Journal of Computer Science and Electrical Engineering
Print ISSN: 2663-1938
Online ISSN: 2663-1946
DOI: https://doi.org/10.61784/jcsee3144

DYNAMIC TRUST EVALUATION AND RISK CONTROL TECHNOLOGY FOR BOUNDARIES OF NOVEL POWER SYSTEMS

Yang Cao*, Yang Su, Peng Zhou, XueFei Tian, ShuXiang Wen, KaiMin Zheng, JinYu Wu
China Southern Power Grid Co., Ltd., Guangzhou 510663, Guangdong, China.
**Corresponding Author: Yang Cao*

Abstract: Against the backdrop of the full advancement of the "dual carbon" strategy, the new power system has gradually formed an operation pattern featuring multi-stakeholder collaboration across source, grid, load and storage. With the ubiquitous access of massive distributed terminals and the open, interconnected network architecture, the traditional static security protection mode can no longer cope with the increasingly complex cybersecurity risks. Targeting practical problems such as blurred power protection boundaries, rapid risk propagation, lagging static rule-based defense and single disposal methods, this paper proposes an integrated security protection scheme combining dynamic trust evaluation and grayscale risk control. Based on multi-dimensional features, a quantitative trust evaluation system is constructed, and a layered grayscale disposal strategy is designed tailored to business scenarios. The full-link system covering data collection, transmission, intelligent analysis, decision issuance and closed-loop optimization is built, and core modules including subject-object identification, trust assessment and intelligent risk control are deeply designed and engineered. Verified by multiple sets of simulation experiments, the system achieves millisecond-level risk identification and policy response, and significantly improves boundary security protection capabilities while ensuring continuous operation of power businesses. It can provide feasible technical ideas and engineering references for the construction of cybersecurity systems for new power systems.
Keywords: New power system; Boundary protection; Dynamic trust; Risk control; Grayscale strategy; Stream computing

1 INTRODUCTION

With the steady implementation of the carbon peak and carbon neutrality goals, China's power industry is undergoing profound systemic transformation. The traditional closed power grid is accelerating its transition to a new power system characterized by a high proportion of new energy and high power electronics. Distributed photovoltaics, wind turbines, energy storage devices and various flexible loads are deployed on a large scale in urban and rural parks, remote areas and other regions. The power network is no longer limited to dedicated intranets, but has gradually evolved into an open form with deep integration of public and private networks and cross-domain interaction of multiple entities. The frequency of cross-network interaction for core businesses such as power dispatching, remote operation and maintenance, and power market trading continues to rise, and security threats including cyber attacks, illegal terminal access, malicious device tampering and business data leakage emerge one after another [1].

The traditional power security protection system has long relied on firewalls, intrusion detection systems and static access devices, and generally adopts the binary protection logic of "permit or block" [2]. Such static protection methods based on fixed rules cannot accurately determine device identity and real-time behavior in scenarios where massive heterogeneous terminals are dynamically accessed. Meanwhile, the iteration speed of rules lags far behind the evolution of attack methods, resulting in weak defense against new threats such as advanced persistent threats (APTs) and zero-day attacks. A more prominent problem is that the one-size-fits-all blocking strategy can easily disrupt normal power businesses and fail to strike a balance between security protection and business continuity [3]. Against this industry background, researching dynamic trust evaluation and adaptive risk control technologies adapted to the boundary characteristics of new power systems, and building a new generation of security protection system with self-perception, independent decision-making and continuous evolution capabilities have become a critical research direction in power cybersecurity.

1.1 Research Status at Home and Abroad

Under the general trend of global energy transformation, developed countries in Europe and the United States launched research on energy internet and power industrial control cybersecurity earlier. Technologies such as dynamic trust, zero trust and adaptive risk control have been gradually applied in power scenarios. In the field of trust evaluation theory, foreign scholars first introduced classical trust models such as subjective trust and Bayesian models into industrial control systems, and assessed credibility based on historical interaction records of devices [4]. Some studies have constructed power network trust graphs with graph neural networks to realize risk traceability through topological

relationships, but most achievements remain at the theoretical simulation stage and have not been engineered with massive streaming data from actual power sites [5].

In terms of risk control, the zero-trust security architecture has become the mainstream direction of the international power industry, adhering to the core concept of "never trust, always verify" and breaking the traditional boundary between internal and external networks. Many overseas power grid operators combine zero trust with micro-segmentation technology to achieve fine-grained terminal permission control, but most existing schemes focus on identity authentication and lack dynamic risk grading capabilities for power business behaviors [6]. In the field of real-time security analysis, foreign industrial control security vendors have launched multiple streaming analysis platforms for real-time log and traffic parsing. However, these platforms generally adopt the mode of offline modeling with static strategies, featuring long model iteration cycles and no support for refined operation and maintenance capabilities such as policy grayscale release and automatic rollback.

