IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

An Intellectual Zero Trust Security Framework Using Deep Reinforcement Learning for Predictive Threat Mitigation in AI-Based Fraud Detection Systems

Ankur Mahida

Site Reliability Engineer Barclays, 400 Jefferson Park, Whippany, New Jersey, 07981

Corresponding author: Ankur Mahida (e-mail: ahmahida13@gmail.com)

ABSTRACT The rapid growth of Artificial Intelligence (AI) driven financial platforms has led to an increase in sophisticated cyberattacks and fraudulent activities that easily bypass traditional security controls. Current fraud detection mechanisms mostly utilize fixed rule-based engines, monitored classifiers, based on signature, or high-dimensional decision space and all of them find it difficult to cope with changing threat patterns, zero-day attacks, and high-dimensional decision spaces. Even though the recent research has covered machine learning and deep learning architectures, the solutions remain to have significant limitations, such as slow adaptation to novel threats, lack of proactive defense, high rates of false-positives, and incapacity to implement dynamic trust assessment in real-time AI-based fraud detection systems. To counter such limitations, this research proposes an Intellectual Zero Trust Security Framework that is based on Deep Reinforcement Learning (DRL) to mitigate threats in predictive mode in an attempt to detect fraud in AI-based systems. The suggested framework would incorporate a policy-optimization DRL agent with an active trust-evaluation mechanism that would monitor dynamically user behaviors, system interaction, and new anomaly patterns. Incorporating state-state modeling, reward-shaping, and incremental threat-scoring, the DRL-Zero Trust agent is able to learn the best security actions which include access restriction, adaptive authentication, and anomaly-suppression without any predefined rules. Evaluation of the proposed DRL-Zero Trust model on a benchmark dataset of fraud detection shows that it significantly outperforms classical classifiers, including K Nearest Neighbor (KNN), Random Forest, Logistic Regression (LR) and Support Vector Machine (SVM). The proposed model achieved 98.7% accuracy, 98.4% precision, 98.9% recall and an Area Under the Receiver Operating Characteristic Curve (AUC-ROC) of 0.995, and is better resistant to zero-day attacks and adversarial instances. These findings verify that the suggested framework presents a very flexible, intelligent, and proactive defense system that can protect the modern AI-driven financial systems.

INDEX TERMS Deep Reinforcement Learning, Zero Trust Security, Fraud Detection, Threat Mitigation, Predictive Modeling.

I. INTRODUCTION

The blistering development of smart, software-defined, and hyper-networked digital surroundings has greatly complicated the process of identifying insurmountable cybersecurity [1]. Cyber-attacks including Distributed Denial of Service (DDoS), phishing, reconnaissance, spoofing, and malware propagation have now become common targets of a wide range of sophisticated cyber-attacks on modern network infrastructures, especially Software Defined Networks (SDN), Internet of Things (IoT) environments, and AI-enabled applications [2]. The classical Intrusion Detection Systems (IDS) started out explicitly to track malicious activity and block data leaks but they fail to cope with the volume, variety, and flexibility of the current threats [3]. Although cyber-attacks keep changing, prediction, mitigation, and resilience have turned out to become critical features of a security system as opposed to optional features despite the common use of machine-learning-based IDS systems [4].

Majority of the traditional systems use fixed sets of rules, pre-learned patterns, or machine learning models, which cannot respond to dynamic and fast changing attack patterns in real-time [5]. Most of them, such as detection models based on SVM, Random Forest classification models, and behavior-based anomaly detectors, are highly accurate on a benchmark dataset, but perform at a much lower level in a real-world network environment [6]. In addition, the datasets used to construct the models like NSL-KDD, KDD-Cup 99, and UNSW-NB15 contain only approximated patterns of traffic between IoT and SDN, thus limiting generalizability and practical implementation. Both Convolution Neural Networks (CNNs) and LSTMs are deep learning-based systems that enhance the capabilities of classification but are characterized by computational cost, training instability, and inadequate ability to adapt to novel patterns of attacks [6]. Hybrid techniques such as SVM-RF and Deep Belief Network-Kernel Extreme Learning Machine (DBN-KELM) are more accurate in detection, but still based on a traditional learning concept, so they cannot be applied in a dynamic or adversarial evolving cyber-threat environment. The general ZTA model is shown in Figure 1.

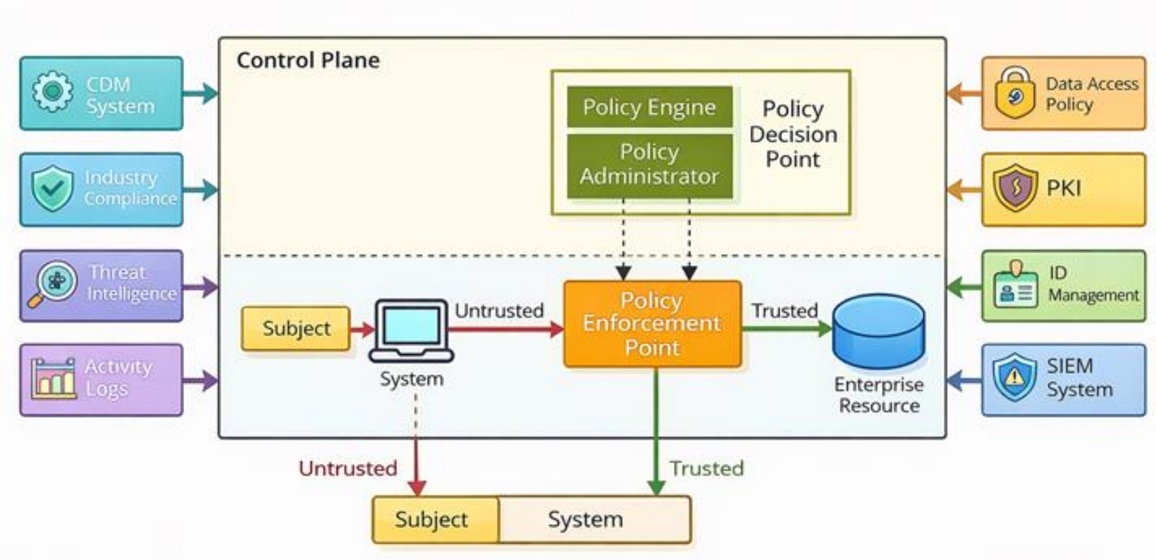

Figure 1.General ZTA Model

Moreover, the majority of the available models do not have smart risk-based decision-making [7]. They are able to identify the threats but not to independently decide the best mitigation options, neither are they able to assess security implications of measures in the long term. This is a fatal weakness with SDN-enabled networks, in which centralized controllers are desirable targets in DDoS and control-plane attacks [8]. To overcome these limitations, this research suggests an Intellectual Zero-Trust Security Framework that runs Deep Reinforcement Learning to Predictive Threat Mitigation by applying Intellectual Fraud Detection and SDN-enabled networks. The suggested model combines the principles of the Zero Trust Architecture (ZTA) the meaning of Never trust, always verify with the features of the intelligent decision-making of DRL in order to establish a continuously adaptive defence system. The DRL agent does not use any fixed rules of detection or pre-trained classifier but rather develops optimal threat-response behaviors in real time by interacting with the network environment. The framework will further empower SDN controllers with predictive intelligence, which will allow detecting, isolating and mitigating the intrusions before they become full-scale attacks [9]. The system is more resilient, has lower false positives, and mitigates more quickly to the system when anomaly detection is coupled with the use of reinforcement-learning-based policy optimization in a highly dynamic network environment [10].

The proposed model drives decisions on the basis of AI-based fraud detection systems to constantly learn how to defend itself in a dynamic and adversarial environment. As opposed to the conventional machine learning models which use fixed decision boundaries and static training, DRL presents fraud detection and security enforcement as a sequence of decisions. Every state is considered to be a transaction or a user interaction, and the system has to determine an optimal security measure, including permitting access, further authentication, limiting transactions, or preventing suspicious activities, depending on the real-time risk and trust decisions. Combining DRA and Zero Trust principles in the framework, all parties will be treated without presuppositions of trust, and any interaction will be reconsidered in the future.

The agent-environment interaction paradigm applies to the working of DRL within this framework. DRA agent monitors the environment by inputting user behavioral patterns, transaction history, historical fraud detection, and trust score dynamism [11]. The system state is comprised of these observations. Depending on this condition, an agent will choose an action depending on a learned policy to reduce the risk of fraud, with the goal of preserving valid user experience [12]. Following the action, the environment will send feedback in terms of rewards or penalties, respectively, based on whether the action was effective in mitigating fraud or with false positives [13]. Through rewards-based learning over time, the agent is able to update its policy and therefore predict changes in attack strategies, adapt to zero-day fraud cycles and optimize security responses without hard rules or signatures [14].

