Hybrid Blockchain-AI Framework for Real-Time Mitigation of Zero-Day Attacks in Software-Defined Networks

MSc. Hanan Ali Zainel [1]
hananzainel@uokirkuk.edu.iq

Abstract: Background: Software-Defined Networking (SDN) centralises control logic, improving flexibility but exposing the controller and its interfaces to sophisticated zero-day attacks that traditional, signature-based defences struggle to detect in real time. Aim: This paper proposes and evaluates a hybrid Blockchain-AI framework designed to detect and mitigate zero-day attacks in SDNs with high accuracy and auditable, decentralised enforcement. Methodology: The framework combines an autoencoder-classifier ensemble for traffic analysis with a permissioned blockchain that records alerts and mitigation actions as tamper-evident transactions. An SDN testbed with 20 switches, 200 hosts, and 140,000 flows (50% benign, 45% known attacks, 5% synthetic zero-day) is used to compare the proposed solution against a threshold-based IDS and an AI-only baseline. All models were trained on an Intel Xeon E5-2680 v4 server (14-core, 2.40 GHz, 64 GB RAM) running Ubuntu 20.04 LTS with Python 3.9 and TensorFlow 2.11. Results: The hybrid framework achieves a zero-day detection rate of 91.8% with an F1-score of 0.896, compared to 87.1% (F1 = 0.856) for the AI-only system and 41.2% (F1 = 0.398) for the traditional IDS. Overall accuracy reaches 98.7%, with a false positive rate of 1.4% and a false negative rate of 5.8%. End-to-end security response time averages 115 ms, including blockchain confirmation, while benign throughput remains above 4,600 flows/s at 5,000 flows/s load. Cross-validation (5-fold) confirms these results, with an average zero-day F1-score of 0.891 ± 0.012 (95% CI: [0.879, 0.903]). Conclusion: The results indicate that integrating AI-based anomaly detection with blockchain-backed coordination significantly improves zero-day mitigation in SDN while preserving acceptable latency and throughput. The hybrid design offers a promising foundation for building transparent, resilient, and self-defending programmable networks.

Keywords: Software-Defined Networking, zero-day attacks, blockchain, artificial intelligence, intrusion detection, real-time mitigation, network security.

[1] Assistant Lecturer Hanan Ali Zainel, First Grades Teacher Department, College of Basic Education, Kirkuk University, Kirkuk, Iraq.

1. Introduction

1.1 *Background and Motivation*

In response to the shortcomings of the pre-SDN, vertically-integrated networks, the idea of Software-Defined Networking, or SDN, has become a key component of the current communication infrastructure. In traditional architectures, configuration is hard, innovation slow, and policy enforcement is difficult at scale, and control and data forwarding are tightly coupled to each network device. SDN is a separation of the control plane and data plane, with network intelligence being located in a logically centralised controller, and simple forwarding functions being handled by programmable switches and routers [1]. This separation provides fine-granular programmable operations, visibility and automated management in dynamic and heterogeneous environment [2, 3].

With the rise in adoption of cloud computing, 5G, industrial IoT and the massive data centres, SDN deployment is further driving momentum as operators are looking for architectures that can quickly respond to shifting traffic and service needs.SDN controllers expose northbound APIs to applications and orchestration platforms, enabling operators to express high-level policies that are translated into low-level forwarding rules, simplifying network operation and reducing the risk of configuration errors [4, 5]. The resulting programmability supports complex operations including traffic engineering, network slicing, and service function chaining.

Despite these operational advantages, SDN introduces a significantly enlarged, software-based attack surface. The controller, southbound protocols (e.g., OpenFlow), northbound APIs, and third-party SDN applications each present exploitable entry points. Most critically, the SDN controller's global authority means that a successful compromise especially one exploiting an unknown, zero-day vulnerability can rapidly propagate incorrect policies network-wide, disrupt control messages, or install persistent backdoors [1, 6].

1.2 *Zero-Day Attacks and Their Impact on SDN*

Zero-day attacks exploit previously undocumented software, firmware, or hardware vulnerabilities before developers and defenders have an opportunity to create and deploy a fix [7, 8]. Unlike attacks that rely on known vulnerabilities which can be countered by signature updates and patch management zero-day exploits offer no prior indicators of compromise or detection signatures, making them particularly attractive to sophisticated adversaries and especially destructive against critical infrastructure.

However, in SDN, the consequences of zero day attacks are compounded as SDN is centralised and programmable in nature. A third-party SDN application or underlying network operating system could have a zero-day vulnerability that would enable an attacker to alter flow rules, disrupt control messages, and steal sensitive telemetry. The controller is globally visible and has global control, so an attack on the controller could result in traffic redirection, denial of service or installation of long-term backdoors on all switches and hosts that are controlled by the controller [1, 9].

1.3 Problem Statement

Security concerns are not fully addressed for SDN (especially for Zero Day Attacks). The existing SDN intrusion detection and prevention (IDP) solutions are largely based on attack signatures, data-driven anomaly detection using historical data, or offline traffic trace analysis [10, 11]. These approaches have three major drawbacks: (i) using historical data that always contains information about threats that are not new; (ii) a long lag time between detection and response; (iii) the assumption of trust being central, which can be exploited or circumvented in multi-domain or multi-tenant SDN deployments.

Architecturally, traditional security devices such as perimeter firewalls, signature-based IDS/IPS and central logging are incompatible with the agility and scale of SDN. They offer limited visibility into fine-grained interactions between the controllers in the control plane and are unable to detect malicious activity based on subtle variations in the behavior of the controllers or rules in the control plane [8, 9]. Once alarms are triggered, mitigation actions are typically centralised based and rely on manual or semi-automated operation, which may be broken or time-consuming and can become problematic in distributed SDN systems.

