk underestimation, and unnecessary takeover, thereby providing a basis for parameter calibration [13,16].

6.4. Authority parameter update

Based on historical handling results, the diagnostic inaccuracy rate, manual override rate, takeover frequency, and actual handling losses are periodically calculated. Offline calibration is performed on uncertainty weights, risk thresholds, safety thresholds, and machine-dominance criterion parameters. The updated parameters are used for subsequent handling after manual review and are not adjusted in real time during the current fault process. Manual authorization rules involving major personal or equipment safety must not be automatically relaxed based on historical data. This forms a controlled parameter update closed loop of "offline statistics—manual review—subsequent parameter update" [8–10,13].

7. Illustration of dynamic authority allocation logic in a typical scenario

7.1. Scenario construction and parameter setting

This paper constructs a fault isolation and service restoration scenario based on the IEEE 33-bus radial distribution network as the network background, to illustrate the impact of changes in diagnostic uncertainty and scheme decision risk on authority configuration. Studies on fault section location, automatic reclosing, and fault repair–restoration coordination have shown that fault sections, switching states, and action

strategies affect isolation, restoration, and subsequent handling [33–35]. The posterior probabilities of candidate faults serve as output parameters of the diagnostic module interface, and the recognition accuracy of a specific AI algorithm is not evaluated. Model parameters are divided into three categories according to their basis of determination: statistical calibration parameters, procedural constraint parameters, and mechanism validation parameters. Uncertainty and loss weights can be determined based on historical samples, expert evaluation, and sensitivity analysis, while safety thresholds and manual authorization boundaries are set according to operating procedures. The scenario parameters in this paper belong to mechanism validation parameters and are used only to illustrate the response characteristics of dynamic authority allocation rules, rather than serving as actual engineering setting values. See **Table 4**.

Table 4. Typical IEEE 33-bus scenarios and corresponding fault–remedial action mapping

Scenario	Candidate faults and main disposal operations	Scenario parameters and weighting results
S1: Low uncertainty – low consequence	Branch fault at nodes 12–13 (f_1,$p_{1,t}$=0.90) or branch fault at nodes 13–14 (f_2,$p_{2,t}$=0.10); local isolation and limited transfer supply via tie switch 18–33.	General load, $_{D,t}$=0.178，$R_{i,t}$=0.15 v_i=0.8, L_1。
S2: Low uncertainty – high consequence	Candidate faults and probabilities same as S1; after isolation, important load transfer supply is implemented via tie switch 18–33.	$U_{D,t}$=0.178，$R_{i,t}$=0.55 v_i=0.80,L_2
S3: High uncertainty – low consequence	Branch fault at nodes 8–9 (f_1,$p_{1,t}$=0.70) or branch fault at nodes 9–10 (f_2,$p_{2,t}$=0.30); local isolation, with tie switch 9–15 as a candidate restoration path.	$U_{D,t}$=0.314，$R_{i,t}$=0.13 v_i=0.80， L_2。
S4: High uncertainty – high consequence	Candidate faults same as S3, $p_{1,t}$=0.60, $p_{2,t}$=0.40; after isolation, important load transfer supply is implemented via tie switch 9–15.	$U_{D,t}$=0.429，$R_{i,t}$=0.82 v_i=0.30， L_3。

Note: The scenario parameters are mechanism-illustrative settings.

7.2. Low uncertainty–low consequence scenario

Take (p_1,p_2)=(0.90,0.10)、$U_{data,t}$=$U_{state,t}$=0.10, yielding H_t=0.469、D_t=0.20 $U_{model,t}$=0.335 5, and thus $U_{D,t}$=(0.10+0.335+0.10)/3=0.178. Take L_{i1}=0.10、L_{i2}=0.60, yielding $R_{i,t}$=0.15. Since $R_{i,t}$<τ_1、$U_{D,t}$<θ_D, and the reversibility condition is satisfied, machine-dominant authority L_1 is adopted.

