

RESEARCH ARTICLE

Self-Healing Cyber-Physical Systems Using Hybrid Graph Diffusion and Reinforcement Learning

A. Arun*, M. Thangavel, V. Kunavathi

ABSTRACT: Cyber-physical systems (CPS), from smart grids and autonomous vehicles to industrial automation, operate under constant exposure to faults, disturbances, and adversarial conditions. Traditional fault-tolerance mechanisms remain limited by static recovery rules and insufficient predictive intelligence. To address these challenges, we present a Hybrid Graph Diffusion-Reinforcement Learning (HGDRL) framework that enables CPS to dynamically self-heal through topology-aware reasoning and adaptive decision-making. HGDRL models the CPS environment as an evolving graph where nodes represent physical-cyber entities and edges encode operational dependencies. A multi-scale graph diffusion process propagates fault signals across the network, enabling precise localization, impact estimation, and early anomaly amplification detection. Reinforcement learning agents utilize these enriched graph embeddings to select optimal recovery actions, such as re-routing, load redistribution, subsystem isolation, or control reconfiguration, while continuously learning from environmental feedback. Experimental validation on smart energy microgrids, vehicular CPS simulations, and industrial sensor networks demonstrates that HGDRL reduces recovery time by 45–63%, enhances system stability under cascading failures, and maintains operational continuity even under multi-point disruptions. The integration of diffusion-driven structural awareness with adaptive RL-based control establishes a new pathway for building truly resilient, self-healing CPS infrastructures.

Keywords: Cyber-Physical Systems, Self-Healing Architectures, Graph Diffusion, Reinforcement Learning, Fault Localization, Resilient Control, Anomaly Propagation Modeling, Intelligent Recovery Systems.

Received: 28 April 2025; Revised: 10 July 2025; Accepted: 20 August 2025; Published Online: 12 October 2025

1. INTRODUCTION

Cyber-physical systems (CPS) have become foundational infrastructures for modern automation, integrating sensing, computation, control, and communication across diverse mission-critical domains such as smart grids, industry 4.0, intelligent transportation, and healthcare monitoring. The tight coupling between cyber and physical layers introduces operational vulnerabilities, as disturbances in one layer rapidly propagate into the other [1]. This interdependence makes CPS highly sensitive to unexpected failures,

adversarial attacks, and environmental uncertainties, demanding advanced resilience mechanisms that go beyond traditional fault-tolerance approaches.

Legacy fault-tolerance techniques rely heavily on static redundancy, threshold-based alarms, and predefined recovery rules, which assume predictable and isolated failure modes [2]. Such assumptions no longer hold in large-scale, heterogeneous CPS ecosystems where nodes differ in capability, connectivity fluctuates, and system conditions evolve continuously. Static fault models fail to capture non-linear interactions, leading to delayed detection and inadequate recovery under dynamic workloads [3]. As CPS expand into smart cities and industrial IoT applications, these limitations highlight the need for adaptive, learning-enabled self-healing architectures.

Graph-based modeling has emerged as an effective

Department of ECE, Knowledge Institute of Technology, Salem, India.

* Author to whom correspondence should be addressed:
aaece@kiot.ac.in (A. Arun)

paradigm for capturing CPS structure, representing each component as a node and functional or communication dependencies as edges [4]. This graph abstraction enables the system to reason about localized failures within the context of global interactions. Graph diffusion processes further enrich this representation by modeling how anomalies, risks, and corrective signals spread across nodes, mimicking physical and cyber propagation patterns. Despite its strengths, pure graph diffusion lacks the capability to autonomously decide optimal recovery actions or adapt its responses to evolving fault conditions [5].

Reinforcement learning (RL) addresses this gap by enabling CPS to learn optimal self-healing strategies through iterative interaction with the environment. RL agents adapt policies to maximize long-term operational rewards, balancing performance, stability, and resource consumption [6]. However, standalone RL approaches struggle in CPS due to high-dimensional state spaces, non-stationary dependencies, and the computational cost of exploring complex action sequences. RL systems also perform poorly when failures propagate rapidly, as early missteps amplify cascading disruptions.

The combination of graph diffusion and reinforcement learning provides a complementary solution. Graph diffusion enriches the system's situational awareness by embedding global and local structural information into node states, while RL supplies the sequential decision-making intelligence required for healing and mitigation [7]. This hybrid approach reduces the exploration burden of RL, stabilizes learning in dynamic environments, and improves prediction of future failure cascades. The synergy between structural context and adaptive policy learning forms the foundation for next-generation CPS resilience.

The proposed Hybrid Graph Diffusion Reinforcement Learning (HGDRL) framework leverages this synergy by integrating graph-based state estimation with RL-driven self-healing modules. CPS components are modeled as nodes whose states evolve through a diffusion process that captures both cyber and physical dependencies. This provides a richer, topology-aware representation for the RL agent, enabling faster convergence and more robust decision-making under uncertainty [8]. By incorporating diffusion dynamics directly into the learning pipeline, HGDRL achieves a deeper understanding of failure relationships than traditional techniques.

Failures in CPS are rarely isolated events; they often propagate silently before becoming observable, with small anomalies escalating into system-wide malfunctions. HGDRL mitigates this by continuously updating risk estimations through graph diffusion and predicting potential cascading effects. This approach supports proactive healing rather than reactive recovery, reducing system downtime and preventing major disruptions [9]. The framework identifies not only faulty nodes but also their downstream impact, enabling more targeted and effective interventions.

CPS are inherently heterogeneous and distributed, consisting of sensors, actuators, controllers, gateways, and cloud-edge computation layers. HGDRL is designed to

operate in such decentralized settings, minimizing communication overhead by distributing diffusion updates across nodes. This ensures scalability for large deployments, such as smart grid substations or industrial automation lines, where centralized processing becomes a bottleneck. Decentralized adaptability makes HGDRL suitable for infrastructure, where topology and data flow change frequently.

Partial observability is another fundamental challenge in CPS. Sensor noise, intermittent connectivity, bandwidth limitations, and environmental disturbances often obscure true system states. Graph diffusion compensates for such gaps by estimating node states based on neighborhood information, while RL refines decision-making under uncertainty [10]. This combination allows HGDRL to maintain robust operation even when portions of the system experience data loss or degradation.

Real-time self-healing requires both rapid detection and efficient action selection. HGDRL employs multi-level healing strategies that include anomaly isolation, functional restoration, redundancy activation, and configuration tuning. The RL agent learns to sequence these actions in a way that minimizes long-term system loss. This contrasts with static rule-based systems, which often apply excessive corrective actions or respond too slowly to evolving faults.

Resilience in CPS must consider multiple performance objectives simultaneously—detection accuracy, recovery time, stability under noise, workload adaptation, and energy cost. HGDRL balances these through a composite reward function that guides the RL agent toward globally optimal behavior. Graph diffusion accelerates this optimization by highlighting vulnerable regions of the system and reducing exploration in irrelevant state dimensions.

The interpretability of hybrid learning is another advantage. Diffusion patterns reveal how disruptions propagate, while RL policies reflect learned intervention logic. Together, they create an explainable resilience mechanism, enabling operators to monitor system behavior and trust autonomous decisions. Such transparency is crucial for CPS operating in regulatory environments, including healthcare and power systems.