Domestic academic and industrial communities have carried out extensive research on new power system security in recent years. Leading enterprises such as State Grid Corporation of China and China Southern Power Grid [7], in collaboration with universities and technology firms, have focused on technical exploration in three major areas: boundary protection, intrusion detection and security situational awareness. At the security architecture level, domestic research teams have localized the zero-trust architecture based on domestic power business practices, built a protection framework adapted to source-grid-load-storage collaboration, and optimized terminal access authentication and dynamic permission management mechanisms. However, a complete, quantifiable trust evaluation system for device behavior has not yet been formed.

In algorithm research, machine learning algorithms such as long short-term memory (LSTM) networks and random forests are widely used for abnormal behavior detection of power equipment, which can effectively identify traffic and time-series anomalies [8]. Nevertheless, most existing algorithms are applied as single models, failing to integrate multi-dimensional features of identity, behavior and environment for comprehensive trust evaluation, nor realizing linkage between evaluation results and risk control strategies. In engineering applications, power network situational awareness platforms have been deployed in many regions of China, supporting risk aggregation, alarm display and post-event traceability [9]. However, these platforms lack active real-time disposal capacity; risk disposal still relies on manual operation, and the static black-and-white list management mode is difficult to adapt to dynamically changing security situations.

A comprehensive review of domestic and foreign research reveals that current technologies still have many shortcomings: trust evaluation is limited to a single dimension and fails to form a comprehensive quantitative model integrating identity, behavior and environmental features; risk disposal methods are rigid, lacking flexible grayscale strategies between full permit and complete block; the model and strategy system lacks a closed-loop iteration mechanism, making it difficult to continuously counter evolving cyber threats [10]. Addressing these gaps, this paper develops a dynamic trust evaluation and grayscale risk control system based on the operation characteristics and security requirements of new power systems.

1.2 Research Content and Innovation Points

Taking boundary security protection of new power systems as the core objective, this paper first sorts out the pain points of system operation and security protection and clarifies overall design requirements. It then builds a layered system architecture and defines the functions and technical solutions of each layer. On this basis, a multi-dimensional dynamic trust evaluation model is constructed, and a matching four-level grayscale risk disposal system is designed. Subsequently, the design and implementation of core engines such as subject-object identification, trust evaluation and risk decision-making are elaborated. Finally, a simulation environment is built to verify the comprehensive performance of the system in terms of identification accuracy, response delay and business compatibility.

This paper makes three main innovations:

1. A dynamic trust evaluation model integrating identity, behavior and environment dimensions is constructed, and multi-algorithm parallel computing is adopted to quantify trust scores, realizing accurate credibility assessment of power terminals and communication sessions.
2. Abandoning the traditional binary protection mode, a four-level grayscale risk control system is designed, with differentiated disposal actions matched to trust levels to balance security protection and business continuity.
3. An integrated closed-loop evolution architecture covering data, models and strategies is established, supporting automatic model retraining, policy grayscale release and automatic fault rollback, enabling the protection system to adapt to threat evolution and business adjustment.

The paper is organized into six chapters: Chapter 1 is the introduction, covering research background, domestic and foreign research status, research content and innovations; Chapter 2 introduces core theories and key supporting technologies including dynamic trust and grayscale risk control; Chapter 3 details the overall system architecture, layered design and deployment integration scheme; Chapter 4 presents the in-depth design of core engines and functional modules; Chapter 5 describes the experimental setup and test results, with data analysis and discussion; Chapter 6 summarizes the work and prospects future research directions.

2 RELEVANT THEORIES AND TECHNICAL FOUNDATIONS

2.1 Dynamic Trust Evaluation Theory

Trust is a comprehensive measure of the identity legitimacy and behavior reliability of network entities. Different from static trust that remains fixed, dynamic trust updates evaluation results in real time as entity behavior and network environment change. This paper defines terminal devices and operation-and-maintenance sessions in the power system as trust subjects, and constructs a comprehensive trust evaluation function as shown in Formula (1):

$$T=f(I,B,E) \tag{1}$$

Where T represents the comprehensive trust score, ranging from 0 to 100; a higher score indicates stronger credibility of the subject. I is the identity feature dimension, covering static attributes such as unique device identifiers, digital certificates and inherent permissions. B is the behavior feature dimension, including dynamic interaction data such as communication frequency, session duration and protocol message characteristics. E is the environment feature dimension, corresponding to contextual information such as access network zone, access time and network operation status.