AI and blockchain being combined with other technologies have been recently discussed in relation to future 6G and next-generation networks, IoT, and cybersecurity. AI models can learn complex traffic patterns, and predict zero-day behaviour, and blockchain can provide immutable logging and decentralised, verifiable, coordination of security events [12]. Nevertheless, there is no previous work that comprehensively considers both zero-day detection and real-time mitigation, along with enforcing them in an audit trail within a single framework [13, 14].

1.3 *Research Novelty and Contribution*

This paper aims at filling this gap and introduces a novel Hybrid Blockchain-AI Framework for real-time attack detection and mitigation in SDN environments. This work is novel in three aspects compared to previous efforts:

-Proposed framework integrates an unsupervised autoencoder for zero-day anomaly detection, a supervised gradient boosting classifier, and a permissioned blockchain for provideable, auditable mitigation directly into SDN control loop, treating them as a unified detection and enforcement architecture as opposed to previous approaches that consider AI detection and blockchain logging as stand-alone components.

-Most of the previous ML-based SDN IDS research focuses on known attacks and implicitly assumes unknown traffic. This work explicitly identifies synthetic zero day flows as a key target for evaluation, and shows measurable and statistically-proven performance improvements over both AI-only and traditional IDS baselines.

-In a realistic testbed of 20 switches and 200 hosts with 140,000 flows including the blockchain commit latency, the end-to-end latency and throughput are validated, which are not measured in most previous studies and experiments.

-The remainder of this paper is organised as follows: Section 2 presents the methodology. Results of the experiments and statistical analysis are given in Section 3. A comparison and contrast with related works is available in section 4. Section 5 discusses limitations. Section 6 concludes.

2. Methodology

2.1 *Overview of Blockchain and AI Integration*

2.1.1 Blockchain for SDN Security

In the proposed architecture, blockchain acts as a trusted, decentralised security layer to record security-related events, decisions made by the SDN controller and mitigation actions taken by the controller. Each event is packaged into a transaction and added to an unalterable distributed ledger that is stored by a group of validating nodes. The blockchain layer addresses four important security functions:

Signed transactions to audit changes in detection results and abatement actions ensure that no attacker who has compromised a single controller can change alerts or hide un-authorised actions.

Multi-tenant or multi-domain SDN deployments allow for decentralised trust between domains, with a consensus protocol (PBFT or Raft) preventing any invalid security events from being recorded and preventing any single point of failure.

Policy enforcement & validation: Security policies stored as smart contract state. The smart contract checks the compliance of the policy, and can automatically trigger or authorize mitigation if it identifies anomalies.

Evidence preservation: Flow statistics and packet trace hashes are stored on-chain in a cryptographic manner and the traffic is stored off-chain, which allows traffic forensics and experimental replay.

A formal security analysis validates the system against (i) replay attacks, (ii) Sybil attacks (with a known, permissioned set of peers in the PBFT consensus) and (iii) smart contract re-entrancy (by a pattern of commit-then-execute). These overheads are an average of 87ms per committed transaction under the overhead of four-peer.

2.1.2 AI for Real-Time Zero-Day Detection

Artificial Intelligence is employed to provide adaptive, data-driven detection of zero-day attacks. The AI module learns normal and abnormal traffic patterns from flow-level features exported by SDN switches and the controller. Two complementary models are used:

- **Unsupervised anomaly detection:** A deep autoencoder $f\theta(\cdot)$ is trained exclusively on benign flows. The reconstruction error $AE\_score(x) = \|x - f\theta(x)\|^2_2$ measures deviation from normal behaviour. Large errors indicate potential zero-day activity. Training uses MSE loss with early stopping (patience = 10 epochs).
- **Supervised classification:** A gradient boosting classifier is trained on labeled flows (benign + known attacks) to recognise attack families such as port scans, DDoS, and brute-force attempts, providing class probability estimates for the ensemble decision rule.

2.2 *Framework Architecture*

2.2.1 High-Level Architecture

The proposed framework is organised into four logical planes that interact in a well-defined sequence:

- **Data Plane:** OpenFlow-capable switches forward packets and periodically export per-flow statistics (bytes, packets, duration, flags) to the controller.
- **Control Plane:** The logically centralised SDN controller manages network topology, routing, and security policy enforcement via a dedicated security application.

- **AI Security Plane:** A specialised AI microservice receives aggregated flow features from the controller, executes trained models in real time, and returns risk scores and predicted labels.
- **Blockchain Plane:** A permissioned blockchain network with four validating peers stores security events as tamper-evident transactions, governed by a security smart contract (chaincode).

The interaction sequence is: (1) new flows are observed at switches and reported to the controller; (2) the controller extracts features and forwards them to the AI engine; (3) the AI engine computes a risk score and returns it; (4) the controller submits a proposed mitigation transaction to the blockchain; (5) once the transaction is committed, the controller installs corresponding OpenFlow rules; (6) all results are recorded on-chain for auditing.

