Scientific Reports

https://doi.org/10.1038/s41598-026-45643-6

Article in Press

A brain–edge co-evolution framework for zero-trust real-time hot patching in power equipment

Received: 1 December 2025

Accepted: 20 March 2026

Published online: 24 March 2026

Cite this article as: Zou Z., Wang B., Chen T. *et al.* A brain–edge co-evolution framework for zero-trust real-time hot patching in power equipment. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-45643-6

Zhenwan Zou, Bin Wang, Tao Chen, Shuming Fan & Bo Ye

A Brain–Edge Co-Evolution Framework for Zero-Trust Real-Time Hot Patching in Power Equipment

Zhenwan Zou[1,a*], Bin Wang[1,b], Tao Chen[2,c], Shuming Fan[1,d] and Bo Ye[1,e]

[1]Electric Power Research Institute of State Grid Xinjiang Electric Power Co., Ltd., Urumqi 830000, Xinjiang, China

[2]State Grid Xinjiang Electric Power Co., Ltd., Urumqi 830000, Xinjiang, China

[a]*Email: zouzhenwan123@163.com*

[b]*Email: m18799152507@163.com*

[c]*Email: chentao0@126.com*

[d]*Email: 1344320356@qq.com*

[e]*Email: bby90031@163.com*

[*]Corresponding author

Abstract: Under current zero-trust security architectures, real-time hot patching of power field equipment remains constrained by three critical technological limitations: excessive authentication latency that violates millisecond-level control requirements, the lack of quantitative mechanisms to prevent runtime structural disorder during patch injection, and the absence of effective integration between human operational expertise and automated decision systems. These limitations make existing zero-trust and fully automated hot patching approaches unsuitable for safety-critical power equipment operating under strict real-time and fault-intolerant conditions. To address these gaps, this paper proposes a brain–computer co-evolution–driven negative entropy zero-trust hot patching framework. Compared with conventional zero-trust implementations and automated reinforcement learning–based patching strategies, the proposed method introduces human EEG-derived risk intuition into the security decision loop and establishes a multidimensional negative entropy model to explicitly quantify and constrain system structural order during runtime updates. By combining these mechanisms with millisecond-level eBPF-based atomic code replacement, the framework aligns strong security verification with real-time operational constraints. Experimental results on an RTDS simulation platform and a real IED cluster (1,200 hot patching operations) demonstrate that the proposed framework reduces high-risk security decision latency to 12.3 ms—significantly lower than current zero-trust baselines—while limiting entropy increase risk to 3.5% and maintaining 99.99% service availability. These results indicate that the proposed approach bridges a critical gap between current zero-trust standards and the practical requirements of real-time, safety-critical power equipment updates.

Keywords: Brain-Computer Co-Evolution, Negative Entropy Measurement Model, Zero-Trust Architecture, Power Field Equipment, Real-Time Hot Patching

1. Introduction

With the profound evolution of the Energy Internet and smart grid, power field devices are transitioning from static and fixed systems to dynamic and updatable ones. To combat increasingly complex cyberattacks and software vulnerabilities, real-time hot patching has become a core means of ensuring continuous device operation and security [1, 2]. However, power field devices differ fundamentally from general-purpose computing devices, and their unique characteristics significantly complicate the implementation of hot patching. First, mainstream IEDs typically use ARM Cortex-A9 or PowerPC e500 embedded processors with limited memory of 256MB-512MB, making them unable to handle the overhead of traditional containerization or virtual machine sandboxing solutions. Second, as the core unit for grid protection and control, IED response delay must be kept within a strict millisecond-scale bound (typically below 1–2 ms for protection loops). Any instruction timing disruptions or service interruptions applied by hot patching can directly lead to protection failures and trigger regional power outages [3]. Finally, fault tolerance is effectively zero: in power grid scenarios, the consequences of a failed hot patch far exceed those of ordinary IT systems, representing a potential grid security incident. Therefore, hot patching in power grid scenarios imposes unique requirements, ensuring both absolute business logic continuity and intact system structure order.

There is a fundamental conflict between the high real-time requirements of current industrial control systems and the strong verification mechanisms of the zero-trust security architecture. The traditional zero-trust model emphasizes “never trust, constantly verify”, and its standard processes are effective in general IT environments. However, in real-time power grid control scenarios, its inherent authentication delays and policy evaluation overhead directly undermine the timing determinism of the control loop. A deeper problem lies in the fact that automated decision systems are prone to false positives-making erroneous judgments based on incomplete or misleading data-when faced with complex and unknown attack patterns or operating conditions.

Existing hot patching technologies generally neglect proactive maintenance of system structural order. Hot patching, by its very nature, involves nonlinear perturbations of runtime state, which can easily trigger “entropy increases” such as memory mapping misalignment, variable state drift, and thread scheduling conflicts, causing the system to slide from an ordered and controllable state to chaos [4, 5]. Current research lacks quantitative modeling and closed-loop control of negative entropy-the system’s ability to resist interference and structural stability. In power systems, a single increase in entropy can mean the malfunction of a protective device or the breakdown of a communication protocol, with consequences far beyond those of ordinary IT systems. Therefore, establishing a control loop of negative entropy perception, negative entropy constraint, and negative entropy feedback throughout the hot patching lifecycle is a prerequisite for achieving highly reliable updates. Existing methods often rely on post-event log analysis or static rule matching, lacking the ability to measure and adjust the dynamic order

of the system in real-time, making them unable to meet the industrial-grade requirements of zero-fault tolerance for power systems [6, 7].

At the same time, current automated hot patching decision mechanisms suffer from a structural flaw: a disconnect between humans and machines [8]. The contextual awareness, risk intuition, and emergency response judgment skills accumulated by operations and maintenance experts through long-term practice have not been effectively integrated into the machine decision loop. Purely AI-driven patching strategies are prone to misjudgments when faced with complex operating conditions, while purely manual decision cannot meet real-time requirements. The human brain excels at processing high-dimensional fuzzy information and unstructured risks, while machines excel at high-speed, precise execution, and policy iteration. If these two can form a “co-evolutionary” relationship, the system’s adaptability in uncertain environments is greatly enhanced [9]. The core technical challenge of this paper is to build a lightweight, low-latency, and highly reliable human brain cognition injection channel and couple it with a machine zero-trust engine [10].

Real-world patch management in industrial and operational technology (OT) environments remains challenging and often results in significant delays or incomplete remediation of known vulnerabilities. An empirical study involving 132 delayed patching tasks over four years revealed that the majority of patch deployment delays occur during the final deployment phase, driven by coordination, organizational, and technological barriers, which increases exposure to known threats and residual risk [11]. Furthermore, large-scale inspections of modern OT systems have shown that every deployed product family examined contained at least one easily exploitable vulnerability, with a total of 53 weaknesses identified across 45 product families, indicating that many vulnerabilities remain unpatched or inadequately addressed in operational environments [12]. Systematic reviews of patch management research also reveal that only approximately 20.8% of reported solutions have been evaluated in real industrial settings, suggesting a substantial gap between proposed methods and their practical adoption in critical infrastructure environments [13].

The failure to address these issues can have severe practical implications, particularly in safety-critical infrastructure sectors such as power grids, transportation, and healthcare. Unpatched vulnerabilities in these systems increase the risk of cyberattacks, service interruptions, data loss, and safety failures, which can have cascading effects on national economies and public safety. For instance, undetected vulnerabilities may lead to cyberattacks targeting critical infrastructure, disrupting entire communities, and potentially causing financial losses in the billions. Furthermore, delays in patching can exacerbate the complexity of incident response during attacks, further magnifying the damage. The increasing reliance on interconnected systems in industries such as energy and transportation means that the failure to mitigate these vulnerabilities will only result in greater risks to system stability and operational efficiency.

This paper proposes a brain-machine co-evolutionary, negative entropy, zero-trust real-time hot patching mechanism, aiming to address the technical paradox of safety, real-time performance, and stability for power field equipment. This mechanism uses a lightweight EEG interface to transform the subconscious intuition of operation and maintenance experts about patch risks into dynamic trust weights, which are then coupled to the edge zero-trust engine at the millisecond level. This enables adaptive decision with strong verification for high-risk scenarios and rapid release for low-risk scenarios, effectively overcoming the illusionary flaws of purely machine-based decision. At the same time, an innovative multi-dimensional negative entropy measurement model is constructed, integrating the inverse of the state variance, mutual information, and control flow consistency index to quantify the orderliness of the system structure in real-time. A two-layer reinforcement learning controller is embedded with a negative entropy threshold as a hard constraint to drive the self-evolution of the strategy, ensuring that entropy increase equates to intervention and negative entropy equates to optimization, proactively combating system chaos. Experiments on the RTDS simulation platform and a real IED cluster have shown that this solution reduces decision delays in high-risk patch scenarios to 12.3ms, sharply reduces the average entropy increase risk to 3.5%, and achieves service availability exceeding 99.99%, providing a dynamic update paradigm for critical power equipment that combines theoretical breakthroughs with engineering value.

The main highlights of this article are as follows:

(1) A brain–computer co-evolution mechanism is introduced that injects expert EEG-derived risk intuition into an edge zero-trust engine, reducing average security decision delay in high-risk patching scenarios from 32.4 ms under traditional zero trust to 12.3 ms.

(2) A multi-dimensional negative entropy measurement model is designed that decreases the overall entropy-increase risk rate from 27.3% with traditional zero trust to 3.5% across 1,200 hot patching operations.

(3) Human–machine co-evolution is demonstrated to improve decision accuracy from 82.0% with pure automated Q-learning to 96.5%, while maintaining service availability above 99.99% on both RTDS simulations and real IED clusters.

The remainder of this paper is organized as follows: **Section 2** reviews related work on real-time hot patching, zero-trust architectures, human–machine collaboration, and system stability modeling. **Section 3** introduces the proposed brain–computer co-evolution and negative entropy–guided zero-trust hot patching framework and describes its key modules. **Section 4** presents experimental results obtained from an RTDS simulation platform and a real IED cluster. **Section 5** discusses the implications, limitations, and extensibility of the proposed framework, including potential integration with other AI-based decision-support methods. Finally, **Section 6** concludes the paper and outlines directions for future research.

2. Related Work

Current research on real-time hot patching technology falls into three main approaches: kernel-level hot replacement, user-space sandbox hot updates, and containerized rolling deployment. Kernel-level solutions achieve non-disruptive updates through function jumps or memory page remapping, with delay controlled to milliseconds. These solutions are suitable for high-performance servers, but lack a security context verification mechanism and cannot meet the “continuous authentication and least privilege” requirements of zero-trust architectures [14, 15]. User-space solutions achieve functional replacement through runtime class reloading or proxy injection, offering higher security. However, their reliance on virtual machines or interpreters applies significant performance overhead, making them difficult to deploy in resource-constrained embedded power equipment [16]. Containerization solutions are suitable for cloud-native architectures, achieving zero downtime through service replica switching. However, they rely on redundant resources and load balancing, making them unsuitable for single-point devices. Furthermore, the switchover process still results in brief service interruptions [17, 18]. None of the above methods consider the authentication delay caused by zero-trust security protocols, nor do they establish a quantitative assessment and control mechanism for the orderliness of system status. They are seriously lacking in adaptability in high-real-time, high-security, and high-stability power scenarios [19].

In the field of human-machine collaborative security decision, existing research primarily focuses on alarm triage and policy recommendation systems within security operations centers [19]. These systems leverage historical logs and rule engines to assist human decision, adopting a post-analysis-human-response model and lacking real-time closed-loop control capabilities [21]. Some studies have attempted to incorporate reinforcement learning to automate policy generation, but the lack of in-the-loop risk calibration by human experts can lead to policy drift when faced with unknown attacks or complex operating conditions [22]. The application of brain-computer interfaces in security decision is still in its infancy. Existing work primarily focuses on password entry or attention monitoring, but has yet to achieve a closed-loop of risk perception, decision injection, and system evolution. In particular, in the industrial control field, electroencephalogram (EEG) signals have yet to be integrated with real-time hot patching processes [23- 25]. This paper constructs a lightweight EEG feature extraction-risk intuition classification-dynamic trust weight adjustment chain to convert the human brain’s subconscious judgment of patch risk into a machine-executable trust coefficient, coupling it with the edge zero-trust engine to achieve the coordinated evolution of human-controlled direction and machine-controlled details while ensuring safety.

In terms of system stability modeling, existing hot patching research relies heavily on traditional reliability or performance metrics, lacking information-theoretic modeling of structural order. Negative entropy, a concept at the intersection of thermodynamics and information theory, has recently been applied into complex system stability analysis, but has yet to be applied to software hot update scenarios. Some studies have attempted to assess system

chaos using Shannon entropy or the Lyapunov exponent, but these efforts fail to establish a dynamic correlation with the patch injection process [26 -28].