In line with power business security classification requirements, four grayscale trust levels are divided based on trust scores, as shown in Formula (2):

$$TrustLevel=\begin{cases} High\ Trust, 80\le T\le 100 \\ Medium\ Trust, 50\le T<80 \\ Low\ Trust, 20\le T<50 \\ Very\ Low\ Trust, 0\le T<20 \end{cases} \tag{2}$$

Different trust levels correspond to differentiated control intensities, providing a quantitative basis for formulating risk strategies.

2.2 Grayscale Risk Control Theory

Grayscale risk control is a flexible disposal philosophy between full permit and complete block, which matches control actions to the trust level of the subject. This paper designs a four-level linkage grayscale disposal strategy: observation, speed limit, reinforcement and isolation. The observation level only mirrors traffic and retains logs without interfering with normal communication. The speed limit level restricts bandwidth and access frequency to reduce attack damage. The reinforcement level activates multi-factor authentication and encryption enhancement for stronger protection. The isolation level directly interrupts risk links and isolates problematic nodes. This gradient disposal mode achieves a dynamic balance between security protection and business operation.

2.3 Key Supporting Technologies

The system is built on mainstream information technologies: the Apache Flink distributed streaming computing framework is adopted to process massive real-time data and ensure low latency. A message bus is built based on Kafka for data distribution and traffic peak shaving. Algorithms including random forest, LSTM and graph neural network are integrated to mine multi-dimensional risk features. Meanwhile, high-availability technologies such as clustering, master-slave redundancy and cross-data-center multi-active deployment are applied to ensure 7×24 stable operation of the system.

3 OVERALL SYSTEM ARCHITECTURE DESIGN

3.1 Architecture Design Principles

Considering the characteristics of distributed deployment, strict real-time requirements and high security standards of new power systems, the system architecture follows five principles: distributed collection compatible with heterogeneous industrial protocols; end-to-end millisecond-level low-latency processing; modular decoupling for easy expansion and maintenance; cluster and multi-data-center architecture to eliminate single points of failure; data-model-strategy linkage for closed-loop self-evolution.

3.2 Layered Overall Architecture

According to data flow and business logic, the system is divided into six layers: probe collection layer, real-time message bus layer, core risk control layer, intelligent analysis layer, strategy execution layer and data persistence layer. All layers work together to form a complete security protection chain, as shown in Figure 1.

The probe collection layer serves as the data entry point. Lightweight probes are deployed on edge nodes such as power gateways, RTUs and intelligent measurement and control terminals. The probes are compatible with mainstream protocols including IEC61850, DNP3, MQTT and Modbus, and adopt traffic mirroring plus differentiated sampling to avoid bandwidth occupation. All data is transmitted via TLS encryption. For network interruption scenarios, the probe supports local ring queue caching and breakpoint resume to ensure data integrity.

The real-time message bus layer is built on a Kafka cluster, with independent topics divided by event, device and alarm types. It supports throughput of millions of messages per second with end-to-end transmission delay below 10ms. ACL

is used for multi-tenant and module permission isolation. Traffic scheduling capabilities support grayscale release of models and strategies, and multi-replica mechanisms prevent data loss.

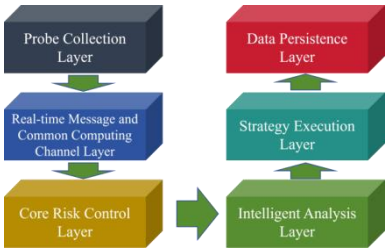

Figure 1 Layered Overall Architecture of the Protection System

As the core control and alarm hub, the core risk control layer consists of two modules: device log management and risk warning orchestration. The device log management module uniformly maintains assets, topology and online status of devices across the network, and supports log parsing, retrieval and traceability. The risk warning module monitors traffic and access behaviors in real time, orchestrates alarms by severity level, and supports multi-channel push via SMS, email and operation-and-maintenance platforms.

The intelligent analysis layer is the core computing unit, including data modeling and online anomaly detection modules. The data module completes data cleaning and feature engineering via offline ETL, and mines implicit risk rules using clustering and statistical algorithms. The online detection module adopts a multi-model parallel architecture, integrating statistical models, machine learning and deep learning algorithms to identify short-term anomalies, complex features and long-sequence/topological risks respectively. Results from multiple models are weighted and fused to output comprehensive trust scores and grayscale levels.