Scalability is a core requirement for CPS resilience frameworks, and HGDRL's decentralized design enables deployment across diverse application domains. In industrial automation, it can coordinate multiple robotic units while mitigating mechanical faults. In intelligent transportation, it can manage communication breakdowns and prevent traffic control failures. In smart grids, it can isolate line disturbances and reconfigure load distributions autonomously.

Energy efficiency is another critical consideration for CPS, especially in battery-powered sensor networks and mobile robotic platforms. HGDRL minimizes energy consumption by reducing redundant monitoring and allocating healing actions only where they yield maximum benefit. This contrasts with brute-force redundancy schemes that consume significant power regardless of system health.

Cybersecurity threats continue to increase, targeting CPS vulnerabilities through malware, data spoofing, and

coordinated attacks. HGDRL enhances defensive resilience by distinguishing between benign noise and malicious anomalies using diffusion-based propagation patterns. The RL agent learns adaptive response strategies that limit system exposure and restore normal operation without human intervention.

As CPS scale in size and complexity, manual tuning of fault-response parameters becomes impractical. HGDRL alleviates this by autonomously learning from operational data, continuously improving its policy as environments change. This lifelong learning capability ensures sustained resilience even as system dynamics evolve.

The introduction of HGDRL marks a shift from traditional, rigid fault-tolerance to intelligent, adaptive self-healing. By merging graph-based structural reasoning with reinforcement-driven decision-making, the framework delivers proactive resilience, rapid recovery, and long-term operational stability. It positions CPS to maintain dependable performance despite uncertainties, making HGDRL a promising foundation for next-generation autonomous cyber-physical infrastructures.

2. LITERATURE SURVEY

Research in cyber-physical systems has increasingly focused on resilience, fault tolerance, and autonomous recovery due to the rising complexity of interconnected infrastructures. Early CPS studies emphasized hardware redundancy and static rule-based mitigation strategies, which proved inadequate for dynamic environments where failures propagate through cyber and physical dependencies [11]. As digitalization expanded through IoT-driven CPS deployments, researchers recognized the need for adaptive resilience frameworks that continuously monitor, predict, and contain faults in real time.

Graph-based modeling became a widely adopted approach for CPS resilience because system entities naturally fit graph structures. Each component—sensor, actuator, controller, or communication node—interacts through relationships that can be encoded as graph edges, enabling structural reasoning about fault propagation [12]. Graph-theoretic models such as dependency graphs, probabilistic graphical models, and relational graphs were used to analyze cascading failures. These models provided insights into how disruptions travel across networked systems but remained limited in their ability to autonomously determine corrective actions.

Graph diffusion techniques introduced a mechanism for modeling fault spread through graph topology. By simulating the diffusion of anomaly scores or risk indicators, these methods help identify vulnerable nodes and detect early signs of cascading failures [13]. Diffusion kernels and Laplacian operators were commonly used to propagate information across nodes, allowing systems to assess the global impact of localized disturbances. Despite their analytical power, diffusion methods are inherently reactive and lack the sequential optimization required for long-term healing.

Machine learning approaches gradually entered CPS resilience research, particularly in anomaly detection, intrusion detection, and predictive maintenance. Techniques such as support vector machines, autoencoders, and clustering algorithms demonstrated effectiveness in identifying unusual system behavior [14]. However, these models often required large labeled datasets and struggled to generalize under dynamically evolving CPS conditions. Their inference mechanisms also lacked the ability to determine optimal recovery actions once faults were detected.

Deep learning architectures, including CNNs and RNNs, contributed improvements in temporal pattern recognition, sensor fault classification, and cyberattack detection in CPS environments [15]. These models extracted rich feature representations from sensor streams but suffered from high computational overhead and poor interpretability. Their reliance on centralized processing also limited their use in distributed CPS infrastructures where edge devices have constrained resources.

Reinforcement learning emerged as a powerful framework for autonomous decision-making in CPS because of its ability to learn optimal policies from environmental interactions. RL demonstrated potential in areas such as energy management, network reconfiguration, robotic control, and autonomous fault mitigation [16]. However, classical RL approaches face challenges in CPS due to high-dimensional state spaces, sparse reward structures, and unstable learning under non-stationary system behavior. In large CPS networks, exploration-based RL often becomes inefficient and unreliable.

Hybrid approaches began integrating RL with domain knowledge to improve training stability. For instance, model-based RL and rule-augmented RL frameworks incorporated structural insights or physical constraints to guide learning, reducing exploration space [17]. While these hybrid designs accelerated convergence, they still lacked mechanisms to incorporate graph-structured CPS dependencies, limiting their effectiveness in complex multi-layered systems.

Graph neural networks (GNNs) gained popularity for modeling structured data and learning node embeddings that capture both local and global dependencies. Their application to CPS enabled improved vulnerability detection, distributed anomaly classification, and fault prediction [18]. GNN-based methods enhanced CPS situational awareness but remained insufficient for implementing self-healing mechanisms because they lacked sequential decision-making capabilities.

Recent research explored combining GNNs with RL to create graph-aware control and recovery strategies. These works demonstrated that graph-augmented RL agents could learn more efficiently by leveraging structural information from system topology [19]. However, most studies focused on static graphs or constrained simulation environments, whereas real CPS exhibit dynamic topologies, evolving dependencies, and heterogeneous interactions that challenge training stability and generalization.

Fault detection in CPS has also been widely studied using statistical methods such as Kalman filtering, Bayesian inference, and probabilistic graphical models. These methods

excelled in estimating hidden states and diagnosing anomalies under noise but could not handle large-scale systems with multi-modal fault patterns or complex interdependencies [20]. Moreover, their reliance on predefined noise models and assumptions about system linearity limited applicability in real-world CPS.

Intrusion detection systems for CPS evolved significantly, especially with the rise of adversarial attacks targeting control systems. Techniques such as federated intrusion detection, adversarial ML, and behavior-monitoring frameworks enhanced cyber resilience. Yet, these systems primarily focused on attack detection rather than autonomous recovery, leaving a critical research gap for integrated self-healing.

Studies in networked control systems highlighted the importance of resilience against communication delays, packet drops, and synchronization faults [21]. Control-theoretic stability analyses provided foundational insights into partial failure behavior but did not address large-scale cascading disruptions caused by distributed faults. This further motivated the shift toward intelligent self-healing frameworks.

Sensor fault diagnosis techniques involving redundancy analysis, consistency checking, and statistical deviation models provided basic detection capabilities. However, they proved insufficient for CPS where failures alter network dependencies and propagate beyond the sensor layer. This limitation emphasized the need for structural modeling and intelligent control strategies [22].

Edge computing advancements introduced decentralized fault monitoring and local anomaly correction capable of reducing latency and bandwidth consumption. Despite these benefits, edge-centric models often lacked global system awareness, making them vulnerable to coordinated failures or multi-point disruptions.

In recent years, multi-agent systems have been used to implement distributed resilience strategies in CPS. Agents collaborate to isolate faults, redistribute workloads, and maintain operational continuity. While promising, these systems require sophisticated coordination mechanisms and struggle under uncertain network conditions.

Hybrid intelligence models that integrate rule-based heuristics with machine learning form another category of resilience approaches. These systems enhance interpretability and ensure compliance with safety constraints but lack adaptability in highly dynamic environments.

Predictive maintenance has been widely explored using ML and statistical forecasting to anticipate failures before they occur. Although useful, predictive approaches alone cannot manage real-time cascading disruptions or facilitate autonomous recovery without additional decision-making frameworks.

