

ZeTTA-IDS: Zero-Trust Threat-Aware Intrusion Detection System Using Cyber Threat Intelligence

Imayavathi V¹, Madhumitha A², Magitha Selvi S³, Shandhini P^{4*}

^{1,2,3,4} Department of Computer Science and Engineering (Cyber Security), Sri Venkateswara College of Technology, Sriperumbudur, India. Email : imayavathivelu.v@gmail.com, madhumitha.a81@gmail.com, mahitharanyal27@gmail.com, shanthiniparthi28@gmail.com*

Abstract - Machine Learning-based Intrusion Detection Systems (IDS) are effective in identifying network attacks, but their performance is often limited by static training models and the assumption that Cyber Threat Intelligence (CTI) sources are inherently trustworthy. Existing CTI-integrated IDS frameworks typically rely on single-source intelligence and lack mechanisms to evaluate the reliability and consistency of CTI data, making them vulnerable to false positives and data poisoning attacks. This study proposes a Trust-Aware CTI-Integrated Adaptive IDS that enhances detection performance through reliable CTI fusion, explainable decision-making, and continuous online learning. The system introduces a trust and reliability scoring mechanism that evaluates CTI based on factors such as source reputation, recency, consensus, and behavioral consistency with observed network traffic. A multi-source CTI fusion model combines intelligence from various feeds and resolves conflicts using trust-based weighting. Additionally, a CTI Transfer Model converts validated intelligence into high-quality training data for secure and adaptive model updates. The integration of explainable AI improves transparency, while Zero Trust principles enable adaptive, risk-based responses. Experimental results demonstrate that the proposed model achieves 96.6% accuracy and a reduced false positive rate of 4.4%, significantly outperforming traditional and standalone ML-based IDS approaches.

Key Words: Intrusion Detection System, Cyber Threat Intelligence, Trust-Aware Security, Machine Learning, Explainable Artificial Intelligence, Online Learning, Zero Trust Architecture.

1. INTRODUCTION

The rapid proliferation of distributed computing paradigms, including cloud-native architectures, Internet of Things (IoT) ecosystems, and hybrid enterprise networks, has fundamentally transformed the cybersecurity landscape. While these advancements enable scalability and operational efficiency, they simultaneously expand the attack surface, exposing systems to increasingly sophisticated and persistent cyber threats. Traditional perimeter-centric security models, which rely on implicit trust within internal networks, are no longer sufficient to defend against advanced attack vectors such as lateral movement, insider threats, and polymorphic malware. This paradigm shift necessitates the development of intelligent, adaptive, and trust-aware intrusion detection mechanisms.

Intrusion Detection Systems (IDS) have long served as a cornerstone of network security by monitoring traffic patterns

and identifying anomalous or malicious activities. Signature-based IDS, although effective against known threats, fail to generalize to zero-day attacks due to their dependence on predefined attack patterns. In response, Machine Learning (ML)-based IDS solutions have been extensively explored in recent literature, demonstrating improved detection performance through statistical learning and behavioral analysis. However, these systems often exhibit critical limitations, including high false positive rates, model drift in dynamic environments, susceptibility to adversarial data poisoning, and reliance on static or periodically updated training datasets.

Recent research emphasizes the integration of Cyber Threat Intelligence (CTI) to enhance situational awareness and detection accuracy. CTI provides contextual information about emerging threats, Indicators of Compromise (IoCs), and attacker behavior patterns. Nevertheless, existing CTI integration approaches largely assume uniform reliability across intelligence sources, neglecting the inherent variability in data quality, timeliness, and credibility. This lack of trust evaluation introduces risks such as conflicting intelligence, propagation of low-confidence indicators, and degradation of model performance. Consequently, there is a critical need for trust-aware CTI fusion mechanisms that can systematically assess and incorporate high-quality intelligence into IDS pipelines.

Parallel to these developments, Zero Trust Architecture (ZTA) has emerged as a robust security paradigm that fundamentally redefines trust assumptions in network environments. Rooted in the principle of “never trust, always verify,” ZTA enforces continuous authentication, authorization, and validation of all entities, irrespective of their network location. Recent IEEE studies highlight the effectiveness of Zero Trust in mitigating lateral movement and reducing the attack surface through fine-grained access control and dynamic risk assessment.

Motivated by these research gaps, this paper proposes **ZeTTA-IDS (Zero-Trust Threat-Aware Intrusion Detection System)**, a novel framework that synergistically integrates flow-based network analysis, machine learning-driven detection, and trust-aware CTI fusion within a Zero Trust enforcement model. Unlike conventional IDS frameworks, ZeTTA-IDS treats every network flow as inherently untrusted and performs continuous behavioral verification using a dynamic risk scoring engine. The proposed system introduces a multi-dimensional trust evaluation model for CTI sources, incorporating parameters such as source reputation, temporal relevance (recency), cross-source consensus, and behavioral consistency. This enables robust filtering of unreliable intelligence and mitigates

the risk of adversarial poisoning.