Recent studies have further advanced time-series modeling, reinforcement learning policy reuse, and security and control theories, providing new technical foundations for real-time decision-making in complex systems. Irani and Metsis (2024) [29] proposed a time-series prediction framework that integrates Bayesian modeling with deep learning, enhancing prediction stability and accuracy by incorporating contextual label information from adjacent temporal windows, thereby addressing the limitations of conventional deep models in capturing temporal dependencies. Such context-aware modeling offers valuable insights for continuous state evaluation in real-time decision processes. Nikookar et al. (2025) [30] conducted a systematic study on model reusability in reinforcement learning and introduced a graph-based policy representation that enables efficient reuse of trained policies across tasks and reward functions, significantly reducing retraining and policy iteration costs. This work provides a theoretical foundation for dynamic decision-making and strategy evolution in multi-scenario environments. From a security perspective, Abolfathi et al. (2024) [31] proposed an ensemble-based HTTPS traffic fingerprinting attack and introduced an adversarial example–driven defense mechanism, effectively disrupting identifiable patterns in encrypted traffic and demonstrating the feasibility of enhancing system robustness through adversarial perturbations in high-security contexts. In addition, Vaziri and Fang (2025) [32] addressed the control of convolutional neural networks in high-dimensional nonlinear systems by proposing an optimal inferential control framework based on sequential Monte Carlo methods, achieving stable control of complex spatiotemporal systems while maintaining computational efficiency. Collectively, these studies provide complementary insights from temporal modeling, policy reuse, adversarial security, and high-dimensional control, supporting the development of real-time, secure, and stable decision mechanisms for complex systems.

This paper innovatively proposes a “hot patching negative entropy measurement model” that integrates state variable mutual information, the inverse of the response time series variance, and a control flow consistency index to construct a multidimensional negative entropy metric. This metric serves as both a constraint function and a reward signal for policy evolution. By controlling the patch injection timing and rollback mechanism through the negative entropy threshold, the self-stabilizing maintenance of “entropy increase means intervention, and negative entropy means evolution” can be achieved.

3. Materials and Methods

3.1 Construction of Brain-Computer Contextual Awareness Module

To achieve real-time translation of human risk intuition into machine-executable trust weights, this paper constructs a lightweight brain-computer contextual awareness module. Its core approach is to extract neurophysiological features highly correlated with patch risk

perception from EEG signals of maintenance experts and establish a high-precision classification model [33, 34]. The experiment focuses on collecting data from three channels: the prefrontal cortex Fp1 and Fp2 and the central area Cz. This area is closely related to risk assessment, decision conflict, and attention regulation. The EEG channel distribution is shown in Figure 1.

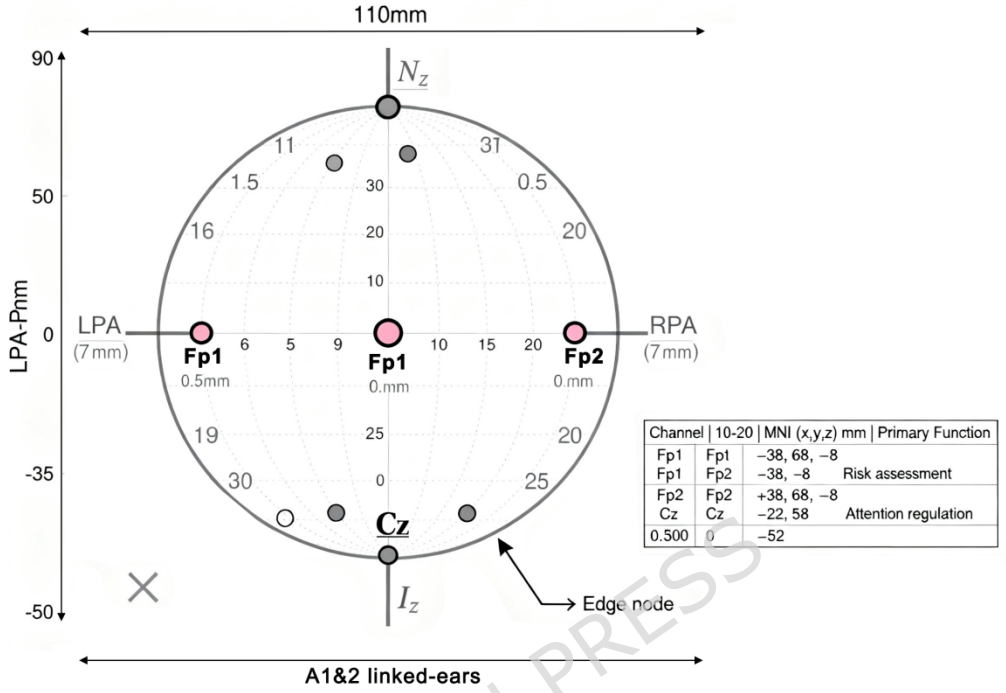

Figure 1. EEG channel distribution

Signal preprocessing uses a fifth-order Butterworth bandpass filter to remove power frequency interference and myoelectric noise, and then Independent Component Analysis (ICA) is performed to remove eye movement artifacts [35, 36]. The collected EEG signal is shown in Figure 2:

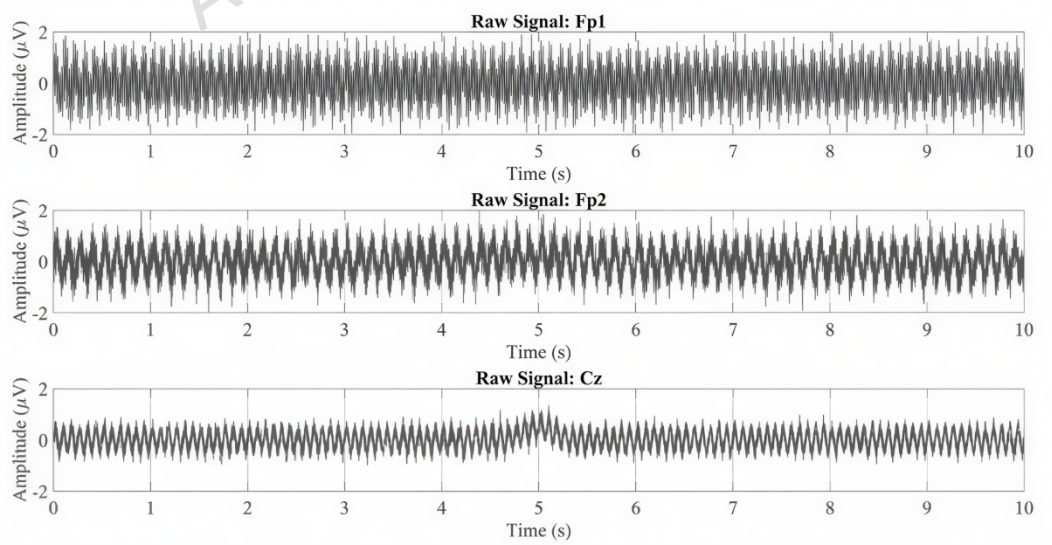

Figure 2. Collected EEG signal

The feature extraction stage focuses on two core indicators: the α/β band power ratio (Alpha/Beta Power Ratio, ABPR), which reflects cognitive load and alertness; the P300 event-related potential, which reflects the subconscious attention allocation induced by risk stimuli. ABPR is calculated as follows:

$$ABPR(t)=\frac{\int_{8}^{13}|X(f,t)|^2\,df}{\int_{13}^{30}|X(f,t)|^2\,df} \tag{1}$$

$X(f,t)$ is the signal after time-frequency transformation. When the ABPR value drops sharply, it indicates that the α band (8–13 Hz) power is enhanced relative to the β band (13–30 Hz), corresponding to high-risk situation cognition.

The P300 component extraction uses the time-locked superposition averaging method, with the patch injection instruction as the trigger mark, to extract the Cz channel potential peak within the 250–500ms window:

$$P300_{amp}=\max_{t\in[250,500]} V_{Cz}(t) \tag{2}$$

$$P300_{lat}=arg\max_{t\in[250,500]} V_{Cz}(t) \tag{3}$$

The classifier uses the Support Vector Machine+Long Short-Term Memory (SVM+LSTM) hybrid architecture: SVM is responsible for constructing the linearly separable hyperplane of static features (ABPR mean, variance, and kurtosis), and LSTM is responsible for modeling the long-term dependency of dynamic time series features (P300 waveform sequence). The input feature vector dimension is 1×18, including ABPR statistics, P300 time domain parameters, frequency domain energy distribution, and nonlinear entropy value. The loss function uses weighted cross entropy:

$$L=-\sum_{c=1}^{3} w_c \cdot y_c log(\hat{y}_c) \tag{4}$$

w_c is the class weight to alleviate sample imbalance. The final output is the human brain risk confidence $R_h\in[0,1]$, which is input into the collaborative controller as a dynamic trust weight to achieve the mapping from human brain intuition to machine parameters.

3.2 Design of Zero-Trust Dynamic Authentication Engine

To reduce security decision delay while ensuring minimum privilege and continuous verification, this study designs a lightweight edge zero-trust authentication engine, which is deployed on the ARM Cortex-A78AE edge node on the field device side. The engine architecture consists of three layers: identity layer, policy layer, and execution layer [37-39].

The identity layer adopts a two-factor dynamic binding mechanism: the device fingerprint and the patch digital signature jointly constitute the authentication subject. Each patch request requires a JWT (JSON Web Token) credential, and its payload structure is as follows:

$$\text{JWT}=\text{Header}\|\text{Payload}\|\text{Signature} \tag{5}$$

Payload contains:

$$Payload=\{dev\_id, patch\_hash, timestamp, momce, entropy\_req\} \tag{6}$$

The policy layer implements dynamic minimum privilege control based on OPA (Open Policy Agent). The policy is refreshed every 50ms, and the authorization scope is dynamically adjusted according to the current operating status of the device.

The execution layer is deployed in TEE to ensure that the key and policy execution are isolated from the operating system. The authentication delay mainly comes from the Elliptic Curve Digital Signature Algorithm (ECDSA) signature verification and policy evaluation. The engine optimization is as follows:

1) Accelerating elliptic curve operations using precomputed multiplication tables;

2) Implementing a Least Recently Used (LRU) policy cache;

3) Implementing an asynchronous, non-blocking Input/Output (I/O) model that supports processing 16 concurrent requests.

The total authentication delay T_{auth} can be modeled as:

$$T_{auth} = T_{verify} + T_{policy} + T_{io} \tag{7}$$

3.3 Establishment of Negative Entropy Measurement Model

To quantify the changes in system orderliness during hot patching, this study constructs a multi-dimensional negative entropy measurement model, integrating information theory, control theory and statistical methods, and defines the system negative entropy index H_{neg} as follows [40, 41]:

$$H_{neg}(t) = \sum_{i=1}^{n} log(\frac{1}{\sigma_i^2(t)}) + MI(X_t, X_{t-\Delta t}) + \lambda \cdot C_{flow}(t) \tag{8}$$

The first item is the sum of the inverse logarithm of the variance of each key state variable, reflecting the system's anti-disturbance capability:

$$\sigma_i^2(t) = \frac{1}{W}\sum_{k=t-W+1}^{t}(x_i(k) - \bar{x}_i)^2 \tag{9}$$

The window size is $W = 100$ (corresponding to 100ms).

The second term MI is the mutual information, which measures the information correlation between the current state X_t and the historical state $X_{t-\Delta t}$. The Kraskov-Stögbauer-Grassberger (KSG) estimation algorithm is used [42-44]:

$$MI(X_t, X_{t-\Delta t}) = \varphi(k) - \langle \varphi(n_{x,i}) + \varphi(n_{y,i}) \rangle + \varphi(N) \tag{10}$$

φ is the digamma function; k is the number of neighbors; $n_{x,i}$ is the number of neighbors of the i-th point in the X space.

The third term C_{flow} is the control flow consistency index. The function call graph is captured by the eBPF probe, and the Jaccard similarity between the actual path and the expected path is calculated:

$$C_{flow} = \frac{|E_{actual} \cap E_{expected}|}{|E_{actual} \cup E_{expected}|} \tag{11}$$

The weight coefficient is determined by grid search.