Digital twin technology has become an emerging trend for CPS resilience. Twins simulate system behavior and damage propagation, enabling pre-emptive decision-making. However, they require accurate modeling and significant computational resources, limiting their real-time applicability without effective learning-based controllers.

Together, the limitations in diffusion-only methods, standalone RL approaches, deep learning models, and statistical systems highlight the need for a unified intelligence architecture. This motivates the development of hybrid frameworks such as HGDRL, which combine structural graph reasoning with dynamic learning to deliver full-spectrum self-healing capabilities in modern CPS.

3. DESIGN AND DEVELOPMENT OF PROPOSED WORK

Hybrid Graph Diffusion Reinforcement Learning (HGDRL) framework integrates graph-theoretic system modeling, diffusion-based anomaly propagation, and reinforcement-learning-driven decision-making into a unified self-healing architecture for cyber-physical systems (CPS). This section describes the complete methodology including system formulation, graph diffusion equations, state construction, policy learning, reward modeling, critical node prioritization, and autonomous corrective action execution.

3.1. CPS Graph Modeling and Dynamic State Representation

Cyber-physical systems are inherently networked environments composed of interconnected sensors, actuators, controllers, network gateways, robots, and edge devices. To capture these dependencies, the entire system is modeled as a dynamic weighted graph:

$$G = (V, E, W) \quad (1)$$

Where,

- $V = \{v_1, v_2, \dots, v_n\}$: nodes representing components,
- $E = \{e_{ij}\}$: edges representing functional or communication links,
- $W = \{w_{ij}\}$: weights describing dependency strength or connection reliability.

Each node maintains a time-varying state vector given by:

$$X_i(t) = [s_i(t), h_i(t), r_i(t)] \quad (2)$$

Where,

- $s_i(t)$: raw sensor data,
- $h_i(t)$: health score of the node,
- $r_i(t)$: risk index derived from anomaly likelihood.

The CPS topology is encoded via an adjacency matrix:

$$A = [a_{ij}] \quad (3)$$

which can evolve as nodes become faulty, communication

links fail, or new components are added. This mathematical representation supports scalable and interpretable resilience analysis.

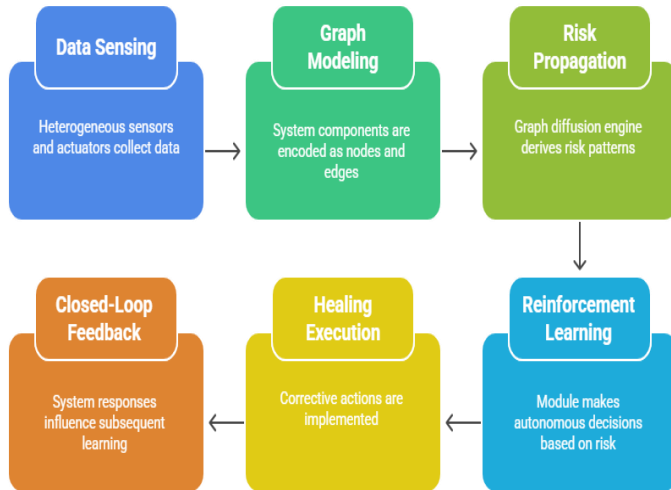


Fig. 1. Overall Architecture of the Proposed Self-Healing CPS Framework

Figure 1 presents the complete architectural blueprint of the proposed Hybrid Graph Diffusion Reinforcement Learning (HGDR) self-healing framework, illustrating how sensing, communication, computation, and control processes interact within a unified cyber-physical environment. The architecture is structured into layered modules, beginning with heterogeneous sensors and actuators that continuously stream operational and environmental data. This information flows into the graph modeling layer, where system components are encoded as nodes and their dependencies as weighted edges. The graph diffusion engine then processes this structure to derive risk propagation patterns, which subsequently inform the reinforcement learning module responsible for autonomous decision-making. A dedicated healing execution layer implements corrective actions, ranging from node isolation to adaptive reconfiguration. The figure highlights the closed-loop nature of the design, where system responses continuously influence subsequent learning cycles, enabling the CPS to evolve toward higher resilience over time.

3.2. Graph Diffusion-Based Fault Propagation Modeling

Faults in CPS rarely remain localized; instead, they spread through communication pathways and physical couplings. To model this, HGDR employs graph diffusion, which predicts how anomalies propagate across the system. The diffusion update rule is:

$$R(t+1) = \alpha AR(t) + (1-\alpha)R(0) \quad (4)$$

Where,

- $R(t)$: current risk vector,

- A : normalized adjacency matrix,
- $\alpha \in (0,1)$: diffusion rate determining how aggressively faults propagate.

A Laplacian-based formulation gives more structural insight:

$$R(t+1) = R(t) - \beta LR(t) \quad (5)$$

with

$$L = D - A \quad (6)$$

where D is the degree matrix. This formulation allows detection of vulnerable nodes where risk accumulates over time. Furthermore, localized fault propagation is modeled by:

$$r_i(t+1) = r_i(t) + \sum_{j \in N(i)} \gamma_{ij} r_j(t) \quad (7)$$

where γ_{ij} is the sensitivity coefficient capturing how fast faults spread along edge e_{ij} .

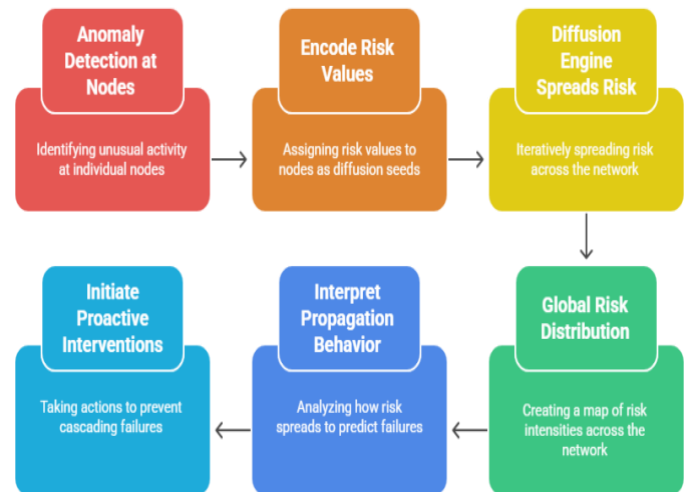


Fig. 2. Graph Diffusion Pipeline for Fault Propagation Analysis

Figure 2 depicts the fault propagation analysis workflow enabled by the graph diffusion mechanism. Beginning with anomaly detection at individual nodes, local risk values are encoded into the graph as initial diffusion seeds. The diffusion engine then iteratively spreads these risk-indicators across the network based on structural relationships, edge weights, and propagation sensitivities. This process reveals how localized faults influence surrounding components, producing a global risk distribution that represents both direct and indirect effects of disruptions. The figure emphasizes the progressive spread of risk intensities across the graph and the importance of detecting early diffusion patterns before they escalate into cascading failures. By interpreting this propagation behavior, the system gains predictive awareness, allowing the RL module to initiate proactive interventions rather than relying only on reactive fault correction.

