RESEARCH ARTICLE

Reliability Evaluation of Social Systems Against Cognitive Warfare: A Hybrid Approach Based on Extended Fault Tree and Multi-Valued Decision Diagram

Yuchang Mo[1,2,*], **Chunyu Miao**[1,3], **Guangxuan Chen**[4], **Tiantian Gai**[1], **Wen Wu**[1] and **Yuhuan Gong**[5]

[1] School of Computer Science and Technology, Zhejiang University of Water Resources and Electric Power, Hangzhou 310018, China
[2] Zhejiang-Kyrgyzstan Joint Laboratory on Artificial Intelligence and Clean Energy, Hangzhou 310018, China
[3] Hangzhou Anheng Information Technology Co., Ltd., Hangzhou, China
[4] Department of Computer and Information Security, Zhejiang Police College, Hangzhou 310053, China
[5] School of Marxism, Zhejiang University of Water Resources and Electric Power, Hangzhou 310018, China

Abstract

As an emerging non-traditional security threat, cognitive warfare undermines social systems by manipulating information dissemination, distorting cognitive consensus, and eroding social trust. The reliability of social systems against cognitive warfare reflects their ability to maintain core functions, resist cognitive infiltration, and restore social order. To quantitatively evaluate this reliability, this paper proposes a hybrid approach integrating the Extended Fault Tree (EFT) and Multi-Valued Decision Diagram (MDD). First, key failure paths and interdependencies are identified, including cascading failures among four core subsystems: information dissemination, cognitive consensus, social trust, and emergency response. Second, EFT models the logical relationships among cognitive attack sources, propagation nodes, and failure consequences, while multi-state nodes and cascade gates characterize graded failures and cascading effects. MDD is then employed to efficiently solve the complex logical model derived from EFT, balancing analytical accuracy and computational efficiency. An empirical case study of cognitive warfare response during a sudden public event verifies the effectiveness and applicability of the proposed approach. The results demonstrate that the hybrid approach provides quantitative decision support for optimizing social cognitive defense systems and offers a methodological reference for national cognitive security governance.

Keywords: social system, cognitive warfare, reliability analysis, extended fault tree (EFT), multi-valued decision diagram (MDD), cascading failure.

1 Introduction

With the deep convergence of globalization and digitalization, the cognitive domain has emerged as

Submitted: 11 June 2026
Revised: 16 August 2026
Accepted: 17 August 2026
Published: 01 October 2026

Vol. 2, **No.** 4, 2026.
10.62762/TSSR.2026.783316

***Corresponding author:**
✉ Yuchang Mo
yuchangmo@sina.com

Citation
Mo, Y., Miao, C., Chen, G., Gai, T., Wu, W., & Gong, Y. (2026). Reliability Evaluation of Social Systems Against Cognitive Warfare: A Hybrid Approach Based on Extended Fault Tree and Multi-Valued Decision Diagram. *ICCK Transactions on Systems Safety and Reliability*, 2(4), 220–234.

© 2026 ICCK (Institute of Central Computation and Knowledge)

a pivotal arena in international security competition. Unlike conventional military engagements that occur within the physical realm, cognitive warfare targets the mental frameworks, value systems, and decision-making processes of individuals, seeking to destabilize social systems from within. This is achieved through the deliberate generation and dissemination of disinformation, the provocation of intergroup tensions, and the subtle infiltration of ideological narratives [1]. Recent events—such as the information and psychological confrontations witnessed in the Russia–Ukraine conflict, the widespread circulation of falsehoods during the global COVID-19 pandemic, and manipulative tactics observed in national elections—have underscored the formidable disruptive potential of cognitive warfare. Beyond inducing public panic and obstructing policy execution, such campaigns can erode social trust, undermine national sovereignty, and pose existential threats to the sustainability of societal structures [2].

Existing studies on cognitive warfare commonly focus on three aspects: operational mechanisms, dissemination pathways, and defensive strategies. Mechanism research draws on psychology, communication, and sociology: cognitive and social psychology show how systematic heuristics/biases (e.g., biased judgment under uncertainty), confirmation bias, and conformity pressures can be exploited to steer information processing and collective judgments [3, 4]; communication research indicates that false information can diffuse faster and more broadly online, often benefiting from novelty and engagement dynamics [5]; and sociotechnical perspectives emphasize that cognitive warfare can exacerbate identity-based divisions and undermine social cohesion, reducing collective action capacity [6]. Dissemination-path studies often model diffusion using complex networks, where scale-free and small-world properties imply that highly connected hubs or influential nodes can disproportionately accelerate spread and amplify impact; decision-diagram-based methods have further been developed to analyze the reliability and complexity of such multi-state networked systems [7, 8]; meanwhile, platform recommendation systems and selective exposure dynamics may reinforce echo chambers or filter-bubble-like patterns for some users and contexts, increasing exposure to bias-consistent narratives and potentially intensifying polarization [9]; new decision-diagram-based approaches have also been proposed to efficiently evaluate the reliability of such complex multi-state dissemination systems [10]. Defensive strategy research typically spans technical, analytical, and educational approaches: on the technical side, machine-learning-based fake news detection synthesizes content, propagation, and social-context signals for automated identification and mitigation [11]; on the analytical side, quantitative methods such as fault tree analysis integrated with fuzzy numbers and TOPSIS [12], dynamic fault tree analysis combined with Monte Carlo simulation for probabilistic safety assessment [13], and reliability and resilience modeling of cascading failures in networked systems [14] provide rigorous tools for evaluating and mitigating the systemic risks posed by cognitive warfare; and on the educational side, media and information literacy frameworks (e.g., UNESCO curricula) aim to improve critical evaluation skills and reduce susceptibility to manipulation at the source [15].

A social system constitutes a dynamic and complex organism shaped by the interplay of individuals, organizations, institutions, and cultural norms. Its resilience to cognitive warfare directly influences national security and societal harmony. Compared with engineered systems, social systems exhibit distinctive complexities: they are open, evolving, nonlinear, and heavily influenced by human agency. Cognitive attacks frequently originate from ambiguous sources, propagate along unpredictable pathways, and exploit heterogeneous cognitive baselines, values, and behavioral tendencies among members of society. Moreover, subsystems within the social fabric are interconnected through multilayered and networked dependencies. Together, these attributes give rise to intricate failure modes under cognitive assault.

Hybrid methodologies combining Fault Tree Analysis (FTA) and Decision Diagrams have gained prominence in engineering reliability studies owing to their robust logical structuring and computational efficiency. FTA employs a top-down approach to establish causal relationships and trace failure propagation pathways, offering intuitive visualization and rigorous deductive reasoning. Decision diagrams—such as Binary Decision Diagrams (BDDs) and Multi-state Decision Diagrams (MDDs)—transform these logical structures into graphical models, enabling efficient evaluation of complex expressions and mitigating the combinatorial explosion that often hampers traditional FTA [16, 17]. For instance, Mo

et al. [18] developed an MDD-based approach for efficient reliability analysis of repairable computing systems subject to scheduled checkpointing, and Mo and Xing [19] developed an MDD-based approach for efficient resource availability analysis of cloud computing systems to reduce SLA violations. Xing and Amari [20] utilized Extended Fault Trees (EFTs) to represent both common-cause and cascading failures, coupling them with MDDs to solve large-scale models efficiently, thus broadening the applicability of hybrid methods.

