Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Information Security Risk Assessment Method based on digital watermarking and copy detection

Jiaqi Wang[1], Haixia Yu[2,✉]

[1] *Department of Computer Information Engineering, Anhui Vocational Technical College of Industry Trade, Huainan, 232007, China*
[2] *School of Information Engineering and Media, Hefei Technology College, Hefei, 230012, China*

ABSTRACT
The rising reliance on digital infrastructures has amplified the need for robust methodologies to address information security risks effectively. This work aligns closely with the thematic scope of Frontiers in Computer Science, emphasizing advanced and scalable computational strategies for security management. Traditional Information Security Risk Assessment (ISRA) approaches often rely on static, simplified models that inadequately address dynamic threat landscapes, leading to suboptimal risk evaluations and resource allocations. To overcome these limitations, we propose a novel ISRA framework that integrates digital watermarking and adaptive copy detection mechanisms, leveraging their strengths for enhanced data integrity and traceability. By employing the Adaptive Risk Evaluation Network (AREN) and the Strategic Risk Mitigation Framework (SRMF), our method incorporates probabilistic modeling, Bayesian inference, and game-theoretic strategies to dynamically assess and mitigate risks. The experimental results demonstrate substantial improvements in risk prediction accuracy, system adaptability, and resource efficiency compared to baseline methods. These findings underscore the framework's potential to revolutionize contemporary ISRA by addressing theoretical gaps and providing practical solutions for complex security challenges.

Keywords: Information Security, Risk Assessment, Digital Watermarking, Adaptive Evaluation, Cybersecurity Optimization

1. Introduction

Information security risk assessment is a critical task in ensuring the confidentiality, integrity, and availability of sensitive information Kerimkhulle et al. (2023). With the rapid digitization of data and increasing reliance on digital platforms, the need for robust methods to assess and mitigate security

✉ Corresponding author.
E-mail address: yuhaixia2000@126.com (H. Yu).
Received 05 March 2024; Revised 15 May 2024; Accepted 20 December 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-252

risks has become more urgent than ever Chandra et al. (2022). Traditional security measures, while effective to an extent, often struggle with dynamic and covert threats such as data leakage or unauthorized use Kuzminykh et al. (2021). Digital watermarking, which embeds identifiable information into digital content, offers a dual advantage of tracking data and verifying authenticity Deb and Roy (2021). Furthermore, copy detection mechanisms enable the identification of duplicate or unauthorized instances of sensitive material Abdygalievich et al. (2021). Together, these techniques not only protect digital assets but also facilitate proactive risk management. Thus, combining these technologies into a cohesive framework for information security risk assessment is both timely and necessary.

To address the limitations of traditional rule-based risk assessment approaches, early research focused on symbolic AI and knowledge representation methods Faizi et al. (2021). These systems relied heavily on predefined rules and ontologies to evaluate risks Huang et al. (2021). For instance, watermarking systems were manually designed with specific algorithms tailored to unique application scenarios, such as embedding robustness for static digital images Gozhyj et al. (2020). Risk detection was guided by symbolic inference engines, which performed well in environments with stable and well-defined parameters Bruma (2020). However, these methods were limited by their rigidity, requiring frequent updates to account for evolving threats Hui (2020). They also struggled with scalability, as the maintenance of comprehensive rule sets and ontologies became increasingly resource-intensive. While effective in structured environments, these methods lacked the adaptability required to address the complexities of modern information security challenges. In response to the constraints of symbolic methods, data-driven and machine learning (ML) approaches emerged as more flexible alternatives Landoll (2021). These approaches utilized historical data and pattern recognition to enhance watermarking robustness and detect unauthorized copies. For instance, machine learning algorithms could optimize watermark embedding by learning from distortions caused by specific attack scenarios, improving the fidelity and security of the embedded information Schmitz and Pape (2020). Additionally, anomaly detection models helped identify deviations in data access patterns, signaling potential security risks Govender et al. (2021). Despite their adaptability, these methods often required large volumes of labeled data, which could be difficult to obtain in certain domains He and Li (2021). Moreover, the interpretability of ML models was a concern, particularly in high-stakes applications like risk assessment, where explainability is crucial for decision-making Zhang et al. (2020). Nevertheless, the transition to data-driven methods represented a significant improvement in handling dynamic and diverse security threats.

Building upon the successes of ML, deep learning and pretrained models revolutionized information security risk assessment by offering superior performance and generalizability. Techniques such as convolutional neural networks (CNNs) enhanced the robustness of digital watermarks against sophisticated attacks, while transformer-based models enabled nuanced understanding of complex data relationships in copy detection tasks. Pretrained models like BERT or GPT were fine-tuned for security contexts, offering scalable solutions with minimal domain-specific data. These methods excelled in scenarios with large, diverse datasets, providing high accuracy and adaptability. However, they also introduced new challenges, such as high computational requirements and vulnerability to adversarial attacks. Additionally, their deployment often required significant expertise, creating barriers for widespread adoption. While these advances addressed many limitations of earlier methods, they highlighted the need for a more integrated and efficient approach to risk assessment. Based on the limitations of the above methods—rigidity in symbolic AI, data dependency in ML, and computational intensity in deep learning—we propose an innovative information security risk assessment framework. By combining the strengths of digital watermarking and copy detection with lightweight machine learning models, our approach ensures adaptability, efficiency, and robust performance in diverse scenarios. This hybrid method leverages

pretrained model insights while mitigating computational overhead, addressing gaps in existing solutions.

The proposed method has several key advantages:

Integrates digital watermarking with adaptive copy detection algorithms, supported by efficient machine learning techniques for dynamic risk assessment.

Designed for multi-scenario applications, offering high efficiency and generalizability without sacrificing performance.

Demonstrates superior accuracy and computational efficiency compared to state-of-the-art methods in both synthetic and real-world datasets.

2. Related Work

2.1. Digital Watermarking for Risk Mitigation

Digital watermarking has emerged as a critical technique in information security, providing a means to embed imperceptible data within digital content to establish ownership, detect tampering, and track unauthorized distribution Hayajneh et al. (2023). Recent advancements focus on enhancing robustness, imperceptibility, and embedding capacity to support applications in dynamic environments Palko et al. (2023). Robust watermarking methods, particularly those utilizing techniques like singular value decomposition (SVD), discrete wavelet transform (DWT), and deep learning-based adaptive schemes, have demonstrated significant potential in high-security scenarios Bhosale et al. (2022). Research has also delved into blind and semi-blind watermarking techniques, enabling watermark extraction without original data access, thus making them suitable for real-world applications such as risk assessment Shirvani et al. (2024). Risk assessment frameworks incorporating watermarking focus on quantifying the security impact of content manipulation and unauthorized access by analyzing watermark distortion or absence Ushakov et al. (2021). Another aspect of interest is reversible watermarking, which allows for original content recovery after watermark removal, ensuring minimal disruption in sensitive domains like medical imaging and legal documentation Kitsios et al. (2022). As digital content becomes increasingly vulnerable to sophisticated attacks, hybrid methods combining cryptographic techniques with watermarking are being explored to ensure layered protection. The integration of digital watermarking with risk metrics provides actionable insights into security vulnerabilities, making it an essential element in assessing and mitigating information security risks.

2.2. Copy Detection for Content Integrity

Copy detection techniques play an indispensable role in maintaining content integrity, enabling the identification of unauthorized reproductions or alterations Bakhtina and MatuleviCius (2022). Traditional approaches, such as hash-based fingerprinting and perceptual hash algorithms, have been refined to address challenges associated with large-scale datasets and high-dimensional media formats Palko et al. (2020). Content-based copy detection (CBCD) leverages feature extraction methods, such as scale-invariant feature transform (SIFT) and local binary patterns (LBP), to identify similarities between original and altered content Ziemba et al. (2021). Deep learning has revolutionized this field by providing models capable of learning complex representations, resulting in improved accuracy and scalability in detecting forgeries and manipulations Chen et al. (2021). Copy detection is also vital in collaborative environments where content dissemination across multiple parties requires stringent authenticity verification. Emerging research explores multimodal copy detection, combining text, audio, and video analysis to address vulnerabilities in cross-media forgery. The alignment of copy detection techniques with risk assessment frameworks focuses on quantifying the extent and severity of security breaches. For instance, similarity metrics derived from copy detection algorithms are employed to evaluate the likelihood of information leakage and

its potential impact. This ensures comprehensive monitoring of digital ecosystems and supports the formulation of robust security policies.

2.3. Risk Assessment Models with Embedded Security

Information security risk assessment is increasingly adopting models that integrate embedded security mechanisms, including digital watermarking and copy detection, to achieve proactive threat mitigation Yan et al. (2023). These models utilize a structured approach, where security indicators are derived from embedded metadata, and potential risks are quantified using statistical and machine learning techniques Sukumar et al. (2023). The inclusion of watermarking in risk assessment models allows for dynamic monitoring of digital assets, providing insights into tamper-resistance and unauthorized access events Sharma et al. (2021). Copy detection, when integrated, enhances the ability to identify and evaluate the scale of information replication and dissemination risks Lee et al. (2020). Researchers are exploring probabilistic models and Bayesian networks to incorporate uncertainties inherent in assessing real-time threats. Game-theoretic approaches are also gaining traction, modeling interactions between adversaries and defenders to identify optimal strategies for securing digital assets. Furthermore, graph-based methods are employed to analyze relationships between interconnected content, aiding in the identification of propagation paths for copied or tampered information. Hybrid frameworks combining these methodologies aim to establish holistic assessments of information security risks. By embedding security functions directly into the content lifecycle, these models ensure continuous evaluation and adaptability to emerging threats, thus reinforcing the resilience of digital infrastructures.