The strategy execution layer is responsible for instruction generation and issuance, including model/rule deployment and real-time stream computing modules. It supports one-click release, rollback and grayscale switching of rules and models. Relying on the streaming engine for real-time decision-making, it generates four-level disposal instructions at millisecond scale.

The data persistence layer adopts a hybrid storage architecture: time-series databases store high-frequency data such as real-time traffic, trust scores and alarms; big data file systems archive logs and historical model data; full metadata and data lineage are managed uniformly for full-link traceability.

3.3 Deployment and Integration Architecture

The system adopts cluster deployment with cross-data-center multi-active configuration. All core services are deployed in clusters, with automatic failover for single-node faults. Primary and backup data centers synchronize data and configurations in real time; when the primary site fails, traffic is switched to the backup site within seconds via routing. Databases and middleware adopt a master-slave architecture, with snapshot plus incremental backup for disaster recovery.

For external integration, risk data is synchronized with the situational awareness system via RESTful API and Kafka dual channels. A two-way interface is established with the remote operation-and-maintenance system to realize strategy issuance and status feedback, achieving linkage between security and operation-and-maintenance businesses.

4 DESIGN AND IMPLEMENTATION OF CORE MODULES

4.1 Subject-Object Identification Engine

The subject-object identification engine is the pre-module for trust evaluation. It integrates multi-source data from probes, CMDB, DHCP/ARP and captures all network entities and links through passive traffic analysis plus active heartbeat detection. The system generates a globally unique identifier for each device and each session based on a hash algorithm, as shown in Formula (3):

$$ID=Hash(MAC\ address + device\ serial\ number + time\ stamp) \tag{3}$$

Meanwhile, a multi-dimensional label system covering business, security and geography is constructed, with labels dynamically updated along with device behavior. The engine stores dynamic topology in a graph database, with a link disconnection rule triggered after 5 minutes of no traffic. The front-end visualization interface supports filtering nodes by risk level and region, and intuitively displays risk propagation paths.

4.2 Risk Control Rule Engine

The rule engine is divided into three categories: static, dynamic and composite rules, which translate manual security experience into executable logic. The full life-cycle of a rule includes four stages: draft, verification, online and offline. A three-level grayscale release process (region → device group → whole network) is adopted, with real-time monitoring of false positive rate and business delay and automatic rollback in case of anomalies.

Millisecond-level rule matching is implemented in the streaming computing framework. The rule matching degree is defined in Formula (4):

$$Match=\frac{Number\ of\ valid\ hits}{Total\ number\ of\ visits} \tag{4}$$

The matching degree is fed into the trust evaluation engine as a core feature, realizing deep integration of rules and AI models.

4.3 Dynamic Trust Evaluation Engine

The engine calculates the identity trust score(T_1), behavior trust score(T_2) and environment trust score (T_3) in parallel, and obtains the comprehensive trust score through weighted fusion:

$$T=w_1T_1+w_2T_2+w_3T_3 \tag{5}$$

Subject to the constraint ($w_1+w_2+w_3=1$), the weights can be dynamically tuned according to power business scenarios.
Model training combines offline and incremental training, with false positive rate and false negative rate as core evaluation metrics:

$$False\ Positive\ Rate=\frac{Number\ of\ normal\ behaviors\ misjudged\ as\ risks}{Total\ number\ of\ normal\ behaviors} \tag{6}$$

$$False\ Negative\ Rate=\frac{Number\ of\ risk\ behaviors\ not\ detected}{Total\ number\ of\ risk\ behaviors} \tag{7}$$

The model supports hot deployment and A/B testing, and the optimal version is put into production after iteration.

4.4 Risk Decision Engine

The engine constructs a decision matrix based on trust level, device importance and business SLA, automatically generates differentiated strategies, and performs instruction conflict detection. Standardized instructions are issued through the message bus, and device execution receipts are returned to the analysis layer, forming a closed loop of "decision – execution – feedback".

4.5 Strategy Issuance and Execution Module

A plug-in southbound architecture is adopted to adapt to various boundary devices such as firewalls, SDN controllers and power gateways. Strategies are issued in batches in grayscale mode, with real-time monitoring of execution status and automatic rollback for abnormal scenarios. Failures in issuance or execution timeout trigger multi-channel alarms to notify operation-and-maintenance personnel.

4.6 Closed-Loop Optimization and Situational Awareness Module

Full-link data is aggregated to build a panoramic situational dashboard, realizing visualization of risk heat map, alarm trends and strategy execution. The system automatically updates features, retrains models and fine-tunes strategy thresholds based on metrics such as false positive rate and delay, completing autonomous iteration and optimization.