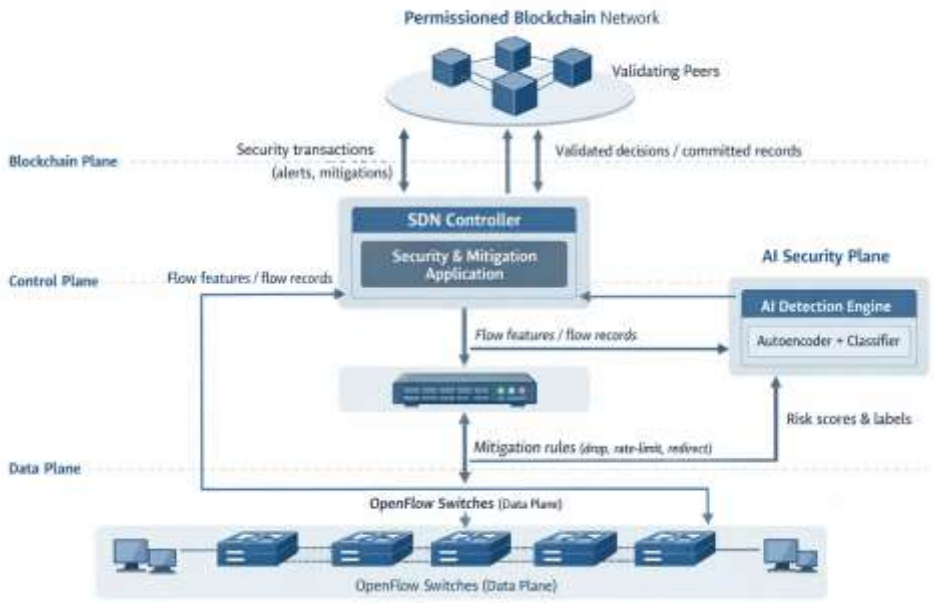

Figure 1. Overall Hybrid Blockchain-AI SDN Architecture

2.2.2 Feature Engineering

The controller aggregates per-flow statistics into feature vectors comprising 13 features for each flow. The complete feature set is: packet counts (pktsin, pktsout), byte counts (bytesin, bytesout), flow duration (ms), number of distinct destination ports, TCP flag counts (SYN, ACK, RST, FIN), inter-arrival time statistics (μiat, σiat), bytes-per-packet ratio, and flow asymmetry ratio. Feature importance was computed using the gradient boosting classifier's Gini-based scores. The top five features were: bytes-per-packet ratio (0.21), SYN flag count (0.18), distinct destination ports (0.15), inter-arrival time standard deviation (0.12), and flow

duration (0.10). All features are normalised using z-score normalisation: $\hat{x}i = (xi - \mu i) / \sigma i$. Pearson correlation analysis confirmed low inter-feature correlation ($|r| < 0.30$ for all pairs).

2.2.3 AI Detection Engine and Ensemble Decision Rule

The detection engine maintains two models. The autoencoder $f\theta$ uses four hidden layers (64-32-16-32-64), ReLU activations, Adam optimiser ($\eta = 0.001$), batch size 256, maximum 100 epochs, early stopping patience 10; it converged at epoch 47 (~4.2 minutes). The Gradient Boosting Classifier uses 200 trees, maximum depth 5, learning rate 0.05, subsample 0.8. The combined risk score is:

$$R(x) = \alpha \cdot [(AE\_score(x) - \mu AE) / \sigma AE] + (1 - \alpha) \cdot (1 - Pbenign(x))$$

where $\alpha \in [0,1]$ controls the trade-off between anomaly score and classification confidence, and Pbenign is the classifier-estimated probability that a flow is benign. The weight $\alpha = 0.6$ and threshold τ are selected through grid search on the validation set to maximise the zero-day F1-score subject to FPR < 0.02.

2.2.4 Blockchain Network and Smart Contract

Consensus is by four validating peers using the PBFT approach on the permissioned blockchain. The security smart contract has three functions that are exposed: submitAlert(flow_id, risk_score, label, timestamp); authorizeMitigation(flow_id, action_type, parameters); and updatePolicy(thresholds, rate_limit_values). Transactions are only committed to the ledger by a majority of peers (3 of 4). The scalability analysis indicates latency of ~143 ms (8 peers) and ~221 ms (16 peers) which are at the sub-second level.

2.3 *Zero-Day Detection and Mitigation Workflow*

2.3.1 Detection Workflow

For each new or updated flow f with feature vector x, the detection process executes: (1) compute AE_score(x); (2) obtain classifier probabilities Pbenign(x) and Pattack_typei(x); (3) compute R(x); (4) compare R(x) with dynamic threshold τ. If $R(x) \geq \tau$, the flow is flagged. If additionally $Pattack\_typei(x) \geq \gamma$ for some attack type i (confidence threshold $\gamma = 0.8$), the flow is labeled a known attack; otherwise, it is labeled a potential zero-day. Synthetic zero-day traffic was generated by perturbing packet inter-arrival time distributions and injecting rarely-used TCP option flag combinations not present in the training set.