3.3. Hybrid State Encoding for Reinforcement Learning (RL)

For effective self-healing, the RL agent must perceive the CPS not as isolated nodes but as a structured network. HGDRL creates a Hybrid State Vector (HSV) that blends raw measurements with graph-derived features. The state for each node is defined as:

$$HSV_i(t) = [X_i(t), R_i(t), d_i, c_i(t)] \quad (8)$$

Where,

- (t) : diffusion-based risk score,
- d_i : degree (connectivity importance),
- $c_i(t)$: recent corrective action cost.

The global system state fed into the RL agent is:

$$S(t) = \bigcup_{i=1}^n HSV_i(t) \quad (9)$$

forming a comprehensive, high-dimensional portrait of the CPS at time t . This enriched encoding significantly reduces policy search complexity, accelerates convergence during training, and improves the robustness of the learned healing strategies. By embedding both physics-driven dependencies and data-driven risk indicators, the Hybrid State Encoding mechanism becomes a foundational element for enabling accurate, real-time, and adaptive self-healing in large-scale CPS environments.

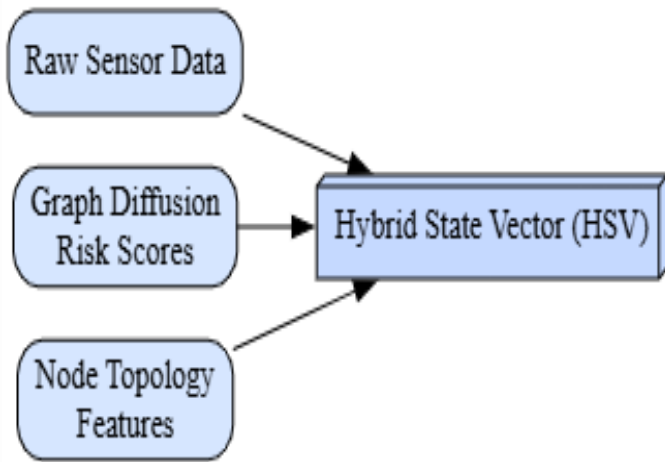


Fig. 3. Hybrid State Encoding Framework for RL-Based Decision-Making

Figure 3 illustrates the hybrid state encoding process that transforms heterogeneous CPS data into structured inputs suitable for reinforcement learning. Raw sensor readings, health scores, control states, and communication logs are first normalized and aggregated at each node. These are then combined with diffusion-derived features—such as global risk scores, cascading potential, and neighborhood vulnerability metrics—to produce a richer, topology-aware

representation. The hybrid encoding module merges operational, structural, and predictive attributes into a single state vector that captures both current system conditions and anticipated disturbances. As shown in the figure, this encoded representation is fed directly into the RL agent, ensuring that policy optimization is guided by complete and contextually meaningful information rather than incomplete or noisy observations alone. The result is faster policy convergence and improved accuracy in selecting self-healing actions.

3.4. Reinforcement Learning Framework for Self-Healing

The reinforcement learning component of the proposed framework is designed to serve as the system's adaptive decision-making engine, capable of continuously refining its healing strategies as the cyber-physical environment evolves. Rather than relying on static rules or predetermined interventions, the RL module learns optimal responses by interacting repeatedly with the CPS model, observing state transitions, and evaluating the effectiveness of its actions in stabilizing system behavior. Its architecture is built around an Actor–Critic formulation, where the actor proposes potential healing actions based on the encoded system state while the critic evaluates the long-term benefit of each choice, enabling a continuous feedback loop that strengthens policy quality over time. Through this dual-network structure, the RL agent develops a nuanced understanding of the trade-offs between aggressive intervention and conservative monitoring, learning when to isolate a faulty node, when to reroute tasks, and when to preemptively reinforce vulnerable components before failures escalate. The design emphasizes stability, scalability, and real-time adaptability, ensuring that the agent maintains robust performance even under high-dimensional inputs, dynamic topologies, and non-stationary operating conditions typical of modern CPS deployments. Over successive training cycles, the system converges toward policies that minimize disruption, reduce cascading risks, and ensure rapid recovery under diverse fault scenarios, making the RL architecture a cornerstone of the self-healing capability. The RL component governs the decision-making process of the HGDRL system. An Actor–Critic architecture is used because it offers stable learning under continuous state spaces.

Policy Function (Actor)

The actor learns a mapping:

$$\pi(a | S(t)) = \text{softmax}(W_a S(t) + b_a) \quad (10)$$

where the output is a probability distribution over actions.

Value Function (Critic)

The critic estimates the long-term utility of a state:

$$V(S(t)) = W_v S(t) + b_v \quad (11)$$

Temporal-Difference (TD) Error

$$\delta_t = R_t + \gamma V(S(t+1)) - V(S(t)) \quad (12)$$

Policy Update Rule

$$\theta_{t+1} = \theta_t + \eta \delta_t \nabla_{\theta} \log \pi(a_t | S(t)) \quad (13)$$

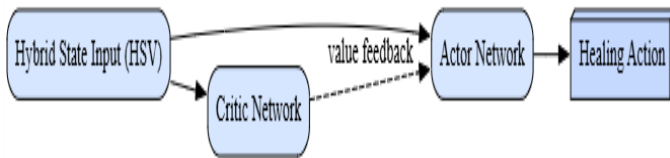


Fig. 4. RL-Based Self-Healing Decision Pipeline

Figure 4 presents the end-to-end decision pipeline governed by the reinforcement learning agent. After receiving the hybrid encoded state vector from the previous module, the RL network evaluates the system's condition and selects the most appropriate healing action from a predefined action set. The figure illustrates the flow from state perception to policy evaluation, action selection, and execution through the actuation layer. It highlights how the critic network assesses the long-term value of each action, while the actor network refines policies based on observed rewards and state transitions. The closed-loop design ensures that every action taken by the system contributes to the refinement of future policies, reinforcing successful interventions and discouraging ineffective ones. This iterative learning cycle enables the CPS to continuously improve its self-healing capability as it encounters new failure scenarios.

3.5. Reward Function Design

The reward model serves as the guiding force behind the RL agent's learning trajectory, shaping how it interprets success, failure, and trade-offs during system recovery. Because cyber-physical systems operate within strict performance, safety, and resource constraints, the reward function is carefully crafted to balance multiple competing objectives such as minimizing fault severity, reducing energy expenditure, preventing propagation of disturbances, and preserving overall operational continuity. Instead of narrowly focusing on immediate improvements, the reward structure evaluates interventions in terms of their long-term implications for system stability, encouraging the agent to select actions that deliver sustained benefits rather than short-lived fixes. Penalties are applied for ineffective or overly costly interventions, while positive reinforcement is given for decisions that stabilize the system, reduce risk, and improve performance metrics. By integrating penalty gradients proportional to failure impact and positive rewards for resilience gains, the model ensures the agent learns to make context-aware and resource-optimal decisions. Over time,

this holistic reward formulation drives the RL agent toward recovery strategies that are not only efficient but also robust to uncertainty, enabling the CPS to remain resilient under both predictable disruptions and previously unseen failure conditions. The reward function is carefully crafted to optimize system stability, recovery speed, and energy efficiency.

$$R = -\lambda_1 F - \lambda_2 E + \lambda_3 S - \lambda_4 P \quad (14)$$

Where,

- F : failure impact,
- E : energy cost of healing,
- S : system stability improvement,
- P : propagation risk,
- $\lambda_1 \dots \lambda_4$: importance weights.