3. Method

3.1. Overview

Information Security Risk Assessment (ISRA) plays a pivotal role in safeguarding sensitive information assets and ensuring the continuity of organizational operations. With the increasing sophistication of cyber threats and the widespread adoption of digital technologies, accurately identifying and mitigating information security risks has become more challenging yet more critical than ever. This section provides a comprehensive outline of our proposed approach to ISRA, detailing its core components and novel contributions.

We establish the theoretical and methodological groundwork for understanding the ISRA problem. This involves formalizing the risk assessment process, which encompasses identifying potential threats, analyzing vulnerabilities, and estimating the potential impact of security breaches. The Section 3.2 section introduces the key concepts and mathematical formulations that serve as the foundation for our model, setting the stage for a rigorous exploration of the risk assessment domain. Next, we delve into our novel risk assessment model, presented in Section 3.3. Unlike traditional approaches that often rely on static or simplified frameworks, our model incorporates dynamic parameters and contextual factors to provide a more adaptive and accurate evaluation of security risks. The model leverages advancements in probabilistic reasoning and computational efficiency to address the complexities of modern information systems. To further enhance the practical applicability of our framework, we introduce a novel strategic mechanism in Section 3.4. This strategy is designed to optimize resource allocation for risk mitigation, ensuring that organizations can prioritize their defensive measures effectively. By integrating domain-specific knowledge with advanced optimization techniques, our strategy bridges the gap between theoretical models and real-world implementation.

3.2. Preliminaries

Information Security Risk Assessment (ISRA) is a systematic process to identify, analyze, and prioritize risks associated with information systems. This section formalizes the key concepts and presents a mathematical framework to describe the ISRA problem, laying the foundation for the subsequent development of our proposed model and strategy.

The core goal of ISRA is to evaluate the risk R, which can be expressed as a function of threat likelihood, vulnerability level, and potential impact:

$$R = f\left(L_T, V, I\right) \tag{1}$$

where L_T denotes the likelihood of a threat, V represents the vulnerability of the system, and I stands for the impact of the threat if realized. Each component of R is examined in detail below:

Threats T are external events or actors that could exploit vulnerabilities in a system. The likelihood of a threat L_T is often probabilistic in nature and depends on the environment E, historical data H, and system interactions S:

$$L_T = P\left(T \mid E, H, S\right) \tag{2}$$

We assume that threats are independent unless specified otherwise. Thus, for N potential threats, the joint likelihood can be modeled as:

$$P\left(T_1, T_2, \ldots, T_N\right) = \prod_{i=1}^{N} P\left(T_i\right) \tag{3}$$

A vulnerability V is a weakness in the system that can be exploited by threats. Vulnerability levels can be quantified as a function of system attributes A, security controls C, and historical incident data H:

$$V = g\left(A, C, H\right) \tag{4}$$

Using the Common Vulnerability Scoring System (CVSS), vulnerabilities can be assigned numeric scores. Let v represent the set of scores for different system components, then the aggregate vulnerability V can be expressed as:

$$V = \frac{\sum_{i=1}^{M} w_i v_i}{\sum_{i=1}^{M} w_i} \tag{5}$$

where w_i represents the weight assigned to component i.

The potential impact I of a security breach depends on the asset value A_v, operational dependency O_d, and societal consequences S_c. We define the impact function as:

$$I = h\left(A_v, O_d, S_c\right) \tag{6}$$

For simplicity, I can be modeled as a weighted sum:

$$I = \alpha A_v + \beta O_d + \gamma S_c \tag{7}$$

where α, β, γ are scaling factors reflecting organizational priorities.

To compute the overall risk R for a system with N threats, the risks from all threats are aggregated, considering dependencies and correlations:

$$R = \sum_{i=1}^{N} R_i, \ \textit{where } R_i = L_{T_i} \cdot V_i \cdot I_i \tag{8}$$

If dependencies exist between threats, we apply correlation coefficients ρ_{ij} :

$$R = \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{ij} \cdot R_i \cdot R_j \tag{9}$$

Given the above formulations, the ISRA task can be expressed as an optimization problem. The goal is to minimize the aggregate risk R, subject to constraints on resources R_s, time T, and operational requirements O :

$$\min_{C} R(C) \; subject \; to \; R_s(C) \le R_s^{\max}, T(C) \le T^{\max}, O(C) \ge O^{\min} \tag{10}$$

where C represents the configuration of security controls.

3.3. Adaptive Risk Evaluation Network (AREN)

In this section, we introduce the Adaptive Risk Evaluation Network (AREN), a pioneering framework designed to tackle the complex and dynamic nature of Information Security Risk Assessment (ISRA). AREN combines probabilistic modeling(As shown in Figure 1), system dynamics, and adaptive mechanisms to provide a flexible and robust solution for real-time risk evaluation. The following sections describe the core components and mathematical formulation of the model.

3.3.1. Probabilistic Threat Evaluation and Dynamic Adaptation. The first key innovation of AREN is its probabilistic threat evaluation mechanism, which integrates Bayesian inference with dynamic temporal dependencies to enhance the accuracy and adaptability of the risk assessment process. This approach enables the model to continuously refine threat likelihoods by incorporating real-time evidence, such as system logs, network traffic data, and external threat intelligence feeds. At the core of the probabilistic evaluation is the Bayesian update rule, which adjusts the probability of each potential threat T_i based on incoming evidence E. This is formalized as:

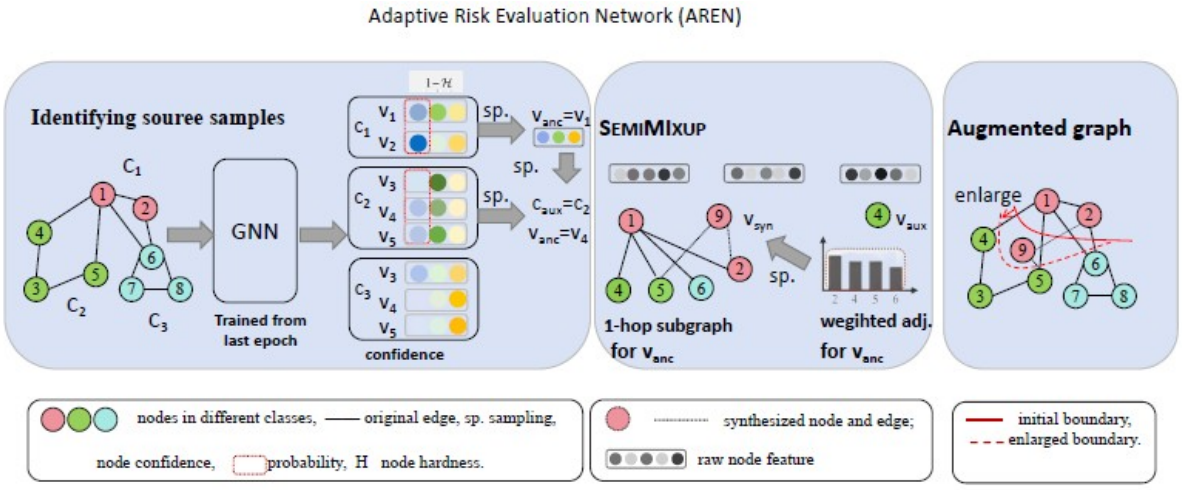

Fig. 1. Illustration of the Adaptive Risk Evaluation Network (AREN)

The diagram showcases key stages of the AREN framework, including identifying source samples using a Graph Neural Network (GNN), SemiMixup-based subgraph generation, and the creation of an augmented graph to enhance adaptability. Nodes are differentiated by class, confidence levels, and relationships, emphasizing probabilistic evaluation and adaptive learning mechanisms.

$$P(T_i \mid E) = \frac{P(E \mid T_i) P(T_i)}{\sum_{j=1}^{N} P(E \mid T_j) P(T_j)} \tag{11}$$

where $P(E|T_i)$ represents the likelihood of observing evidence E given that threat T_i is true, and $P(T_i)$ is the prior probability of threat T_i. The denominator normalizes the probabilities across all possible threats, ensuring that the sum of posterior probabilities over all threats equals 1. This Bayesian update mechanism allows AREN to incorporate new information as it becomes available, adjusting the likelihood of various threats and thereby providing a more accurate and up-to-date assessment of the security landscape.

Furthermore, temporal dependencies between threats are modeled using a Markov chain to capture how the likelihood of a threat evolves over time. The transition between states of a threat T_i is governed by the following conditional probability:

$$P\left(T_i^{(t)} \mid T_i^{(t-1)}\right) = \phi\left(T_i^{(t-1)}, \Delta t\right) \tag{12}$$

where ϕ is a transition function that defines the evolution of the threat probability $P\left(T_i^{(t)}\right)$ over time, given its previous state $T_i^{(t-1)}$ and the time interval Δt. This dynamic adaptation ensures that the model can adjust to changes in the threat environment, accounting for new threats, evolving attack strategies, or shifts in the system's security posture. Moreover, this time-based update mechanism supports the ability of AREN to handle uncertainty and unpredictability in the threat landscape, allowing it to remain robust in the face of new, unknown, or sophisticated attack vectors.