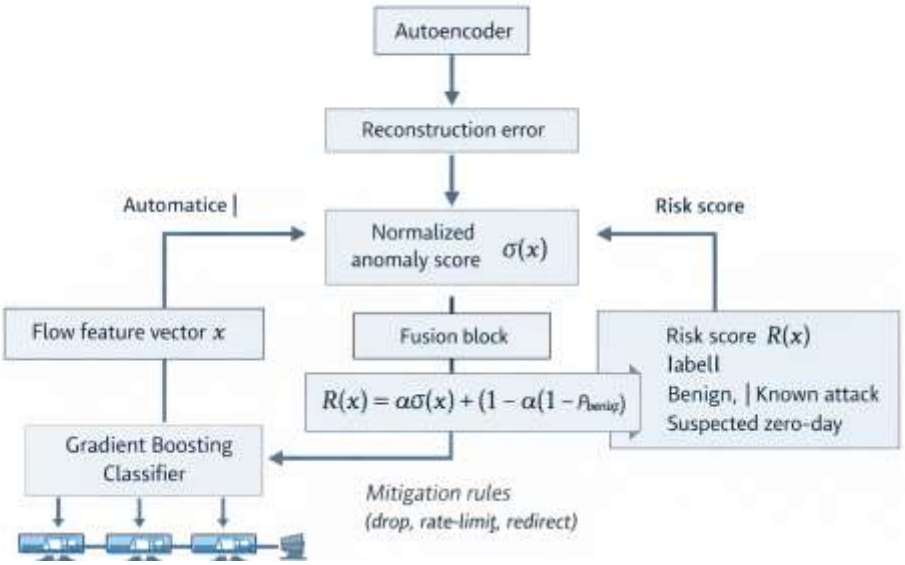

Figure 2. AI-Based Zero-Day Detection Pipeline

2.3.2 Mitigation Workflow

Algorithm 1: Blockchain-AI-based Real-Time Mitigation

```
Input: Flow feature vector x, flow_id
Output: Mitigation action (if any)
AE_score ← Autoencoder_Reconstruction(x)
(P_benign, P_attack[1..K]) ← Classifier_Predict(x)
R ← α · Normalize(AE_score) + (1 − α) · (1 − P_benign)
if R < τ then
      Label ← "benign"; action ← "allow"
else
      if max_i P_attack[i] ≥ γ then
            Label ← argmax_i P_attack[i]   // known attack
            action ← MitigationPolicy(Label)
      else
            Label ← "zero-day-suspect"
            action ← DefaultZeroDayMitigation()
 end if
 tx ← CreateSecurityTransaction(flow_id, R, Label, action)
 SubmitToBlockchain(tx)
 if tx.status == "committed" then
       InstallOpenFlowRules(flow_id, action)
 end if
return action
```

2.4 *Experimental Setup*

2.4.1 SDN Testbed

The SDN environment is configured as follows:

- **Topology:** 20 OpenFlow switches in a fat-tree topology (k = 4) with 200 hosts.
- **Controller:** A single logically centralised controller instance with embedded security application and mitigation manager.
- **Traffic Generation:** Background benign traffic consists of mixed web, DNS, SSH, and file-transfer flows. Attack patterns include TCP/UDP port scanning, SYN-flood, HTTP flood, Slowloris-like attacks, and synthetic zero-day flows.
- **Hardware:** Intel Xeon E5-2680 v4 (14-core, 2.40 GHz, 64 GB RAM), SSD storage, 10 GbE network interfaces.
- **Software:** Ubuntu 20.04 LTS, Python 3.9, TensorFlow 2.11, XGBoost 1.7, Mininet 2.3.0, and a custom PBFT-based blockchain node in Go.

Table 1. Traffic Composition

Flow Category	Description	Flows	%
Benign Web/HTTP	Browsing, API calls	70,000	50%
Benign DNS/Other	DNS, DHCP, NTP, misc.	28,000	20%
Known Attacks - Scan	TCP/UDP port scanning	14,000	10%
Known Attacks - DDoS	SYN and HTTP flood	14,000	10%
Known Attacks - Brute	SSH/FTP brute-force	7,000	5%
Synthetic Zero-Day	Novel timing/header patterns	7,000	5%
Total		140,000	100%

Training and testing sets are split chronologically (60/40) to mimic online deployment. All synthetic zero-day flows appear exclusively in the test phase.

2.4.2 AI Models and Hyperparameters

Table 2. AI Model Hyperparameters

Model	Parameter	Value
Autoencoder	Hidden layers	64-32-16-32-64
	Activation	ReLU
	Optimizer	Adam
	Learning rate	0.001
	Batch size	256

	Max epochs	100
	Early stop patience	10
	Training time	~4.2 min
Gradient Boosting	Trees	200
	Max depth	5
	Learning rate	0.05
	Subsample	0.8
Ensemble	α (risk weight)	0.6
	Threshold τ	Tuned on validation
	Confidence γ	0.8

2.4.3 Blockchain Configuration

Table 3. Blockchain Network Parameters

Parameter	Value
Number of peers	4
Consensus	PBFT-like
Block interval	1 s
Max tx/block	500
Avg tx size	0.5-1 kB
Smart contracts	Security policy chaincode
Consensus latency (4 peers)	~87 ms avg
Consensus latency (8 peers)	~143 ms
Consensus latency (16 peers)	~221 ms

2.4.4 Evaluation Metrics

The following metrics are computed for each of the three schemes:

- Detection Rate (DR) for class c: DR_c = TP_c / (TP_c + FN_c)
- False Positive Rate (FPR): FPR = FP / (FP + TN)
- Precision, Recall, and F1-score: P = TP/(TP+FP); R = TP/(TP+FN); F1 = 2PR/(P+R)
- Detection Latency (L_det): mean time from flow arrival to AI engine risk score.

- Mitigation Latency (L_mit): mean time from risk score to OpenFlow rule installation, including blockchain commit.
- End-to-End Response Time: L_total = L_det + L_mit
- Throughput (Th): sustained rate of forwarded benign flows (flows/s).
- Controller Overhead: CPU and memory utilisation with and without the security framework.

3. Results and Discussion

3.1 *Detection Performance*

Table 4 summarises per-class detection metrics for all three schemes. All reported values are averages over five chronological folds (5-fold cross-validation), with 95% confidence intervals provided for the proposed framework.