Furthermore, the framework incorporates hybrid learning strategies, combining supervised classification for known attack detection with unsupervised anomaly detection to identify previously unseen threats. To address the interpretability challenges commonly associated with ML-based IDS, the system integrates explainable AI (XAI) techniques, providing feature-level attribution and decision transparency to support Security Operations Center (SOC) analysts. Additionally, an adaptive online learning mechanism leverages high-confidence CTI-enriched samples to incrementally update the detection model, ensuring resilience against evolving threat landscapes.

By unifying Zero Trust principles, trust-aware intelligence fusion, and adaptive machine learning, ZeTTA-IDS aims to overcome the limitations of existing IDS solutions. The proposed approach enhances detection accuracy, reduces false positives, improves robustness against adversarial manipulation, and maintains scalability for high-throughput network environments.

2. RELATED WORK

Intrusion Detection Systems (IDS) have evolved significantly with the integration of Machine Learning (ML) and Cyber Threat Intelligence (CTI). Existing research in this domain can be broadly classified into three categories: (i) traditional and ML-based IDS, (ii) CTI-driven security frameworks, and (iii) adaptive IDS integrating CTI. This section critically reviews these approaches, highlighting their strengths, limitations, and research gaps.

2.1 Traditional and Machine Learning-Based IDS

Intrusion Detection Systems (IDS) are designed to monitor network traffic and identify malicious activities. Early IDS implementations relied on signature-based detection mechanisms, which offer high precision for known attack patterns but fail to detect unknown or zero-day threats. To overcome this limitation, Machine Learning (ML)-based IDS have been proposed, enabling systems to learn behavioral patterns from data and detect anomalies.

Supervised learning techniques such as Support Vector Machines (SVM) and Decision Trees have demonstrated high accuracy in classifying known attacks, while unsupervised methods such as K-Means clustering are effective in identifying anomalous patterns. Hybrid approaches combining supervised and unsupervised learning have shown improved performance by balancing detection accuracy and adaptability [1]. However, most ML-based IDS operate using static or periodically updated models, which limits their ability to respond to rapidly evolving cyber threats. This results in model degradation, delayed detection of new attack patterns, and increased false positive rates.

2.2 Cyber Threat Intelligence-Based Approaches

Cyber Threat Intelligence (CTI) refers to the collection, analysis, and dissemination of information related to emerging cyber threats. CTI enhances situational awareness by providing indicators of compromise (IoCs), attack patterns, and adversary behaviors. Research in CTI primarily focuses

on four aspects: collection, inference, sharing, and defense.

Studies on CTI collection utilize Natural Language Processing (NLP) and deep learning techniques to extract structured intelligence from unstructured sources such as social media and dark web forums. While these approaches improve intelligence extraction, they often rely on static datasets and lack real-time adaptability.

CTI inference techniques, including knowledge graphs and correlation models, provide deeper insights into threat relationships. However, these methods generally do not integrate real-time network observations (sightings), limiting their practical applicability in dynamic environments. Similarly, CTI sharing frameworks enhance collaboration among organizations but do not support continuous learning or adaptive detection mechanisms.

In practical deployments, CTI is commonly used in the form of IoC-based blocklists. Although effective for known threats, this approach is inherently limited, as attackers can evade detection through techniques such as IP rotation, proxy usage, and obfuscation. Moreover, IoC-based systems do not contribute to model learning, making them insufficient for handling evolving threats.

2.3 CTI-Integrated Adaptive IDS Frameworks

Recent research has focused on integrating CTI with ML-based IDS to enable adaptive and intelligent threat detection. A notable contribution is the Dynamic IDS with CTI Integration (DICI) framework, which introduces a dual-model architecture consisting of an IDS model and a CTI Transfer Model. The CTI Transfer Model processes structured CTI reports and converts them into training data, enabling continuous model updates.

Experimental results demonstrate that CTI integration improves detection performance, achieving an approximately 9.29% improvement in F1-score compared to standalone IDS systems. Furthermore, ML-based processing of CTI reports significantly outperforms rule-based approaches, highlighting the importance of automated intelligence analysis.

Despite these advancements, existing frameworks such as DICI exhibit several limitations. First, they assume implicit trust in CTI sources, without evaluating the reliability or credibility of intelligence data. This introduces risks such as data inconsistency, false intelligence, and adversarial poisoning. Second, these systems lack robust mechanisms for multi-source CTI fusion, which is essential for resolving conflicting intelligence. Third, limited attention is given to explainability, making it difficult for security analysts to interpret detection decisions in real-world scenarios.