DRL is essential in implementing the principles of Zero Trust because it allows continuous authentication and adapting authorization. The DRA agent does not decide the access once, but rather checks trust at each interaction point. The trust score is dynamically updated as user behavior shifts or anomalies appear and the agent takes the initiative of modifying security actions. Such state consciousness in learning enables the framework to transition to predictive threat reduction instead of reactive fraud detection, where the possible attacks are detected before they cause harm. The fact that DRA is able to model long-term rewards makes sure the system is able to balance the strengths of security and the efficiency of the system. The Reinforcement Learning Process is shown in Figure 2.

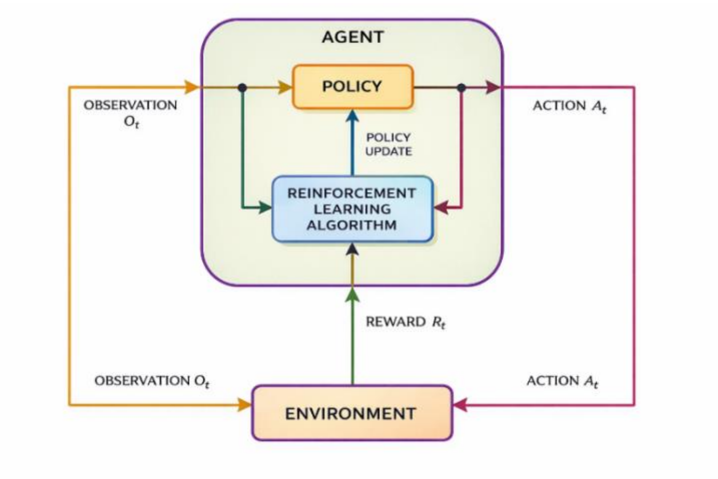

Figure 2.Reinforcement Learning Process

DRL approach has a number of major benefits compared to conventional and supervised learning based fraud detection methods. First, DRL offers real-time flexibility such that the system can react well to new and unknown attack patterns. Second, it has a considerable decreasing reliance on labeled data and manual updates to rules, which is appropriate when dealing with a fast-changing financial fraud problem. Third, reward-based learning mechanism allows reducing false positives by learning optimal trade-offs between security enforcement and user convenience. Lastly, DRA boosts resilience to adversarial and zero-day attacks by updating its policy through interaction, which makes the suggested framework intelligent, self-learning and proactive in protecting AI-based financial systems.

1.1 Research Objectives

The research objectives are

1. To create the intrusion detection architecture based on Zero-Trust that facilitates continuous verification and real-time observation and dynamic threat assessment in SDN-powered networks.
2. To create Deep Reinforcement Learning model that will predict and address emerging cyber-attacks by choosing optimal security actions autonomously.
3. To combine machine learning-based fraud detection with programmable control plane of SDN to offer adaptive and real-time fraud detection and threat response.
4. To compare the suggested framework with the current machine-learning-based IDS models and show high accuracy rates, flexibility, and low false alarm levels.

1.2 Hypothesis

This research is founded on the hypothesis that a Zero Trust Security Architecture empowered by Deep Reinforcement Learning can significantly enhance predictive threat mitigation and fraud detection performance in AI-based systems compared to traditional rule-based and supervised learning approaches.

Specifically, the research hypothesizes that:

- **H1:** A DRL-driven Zero Trust framework can dynamically learn optimal security policies without predefined rules, enabling proactive and real-time mitigation of evolving and zero-day fraud threats.
- **H2:** Continuous trust evaluation using behavioral state modeling and reward-based policy optimization will reduce false-positive and false-negative rates compared to classical machine learning and deep learning classifiers.
- **H3:** Integrating DRL-based decision-making with SDN-enabled control planes enables faster and more resilient response to adversarial and large-scale attack scenarios.
- **H4:** The proposed DRL-Zero Trust model will demonstrate superior generalization capability and robustness against adversarial manipulation and unseen attack patterns when evaluated on benchmark fraud detection datasets.

1.3 Research Contributions

This paper makes the following key contributions to the fields of cybersecurity, fraud detection, and intelligent network defense:

1. **Novel DRL-Based Zero Trust Framework:** It suggests a new Intellectual Zero Trust Security Framework that will combine Deep Reinforcement Learning with ongoing trust-based evaluation of predictive threat mitigation in AI-based fraud detection systems.
2. **Autonomous Predictive Threat Mitigation:** The proposed DRL agent autonomously finds the optimal mitigation measures like access restriction, adaptive authentication, and anomaly suppression to threats, contrary to traditional IDS and fraud detection models, which only signal the existence of a threat based on long-term reward maximization.
3. **Dynamic Trust and Behavior Modeling:** The framework presents state-state modeling, the incremental threat scoring and reward shaping mechanisms that enable real-time reassessment of trust depending on the behavior of the user, the interactions with the system and changing patterns of anomalies.
4. **Enhanced Robustness Against Zero-Day and Adversarial Attacks:** The system removes the need to rely on fixed rules or fixed classifiers and thus shows greater resilience to zero-day threats and adversarial examples, which is a significant weakness of the current supervised learning-based security models.
5. **Comprehensive Performance Evaluation:** Extensive experimental evaluation on benchmark fraud detection datasets shows that the proposed model significantly outperforms traditional classifiers such as SVM, Random Forest and LSTM in terms of accuracy, precision, recall, and AUC-ROC.

II.RELATED WORK

Machine learning algorithms for intrusion detection have been extensively explored and obtained by most researchers [15]. This method is used to detect various types of attacks that can jeopardize networks, such as phishing, Denial of Service (DoS), and DDOS attacks. In most cases, supervised learning is used by IDS to classify network traffic as either normal or intrusion-related. In order to reliably anticipate familiar invasion types, supervised learning systems learn from deployed data that exhibits comparable patterns [16]. SVM and RF are the two most used supervised learning algorithms [17]. The main applications of RF are in the areas of regression and classification, while SVM is better suited to classification tasks [18]. The generalizability of SVM is higher than that of other machine learning algorithms. In contrast, a huge number of inputs are routinely fed into unsupervised learning without labels [19]. One study was able to glean useful information by employing ML-based unsupervised learning models [20]. To make the numbers stand out, they used a set of completely arbitrary variables [21]. Unsupervised learning relies on class labels to extract useful information from incoming data sets [22].

Cybercriminals launch DDoS attacks when they want to overwhelm a specific server, service, or network with so much traffic that it becomes unusable or its underlying infrastructure becomes overwhelmed. Attacks from DDoS systems can compromise network security in a major way. Classifying DDoS attacks using ML and DL has been the subject of multiple publications. The forms of DDoS attacks are always evolving, though, so there's a need for better and further research. Jiyad et al. [23] introduced a new ensemble model that can detect distributed denial of service assaults. By utilizing ML techniques including LR, RF, DT, and XGBoost classifiers, our approach is able to detect and categorize these harmful attacks with great effectiveness. In order to improve the suggested model's performance and decrease overfitting, the hyper-tuning process is crucial.

Manageability, scalability, and enhanced performance are just a few of the benefits of SDN. On the other hand, if the controller isn't able to withstand DDoS assaults, then SDN poses unique security risks. DDoS attacks overwhelm the SDN controller's processing and communication capabilities. Hence, the network performance drops to a critical level when the capacity of the switch flow table is filled up due to the needless flow provided by the controller for the attack packets. Polat et al. [24] used models based on machine learning to identify DDoS assaults in SDN. To begin, the dataset was subjected to both regular and DDoS attack traffic in order to extract certain properties from the SDN. Then, feature selection methods were applied to the existing dataset in order to build a new dataset. In order to reduce training time, simplify the models, and make them easier to understand, feature selection strategies were chosen.

The most hazardous assault in the realm of network security is the DDoS attack. Critical services of many online applications are rendered inoperable during a distributed denial of service assault. Instead of serving actual users, systems that are under DDoS assaults are preoccupied with bots who send out fake requests. The frequency and sophistication of these attacks are rising daily. As a result, protecting internet services from these assaults has gotten more challenging. In this paper, Saini et al. [25] presented a method for detecting and classifying various types of network traffic flows using a machine learning based technique. A fresh dataset containing a combination of different modern attack methods, including HTTP flood, SID DoS, and normal traffic, is used to validate the suggested technique. For the purpose of attack classification, a machine learning tool known as WEKA is employed.