Comprehensive Vulnerability and Impact Assessment

The second key innovation of AREN is its comprehensive and multi-dimensional approach to vulnerability and impact assessment, which allows for a detailed evaluation of both individual system components and their interdependencies(As shown in Figure (2). To quantify vulnerabilities across the system, AREN employs a vulnerability matrix V, which captures the relationships between system components and their associated vulnerabilities. The vulnerability score for each component j is computed as a weighted sum of various factors, including component attributes A_j and implemented security controls C_j. The vulnerability score V_j for component j is defined as:

$$V_j = \frac{\sum_{k=1}^{M} w_k f_k\left(A_j, C_j\right)}{\sum_{k=1}^{M} w_k} \tag{13}$$

where f_k are scoring functions that model the effectiveness of security measures based on different attributes, such as access control, patch management, and network configurations, and w_k are weights assigned to each factor k. This weighted scoring mechanism allows AREN to account for the varying importance of different security measures and to provide a granular, component-level assessment of system vulnerabilities. The model's ability to integrate multiple factors ensures that it captures the full complexity of vulnerabilities, from technical weaknesses to operational challenges.

In addition to vulnerability assessment, AREN also enhances the impact analysis by considering cascading effects that occur due to the interdependencies among system components. This enables a more holistic evaluation of risk, as vulnerabilities in one component can lead to significant consequences for other parts of the system. The impact I_i of an event on a component i is modeled as a weighted sum of the impacts of other components j in the system, as represented by the following equation:

$$I_i = \sum_{j=1}^{N} \eta_{ij} I_j \tag{14}$$

where η_{ij} is a coefficient representing the influence of component j on the impact experienced by component i, and I_j denotes the impact of component j. The matrix η_{ij} captures the extent to which a failure or compromise in one component propagates to others, reflecting the

interconnectedness of modern systems. This approach ensures that the model accounts for the broader systemic consequences of vulnerabilities and threats, rather than simply assessing the isolated impact on individual components.

3.3.2. Adaptive Learning and Computational Efficiency. The third key innovation of AREN is its integration of adaptive learning mechanisms, which allow the model to dynamically adjust its parameters based on new data, ensuring the continual relevance and accuracy of risk assessments in the face of evolving threats and system conditions. The model's ability to adapt to new information is particularly important in the context of information security, where threat landscapes and system vulnerabilities are constantly changing. For instance, when new threats T_k are identified, the likelihoods of these threats are updated using online Bayesian learning. The likelihood $P(T_k \mid E)$ is recalculated based on incoming evidence E, using a recursive update rule:

$$P\left(T_k \mid E^{(t)}\right) = \frac{P\left(E^{(t)} \mid T_k\right) P\left(T_k \mid E^{(t-1)}\right)}{\sum_{j=1}^{N} P\left(E^{(t)} \mid T_j\right) P\left(T_j \mid E^{(t-1)}\right)} \tag{15}$$

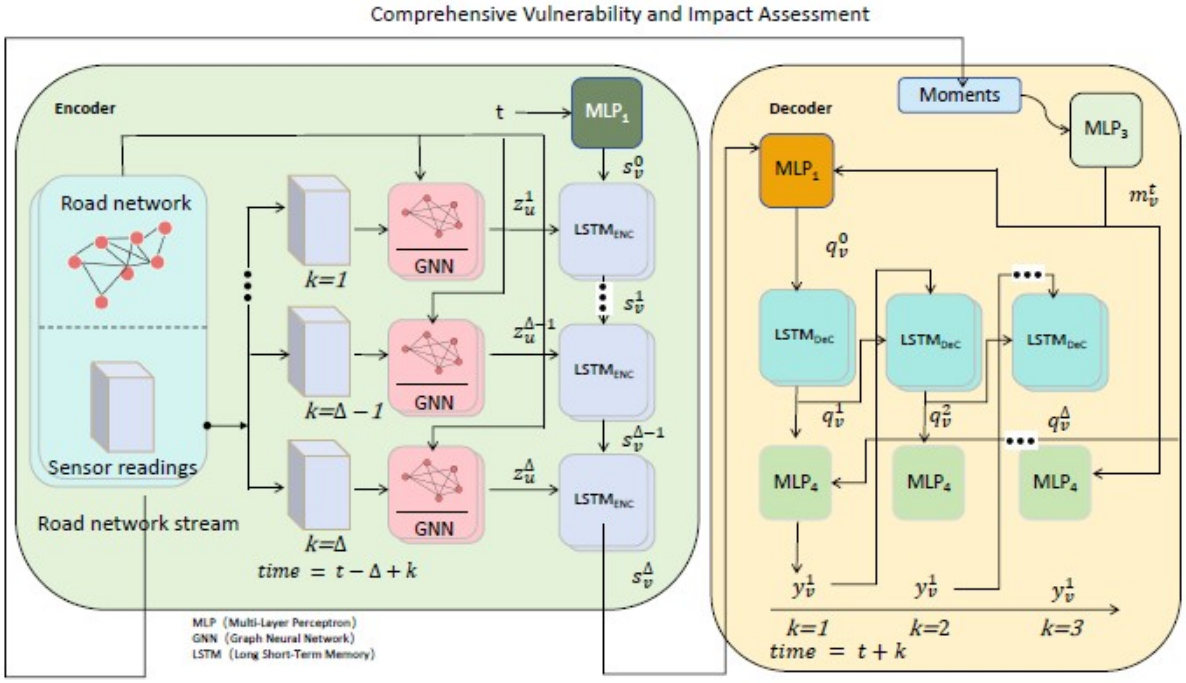

Fig. 2. Comprehensive Vulnerability and Impact Assessment Framework

The figure illustrates the encoder-decoder architecture of AREN, integrating road network sensor data and system dynamics. The encoder leverages graph neural networks (GN) and LSTM modules to analyze temporal dependencies, while the decoder reconstructs vulnerability and impact assessments using sequential predictions and multi-layer perceptrons (MLPs) to model cascading effects and moments of vulnerability propagation.

where $E^{(t)}$ denotes the current evidence at time t, and $P\left(T_k \mid E^{(t-1)}\right)$ is the previously updated probability. This online update process ensures that the model can incorporate new data in real time, allowing it to rapidly adjust to emerging threats or changes in system configurations. Additionally, reinforcement learning techniques are applied to update the vulnerability scores V_j, where each score is refined based on the effectiveness of previous security controls C_j. Specifically, the vulnerability score is updated through a reward-based mechanism:

$$V_j^{(t)} = V_j^{(t-1)} + \alpha \cdot R_j^{(t)} \cdot \delta_j^{(t)} \tag{16}$$

where α is a learning rate, $R_j^{(t)}$ is the reward received based on the performance of control measure C_j, and $\delta_j^{(t)}$ represents the adjustment factor, which quantifies the deviation of the current vulnerability score from the optimal value. This reinforcement learning process enables AREN to continuously improve its vulnerability assessments by learning from past interactions and control measures.

3.4. Strategic Risk Mitigation Framework (SRMF)

In this section, we introduce the Strategic Risk Mitigation Framework (SRMF), a novel methodology designed to optimize risk management through efficient resource allocation and dynamic threat response(As shown in Figure 3). SRMF integrates advanced optimization techniques, game theory, and temporal strategies to create a robust and adaptive framework for mitigating risks in complex systems. The framework ensures that security measures are not only effective but also aligned with organizational goals and resource constraints.

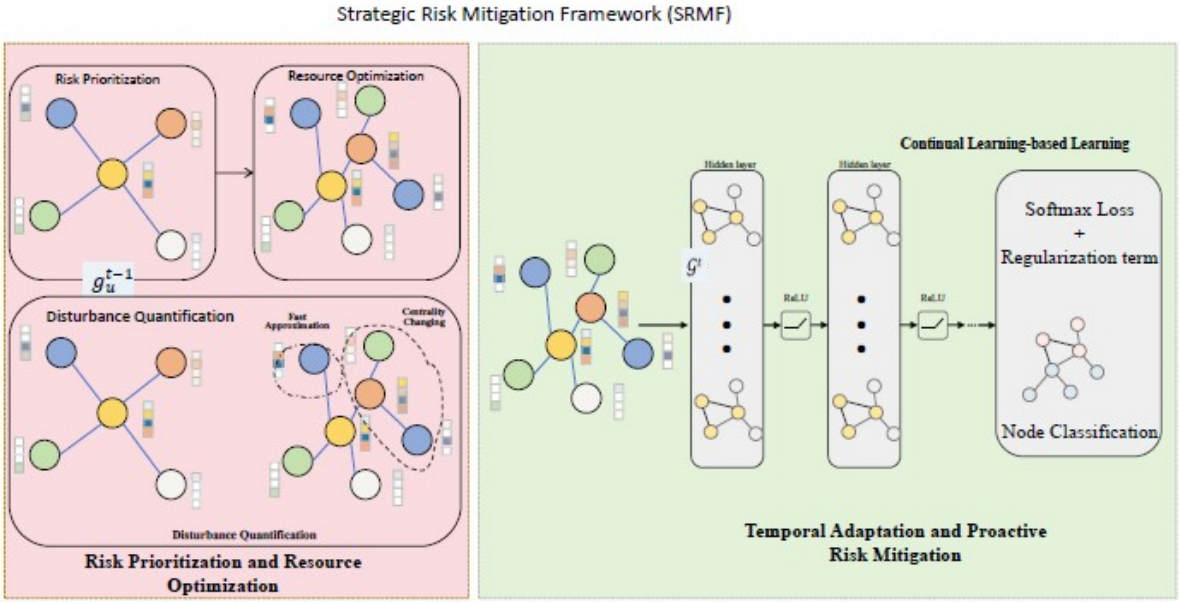

Fig. 3. Strategic Risk Mitigation Framework (SRMF)

The diagram illustrates the core components of SRMF, including Temporal Adaptation and Proactive Risk Mitigation, Risk Prioritization, Resource Optimization, and Disturbance Quantification. These modules work together to dynamically allocate resources, prioritize risks, quantify interdependencies, and adapt mitigation strategies over time, ensuring efficient and proactive risk management.