To address these challenges, this paper proposes a novel hybrid EFT–MDD framework for the quantitative assessment of social system reliability under cognitive warfare. The proposed approach integrates the expressive power of EFTs in modeling complex, graded failure processes with the computational efficiency of MDDs in handling large logical models. The remainder of this paper is organized as follows: Section 2 defines the research scope and problem, detailing the structure and failure mechanisms of social systems under cognitive attack. Section 3 presents the design of the EFT–MDD hybrid approach, including its basic idea, failure factor identification and quantification, EFT construction, and MDD-based solution procedures. Section 4 applies the proposed approach in an empirical case study, validating its effectiveness and performing sensitivity analysis to pinpoint critical reliability determinants. Finally, Section 5 summarizes the principal findings and outlines directions for future research.

2 System Description and Problem Definition

2.1 System Description

The social system exposed to cognitive warfare constitutes a complex, adaptive entity shaped by the dynamic interactions and interdependencies among diverse components. Functionally, it can be disaggregated into four core subsystems, each operating with a degree of autonomy while remaining tightly integrated within the broader system architecture. 1) The information dissemination subsystem encompasses official media outlets, commercial social media platforms, self-publishing content creators (we-media), and professional journalism organizations. Its primary responsibilities include the issuance, propagation, and verification of information flows. 2) The cognitive consensus subsystem is composed of individuals, collective groups, civil society organizations, and cultural norms, serving as the foundation for the formation and maintenance of shared societal values and mutual understandings. 3) The social trust subsystem integrates trust in governmental institutions, interpersonal relationships, and organizational credibility, underpinning cooperative behavior and enabling effective policy execution. 4) Finally, the emergency response subsystem consists of government emergency management agencies, specialized rescue units, non-governmental organizations, and media actors, tasked with monitoring potential threats, issuing early warnings, coordinating interventions, and facilitating post-incident recovery.

Cognitive warfare does not manifest as an isolated disruption but unfolds through a dynamic failure progression characterized by multi-source initiation, cascading propagation, and system-wide breakdown. Here, multi-source initiation represents the entry point of hostile influence activities, whereby cognitive attacks penetrate the social system via diverse channels and formats, triggering localized disturbances. Cascading propagation constitutes the central phase, wherein initial failures diffuse across subsystems through structural interconnections, generating sequential chain reactions. In practice, the four subsystems form an intricate web of dependencies; although this connectivity enhances operational efficiency under normal conditions, it simultaneously creates pathways for failure amplification. System-wide breakdown denotes the ultimate outcome, occurring when cumulative cascading failures surpass the resilience threshold of the social system. At this stage, essential functions become incapacitated, evidenced by the erosion of cognitive consensus, social fragmentation, and the paralysis of public policy implementation.

2.2 Problem Definition

The reliability of a social system in countering cognitive warfare serves as a critical metric for assessing its capacity to sustain essential functions under adversarial cognitive pressure. This concept must encapsulate both the distinctive nature of cognitive warfare and the inherent complexity of social systems. Formally, reliability can be defined as the probability that a social system will continuously uphold its core functions—including the maintenance of stable cognitive consensus, the preservation of orderly social relations, and the effective implementation of public policies—within a defined timeframe and in the presence of specified cognitive warfare threats.

Table 1. Notation of failure factor categories.

Symbol category	Meaning	Examples
A	Attack-source factors	A_1–A_3: Disinformation attack, ideological infiltration, social antagonism incitement
B	Subsystem failure factors	B_1–B_4: Information dissemination, cognitive consensus, social trust, emergency response failures
C	Propagation factors	C_1–C_3: Information diffusion, cognitive bias propagation, trust collapse propagation
D	Consequence factors	D_1–D_3: Social disorder, policy implementation failure, cognitive security damage
TE	Top event	Failure of social systems in responding to cognitive warfare

Five assumptions are made in this paper. First, the intensity and modality of cognitive warfare attacks are treated as known quantities. Second, each subsystem's operational status is categorized into three discrete states: normal, partial failure, and complete failure. Third, the analysis emphasizes direct inter-subsystem influence pathways—for example, a failure in the information dissemination subsystem may directly precipitate a failure in the cognitive consensus subsystem—thereby enhancing computational tractability. Fourth, the efficiency of emergency response actions is assumed to follow a predefined probability distribution. Fifth, for probability calculation, basic failure events without direct causal dependencies are assumed to be statistically independent. For dependent events caused by subsystem interactions and cascading propagation, conditional probabilities are adopted and represented through cascade gates in the EFT model.

3 The Hybrid Analysis Approach

3.1 Basic Ideas

The reliability analysis of social systems against cognitive warfare is essentially a probability evaluation problem of complex multi-state systems, and its core challenge is to depict the nonlinear failure relationship of multi-source triggering and cascading propagation under cognitive attacks. The hybrid EFT–MDD approach proposed in this paper is constructed based on the following three basic ideas:

1) The logical modeling method of fault tree analysis. This method decomposes system failure events layer by layer into subsystem failures and basic failure events through top-down logical relationship decomposition, which is suitable for clarifying the failure causal chain. In view of the multi-state failure characteristics of social systems in cognitive warfare scenarios, this paper extends the traditional fault tree by introducing multi-state nodes and cascade gates to adapt to the gradient failure mode of normal-partial failure-complete failure.

2) The efficient evaluation method of decision diagrams. As a generalization of BDD to multi-valued variables, MDD effectively addresses the computational bottleneck of combinatorial explosion in traditional fault tree analysis. The core advantage of this method is that it can convert complex logical expressions into compact graphical models, and realize efficient evaluation of reliability through path traversal.

3) The multi-attribute decision-making and probability estimation method. In order to handle the fuzziness and subjectivity of some failure parameters in social systems, Analytic Hierarchy Process (AHP) and expert scoring method are combined to convert qualitative evaluation into quantitative probability parameters, providing data support for reliability analysis. Meanwhile, drawing on the parameter estimation methods in engineering reliability, the probability distribution characteristics of each failure event are determined through historical data fitting and scenario simulation.