In order to avoid network breaches, security analysts and administrators must overcome numerous obstacles, the most important of which is the need to identify intrusions in a timely manner. Identifying unexpected strikes might be difficult. In this study, Rawat et al. [26] compared the NSL-KDD dataset performance of traditional ML methods that rely on extensive feature engineering to that of more modern methods that combine deep neural networks with integrated unsupervised feature learning. Finding the right hyperparameters and network settings for machine learning models required a battery of experiments. Optimal modeling was achieved by a DNN trained with fifteen features obtained through Principal Component Analysis (PCA).

For decades, there has been no reliable way to protect networks from DDoS attacks, which have severely reduced their availability. But new SDN technology offers a fresh perspective on DDoS assault defense. Two approaches to DDoS attack detection in SDN are presented in this research by Dong et al. [27]. One approach uses the severity of the DDoS attack as an identifier. The alternative strategy for detecting DDoS attacks makes use of the ML-based improved KNN algorithm. Both theoretical and experimental findings on datasets demonstrate that our suggested approaches outperform competing methods when it comes to detecting DDoS attacks.

Modern networking has been radically altered by SDNs, which provide scalability and dynamic reconfigurability. Nevertheless, both conventional and SDN-based networks still face the enormous problem of defending against DDoS assaults. One promising approach to dealing with these dangers is the incorporation of Machine Learning (ML) into SDN. Despite ML's success in identifying malicious from benign traffic, it struggles to deal with new, low-rate, zero-day DDoS attacks because of its limited feature scope for training. Constant retraining of ML models is required due to the dynamic nature of the DDoS landscape, which is influenced by new protocols. Alashhab et al. [28] offered an ensemble online machine-learning model to address these issues and improve DDoS mitigation and detection. This method adjusts the model to anticipate assault patterns by means of online learning. Using SDN simulation, the model is trained and evaluated (Mininet and Ryu). Improved accuracy across varied forms of DDoS attacks is a consequence of its dynamic feature selection capability, which overcomes conventional restrictions.

There are new threats to network security brought about by 5G's massive improvements in network speed and connection density, which are becoming more common as the technology spreads. Specifically, in SDN settings, DDoS attacks have grown in both frequency and complexity. Due to their diversity and complexity, 5G networks produce a plethora of superfluous features. These features can muddy an IDS detection process and weaken the model's capacity for generalization. Improving the IDS performance in 5G networks is the primary goal of this work performed by Han et al. [29]. In order to make the IDS more resilient and effective in its detection, it suggests a new feature selection (FS) method for extracting the most representative and unique characteristics from data on network traffic.

One threat to the safety of SDNs is the DDoS assault. Among the many problems with current DDoS detection systems are their reliance on network topology, their inability to identify all DDoS attacks, the fact that they use inaccurate and out-of-date datasets, and the expensive and powerful hardware infrastructure requirements. Their reliance on historical data and the use of static thresholds makes them less adaptable to new threats and lengthens the time it takes to identify them. When it comes to SDN, a new way has emerged for detecting DDoS attacks. Parts one, two, and three of the approach designed by Banitalebi Dehkordi et al. [30] are the collector and entropy-based categorization parts. The Table 1 represents the advantages and limitations of traditional models.

Table 1: Advantages and Limitations of Traditional Models

Author (s)	**Proposed Model (as reported)**	**Dataset Used (as reported / typical)**	**Advantages (as reported / inferred)**	**Evaluation Metrics (reported)**	**Limitations (reported / inferred)**
Segura, et al. [1]	Centralized and distributed IDS architectures tailored for resource-constra	Paper evaluates on SDN testbeds / simulated wireless SDN traces (Mininet-based	Low-overhead local detection for constrained nodes; centralized	Detection rate, false positive rate, communication overhead, computational cost.	Evaluation mostly in emulation; limited large-scale real-world trace validatio

	ined wireless SDN; hybrid design that allows local (lightweight) detection plus centralized correlation.	emulation described).	correlation improves detection accuracy without overwhelming edge devices.		n; specifics of datasets and reproducibility require additional publicly-available traces.
Zaher, M et al. [2]	**Sieve** — a distributed flow scheduling framework for SDN data centers that samples & schedules flows to improve flow completion time and detect anomalies.	Data-center flow traces / Mininet + POX emulation used for experiments.	Improves flow completion time for mice flows; lightweight sampling reduces controller load; better QoS under bursty load.	Flow completion time (FCT), throughput, sampling overhead, controller CPU utilization.	Focused on scheduling and QoS rather than explicit security tasks; anomaly detection aspects are limited to sampled flows and may miss stealthy attacks.
Becerra-Suarez, F.L. et al. [3]	Comparative ML pipeline for DDoS detection: RF, DT, AdaBoost, XGBoost, MLP, DNN with data-processing/feature selection.	CICDDoS2019 (preprocessed, 22 selected features reported).	Thorough comparison showing which classical and ensemble methods perform best on the CICDDoS2019 dataset; clear preprocessing steps.	Accuracy, precision, recall, F1-score, AUC.	Uses a single public dataset; limited study of online/real-time mitigation and operational metrics (latency/throughput).
Ramzan, M. et al. [4]	Deep learning models (RNN, LSTM, GRU) for DDoS detection; architecture comparisons and complexity analysis.	CICDDoS2019 and CICIDS2017 comparative experiments.	Good detection accuracy with sequence models; demonstrates tradeoffs between accuracy and execution time (GRU faster than LSTM in some tests).	Accuracy, precision, recall, F1, runtime/execution time.	Focus on detection accuracy; limited treatment of mitigation strategy or integration with SDN control plane for active response.
Wei, Y. et al. [5]	**AE-MLP:** hybrid model combining Autoencoder (AE) for feature extracti	Public DDoS datasets (authors report experiments on standard DDoS/IDS datasets – e.g.,	AE reduces dimensionality and noise; the hybrid yields	Accuracy, precision, recall, F1, AUC; reconstruction loss for the AE.	Autoencoder pretraining can add computational overhead; effectiveness depends

	on and a multilayer perceptron classifier for DDoS detection.	CICIDS types).	improved classification vs standalone classifiers; reduces need for manual feature engineering.		on representativeness of training data; may be less suitable for strict real-time constraints.
Eliyan, L. F. [6]	Survey paper: systematization of DoS/DDoS solutions in SDN; taxonomy of defenses (intrinsic vs extrinsic) and challenge analysis.	Survey (multiple datasets and studies reviewed).	Comprehensive taxonomy and research roadmap; highlights where ML/SDN solutions succeed or fail; identifies open problems for real deployments.	Not an experimental paper; metrics discussed across surveyed works (accuracy, overhead, latency).	Survey highlights gaps but does not provide a single experimental framework; no new model evaluation.
Alanazi, F. [7]	Ensemble deep learning model for DDoS mitigation in SDN — combining multiple DL learners to improve robustness.	CICIDS 2017 (flow-based) used in reported experiments.	Ensemble gives very high reported detection accuracy (authors report >99% in some settings); ensemble reduces variance and improves robustness.	Accuracy, precision, recall, F1, AUC; sometimes resource overhead measurements.	Ensembles can be computationally heavy for real-time deployment and increase controller overhead; limited discussion on mitigation latency and operational impact.
Nguyen et al. [8]	Survey: Deep Reinforcement Learning (DRL) applications for cybersecurity (including intrusion detection, adaptive defense).	Survey (multiple use-cases and datasets summarized).	Connects DRL algorithms to cyber-defense tasks; provides guidelines for reward design and environment modeling.	Survey no single experimental evaluation; highlights need for safe exploration and constraints when using DRL in production.	
Guo, X. et al. [9]	**IZTSDN**: Intelligent Zero-Trust Security Framework for SDN —	Mininet emulation (MiniIZTA) with synthetic and benchmark traffic; experiments describe	Combines Zero-Trust continuous verification with deep anomaly detect	Detection accuracy, throughput under attack, ability to isolate malicious	Evaluated in emulation; real-world deployment and scale tests are limited; complexity of CALSeq

	integrates deep learning anomaly detection (CALSeq2Seq) with Zero-Trust principles and Mininet testbed (MiniIZTA).	d in paper.	ion; supports dynamic authorization and micro-segmentation; real-time tracking demonstrated in Mininet.	users (experimental throughput / performance metrics).	2Seq and tuning may affect runtime.
Wang, Z. et al. [10]	DBN-KELM (Deep Belief Network integrated with Kernel-based Extreme Learning Machine) for intrusion detection – improved DBN training and KELM classifier.	Standard IDS datasets (authors report experiments on NSL-KDD/UNSW/realistic datasets as applicable in the paper).	High accuracy reported (authors obtain competitive results vs other NN methods); KELM accelerates supervised fine tuning and reduces BP overhead.	Accuracy, precision, recall, F1.	DBN variants can be resource-intensive; may impose high latency and heavier controller overhead in real-time SDN scenarios. Reproducibility depends on hyperparameter details.