5 EXPERIMENTAL VERIFICATION AND RESULT ANALYSIS

5.1 Experimental Environment and Test Scheme

This simulation experiment replicates the operating environment of a new power system. Hardware includes distributed photovoltaics, energy storage terminals, power gateways and high-performance server clusters. The software environment is CentOS 7.9 with components including Kafka, Flink, InfluxDB and TensorFlow. The test dataset contains 130,000 samples in total, including 100,000 pieces of normal power business traffic and risk traffic simulating port scanning, illegal access and APT attacks.
Three comparison schemes are set: traditional static IDS, single LSTM model, and the multi-dimensional trust risk control system proposed in this paper. Tests are conducted from four dimensions: identification accuracy, response delay, business compatibility and system stability.

5.2 Experimental Results

To present experimental data and conclusions intuitively, this paper uses visual charts to display test results. According to the characteristics of experimental indicators, four types of charts are presented: bar chart of risk identification indicators, line chart of concurrency vs. response delay, 72-hour system resource monitoring curve, and whole-network risk distribution heat map.
As shown in Figure 2, the bar chart compares the risk identification capabilities of traditional static IDS, single LSTM model and the proposed system, using three core metrics: accuracy, false positive rate and false negative rate. Accuracy represents the system's ability to correctly distinguish normal from risky traffic; false positive rate is the proportion of normal behaviors misclassified as risky; false negative rate is the proportion of real attacks that are not detected. These three metrics are the core criteria for evaluating the identification performance of security protection systems. The raw

experimental data comes from statistics on 130,000 test samples, and the bar chart intuitively highlights the performance gaps between different schemes.

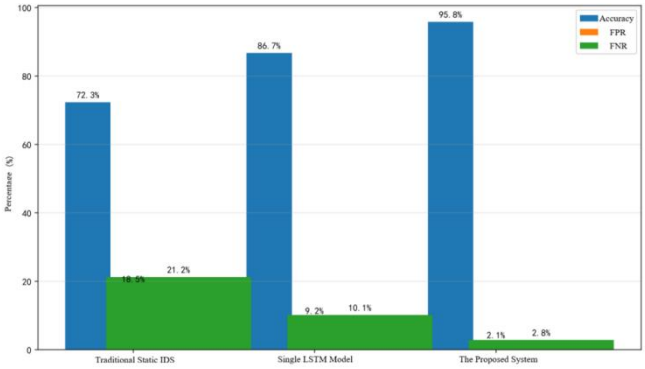

Figure 2 Comparison of Risk Identification Indicators among Different Schemes

As shown in the Figure 3, the new power system faces scenarios of massive concurrent terminal access, making system response delay a key indicator for ensuring real-time power business. This test selects four levels of concurrent requests: 1000, 5000, 10000 and 50000, simulating operating scenarios from normal load to high load. The end-to-end average delay from data collection, trust calculation to strategy execution is counted for each concurrency level. The chart clearly reflects the trend of delay with increasing concurrency, verifying the real-time processing capability of the system under high-concurrency scenarios.

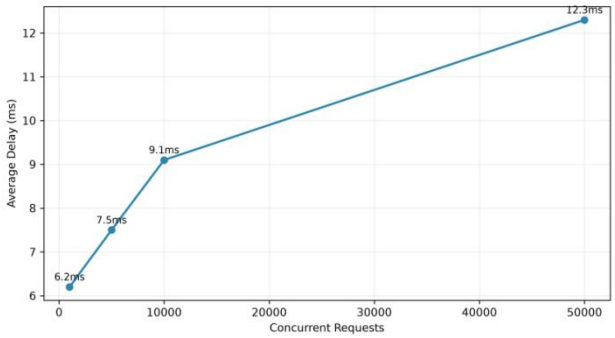

Figure 3 Concurrency-Response Delay Line Chart

Power systems require security protection platforms to operate 7×24 without interruption. As shown in the 72-hour system resource monitoring curve, a continuous 72-hour stress test is conducted to simulate long-term full-load operation. Figure 4 shows two core resource metrics, server CPU usage and memory usage, during the test period. The curve trends are used to judge whether there are problems such as resource leakage, sudden load spikes and service downtime, so as to verify the cluster architecture, disaster recovery mechanism and overall operational stability.