3.2 Identification and Quantification of System Failure Factors

Based on the structural composition and cognitive warfare failure characteristics of social systems, a four-level failure factor system including attack source factors, subsystem failure factors, propagation factors and consequence factors is constructed. The notation categories and representative variables used in the

proposed EFT model are summarized in Table 1, which helps distinguish different types of failure factors and their roles in the reliability analysis framework. Here, attack source factors focus on the external trigger conditions of cognitive warfare, including three types: disinformation attack (A_1), ideological infiltration (A_2) and social antagonism incitement (A_3). Subsystem failure factors cover the internal failure events of four core subsystems: information dissemination (B_1), cognitive consensus (B_2), social trust (B_3) and emergency response (B_4), such as screening failure (B_{11}) and official information lag (B_{12}) of the information dissemination subsystem. Propagation factors depict the failure propagation paths among subsystems, including information diffusion (C_1), cognitive bias propagation (C_2) and trust collapse propagation (C_3). Consequence factors correspond to the overall failure states of social systems, such as social disorder (D_1) and policy implementation failure (D_2).

For the probability quantification of system failure factors, three methods are used. For failure factors with historical data support (e.g., disinformation propagation rate, policy implementation success rate), the parameter estimation method is used to fit the probability distribution; for qualitative factors lacking historical data (e.g., the impact of insufficient public media literacy), the expert scoring method combined with triangular fuzzy numbers is used to convert into probability values; for the propagation probability among subsystems, it is calculated based on the conditional probability formula: let event X be the failure of the preceding subsystem and event Y be the failure of the subsequent subsystem; then the propagation probability is given by

$$P(Y|X) = P(X \cap Y)/P(X),$$

where $P(X \cap Y)$ is determined by joint expert evaluation. In the proposed framework, independent basic events are quantified using their marginal state probabilities, whereas dependent propagation events are quantified using conditional probabilities. Therefore, the dependency relationships among subsystems are explicitly incorporated into the EFT structure rather than ignored in the probability calculation.

3.3 Extended Fault Tree Modeling

The modeling of EFT follows the hierarchical decomposition logic of top event–intermediate event–bottom event. Here, the top event (TE) represents the failure of social systems in responding to cognitive warfare, defined according to the failure states of consequence factors, and its state is a binary variable (0 = normal, 1 = failure). The intermediate events (IE) include subsystem failure factors (B_1–B_4) and propagation factors (C_1–C_3), described by three-state variables: 0 = normal, 1 = partial failure, 2 = complete failure. Partial failure means that the subsystem functions are partially damaged but the core capabilities are not lost; complete failure means that the core functions of the subsystem are completely paralyzed. The bottom events (BE) include attack source factors (A_1–A_3) and internal failure factors of subsystems (B_{11}–B_{12}, etc.). Considering that cognitive warfare attacks may produce different levels of impact on social systems, the attack-source factors are modeled as three-state variables: 0 = no effective attack, 1 = partial impact, and 2 = complete impact. The internal failure factors of subsystems also use three-state variables (0 = normal, 1 = partial failure, 2 = complete failure). This multi-state representation enables the proposed model to capture the gradual escalation characteristics of cognitive warfare.

In addition to the traditional AND gate and OR gate, a cascade gate (CG) is introduced to depict the cascading failure propagation among subsystems. The input of the cascade gate is the failure state and weight of the preceding subsystem, and the output is the failure state of the subsequent subsystem. Its mathematical expression is:

$$Y = \begin{cases} 0, & \sum_{i=1}^{n} w_i X_i < T_1, \\ 1, & T_1 \le \sum_{i=1}^{n} w_i X_i < T_2, \\ 2, & \sum_{i=1}^{n} w_i X_i \ge T_2. \end{cases}$$

where X_i is the failure level (0–2) of the i-th input event, w_i is the weight of the i-th input event ($\sum w_i = 1$), T_1 and T_2 are the trigger thresholds for partial failure and complete failure respectively (set according to expert experience and scenario characteristics, e.g., $T_1 = 0.8$, $T_2 = 1.5$), and Y is the failure level (0–2) of the output event.

The multi-state extended expressions of traditional logical gates are as follows. For a multi-state AND gate (MAND), the output event reaches a certain failure level only when all input events reach or exceed that

failure level, i.e.,

$$Y = \min(X_1, X_2, \ldots, X_n).$$

For a multi-state OR gate (MOR), the output event reaches a certain failure level as long as one input event reaches or exceeds that failure level, i.e.,

$$Y = \max(X_1, X_2, \ldots, X_n).$$

A practical EFT modeling procedure consists of the following four steps. 1) Determine the top event: clarify the specific definition and judgment criteria for the failure of social systems in responding to cognitive warfare; 2) Decompose intermediate events: decompose the top event layer by layer into subsystem failure events and propagation factors; 3) Identify bottom events: sort out the internal failure events of each subsystem and cognitive attack sources, and form a list of bottom events; 4) Establish logical relationships: connect each event with multi-state AND gates, OR gates and cascade gates to form a complete EFT. For example, the failure of the information dissemination subsystem (B_1) is determined by attack source factors (A_1, A_2, A_3), information screening failure (B_{11}) and official information lag (B_{12}) through a multi-state OR gate, i.e.,

$$B_1 = \mathrm{MOR}(A_1, A_2, A_3, B_{11}, B_{12}).$$

The failure of the cognitive consensus subsystem (B_2) is determined by the failure of the information dissemination subsystem (B_1), information diffusion (C_1), insufficient media literacy (B_{21}) and weak cognitive consensus (B_{22}) through a cascade gate, i.e.,

$$B_2 = \mathrm{CG}(B_1, C_1, B_{21}, B_{22}; w_1, w_2, w_3, w_4; T_1, T_2).$$

3.4 Multi-Valued Decision Diagram Solution

MDD is a discrete function decomposition tool based on graphical representation, which is suitable for logical reasoning and probability calculation of multi-state systems. The core of converting EFT to MDD is to map the logical relationships of the fault tree into nodes and edges of MDD, where each non-leaf node of MDD corresponds to a failure event (variable), each edge corresponds to a state (0–2) of the variable, and the leaf nodes correspond to the final state of the system (0 = normal, 1 = failure).

The conversion process follows the recursive decomposition principle. For any logical expression $F(X_1, X_2, \ldots, X_n)$ (X_i is a multi-state variable), its MDD decomposition formula is:

$$F(X_1, X_2, \ldots, X_n) = \sum_{s_1=0}^{k_1-1} I(X_1 = s_1) \cdot F(s_1, X_2, \ldots, X_n)$$

where k_i is the number of states of variable X_i ($k_i = 2$ or 3 in this paper), $I(\cdot)$ is the indicator function ($I(\text{true}) = 1$, $I(\text{false}) = 0$), and $F(s_1, s_2, \ldots, s_n)$ is the final state (0 or 1) of the system when the variables take a specific state combination.