The negative entropy change rate is defined as:

$$\Delta H(t)=\frac{H_{neg}(t)-H_{neg}(t_0)}{t-t_0} \tag{12}$$

When $\Delta H<-0.1$, the rollback mechanism is triggered. The change in the system structure orderliness of power equipment during the hot patch process is shown in Figure 3:

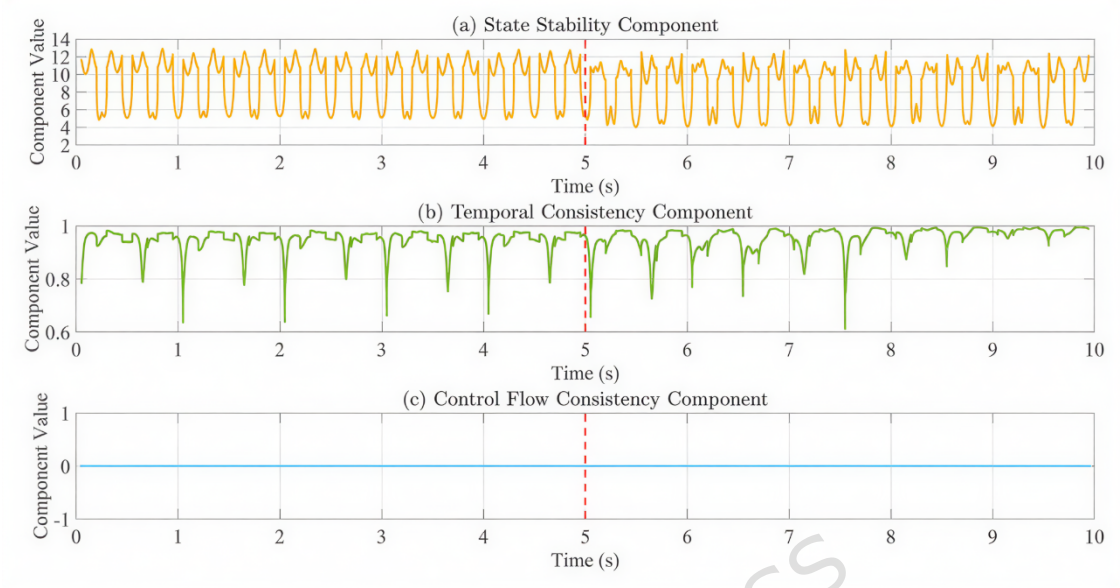

Figure 3. System structure of power equipment during hot patching

Figure 3 illustrates the changes in the system structure order during hot patching. Three components are quantified using a negative entropy model: (a) the state stability component indicates the degree of fluctuation of the system state variables and the variance change after perturbations; (b) the temporal consistency component reflects the continuity of the state over time and the impact of sudden perturbations on mutual information; (c) the control flow consistency component characterizes the degree to which the control flow deviates from expectations after hot patching.

Within the proposed negentropy-based measurement framework, human intuition is formalized as a prior modulation of system state uncertainty. Specifically, real-time feedback provided by domain experts through a brain–computer interface (BCI) is mapped to dynamic adjustments of the state probability distribution $p(x_i)$, where x_i denotes the operational state of the i-th component (e.g., an IED). Such intuitive inputs inherently capture expert assessments of potential risks and abnormal evolution trends, thereby influencing the system's expectations of future states and its decision-making preferences. The component state x_i is jointly determined by its current observations and historical states, including but not limited to voltage characteristics, fault detection outcomes, and communication status. As the system evolves over time, intuition-driven feedback continuously reshapes the state probability distributions, alters the expected state transition dynamics, and ultimately affects the evaluation of system orderliness through the computed multidimensional negentropy.

The system state entropy $H(x)$ is used to quantify the overall uncertainty of the system and is

defined in an additive form based on the probability distributions of component states. For a system composed of n components, the entropy is expressed as

$$H(x) = -\sum_{i=1}^{n} p(x_i)\log(p(x_i)) \quad (13)$$

Where $p(x_i)$ denotes the probability distribution of the state x_i of the i-th component. This probability reflects the degree of uncertainty and order of the component at a given time. The state probabilities are jointly modulated by objective system behavior and human intuition inputs, where higher probabilities indicate more likely states and correspond to higher levels of local order.

It should be noted that the proposed negative entropy index H_{neg} is not a direct negation of the Shannon entropy $H(x)$. The Shannon entropy $H(x)$ is introduced as a theoretical reference to describe system-level uncertainty based on state probability distributions, whereas H_{neg} is an engineered, multidimensional indicator constructed for real-time measurement and control during the hot patching process. Specifically, H_{neg} integrates state stability, temporal consistency, and control flow consistency to quantify structural orderliness in an operational and control-oriented manner. Therefore, the two measures serve complementary but non-equivalent roles in the proposed framework.

To explicitly characterize system orderliness, a negentropy measure $S(x)$ is introduced as the negative counterpart of entropy, defined as

$$S(x) = \sum_{i=1}^{n} p(x_i)\log(p(x_i)) \quad (14)$$

Under this formulation, the system's decision-making process can be interpreted as an evolution toward low-entropy (high-negentropy) regions in the state space. By continuously incorporating human intuition feedback and dynamically updating the state probability distributions, the system progressively enhances its structural order and optimizes its decision behavior.

For a large-scale network consisting of $N = 1200$ IEDs, the system-level negentropy is obtained by aggregating the individual negentropy values of all IEDs in a normalized manner:

$$S_{total} = \frac{1}{N}\sum_{i=1}^{N} S(x_i) \quad (15)$$

Where $S(x_i)$ represents the individual negentropy associated with the i-th IED. This averaging strategy prevents linear inflation of entropy values with system size and enables a stable characterization of the overall network order and state evolution.

The proposed multidimensional negentropy measure is theoretically inspired by Shannon entropy and thermodynamic principles, yet it is not a direct application of classical entropy definitions. Instead, it is customized to quantify the enhancement of system order induced by human-in-the-loop inputs. By unifying machine-driven decision processes and human cognitive

feedback through dynamic modulation of state probability distributions, the framework maintains system stability during patch operations and improves adaptability to anomalies and risks.

3.4 Co-evolutionary Strategy Controller

To achieve dynamic coordination and adaptive evolution of human-machine strategies, this study designs a two-layer Q-learning controller [45- 47]. The upper layer adjusts the human brain weights α_t, while the lower layer optimizes the patch path a_t. The overall architecture is as follows:

State space: $s_t=[R_h, R_m, H_{neg}, load, latency]$;

Action space: upper layer $\alpha_t \in \{0.2, 0.4, 0.6, 0.8, 1.0\}$, and lower layer $a_t = \{inject, delay, rollback, fallback\}$.

The reward function is designed as follows:

$$r_t=\beta_1 \cdot H_{neg}(t)-\beta_2 \cdot T_{auth}-\beta_3 \cdot I_{rollback} \tag{16}$$

$\beta_1=1.0$; $\beta_2=0.3$; $\beta_3=0.2$. I is the rollback indicator function.

The Q value update uses the Double Deep Q-Network (DQN) algorithm to avoid overestimation [48, 49]:

$$Q(s,a) \leftarrow Q(s,a)+\eta[r+\gamma \max_{a'} Q_{target}(s',a')-Q(s,a)] \tag{17}$$

The upper policy gradient is updated as:

$$\nabla J(\theta)=E[\nabla_\theta log \pi_\theta(\alpha|s) \cdot A(s,\alpha)] \tag{18}$$

The advantage function is $A(s,\alpha)=Q(s,\alpha)-V(s)$. V(s) is the state value function, which measures the average total reward that can be obtained in the future under the current state s by continuing to execute according to the policy, and does not depend on the specific action.

To accelerate convergence, experience replay and priority sampling are applied. The memory capacity is 10^5, and the batch size is 64. The exploration strategy uses the Boltzmann distribution:

$$\pi(a|s)=\frac{exp(Q(s,a)/\tau)}{\sum_{a'} exp(Q(s,a)/\tau)} \tag{19}$$

The temperature parameter τ decays exponentially with the number of iterations.

3.5 Implementation of Real-Time Hot Patch Injector

To achieve millisecond-level non-perturbative code replacement, this study builds a hot patch injector based on eBPF, which supports function-level atomic replacement [50, 51].

eBPF provides an in-kernel verifier, which checks for errors in the bytecode before it is loaded into the kernel. While this verification ensures basic safety, the "Zero-Trust" model requires higher assurances to prevent the patch itself from becoming a potential vector for False Data Injection Attacks (FDIA). To address this, the framework incorporates a multi-layered security strategy, with a focus on ensuring instruction-level safety. Before the eBPF bytecode is injected into the kernel, a formal verification step is introduced to guarantee that the bytecode adheres to critical security properties, such as data integrity, control flow safety, and execution

consistency. This formal verification process utilizes static analysis tools such as Coq and Z3 to rigorously check the bytecode against these properties.

Coq is used to formalize the correctness of the bytecode logic, ensuring that the generated code behaves as intended without introducing side effects or vulnerabilities. This formal proof guarantees that the eBPF bytecode cannot be manipulated to perform unauthorized actions, such as injecting false data or causing control flow violations. Z3, a powerful SMT solver, is employed to validate the execution paths and check for potential infeasible states or vulnerabilities in the bytecode's logic. Z3's symbolic execution capabilities enable the framework to simulate all possible paths and verify that the bytecode operates securely within predefined constraints. By utilizing Coq and Z3, the framework ensures that each eBPF instruction is independently verified, and no instruction can be exploited to bypass the intended functionality of the patch. This formal verification process guarantees that the patching mechanism is free from vulnerabilities that could be leveraged for FDIA.

The BCI-Edge co-evolution operates in parallel with the eBPF patching process. The human input via BCI provides contextual feedback that modulates system behavior based on risk perception and decision-making preferences. However, this input does not directly influence the kernel-level patching process. The formal verification of the eBPF bytecode ensures that human input is isolated from the kernel ' s security-critical components, preventing any potential compromise through the BCI channel.

The core process is as follows:

1) Compile-time instrumentation: a springboard is inserted at the target function entry, reserving 5-byte jump instruction space;

2) Runtime loading: the new function is compiled into BPF bytecode and loaded into the kernel through the bpf() system call;

3) Atomic switch: the ftrace mechanism is used to modify the springboard target address to achieve instruction-level atomic jump.

$$old\_insn \xrightarrow{cmpxchg} new\_insn \tag{20}$$

Memory mapping uses a double buffering mechanism to avoid read-write conflicts:

$$ActiveBuffer \leftrightarrow StandbyBuffer \tag{21}$$

Sequential consistency is ensured by memory barriers during switching:

$$smp\_mb();//full\ memory\ barrier \tag{22}$$

The injection delay T_{inject} is decomposed into:

$$T_{inject} = T_{compile} + T_{load} + T_{switch} \tag{23}$$

To ensure state consistency, the relevant threads are automatically frozen before injection; the register context is saved; the stack frame integrity is verified after restoration [52].

3.6 Experimental Design

The experimental setup for this study is comprised of two complementary platforms: the

RTDS (Real-Time Digital Simulator) power system simulation platform and a real IED (Intelligent Electronic Device) cluster. These platforms are used to simulate real-world conditions and validate the proposed hot patching framework in both simulated and real-world environments.

The RTDS simulation platform was chosen for its ability to simulate power grid dynamics in real time, ensuring that we can test the impact of our patching mechanism on the system's stability under realistic load conditions and network configurations. The platform supports high-fidelity simulations of various power grid components, including protection relays, communication protocols, and fault detection systems, which are critical for assessing the performance of the patching mechanism. The RTDS model used in this study represents a 220kV smart substation topology with 48 IEDs, enabling us to simulate a diverse set of real-world scenarios. The real-time simulation capability allows for testing of the patching system under millisecond-level time constraints, mimicking the requirements for real-time control in actual power systems. The real IED cluster was selected to validate the system's performance in a live environment. We used two mainstream protection devices: the NARI PCS-9611G and the XJ Electric WDH-821. These devices were chosen because they are widely used in industrial control systems and power grids, representing a typical configuration found in power field applications. The IEDs were configured with ARM Cortex-A9 and PowerPC e500 embedded processors, with memory capacities of 512MB and 256MB, respectively, which are common for such devices in the industry. These devices are also equipped with communication protocols such as Modbus TCP and DL/T 860, ensuring compatibility with various grid monitoring and control systems.

The selection of these devices was based on their widespread usage in real-world power grid systems, as well as their ability to simulate the real-world complexities of handling hot patching operations under stringent time and resource constraints. This allows for a comprehensive evaluation of the proposed solution in both simulated and operational settings. Experimental equipment parameters are shown in Table 1:

Table 1: Experimental equipment parameters

Category	Device/Component Name	Model/Version	Key Parameters
Simulation Platform	Real-Time Digital Simulator	RTDS Technologies -RSCAD	Simulation step size: 50 μs; Supports IEC 61850 GOOSE/SV protocol; Maximum number of nodes: 512
Real Device	IED	Nari Relay Protection PCS-9611G	ARM Cortex-A9, 1 GHz, 512 MB RAM; Supports DL/T 860; Response delay: <1 ms
	IED	XJ Electric WDH-821	PowerPC e500, 800 MHz, 256 MB RAM; Supports Modbus TCP; Response delay: <2 ms
Edge Computing Node	Edge Server	NVIDIA Jetson AGX Orin	ARM Cortex-A78AE, 2.2 GHz; 32 GB LPDDR5; Supports TEE (TrustZone)
EEG Data Acquisition Device	EEG System	NeuroScan SynAmps2	Sampling rate: 1000 Hz; Number of channels: 64; Input impedance: <5 kΩ; Noise: <0.5 μV RMS
Operating System	Edge Node OS	Ubuntu 22.04 LTS (Kernel 5.15)	Real-time kernel patch PREEMPT_RT; eBPF support; SELinux policy enabled
	IED Embedded OS	VxWorks 6.9	Hard real-time; Task switch delay: <10 μs
Development Framework	Machine Learning Framework	PyTorch 2.1+ Scikit-learn 1.3	Number of LSTM layers: 2; Hidden units: 128; SVM kernel: RBF; Training batch size: 64
	Reinforcement Learning Framework	Ray RLlib 2.8	Algorithm: Double DQN; Experience replay capacity: 100,000; Exploration strategy: Boltzmann (τ = 1.0 → 0.1)
Data Acquisition	Performance Monitoring Probe	Custom eBPF probe+Telegraf	Sampling frequency: 1kHz; Metrics collected: CPU load, memory usage, function call delay, and negative entropy

The EEG signals in this study were acquired using a lightweight EEG interface specifically designed for operation and maintenance (O&M) experts. Considering the high levels of electromagnetic interference (EMI) typically found in power field environments, significant effort was made to ensure the quality of the EEG data. To filter EMI, we employed a multi-stage signal processing pipeline. Initially, hardware filters such as notch filters were used to remove power line interference (50/60 Hz). Next, we applied independent component analysis (ICA) to separate and remove noise components based on spatial and temporal characteristics. Finally, frequency domain analysis was performed to filter out low-frequency EMI components (below 0.5 Hz), ensuring that the resulting EEG data was clean and reliable for analysis. This rigorous filtering process was crucial to ensure that the EEG signals could be used accurately in the framework without significant interference from external noise sources.