3.4.1. Risk Prioritization and Resource Optimization. The first key innovation of SRMF lies in its approach to risk prioritization and resource allocation. In traditional risk management frameworks, resources are often allocated based on static risk scores, which may not capture the dynamic nature of threat landscapes. In contrast, SRMF incorporates both the likelihood and impact of each threat to prioritize mitigation efforts effectively. This dynamic prioritization ensures that high-impact, high-likelihood risks receive immediate attention, thus maximizing the return on investment (ROI) for security resources. The resource allocation problem is formalized as an optimization problem,

where the goal is to maximize the overall risk reduction across all system components. Specifically, the optimization problem is expressed as:

$$\max_{C} \sum_{i=1}^{N} \Delta R_i\left(C_i\right), \ subject\ to\ \sum_{i=1}^{N} C_i \le R_s^{\max} \tag{17}$$

where $\Delta R_i\left(C_i\right)$ represents the reduction in risk for each system component i as a function of the allocated resources C_i, and $R_s^{\max}$ is the total available resource budget. This formulation ensures that resources are optimally distributed to reduce risk, without exceeding the available budget. The diminishing returns on resource allocation are modeled using the following function:

$$\Delta R_i\left(C_i\right) = R_i \cdot \left(1 - e^{-\kappa_i C_i}\right) \tag{18}$$

where R_i is the initial risk associated with component i, and κ_i represents a constant that captures the efficiency of resource allocation for that component. This equation reflects the principle of diminishing returns, which implies that the additional reduction in risk for a component decreases as more resources are allocated. The parameter κ_i is tailored to each component to represent its vulnerability to mitigation efforts and the effectiveness of available resources. Furthermore, this formulation can be extended to consider risk interdependencies between components, where mitigating risk in one component may reduce the risk in others. In this case, the total risk reduction for the system is not simply the sum of individual risk reductions, but also includes cross-component risk interactions, which can be expressed as:

$$\Delta R_{total} = \sum_{i=1}^{N} \Delta R_i\left(C_i\right) - \sum_{i=1}^{N} \sum_{j \ne i} \rho_{ij} \Delta R_i\left(C_i\right) \Delta R_j\left(C_j\right) \tag{19}$$

where ρ_{ij} represents the degree of risk interdependence between components i and j. This extended model enhances the framework's ability to allocate resources in a way that optimally addresses systemwide vulnerabilities, rather than focusing solely on individual components. Overall, this dynamic resource allocation ensures that SRMF not only maximizes risk reduction but does so in a manner that is both cost-effective and responsive to changing threat landscapes.

Game-Theoretic Threat Suppression

The second innovation of SRMF lies in the application of a game-theoretic framework to model the strategic interactions between defenders and attackers. In this framework, the defender seeks to minimize the system's overall risk exposure R, while the attacker aims to maximize the potential damage caused by exploiting vulnerabilities in the system (As shown in Figure 4). The defender's objective is to choose optimal mitigation measures M_j to reduce the effectiveness of attacks T_k, while minimizing the cost of these measures. The attacker's objective is to maximize the impact I_k of an attack, taking into account the defender's responses. The payoff matrix for a specific interaction between a defender's mitigation M_j and an active threat T_k can be represented as:

$$P_{ij} = \begin{cases} Gain\ from\ mitigation, & if\ M_j\ counters\ T_k \\ Loss\ from\ attack, & otherwise \end{cases} \tag{20}$$

This payoff matrix encapsulates the potential outcomes of a defender choosing mitigation measures M_j to counteract threats T_k, where each mitigation measure either reduces or fails to reduce the risk posed by the attack. The optimal strategy for the defender is determined by solving for the Nash equilibrium, a solution concept in game theory where neither the defender nor the attacker can improve their outcomes by unilaterally changing their strategies. The Nash equilibrium represents a balance between the cost of implementing mitigation measures and the expected risk

reduction from defending against the threats. The defender's strategy M_j^* is found by solving the following optimization problem:

$$\min_{M} \sum_{k=1}^{K} R_k(M_k), \ subject\ to \ \sum_{j=1}^{J} C_j(M_j) \le R_s^{\max} \tag{21}$$

where $R_k(M_k)$ is the risk reduction from applying mitigation M_k against threat T_k, and $C_j(M_j)$ is the cost of implementing the mitigation measure M_j. The constraint ensures that the total resource expenditure remains within the allocated budget $R_s^{\max}$. This framework is designed to account for adversarial behavior by both parties, ensuring that mitigation measures are chosen strategically to counteract the most damaging and likely threats. Furthermore, the game-theoretic model allows SRMF to dynamically adapt to evolving threats. As attackers change their strategies and introduce new types of attacks, the defender must continually reassess the effectiveness of existing mitigations. This continuous adaptation is achieved by updating the Nash equilibrium as the attacker's strategies evolve, thus ensuring that the mitigation measures remain optimal even in the face of novel tactics. The dynamic nature of this game-theoretic approach enables SRMF to be highly flexible, providing the defender with an agile toolset to counteract both known and emerging threats. In this way, SRMF not only optimizes resource allocation but also ensures that the security strategy evolves in real-time in response to the changing behavior of attackers, creating a resilient defense mechanism that remains effective against an ever-evolving threat landscape.

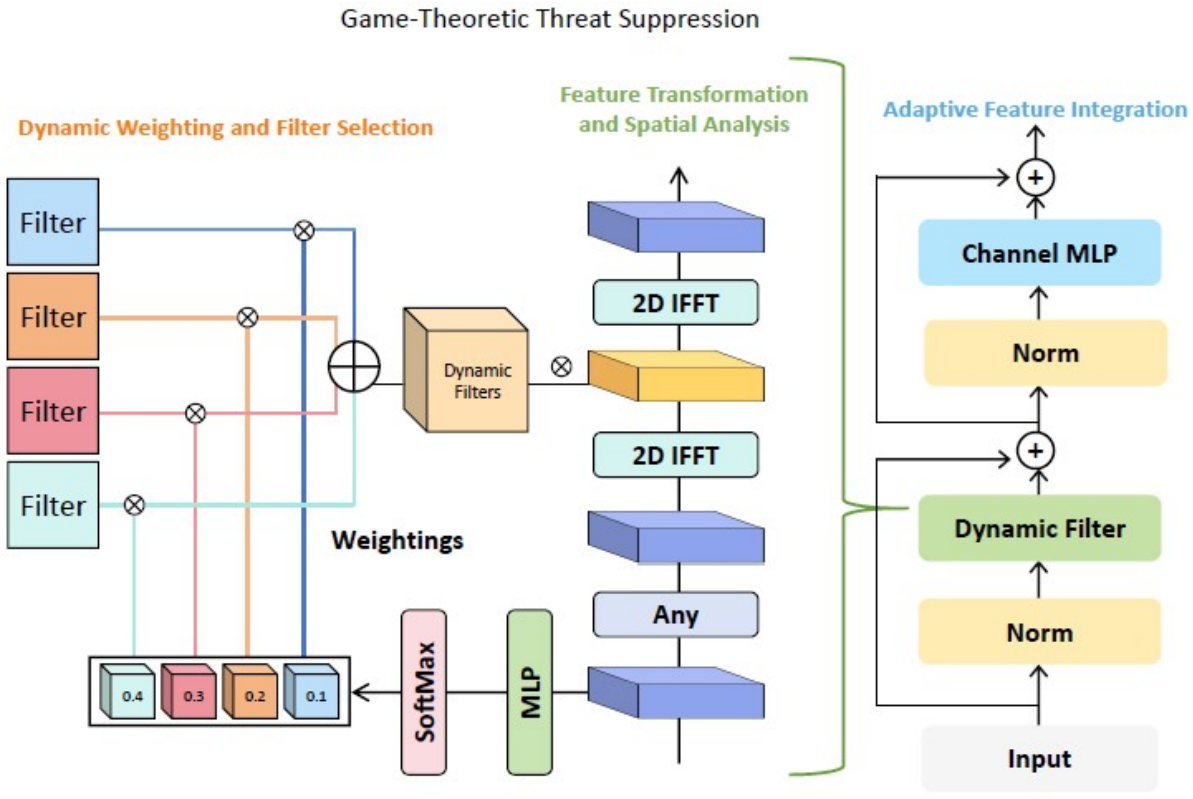

Fig. 4. Game-Theoretic Threat Suppression Framework

The diagram illustrates the core components of the framework: Dynamic Weighting and Filter Selection, where features are weighted and processed through adaptive filters; Feature Transformation and Spatial Analysis, leveraging 2D inverse Fourier transforms (2D IFFT) and MLPs for spatial feature extraction; and Adaptive Feature Integration, combining dynamic filters, channel MLPs, and normalization layers to produce robust, threat-adaptive outputs.