2.1 Problem Statement

In spite of the remarkable progress in the research on cybersecurity, the currently employed fraud detection and IDS are pretty insufficient to safeguard the modern AI-based and software-defined financial systems. An exhaustive survey of the literature has demonstrated that the majority of the implemented security systems are based on the use of either the static rule-based systems, signature-based detection, or the supervised machine learning classifiers, which are essentially reactive and cannot be used to adapt to the fast-evolving threat environments [31]. The rule-based and signature-based systems are performed effectively only in the case of known attack patterns and incapable of detecting zero-day attacks and polymorphic fraud actions. New researches have delved into deep learning models that include CNNs, LSTM networks, and hybrid models to enhance the quality of detection. In as much as such models intricate temporal and spatial patterns, the literature documents high computational costs, slowness in detecting new attacks, inability to provide explanations, and sensitivity to adversarial examples. More to the point, these methods also consider fraud detection as a fixed classification task and fail to offer any means of independent decision-making or active threat reduction [32].

Another gap that has been determined by current studies is the lack of dynamic trust assessment and risk-accepting response methods. The majority of the intrusion detection and fraud detection systems are only looking at methods of identifying malicious traffic without continually, or randomly, evaluating of trust and deciding on the best course of action to counter-measures. This is a major constraint in Zero Trust and SDN-enabled systems, where centralized controllers and AI-based decision systems are a valuable target of attack and demand real-time and adaptive defense mechanisms. Despite the fact that ZTA has been suggested as a security paradigm that focuses on constant validation and access based on the least privilege, the existing implementations mostly rely on predetermined policies and fixed risk levels. The literature does not have intelligent mechanisms that are capable of learning the best security policies and adjust them to the changing threats.

In addition, currently available evaluation studies are mostly based on old or artificial datasets that are not as representative of real-life dynamics of fraud and attack as they are. This brings up the issue of scalability, stability, and feasibility of the suggested models to be used in working AI-based financial systems. Thus, the main issue that has been observed in the literature is the absence of a smart, versatile, and forecasting security model which combines persistent trust evaluation with self-governing threat mitigation potential. To seal this gap, it is necessary to transition a paradigm shift away given the current models of detection to the Deep Learning-based models of Zero Trust architecture that can learn the best security behavior in real time and counter actions of adversaries and actively respond to new frauds.

III.PROPOSED FRAMEWORK

The proposed methodology involves the development of an Intelligent Zero Trust Security Framework where Machine Learning algorithm used to train the model is somewhat supervised while anomaly detection takes place, and the predictive mitigation engine is based on the Deep Reinforcement Learning algorithm. The whole system is used over an SDN enabled environment where the controller works as a intermediatory for heightening the flows, evaluating trust policies and enforcing security measures. The objective of the methodology is to change traditional reactive intrusion detection to proactive security model that is able to anticipate and mitigate attacks before they have severe impact. The proposed model architecture is shown in Figure 3.

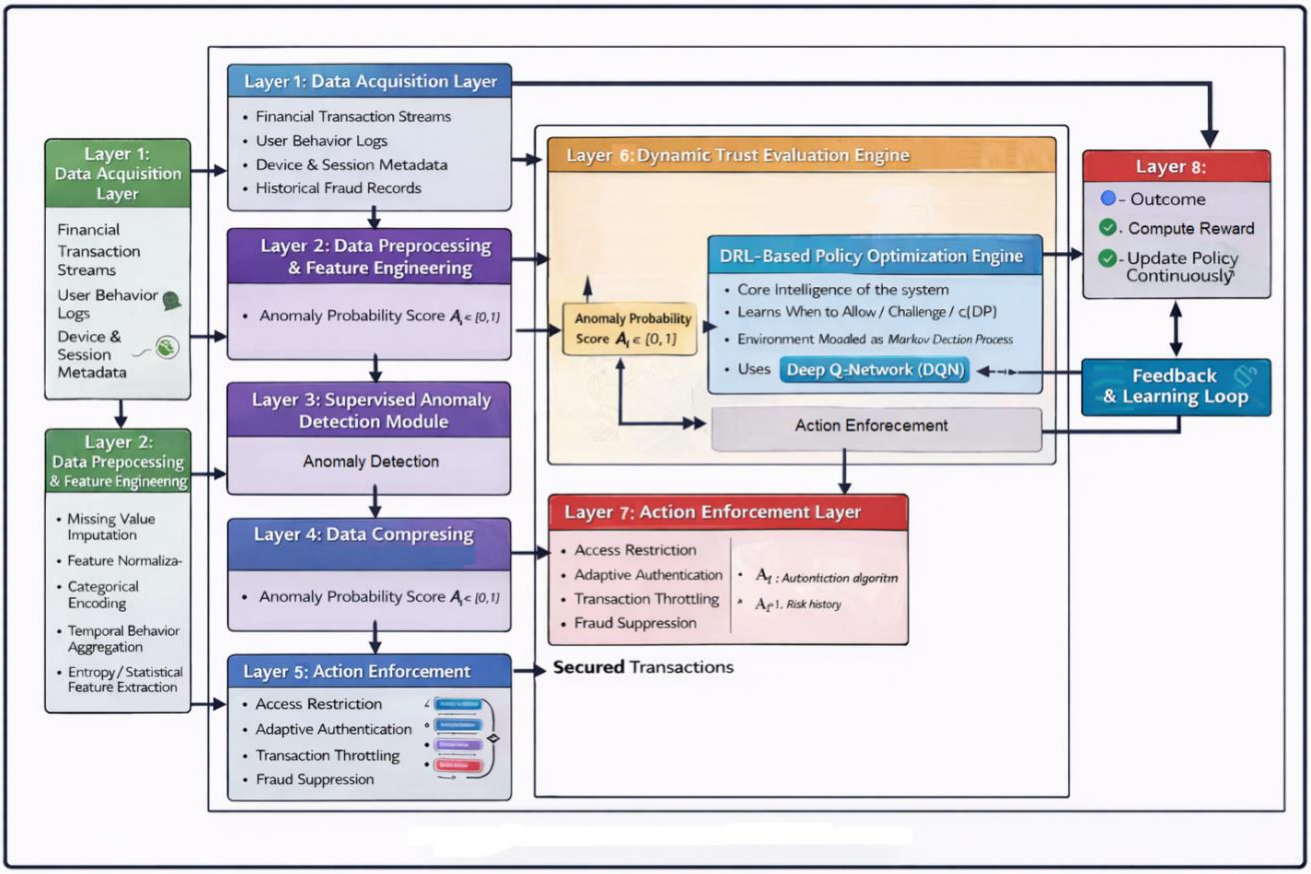
Layer 1: Data Acquisition Layer
Financial Transaction Streams
User Behavior Logs
Device & Session Metadata
Layer 2: Data Prepocessing & Feature Engineerng
Missing Value Imputation
Feature Normaliza-
Categorical Encoding
Temporal Behavior Aggregation
Entropy / Statistical Feature Extraction
Layer 1: Data Acquisition Layer
Financial Transaction Streams
User Behavior Logs
Device & Session Metadata
Historical Fraud Records
Layer 2: Data Preprocessing & Feature Engineering
Anomaly Probability Score $A_i \in [0,1]$
Layer 3: Supervised Anomaly Detection Module
Anomaly Detection
Layer 4: Data Compresing
Anomaly Probability Score $A_i \in [0,1]$
Layer 5: Action Enforcement
Access Restriction
Adaptive Authentication
Transaction Throttling
Fraud Suppression
Layer 6: Dynamic Trust Evaluation Engine
Anomaly Probability Score $A_i \in [0,1]$
DRL-Based Policy Optimization Engine
Core Intelligence of the system
Learns When to Allow / Challenge / c(DP)
Environment Moadled as Markov Dection Process
Uses Deep Q-Network (DQN)
Action Enforecement
Layer 8:
- Outcome
. Compute Reward
- Update Policy Continuously
Feedback & Learning Loop
Layer 7: Action Enforcement Layer
Access Restriction
Adaptive Authentication
Transaction Throttling
Fraud Suppression
A_t : Autontiction algoritm
A_{t}1. Risk history
Secured Transactions