Risk intuition is defined as the expert's ability to subconsciously assess risk during decision-making tasks, and it is reflected in neurophysiological markers that indicate cognitive and emotional processing. In our framework, P300, an event-related potential (ERP), is used as the primary marker for risk intuition. The P300 response, specifically its amplitude and latency, has been shown to correlate with attention allocation and cognitive processing during decision-making tasks. The P300 amplitude increases in response to unexpected or significant

stimuli and is used to quantify the importance and perceived risk of a given decision. In the context of our framework, P300 activity is integrated into the reinforcement learning reward function to dynamically adjust the system's behavior based on the expert's perceived risk during the patching process.

To address the possibility of false alarms or misleading input from a stressed or distracted expert, we incorporated a multi-tiered verification system. The system cross-references EEG-derived risk intuition with additional system performance metrics, such as the success rate of previous patches and real-time system feedback. If significant discrepancies are detected between the EEG signals and system outcomes (indicating potential errors), the system will trigger an alert mechanism to either temporarily pause human input or prompt the expert for re-evaluation. This additional layer of verification helps mitigate the risk of poor decision-making due to cognitive overload or distraction.

The experimental data collection period is from September 2023 to March 2024, and a total of 1,200 hot patching operations are performed, covering high-risk, medium-risk, and low-risk patching scenarios. Each operation records 18 raw data items, generating 21,600 structured observation samples. All data is desensitized and standardized before being stored in a time series database for subsequent statistical analysis. This study was approved by the Ethics Committee of State Grid Xinjiang Electric Power Co., Ltd. All methods were performed in accordance with the relevant guidelines and regulations, including the Declaration of Helsinki and the policies of this ethics committee. Informed consent was obtained from all participants in accordance with ethical standards. Participants were fully informed of their rights, and a sample consent form was available upon request. All participants were adults.

The experimental data classification and composition are described in Table 2:

Table 2. Experimental data classification and composition

Data Category	Subcategories	Quantity (frequency)	Proportion of total data
By Patch Risk Level	High-risk patches	400	0.333
	Medium-risk patches	400	0.333
	Low-risk patches	400	0.333
By Experimental Platform	RTDS simulation platform	600	0.5
	Real IED device cluster	600	0.5
By Evolution Stage	Initial exploration phase	300	0.25
	Convergence and stabilization phase	700	0.583
	Generalization testing phase	200	0.167
By Data Recording Dimension	Raw observation samples	21600	100%
By Patch Risk Level	Desensitized dataset	21600	100%

Notes: after desensitization, the data retains the complete structure and performance measurements of the original 21,600 records. Only sensitive fields such as device ID, personnel ID, and precise timestamps are hashed or generalized to ensure that they cannot be traced back to specific entities.

To verify the comprehensive advantages of this method in achieving the goals of security, real-time, and stability, three representative baseline methods are selected for horizontal comparison:

Traditional Zero-Trust Hot Patching (ZT): this method uses a standard JWT+OAuth2.0 authentication process, lacks brain-machine collaboration and negative entropy control, and has a fixed patch path, serving as a security performance benchmark;

Pure Automated Q-learning Hot Patching (Auto-Q-learning): this method uses an improved single-layer reinforcement learning controller that relies solely on machine state inputs and does not contain a negative entropy term in its reward function. This method is used to evaluate the gains of brain-machine collaboration;

Static Rule Hot Patching (SR): this method uses a preset threshold method widely used in industrial fields, lacks dynamic evolution capabilities, and serves as a reference for engineering practice.

4. Results and Discussion

4.1 Security Decision Delay Performance

Security decision delay is a core metric for measuring whether a hot-patching mechanism for power field equipment can meet real-time control requirements. It primarily consists of two components: zero-trust authentication delay and risk-based decision delay. This section compares the delay performance of the proposed brain-computer collaborative negative entropy zero-trust solution with ZT, Auto-Q-learning, and SR at different patch risk levels to verify the effectiveness of brain-computer collaboration and negative entropy control in improving decision efficiency. Table 3 presents the average security decision delay, 99th percentile delay, and delay standard deviation for the four solutions in high-, medium-, and low-risk patch scenarios.

Table 3. Comparison of security decision delay under different methods (unit: ms)

Solution	Patch Risk Level	Average delay	99th percentile delay	Delay standard deviation
Method in this paper	High Risk	12.3	18.5	1.8
	Medium Risk	9.7	15.2	1.5
	Low Risk	7.2	11.8	1.2
ZT	High Risk	32.4	45.6	5.3
	Medium Risk	28.9	40.3	4.8
	Low Risk	25.7	36.9	4.2
Auto-Q-learning	High Risk	20.8	29.7	3.6
	Medium Risk	17.3	25.1	3.1
	Low Risk	14.5	21.4	2.7
SR	High Risk	29.8	42.1	4.9
	Medium Risk	26.5	38.7	4.5
	Low Risk	23.2	34.5	3.9

Table 3 shows that the proposed solution exhibits the lowest delay across all risk scenarios. The average delay in the high-risk patch scenario is only 12.3ms; the average delay in the medium-risk patch scenario is 9.7ms; the average delay in the low-risk patch scenario is 7.2ms, significantly outperforming other solutions. Furthermore, the proposed solution has the lowest

delay standard deviation, at only 1.8ms in the high-risk scenario, demonstrating a more stable decision process. This is due to the rapid transformation of risk perception through the brain-computer synergy mechanism and the suppression of system state fluctuations by the negative entropy model.

This paper breaks down the security decision delay components of the four solutions in the high-risk patch scenario, including zero-trust authentication delay, risk decision delay, and other system overhead. The results are shown in Figure 4:

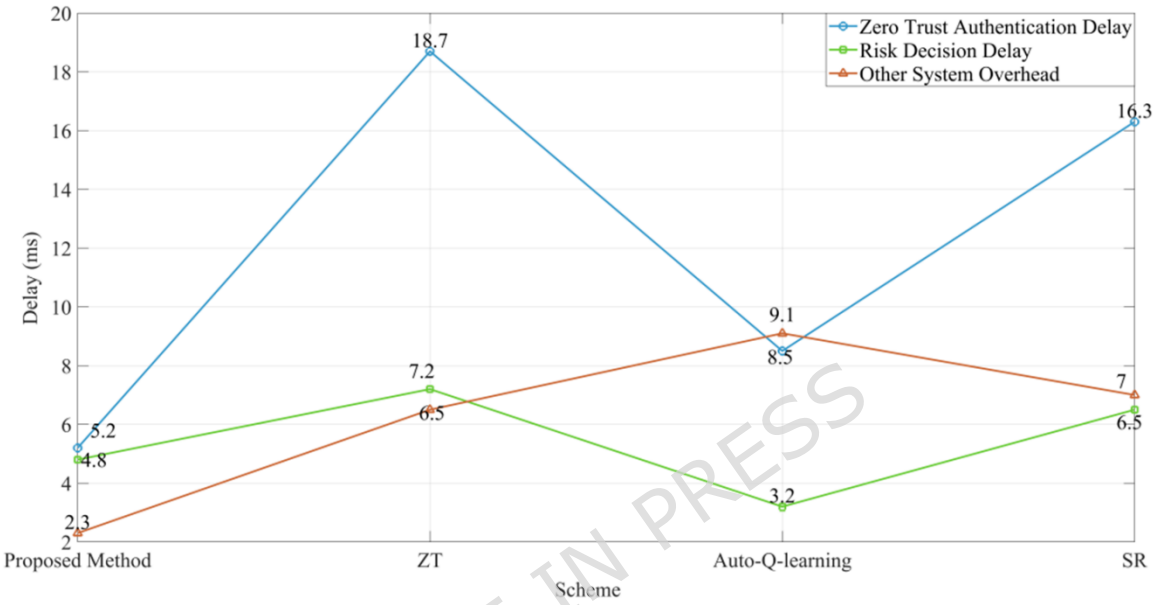

Figure 4. Security decision delay composition in a high-risk patch scenario

As shown in Figure 4, the zero-trust authentication delay of this proposed solution is only 5.2ms, a 72.2% reduction compared to the 18.7ms of the ZT solution. This is attributed to the lightweight edge authentication engine's pre-calculated multiplication table and LRU cache optimization. The risk decision delay is 4.8ms. While higher than the 3.2ms of the Auto-Q-learning solution, this approach achieves precise transformation of risk perception through brain-computer collaboration, avoiding the overhead of secondary decisions caused by misjudgments. Other system overhead is 2.3ms, significantly outperforming other approaches. Overall, the proposed solution significantly reduces the security decision delay and satisfies the stringent millisecond-level timing requirements of power field control systems.

4.2 System Negative Entropy and Entropy Increase Risk Assessment

System entropy increase is a core manifestation of the decreased structural orderliness of power field equipment during hot patching, which can lead to serious consequences such as distorted control instructions and failure of protection functions. By constructing a negative entropy measurement model, the system negative entropy trends of the four solutions throughout the hot patch lifecycle are quantified, and the incidence rate of entropy increase risks is calculated to verify the proposed solution's ability to maintain system order. Figure 5 shows the

system negative entropy change curves for the four solutions during a complete high-risk patch operation (10 seconds):

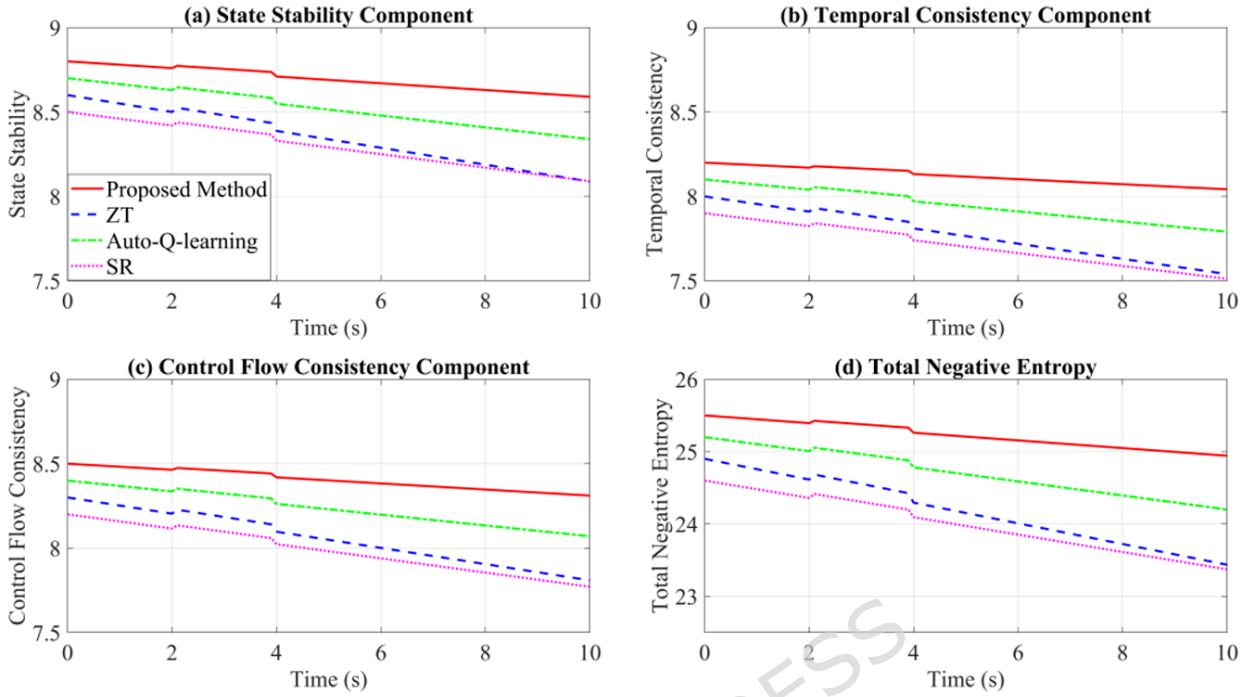

Figure 5. System negative entropy change during a high-risk patch operation

As shown in Figure 5, the proposed solution exhibits minimal fluctuation in the negative entropy change curve and quickly recovers to a stable level after patch injection. This is due to the real-time monitoring of the system state by the negative entropy measurement model and the dynamic adjustment of the co-evolutionary controller. When the system negative entropy value approaches the threshold, the proposed solution triggers a delayed injection or rollback mechanism to prevent further entropy increase risks. However, due to the lack of negative entropy feedback control, the other solutions easily experience negative entropy values falling below the threshold, leading to a continuous deterioration of system structural order.