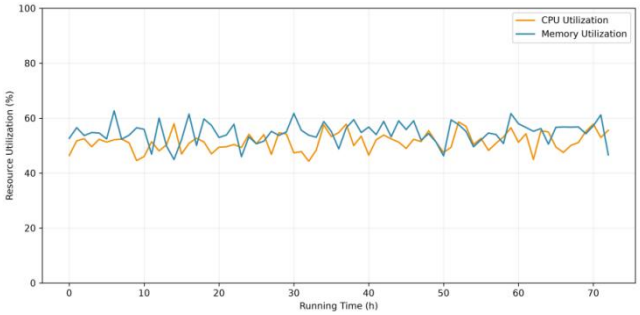

Figure 4 72-Hour System Resource Monitoring Curve

To verify the system's situational visualization and regional risk perception capabilities, the whole-network risk distribution heat map experiment follows the power zoning management mode, dividing the test environment into 3 major regions and 9 operation zones to simulate terminal access and risk distribution in different areas. Figure 5 uses color depth to represent regional risk level: larger values and darker colors indicate higher security risk in the zone. It intuitively presents the characteristics of whole-network risk distribution, and supports practical functions such as risk node drill-down and traceability, fitting the actual operation-and-maintenance scenarios of power systems.

Figure 5 Heat Map of Power Whole-Network Risk Distribution

5.3 Experimental Results

Combined with four types of visual charts and measured data, in-depth analysis is conducted from five dimensions: risk identification capability, real-time response performance, long-term operation stability, business adaptability and situational awareness capability. Comprehensive discussion is also carried out in combination with new power system business scenarios, shortcomings of traditional schemes and technical advantages of the proposed scheme.

5.3.1 Risk identification capability

Data from Figure 2 shows significant differences in identification metrics among the three schemes: traditional static IDS achieves only 72.3% accuracy, with 18.5% false positive rate and 21.2% false negative rate; the single LSTM model improves accuracy to 86.7%, with 9.2% false positive rate and 10.1% false negative rate; the system proposed in this paper reaches 95.8% accuracy, with false positive rate and false negative rate reduced to 2.1% and 2.8% respectively.

Traditional static IDS relies on a fixed rule base and can only detect known attacks with clear signatures. It is completely ineffective against new threats such as zero-day attacks, APT attacks and disguised illegal access. Meanwhile, normal communication messages from a large number of heterogeneous devices in power systems are easily misjudged by rules, leading to high false positive and false negative rates. The single LSTM model only analyzes time-series traffic features with a single dimension, and cannot incorporate key information such as device identity, network environment and topological correlation. It has limited ability to identify complex cross-link and multi-node collaborative risks, resulting in obvious bottlenecks in identification performance.

The multi-dimensional dynamic trust evaluation system constructed in this paper integrates three major features of identity, behavior and environment, and adopts parallel analysis of statistical models, machine learning and deep learning. On one hand, explicit abnormal behaviors are detected through the rule engine; on the other hand, implicit risk patterns are mined via AI algorithms, and risk propagation relationships between nodes are perceived through dynamic topology. The architecture of multi-dimensional feature fusion and multi-algorithm complementation greatly improves the detection capability for complex threats. The dynamic model iteration mechanism continuously adapts to new attack methods, effectively reducing false positives and false negatives. From an engineering perspective, a low false positive rate reduces the workload of operation-and-maintenance personnel handling invalid alarms, and a low false negative rate strengthens the security defense of power networks, fully meeting the high security requirements of new power systems.

5.3.2 Real-time response performance

As shown in Figure 3, when the number of concurrent requests is 1000, 5000, 10000 and 50000, the end-to-end average delays of the system are 6.2ms, 7.5ms, 9.1ms and 12.3ms respectively. Delay rises gently with increasing concurrency, and even under the high-concurrency scenario of 50000 requests, the overall delay remains within 15ms.

The new power system features ubiquitous access of massive distributed terminals. Devices such as photovoltaics, energy storage and distribution terminals initiate communication simultaneously, creating huge concurrency pressure. Meanwhile, businesses such as power dispatching and fault alarm have strict delay requirements, and millisecond-level response is a mandatory indicator for security systems. The proposed system is built on the Apache Flink streaming computing framework and Kafka message bus, adopting streaming processing throughout the entire chain of data collection, transmission, computing and instruction issuance, replacing the traditional offline analysis architecture. The probe layer uses lightweight collection and local preprocessing to reduce raw data transmission; the message bus performs traffic peak shaving and asynchronous distribution; the core computing module schedules tasks in parallel to ensure processing efficiency under high concurrency.

Compared with the second-level or even minute-level processing delay of traditional security platforms, the millisecond-level response capability of this system enables full-process linkage of risk discovery, decision-making and disposal. It can intervene at the early stage of attacks to prevent rapid risk spread across the power network. Whether in daily operation-and-maintenance scenarios or high-load scenarios such as sudden failures and attack outbreaks, the real-time performance of the system stably meets standards, adapting to the full-time operation requirements of power businesses.