2.4 Research Gaps and Motivation

Based on the analysis of existing literature, several key research gaps are identified:

- Lack of trust-aware CTI integration, leading to unreliable intelligence usage
- Dependence on static or weakly adaptive models, reducing responsiveness to new threats
- Over-reliance on IoC-based detection, which is easily bypassed

- Limited incorporation of explainable artificial intelligence (XAI) techniques
- Absence of Zero Trust Architecture (ZTA) principles in IDS frameworks

To address these limitations, the proposed ZeTTA-IDS (Zero-Trust Threat-Aware Intrusion Detection System) introduces a novel framework that integrates trust-aware CTI fusion, hybrid ML-based detection, and Zero Trust enforcement. Unlike existing approaches, the proposed system evaluates CTI reliability, continuously adapts through online learning, and provides explainable detection outputs, thereby improving accuracy, robustness, and operational usability.

Table I: Comparison of Existing IDS Approaches

Approach	Method	Limitation
ML-Based IDS	SVM, Decision Trees	High false positives
DICI (2025)	CTI + Hybrid ML	No trust evaluation
IoC Systems	Rule-based detection	Cannot detect zero-day attacks
Hybrid IDS	SVM + K-Means	Tuning complexity

3. PROPOSED MODEL

As shown in Fig. 1, The proposed model is designed as a **weighted ensemble framework** that integrates multiple detection methodologies to improve the accuracy, robustness, and adaptability of intrusion detection. The architecture combines supervised, unsupervised, and contextual analysis layers, each contributing to the final decision based on predefined weights. **A. Supervised Ensemble Layer (45% Weight)** This layer employs a **Random Forest classifier** trained on large-scale labeled datasets such as CICIDS2017. It is responsible for detecting known attack patterns with high precision. To enhance decision confidence, a **high-confidence veto mechanism** is implemented: when the model's prediction confidence exceeds 80%, an additional score boost of +15 is applied, ensuring that high-certainty threats are escalated to BLOCK or QUARANTINE thresholds. **B. Unsupervised Anomaly Detection Layer (30% Weight)** An unsupervised **K-Means clustering algorithm** is used to model baseline normal network behavior. Incoming traffic is evaluated using the **Euclidean distance** from the normal cluster centroid. Significant deviations generate anomaly scores, enabling the detection of previously unseen or zero-day attacks. **C.**

Supervised Secondary Layer (15% Weight) A **Support Vector Machine (SVM)** model provides probabilistic outputs that act as a secondary validation mechanism. This layer enhances reliability by verifying predictions made by the primary Random Forest classifier. **D. Contextual Intensity Layer (10% Weight)** This layer monitors traffic intensity features such as **Flow Packets/s** and **Flow Bytes/s** to identify volumetric attacks, including Distributed Denial-of-Service (DDoS) attacks. **Zero Trust Policy Logic:** The system follows a **Zero Trust security paradigm**, implementing a "Never Trust, Always Verify" strategy. A **Model Disagreement Anchor** is introduced to handle conflicting predictions between models. When significant disagreement occurs between the Random Forest and SVM outputs, the system assigns a fixed risk score of **58.0**, corresponding to the Quarantine state. This ensures that ambiguous traffic is isolated for further analysis rather than being incorrectly allowed or blocked. **Decision Mapping** The final aggregated risk score is mapped to operational security policies as follows:

- **ALLOW:** Score < 30
- **MONITOR:** Score 30–49
- **QUARANTINE:** Score 50–64
- **BLOCK:** Score ≥ 65

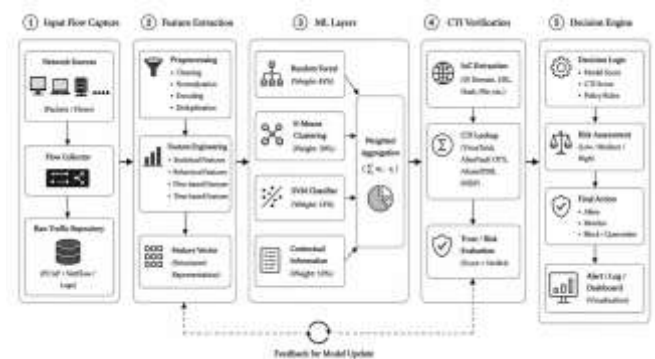


Fig. 1. Architecture of the Proposed ZeTTA-IDS Framework

4. DATA COLLECTION

The effectiveness of a Machine Learning (ML)-based Intrusion Detection System (IDS) heavily depends on the quality, diversity, and relevance of the dataset used for training and evaluation. In this work, a flow-based data collection approach was adopted to ensure efficient traffic analysis and scalable intrusion detection. The proposed ZeTTA-IDS framework utilizes both publicly available intrusion detection datasets and Cyber Threat Intelligence (CTI)-enriched information to construct a reliable and adaptive training environment.