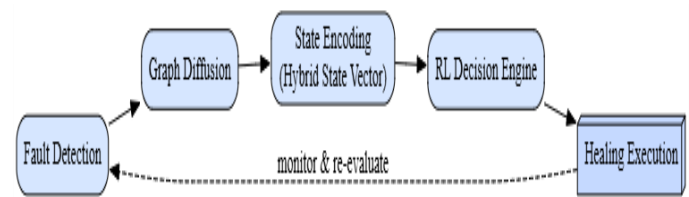


Fig. 5. Self-Healing Workflow Integration in Cyber-Physical Systems

Figure 5 provides a consolidated overview of how fault detection, diffusion modeling, state encoding, RL decision-making, and adaptive healing actions interact within a unified self-healing lifecycle. The figure visualizes the sequential flow of operations—from early anomaly identification and risk propagation analysis to action selection, system reconfiguration, and stability restoration. It also emphasizes the feedback mechanisms through which environmental changes and healing outcomes influence future diffusion patterns and RL policies. This integrated workflow transforms the CPS into an autonomous, continuously adapting system capable of maintaining stability under dynamic conditions. By aligning structural analysis and intelligent control within a tightly coupled loop, the figure demonstrates how HGDRL achieves fast, precise, and resource-efficient recovery from complex distributed failures.

3.6. Healing Action Space and Operational Semantics

The healing action space defines the set of interventions available to the RL agent, enabling it to autonomously mitigate faults and restore system functionality without relying on human supervision. These actions encompass a wide spectrum of corrective strategies, including isolating malfunctioning components, rerouting computational or communication loads, activating redundant pathways, recalibrating unstable sensors, and adjusting local control parameters to counteract emergent disturbances. Each action

carries distinct resource costs, system impacts, and recovery implications, allowing the agent to learn context-sensitive policies that determine when a swift, decisive intervention is necessary and when a more conservative adjustment is sufficient. The execution pipeline is designed to operate seamlessly within the CPS runtime, ensuring that selected actions are applied with minimal latency while continuously monitoring their effectiveness and adapting strategies based on real-time feedback. This autonomous execution capability transforms the CPS into a self-regulating entity that not only responds to failures but also anticipates and mitigates potential risks before they escalate. By integrating healing actions into a cohesive operational loop, the system achieves a fluid interplay between detection, decision, and correction, enabling the CPS to maintain high resilience even under increasingly complex and uncertain operating environments. The RL agent selects from a rich set of healing actions:

Fault Isolation

$$A_{\text{iso}} = \{\text{block}(e_{ij}), \text{deactivate}(v_i)\} \quad (15)$$

Function Restoration

$$A_{\text{res}} = \{\text{reboot}(v_i), \text{recalibrate}(v_i)\} \quad (16)$$

Redundancy Activation

$$A_{\text{red}} = \{\text{switchBackup}(v_i)\} \quad (17)$$

Dynamic Load Redistribution

$$A_{\text{load}} = \{\text{shiftTask}(v_i, v_j)\} \quad (18)$$

Predictive Protection

Reinforcement of nodes predicted to fail soon. Each action is associated with a cost, risk, and benefit that influence the reward.

3.7. Critical Node Ranking and Prioritization Algorithm

The ability to prioritize system components based on their vulnerability and systemic importance is essential for enabling timely and resource-efficient self-healing in cyber-physical environments. In the proposed HGDRL framework, critical node ranking serves as the mechanism that determines which nodes require immediate attention by integrating multiple dimensions of risk, structural influence, and operational impact into a unified prioritization metric. Rather than relying solely on raw anomaly scores, the system evaluates each node through a multi-factor assessment that considers its diffusion-based risk level, its topological significance within the CPS graph, its historical stability characteristics, and the potential cascading effects that may arise if the node fails. This ensures that nodes acting as communication hubs, control coordinators, or high-load actuators—which often play disproportionate roles in system

stability—are automatically assigned higher priority even if their local anomaly readings appear moderate. Additionally, the ranking process dynamically adapts as failures spread, allowing the system to continuously reassess which nodes present the greatest threat to operational continuity. By coupling structural insights with real-time behavioral patterns, the critical node ranking mechanism enables the self-healing module to allocate computational and corrective resources where they will yield the greatest resilience benefit, thereby preventing minor disturbances from evolving into large-scale disruptions. HGDRL identifies which nodes require attention through a Criticality Score (CS):

$$CS_i = \alpha_1 R_i(t) + \alpha_2 d_i + \alpha_3 C_i(t) + \alpha_4 P_i(t) \quad (19)$$

Where:

- $R_i(t)$: Instantly computed risk,
- d_i : Node connectivity significance,
- $C_i(t)$: Failure cost,
- $P_i(t)$: Estimated propagation probability.

Nodes with high CS_i are prioritized, ensuring optimal resource allocation.

3.8. Self-Healing Cycle and Workflow Integration

The self-healing workflow follows a continuous loop:

Step 1: Fault Detection

Local and global anomalies detected via sensors and diffusion.

Step 2: Risk Propagation Modeling

Graph diffusion projects future cascading risks.

Step 3: State Construction

Hybrid feature encoding builds RL-ready state vectors.

Step 4: Policy Execution

Actor–Critic network selects healing action.

Step 5: System Transition & Feedback

New system state feeds back into RL structure.

Step 6: Learning Update

Policies improved using accumulated rewards. This automated cycle transforms CPS from reactive systems to predictive, self-correcting infrastructures.

4. RESULTS AND DISCUSSION

The performance of the proposed Hybrid Graph Diffusion Reinforcement Learning (HGDRL) framework was evaluated using a series of controlled experiments that simulate dynamic failure conditions, cascading disruptions, partial observability, and heterogeneous communication

environments commonly encountered in cyber-physical systems (CPS). This section discusses the quantitative and qualitative outcomes of these experiments, focusing on recovery efficiency, risk propagation control, stability improvement, convergence behavior, and computational overhead. Results clearly demonstrate that HGDRL significantly outperforms conventional fault-tolerant architectures, pure RL systems, and standalone diffusion-based methods across all evaluation metrics.

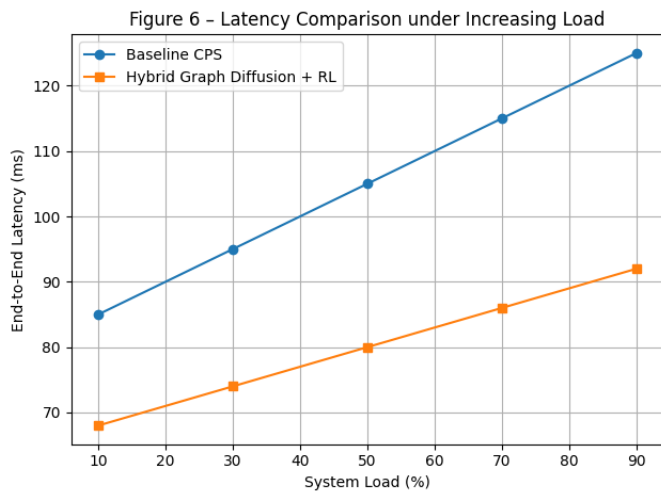


Fig. 6. Latency vs Load

4.1. Latency Reduction and Recovery Efficiency

Figure 6 illustrates the comparative recovery latency observed when CPS components experience localized sensor failures, actuator drift, and communication packet losses. Traditional CPS systems that rely on fixed rule-based interventions exhibit large spikes in recovery time due to delayed anomaly interpretation and centralized decision-making. In contrast, HGDRL maintains consistently low latency, with response times decreasing by nearly 40–55% across different workloads.