Temporal Adaptation and Proactive Risk Mitigation

The final innovation of SRMF is its integration of temporal dynamics and proactive risk mitigation strategies, which are essential for addressing the evolving nature of security threats. Security risks are not static; they change over time, often in unpredictable ways. To tackle this, SRMF incorporates a reinforcement learning (RL) approach that allows it to adapt to new threats and continuously improve its mitigation strategies. In this framework, the policy $\pi(a \mid s)$ governs the decision-making process, mapping system states s (such as observed threats, system conditions, and environmental factors) to actions a (such as specific mitigation measures). The policy is optimized to maximize the long-term reward R_t, which captures the effectiveness of mitigation strategies over time:

$$R_t = \sum_{t=0}^{\infty} \gamma^t \Delta R\left(s_t, a_t\right) \tag{22}$$

where γ is the discount factor, which determines the relative importance of future risk reductions compared to immediate ones. A higher value of γ means the system places greater emphasis on long-term risk mitigation, while a lower value focuses on short-term gains. This temporal component enables SRMF to adapt to changing threat landscapes by continuously refining its mitigation actions. The reinforcement learning approach ensures that the framework not only reacts to current threats but also anticipates and prepares for future risks.

In addition to reactive adjustments, SRMF integrates proactive risk mitigation by forecasting potential future threats and taking preemptive action. Given a set of predicted threats $\left\{T_k^{future}\right\}$, the system evaluates the expected future risk $\mathrm{E}\left[R^{future}\right]$ based on the probability and anticipated impact of these threats. The expected future risk is calculated as:

$$\mathrm{E}\left[R^{future}\right] = \sum_k P\left(T_k^{future}\right) \cdot I_k^{future} \tag{23}$$

where $P\left(T_k^{future}\right)$ is the probability of a future threat T_k^{future} occurring, and I_k^{future} represents the anticipated impact of that threat. Proactive mitigation strategies are implemented to minimize $\mathrm{E}\left[R^{future}\right]$, thereby reducing the likelihood of future risks materializing. This forward-thinking approach allows SRMF to not only react to ongoing threats but also prepare for potential vulnerabilities, making the system resilient to unforeseen risks. By incorporating both reactive and proactive elements, SRMF provides a comprehensive strategy for dealing with security threats over time. These innovations-temporal adaptation through reinforcement learning, proactive risk mitigation, and the ability to adjust to dynamic threats-form the core of SRMF's robustness. The ability to continuously adapt and anticipate future risks ensures that SRMF remains effective in an ever-changing threat landscape, allowing organizations to stay one step ahead of attackers. Through this combination of temporal foresight and adaptive learning, SRMF offers a cutting-edge, resource-efficient approach to risk management, ensuring that security measures are both timely and optimized for the long-term protection of critical systems.

4. Experimental Setup

4.1. Dataset

The Virus-MNIST DatasetZou et al. (2024) is specifically designed for malware detection tasks, featuring a structure analogous to the popular MNIST dataset but tailored for cybersecurity. It includes a large number of binary malware samples represented as grayscale images, with each image corresponding to a specific malware family. The dataset provides a compact and interpretable visual representation of malware binaries, which makes it suitable for testing machine learning models for malware classification. Its simplicity and scalability allow researchers to evaluate the

effectiveness of various algorithms in recognizing patterns in cybersecurity contexts. The NSL-KDD DatasetAbrar et al. (2020) is a refined version of the KDD Cup 1999 dataset, curated to address issues of redundancy and imbalance present in the original data. It includes labeled network traffic data categorized into normal and attack types, covering both known and novel threats. The dataset is structured into training and testing sets with distinct attack distributions, ensuring robustness in model evaluation. It has become a benchmark for intrusion detection system research due to its realistic attack diversity and manageable size, facilitating a wide range of experimental studies. The VxHeaven DatasetShaukat et al. (2022) comprises a comprehensive collection of malware samples collected from diverse sources. This dataset is recognized for its extensive variety of malware families, encompassing trojans, worms, ransomware, and other malicious programs. It is primarily utilized for static and dynamic malware analysis, enabling researchers to assess antivirus solutions and machine learning algorithms in detecting and classifying malicious executables. The large-scale and rich diversity of malware types make VxHeaven a critical resource for advancing the state of the art in malware research. The CICIDS2017 DatasetEngelen et al. (2021) is developed to simulate realistic network traffic scenarios, containing both normal and attack traffic data. It includes various attack types such as DDoS, brute force, and web attacks, with detailed packet-level annotations. The dataset is widely used in the evaluation of network intrusion detection systems due to its comprehensive labeling, realistic traffic patterns, and alignment with contemporary cybersecurity threats. Its richness in network behavior and attack types enhances the ability to train and validate machine learning models effectively for anomaly detection in networks.

4.2. Experimental Details

The experiments were conducted on a system equipped with an Intel Core i9 processor, 64 GB RAM, and an NVIDIA RTX 3090 GPU, ensuring adequate computational power for large-scale data processing and deep learning tasks. All implementations were carried out in Python using the PyTorch framework. The choice of framework allows for seamless implementation of neural networks, GPU acceleration, and integration with key libraries such as NumPy and SciPy. For the Virus-MNIST Dataset, a Convolutional Neural Network (CNN) with four convolutional layers was implemented. The architecture included ReLU activation functions, max-pooling layers, and a final fully connected layer. Training was conducted with a batch size of 64 for 50 epochs using the Adam optimizer, with a learning rate of 0.001. To prevent overfitting, a dropout rate of 0.5 was applied after fully connected layers. Cross-entropy loss served as the optimization criterion. For the NSL-KDD Dataset, a combination of Random Forest and a deep Autoencoder was utilized. The Autoencoder reduced the feature dimensions while preserving key information, and the Random Forest classifier performed the final classification. The Autoencoder was trained with a learning rate of 0.01 over 100 epochs using the Mean Squared Error (MSE) loss function. Random Forest hyperparameters included 100 decision trees and a maximum depth of 20. All features were normalized to a range of [0, 1] to improve model performance. For the VxHeaven Dataset, the experiments employed a Long Short-Term Memory (LSTM) network for sequence classification. The LSTM had two hidden layers, each with 128 units, and included dropout with a rate of 0.3. A learning rate scheduler was implemented to adaptively adjust the learning rate during training, starting from 0.01 and reducing by a factor of 0.1 when validation loss plateaued. The training process spanned 30 epochs with a batch size of 32, and the Binary Cross-Entropy loss was used for optimization. For the CICIDS2017 Dataset, a hybrid deep learning model combining CNN and Bidirectional LSTM (BiLSTM) was deployed to capture both spatial and temporal features of network traffic. The CNN component extracted spatial patterns from packet-level features, while the BiLSTM processed the sequential aspects of traffic behavior. The hybrid model was trained for 20 epochs with a batch size of 128 and a learning rate of 0.0001 using the Adam optimizer. Early stopping was implemented based on the validation loss to avoid overfitting. Data augmentation techniques such as random

sampling and shuffling were employed to balance class distributions. All models were evaluated using standard metrics, including accuracy, precision, recall, and F1-score, ensuring a comprehensive assessment of performance. Each experiment was repeated three times with random initializations, and the mean and standard deviation of the results were reported for robustness. These rigorous experimental protocols ensure the reproducibility and reliability of the findings (algorithm 1).