7.3. Low uncertainty–high consequence scenario

Keeping the S1 diagnostic condition and v_i=0.80 unchanged, take L_{i1}=0.50、L_{i2}=1.00, yielding $R_{i,t}$=0.55. Machine dominance is excluded; when h_t=1 and no safety red line is triggered, the AI recommendation–human confirmation authority L_2 is adopted. Compared with S1, $U_{D,t}$ remains unchanged while risk increases, indicating that diagnostic certainty does not necessarily mean suitability for autonomous machine handling.

7.4. High uncertainty–low consequence scenario

Take (p_1,p_2)=(0.70,0.30)、$U_{data,t}$=$U_{state,t}$=0.10, yielding H_t=0.881、D_t=0.60 $U_{model,t}$=0.741 $U_{D,t}$=0.314. Take L_{i1}=0.10、L_{i2}=0.20, yielding $R_{i,t}$=0.13. Although decision risk is low, diagnostic uncertainty has exceeded the machine-dominant threshold; when h_t=1, the AI recommendation–human confirmation authority L_2 is adopted. Compared with S1, high diagnostic uncertainty can likewise restrict autonomous machine execution.

7.5. High uncertainty–high consequence scenario

Take $(p_1,p_2)=(0.60,0.40)$、$U_{data,t}=U_{state,t}=0.20$, yielding $H_t=0.971$、$D_t=0.80$ $U_{model,t}=0.886$ $U_{D,t}=0.429$. Take $L_{i1}=0.70$、$L_{i2}=1.00$ $v_i=0.30$, yielding $R_{i,t}=0.82$. From Eq. (15), $\Omega_{i,t}^{R}=\{L_3\}$ is obtained, and after intersection via Eq. (18), it finally remains $\Omega_{i,t}=\{L_3\}$. The system enters the human-dominant–AI-assisted mode and can be re-evaluated after supplementing evidence according to Section 5.2.

7.6. Comparison of different decentralization strategies

Using S1–S4 as unified conditions, compare full human handling, fixed AI recommendation, authorization solely by the $U_{D,t}$ threshold, and the proposed method. Full human handling always adopts L_3; fixed AI recommendation always adopts L_2; when authorization is based solely on uncertainty, L_1 is adopted if $U_{D,t} \le \theta_D$; otherwise, L_2 is adopted. See **Table 5**.

Table 5. Authority allocation results of different human–AI decision strategies under typical scenarios

Scenario	UD,t	Ri,t	Manual full-process handling	AI fixed recommendation	UD threshold only	Proposed method
S1	0.178	0.15	L3	L2	L1	L1
S2	0.178	0.55	L3	L2	L1	L2
S3	0.314	0.13	L3	L2	L2	L2
S4	0.429	0.82	L3	L2	L2	L3

Table 5 indicates that fixed authority strategies struggle to simultaneously adapt to diagnostic uncertainty and variations in handling consequences. In S2, authorization based solely on $U_{D,t}$ still retains L_1, ignoring the relatively high decision risk; S3 shows that low decision risk cannot offset high diagnostic uncertainty; in S4, the two types of factors act together, and the authority is reduced to L_3.

This comparison only illustrates the logical differences in authority configuration among different rules, and does not demonstrate statistical superiority in restoration efficiency, safety performance, or economic performance.

8. Conclusions and prospects

This paper addresses the problem that machine handling authority is difficult to allocate reasonably under conditions of uncertainty in distribution network fault diagnosis, and proposes a human-machine dynamic authority allocation decision model. The model quantifies integrated diagnostic uncertainty from data evidence, model output, and state information, combines the posterior probabilities of candidate faults with the handling losses of schemes under different true states to construct scheme-level decision risk; determines permitted authority, actual authority, and final schemes based on risk thresholds, safety constraints, operational reversibility, and personnel intervention conditions, and forms a feedback closed loop through supplementary evidence, re-diagnosis, and offline parameter calibration. Comparisons of constructed scenarios and strategies show that machine-dominant mode is adopted under low uncertainty and low risk; AI recommendation–human confirmation is adopted when only diagnostic uncertainty or decision risk limits machine autonomous authority; and human-dominant–AI assistance is adopted when risk thresholds further restrict authority under high uncertainty and high risk scenarios. Diagnostic credibility does not equate to suitability for autonomous machine handling; human-machine authority should be jointly determined by diagnostic uncertainty, scheme decision risk, and safety constraints.