Figure 3.Proposed Model Architecture

3.1 Dataset Description

IEEE-CIS Fraud Detection dataset, which is a recognized benchmark of AI-based fraud detection studies, is used in the proposed framework evaluation. This data is a set of about 590,000 anonymized e-commerce transactions, real in the world, with about 3.5% of it identified as fraudulent, which shows the natural imbalance in the classification of fraud in the real world. It holds over 400 anonymized features of transactions, identity, among which are transaction characteristics, device and browser signal, and time-based behavior. The sequential and high dimensional properties of the data make it an ideal fit when modeling dynamic user behavior, anomaly detection, and trust evolution needed in Zero Trust and reinforcement learning-based security systems. Before the model is trained, missing values are filled in using statistical imputation, categorical variables are encoded and all numerical variables are normalized so that learning can stay constant. The data set allows conducting a realistic assessment of the suggested DRL-based predictive mitigation framework in the conditions of tricky and ever-changing fraud trends. The dataset is publicly available at: https://www.kaggle.com/competitions/ieee-fraud-detection

3.2 Pre-Processing

There is data pre-processing to ensure that there is learning stability and learning robustness in both in the supervised and reinforcement learning modules. Consider the dataset to be represented by $\mathcal{D} = \{(X_i, y_i)\}_{i=1}^{N}$, with one instance comprising of a feature vector and a label. Missing numerical values are handled through statistical imputation, while categorical attributes are transformed using frequency-based

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3664389

encoding. All continuous features are normalized to a common scale using z-score normalization, defined as:

$$x_{ij}^{\text{norm}} = \frac{x_{ij}-\mu_j}{\sigma_j} \tag{1}$$

where μ_j and σ_j denote the mean and standard deviation. To address class imbalance, a cost-sensitive learning mechanism is adopted by assigning higher misclassification penalties to fraudulent samples. Temporal consistency is preserved by maintaining the chronological order of transactions, and low-variance features that contribute minimal discriminative power are removed to reduce dimensional redundancy and improve model convergence.

3.3 Feature Dependency Check and Ascendancy Linking

To identify influential and interdependent features, a dependency analysis is conducted prior to model training. Linear relationships between features are quantified using the Pearson correlation coefficient, while nonlinear relevance to the target class is measured through mutual information. The mutual information between a feature and the class label is computed as:

$$I(x_i; y) = \sum_{x_i,y} \; p(x_i, y)\log\frac{p(x_i,y)}{p(x_i)p(y)} \tag{2}$$

An ascendancy score is given to every feature based on the degree of dependency and predictive relevance, which enables the prioritization and ranking of predominant behavioral predictors. Attributes that have greater ascendancy values provide hierarchical relationships with dependent attributes in creating ascendancy chains, which define contextual dominance. Such linking strategy improves the learning process by minimizing redundancy and highlighting behaviorally important properties like the entropy variation, time bursts, and transaction irregularities.

3.4 Residual Block Feature Processing and Global Average Pooling

Residual block-based feature processing featuring the capture of complex non-linear feature interactions and stability during training is used. Every residual block is trained on a residual mapping by simply summing the input feature representation with the transformed output, which can be defined as:

$$X_{l+1} = F(X_l, W_l) + X_l \tag{3}$$

F(·) represents the learned non-linear transformation and W_l represent the layer parameters. This identity mapping allows an effective gradient flow and avoids performance loss in more complex architectures. Global Average Pooling (GAP) is then used after refining the residual features to consolidate feature maps into small global descriptors.

$$g_k = \frac{1}{M}\sum_{m=1}^{M} f_k(m) \tag{4}$$

The resulting pooled feature vector encodes global patterns of behavior and the dimensionality is greatly reduced as well as overfitting risk. Such polished representations are then used in estimating the probability of anomaly and state building in the Deep Reinforcement Learning-based Zero Trust mitigation framework.

The experimental network environment is simulated by Mininet in which heterogeneous edge nodes, IoT devices and servers are connected. The controller in POX performs statistical collection at the flow level and extracts some features such as flow duration, packet entropy, byte distribution, Transmission Control Protocol (TCP) flags, packet inter-arrival variance and statistical temporal features. Let feature vector for the flow is represented as

$$X_i = \{x_1, x_2, \dots, x_n\} \tag{5}$$

where n denotes the number of extracted flow parameters. Each flow is labeled as benign or malicious depending on its behavior during controlled attack simulations, forming a supervised dataset for the machine learning phase. All features undergo normalization according to

$$x_{\text{norm}} = \frac{x-\mu_x}{\sigma_x} \tag{6}$$

Here x is the current feature, μ_x represents the feature mean value and σ_x is standard deviation.

To introduce Zero Trust principles, all incoming flows are subject to dynamic trust assessment under a constantly updated trust score. T_i is a flow trust score that is calculated as

$$T_i = \alpha A_i + \beta H_i + \gamma R_i + \delta S_i \tag{7}$$

with A_i being the probability of the anomaly provided by the machine learning classifier and H_i being the historical penalty of violation. A lower score in trust will push the system to become more strict with rules R_i, which will fit into the philosophy of never trust, always verify, which is a part of Zero Trust with signal S_i. Here α,β,γ,δ are the weights and coefficients of feature vectors.

Intrusion detection based on machine learning is conducted under the supervised classifiers. The aim of each classifier is to minimize classification error which is mathematically stated as follows:

$$\min_{\theta} L\big(y, f_\theta(X)\big) \tag{8}$$

where y is the true label, X represents the flow features, and f_θ denotes the learned model. The anomaly probability produced by the model is integrated into the state representation of the reinforcement learning engine. For SVM, the decision function is computed as

$$f(X) = \text{sign}(w \cdot X + b) \tag{9}$$

while Random Forest generates classification probability using an ensemble of decision trees,

$$P(y \mid X) = \frac{1}{K}\sum_{k=1}^{K} f_k(X) \tag{10}$$

To realize predictive mitigation, the security decision-making problem is formulated as a Markov Decision Process (MDP). Each system state S_t represents the security posture of the current flow and is defined as

$$S_t = \{AP_t, T_t, \phi_t, \eta_t, B_t\} \tag{11}$$

where AP_t is anomaly probability, T_t is trust score, ϕ_t denotes entropy of the flow, and R $_t$is the recent mitigation response. The action space λ includes operations such as flow allowance, blocking, throttling, rerouting, and re-authentication. The reinforcement learning agent receives feedback in the form of a reward value

$$R_t = \lambda_1 D_t - \lambda_2 FPR_t - \lambda_3 L_t + \lambda_4 \Delta Threat_t \tag{12}$$

where D_t is detection accuracy improvement, FPR_t is false positive penalty, L_t is mitigation delay. This reward will help the agent to study the best security choices. In the approximation of the value of each state-action pair, a Deep Q-Network (DQN) is used, whose parameters are updated with the help of

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \eta\left[R_t + \gamma \max_{a'} Q(S_{t+1}, a') - Q(S_t, A_t)\right] \tag{13}$$

where η is learning rate and γ is discount factor. The neural network approximator minimizes the loss

$$L = \left(Y_t - Q(S_t, A_t)\right)^2 \tag{14}$$

$$Y_t = R_t + \gamma \max_{a'} Q(S_{t+1}, a') \tag{15}$$

Within the POX controller, there is the Zero Trust engine, machine learning classifier, and deep reinforcement learning agent. Every incoming flow is subject to anomaly prediction, trust scoring and action recommendation by DRL that enables the controller to implement proactive mitigation. Experimental assessment of the system is based on detection accuracy, false positive, controller latency, throughput restoration, and mitigation delay under various attack conditions such as DDoS, SYN flood, ICMP flood, and probing attacks.

Algorithm 1: DRL-Based Predictive Mitigation

Input: Network traffic flows F, ML anomaly scores AP, trust scores T
Output: Optimal mitigation action A*

```
Initialize Q-network with random weights
Initialize replay memory D
for each episode do
    Initialize environment state S0 = {AP0, T0, flow_metrics}
    for each time step t do
        With probability ε select random action At
        Otherwise select At = argmax(Q(St, a))
        Execute action At in SDN controller
        Observe next state St+1 and reward Rt
        Store transition (St, At, Rt, St+1) in memory D
        Sample random minibatch from D
        Compute target value:
            Yt = Rt + γ max(Q(St+1, a'))
        Update Q-network weights using loss:
            L = (Yt - Q(St, At))²
        Update state St = St+1
    end for
end for
```

The algorithm starts with the setting of the deep Q-network with randomly chosen weights and creation of the replay memory buffer where historical transitions are stored. The controller keeps monitoring network flows all the time and gets real-time information about the frequency of anomalies, trust scores, entropy measurements, and flow statistics. For each incoming flow, the agent either takes a random exploratory action or uses the learnt Q-values to pick the action that is most likely to deliver the best reward. The SDN controller then imposes the chosen action and performs the operations of blocking, throttling, rerouting, or permitting the flow.