Each path from the root node to the leaf node "normal (0)" in MDD corresponds to a state combination representing normal system operation. For independent basic events, the probability of a path can be calculated by multiplying the corresponding state probabilities of the nodes along the path. For dependent events, the conditional probabilities embedded in the cascade gates are used during path probability evaluation. Therefore, the reliability of social systems in responding to cognitive warfare is obtained by summing the probabilities of all normal paths, and its mathematical expression is:

$$R = \sum_{p \in P_0} \prod_{v \in p} P(v = s_v)$$

where P_0 is the set of all paths from the root node to the leaf node "normal (0)", v is the node on path p, s_v is the state of node v on path p, and $P(v = s_v)$ is the probability that node v is in state s_v. For nodes involved in cascading dependencies, $P(v = s_v)$ is replaced by the corresponding conditional probability determined by the preceding events.

For the probability calculation of multi-state nodes, take the three-state variable X as an example, its state probability satisfies

$$P(X = 0) + P(X = 1) + P(X = 2) = 1$$

where $P(X = 0)$ is the normal probability of the node, $P(X = 1)$ is the partial failure probability, and $P(X = 2)$ is the complete failure probability, all of which are determined by the quantification method in Section 3.2.

Assume that the EFT contains n bottom events, including m three-state variables (attack source factors) and $n - m$ three-state variables (subsystem failure events and propagation factors). The number of MDD nodes is $O(k^n)$, where k is the average number

of states, $k = 3$ in this paper. The actual number of nodes can be significantly reduced through heuristic variable ordering. The time complexity of reliability calculation is $O(N)$, where N is the number of MDD nodes, because each node only needs to be traversed once to complete the probability accumulation.

4 Case Study

4.1 Case Description

This section examines a cognitive warfare incident that occurred during a regional public health emergency response in 2023, using it as an empirical case to evaluate the effectiveness of the proposed hybrid analytical approach. The selected scenario is particularly suitable for analysis due to its pronounced characteristics of cognitive aggression and the dynamic interplay between external disruption and systemic societal reactions. Amid the unfolding crisis, external adversaries capitalized on widespread public concern and uncertainty, launching coordinated influence operations across multiple digital platforms, including social media networks and transnational online forums. These campaigns exhibited a layered pattern involving the deliberate creation of false narratives, emotionally charged storytelling, and subtle ideological penetration. For instance, fabricated claims—such as assertions of ineffective containment strategies, imminent collapse of medical services, and manipulation of official statistics—were strategically amplified through edited multimedia content and anonymous sources to increase their perceived legitimacy. Simultaneously, these efforts sought to exacerbate public dissatisfaction with government measures, foster social division, obstruct emergency operations, and ultimately undermine the stability of the regional sociopolitical order.

The social system under investigation is structured around the region's public health emergency management framework and comprises four interrelated subsystems: information diffusion, cognitive alignment, institutional trust, and emergency coordination. Each component operates within clearly defined parameters and fulfills distinct functional roles. The information diffusion subsystem incorporates official government portals, state-affiliated digital communication channels, mainstream media outlets, dominant social media platforms, and specialized online sentiment monitoring entities. The cognitive alignment subsystem spans the general population, grassroots organizations, sector-specific associations, and academic expert communities within the region. Institutional trust forms the nucleus of the social trust subsystem, encompassing public confidence in governmental bodies, healthcare providers, journalistic institutions, and scientific research organizations. Meanwhile, the emergency coordination subsystem includes municipal emergency management agencies, public health departments, law enforcement units, communications offices, and contracted technical support services.

The overarching objective of the system's cognitive defense strategy is to ensure the uninterrupted implementation of containment policies, safeguard societal cohesion, and preserve reasoned public discourse despite persistent adversarial cognitive pressure. To quantify resilience, a minimum system reliability threshold of 0.85 has been established, meaning the system must retain its critical functions with at least an 85% probability under sustained attack conditions.

4.2 Data Collection and Parameter Quantification

To ensure the accuracy and validity of the model parameters, a multi-source data acquisition framework was employed, integrating literature review, expert evaluation, questionnaire surveys, and historical data analysis.

First, publicly available documents from the targeted region during the epidemic containment period were collected, including government work reports, official announcements, and public opinion monitoring records. From these sources, quantitative indicators such as the volume of disinformation incidents, the success rate of policy implementation, and the frequency of public complaints were extracted. In parallel, comparative parameters were obtained by reviewing documented cognitive warfare responses in analogous domestic public health emergencies, thereby establishing reference benchmarks for model calibration.

Second, a panel of 15 experts was convened to assess the likelihood of individual failure events and to evaluate parameters associated with logical gate configurations. The panel comprised five frontline practitioners from government emergency management agencies, five university professors specializing in relevant disciplines, three senior analysts from public opinion research institutions, and two cybersecurity professionals. Their assessments were grounded in the specifics of the case under investigation.

Table 2. Failure probability parameters of bottom events.

Bottom Event	Normal Probability	Partial Failure Probability	Complete Failure Probability
A_1 (Disinformation attack)	0.10	0.60	0.30
A_2 (Ideological infiltration)	0.30	0.50	0.20
A_3 (Social antagonism incitement)	0.20	0.50	0.30
B_{11} (Information screening failure)	0.40	0.40	0.20
B_{12} (Official information lag)	0.30	0.50	0.20
B_{21} (Insufficient media literacy)	0.20	0.60	0.20
B_{22} (Weak cognitive consensus)	0.30	0.50	0.20
B_{31} (Insufficient government trust)	0.20	0.50	0.30
B_{32} (Interpersonal trust breakdown)	0.30	0.40	0.30
B_{41} (Monitoring and early warning failure)	0.50	0.30	0.20
B_{42} (Slow emergency disposal)	0.40	0.40	0.20
C_1 (Information diffusion)	0.30	0.50	0.20
C_2 (Cognitive bias propagation)	0.25	0.55	0.20
C_3 (Trust collapse propagation)	0.20	0.40	0.40
D_1 (Social disorder)	0.80	0.15	0.05
D_2 (Policy implementation failure)	0.75	0.20	0.05
D_3 (Cognitive security damage)	0.70	0.20	0.10

Third, an online questionnaire survey was administered using stratified sampling across a sample of 1,200 residents within the region. Of these, 1,086 responses were deemed valid, yielding an effective response rate of 90.5%. Survey items captured respondents' understanding of containment policies, trust in official communications, and levels of media literacy. These data served to validate the plausibility of identified failure events and to refine the quantification of related parameters.

Following data integration, the Analytic Hierarchy Process (AHP) was applied to determine the relative weights of all input events for each cascade gate. A two-tier hierarchical structure was developed: the top-level objective was defined as the cascade gate event; the bottom layer comprised the input events of this cascade gate. Using the standard 1–9 pairwise comparison scale, the expert panel evaluated the relative importance of each element, constructed corresponding judgment matrices, and performed consistency checks. All resulting consistency ratios (CR) were below 0.1, satisfying the required threshold for logical coherence. Eigenvectors associated with the maximum eigenvalues ($\lambda_{\max}$) of the matrices were then calculated to derive the final input event weights. Additionally, combined with expert experience and the actual situation of the case, the partial failure trigger threshold T_1 and the complete failure trigger threshold T_2 are set for each cascade gate.