5.3.3 Long-term operation stability

Figure 4 shows the CPU and memory usage curves of the server under continuous 72-hour full-load stress test. During the test period, CPU usage stays stably in the range of 40%–60%, and memory usage remains at 45%–65%. Both curves show very small fluctuations, with no abnormal spikes, drops or resource overflow. The system runs continuously without interruption or downtime.

As critical infrastructure, power systems require security protection platforms to operate 7×24 uninterrupted. System stability, disaster recovery capability and resource management are core prerequisites for practical deployment. The proposed system adopts cluster deployment, master-slave redundancy and cross-data-center multi-active architecture. All core services, middleware and databases support multi-node hot standby, and single-node faults trigger automatic load migration and failover. Mechanisms such as local caching, breakpoint resume and data multi-replicas prevent data loss or service paralysis caused by network jitter and hardware failures.

In terms of resource scheduling, the system supports dynamic elastic scaling, automatically adding or removing computing nodes according to real-time concurrency to balance server load and avoid single-node overload. The 72-hour continuous stress test proves that the system's cluster architecture, resource scheduling strategy and disaster recovery design are mature and reliable, capable of stably carrying massive business traffic for a long time and complying with the operation and maintenance standards of key information systems in the power industry.

5.3.4 Business compatibility and flexible disposal

A special test was conducted for the four-level grayscale control strategy, and the impact rate on normal power business under different strategies was counted: 0% for the observation level, 1.2% for the speed limit level, 3.5% for the reinforcement level, and 0% for the isolation level (which only blocks risk nodes).

Traditional power security devices generally adopt the binary "permit/block" disposal mode, directly cutting off communication links once a risk is identified. In new power systems with blurred boundaries and heterogeneous terminals, this mode easily misjudges legitimate behaviors such as normal operation and maintenance and temporary access, causing business interruption and affecting core work such as power supply and dispatching. This paper innovatively designs a four-level grayscale risk control system (observation, speed limit, reinforcement, isolation), matching differentiated disposal strategies to trust levels for flexible control:

1. Observation level: Only performs traffic mirroring and log retention, with zero intervention to ensure all normal businesses run uninterrupted.
2. Speed limit level: Restricts bandwidth and access frequency for low-credibility subjects, reducing risk impact while preserving basic communication capability.
3. Reinforcement level: Enhances protection through secondary authentication and encryption, suitable for medium-risk scenarios.
4. Isolation level: Only cuts off links for extremely high-risk nodes, achieving precise risk isolation.

The grayscale strategy completely abandons the "one-size-fits-all" protection mode and strikes a balance between security protection and business continuity. For complex scenarios such as distributed terminals in remote mountainous areas and temporary operation-and-maintenance access, the flexible disposal method minimizes the interference of security control on normal power operation, with strong engineering practicability.

5.3.5 Situational visualization and global perception

The whole-network risk distribution heat map in Figure 5 divides the test area into 3 major regions and 9 zones, representing the risk level of each zone through numerical values and color depth for intuitive location of high-risk areas. Combined with supporting functions such as topology visualization, risk drill-down and traceability, operation-and-maintenance personnel can view risk nodes, attack paths and behavior characteristics with one click, realizing full-process traceability from global overview to single-point details.

The new power system covers a wide geographical area with a large number of nodes. Traditional security platforms can only display isolated alarms without forming a global situational view, making it difficult for operation-and-maintenance personnel to grasp the overall security situation. The proposed system integrates device topology, traffic data and risk events across the network, and builds an integrated situational awareness dashboard supporting multi-dimensional display of risk heat, alarm trends, strategy execution and device status. In daily operation and maintenance, managers can quickly identify high-risk areas through the heat map and conduct proactive inspection. During cyber attack outbreaks, they can track risk propagation paths in real time and coordinate devices across the region for joint disposal. Experiments prove that the visualization and global perception capabilities of the system fit the power operation-and-maintenance work mode, and can improve security operation efficiency and emergency response capability.

5.3.6 Comprehensive comparison and engineering value summary

Based on experimental results across five dimensions, the proposed scheme outperforms traditional static IDS and single AI models in all aspects: it achieves a qualitative improvement in identification accuracy, meets the millisecond-level real-time requirements of power business, runs stably for a long time, adapts to complex business scenarios with grayscale strategies, and empowers operation and maintenance with global situational awareness.