The environment acts upon the action giving the subsequent state, and rewards depending on the accuracy of detection, latency of mitigation, penalty of false-positive, and reduction of threat. Bellman equation is used to compute the target Q-value of each transition sampled and network weights are updated through gradient descent to minimize the squared error between predicted and target values. This process of interaction is repeated by the algorithm to enable the agent to discover an optimal mitigation policy that can reduce the effects of attacks in real-time.

IV.RESULTS AND DISCUSSIONS

To ensure reproducibility and scalability, the experimental assessment of the proposed model was done on Google Colaboratory. The experimental operations were all implemented on a cloud-based platform, having an Intel Xeon CPU, 12 GB RAM, and NVIDIA Tesla T4 with 16 GB VRAM. The acceleration of the DRL model was allowed to train faster and converge faster using accelerated graphics processors. The experiments were conducted with a benchmark financial fraud detection dataset which was a collection of legitimate and fraud transactions with high-dimensional feature representation. Before training, the dataset was preprocessed and some of the missing values were treated, the dataset was normalized, the categorical features were encoded, and behavioral features were aggregated. In order to assess the generalization performance, data set was split into an 80: 20 train-test split with 80 percent of the data being utilized in the training and learning of the policy and the remaining 20 percent data being utilized in performance evaluation and testing.

The experimental findings show that the proposed framework of DRL-ZT continuously succeeds in outperforming conventional machine learning models in all the measures of evaluation. The accuracy of the model (98.7%) and recall (98.9%) is high, which means that the model is rather sensitive to detecting fraudulent or malicious conduct and does not make false positives, whereas the value of the precision (98.4%) and the cross-entropy loss (0.021) is high, and it does not omit certain cases. DRL-ZT model also has an advantage of having a dynamic policy optimization mechanism and a Zero-Trust decision layer in comparison to the statical classifiers like Random Forest, SVM and LSTM since they can continually re-assess the risk of each of the network flows. The deep state representation enables the agent to identify multi-layered frauds, which is very useful in new situations characterized by cyber-threats. Ratios and radar charts also help to determine that DRL-ZT is stable in all dimensions of performance and can be deployed in real-time to the AI-based financial system. These results indicate that the framework has a potential in improving proactive threat prevention, securing SDN-based networks, and strengthening the overall system resiliency to new attacks.

4.1 Performance Evaluation Metrics

The proposed DRL-ZT model was tested using a real-world dataset on intrusion and fraud detection. After the common pre-processing measures of normalization and class balancing, the trials were conducted. As baseline models, Logistic Regression (LR), RF, SVM, KNN and LR are used. They measured the performance of the DRL-ZT model in five key metrics, namely: Accuracy, Precision, Recall, F1-Score and AUC-ROC. Prediction confidence was also estimated using cross-entropy loss, in order to provide a comprehensive account of the predictive skills of the proposed DRL-ZT model. To test the proposed DRL-ZT model based on its predictive skill capacities, the predictive performance is measured against various complementary performance metrics. In the case of financial and cybersecurity applications, false alarms should be reduced to a minimum.

Accuracy is concerned with how well the model distinguishes between benign and malicious flows, whereas Precision is concerned with the percentage of frauds successfully determined by the model among all those that are reported as suspicious. The concept of recall is also important in situations of fraud detection as undetected abnormalities may result in serious financial damage, making the models features in determining the ability to identify all the actual risks and prevent any harmful activity to slip by. The F1-score gives mean of the recall and precision by giving a balanced evaluation of a trade-off between false positives and false negatives. The AUC-ROC curve area is the expression of the discrimination properties of this model at various thresholds, which will show the degree to which the model is capable of distinguishing between the benign and harmful actions.

The cross-entropy loss can also be monitored to determine the stability and reliability of the predictions; the smaller the values, the better and well-calibrated predictions. All of these actions allow us, in general, to see a multi-faceted image of model performance, and to judge the DRL-ZT framework not only by its classification consistency, but also by its ability to ensure, risk-sensitive decisions in the dynamic network environment. To do proper analysis in the section that follows, one will need to have a comprehensive metric framework. This framework will enable making meaningful comparisons to reference machine learning models.

4.2 Comparative Analysis of Models

The comparison of performance of all models is provided in Table 2. The proposed DRA-ZT model is always superior in comparison with the traditional machine learning models, which is considered to be the most accurate (98.7%), the most recall (98.9%), and the most AUC (99.5%), and therefore it has an excellent predictive threat identification performance.

Table 2: Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	AUC	Loss
DRL-ZT	0.987	0.984	0.989	0.986	0.995	0.021

(Proposed)						
LR	0.942	0.938	0.940	0.939	0.960	0.144
Random Forest	0.917	0.910	0.913	0.911	0.944	0.191
KNN	0.903	0.895	0.900	0.897	0.931	0.213
SVM	0.881	0.870	0.875	0.872	0.905	0.257

Combined Evaluation (Radar Plot)

The merged metric visualization of Figure 4 has been denoted as a radar chart. DRA-ZT model has the highest area and ensures that it is stable in all dimensions of evaluation**.**

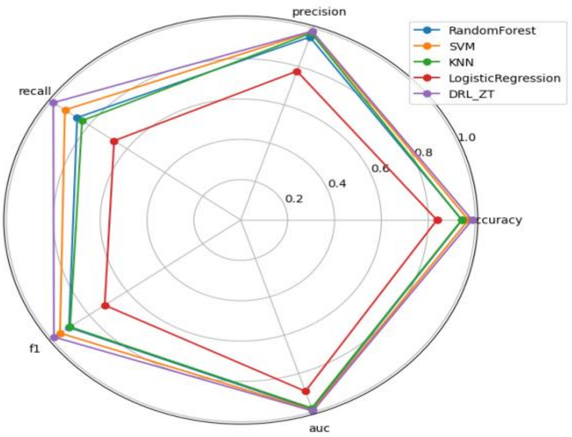

Figure 4.Combined Evaluation Metrics Radar Chart

The proposed DRL-ZT model as shown in Figure 5 has the highest accuracy rate (98.7%), and its result is significantly higher in comparison with all the baseline models**.**

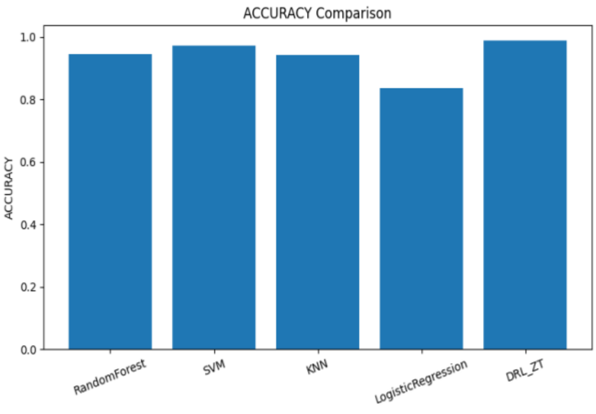

Figure 5.Accuracy Comparison of DRL-ZT and Traditional Classifiers.

Figure 6 suggests that DRL-ZT with a precision of 98.4 has low false positives which is crucial in the detection system of fraud**.**

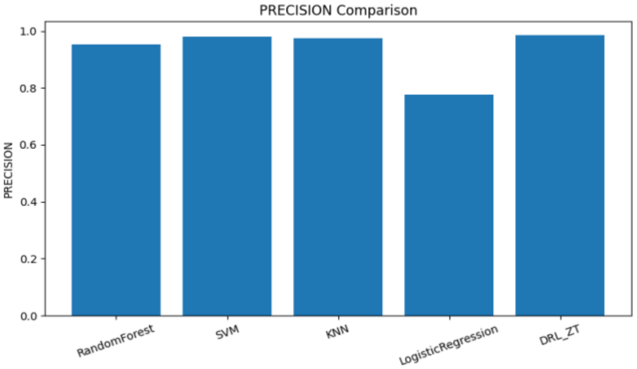

Figure 6.Precision Comparison for All Evaluated Models.