For failure events with historical data support, such as disinformation propagation rate (A_1) and official information release timeliness (the inverse event of B_{12}), the parameter estimation method is used to fit the exponential distribution or normal distribution. On the other hand, for qualitative events lacking historical data, such as insufficient public media literacy (B_{21}) and insufficient government trust (B_{31}), the expert scoring method combined with triangular fuzzy numbers is used to convert into probability values. The membership function of triangular fuzzy numbers is:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0, & x < a, \\ \dfrac{x-a}{m-a}, & a \le x \le m, \\ \dfrac{b-x}{b-m}, & m < x \le b, \\ 0, & x > b. \end{cases}$$

where a, m, and b are the minimum, most likely, and maximum values of expert evaluation respectively, and the probability value is obtained through fuzzy number defuzzification. The final failure probability parameters of each bottom event are shown in Table 2.

4.3 EFT Modeling and Verification

According to the case description and parameter settings, an EFT model as shown in Figure 1 is

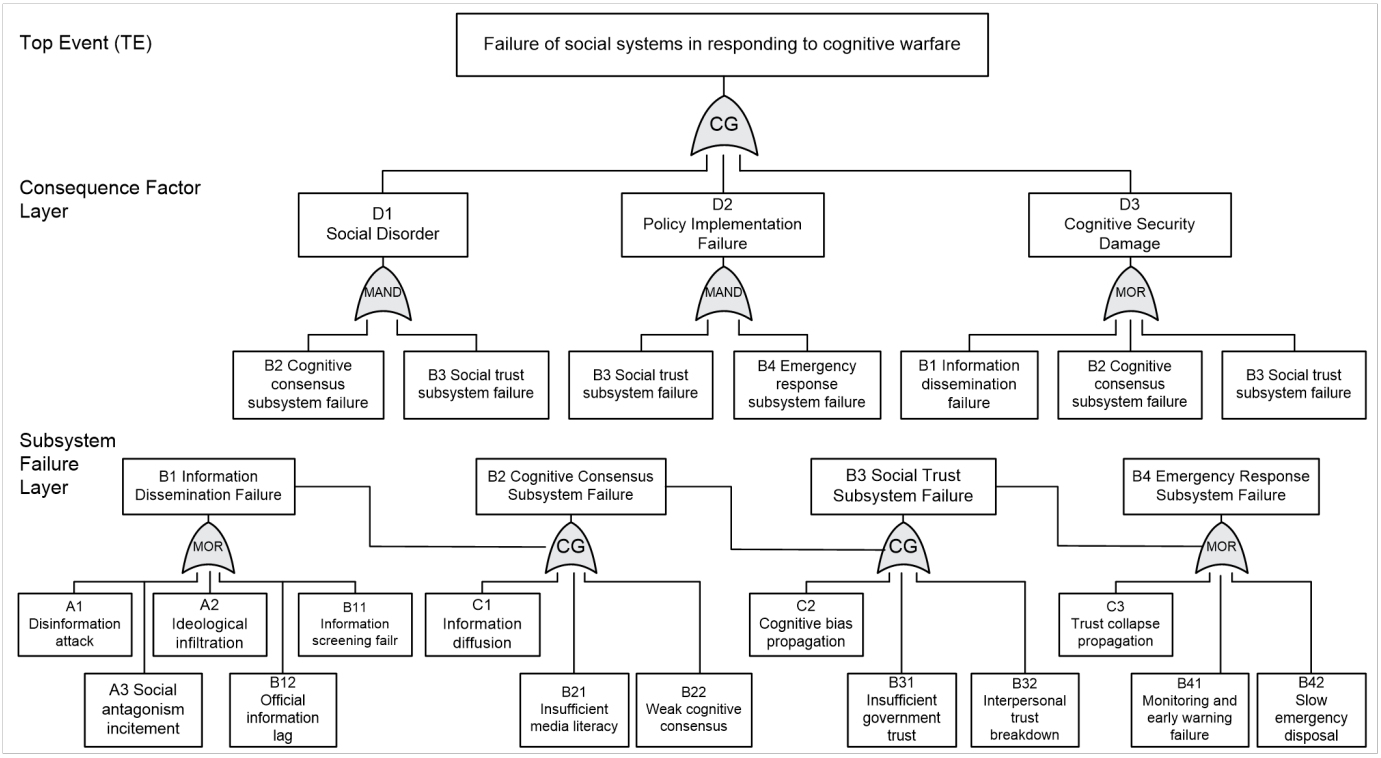


Figure 1. Extended Fault Tree (EFT) model of social system reliability under cognitive warfare.

constructed following the hierarchical decomposition logic.

1. Top Event (TE): The failure of social systems in responding to cognitive warfare. The top event is determined by the states of consequence factors. In this case,

$$TE = \mathrm{CG}\left(D_1, D_2, D_3; w_1 = \tfrac{1}{3}, w_2 = \tfrac{1}{3}, w_3 = \tfrac{1}{3}; T_1 = \tfrac{2}{3}\right).$$

The system is judged to fail (TE = 1) if the weighted aggregate of consequence factor states satisfies $\sum w_i D_i \geq T_1 = \frac{2}{3}$, which corresponds to any single consequence factor reaching state 2, or two or more consequence factors reaching state 1.

2. Consequence factor layer: The consequence factor layer includes three three-state variables: D_1 (social disorder), D_2 (policy implementation failure) and D_3 (cognitive security damage), which are the direct trigger conditions of the top event.

$$D_1 = \mathrm{MAND}(B_2, B_3).$$

Social disorder is triggered only when both the cognitive consensus subsystem (B_2) and the social trust subsystem (B_3) fail;

$$D_2 = \mathrm{MAND}(B_3, B_4).$$

Policy implementation failure is caused only when both the social trust subsystem (B_3) and the emergency response subsystem (B_4) fail;

$$D_3 = \mathrm{MOR}(B_1, B_2, B_3).$$

Cognitive security damage is caused if any of the information dissemination (B_1), cognitive consensus (B_2) and social trust (B_3) subsystems fail.

3. Subsystem failure layer: The subsystem failure layer includes four three-state variables: B_1, B_2, B_3 and B_4.

$$B_1 = \mathrm{MOR}(A_1, A_2, A_3, B_{11}, B_{12}).$$

Attack source factors (A_1, A_2, A_3) are exogenous trigger events, and the failure of the information dissemination subsystem is caused if any of them occurs (partial or complete failure), or if any of the information screening failure (B_{11}) and official information lag (B_{12}) occurs;

$$\begin{aligned} B_2 = \mathrm{CG}\big(&B_1, C_1, B_{21}, B_{22}; \\ &w_1 = 0.30,\ w_2 = 0.20,\ w_3 = 0.35,\ w_4 = 0.15; \\ &T_1 = 0.8,\ T_2 = 1.5\big). \end{aligned}$$

The failure of the cognitive consensus subsystem is caused when the weighted sum of the failure

Table 3. Sensitivity analysis results of bottom events.