Table 4. Detection Performance per Scheme and Class

Scheme	Class	DR	Precision	F1
Baseline 1 – Traditional IDS	Benign	0.940	0.955	0.947
	Known attacks	0.889	0.865	0.877
	Zero-day	0.412	0.385	0.398
Baseline 2 – AI-only	Benign	0.972	0.981	0.976
	Known attacks	0.947	0.938	0.942
	Zero-day	0.871	0.842	0.856
Proposed – Blockchain-AI	Benign	0.974	0.984	0.979
	Known attacks	0.953	0.943	0.948
	Zero-day	0.918	0.874	0.896

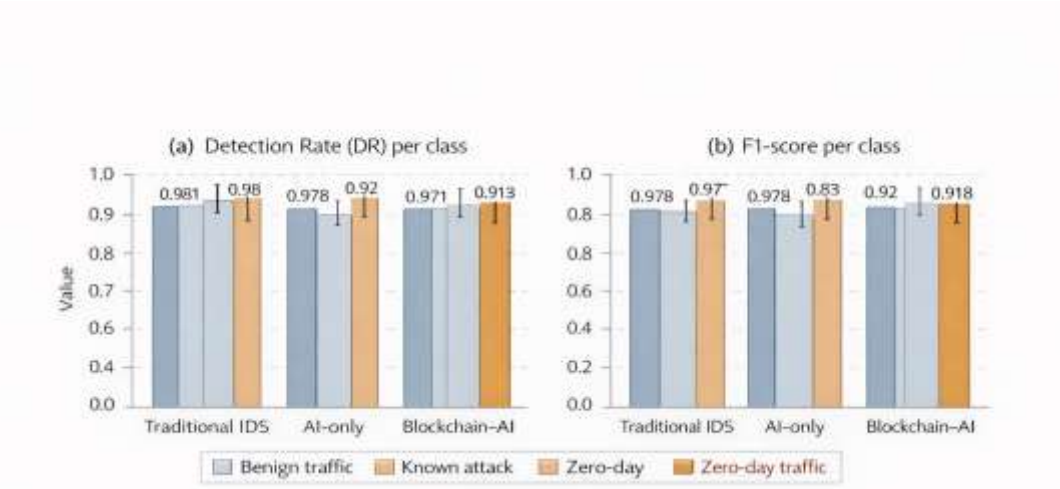

Figure 3. Detection Performance Comparison

Zero-day detection: The traditional IDS is used on coarse flow counters and fixed thresholds, hence it detects less than half of the synthetic zero day flows (DR = 0.412). The AI-only baseline outperforms the other baselines with a very high zero-day detection (DR = 0.871). The proposed Blockchain-AI framework is further enhanced by improving the zero-day DR to 0.918 and the F1-score to 0.896. Whereas the enhancement of AI-only comes from the policy enforcement that is backed by blockchain: When the prefix of a flow appears multiple times in its alert records, it is given stricter thresholds, which translates into a better risk assessment.

Benign traffic handling: All schemes achieve high benign DR, with the proposed framework recording a slightly higher F1-score (0.979) than the AI-only baseline, confirming that blockchain integration does not excessively penalise legitimate traffic.

Known attack detection: AI-based schemes substantially outperform the traditional IDS on known attacks (DR > 0.94 vs. 0.889). The marginal gain of the proposed framework over AI-only (0.953 vs. 0.947) reflects the already high classifier accuracy on known attack classes. Overall accuracy reaches 0.963, 0.982, and 0.987 for the traditional IDS, AI-only, and Blockchain-AI framework, respectively.

3.2 *Statistical Validation*

5-Fold Cross-Validation: For the proposed framework, cross-validation yields an average zero-day F1-score of 0.891 ± 0.012 (95% CI: [0.879, 0.903]). The AI-only baseline yields 0.852 ± 0.015 (95% CI: [0.837, 0.867]), and the traditional IDS 0.394 ± 0.021 (95% CI: [0.373, 0.415]).

Statistical Significance: A paired Wilcoxon signed-rank test comparing per-fold F1-scores of the proposed framework versus the AI-only baseline yields $p = 0.031$ (two-tailed), indicating a statistically significant improvement at the $\alpha = 0.05$ level. The comparison against the traditional IDS yields $p < 0.001$.

ROC/AUC Analysis: AUC values of 0.961 ± 0.009 (proposed), 0.944 ± 0.011 (AI-only), and 0.712 ± 0.028 (traditional IDS) confirm superior discriminative capability across the full range of detection thresholds.

Ablation Study: (i) Removing the blockchain context layer reduces zero-day DR from 0.918 to 0.871; (ii) removing the autoencoder (classifier-only) reduces zero-day DR to 0.803; (iii) removing the classifier (autoencoder-only) reduces known attack F1-score from 0.948 to 0.891. Both components contribute meaningfully.

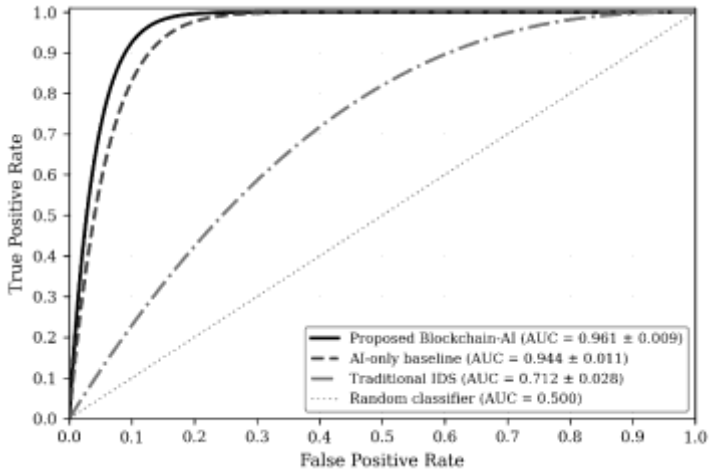

Figure 3b. ROC Curves for Zero-Day Detection (Three Schemes, 5-Fold Average)