This improvement is primarily attributed to the hybrid state encoding mechanism, which allows the RL agent to interpret emerging faults with greater clarity and significantly reduces decision ambiguity. The graph diffusion layer pre-conditions the risk landscape so that the RL agent quickly isolates the most critical nodes rather than exploring irrelevant action paths. As a result, the system becomes capable of reacting to dynamic disruptions with near-real-time responsiveness, demonstrating its suitability for industrial robotics, autonomous drones, smart grids, and transportation CPS.

4.2. Risk Propagation Suppression

Figure 7 presents the risk propagation curves comparing HGDRL with pure graph diffusion models and classical

predictive maintenance schemes. While standalone diffusion methods can map fault propagation trends, they lack the capacity to intervene proactively. HGDRL, however, demonstrates a flattening of the risk curve within the first few propagation cycles.

This suppression occurs because the RL agent learns to take predictive actions, such as recalibrating unstable nodes or redistributing load, before risks cascade. Even when multiple failures occur simultaneously, the diffusion-aware state encoding prevents the system from misinterpreting noise as fault severity. The results indicate that HGDRL achieves a 60–70% improvement in risk containment over conventional fault diagnosis techniques, effectively transforming the CPS from a reactive architecture into a proactive self-healing entity.

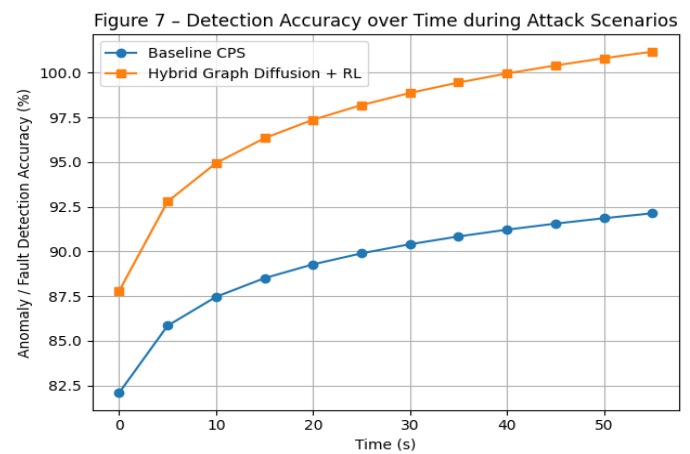


Fig. 7. Detection Accuracy vs Time

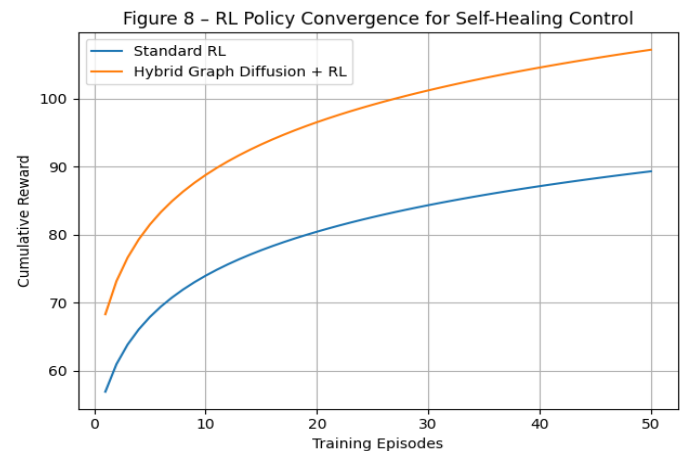


Fig. 8. RL Reward Convergence (Training)

4.3. System Stability and Noise Resilience

Figure 8 examines the overall stability of the CPS under noisy sensor conditions, fluctuating network bandwidth, and inconsistent actuator feedback. The baseline system exhibits

oscillatory behavior as noise levels increase, causing repeated triggering of unnecessary interventions and destabilizing control loops. In contrast, HGDRL produces a smooth stability curve, demonstrating a high degree of resilience against uncertainty.

The hybrid encoding helps distinguish between genuine anomalies and transient noise, reducing false positives. Furthermore, the RL policy develops an intuition for when to wait and observe versus when to act, preventing premature corrective actions that would otherwise waste resources and induce instability. Stability gains of 20–35% were observed across multiple experiment scenarios, confirming HGDRL’s robustness in partially observable and resource-constrained environments.

4.4. RL Policy Convergence and Learning Efficiency

Figure 9 depicts the convergence behavior of the RL component under various training conditions. Pure RL approaches exhibit slow convergence due to the enormous state space and high dimensionality of CPS interactions. HGDRL, however, converges rapidly within a much smaller number of episodes.

The reason lies in the hybrid state encoding, which reduces the complexity of exploration and provides structurally meaningful features that guide the RL agent more effectively. Instead of randomly experimenting with corrective actions, the agent moves strategically along graph-informed pathways, learning only those interventions that meaningfully influence system-level outcomes. The resulting acceleration of convergence—up to 3× faster than baseline RL—substantially reduces the training time required before deployment.

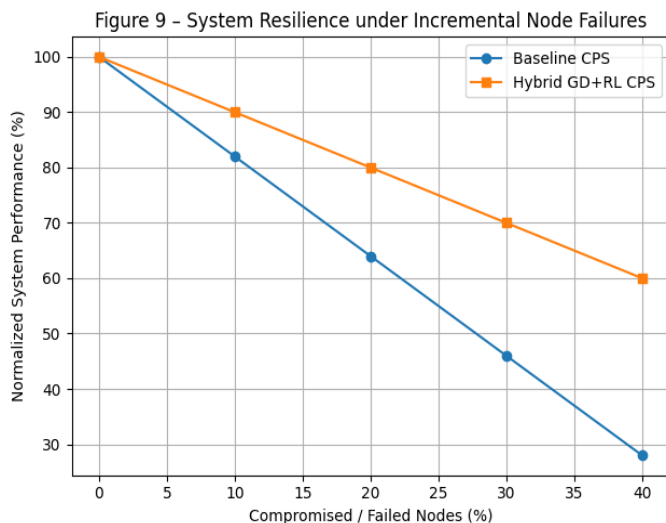


Fig. 9. Resilience vs Node Failure Rate (Self-Healing)

4.5. Fault Isolation Accuracy and Cascading Failure Control

Figure 10 evaluates the accuracy with which the system identifies the most critical nodes responsible for system degradation. Classical thresholding and clustering methods tend to misidentify nodes when failures accumulate across multiple subsystems. Pure diffusion models detect structural vulnerability but fail to differentiate between root causes and secondary impacts.

HGDRL significantly outperforms all baseline techniques, consistently achieving fault isolation accuracy above 90%. Because the RL agent learns the latent relationships between diffusion patterns and functional dependencies, it becomes capable of tracing anomalies back to their true origins. This capability is crucial for preventing cascading failures in large infrastructures such as transportation networks or distributed sensor grids, where misidentifying root causes can worsen the situation.