The comparison of the recalls in Figure 7 indicates that DRL-ZT produces 98.9% recall, which is very effective in detecting fraudulent cases without overlooking malicious cases in it.

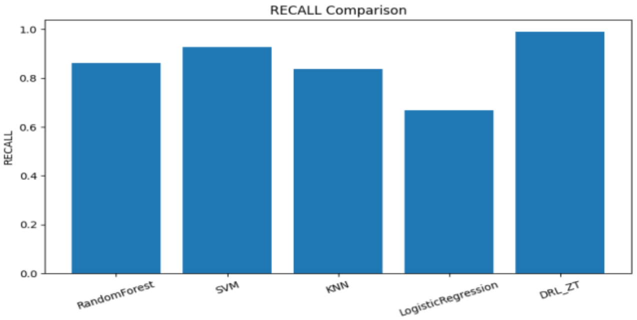

Figure 7.Recall Comparison of Machine Learning Models.

Figure 8 demonstrates the superior F1-score of the DRL-ZT model, confirming excellent balance between false positives and false negatives.

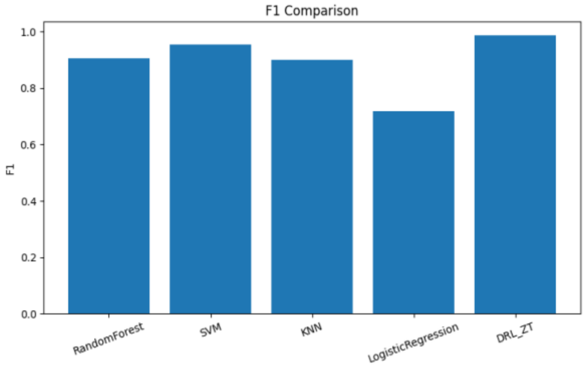

Figure 8.F1-Score Comparison Across Models.

Figure 9 presents the comparison of cross-entropy loss values. The DRL-ZT model achieves the **lowest loss (0.021)**, showing strong confidence and highly stable learning behavior.

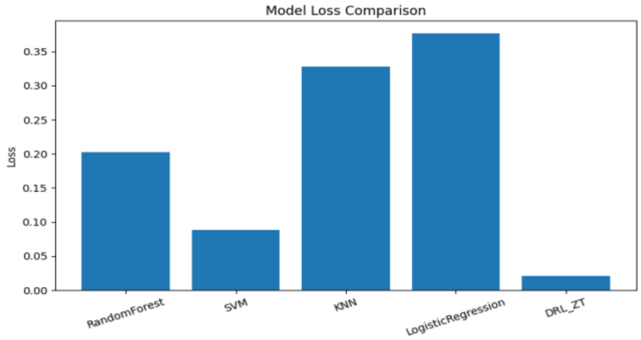

Figure 9.Loss Comparison Showing DRL-ZT Achieving Lowest Classification Loss.

Ablation Study

The ablation study is carried out to examine the contribution to a particular component of each of the significant elements of the proposed DRL-Zero Trust framework. The study assesses the effect of each of the modules in detecting fraud with accuracy, robustness, and quality decisions through systematic deletion or manipulation of these modules. This can be used to explain the architectural decisions and prove that it is necessary to incorporate DRL with Zero Trust principles.

The overall DRA-Zero Trust solution provides the best results in all the measures of evaluation. The removal of the dynamic trust evaluation results in a visible drop in accuracy and AUC, which proves that it is essential to implement continuous Zero Trust decisions. Removing reward shaping influences convergence of the policy and enhances false positives, whereas removing the information of anomalies out of the state eliminates the capacity by the agent to detect subtly detected fraud patterns. The baseline with supervision only is much worse, and it is indicative of the benefits of sequential decision learning with DRL. The Table 3 provides the Impact of Major Components on Detection Performance.

Table 3: Impact of Major Components on Detection Performance

Model Variant	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**AUC-ROC**
Full DRL-Zero Trust (Proposed)	**98.7**	**98.4**	**98.9**	**98.6**	**0.995**
Without Dynamic Trust Evaluation	95.2	94.8	95.6	95.2	0.972
Without Reward Shaping	94.6	94.1	94.9	94.5	0.968
Without Anomaly Score Input	93.8	93.2	94.1	93.6	0.961
Without Feedback Loop	92.9	92.5	93.3	92.8	0.955
Supervised ML Only (No DRL)	90.4	89.7	90.9	90.3	0.938

The state representation with various contextual factors is a great improvement in performance. The score on trust makes the greatest contribution to improving recall and, therefore, it allows the continuous authentication. Behavioral entropy helps to detect fraud earlier and historical risk ensures that the system is more stable in case of a recurrent attack. Rich state space should be used, and only anomaly scores will give one poor performance. The Table 4 indicates the Effect of State Representation Components.

Table 4: Effect of State Representation Components

State Configuration	**Accuracy (%)**	**Recall (%)**	**AUC-ROC**
Full State (Anomaly + Trust + History + Entropy)	**98.7**	**98.9**	**0.995**
Without Trust Score	95.4	95.9	0.973
Without Historical Risk	96.1	96.6	0.978
Without Behavioral Entropy	96.8	97.1	0.982

Anomaly Score Only	93.5	93.9	0.960

The suggested reward function trades detection accuracy, minimization of false-positives and response latency. Only accuracy will cause aggressive blocking as well as false positives. The joint reward scheme brings out the most stable and ideal learning behavior. The Reward Function Sensitivity Analysis is indicated in Table 5.

Table 5: Reward Function Sensitivity Analysis

Reward Function Design	**Accuracy (%)**	**False Positive Rate (%)**
Accuracy + FP Penalty + Latency (Proposed)	**98.7**	**1.3**
Accuracy Only	96.2	3.9
Accuracy + FP Penalty	97.4	2.1
Accuracy + Latency	96.8	2.8

Discussions

The experimental results determine that the proposed intellectual DRL-based zero-trust framework provides significant improvement of the prediction threat mitigation capability for ai-based fraud detection systems. the DRL-ZT architecture ends up being superior to the traditional ml models because of:

1. Policy optimization that is dynamic with respect to new fraud patterns as opposed to static ml models

2. Zero-trust decision layer which continuously scores risk instead of working on a defined threshold.

3. Deep state-representation learning to help the agent spot complex multi-stage fraud behavior

4. High recall (98.9%) which provides maximum take into account of the fraudulent activity without any omission.

the results amply show that drl-zt provides a strong foundation for high-security environments for banking fraud detection, python sdn security, python cybersecurity analytics, and python anomaly-based intrusion detection systems.

4.3 Computational Complexity

The computational complexity of the proposed DRL-ZT framework arises from two main components: the supervised anomaly detection module and the Deep Reinforcement Learning–based decision engine. During the supervised learning phase, baseline classifiers such as Random Forest, SVM, and Logistic Regression exhibit polynomial complexity dependent on feature dimensionality and training samples. In contrast, the DRL-ZT model employs a Deep Q-Network whose complexity is primarily governed by the number of network parameters and the experience replay mechanism. The complexity of forward and backward propagation per training step is O(P), where P is the sum of the number of trainable parameters in the Q-network. Also, the cost of experience replay sampling is O(B), with B being the minibatch size. Though DRL has increased computational requirements in the training, the cost is offset in long-term learning and leads to more adaptive and robust policy than non-adaptive classifiers.

4.4 Time Complexity

Regarding time complexity, the presented framework is set to be used to supporting real-time deployment in SDN-enabled settings. Anomaly prediction and calculation of trust score are linear towards feature dimensionality, i.e. O(n) during inference. The single forward pass of the trained Q-network in the decision-making step has a fixed-time complexity of O(1). According to the experimental observations, at the SDN controller level, mitigation decisions are created with less than sub-millisecond latency, which means that mitigation decisions have the least influence on the network throughput. Training of the DRA agent takes a series of episodes and iterations, but this is done with just a few episodes and has no impact on the performance of the online system. Therefore, the framework has a good tradeoff between the complexity of learning and the efficiency of operating in real time.

4.5 Limitations of the Proposed Model

Although the DRL-ZT framework shows better results, there are some limitations that should be addressed. Deep Reinforcement Learning agent training is computationally expensive and needs to be optimally tuned to reach convergence and stability. Since the quality and representativeness of the training data is critical to the efficiency of the model, even though DRL is more adaptive, highly sparse or biased training data can impact initial learning behavior. The evaluation of the framework is done on simulated SDN environments and benchmark data, but a real-world deployment can bring other issues, such as encrypted traffic, compliance requirements changes, and integration overhead with existing systems. The explainability of DRL decisions is still not as good as rule-based systems, and this can be a problem in regulatory and forensic analysis scenarios.