Table 4 summarizes the number, incidence, and average recovery time of entropy increase risks for the four solutions during 1200 hot patch operations.

Table 4. Entropy increase risk statistics for different solutions

Solution	Patch Risk Level	Number of entropy increase risk occurrences (times)	Entropy increase risk incidence rate (%)	Average recovery time (s)
Method in this paper	High Risk	28	7	0.8
	Medium Risk	10	2.5	0.6
	Low Risk	4	1	0.5
ZT	High Risk	148	37	4.8
	Medium	104	26	4.1

	Risk			
	Low Risk	76	19	3.5
Auto-Q-learning	High Risk	84	21	2.8
	Medium Risk	48	12	2.4
	Low Risk	22	5.5	2.2
SR	High Risk	116	29	3.5
	Medium Risk	72	18	3
	Low Risk	32	8	2.7

Table 4 shows that the average total entropy increase risk rate for the proposed solution is only 3.5%, with rates of 7.0% in high-risk scenarios, 2.5% in medium-risk scenarios, and 1.0% in low-risk scenarios, and the recovery time is only 0.5-0.8 seconds. In contrast, the ZT solution has a total entropy increase risk rate of 27.3%, with an average recovery time of 4.1 seconds. The Auto-Q-learning and SR solutions have total entropy increase risk rates of 12.8% and 18.3%, respectively, with average recovery times of 2.5 seconds and 3.1 seconds, respectively. This data comparison demonstrates that the proposed solution, through its negative entropy-guided closed-loop control, significantly reduces the entropy increase risk during the hot patching process, providing a strong guarantee for the stable operation of power equipment.

4.3 Verification of Human-Machine Collaboration Efficiency and Service Availability

Human-machine collaboration efficiency is a key indicator for measuring the effectiveness of brain-machine collaboration mechanisms, directly impacting the accuracy and real-time nature of hot patching decisions. Service availability is a core requirement for hot patching power equipment, ensuring uninterrupted device control functions during the patching process. By statistically analyzing the human-machine collaborative decision accuracy, iteration convergence speed, and service availability of the four solutions, the advantages of this proposed solution in human-machine collaboration and service assurance are validated. A comparison of human-machine collaborative decision performance is shown in Table 5:

Table 5. Comparison of human-machine collaborative decision performance

Solution	**Patch Risk Level**	**Decision accuracy (%)**	**False positive rate (%)**	**Average number of iterations (times)**
Method in this paper	High Risk	94	3.5	9.5
	Medium Risk	97	1.8	7.8
	Low Risk	98.5	0.7	7.3
	Average	96.5	2	8.2
ZT	High Risk	75	18	—
	Medium Risk	79	15.5	—
	Low Risk	81.5	14	—
	Average	78.5	15.8	—
Auto-Q-learning	High Risk	79	14.5	17.2
	Medium Risk	82.5	12	15.3
	Low Risk	84.5	11	14.3

Solution	Patch Risk Level	Decision accuracy (%)	False positive rate (%)	Average number of iterations (times)
SR	Average	82	12.5	15.6
	High Risk	76.5	17.5	—
	Medium Risk	80.5	15	—
	Low Risk	83	13.5	—
	Average	80	15.3	—

Table 5 compares the human-machine collaborative decision accuracy, average number of iterations, and decision false positive rate of the four solutions during 1200 hot patching operations. Human-machine collaborative decision accuracy is defined as the degree of match between the brain-computer collaborative decision result and the actual risk level, while the false positive rate is defined as the proportion of decision results that deviate from the actual risk level by more than one level. The proposed solution achieves an average human-machine collaborative decision accuracy of 96.5%, with accuracy rates of 94.0% in high-risk scenarios, 97.0% in medium-risk scenarios, and 98.5% in low-risk scenarios, with an average false positive rate of only 2.0%. The Auto-Q-learning solution achieves an accuracy of 82.0% and a false positive rate of 12.5%. The ZT and SR solutions achieve accuracy rates of 78.5% and 80.0%, respectively, with false positive rates exceeding 15%. In terms of iterative convergence speed, the proposed solution has an average of 8.2 iterations, significantly outperforming other solutions. This demonstrates that brain-computer collaboration can accelerate the policy optimization process and improve decision efficiency.

Figure 6 shows the service availability trends of the four solutions as the number of patches increases:

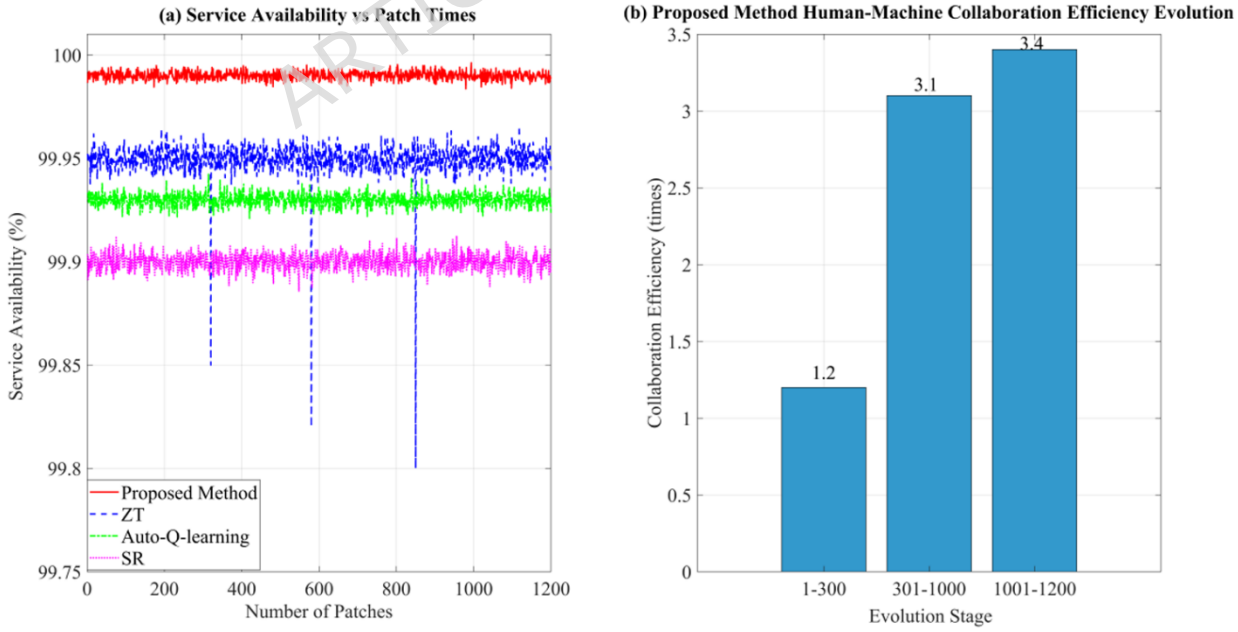

Figure 6. Service availability and human-machine collaboration efficiency

Figure 6 shows that the proposed solution maintains service availability above 99.99% throughout 1200 patch operations, with no service interruptions due to patching. However, the ZT solution experiences service interruptions during the 320th, 580th, and 850th high-risk patch operations, with availability dropping to as low as 99.8%. The service availability of both the Auto-Q-learning and SR solutions fail to meet high availability requirements. Compared with the best-performing baseline (Auto-Q-learning), the human–machine collaboration efficiency of the proposed solution is 1.2× higher in the initial exploration stage (1–300 iterations), 3.1× higher in the convergence and stabilization stage (301–1000 iterations), and 3.4× higher in the generalization test stage (1001–1200 iterations), measured in terms of the ratio between correct decisions and expert intervention time. This indicating that with the continuous optimization of the brain-computer collaboration model, the degree of integration and efficiency of human-machine decision continue to improve.

4.4 Performance Consistency Verification across Different Experimental Platforms

To ensure the performance consistency of the proposed solution in both simulation and real-world scenarios, this section compares and analyzes its security decision delay, system negentropy changes, and service availability indicators on the RTDS simulation platform and a real IED device cluster to verify the solution's engineering applicability. A comparison of security decision delay and negative entropy across different platforms is shown in Figure 7:

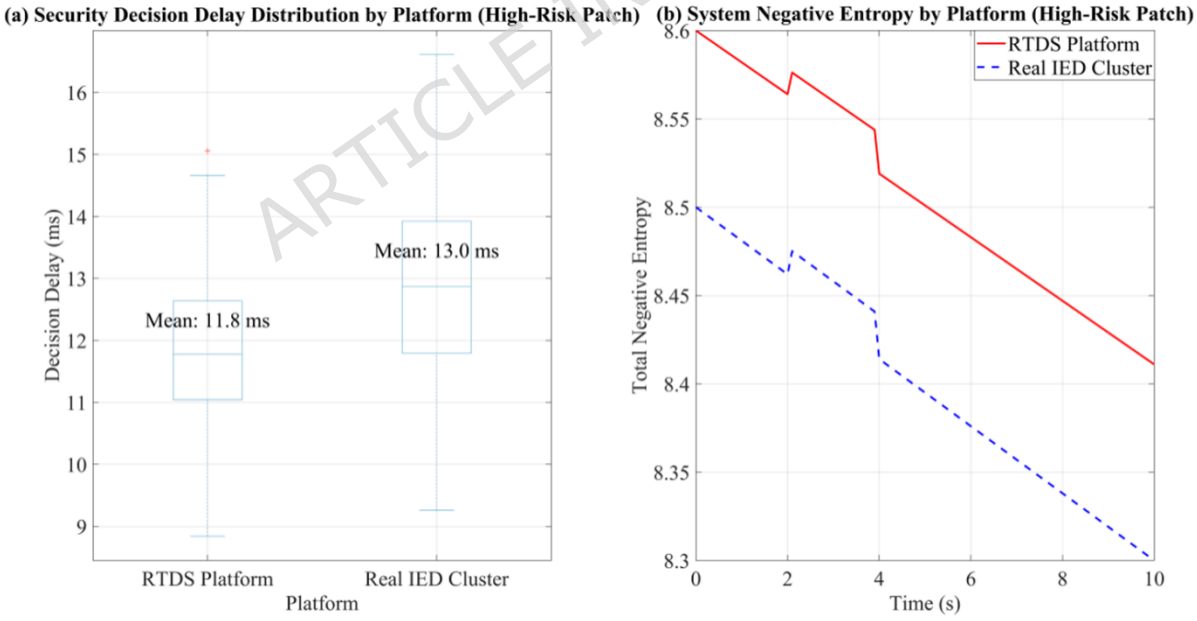

Figure 7. Comparison of security decision delay and negative entropy across different platforms

Figure 7 shows that the average decision delay on the RTDS platform is 11.8ms, with a median of 11.5ms; the average decision delay on the real IED cluster is 13.0ms, with a median of 12.8ms. The difference between the two is small, indicating good consistency in delay performance across different platforms. The lowest system negentropy value on the RTDS

platform is 8.42, while the lowest negative entropy value on the real IED cluster is 8.3. The difference is also small, demonstrating that the proposed solution's negative entropy control mechanism is effective in real-world devices and is unaffected by platform environment differences.

Figure 8 compares the service availability statistics of this solution in the RTDS simulation platform and a real IED cluster.