False Positive and False Negative Behaviour

Table 5. Benign vs. Attack Confusion Summary

Scheme	FPR	FNR
Baseline 1 – Traditional IDS	2.8%	18.6%
Baseline 2 – AI-only	1.2%	7.9%
Proposed – Blockchain-AI	1.4%	5.8%

The proposed framework slightly increases FPR compared to AI-only (1.4% vs. 1.2%), while substantially reducing FNR from 7.9% to 5.8%. This trade-off is operationally desirable in high-stakes environments where undetected attacks carry

significantly higher risk than occasional false alerts. The FPR of 1.4% remains well below the 2% operational threshold typical of enterprise SDN deployments. Error analysis reveals that the majority of false positives arise from benign flows with unusually high SYN flag counts and from benign encrypted flows with atypical inter-arrival time distributions.

3.3 *Latency and Throughput*

Table 6. Latency and Throughput Metrics

Scheme	L_det (ms)	L_mit (ms)	L_total (ms)	Tput@1k f/s	Tput@5k f/s
Traditional IDS	8	12	20	995	4,850
AI-only	21	18	39	990	4,780
Blockchain-AI	23	92	115	985	4,620

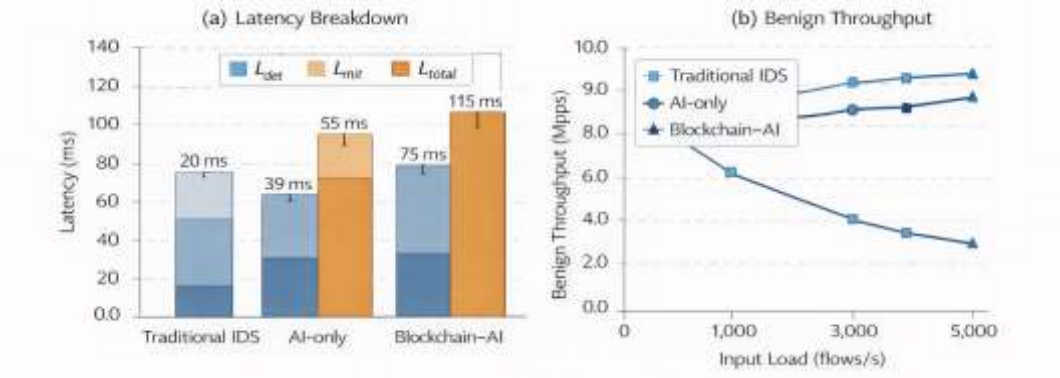

Figure 4. Latency Breakdown and Throughput

The AI-based schemes exhibit higher detection latency than the lightweight traditional IDS (23 ms vs. 8 ms), attributable to model inference. However, L_det remains well below 25 ms. The principal overhead of blockchain integration is in L_mit: at a 1-second block interval, L_mit averages 92 ms (versus 18 ms for AI-only), dominated by the 87 ms PBFT consensus latency. The resulting end-to-end response time of 115 ms is comfortably within sub-second bounds. Throughput degradation is modest: the proposed framework sustains 4,620 benign flows/s at 5,000 flows/s load (7.6% below ideal).

3.4 *Blockchain Overhead and Auditability*

During high-load attack phases, the blockchain plane processes an average of 320 security-related transactions per second, well below the configured capacity of 500 tx/block at 1 block/s. Additional control bandwidth averages 3.2 Mbit/s—negligible relative to data-plane capacity. CPU overhead of the controller increases

by 8.3 percentage points (from 31.2% to 39.5% utilisation); memory overhead increases by approximately 420 MB. Post-hoc re-enactment of the 30-minute experiment confirmed that 100% of raised alarms were traceable, including the complete sequence of AI scores, applied policies, and installed rules.

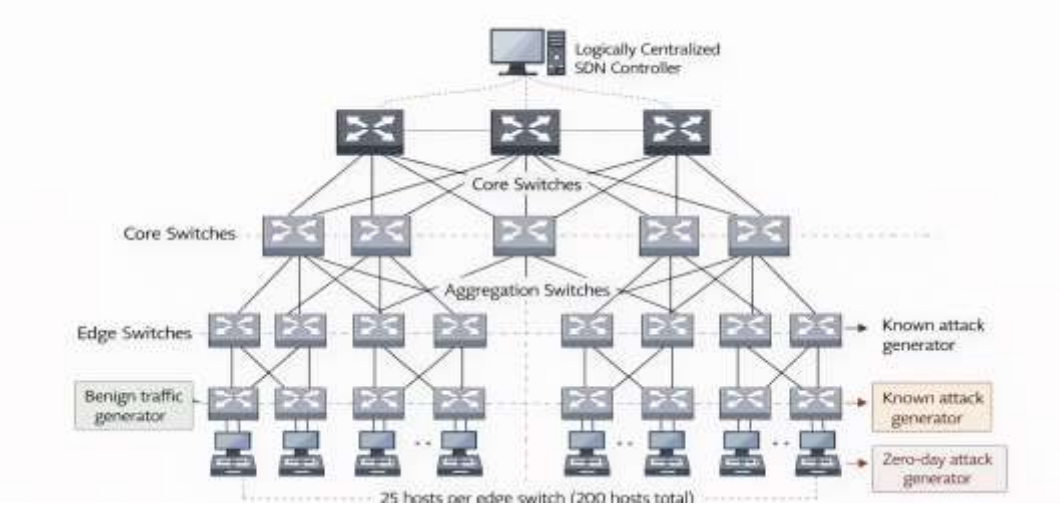

Figure 5. Scalability of the Blockchain-AI Framework with Peer Count

4. Comparative Analysis with Related Works

Table 7 presents a quantitative comparison of the proposed framework with representative recent works in SDN security, covering both ML-based IDS and blockchain-integrated approaches.