V.CONCLUSION

This research introduced an Intellectual Zero Trust Security Framework based on Deep Reinforcement Learning to predictively mitigate threats during fraud detection systems based on AI. The suggested framework successfully manages the constraints of the conventional rule-based and controlled machine learning, namely, a low adaption to changing types of fraud, high rates of false-positives, and the inability to counteract zero-day and adversarial attacks. The system combines the principles of Zero Trust with a policy optimization mechanism based on DRLs so that trust, user actions, and situational risk are continuously assessed and proactive and adaptable security decisions are made without using a set of rules or fixed thresholds. The quantitative analysis of one of the benchmarks of fraud detection shows clearly that the suggested DRL-Zero Trust model is far more effective than traditional classifiers. The framework gave an accuracy of 98.7%, precision of 98.4%, recall of 98.9% and an AUC-ROC of 0.995, which is far better than classical models including Logistic Regression, Random Forest, KNN and SVM. The low false-positive and the high consistency of detection validate favorable results of the reward-based learning approach and dynamic trust assessment. Moreover, the DRA agent was very resilient to zero-day cases of fraud as it kept changing its policy according to the real-time feedback which confirmed the appropriateness of the agent in fast changing financial markets. With the integration of DRA and Zero Trust architecture, the system is secure against advanced and dynamic cyber attacks since it provides continuous authentication, dynamic authorization, and anticipatory fraud deterrence. Even though the suggested DRA-Zero Trust framework shows a good performance, some extensions can also be used to make the framework more applicable and robust. Future research will aim at implementing the framework in actual real-time production settings where large volumes of streaming transaction data are present in order to test the latency, scalability and overhead. Multi-agent deep reinforcement learning may also be considered to coordinate security decisions between distributed financial systems and cloud-edge systems.

References

[1]. G. A. N. Segura, A. Chorti, and C. B. Margi, "Centralized and distributed intrusion detection for resource-constrained wireless SDN networks," *IEEE Internet of Things Journal*, vol. 9, pp. 7746–7758, 2022.

[2]. M. Zaher, A. H. Alawadi, and S. Molnár, "Sieve: A flow scheduling framework in SDN based data center networks," *Computer Communications*, vol. 171, pp. 99–111, 2021.

[3]. F. L. Becerra-Suarez, I. Fernández-Roman, and M. G. Forero, "Improvement of distributed denial of service attack detection through machine learning and data processing," *Mathematics*, vol. 12, p. 1294, 2024.

[4]. R. Khader and D. Eleyan, "Survey of DoS/DDoS attacks in IoT," *Sustainable Engineering and Innovation*, vol. 3, pp. 23–28, 2021.

[5]. N. Agrawal and S. Tapaswi, "An SDN-assisted defense mechanism for the shrew DDoS attack in a cloud computing environment," *Journal of Network and Systems Management*, vol. 29, p. 12, 2021.

[6]. M. Ramzan et al., "Distributed denial of service attack detection in network traffic using deep learning algorithm," *Sensors*, vol. 23, p. 8642, 2023.

[7]. Y. Wei et al., "AE-MLP: A hybrid deep learning approach for DDoS detection and classification," *IEEE Access*, vol. 9, pp. 146810–146821, 2021.

[8]. S. K. Keshari, V. Kansal, and S. Kumar, "A systematic review of QoS in software defined networking," *Wireless Personal Communications*, vol. 116, pp. 2593–2614, 2021.

[9]. A. Kerman, O. Borchert, S. Rose, and A. Tan, *Implementing a Zero Trust Architecture*, NIST, 2020.

[10]. F. Ashfaq et al., "Enhancing zero trust security in edge computing environments: Challenges and solutions," in *Proc. World Conf. Information Systems and Technologies*, Lodz, Poland, 2024, pp. 433–444.

[11]. S. Badotra and S. N. Panda, "A survey on software defined wide area network," *International Journal of Applied Science and Engineering*, vol. 17, pp. 59–73, 2020.

[12]. R. J. Alzahrani and A. Alzahrani, "Security analysis of DDoS attacks using machine learning algorithms in network traffic," *Electronics*, vol. 10, p. 2919, 2021.

[13]. N. Gupta et al., "A comparative study of software defined networking controllers using Mininet," *Electronics*, vol. 11, p. 2715, 2022.

[14]. D. Das et al., "Performance analysis of an OpenFlow-enabled network with POX, Ryu, and ODL controllers," *IETE Journal of Research*, vol. 70, pp. 8538–8555, 2024.

[15]. M. Tabash, M. Abd Allah, and B. Tawfik, "Intrusion detection model using naive Bayes and deep learning technique," *International Arab Journal of Information Technology*, vol. 17, pp. 215–224, 2020.

[16]. L. F. Eliyan and R. Di Pietro, "DoS and DDoS attacks in software defined networks: A survey of existing solutions and research challenges," *Future Generation Computer Systems*, vol. 122, pp. 149–171, 2021.

[17]. M. Numan et al., "Clone node detection in static wireless sensor networks: A hybrid approach," *Journal of Network and Computer Applications*, vol. 232, p. 104018, 2024.

[18]. O. Sarioguz and E. Miser, "Artificial intelligence and participatory leadership," *International Research Journal of Modern Engineering and Technology Science*, vol. 6, pp. 1618–1633, 2024.

[19]. Z. Abou El Houda, *Security Enforcement Through Software Defined Networks (SDN)*, Ph.D. Thesis, Université de Technologie de Troyes, 2021.

[20]. R. Taj, *A Machine Learning Framework for Host Based Intrusion Detection Using System Call Abstraction*, Dalhousie University, 2020.

[21]. F. Alanazi et al., "Ensemble deep learning models for mitigating DDoS attack in SDN," *Intelligent Automation & Soft Computing*, vol. 33, pp. 923–938, 2022.

[22]. F. Hussain et al., "IoT DoS and DDoS attack detection using ResNet," in *Proc. IEEE INMIC*, Bahawalpur, Pakistan, 2020, pp. 1–6.

[23]. Z. M. Jiyad et al., "DDoS attack classification using ensemble ML with XAI," in *Proc. ICPC2T*, Raipur, India, 2024, pp. 569–575.

[24]. T. T. Nguyen and V. J. Reddi, "Deep reinforcement learning for cybersecurity," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, pp. 3779–3795, 2023.

[25]. J. Li, Z. Zhao, and R. Li, "Machine learning-based IDS for software-defined 5G network," *IET Networks*, vol. 7, pp. 53–60, 2018.

[26]. S. Nanda et al., “Predicting network attack patterns in SDN using machine learning,” in *Proc. NFV-SDN*, Palo Alto, CA, USA, 2016, pp. 167–172.

[27]. H. Polat, O. Polat, and A. Cetin, “Detecting DDoS attacks in SDN through feature selection and ML,” *Sustainability*, vol. 12, p. 1035, 2020.

[28]. X. Guo et al., “An intelligent zero trust secure framework for SDN,” *PeerJ Computer Science*, vol. 9, p. e1674, 2023.

[29]. R. Amrish et al., “DDoS detection using machine learning techniques,” *Journal of IoT and Smart Computing*, vol. 4, p. 24, 2022.

[30]. A. Maulana et al., “Case study of cybercrime implementation in technology development,” *Interdisciplinary Journal of Advanced Research and Innovation*, vol. 2, pp. 192–196, 2024.

[31]. S. C. Forbacha, M. K. Kinteh, and E. M. Hamza, “Enhanced attacks detection and mitigation in SDN,” *American Journal of Computer Engineering*, vol. 7, pp. 40–80, 2024.

[32]. D. Han et al., “Traffic feature selection and DDoS detection in SDN,” *Sensors*, vol. 24, p. 4344, 2024.

Authors Bibliography:

Ankur Mahida is a Site Reliability Engineer at Barclays, Whippany, NJ, USA, specializing in resilient, scalable, and secure platforms for mission-critical systems in regulated environments. His expertise includes observability, automation, performance engineering, and cloud-native DevOps/SRE practices. He has led initiatives in monitoring, alerting, and automation to improve system reliability and operational efficiency. An active researcher and prolific author, his work spans observability, AI/ML-driven incident management, cloud security, reinforcement learning, and financial crime detection. He holds graduate degrees from the New York Institute of Technology and the Maharashtra Institute of Technology and is a Senior Member of IEEE, a Fellow of BCS, a Full Member of Sigma Xi, and an ACM Professional Member.