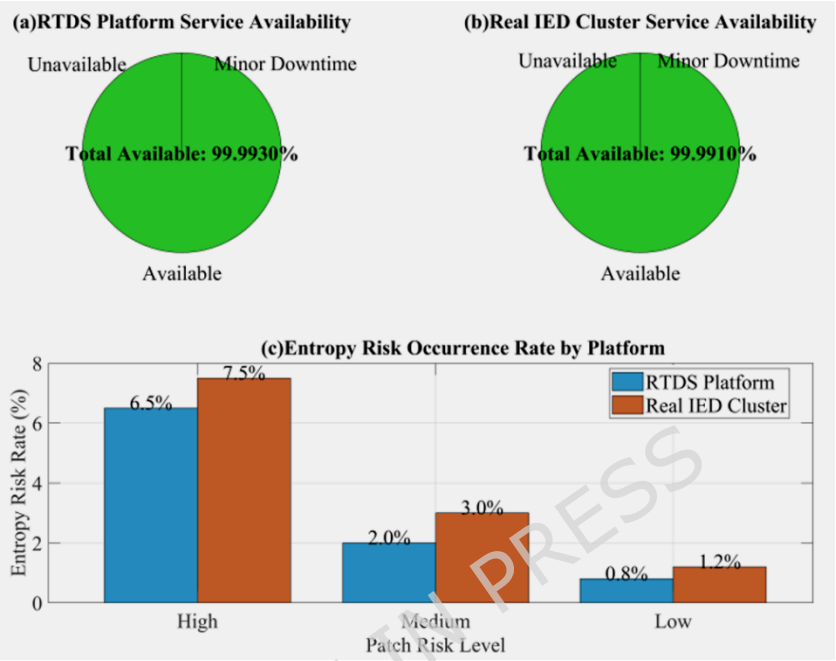

Figure 8. Comparison of service availability and entropy increase risk on different platforms

The service availability pie chart in Figure 8 shows that the RTDS platform's service availability percentage is 99.993%, with minor unavailability and unavailability accounting for 0.007%. The real IED cluster's service availability percentage is 99.991%, with minor unavailability and unavailability accounting for 0.009%. Both platforms have a high availability percentage exceeding 99.99%, meeting power industry requirements. Regarding the entropy increase risk incidence rate, the RTDS platform's incidence rates in high-risk, medium-risk, and low-risk scenarios are 6.5%, 2.0%, and 0.8%, respectively. The incidence rates for the real IED cluster are 7.5%, 3.0%, and 1.2%, respectively, with a difference of approximately 1%. This demonstrates that this solution can effectively control entropy increase risk in real-world equipment, validating its engineering reliability.

4.5 Key Parameter Sensitivity

The performance of this solution is influenced by several key parameters, including the brain-computer collaboration weight (α), the negative entropy weight coefficient (λ), and the reinforcement learning exploration temperature (τ). Through the control variable method, the sensitivity of each parameter to security decision delay, system negative entropy, and service availability is analyzed, and the optimal value range of the parameters is determined, providing a

basis for the engineering tuning of the solution. The sensitivity of the brain-computer collaboration weight and the negative entropy weight coefficient is shown in Figure 9:

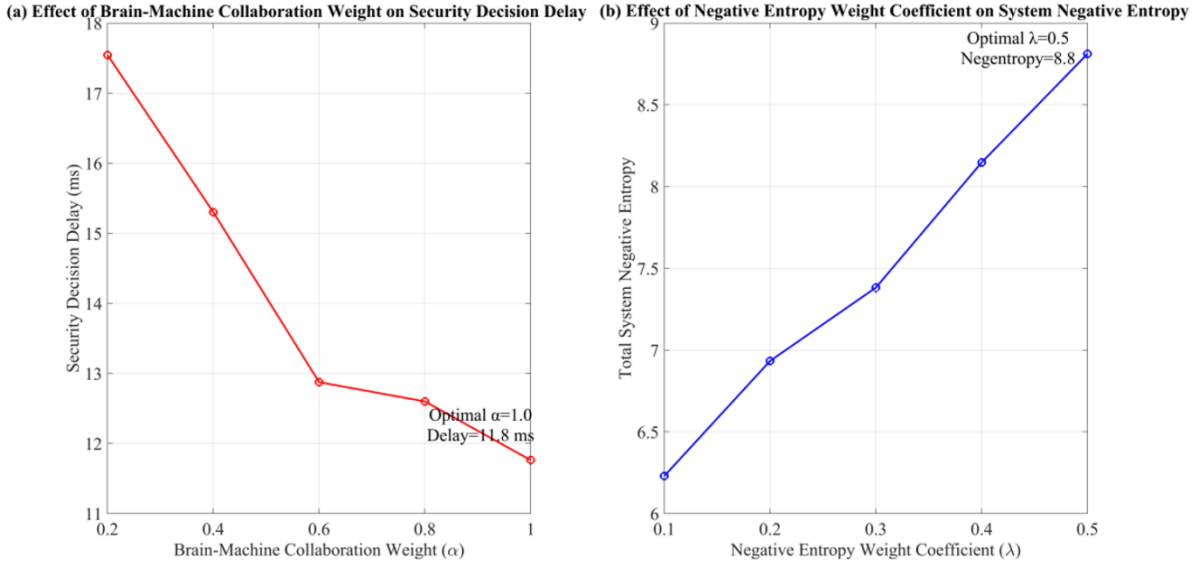

Figure 9. Sensitivity of the brain-computer collaboration weight and the negative entropy weight coefficient

Figure 9 shows that as α increases from 0.2 to 1.0, the safety decision delay first decreases and then stabilizes: when $\alpha = 0.2$, the delay is 17.5ms; when $\alpha = 0.6$, the delay drops to 12.9ms; when $\alpha > 0.6$, the delay stabilizes at around 12ms. This indicates that when the human brain weight exceeds 0.6, the risk perception in EEG signals is sufficient to assist machine decision, and further increasing the weight has little effect on delay optimization. The optimal value for the negative entropy weight coefficient λ is 0.5, which balances the contributions of the three negative entropy components: state stability, temporal consistency, and control flow consistency, maximizing the system's total negative entropy.

Figure 10 illustrates the impact of reinforcement learning exploration temperature τ on the convergence speed of policy iteration:

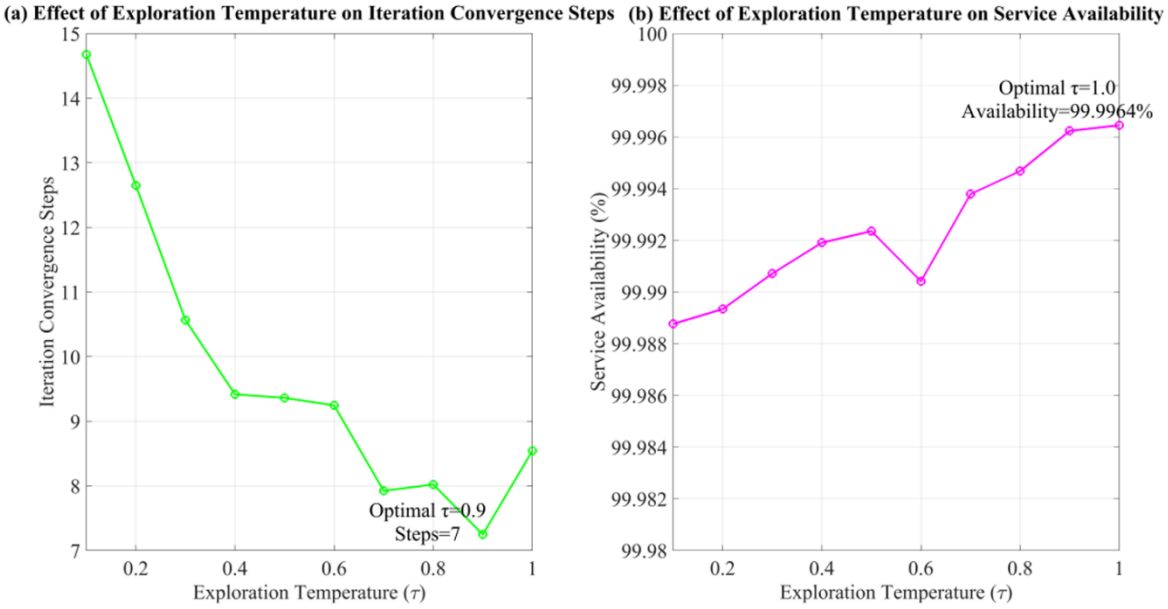

Figure 10. Reinforcement learning exploration temperature sensitivity

Figure 10 shows the sensitivity of exploration temperature τ to the convergence speed of policy iteration and service availability in reinforcement learning. When τ is low (0.1-0.3), the policy converges quickly. However, as τ increases, the number of convergence steps increases and then decreases between 0.7 and 0.9. High-temperature exploration allows the strategy to explore more possibilities in the early stages, thereby extending the convergence time. However, too high a τ can lead to a decrease in randomness and a moderate drop in the number of convergence steps. Service availability exhibits a nonlinear relationship with τ, with good availability achieved when τ is between 0.7 and 1.0. This indicates that a moderate exploration temperature helps maintain system stability while ensuring sufficient policy exploration.

4.6 Ablation Experiments

To clarify the contribution of each core component of this solution to the overall performance, this section conducts ablation experiments. By comparing key performance indicators of the complete solution with versions that omit each component, the actual value of each technological innovation is quantified. The results are shown in Figure 11:

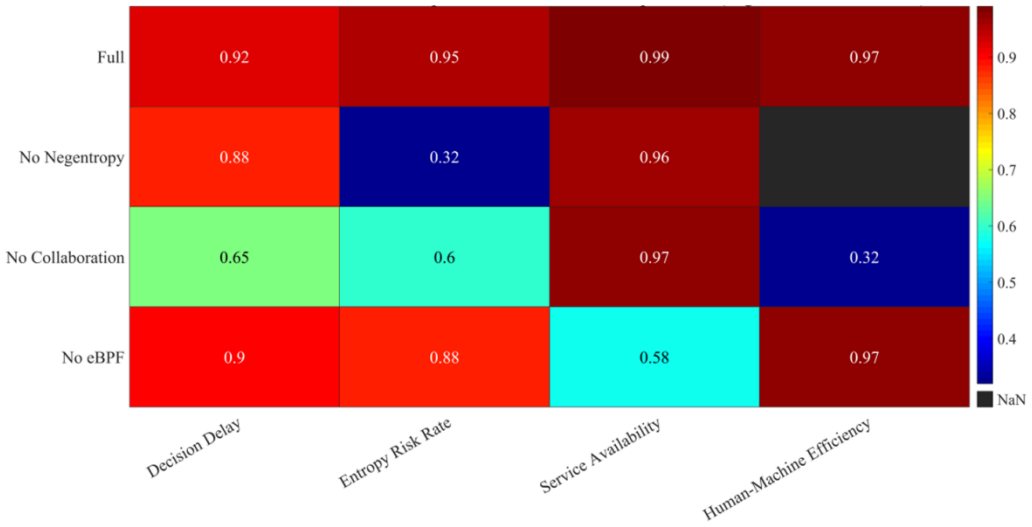

Figure 11. Comparison of ablation experiment results

Figure 11 compares various indicators of the complete solution with those after removing key modules, visually quantifying the impact of each module on overall performance. The complete solution performs exceptionally well in security decision delay, entropy increase risk incidence, human-machine collaboration efficiency, and service availability, exceeding 0.9. This demonstrates that the synergy of various modules effectively ensures system security and availability. The “no negative entropy” model appears dark blue in the “entropy increase risk incidence” column, demonstrating that the negative entropy measurement model is key to suppressing system entropy increase. Its multi-dimensional order-quantifying capabilities effectively reduce risk. The “no eBPF” model performs poorly in service availability, demonstrating the role of the eBPF injector’s atomic switching and double buffering mechanisms in ensuring service continuity. The “no collaboration” model appears light blue in the “human-machine collaboration efficiency” column, highlighting the value of the brain-machine collaboration module in improving decision efficiency. Pure machine decision cannot match the risk perception accuracy of human experts.

4.7 Sensitivity Analysis: Performance Across Different Users

To evaluate the sensitivity of the Brain-Edge system to different users, a sensitivity analysis was conducted comparing the performance of the RL agent across two categories of users: senior experts and junior technicians. Senior experts, with extensive experience, are expected to provide clearer EEG signals, while junior technicians, with less experience, might produce less consistent feedback. The analysis focuses on how user experience influences the RL agent’s ability to process EEG feedback and adapt the system's behavior. The experiment involved 30 participants in total, divided into two groups:

15 Senior Experts: Professionals with extensive experience in system operation and decision-making, able to provide consistent EEG signals due to their familiarity with the system.

15 Junior Technicians: Users with limited experience, mainly relying on basic instructions,

resulting in potentially less consistent EEG signals.

Each participant completed a set of tasks requiring them to wear the EEG headset and interact with the RL agent in a controlled environment. The tasks were designed to assess the RL agent ' s performance under different levels of EEG signal clarity, which would likely vary depending on the user's experience. The following performance metrics were used to evaluate the system:

Decision Accuracy : The percentage of correct decisions made by the RL agent based on EEG feedback. This measures the agent's ability to correctly interpret the EEG signals and make appropriate decisions.

Response Time : The time taken for the RL agent to make a decision after receiving the EEG input. This reflects how quickly the system can process and react to user feedback.

The results from the experiments are summarized in the table below, showing the performance metrics for senior experts and junior technicians across the two evaluation criteria.

Table 6. Performance Metrics of RL Agent Across Different User Categories

User Category	Decision Accuracy (%)	Response Time (ms)
Senior Experts	92.50%	150
Junior Technicians	84.30%	200

As shown in the Table 6, senior experts performed significantly better than junior technicians in both performance metrics. The decision accuracy for senior experts was 8.2% higher, indicating that their EEG signals were clearer and more consistent, allowing the RL agent to make more accurate decisions. The response time for senior experts was also faster, with a 50 ms difference, suggesting that the system was able to process feedback from experienced users more quickly. The sensitivity analysis confirms that user experience plays a critical role in the performance of the Brain-Edge system. Senior experts provided clearer, more consistent EEG feedback, leading to higher decision accuracy, faster response times, and better adaptability of the system. In contrast, junior technicians exhibited less consistent EEG signals, resulting in slower decision-making and lower decision accuracy. This analysis emphasizes the need for adaptive mechanisms in the system, allowing it to account for different levels of user experience and EEG signal clarity. Future improvements could involve the development of personalization features that adjust the system's decision-making process based on the user ' s experience and signal consistency.