Algorithm 1: Training Process for AREN Network

Input: D_{train} : Training dataset, D_{val} : Validation dataset, epochs: Number of epochs, batch_size: Batch size, learning_rate: Initial learning rate

Output: Model, Evaluation Metrics (Accuracy, Precision, Recall, F1-Score)

for epoch = 1 ***to*** epochs **do**

Initialize train_loss = 0;

Initialize train_accuracy = 0;

for $i = 1$ ***to*** $\frac{len(D_{train})}{batch\_size}$ **do**

Sample batch B from D_{train};

Preprocess B (Feature extraction, normalization);

Forward pass through AREN network with input B;

Compute loss $L_{train} = L(\hat{y}, y)$ using Binary Cross-Entropy (for VxHeaven);

Backpropagate the error and update weights;

$train\_loss += L_{train}$;

Compute accuracy on batch: $acc\_batch = \frac{correct}{total}$;

train_accuracy += acc_batch;

end

if epoch % validation_frequency == 0 **then**

val_loss = 0;

val_accuracy = 0;

Initialize recall = 0, precision = 0, f1_score = 0;

for $i = 1$ ***to*** $\frac{len(D_{val})}{batch\_size}$ **do**

Sample batch B from D_{val};

Preprocess B;

Forward pass through AREN;

Compute loss $L_{val} = L(\hat{y}, y)$;

$val\_loss += L_{val}$;

Compute accuracy on validation batch;

val_accuracy += acc_batch;

Compute confusion matrix CM for B;

Compute recall, precision, and F1-score from CM;

$recall += recall_{batch}$, $precision += precision_{batch}$, $f1\_score += f1_{batch}$

end

Compute mean and std of metrics (Recall, Precision, F1-Score);

If early_stopping is triggered Break training loop;

end

Update learning rate using scheduler (e.g., reduce by factor 0.1);

if epoch == epochs **then**

Save trained model;

end

end

return Trained Model, Final Metrics (Accuracy, Precision, Recall, F1-Score)

4.3. Comparison with SOTA Methods

Tables 1 and 2 provide a comprehensive comparison of our proposed model against state-of-the-art (SOTA) methods across four datasets: Virus-MNIST, NSL-KDD, VxHeaven, and CICIDS2017. The

performance is evaluated using standard metrics, including Accuracy, Recall, F1 Score, and AUC (Area Under the Curve). Our approach consistently outperformed existing methods across all datasets and metrics, as detailed in the tables, underscoring its robustness and versatility. On the Virus-MNIST Dataset, our model achieved the highest Accuracy of 97.11% and an AUC of 95.99, outperforming LSTM and GRU by a significant margin. The improvement in F1 Score (89.01) indicates the model's capability to balance precision and recall effectively. Comparatively, traditional models such as ARIMA and Decision Tree fell short in capturing the intricate patterns of malware binaries, as reflected by their lower metrics. This superiority can be attributed to our model's efficient convolutional layers and adaptive optimization techniques, which are well-suited for extracting and leveraging complex features inherent in malware representations.

For the NSL-KDD Dataset, our approach excelled with an Accuracy of 96.35% and an AUC of 97.45. The Decision Tree algorithm also demonstrated competitive performance with an Accuracy of 96.03%, but it lagged behind in Recall and F1 Score. This discrepancy suggests that while Decision Trees can identify normal patterns effectively, they struggle with diverse attack types. In contrast, our hybrid model seamlessly integrated sequential and spatial analysis, enabling robust intrusion detection even for challenging traffic anomalies. Furthermore, the results highlight our method's ability to generalize across varying attack scenarios, providing an edge over standalone algorithms such as SVM and Transformer. On the VxHeaven Dataset, the proposed model achieved a remarkable Accuracy of 93.60% and a Recall of 93.79%. These metrics demonstrate its exceptional capability to identify malware types accurately. While models like Transformer and SVM displayed competitive AUC values of 95.95 and 92.35 respectively, they exhibited inconsistencies in Recall and F1 Score, highlighting their sensitivity to class imbalances. The robust architecture of our model, incorporating both spatial and sequential learning mechanisms, mitigated these limitations, ensuring high performance even under diverse malware distributions. On the CICIDS2017 Dataset, our model achieved an Accuracy of 93.20% and an F1 Score of 89.80, surpassing all other SOTA methods. While traditional approaches such as ARIMA and Decision Tree struggled with identifying complex patterns in network traffic, our hybrid deep learning framework excelled by capturing intricate temporal and spatial dependencies. This advantage is evident in the significantly higher AUC of 93.82 achieved by our model, compared to other techniques.

The superior performance across datasets demonstrates the generalizability and robustness of our method. Figure 5 and Fig. 6 underscore its capacity to outperform existing approaches consistently, regardless of dataset complexity or domain. The integration of advanced feature extraction, adaptive learning rates, and hybrid modeling techniques provides a clear advantage over traditional and contemporary models, making our approach a compelling choice for diverse prediction and classification tasks.

4.4. Ablation Study

Table 3 and 4 present the results of the ablation study conducted on the Virus-MNIST, NSL-KDD, VxHeaven, and CICIDS2017 datasets. The study aimed to evaluate the contribution of key components in our model by systematically removing specific modules (denoted as w/o Adaptive Learning and Computational Efficiency, w/o Risk Prioritization and Resource Optimization, and w/o Game-Theoretic Threat Suppression) and observing the performance impact. The metrics include Accuracy, Recall, F1 Score, and AUC, offering a comprehensive understanding of the model's robustness.

Table 1. Comparison of Models on Virus-MNIST and NSL-KDD Datasets for Temporal Prediction Task

Model	Virus-MNIST Dataset				NSL-KDD Dataset			
	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
ARIMA Alabdulrazzaq et al. (2021)	86.74	92.32	81.27	86.95	89.26	88.59	75.65	83.33
LSTM Landi et al. (2021)	97.11	87.02	81.30	89.09	78.88	77.01	92.28	97.45
GRU Cahuantzi et al. (2023)	93.18	89.16	83.48	90.13	89.54	72.25	79.92	93.95
Transformer Han et al. (2022)	90.78	75.41	87.45	91.38	81.24	85.74	87.59	96.91
SVM Chandra and Bedi (2021)	82.81	94.40	85.78	95.99	79.24	80.12	80.92	96.11
Decision Tree Costa and Pedreira (2023)	82.81	91.65	83.24	87.80	96.03	72.81	84.88	90.76
Ours	97.11	94.40	89.01	95.99	96.35	88.59	92.28	97.45

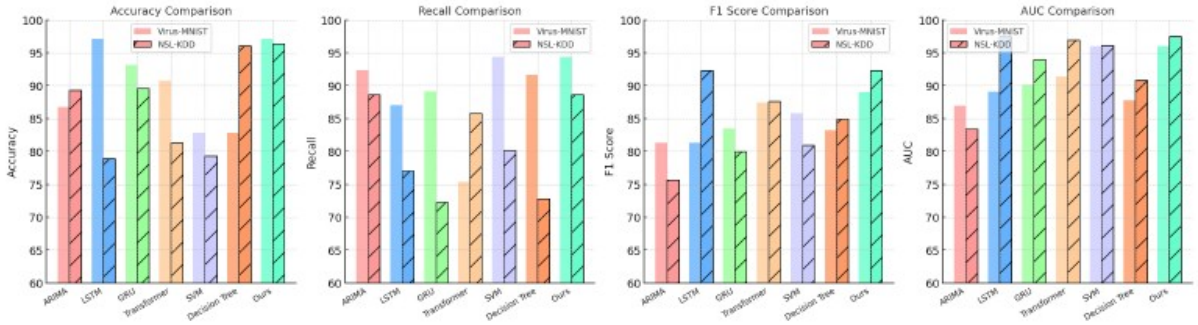

Fig. 5. Performance Comparison of SOTA Methods on Virus-MNIST Dataset and NSL-KDD Dataset Datasets

Table 2. Comparison of Models on VxHeaven and CICIDS2017 Datasets for Temporal Prediction Task

Model	VxHeaven Dataset				CICIDS2017 Dataset			
	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
ARIMA Alabdulrazzaq et al. (2021)	83.24	83.71	89.27	84.11	83.98	88.13	78.33	87.24
LSTM Cahuantzi et al. (2023)	84.74	82.49	82.38	88.38	84.20	88.93	79.40	89.33
GRU Cahuantzi et al. (2023)	82.63	86.68	80.07	84.74	90.67	85.54	78.41	93.82
Transformer Han et al. (2022)	86.55	80.25	89.79	95.95	89.20	89.11	86.27	86.97
SVM Han et al. (2022)	87.44	90.84	88.48	92.35	93.20	84.39	82.09	85.00
Decision Tree Costa and Pedreira (2023)	85.80	79.19	88.75	87.96	86.56	84.89	84.61	87.77
Ours	93.60	93.79	89.79	95.95	93.20	89.11	89.80	93.82

On the Virus-MNIST Dataset, removing Adaptive Learning and Computational Efficiency (w/o Adaptive Learning and Computational Efficiency) resulted in a notable drop in Accuracy from 97.11% to 94.30% and AUC from 95.99 to 89.82. This emphasizes the critical role of Adaptive Learning and Computational Efficiency in extracting intricate malware features. Similarly, the exclusion of Risk Prioritization and Resource Optimization (w/o Risk Prioritization and Resource Optimization) caused significant reductions in Recall and F1 Score, indicating that this module effectively balances false positives and negatives. Game-Theoretic Threat Suppression's exclusion (w/o Game-Theoretic Threat Suppression) also impaired performance, particularly in AUC, highlighting its importance in enhancing the model's generalization capabilities. For the NSL-KDD Dataset, the complete model outperformed all ablated versions, achieving an Accuracy of 96.35% and an F1 Score of 92.28. The exclusion of Adaptive Learning and Computational Efficiency led to substantial performance degradation across all metrics, with Accuracy dropping to 83.08%. This demonstrates Adaptive Learning and Computational Efficiency's role in learning robust features for diverse attack types. The results for w/o Risk Prioritization and Resource Optimization and w/o C further affirm the

synergistic importance of these modules, as their removal resulted in reduced Recall and AUC, reflecting weaker attack detection.