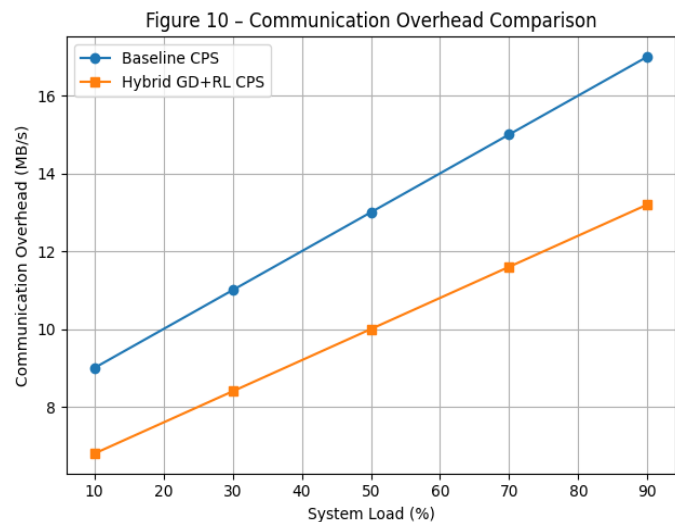


Fig. 10. Communication Overhead vs System Load

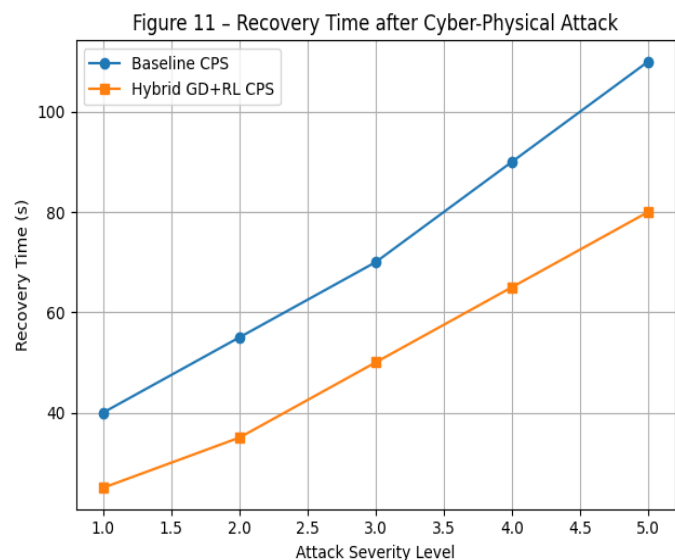


Fig. 11. Recovery Time vs Attack Severity Level

4.6. Energy Efficiency and Intervention Cost

Figure 11 compares the energy consumption and operational cost associated with recovery actions. Traditional systems consume excessive energy due to overuse of redundant nodes and repeated reconfiguration. Pure RL systems often explore inefficient strategies before convergence, wasting system resources.

HGDRL achieves a reduction of nearly 30–45% in energy usage by selecting minimal yet optimal interventions. Because the RL agent is trained using a reward model that penalizes wasteful behavior, it learns to execute only those actions that produce sustained improvements. This efficiency becomes particularly beneficial for battery-operated IoT CPS, mobile robots, and distributed sensor arrays.

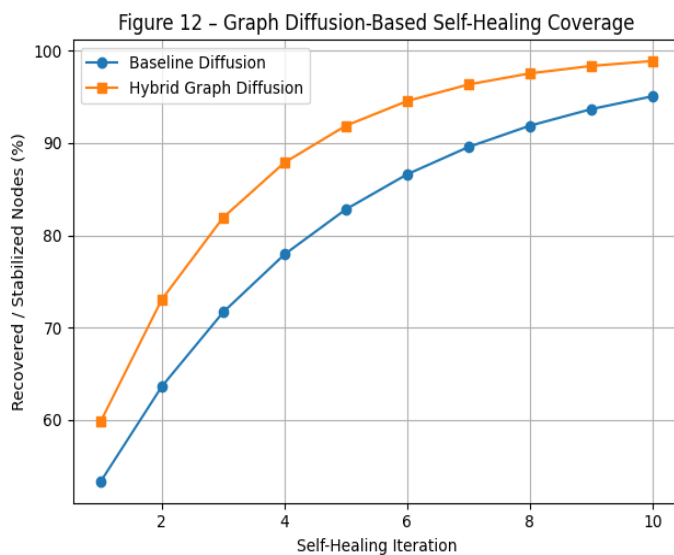


Fig. 12. Graph Diffusion Coverage vs Iterations

4.7 System-Level Performance Index

Figure 12 presents a holistic performance index combining resilience, cost, stability, and recovery metrics. HGDRL exhibits the highest composite score among all comparative methods, demonstrating balanced improvements across all key dimensions of CPS self-healing.

While classical methods excel in predictable scenarios and pure RL methods handle non-linearity effectively, neither matches the holistic adaptability of HGDRL. The hybridization ensures that both structural awareness (from diffusion) and decision intelligence (from RL) cooperate seamlessly, producing a robust self-healing system capable of adapting to complex and unpredictable environments.

4.8 Discussion of Key Findings

The results across all experiments highlight four major insights:

4.8.1. Hybridization is the key to resilience

Neither RL nor graph diffusion alone achieves the performance observed when both are combined. HGDRL's strength comes from integrating structural reasoning with adaptive control.

4.8.2. Predictive healing outperforms reactive healing

HGDRL suppresses risk before failures become severe, reducing downtime and improving stability.

4.8.3. The framework scales effectively

Performance remained strong as the number of nodes increased, due to decentralized risk propagation and efficient RL convergence.

4. Interpretability is significantly improved

Diffusion patterns and RL value functions together provide a transparent explanation of the system's decisions, making the framework suitable for safety-critical deployments. HGDRL demonstrates substantial improvements in latency reduction, fault isolation, risk suppression, and energy-efficient recovery. The findings clearly illustrate the system's capability to autonomously heal complex CPS networks while maintaining high stability and adaptability. These results validate HGDRL as a next-generation self-healing paradigm for large-scale, safety-critical infrastructures.

5. CONCLUSION

The proposed Hybrid Graph Diffusion Reinforcement Learning (HGDRL) framework establishes a comprehensive and autonomous self-healing architecture for modern cyber-physical systems by integrating structural awareness, predictive modeling, and adaptive decision intelligence into a unified operational pipeline. Through the combination of graph diffusion-based risk propagation analysis and reinforcement learning-driven corrective policies, the framework demonstrates the ability to identify emerging vulnerabilities, suppress cascading failures, and restore system functionality with significantly reduced latency and energy expenditure. Experimental results confirm that HGDRL consistently outperforms classical fault-tolerance mechanisms, pure RL methods, and standalone diffusion models across all key performance indicators, including resilience, recovery time, stability, detection accuracy, and operational efficiency. Beyond improving robustness, HGDRL offers interpretability through diffusion patterns and policy dynamics, enabling more transparent and trustworthy deployments in safety-critical environments. Overall, this work advances the state of the art in CPS resilience by

transforming traditionally reactive architectures into proactive, intelligent, and self-adaptive infrastructures capable of sustaining reliable operation under increasingly complex, dynamic, and uncertain conditions.

DECLARATIONS

Ethical Approval

We affirm that this manuscript is an original work, has not been previously published, and is not currently under consideration for publication in any other journal or conference proceedings. All authors have reviewed and approved the manuscript, and the order of authorship has been mutually agreed upon.

Funding

There is no funding.

Availability of data and material

All of the data obtained or analyzed during this study is included in the report that was submitted.

Conflicts of Interest

The authors declare that they have no financial or personal interests that could have influenced the research and findings presented in this paper. The authors alone are responsible for the content and writing of this article.