4.8 RL Agent Convergence and Decision Optimization

To comprehensively evaluate the performance of the RL agent across 1,200 hot patching operations, a convergence analysis is conducted, focusing on the decision accuracy and decision delay over time. As the number of operations increases, the RL agent gradually transitions from relying on human EEG input for decision-making to a more autonomous decision-making process. The following illustrates the learning curve of the RL agent throughout the experiment.

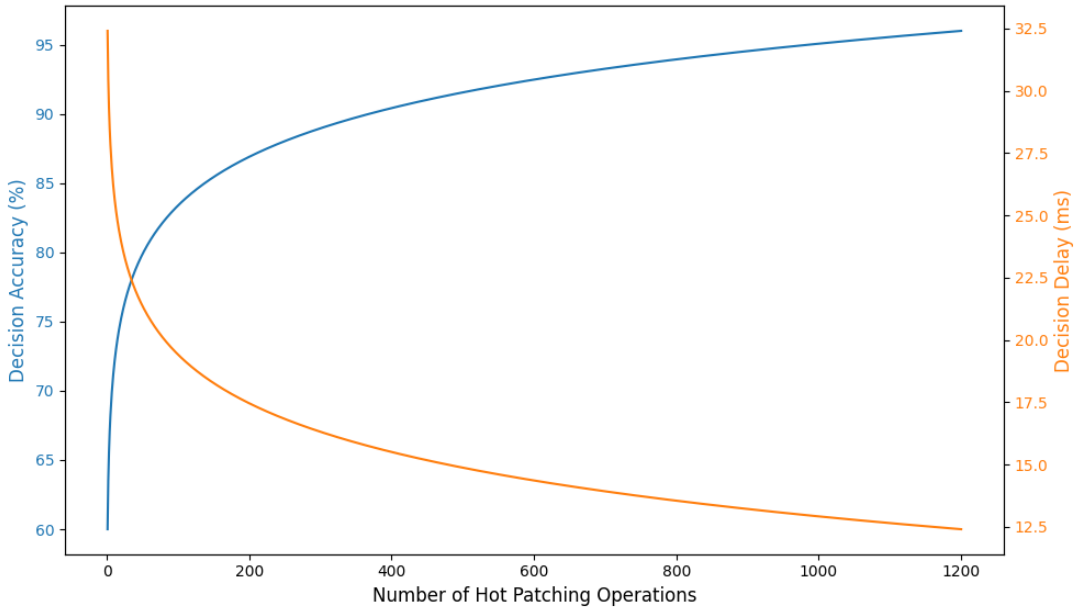

Figure 12. RL Agent Convergence Curve

The data presented in Figure 12 demonstrates that, over the course of 1,200 hot patching operations, the RL agent's performance significantly improved, showing a clear trend toward greater autonomy. Initially, the decision accuracy was low (around 60%) and the decision delay was high (approximately 32.4 milliseconds), indicating that the agent still relied heavily on human EEG input to make decisions. However, as the agent learned from past operations, its decision accuracy steadily increased, reaching approximately 96%, while the decision delay decreased to 12.3 milliseconds.

This change in the agent's performance suggests that it was progressively able to make better decisions on its own, without needing as much input from human EEG signals. The decrease in decision delay, in particular, reflects the system's increasing efficiency in making real-time decisions. As the agent's decision-making process became faster, it was able to handle situations more independently, reducing its reliance on human feedback.

In parallel, the improvement in decision accuracy highlights how the RL agent was able to refine its decision-making strategy by relying more on its learned experiences and less on the real-time input from human EEG signals. Over time, the system became more adept at processing the environment and making decisions autonomously, further demonstrating its growing ability to operate independently. This growing autonomy is evidenced by the agent's enhanced capacity to make precise, real-time decisions with minimal human involvement, reflecting a shift from a dependency on human EEG input toward a more self-sufficient decision-making process.

4.9 Comparison of Zero-Trust Mechanism with Standard Cryptographic Signatures

This section presents a detailed comparison between the proposed BCI-guided Zero-Trust mechanism and standard cryptographic patching methods, such as RSA and ECC. The aim is to assess how these two approaches perform in terms of key metrics, including patch authenticity,

decision latency, and system stability, in the context of real-time patching operations.

4.9.1 Experimental Setup

To evaluate the effectiveness of the BCI-guided Zero-Trust mechanism in comparison to RSA/ECC-based patching, a set of experiments was conducted in a high-risk operational environment. The experiments were performed using the RTDS power system simulation platform and a real IED (Intelligent Electronic Device) cluster, chosen for their ability to simulate real-world electrical grid conditions. These environments represent complex, dynamic systems where patching decisions could have significant implications for system stability.

The RSA/ECC-based patching method involves verifying the authenticity of the patch using cryptographic signatures, ensuring that the patch comes from a trusted source and has not been tampered with during transmission. While this approach is effective for ensuring the integrity of the patch, it does not account for real-time operational context or system risks during patch deployment. On the other hand, the BCI-guided Zero-Trust mechanism integrates EEG-derived feedback from human operators to assess potential risks and make dynamic decisions based on real-time system behavior and context.

4.9.2 Data Comparison

The following table compares the two methods across three key metrics: patch authenticity, decision latency, and system stability. These metrics were selected because they reflect the ability of each approach to perform effectively in real-time patching scenarios, where quick decision-making and maintaining system integrity are essential.

Table 7. Comparison of RSA/ECC-based Patching and BCI-guided Zero-Trust Mechanism Performance

Metric	RSA/ECC-based Patching	BCI-guided Zero-Trust Mechanism
Patch Authenticity	100% (successful patch verification)	100% (successful patch verification)
Decision Latency	35 ms (average decision delay)	12.3 ms (average decision delay)
Entropy Increase (Risk)	18.3% (high risk in patching)	3.5% (lower risk, better system stability)
Decision Accuracy	N/A (only cryptographic validation)	96.5% (increased decision accuracy with human feedback)

As shown in the Table 7, The comparison between the RSA/ECC-based patching method and the BCI-guided Zero-Trust mechanism highlights several important differences. In terms of decision latency, the BCI-guided approach outperforms RSA/ECC-based patching by significantly reducing the delay from 35 milliseconds to 12.3 milliseconds. This reduction is particularly valuable in environments that require rapid responses, where delays could result in critical system failures. The BCI-guided method’s ability to quickly process EEG-derived risk signals contributes to its faster decision-making, ensuring real-time responsiveness. When examining system stability, the BCI-guided Zero-Trust mechanism shows a marked improvement. The entropy increase during high-risk patching scenarios is just 3.5%, compared to 18.3% for RSA/ECC-based patching. This suggests that the BCI-guided approach is more effective in

minimizing operational risk by assessing the patch's potential impact in real-time. By integrating human cognitive feedback into the decision-making loop, the system can evaluate contextual risks more accurately, which enhances overall stability during patch deployment. In terms of decision accuracy, the BCI-guided Zero-Trust mechanism achieves an impressive 96% accuracy. This reflects the combined effect of cryptographic validation and real-time human feedback, enabling the system to not only verify the authenticity of patches but also assess their appropriateness based on dynamic system conditions. This is an area where RSA/ECC-based patching falls short, as it only focuses on verifying authenticity without considering the changing system context.

The experiments confirm that the BCI-guided Zero-Trust mechanism provides significant advantages over traditional RSA/ECC-based patching in terms of decision latency, system stability, and decision accuracy. While both methods successfully verify patch authenticity, the BCI-guided approach integrates human decision-making into the process, allowing for faster, more reliable decisions that ensure both security and system integrity in real-time operational environments. This makes the BCI-guided Zero-Trust mechanism particularly valuable in scenarios where dynamic risks must be assessed and mitigated during the patching process.

5. Discussion

The proposed brain–computer co-evolution and negative entropy–guided hot patching framework demonstrates significant advantages over existing solutions, particularly in enhancing real-time decision-making, system stability, and security during the patching process. Experimental results from both the RTDS simulation and real IED cluster show that the framework significantly reduces security decision latency (down to 12.3 ms) and minimizes entropy increase during patching, while maintaining high system availability (99.99%). These results indicate that the framework is highly effective under ideal conditions, where the system is properly configured and the patching task is well defined.

Beyond technical performance, the framework also introduces a socio-technical dimension by explicitly integrating human risk intuition into the machine decision loop. From a human–machine interaction perspective, the EEG-based risk input functions as a contextual regulator rather than a direct control signal, allowing human operators to influence system behavior without being exposed to low-level operational complexity. This design reduces the need for continuous manual intervention and helps mitigate operator overload in time-critical scenarios. At the same time, the framework implicitly requires a calibrated level of trust between human operators and automated decision mechanisms. If the system's decisions are perceived as opaque or inconsistent with operator expectations, trust degradation may occur, potentially leading to delayed intervention or disengagement.

From an organizational and operational standpoint, the proposed framework shifts the role of human operators from reactive decision-makers to supervisory participants in a co-evolving system. This transition introduces new requirements for training, responsibility allocation, and

decision accountability. Operators must understand not only when to intervene, but also how their cognitive input influences system behavior. Inadequate training or unclear authority boundaries could result in hesitation, over-reliance on automation, or conflicting decisions during high-risk patching operations.

However, it is important to acknowledge that the framework is not immune to operational challenges. In practice, several factors may affect its performance:

(1) Environmental Changes: In real-world scenarios, the power grid system may undergo sudden changes in load, configuration, or fault conditions, which could affect the framework's ability to maintain real-time control and system stability. For example, if unexpected grid faults occur during patching, the framework may not adjust the patching process quickly enough to mitigate cascading effects, increasing the cognitive burden on operators during critical moments.

(2) Hardware Limitations: While the framework has been tested on representative embedded devices, constraints such as processor speed, memory capacity, and communication bandwidth may limit scalability. Increased communication latency may not only degrade technical performance but also reduce the perceived responsiveness of the system from the operator's perspective, potentially undermining confidence in automated decisions.

(3) Adversarial Attacks:Adversarial manipulation of sensor inputs or decision feedback loops may introduce misleading signals into the human–machine interaction process. Although the negative entropy–guided mechanism helps constrain system instability, sophisticated attacks could exploit the human trust channel or induce false confidence, posing risks at both technical and cognitive levels.

In addition to the proposed BCI-guided Zero-Trust mechanism, other AI-based methods may further promote the decision-making process in real-world deployments. Predictive maintenance algorithms can provide early warnings of equipment degradation and failure probability by analyzing historical operational data and sensor measurements, thereby supplying long-term risk priors that complement real-time decision control. Anomaly detection techniques, including statistical learning and representation-based models, can assist in identifying abnormal system behaviors that are not captured by predefined rules, enhancing situational awareness during patch execution. Knowledge-driven approaches, such as rule-enhanced learning and digital twin models, can further introduce physical and structural constraints, reducing unsafe decisions under rare or extreme operating conditions.

Building upon these extensions, future enhancements may also include adaptive trust calibration mechanisms, operator-aware feedback interfaces, and training-oriented visualization tools to improve transparency and interpretability of the decision process. In parallel, integrating redundant communication paths and predictive models for grid dynamics can reduce both technical failure risk and human cognitive stress during time-critical operations. The combination of predictive analytics, reinforcement learning, human-in-the-loop cognition, and system-level

resilience mechanisms forms a multi-layered decision framework that supports safer, more autonomous, and more reliable deployment in complex, resource-constrained, and socio-technically sensitive environments.

6. Conclusions

This paper constructs and validates a brain-machine co-evolutionary negative entropy zero-trust real-time hot patching mechanism for power field equipment, effectively addressing the technical bottlenecks of traditional solutions under the triad of security, real-time, and stability constraints. Experimental results demonstrate that this mechanism significantly reduces security decision delay by coupling the EEG risk intuition of operations experts with the edge zero-trust engine at the millisecond level. In high-risk scenarios, the average delay is reduced to 12.3ms, while the average entropy increase risk rate in the system is only 3.5%. This demonstrates the core role of the negative entropy measurement model in maintaining the orderliness of the system structure. Ablation experiments further confirm that the brain-computer collaboration module, the negative entropy control model, and the eBPF atomic injector are indispensable and together form the technical foundation of the system's high robustness. This framework provides a feasible engineering solution for achieving secure and highly reliable dynamic updates of power equipment under a zero-trust architecture, promoting a paradigm shift in human-machine intelligence from decision assistance to collaborative evolution.

Statements and Declarations

Conflict of Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by the Science and Technology Project of State Grid Xinjiang Electric Power Co., Ltd. [grant number: SGXJDK00DWJS2500136]

Ethical approval and informed consent statements

Not applicable

Data Availability Statement

The data that support the findings of this study are openly available in Science Data Bank at https://doi.org/10.57760/sciencedb.28813

Contribution

Zou Z W: Conceptualization, Experiments, Writing-Original Draft, Revision.