Addressing the "distributed, open, dynamic" characteristics of new power systems, this system solves four major pain points of traditional protection systems: blurred boundaries, lagging rules, rigid disposal and difficult operation and maintenance. Meanwhile, the system supports seamless integration with power situational awareness and remote operation-and-maintenance platforms, and can be quickly embedded into the existing power information system without large-scale transformation of original equipment and businesses. The experiments fully verify the scientificity,

advancement and feasibility of this technical scheme, which can provide reference for boundary security construction in various power scenarios such as provincial power grids, regional distribution networks and new energy power stations.

5.4 Experimental Summary

This set of simulation experiments comprehensively verifies the dynamic trust evaluation and grayscale risk control system proposed in this paper from five dimensions: risk identification, response delay, operation stability, business compatibility and situational awareness. Experimental data shows that relying on the multi-dimensional trust model, multi-algorithm fusion analysis, streaming computing architecture and grayscale disposal strategy, the system significantly outperforms traditional security schemes in core indicators such as identification accuracy, real-time performance and stability.

The system can accurately identify both known attacks and new complex threats, while maintaining millisecond-level response under high-concurrency scenarios. The 72-hour continuous operation test confirms its 7×24 operation capability meeting power industry standards. The four-level grayscale strategy effectively balances security protection and business continuity, and the global visualization function greatly improves operation-and-maintenance efficiency. Overall, this scheme fully adapts to the security protection requirements of new power systems, with mature technology and strong practicability, and is qualified for large-scale engineering deployment.

6 CONCLUSION AND PROSPECT

Aiming at the security characteristics of open interconnection and blurred boundaries in new power systems, this paper conducts comprehensive research on dynamic trust evaluation and grayscale risk control technologies. It first sorts out the industry background and domestic and foreign research status, and clarifies existing technical gaps and research directions. Then it expounds the core theories of dynamic trust and grayscale risk control and provides complete quantitative formulas. Next, a six-layer system architecture is designed with full-link planning, and each core engine is disassembled and designed in detail. Finally, system performance is verified through simulation experiments. The results prove that the proposed scheme can effectively improve the boundary security protection level of power systems and adapt to distributed power terminal operation scenarios.

Future research can be deepened in four directions: first, introduce large language models to improve the detection capability for unknown attacks and complex attack chains; second, develop lightweight algorithms and probes to adapt to edge micro power terminals with extremely limited computing power; third, study cross-regional trust linkage technology to realize global collaborative risk control; fourth, expand application scenarios to cover terminal equipment such as charging piles and user-side distributed power sources.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Lu Yujiao, You Qingshan, Lin Xinyan, et al. Robust optimal economic dispatch method for active distribution network with distributed energy storage. Journal of Lanzhou University of Technology, 2020, 46(6): 112-118.

[2] Yang Junhu, Han Xiaoqing, Gao Wenjun. Research on frequency regulation characteristics of droop-controlled microgrid. Shanxi Electric Power, 2013(5): 1-5.

[3] Yu Zhuoyan, Guo Jingqian. Research on active-reactive power coordinated optimal dispatch strategy for active distribution network. China Plant Engineering, 2026(7): 102-104.

[4] Liu Nan, Kou Zhiyu. Multi-objective optimal dispatch of microgrid energy based on bat algorithm. Electronic Design Engineering, 2026, 34(9): 112-115+121.

[5] Su Mingjun. Voltage-stabilized transmission optimization scheme based on distributed generation access. Popular Utilization of Electricity, 2024, 39(11): 32-33.

[6] Wu Xiaobo. Virtual synchronous generator technology in distributed generation. North China Power, 2016(4): 50-53.

[7] Yang Xu, Chen Xi, Gao Jingjing, et al. Large-scale cloud-network-edge-end resource scheduling algorithm integrating tabu search and NSGA-II. Computer Integrated Manufacturing Systems, 1-22[2026-06-08].

[8] Wu Jun, Mu Guohui. Application of distributed battery energy storage in power grid — Review of Optimal Configuration and Dispatch Technology of Distributed Battery Energy Storage System. Battery Bimonthly, 2021, 51(6): 653-654.

[9] He Wenhua, Ding Guili, Han Wei, et al. Research on demand response incentive strategy optimization model considering both power grid carbon benefits and user satisfaction. Distribution & Utilization, 2023, 40(10): 95-105.

[10] Duan Ketong, Han Zijiao, Dong Yannan, et al. Black start control strategy for PV-storage microgrid based on improved micro-source voltage and virtual synchronous generator. Electrical & Energy Management Technology, 2025(10): 28-36.