Bottom Event	Original Complete Failure Probability	Adjusted Complete Failure Probability	Original Reliability R	Adjusted Reliability R'	Reliability Change Rate $\Delta R/R$	Sensitivity Ranking
B_{21}	0.20	0.22	0.78	0.7176	0.08	1
B_{31}	0.30	0.33	0.78	0.7332	0.06	2
A_1	0.30	0.33	0.78	0.7410	0.05	3
C_2	0.20	0.22	0.78	0.7488	0.04	4
B_{11}	0.20	0.22	0.78	0.7566	0.03	5
B_{22}	0.20	0.22	0.78	0.7614	0.024	6
C_3	0.40	0.44	0.78	0.7644	0.02	7
A_3	0.30	0.33	0.78	0.7662	0.0176	8
B_{32}	0.30	0.33	0.78	0.7680	0.0154	9
C_1	0.20	0.22	0.78	0.7698	0.0131	10
B_{12}	0.20	0.22	0.78	0.7716	0.0103	11
B_{41}	0.20	0.22	0.78	0.7734	0.0082	12
B_{42}	0.20	0.22	0.78	0.7743	0.0073	13
A_2	0.20	0.22	0.78	0.7752	0.0062	14
D_2	0.05	0.055	0.78	0.7770	0.0038	15
D_1	0.05	0.055	0.78	0.7779	0.0027	16
D_3	0.10	0.11	0.78	0.7788	0.0015	17

of the information dissemination subsystem (B_1), information diffusion (C_1), insufficient media literacy (B_{21}) and weak cognitive consensus (B_{22}) reaches the threshold;

$$B_3 = \mathrm{CG}\big(B_2, C_2, B_{31}, B_{32};\ w_1 = 0.35,\ w_2 = 0.25,\ w_3 = 0.30,\ w_4 = 0.10;\ T_1 = 0.8,\ T_2 = 1.5\big).$$

The failure of the social trust subsystem is caused when the weighted sum of the failure of the cognitive consensus subsystem (B_2), cognitive bias propagation (C_2), insufficient government trust (B_{31}) and interpersonal trust breakdown (B_{32}) reaches the threshold;

$$B_4 = \mathrm{MOR}(B_3, C_3, B_{41}, B_{42}).$$

The failure of the emergency response subsystem is caused if any of the failure of the social trust subsystem (B_3), trust collapse propagation (C_3), monitoring and early warning failure (B_{41}) and slow emergency disposal (B_{42}) occurs.

Five experts in the fields of reliability engineering and public management are invited to verify the rationality of the EFT, with verification indicators including logical relationship correctness, event coverage integrity and parameter setting rationality. A Likert 5-point scale (1=completely unreasonable, 5=completely reasonable) is used for scoring. The results show that the average score of each verification indicator is greater than 4.2 and the standard deviation is less than 0.5, indicating that the EFT model has good rationality and effectiveness.

4.4 MDD Conversion and Reliability Calculation

According to the logical relationship of the EFT model, the recursive decomposition method is used to convert it into an MDD model. During the conversion process, the shared subgraph technology and pruning technology of the decision diagram are used to remove redundant nodes and paths and simplify the model structure. The final generated MDD model contains 328 non-leaf nodes and 2 leaf nodes (0=normal, 1=failure). The number of nodes is significantly reduced compared with the combined state number of the traditional fault tree (3^{17}), which verifies the advantage of MDD in compressing the model scale.

Based on the path traversal algorithm of the MDD model, the sum of the probabilities of all paths from the root node to the leaf node "normal (0)" is calculated to obtain the reliability of the social system in responding

to cognitive warfare. The path probability is the product of the state probabilities of each node, and for multi-state nodes, their state probabilities satisfy $\Pr(X=0)+\Pr(X=1)+\Pr(X=2)=1$. Path traversal and probability calculation are realized through programming, and the system reliability is obtained as:

$$R=\sum_{p\in P_0}\prod_{v\in p}\Pr(v=s_v)=0.78$$

where P_0 is the set of all paths from the root node to the leaf node "normal (0)". The calculation results show that the system reliability $R=0.78<0.85$, indicating that the social system fails to reach the set reliability threshold in this cognitive warfare scenario and has a failure risk.

4.5 Sensitivity Analysis and Key Factor Identification

To identify the key factors affecting system reliability, the single-factor sensitivity analysis method is adopted: the complete failure probability of each bottom event is increased by 10%, other parameters remain unchanged, and the change rate of system reliability $\Delta R/R$ is calculated. The larger the change rate, the more significant the impact of the factor on system reliability. The sensitivity analysis results are shown in Table 3.

It can be seen from Table 3 that there are significant differences in the sensitivity of each bottom event. The top 5 bottom events in the sensitivity ranking are: B_{21} (insufficient media literacy, $\Delta R/R=0.08$), B_{31} (insufficient government trust, $\Delta R/R=0.06$), A_1 (disinformation attack, $\Delta R/R=0.05$), C_2 (cognitive bias propagation, $\Delta R/R=0.04$), B_{11} (information screening failure, $\Delta R/R=0.03$). This indicates that insufficient media literacy in the cognitive consensus subsystem, insufficient government trust in the social trust subsystem and disinformation attack are the core factors affecting the reliability of social system cognitive defense, and also the key targets for system optimization.

4.6 Method Effectiveness Verification

The Monte Carlo simulation method is used to verify the effectiveness of the proposed hybrid approach. Based on the parameter settings in the case, a Monte Carlo simulation program is implemented with the simulation times set to 10^5, 10^6, 10^7 and 10^8 respectively. In each simulation, the state of each bottom event is randomly generated (based on its failure probability distribution), and the normal operation of the system is judged according to the logical relationship of the EFT. The simulation results are shown in Table 4. With the increase of the simulation times, the average reliability R_{avg} obtained by Monte Carlo simulation gradually converges to the result calculated by the proposed method ($R=0.78$). When the simulation times reach 10^8, the absolute error is only 1.2×10^{-5}, indicating that the proposed approach has high accuracy. At the same time, the MDD model evaluation time is only 0.8 s (measured independently), which is much lower than that of Monte Carlo simulation (679.1 s for 10^8 simulations as shown in Table 4), verifying the computational efficiency advantage of the proposed method.

5 Discussion

The calculation results of the empirical case show that the reliability of the social system in this region in responding to cognitive warfare is 0.78, which is lower than the set threshold of 0.85, indicating that the core function maintenance capacity of the social system is significantly insufficient in this cognitive warfare scenario and there is a certain risk of system failure. Combined with the sensitivity analysis results, the key factors affecting system reliability can be summarized into three aspects, and corresponding optimization suggestions are put forward accordingly.