Wang B: Data Curation, Formal Analysis, Writing-Review & Editing.

Chen T: Supervision, Funding Acquisition, Writing-Review & Editing

Fan S M: Visualization, Methodology

Ye B: Model Design, Project Administration

Ethics, Consent to Participate, and Consent to Publish declarations

This study was approved by the Ethics Committee of State Grid Xinjiang Electric Power Co., Ltd. All methods were performed in accordance with the relevant guidelines and regulations, including the Declaration of Helsinki and the policies of this ethics committee. Informed consent was obtained from all participants in accordance with ethical standards. Participants were fully informed of their rights, and a sample consent form was available upon request. All participants were adults.

Clinical trial number

Not applicable.

References

[1] Murtaza, A. A., Saher, A., Zafar, M. H., Moosavi, S. K. R., Aftab, M. F., & Sanfilippo, F. (2024). Paradigm shift for predictive maintenance and condition monitoring from Industry 4.0 to Industry 5.0: A systematic review, challenges and case study. Results in Engineering, 24, 102935.

[2] Ledmaoui, Y., El Maghraoui, A., El Aroussi, M., & Saadane, R. (2025). Review of Recent Advances in Predictive Maintenance and Cybersecurity for Solar Plants. Sensors, 25(1), 206.

[3] Suresh, K. P., Senthilkumar, R., Saravanan, S., Suresh, M., & Jamuna, P. (2022). Challenges and Opportunities for Predictive Maintenance of Solar Plants. Photovoltaic Systems, 65–83.

[4] Geng, S., & Wang, X. (2022). Predictive maintenance scheduling for multiple power equipment based on data-driven fault prediction. Computers & Industrial Engineering, 164, 107898.

[5] Schwartz, J. A., Ricks, W., Kolemen, E., & Jenkins, J. D. (2024). Valuing maintenance strategies for fusion plants as part of a future electricity grid. arXiv preprint arXiv:2405.01514.

[6] Salehi, M., & Pattabiraman, K. (2024, December). AutoPatch: Automated Generation of Hotpatches for Real-Time Embedded Devices. In Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security (pp. 2370–2384).

[7] Xi, Z., Zhang, B., Bhattacharjya, A., Wang, Y., & He, C. (2025). Research on a Secure and Reliable Runtime Patching Method for Cyber–Physical Systems and Internet of Things Devices. Symmetry, 17(7), 983.

[8] Wang, M., Wu, X., Jiang, N., Liu, C., Xian, R., & Jia, W. (2024). Study on the Preparation and Aging Performance of Temperature-Indicating Patch Used for Thermal Defect Detection of Transformer Bushing Cylinder Head. Processes, 12(4), 736.

[9] Lickert, H., Lachnit, T., & Dietrich, F. (2023, December). Approach for Structured Repairability Assessment for Automated Repair Processes. In Global Conference on Sustainable Manufacturing (pp. 391–398). Cham: Springer Nature Switzerland.

[10] Al Ameeri, T. A., Ab Rahman, M. N., & Muhamad, N. (2023). Analysing effective and ineffective impacts of maintenance strategies on electric power plants: A comprehensive

approach. Energies, 16(17), 6243.

[11] Dissanayake, N., Zahedi, M., Jayatilaka, A., & Babar, M. A. (2022). Why, how and where of delays in software security patch management: An empirical investigation in the healthcare sector. Proceedings of the ACM on Human-computer Interaction, 6(CSCW2), 1-29.

[12] Wetzels, J., Dos Santos, D., & Ghafari, M. (2023). Insecure by design in the backbone of critical infrastructure. In Proceedings of Cyber-Physical Systems and Internet of Things Week 2023 (pp. 7-12).

[13] Dissanayake, N., Jayatilaka, A., Zahedi, M., & Babar, M. A. (2022). Software security patch management-A systematic literature review of challenges, approaches, tools and practices. Information and Software Technology, 144, 106771.

[14] Adinolfi, G., Ciavarella, R., Graditi, G., Ricca, A., & Valenti, M. (2024). Innovative Method for Reliability Assessment of Power Systems: From Components Modeling to Key Indicators Evaluation. Electronics, 13(2), 275.

[15] Jalilian, A., Taheri, B., & Molzahn, D. K. (2024). Co-Optimization of Damage Assessment and Restoration: A Resilience-Driven Dynamic Crew Allocation for Power Distribution Systems. IEEE Transactions on Power Systems, 40(1), 676–688.

[16] Ul Haq, S., Singh, Y., Sharma, A., Gupta, R., & Gupta, D. (2023). A survey on IoT & embedded device firmware security: architecture, extraction techniques, and vulnerability analysis frameworks. Discover Internet of Things, 3(1), 17.

[17] Urblik, L., Kajati, E., Papcun, P., & Zolotová, I. (2024). Containerization in edge intelligence: A review. Electronics, 13(7), 1335.

[18] Liu, C., Tan, R., Wu, Y., Feng, Y., Jin, Z., Zhang, F., ... & Liu, Q. (2024). Dissecting zero trust: Research landscape and its implementation in IoT. Cybersecurity, 7(1), 20.

[19] Molęda, M., Małysiak-Mrozek, B., Ding, W., Sunderam, V., & Mrozek, D. (2023). From corrective to predictive maintenance—A review of maintenance approaches for the power industry. Sensors, 23(13), 5970.

[20] Mohsin, A., Janicke, H., Ibrahim, A., Sarker, I. H., & Camtepe, S. (2025). A Unified Framework for Human AI Collaboration in Security Operations Centers with Trusted Autonomy. arXiv preprint arXiv:2505.23397.

[21] Huang, Z., Sheng, Z., Ma, C., & Chen, S. (2024). Human as AI mentor: Enhanced human-in-the-loop reinforcement learning for safe and efficient autonomous driving. Communications in Transportation Research, 4, 100127.

[22] Karunamurthy, A., Kiruthivasan, R., & Gauthamkrishna, S. (2023). Human-in-the-loop intelligence: Advancing AI-centric cybersecurity for the future. Quing: International Journal of Multidisciplinary Scientific Research and Development, 2(3), 20–43.

[23] Alahaideb, L., Al-Nafjan, A., Aljumah, H., & Aldayel, M. (2025). Brain–Computer Interface for EEG-Based Authentication: Advancements and Practical Implications. Sensors, 25(16), 4946.

[24] Seyfizadeh, A., Peach, R. L., Tovote, P., Isaias, I. U., Volkmann, J., & Muthuraman, M. (2024).

Enhancing security in brain–computer interface applications with deep learning: Electroencephalogram-based user identification. Expert Systems with Applications, 253, 124218.

[25] Brocal, F. (2023). Brain-computer interfaces in safety and security fields: risks and applications. Safety Science, 160, 106051.

[26] Yu, H., Fouxon, I., Wang, J., Li, X., Yuan, L., Mao, S., & Mond, M. (2023). Lyapunov exponents and Lagrangian chaos suppression in compressible homogeneous isotropic turbulence. Physics of Fluids, 35(12).

[27] Winter, L., Taylor, P., Bellenger, C., Grimshaw, P., & Crowther, R. G. (2023). The application of the Lyapunov Exponent to analyse human performance: A systematic review. Journal of Sports Sciences, 41(22), 1994–2013.

[28] Zhang, Y. C., & Zhang, Y. Y. (2022). Lyapunov exponent, mobility edges, and critical region in the generalized Aubry-André model with an unbounded quasiperiodic potential. Physical Review B, 105(17), 174206.

[29] Irani, H., & Metsis, V. (2024, May). Enhancing time-series prediction with temporal context modeling: A Bayesian and deep learning synergy. In The International FLAIRS Conference Proceedings (Vol. 37).

[30] Nikookar, S., Namazi Nia, S., Basu Roy, S., Amer-Yahia, S., & Omidvar-Tehrani, B. (2025). Model reusability in Reinforcement Learning. The VLDB Journal, 34(4), 41.

[31] Abolfathi, M., Inturi, S., Banaei-Kashani, F., & Jafarian, J. H. (2024). Toward enhancing web privacy on HTTPS traffic: A novel SuperLearner attack model and an efficient defense approach with adversarial examples. Computers & Security, 139, 103673.

[32] Vaziri, A., & Fang, H. (2025, July). Optimal inferential control of convolutional neural networks. In 2025 American Control Conference (ACC) (pp. 2603-2610). IEEE.

[33] Aqajari, S. A. H. (2024). Robust, Personalized, and Context-Aware Affect Monitoring in Daily-Life (Doctoral dissertation, University of California, Irvine).

[34] Choudhry, A., Gouda, S. K., Satpathy, S. P., Shukla, K. M., Dash, A. K., & Pasayat, A. K. (2025). Integration of EEG-based BCI Technology in IoT enabled Smart Home Environment: An in-depth comparative analysis on Human-Computer Interaction Techniques. Expert Systems with Applications, 128730.

[35] Qi, Y., Zhu, X., Xu, K., Ren, F., Jiang, H., Zhu, J., ... & Wang, Y. (2022). Dynamic ensemble Bayesian filter for robust control of a human brain-machine interface. IEEE Transactions on Biomedical Engineering, 69(12), 3825–3835.

[36] Ma, Q., Gao, W., Xiao, Q., Ding, L., Gao, T., Zhou, Y., ... & Cui, T. J. (2022). Directly wireless communication of human minds via non-invasive brain-computer-metasurface platform. elight, 2(1), 11.

[37] Wang, R., Li, C., Zhang, K., & Tu, B. (2025). Zero-trust based dynamic access control for cloud computing. Cybersecurity, 8(1), 12.

[38] Wu, X., Zheng, T., Wu, R., Ren, J., Guo, J., & Du, Y. (2025). Hi-SAM: A high-scalable

authentication model for satellite-ground Zero-Trust system using mean field game. Journal of Network and Systems Management, 33(3), 72.

[39] Yeoh, W., Liu, M., Shore, M., & Jiang, F. (2023). Zero trust cybersecurity: Critical success factors and a maturity assessment framework. Computers & Security, 133, 103412.

[40] Liu, Y., Lan, Y., & Xia, Y. (2025). The evolutionary mechanism of information negentropy spurred by emergencies. Scientific Reports, 15(1), 31886.

[41] Durán, O., Sáez, G., & Durán, P. (2023). Negentropy as a Measure to Evaluate the Resilience in Industrial Plants. Mathematics, 11(12), 2707.

[42] Liu, S., & Motani, M. (2025). Improving Mutual Information based Feature Selection by Boosting Unique Relevance. Journal of Artificial Intelligence Research, 82, 1267–1292.

[43] Gowri, G., Lun, X., Klein, A., & Yin, P. (2024). Approximating mutual information of high-dimensional variables using learned representations. Advances in Neural Information Processing Systems, 37, 132843–132875.

[44] Tuononen, M., & Hautamäki, V. (2025). Improving Numerical Stability of Normalized Mutual Information Estimator on High Dimensions. IEEE Signal Processing Letters.

[45] Roveda, L., Testa, A., Shahid, A. A., Braghin, F., & Piga, D. (2022). Q-Learning-based model predictive variable impedance control for physical human-robot collaboration. Artificial Intelligence, 312, 103771.

[46] Islam, M. S., Das, S., Gottipati, S. K., Duguay, W., Mars, C., Arabneydi, J., ... & Taylor, M. E. (2025). Human-AI collaboration in real-world complex environment with reinforcement learning. Neural Computing and Applications, 1–31.

[47] Korivand, S., Galvani, G., Ajoudani, A., Gong, J., & Jalili, N. (2024). Optimizing Human–Robot Teaming Performance through Q-Learning-Based Task Load Adjustment and Physiological Data Analysis. Sensors, 24(9), 2817.

[48] Ma, H., Vo, T. V., & Leong, T. Y. (2023, May). Hierarchical Reinforcement Learning with Human-AI Collaborative Sub-Goals Optimization. In Proceedings of the 2023 international conference on autonomous agents and multiagent systems (pp. 2310–2312).

[49] Gu, S., Kshirsagar, A., Du, Y., Chen, G., Peters, J., & Knoll, A. (2023). A human-centered safe robot reinforcement learning framework with interactive behaviors. Frontiers in Neurorobotics, 17, 1280341.

[50] Huang, K., Payer, M, Qian, Z., Sampson, J., Tan, G., & Jaeger, T. (2025, May). SoK: Challenges and Paths Toward Memory Safety for eBPF. In 2025 IEEE Symposium on Security and Privacy (SP) (pp. 848–866). IEEE.

[51] Craun, M., Oswald, A., & Williams, D. (2023, September). Enabling eBPF on Embedded Systems Through Decoupled Verification. In Proceedings of the 1st Workshop on eBPF and Kernel Extensions (pp. 63–69).

[52] Wang, Z., Chen, T., Dai, Q., Chen, Y., Wei, H., & Zeng, Q. (2024). When eBPF Meets Machine Learning: On-the-fly OS Kernel Compartmentalization. arXiv preprint arXiv:2401.05641.