This paper uses an IEEE 33-node distribution network as the network background to construct scenarios illustrating the decision logic, and the relevant weights and thresholds have not yet been statistically calibrated using actual handling data. Future work can calibrate model parameters by combining real faults, dispatch operations, and operator interaction records, test its adaptability under different grid structures, distributed resource penetration levels, and operating procedures, and evaluate its impact on handling time, frequency of manual intervention, and risk control through digital simulation, training simulation, and field exercises.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Zhang Q, Ma Z, Zhu Y, et al., 2021, A Two-Level Simulation-Assisted Sequential Distribution System Restoration Model With Frequency Dynamics Constraints. IEEE Transactions on Smart Grid, 12(5): 3835–3846.

[2] Yang X, Xu C, Wen J, et al., 2023, Cooperative Repair Scheduling and Service Restoration for Distribution Systems With Soft Open Points. IEEE Transactions on Smart Grid, 14(3): 1827–1842.

[3] Wang G, Lei X, Wu H, et al., 2021, A Comprehensive Network Restoration Model for Active Distribution Network Considering Forecast Uncertainty. IEEE Access, 9: 130997–131005.

[4] Heidari-Akhijahani A, Butler-Purry K, 2024, A Review on Black-Start Service Restoration of Active Distribution Systems and Microgrids. Energies, 17(1): 100.

[5] Yuan Y, Dehghanpour K, Wang Z, et al., 2022, Multisource Data Fusion Outage Location in Distribution Systems via Probabilistic Graphical Models. IEEE Transactions on Smart Grid, 13(2): 1357–1371.

[6] Gao Y, Su X, Zhang H, et al., 2024, A Highly Fault-Tolerant Distribution Network Fault Diagnosis Method Based on Data Check and Graph Convolution Neural Network. Advanced Technology of Electrical Engineering and Energy, 43(2): 95–104.

[7] Ji L, Tian X, Wei Z, et al., 2025, Intelligent Fault Diagnosis in Power Distribution Networks Using LSTM-DenseNet Network. Electric Power Systems Research, 239: 111202.

[8] Parasuraman R, Sheridan T, Wickens C, 2000, A Model for Types and Levels of Human Interaction With Automation. IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 30(3): 286–297.

[9] Lee J, See K, 2004, Trust in Automation: Designing for Appropriate Reliance. Human Factors, 46(1): 50–80.

[10] Santoni de Sio F, Van den Hoven J, 2018, Meaningful Human Control Over Autonomous Systems: A Philosophical Account. Frontiers in Robotics and AI, 5: 15.

[11] Westphal M, Vössing M, Satzger G, et al., 2023, Decision Control and Explanations in Human-AI Collaboration: Improving User Perceptions and Compliance. Computers in Human Behavior, 144: 107714.

[12] Johnson C, Demir M, McNeese N, et al., 2023, The Impact of Training on Human-Autonomy Team Communications and Trust Calibration. Human Factors, 65(7): 1554–1570.

[13] Chiou E, Lee J, 2023, Trusting Automation: Designing for Responsivity and Resilience. Human Factors, 65(1): 137–165.

[14] Jui J, Hettiarachchi I, Bhatti A, et al., 2025, A Recent Review on Subjective and Objective Assessment of Trust in

Human Autonomy Teaming. IEEE Transactions on Human-Machine Systems, 55(5): 819–833.

[15] Van den Bosch K, Van Diggelen J, Verdult S, et al., 2025, Measuring Meaningful Human Control in Human-AI Teaming: Effects of Team Design in AI-Assisted Pandemic Triage. AI and Ethics, 5(3): 3329–3353.

[16] McGrath M, Duenser A, Lacey J, et al., 2025, Collaborative Human-AI Trust (CHAI-T): A Process Framework for Active Management of Trust in Human-AI Collaboration. Computers in Human Behavior: Artificial Humans, 6: 100200.