First, the cognitive consensus subsystem is the core link affecting system reliability, especially the insufficient media literacy of the public (B_{21}), which has the most significant impact on system reliability (reliability change rate 8%). This is because in the scenario of sudden public health events, the public's ability to distinguish disinformation directly determines the speed and scope of cognitive distortion. If the public lacks basic media literacy, it is easy to be misled by emotional disinformation, which will accelerate the collapse of cognitive consensus and trigger social panic. Therefore, it is necessary to strengthen public media literacy education, carry out targeted publicity and guidance through community lectures, online popular science and other forms, improve the public's ability to identify disinformation and rational cognitive level, and build the first line of defense against cognitive warfare from the source.

Second, the social trust subsystem, especially the insufficient government trust (B_{31}), is the key factor restricting system reliability (reliability change rate 6%). Government trust is the foundation of social trust system. In the process of cognitive warfare,

Table 4. Monte carlo simulation verification results.

Simulation Times	Result 1	Result 2	Result 3	R_{avg}	$\lvert R_{\text{avg}} - R \rvert$	Calculation Time (s)
10^5	0.7786	0.7812	0.7824	0.7807	7.0×10^{-4}	0.72
10^6	0.7795	0.7803	0.7811	0.7803	3.0×10^{-4}	6.8
10^7	0.7799	0.7800	0.7802	0.7800	$< 1.0 \times 10^{-4}$	66.3
10^8	0.77999	0.78000	0.78001	0.78000	1.2×10^{-5}	679.1

the public's trust in the government directly affects their acceptance of official information and prevention and control policies. If the public's trust in the government is insufficient, even if the government releases accurate information and effective policies, it is difficult to gain widespread recognition, which will provide opportunities for the spread of disinformation and ideological infiltration. Therefore, it is necessary to improve the transparency of government work, timely release authoritative information, respond to public concerns, and enhance the public's trust in the government through standardized governance and efficient service.

Third, disinformation attacks (A_1) in attack source factors are important external factors affecting system reliability (reliability change rate 5%). Disinformation is the main carrier of cognitive warfare, and its rapid spread will directly damage the information dissemination order and disrupt the public's cognitive judgment. Therefore, it is necessary to strengthen the construction of information screening and supervision mechanisms, give play to the role of social media platforms in information management, use technical means such as artificial intelligence and big data to realize automatic detection, interception and disposal of disinformation, block the propagation channels of cognitive attack information, and reduce the impact of external cognitive attack sources.

In addition, the cognitive bias propagation (C_2) and information screening failure (B_{11}) also have a certain impact on system reliability. For cognitive bias propagation, it is necessary to strengthen the guidance of public opinion, break the "information cocoon" effect, and guide the public to obtain information from multiple channels to avoid cognitive polarization. For information screening failure, it is necessary to improve the information screening capacity of the information dissemination subsystem, establish a multi-level information verification mechanism, and ensure the accuracy and authority of information dissemination.

6 Conclusions and Future Work

This paper addresses the challenge of quantitatively evaluating the reliability of social systems in countering cognitive warfare by proposing a hybrid analytical framework that integrates Extended Fault Tree analysis with Multi-Valued Decision Diagrams. The effectiveness and practical applicability of the proposed method are demonstrated through an empirical case study. The principal findings are summarized as follows:

1) The reliability of a social system against cognitive warfare constitutes a critical measure of its capacity to withstand adversarial cognitive operations and maintain essential functions. This reliability is influenced by multiple factors, including the nature of attack sources, subsystem failures, propagation pathways, and resultant consequences, and is characterized by complex dynamics such as multi-source initiation and cascading propagation. The four core subsystems—information dissemination, cognitive consensus formation, social trust maintenance, and emergency response coordination—are highly interdependent; failure in any single subsystem may initiate a chain of failures that compromise the integrity of the entire system.

2) The hybrid EFT–MDD approach offers an effective solution for the quantitative assessment of social system reliability in cognitive warfare contexts. By incorporating multi-state nodes and cascade gates (CG), EFT accurately captures the graded failure modes and the cascading effects within the system. MDD complements this by enabling efficient computation of the complex logical structures generated by EFT, thereby avoiding the combinatorial explosion problem prevalent in traditional fault tree analysis and achieving a balance between analytical precision and computational feasibility.

3) Empirical analysis reveals that the reliability of the social system in the case study region, when exposed to cognitive warfare, is estimated at 0.78—below the predefined resilience threshold of 0.85. Key

determinants undermining system reliability include inadequate public media literacy, diminished trust in government institutions, and intensive disinformation campaigns. Targeted enhancement of these areas is shown to significantly strengthen the cognitive defense capabilities of the social system.

Despite these contributions, several limitations warrant consideration and offer avenues for future research. The current analytical scope is confined to four core subsystems, excluding indirect influences from economic, ecological, and other peripheral systems. Future studies could expand system boundaries to incorporate multidimensional determinants, thereby constructing a more holistic reliability evaluation framework. Furthermore, the model assumes full knowledge of the intensity and modalities of cognitive warfare attacks, whereas real-world scenarios are often marked by high levels of concealment and uncertainty. Subsequent research might integrate interval analysis and fuzzy probability theory to develop reliability assessment methodologies capable of addressing uncertain attack environments.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

Chunyu Miao is affiliated with the Hangzhou Anheng Information Technology Co., Ltd., Hangzhou, China. The authors declare that this affiliation had no influence on the study design, data collection, analysis, interpretation, or the decision to publish, and that no other competing interests exist.

AI Use Statement

The authors declare that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

This study involved expert evaluations and an anonymous questionnaire survey. According to the institutional guidelines applicable to anonymous, non-interventional, and minimal-risk survey research, formal ethical approval was not required for this study. Participation was entirely voluntary, and informed consent was obtained from all participants before data collection. No personally identifiable or sensitive personal information was collected. All responses were anonymized and used exclusively for academic research.

References

[1] Green, S. (2008). *Cognitive Warfare* (M.S. thesis). Joint Military Intelligence College, Washington, DC, USA. https://theaugeanstables.com/wp-content/uploads/2009/10/Stuart-Green-LTC-USN-Cognitive-Warfare.pdf

[2] Danyk, Y., & Briggs, C. M. (2023). Modern cognitive operations and hybrid warfare. *Journal of Strategic Security, 16*(1), 35-50. [CrossRef]

[3] Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science, 185*(4157), 1124-1131. [CrossRef]

[4] Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology, 2*(2), 175-220. [CrossRef]

[5] Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146-1151. [CrossRef]

[6] Claverie, B., & du Cluzel, F. (2022). *The Cognitive Warfare Concept*. NATO-ACT Innovation Hub. https://innovationhub-act.org/wp-content/uploads/2023/12/CW-article-Claverie-du-Cluzel-final_0.pdf

[7] Shrestha, A., Xing, L., & Dai, Y. (2009). Decision diagram based methods and complexity analysis for multi-state systems. *IEEE Transactions on Reliability, 59*(1), 145-161. [CrossRef]