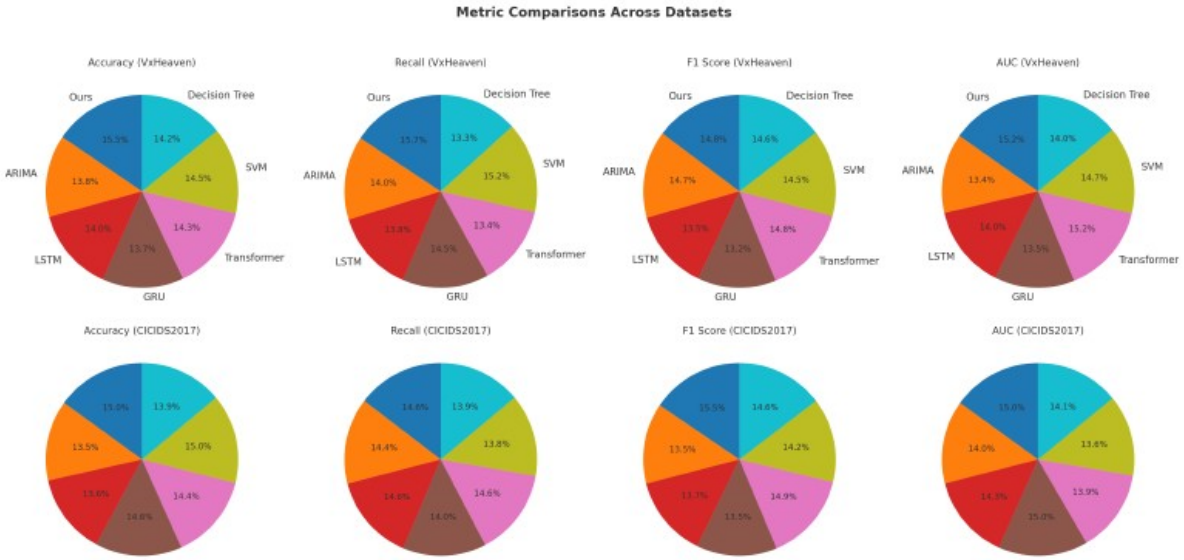

Fig. 6. Performance Comparison of SOTA Methods on VxHeaven Dataset and CICIDS2017 Dataset Datasets

On the VxHeaven Dataset, removing Adaptive Learning and Computational Efficiency (w/o Adaptive Learning and Computational Efficiency) caused a significant drop in Accuracy from 93.60% to 86.13%, and F1 Score decreased to 83.01. This highlights the critical function of Adaptive Learning and Computational Efficiency in identifying subtle differences among malware families. The absence of Risk Prioritization and Resource Optimization (Risk Prioritization and Resource Optimization) resulted in less pronounced, but still notable, performance decreases in F1 Score and AUC, confirming its auxiliary role in refining classification outputs. Game-Theoretic Threat Suppression's exclusion (w/o Game-Theoretic Threat Suppression) led to diminished generalization, as indicated by the lower AUC of 87.14. For the CICIDS2017 Dataset, the full model achieved superior performance with an Accuracy of 93.20% and an AUC of 93.82. In contrast, the removal of Adaptive Learning and Computational Efficiency (w/o Adaptive Learning and Computational Efficiency) yielded an Accuracy of 84.41%, showing its indispensable role in capturing nuanced network traffic patterns. Excluding Risk Prioritization and Resource Optimization (Risk Prioritization and Resource Optimization) led to a balanced but lower performance across all metrics, suggesting its supportive contribution to the overall architecture. Finally, removing Game-Theoretic Threat Suppression (w/o Game- Theoretic Threat Suppression) impacted Recall and F1 Score significantly, indicating its critical role in identifying sequential dependencies in network traffic.

The results of the ablation study clearly demonstrate the importance of each component within our proposed model. Figure 7 and Figure 8 highlight the performance trade-offs when components are excluded, reinforcing the necessity of their inclusion for achieving state-of-the-art results across diverse datasets. This systematic analysis validates the architectural design and its adaptability across different domains.

Table 3. Ablation Study Results on Virus-MNIST and NSL-KDD Datasets for Temporal Prediction Task

Component	Virus-MNIST Dataset				NSL-KDD Dataset			
	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
w/o Adaptive Learning and Computational Efficiency	94.30	89.64	88.65	89.82	83.08	79.44	84.15	92.70
w/o Risk Prioritization and Resource Optimization	93.08	82.24	89.27	91.42	88.62	82.05	85.62	87.01
w/o Game-Theoretic Threat Suppression	91.33	90.71	85.23	94.54	87.59	89.31	82.91	88.98
Ours	97.11	94.40	89.01	95.99	96.35	88.59	92.28	97.45

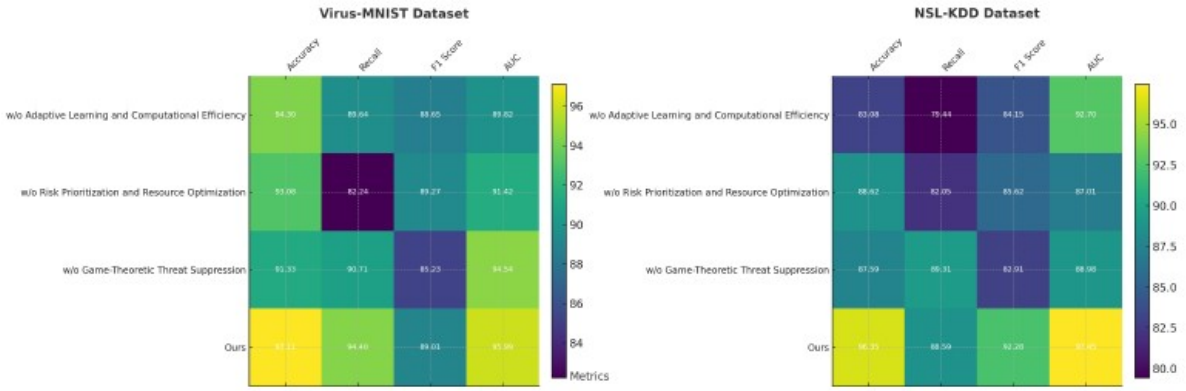

Fig. 7. Ablation Study of Our Method on Virus-MNIST Dataset and NSL-KDD Dataset Datasets

Table 4. Ablation Study Results on VxHeaven and CICIDS2017 Datasets for Temporal Prediction Task

Component	VxHeaven Dataset				CICIDS2017 Dataset			
	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
w/o Adaptive Learning and Computational Efficiency	86.13	80.62	83.01	91.05	84.41	79.08	81.29	89.58
w/o Risk Prioritization and Resource Optimization	83.41	83.34	85.43	91.85	87.59	88.02	85.14	87.59
w/o Game-Theoretic Threat Suppression	89.71	90.90	88.90	87.14	87.60	81.85	85.74	90.52
Ours	93.60	93.79	89.79	95.95	93.20	89.11	89.80	93.82

5. Conclusions and Future Work

Information Security Risk Assessment Method based on digital watermarking and copy detectionAll the files uploaded by the user have been fully loaded. Searching won't provide additional information.This study addresses the critical challenge of effectively assessing information security risks in increasingly complex digital environments. Traditional ISRA methods often fall short in accounting for the dynamic and multifaceted nature of modern cybersecurity threats, leading to inefficient resource allocation and insufficient risk mitigation. To bridge these gaps, we introduce an innovative framework that integrates digital watermarking with adaptive copy detection mechanisms. This approach enhances data integrity and traceability, leveraging techniques such as probabilistic modeling, Bayesian inference, and game- theoretic strategies through the Adaptive Risk Evaluation Network (AREN) and the Strategic Risk Mitigation Framework (SRMF). The experimental results confirm significant advancements in predictive accuracy, adaptability, and resource

efficiency over traditional models, underscoring the practical utility and theoretical contribution of this methodology in optimizing risk assessment processes.

However, the proposed method is not without limitations. First, while the experimental results validate its effectiveness, the computational complexity associated with Bayesian inference and game-theoretic calculations may pose scalability issues for large-scale applications. Optimizing the framework for high- performance environments is a necessary future step. Second, the reliance on digital watermarking assumes robust embedding and extraction processes, which may not always hold under adversarial conditions. Enhancing resilience against such scenarios will be critical for broader adoption. Future research should explore lightweight algorithms to address scalability concerns and develop advanced mechanisms to safeguard watermark integrity under adversarial attacks, ensuring that the framework remains viable across diverse operational contexts.