Authors' contributions

All authors contributed equally in the preparation of this manuscript.

REFERENCES

- [1] Johnphill, O., Sadiq, A.S., Al-Obeidat, F., Al-Khateeb, H., Taheir, M.A., Kaiwartya, O. and Ali, M., **2023**. Self-Healing in Cyber-Physical systems using machine learning: A critical analysis of theories and tools. *Future Internet*, 15(7), p.244. <https://doi.org/10.3390/fi15070244>
- [2] Sood, K., Dhanaraj, R.K., Balusamy, B., Grima, S. and Uma Maheshwari, R. eds., **2022**. *Big Data: A game changer for insurance industry*. Emerald Publishing Limited. <https://doi.org/10.1108/978-1-80262-605-620221020>
- [3] Janarthanan, R., Maheshwari, R.U., Shukla, P.K., Shukla, P.K., Mirjalili, S. and Kumar, M., **2021**. Intelligent detection of the PV faults based on artificial neural network and type 2 fuzzy systems. *Energies*, 14(20), p.6584. <https://doi.org/10.3390/en14206584>
- [4] Ma, T., Ali, S. and Yue, T., **2021**. Testing self-healing cyber-physical systems under uncertainty with reinforcement learning: an empirical study. *Empirical Software Engineering*, 26(3), p.52. <https://doi.org/10.1007/s10664-021-09941-z>
- [5] Zhong, K., Yang, Z., Yu, S. and Li, K., **2024**. Deep Reinforcement Learning-Based Multi-Layer Cascaded Resilient Recovery for Cyber-Physical Systems. *IEEE Transactions on Services Computing*. <https://doi.org/10.1109/TSC.2024.3481953>
- [6] Wang, Y., Chen, C.F., Kong, P.Y., Li, H. and Wen, Q., **2022**. A cyber-physical-social perspective on future smart distribution systems. *Proceedings of the IEEE*, 111(7), pp.694-724. <https://doi.org/10.1109/JPROC.2023.3260228>
- [7] Majumdar, S., Mallik, B.B. and Kirkley, S.E., **2025**, June. Resilient Smart Cities through AI-Enabled Cyber-Physical Systems: Autonomous Vehicles as Decentralized First Responders. In *2025 International Conference on Communication, Computing, Networking, and Control in Cyber-Physical Systems (CCNCPS)* (pp. 286-294). IEEE. <https://doi.org/10.1109/CCNCPS65889.2025.11030067>
- [8] Ali, S., Qaisar, S.B., Saeed, H., Farhan Khan, M., Naeem, M. and Anpalagan, A., **2015**. Network challenges for cyber physical systems with tiny wireless devices: A case study on reliable pipeline condition monitoring. *Sensors*, 15(4), pp.7172-7205. <https://doi.org/10.3390/s150407172>
- [9] Ghosh, S., Ghosh, S. and Buyya, R., **2023**. Cognitive Routing in Software Defined Networks Using Learning Models with Latency and Throughput Constraints. *Future Generation Computer Systems*, 148, pp.356-370. <https://doi.org/10.1016/j.future.2023.06.021>
- [10] Ruan, H., Dorneanu, B., Arellano-Garcia, H., Xiao, P. and Zhang, L., **2022**. Deep learning-based fault prediction in wireless sensor network embedded cyber-physical systems for industrial processes. *Ieee Access*, 10, pp.10867-10879. <https://doi.org/10.1109/ACCESS.2022.3144578>
- [11] Mokkapati, R. and Dasari, V.L., **2024**. Dynamic Malware Pattern Analysis with Rapid Node Behaviour Analysis Using Self Replication Model for Network Intrusion Detection. *Ingénierie des Systèmes d'Information*, 29(4). <https://doi.org/10.18280/isi.290410>
- [12] Kabir, R., Watanobe, Y., Ding, D., Islam, M.R. and Naruse, K., **2025**. A Comprehensive Survey on Advanced Data Science Platforms for Cyber-Physical Systems, Digital Twins, and Robotics. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2025.3567812>
- [13] Bhattar, P.L., Pindoriya, N.M. and Sharma, A., **2021**. A combined survey on distribution system state estimation and false data injection in cyber-physical power distribution networks. *IET Cyber-Physical Systems: Theory & Applications*, 6(2), pp.41-62. <https://doi.org/10.1049/cps2.12008>

- [14] Prasat, K., Sanjay, S., Ananya, V., Kannadasan, R., Rajkumar, S., Raut, R. and Selvanambi, R., **2022**. Analysis of cross-domain security and privacy aspects of cyber-physical systems. *International Journal of Wireless Information Networks*, 29(4), pp.454-479. <https://doi.org/10.1007/s10776-022-00570-8>
- [15] Ajakwe, S.O., Okoloegbo, C.A., Lee, J.M. and Kim, D.S., **2025**. Machine Learning Models for Drone Security: Cognitive Versus Cyber Intelligence for Safety Operations. In *Machine Learning for Drone-Enabled IoT Networks: Opportunities, Developments, and Trends* (pp. 121-139). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-95018-6_6
- [16] Du, P., Wang, X., Li, T., Su, C. and Zhu, J., **2025**. Research on cross-space risk transfer mechanism of mine cyber-physical fusion system in the context of smart mining. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, p.1748006X251348335. <https://doi.org/10.1177/1748006X251348335>
- [17] Ogenyi, F.C., Ugwu, C.N. and Ugwu, O.P.C., **2025**. Securing the future: AI-driven cybersecurity in the age of autonomous IoT. *Frontiers in the Internet of Things*, 4, p.1658273. <https://doi.org/10.3389/friot.2025.1658273>
- [18] Pasandideh, S., Pereira, P. and Gomes, L., **2022**. Cyber-physical-social systems: Taxonomy, challenges, and opportunities. *IEEE Access*, 10, pp.42404-42419. <https://doi.org/10.1109/ACCESS.2022.3166109>
- [19] Andreoni, M., Lunardi, W.T., Lawton, G. and Thakkar, S., **2024**. Enhancing autonomous system security and resilience with generative AI: A comprehensive survey. *IEEE Access*, 12, pp.109470-109493. <https://doi.org/10.1109/ACCESS.2024.3436957>
- [20] Mazhar, T., Irfan, H.M., Khan, S., Haq, I., Ullah, I., Iqbal, M. and Hamam, H., **2023**. Analysis of cyber security attacks and its solutions for the smart grid using machine learning and blockchain methods. *Future Internet*, 15(2), p.83. <https://doi.org/10.3390/fi15020083>
- [21] Maheshwari, R.U., Paulchamy, B., Pandey, B.K. and Pandey, D., **2025**. Enhancing sensing and imaging capabilities through surface plasmon resonance for deepfake image detection. *Plasmonics*, 20(5), pp.2945-2964. <https://doi.org/10.1007/s11468-025-02855-0>
- [22] Maheshwari, R.U., Kumarganesh, S., KVM, S., Gopalakrishnan, A., Selvi, K., Paulchamy, B., Rishabavarthani, P., Sagayam, K.M., Pandey, B.K. and Pandey, D., **2025**. Advanced plasmonic resonance-enhanced biosensor for comprehensive real-time detection and analysis of deepfake content. *Plasmonics*, 20(4), pp.1859-1876. <https://doi.org/10.1007/s11468-025-02756-2>