[8] Mo, Y., Xing, L., & Amari, S. V. (2014). A multiple-valued decision diagram based method for efficient reliability analysis of non-repairable phased-mission systems. *IEEE Transactions on Reliability, 63*(1), 320-330. [CrossRef]

[9] Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly, 80*(S1), 298-320. [CrossRef]

[10] Xing, L., & Dai, Y. S. (2008). A new decision-diagram-based method for efficient analysis on multistate systems. *IEEE Transactions on Dependable and Secure Computing, 6*(3), 161-174. [CrossRef]

[11] Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter, 19*(1), 22-36. [CrossRef]

[12] Wang, H., Lu, X., Du, Y., Zhang, C., Sadiq, R., & Deng, Y. (2017). Fault tree analysis based on TOPSIS and triangular fuzzy number. *International journal of system assurance engineering and management, 8*(Suppl 4), 2064-2070. [CrossRef]

[13] Rao, K. D., Gopika, V., Rao, V. S., Kushwaha, H. S., Verma, A. K., & Srividya, A. (2009). Dynamic fault tree analysis using Monte Carlo simulation in probabilistic safety assessment. *Reliability Engineering & System Safety, 94*(4), 872-883. [CrossRef]

[14] Xing, L. (2020). Cascading failures in Internet of Things: Review and perspectives on reliability and resilience. *IEEE Internet of Things Journal, 8*(1), 44-64. [CrossRef]

[15] Wilson, C., Grizzle, A., Tuazon, R., Akyempong, K., & Cheung, C. K. (2011). *Media and Information Literacy Curriculum for Teachers*. UNESCO. https://www.unesco.org/en/articles/media-and-information-literacy-curriculum-teachers

[16] Rauzy, A. (1993). New algorithms for fault trees analysis. *Reliability Engineering & System Safety, 40*(3), 203-211. [CrossRef]

[17] Xing, L. (2005). Reliability modeling and analysis of complex hierarchical systems. *International Journal of Reliability, Quality and Safety Engineering, 12*(06), 477-492. [CrossRef]

[18] Mo, Y., Xing, L., Lin, Y. K., & Guo, W. (2018). Efficient analysis of repairable computing systems subject to scheduled checkpointing. *IEEE Transactions on Dependable and Secure Computing, 18*(1), 1-14. [CrossRef]

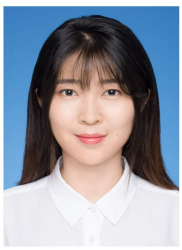
[19] Mo, Y., & Xing, L. (2021). Efficient analysis of resource availability for cloud computing systems to reduce SLA violations. *IEEE Transactions on Dependable and Secure Computing, 19*(6), 3699-3710. [CrossRef]

[20] Xing, L., & Amari, S. V. (2015). *Binary decision diagrams and extensions for system reliability analysis*. John Wiley & Sons. [CrossRef]

Yuchang Mo received the B.E. (2002), M.S. (2004), and Ph.D. (2008) degrees in Computer Science from Harbin Institute of Technology, Harbin, China. He is currently a Distinguished Professor with the School of Computer Science and Technology, Zhejiang University of Water Resources and Electric Power, Hangzhou, China. He was a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Massachusetts (UMass) Dartmouth, USA from September 2012 to August 2013. Prof. Mo was the Program Chair of the 5th Conference on Dependable Computing in China. He is currently a Senior Expert at the Reliability Research Society of China and a Senior Expert at the Technical Committee on Fault Tolerant Computing of China. His current research focuses on the reliability analysis of complex systems and networks using combinatorial models. His research has been supported by the National Science Foundation of China. (Email: yuchangmo@sina.com)

Chunyu Miao, Senior Engineer, Chief Talent Officer, Senior Vice President and Dean of Digital Talent Research Institute at DBAPPSEC, the Hangzhou Anheng Information Technology. He serves as visiting professor and industrial supervisor for postgraduate students at several universities. As Hangzhou leading high-skilled talent with more than 15 years of industrial experience in cybersecurity, his research focuses on cyber defense, cyberspace trust system and digital security talent training. He has presided over and participated in six national/provincial research projects, published over 50 papers, compiled 10 monographs, and obtained more than 30 invention patents and software copyrights. He has delivered special lectures for government departments and won many provincial and ministerial scientific and teaching awards. (Email: chunyumiao@sina.com)

Guangxuan Chen, Deputy Dean of School of Information and Cybersecurity, Zhejiang Police College. He leads the secondary discipline of Cyberspace Security and the major of Cybersecurity and Law Enforcement. He completed postdoctoral research at Institute of Software, Chinese Academy of Sciences and was a visiting scholar at University of Lausanne, Switzerland. He is a specially-appointed expert of Electronic Forensics Expert Committee and core member of multiple provincial research platforms. His research interests cover cyberspace security, cybercrime investigation and electronic evidence forensics. He has published over 80 papers, authored 6 textbooks/monographs and participated in drafting one information security standard. He presides over multiple national and provincial research projects and takes part in five national key research projects as a key participant. (Email: Guangxuanchen@sina.com)



Tiantian Gai, Lecturer, School of Computer Science and Technology, Zhejiang University of Water Resources and Electric Power. She received all her degrees from Shanghai Maritime University and conducted joint research at De Montfort University funded by China Scholarship Council. Her research focuses on large language models, intelligent decision-making, group consensus decision, social network analysis and trust modeling. As the first author, she has published high-level papers in top international journals including IEEE transactions and Nature portfolio journals, with over 770 citations. She was awarded National Postgraduate Scholarship for both master and doctoral studies, presided over university innovative talent project and participated in several NSFC projects. (Email: Tiantiangai@sina.com)

Wen Wu received the B.Sc. degree in electronic and information engineering from Wuhan College of Arts and Science, Wuhan, China, in 2016, and the M.Sc. degree in computer applied technology from Hubei University, Wuhan, China, in 2019. He received the Ph.D. degree in computer science and technology from Hangzhou Dianzi University, Hangzhou, China, in 2025. From 2019 to 2022, he was a Lecturer in the Department of Computer Science at the Xinjiang Institute of Technology, Xinjiang, China. Since 2025, he has been with the School of Computer Science and Technology, Zhejiang University of Water Resources and Electric Power, Hangzhou, China, as a Faculty Member. In addition, since 2024, he has been a Research Assistant with the City University of Hong Kong, Hong Kong, China. His research interests include deep learning and computer vision. (Email: wuwen@sina.com)

Yuhuan Gong, Associate Professor, School of Marxism, Zhejiang University of Water Resources and Electric Power. She holds a master's degree in philosophy from Harbin Institute of Technology. She has worked in multiple universities and long engaged in research on Marxist theory, digital ideology and cognitive security. Her research spans technology risk analysis, digital ideological security, information cognitive warfare and digital trust system construction, providing theoretical support for cognitive warfare governance in cyberspace. She has presided over a number of provincial and university-level research projects and published papers in core journals of philosophy of science and technology management. (Email: yuhuangong@sina.com)