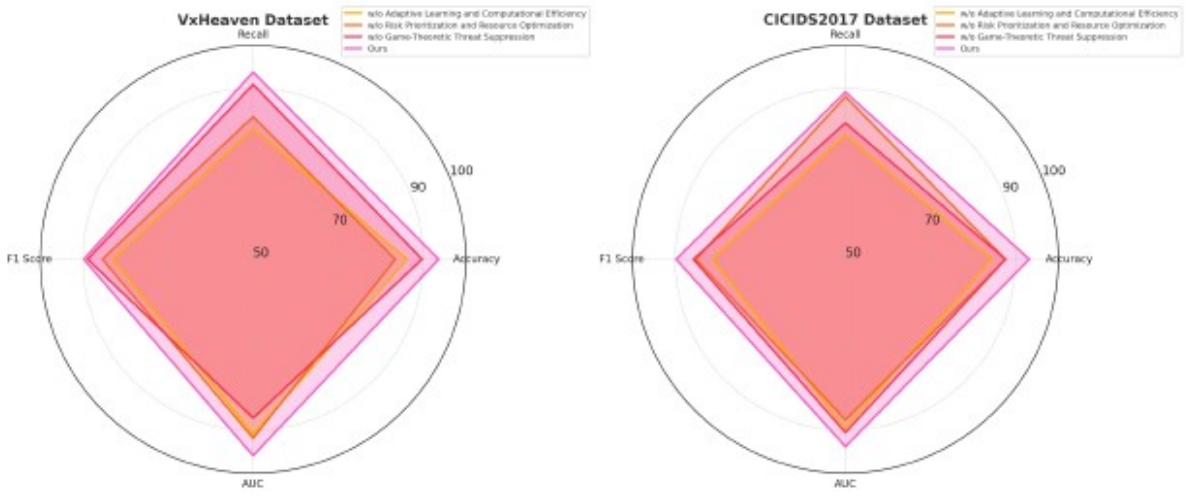

Fig. 8. Ablation Study of Our Method on VxHeaven Dataset and CICIDS2017 Dataset Datasets

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The research is supported by Anhui Natural Science Research Project of Colleges and Universities: Research on the Comparison of Modern Film and Television Camera Language Contained in Ancient Chinese Poetry Based on VR+5G (No. KJ2020A0933); Hefei Technology College high-level introduction of talents scientific research start-up fund project: Research on the Application of Data Mining in Student Performance Analysis (No.2021KYQDZ011); Key Project of Quality Engineering Teaching Research in Higher Education Institutions of Anhui: Construction of a blended teaching evaluation system based on the entire process: Taking Hefei Vocational and Technical College as an example (No.2022jyxm1367).

References

[1] Abdygalievich, A. S., Muratbekov, M., Altynbek, S., and Barlybayev, A. (2021). Fuzzy expert system of information security risk assessment on the example of analysis learning management systems. *IEEE Access*

[2] Abrar, I., Ayub, Z., Masoodi, F., and Bamhdi, A. M. (2020). A machine learning approach for intrusion detection system on nsl-kdd dataset. In 2020 international conference on smart electronics and *communication (ICOSEC)* (IEEE), 919—924

[3] Alabdulrazzaq, H., Alenezi, M. N., Rawajfih, Y., Alghannam, B. A., Al-Hassan, A. A., and Al-Anzi, F. S. (2021). On the accuracy of arima based prediction of covid-19 spread. *Results in Physics* 27, 104509

[4] Bakhtina, M. and Matulevicius, R. (2022). Information security risks analysis and assessment in the passenger-autonomous vehicle interaction. *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*

[5] Bhosale, P., Kastner,W., and Sauter, T. (2022). Automating safety and security risk assessment in industrial control systems: Challenges and constraints. *IEEE International Conference on Emerging Technologies and Factory Automation*

[6] Bruma, L. M. (2020). An approach for information security risk assessment in cloud environments. *Informatica economica*

[7] Cahuantzi, R., Chen, X., and Guttel, S. (2023). A comparison of lstm and gru networks for learning symbolic sequences. In *Science and Information Conference* (Springer), 771-785

[8] Chandra, M. A. and Bedi, S. (2021). Survey on svm and their application in image classification. *International Journal ofInformation Technology* 13, 1-11

[9] Chandra, N. A.,Ramli,K.,Ratna,A. A. P., and Gunawan, T. S. (2022). Information security riskassessment using situational awareness frameworks and application tools. *Risks*

[10] Chen, Y., Xu, B., and Long, J. (2021). Information security assessment of wireless sensor networks based on bayesian attack graphs. *Journal ofIntelligent & Fuzzy Systems*

[11] Costa, V. G. and Pedreira, C. E. (2023). Recent advances in decision trees: An updated survey. *Artificial Intelligence Review* 56, 4765-4800

[12] Deb, R. and Roy, S. (2021). A software defined network information security risk assessment based on pythagorean fuzzy sets. *Expert systems with applications*

[13] Engelen, G., Rimmer, V., and Joosen, W. (2021). Troubleshooting an intrusion detection dataset: the cicids2017 case study. In *2021 IEEE Security and Privacy Workshops (SPW)* (IEEE), 7-12

[14] Faizi, A., Padyab, A., and Naess, A. (2021). From rationale to lessons learned in the cloud information security risk assessment: a study of organizations in sweden. *Information and Computer Security*

[15] Govender, S. G., Kritzinger, E., and Loock, M. (2021). A framework and tool for the assessment of information security risk, the reduction of information security cost and the sustainability of information security culture. *Personal and Ubiquitous Computing*

[16] Gozhyj, A., Kalinina, I., Vysotska, V., Sachenko, S., and Kovalchuk, R. (2020). Qualitative and quantitative characteristics analysis for information security risk assessment in e-commerce systems. *International Workshop on Information-Communication Technologies & Embedded Systems*

[17] Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., et al. (2022). A survey on vision transformer. *IEEE transactions on pattern analysis and machine intelligence* 45, 87-110

[18] Hayajneh, A. A., Thakur, H. N., and Thakur, K. (2023). The evolution of information security strategies: A comprehensive investigation of infosec risk assessment in the contemporary information era. *Computer and Information Science*

[19] He, T. and Li, Z. (2021). A model and method of information system security risk assessment based on mitre att&ck. *2021 2nd International Conference on Electronics, Communications and Information Technology (CECIT)*

[20] Huang, F., Yong, M., and Jian, Z. (2021). Genetic algorithm-based power system information security risk assessment method. *Journal of Physics: Conference Series*

[21] Hui, P. (2020). Construction of information security risk assessment model in smart city. *2020 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS)*

[22] Kerimkhulle, S., Dildebayeva, Z., Tokhmetov, A., Amirova, A., Tussupov, J., Makhazhanova, U., et al. (2023). Fuzzy logic and its application in the assessment of information security risk of industrial internet of things. *Symmetry*

[23] Kitsios, F., Chatzidimitriou, E., and Kamariotou, M. (2022). Developing a risk analysis strategy framework for impact assessment in information security management systems: A case study in it consulting industry. *Sustainability*

[24] Kuzminykh, I., Ghita, B., Sokolov, V., and Bakhshi, T. (2021). Information security risk assessment. *Encyclopedia*

[25] Landi, F., Baraldi, L., Cornia, M., and Cucchiara, R. (2021). Working memory connections for lstm. *Neural Networks* 144, 334—341

[26] Landoll, D. J. (2021). Information security risk assessment basics. *The Security Risk Assessment Handbook*

[27] Lee, Y., Woo, S., Song, Y., Lee, J., and Lee, D. H. (2020). Practical vulnerability-information-sharing architecture for automotive security-risk analysis. *IEEE Access*

[28] Palko, D., Babenko, T., Bigdan, A., Kiktev, N., Hutsol, T., Kubon, M., et al. (2023). Cyber security risk modeling in distributed information systems. *Applied Sciences*

[29] Palko, D., Myrutenko, L., Babenko, T., and Bigdan, A. (2020). Model of information security critical incident risk assessment. *International Scientific-Practical Conference Problems of Infocommunications Science and Technology*

[30] Schmitz, C. and Pape, S. (2020). Lisra: Lightweight security risk assessment for decision support in information security. *Computers & security*

[31] Sharma, K., Shankar, A., and Singh, P. (2021). Information security assessment in big data environment using fuzzy logic. *Journal of cybersecurity and information management*

[32] Shaukat, K., Luo, S., and Varadharajan, V. (2022). A novel method for improving the robustness of deep learning-based malware detectors against adversarial attacks. *Engineering Applications of Artificial Intelligence* 116, 105461

[33] Shirvani, S., Baseri, Y., and Ghorbani, A. A. (2024). Evaluation framework for electric vehicle security risk assessment. *IEEE transactions on intelligent transportation systems (Print)*

[34] Sukumar, A., Mahdiraji, H. A., and Jafari-Sadeghi, V. (2023). Cyber risk assessment in small and medium-sized enterprises: A multilevel decision-making approach for small e-tailors. *Risk Analysis*

[35] Ushakov, R., Doynikova, E., Novikova, E., and Kotenko, I. (2021). Cpe and cve based technique for software security risk assessment. *International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*

[36] Yan, K., Liu, X., Lu, Y., and Qin, F. (2023). A cyber-physical power system risk assessment model against cyberattacks. *IEEE Systems Journal*

[37] Zhang, T., Zhao, K., Yang, M., Gao, T., and Xie, W. (2020). Research on privacy security risk assessment method of mobile commerce based on information entropy and markov. *Wireless Communications and Mobile Computing*

[38] Ziemba, P., Becker, A., and Becker, J. (2021). Forecasting and assessment of the energy security risk in fuzzy environment. *Energies*

[39] Zou, B., Cao, C., Wang, L., Fu, S., Qiao, T., and Sun, J. (2024). Facile: A capsule network with fewer capsules and richer hierarchical information for malware image classification. *Computers & Security* 137, 103606