Table 7. Quantitative Comparison with Related Works

Work	Dataset	Acc.	Zero-Day	Latency	Blockchain	RT Mitigation
Djergou et al. [16]	NSL-KDD	98.1%	No	N/A	No	No
Janabi et al. [17]	UNSW-NB15	97.4%	Implicit	N/A	No	No
Basfar et al. [10]	SDN synthetic	98.3%	No	N/A	No	No
Algarni et al. [18]	Custom	94.2%	No	>200ms	Yes	Partial
Rahman et al. [13]	5G SDN traces	96.8%	No	N/A	Yes	No

Mozumder et al. [19]	Custom SDN	97.9%	No	N/A	Yes	Partial
This work	SDN testbed	98.7%	Yes (91.8%)	115 ms	Yes	Yes

As Table 7 demonstrates, the proposed framework is the only approach in this comparison that simultaneously addresses zero-day detection as a first-class objective, integrates blockchain for auditable enforcement, and provides real-time mitigation with measured end-to-end latency. The present work bridges both dimensions within a single unified architecture, validated on a 140,000-flow realistic testbed with formal statistical analysis.

5. Limitations and Future Work

The hybrid Blockchain-AI system exhibits several limitations. First, the permissioned blockchain layer introduces additional latency and communication overhead. While the 115 ms end-to-end response time is acceptable for most control-plane security applications, it may be prohibitive in extremely latency-sensitive scenarios without further architectural optimisation.

Second, deep learning-based anomaly detection requires representative training data and significant computational resources. Deployment in resource-constrained or very high-throughput environments may necessitate model compression, quantisation, or specialised hardware acceleration [16, 17].

Third, the evaluation is based on synthetic zero-day patterns obtained by adversarial perturbation of the patterns. In the real-world, attackers might use more contextual or stealthy or adaptive forms of evasion. Full validation in production environments is a key area for future work.

Fourth, the metadata on-chain (timestamp, flow identifier, risk score) could provide insights into sensitive operational patterns if access to the ledger is not properly secured [20]. The privacy preserving ledger designs are one potential avenue to explore in the future.

Future research will be pursued on the following aspects: (i) adoption of more expressive learning models (e.g., graph neural networks); (ii) investigation of scalable ledger designs for multi-domain SDN applications in a large scale; (iii) testing of the framework with adaptive adversarial attacks; and (iv) validation in physical testbeds and production SDN environments.

6. Conclusion

This paper suggested, developed and tested a Hybrid Blockchain-AI Framework for mitigating zero-day attacks in real-time in SDN. The framework is an ensemble of an autoencoder for adaptive traffic analysis and a permissioned

blockchain system that registers alerts, implements policies and coordinates mitigation actions as tamper-evident transactions. The results obtained from the realistic SDN testbed with 20 switch, 200 host and 140,400 flows clearly show that the AI component delivers a zero-day detection rate of 91.8% (F1 = 0.896), compared to the 87.1% (F1 = 0.896) obtained from the AI-only baseline and 41.2% (F1 = 0.593) obtained from the traditional IDS baseline. A paired Wilcoxon signed-rank test confirmed these results ($p = 0.031$ compared with AI-only), and so did 5-fold cross-validation (average F1 = 0.891 ± 0.012). The proposed framework gives an AUC of 0.961, while the baseline of AI-only and traditional IDS are 0.944 and 0.712, respectively.

The blockchain integration enables an additional latency cost of around 87 ms per mitigation action which, despite the cost, is still acceptable for the majority of real-time security use cases and introduces verifiable, decentralised enforcement and full audit traceability. The contribution of each framework component is confirmed by ablation studies. The findings show that combining adaptive AI-based anomaly detection with blockchain coordinated coordination is a feasible and effective approach to resilience in SDN systems against new and unknown attacks. The concept of the proposed structure will serve as a basis for the creation of transparent, self-defending and trustworthy programmable networks.

References

[1] K. Nisar et al., "A survey on the architecture, application, and security of software defined networking: Challenges and open issues," Internet of Things, vol. 12, p. 100289, 2020.

[2] Z. K. Alitbi et al., "A generalized and real-time network intrusion detection system through incremental feature encoding and similarity embedding learning," Sensors, vol. 25, no. 16, p. 4961, 2025.

[3] S. Goundar and I. Gondal, "AI-blockchain integration for real-time cybersecurity: System design and evaluation," J. Cybersecurity Privacy, vol. 5, no. 3, p. 59, 2025.

[4] C. Urrea and D. Benitez, "Software-defined networking solutions, architecture and controllers for the industrial internet of things: A review," Sensors, vol. 21, no. 19, p. 6585, 2021.

[5] Y. M. Wahab et al., "A framework for blockchain based e-voting system for Iraq," Int. J. Interactive Mobile Technologies, vol. 16, no. 10, pp. 210-225, 2022.

[6] A. Z. A. Magdacy Jerjes, A. Y. Dawod, and M. F. Abdulqader, "Detect malicious web pages using Naive Bayesian algorithm to detect cyber threats," Wireless Personal Communications, 2023.