[17] Srivastava D, Lilly J, Feigh K, 2025, Exploring the Role of Judgement and Shared Situation Awareness When Working With AI Recommender Systems. Cognition, Technology & Work, 27(1–2): 1–18.

[18] Lin J, Tomlin N, Andreas J, et al., 2024, Decision-Oriented Dialogue for Human-AI Collaboration. Transactions of the Association for Computational Linguistics, 12: 892–911.

[19] Liang K, Wang H, Pozo D, et al., 2024, Power System Restoration With Large Renewable Penetration: State-of-the-Art and Future Trends. International Journal of Electrical Power & Energy Systems, 155: 109494.

[20] Hu Y, Qian F, Li H, et al., 2024, Island Fault Recovery Method of Distribution Network Based on Flexible Interconnection Device With Energy Storage. Distribution & Utilization, 41(2): 21–27.

[21] Saaklayen M, Liang X, Faried S, et al., 2024, Soft Open Point-Based Service Restoration Coordinated With Distributed Generation in Distribution Networks. IEEE Transactions on Industry Applications, 60(2): 2554–2566.

[22] Sun L, Huang Z, Yi K, et al., 2024, Coordinated Optimization of Repair Scheduling and Service Restoration for Distribution Network Considering the Damaged Roads. IEEE Transactions on Industry Applications, 60(4): 5407–5422.

[23] IEC 60870-5-104, 2016, Telecontrol Equipment and Systems—Part 5-104: Transmission Protocols—Network Access for IEC 60870-5-101 Using Standard Transport Profiles.

[24] Wu X, Zhao H, Xu W, et al., 2024, Fault Diagnosis of the Distribution Network Based on the D-S Evidence Theory Bayesian Network. Frontiers in Energy Research, 12: 1422639.

[25] Ibrahim A, Sadanandan S, Ghaoud T, et al., 2024, Incipient Fault Detection in Power Distribution Networks: Review, Analysis, Challenges, and Future Directions. IEEE Access, 12: 112822–112838.

[26] Guo Z, Ji X, Wang H, et al., 2024, Active Distribution Network Fault Diagnosis Based on Improved Northern Goshawk Search Algorithm. Electronics, 13(7): 1202.

[27] Xu C, Cheng X, Zhuang X, et al., 2025, Fault Diagnosis of Distribution Network Based on Time Constraint Intuition Fuzzy Petri Nets. Measurement: Energy, 5: 100034.

[28] Xie D, Xu Y, Zhang C, 2023, Robust Service Restoration of Distribution Systems Towards Coordinated Parallel Network Reconstruction and Cold Load Pick-Up. International Journal of Electrical Power & Energy Systems, 152: 109187.

[29] Igder M, Liang X, Mitolo M, 2022, Service Restoration Through Microgrid Formation in Distribution Networks: A Review. IEEE Access, 10: 46618–46632.

[30] Gawlikowski J, Tassi C, Ali M, et al., 2023, A Survey of Uncertainty in Deep Neural Networks. Artificial Intelligence Review, 56(Suppl 1): 1513–1589.

[31] Shannon C, 1948, A Mathematical Theory of Communication. Bell System Technical Journal, 27(3): 379–423.

[32] Baran M, Wu F, 1989, Network Reconfiguration in Distribution Systems for Loss Reduction and Load Balancing. IEEE Transactions on Power Delivery, 4(2): 1401–1407.

[33] Ji X, Zhang S, Zhang Y, et al., 2021, Fault Section Location for Distribution Network Based on Improved Electromagnetism-Like Mechanism Algorithm. Automation of Electric Power Systems, 45(22): 157–166.

[34] Saad M, Kim C, Munir N, 2022, Single-Phase Auto-Reclosing Scheme Using Particle Filter and Convolutional Neural Network. IEEE Transactions on Power Delivery, 37(6): 4775–4785.

[35] Yang L, Qin Y, 2023, Research on Rolling Co-Optimization of Fault Repair and Service Restoration in Distribution Network Based on Combined Drive Methodology. Electric Power Systems Research, 220: 109373.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.