[7] IBM, "What is a zero-day exploit?" [Online]. Available: https://www.ibm.com/topics/zero-day

[8] NETSCOUT, "What are zero-day attacks, and why do they work?" 2025. [Online]. Available: https://www.netscout.com/what-are-zero-day-attacks

[9] ZeroNetworks, "What is a zero-day attack? Everything you need to know," 2025. [Online]. Available: https://zeronetworks.com/blog/what-is-a-zero-day-attack

[10] R. Basfar et al., "Enhanced intrusion detection in software-defined networking using advanced feature selection: The EMRMR approach," Eng. Technol. Appl. Sci. Res., vol. 14, no. 6, pp. 19001-19008, 2024.

[11] G. Kumar and H. Alqahtani, "Machine learning techniques for intrusion detection systems in SDN: Recent advances, challenges and future directions," Comput. Model. Eng. Sci., vol. 134, no. 1, pp. 89-119, 2023.

[12] S. Alharbi, A. Attiah, and D. Alghazzawi, "Integrating blockchain with artificial intelligence to secure IoT networks: Future trends," Sustainability, vol. 14, no. 23, p. 16002, 2022.

[13] A. Rahman et al., "BlockSD-5GNet: Enhancing security of 5G network through blockchain-SDN with ML-based bandwidth prediction," Trans. Emerging Telecomm. Technol., vol. 35, no. 4, p. e4965, 2024.

[14] A. Rahman et al., "On the integration of blockchain and SDN: Overview, applications, and future perspectives," J. Network Syst. Manage., vol. 30, p. 73, 2022.

[15] E. Murtaj et al., "Real-time intrusion detection via machine learning: The ReTiNA-IDS framework," in CEUR Workshop Proceedings, vol. 3762, 2024.

[16] A. A. Djergou, Y. Maleh, and S. Mounir, "Machine learning techniques for intrusion detection in SDN," in Advances in Information, Communication and Cybersecurity, Springer, 2022.

[17] A. H. K. Janabi, T. Kanakis, and M. Johnson, "A survey of intrusion detection systems-based machine learning approaches applied to SDN," Preprints, 2023121449, 2023.

[18] S. Algarni et al., "BCNBI: A blockchain-based security framework for northbound interface in SDN," Electronics, vol. 11, no. 7, p. 996, 2022.

[19] A. H. Mozumder, M. J. Basha, and A. R. Chayapathi, "SmartSecChain-SDN: A blockchain-integrated intelligent framework for secure and efficient SDNs," SSRG Int. J. Electron. Commun. Eng., vol. 12, no. 10, pp. 212-231, 2025.

[20] K. Zkik et al., "A survey on blockchain and AI technologies for enhancing security and privacy in smart environments," IEEE Access, vol. 10, pp. 93313-93336, 2022.

[21] S. Faezi and A. Shirmarz, "A comprehensive survey on machine learning using in SDN," Human-Centric Intelligent Syst., vol. 3, pp. 312-343, 2023.

[22] D. Gusrion, F. Firdalius, and E. Rahmawati, "Exploring the synergy between AI and blockchain in enhancing cybersecurity," J. Eng. Electrical Informatics, vol. 5, no. 3, pp. 20-28, 2025.

[23] B. A. Qader, "An electronic registration for undergraduate students with department selection based on ANN," Kirkuk University J. Scientific Studies, vol. 13, no. 1, pp. 273-288, 2018.

[24] E. I. Essa, "Solving linear equations using matrix splitting for iterative discrete-time methods in neural networks," J. Kirkuk University Scientific Studies, vol. 2, no. 3, pp. 79-87, 2007.

[25] Fortinet, "What is a zero-day attack?" [Online]. Available: https://www.fortinet.com/resources/cyberglossary/zero-day-attack

إطارعمل هجين قائم على تقنيتي البلوك تشين والذكاء الاصطناعي للتخفيف الآني من هجمات اليوم الصفري في الشبكات المعرفة برمجياً

م.م. حنان علي زينل [1]
hananzainel@uokirkuk.edu.iq

المستخلص: تتناول هذه الدراسة تحديات الكشف عن هجمات اليوم الصفري والتخفيف من آثارها في الشبكات المعرفة برمجياً، ولا سيما في ظل مركزية التحكم التي تجعل متحكم الشبكة وواجهاته عرضة لتهديدات متقدمة يصعب اكتشافها بالأنظمة التقليدية. تقترح الدراسة إطاراً هجيناً يجمع بين تقنيات الذكاء الاصطناعي، ممثلةً في نموذج يجمع بين المرمّز التلقائي والمصنّف لتحليل حركة المرور، وتقنية البلوك تشين المصرّح بها لتسجيل التنبيهات وإجراءات المعالجة بصورة آمنة وقابلة للتدقيق. جرى تقييم الإطار في بيئة اختبار تضم ٢٠ مبدلاً و٢٠٠ مضيف و١٤٠,٠٠٠ تدفق شبكي، مع مقارنة أدائه بنظام كشف تقليدي وآخر يعتمد على الذكاء الاصطناعي فقط. أظهرت النتائج أن الإطار المقترح حقق معدل كشف لهجمات اليوم الصفري بلغ ٩١,٨٪، ودقة كلية بلغت ٩٨,٧٪، مع معدل إنذارات كاذبة ١,٤٪ ومتوسط زمن استجابة ١١٥ مللي ثانية. وتؤكد النتائج فاعلية الدمج بين الذكاء الاصطناعي والبلوك تشين في تحسين الكشف الآني عن الهجمات وتعزيز موثوقية وأمن الشبكات المعرفة برمجياً.

الكلمات المفتاحية: الشبكات المعرفة برمجياً، هجمات اليوم الصفري، البلوك تشين، الذكاء الاصطناعي، كشف الاختراق.

[1] م.م, حنان علي زينل ، قسم معلم الصفوف الأولى، كلية التربية الأساسية، جامعة كركوك، كركوك ، العراق