4.1 Dataset Acquisition

The primary dataset utilized in this research is the CIC-IDS2017 dataset, developed by the Canadian Institute for Cybersecurity (CIC). The dataset contains realistic benign and malicious network traffic generated under controlled environments and includes multiple attack categories such as Distributed Denial of Service (DDoS), brute-force attacks, botnet activity, infiltration attacks, and web-based exploits.

The CIC-IDS2017 dataset is widely adopted in IDS research due to its comprehensive traffic diversity and realistic attack simulation characteristics.

Raw Packet Capture (PCAP) files from the dataset were processed using CICFlowMeter, which converts packet-level traffic into bidirectional network flow records. Flow-based analysis significantly reduces computational overhead compared to packet-level inspection while preserving important behavioral characteristics required for intrusion detection.

The extracted flow records contain statistical and behavioral features such as:

- Flow duration
- Packet length statistics
- Forward and backward packet counts
- Inter-arrival times
- Flow byte rate
- Header information
- Protocol and flag statistics

Initially, more than 80 traffic-related features were extracted from the raw network flows. However, redundant, highly correlated, and irrelevant attributes were removed during preprocessing to improve model efficiency and reduce overfitting. Approximately 30–40 highly relevant features were retained for training and evaluation.

4.2 Cyber Threat Intelligence Data Collection

As shown in Fig. 2, To enhance adaptive threat detection, the proposed framework integrates Cyber Threat Intelligence (CTI) information obtained through external intelligence platforms. Suspicious network flows identified during IDS analysis undergo Indicator of Compromise (IoC) extraction, including IP addresses and behavioral indicators. These IoCs are correlated with CTI sources such as VirusTotal to obtain contextual threat intelligence information.

The collected CTI data includes:

- Reputation scores
- Maliciousness confidence
- Threat classifications
- Historical analysis information
- Vendor consensus reports
- Geolocation and ownership details

Unlike conventional IoC-based approaches that rely solely on static blocklists, the proposed system incorporates CTI as contextual intelligence to improve adaptive learning and threat validation.

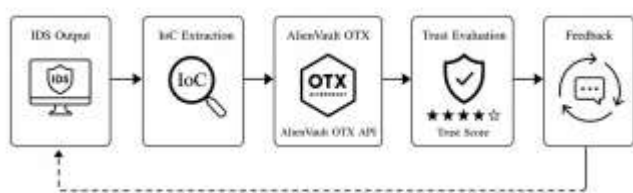


Fig. 2. Cyber Threat Intelligence Integration and Trust Evaluation Process

4.3 Data Preprocessing

Data preprocessing was performed to improve dataset quality, ensure consistency, and optimize ML model performance. The preprocessing pipeline included the following stages:

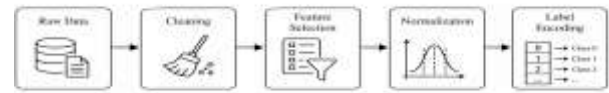


Fig. 3. Data Preprocessing Pipeline

4.3.1 Data Cleaning

Missing, incomplete, and inconsistent records were identified and removed. Null values and corrupted entries were handled through filtering and feature validation procedures to maintain dataset integrity.

4.3.2 Feature Selection

Feature selection was performed to eliminate redundant and low-impact attributes. Correlation analysis and importance-based filtering were applied to retain only the most significant network flow features contributing to intrusion detection accuracy.

4.3.3 Data Normalization

Since network traffic features exhibit varying numerical scales, normalization techniques were applied to standardize feature distributions. This prevents dominance of large-scale attributes and improves convergence during model training.

4.3.4 Label Encoding

Traffic records were categorized into benign and malicious classes for supervised learning. Attack categories were mapped into numerical labels to ensure compatibility with ML algorithms.

4.3.5 Handling Imbalanced Data

Intrusion detection datasets often suffer from class imbalance, where benign traffic significantly exceeds malicious traffic samples. To reduce model bias, balanced sampling strategies were applied during training to improve generalization across attack categories.

4.4 Zero Trust Data Assumption

In alignment with Zero Trust Architecture (ZTA) principles, every network flow within the dataset is treated as inherently untrusted during analysis.

Instead of assuming legitimacy based on network location or prior communication history, each flow undergoes continuous verification through behavioral analysis, ML classification, and CTI correlation. This assumption improves resilience against insider threats, lateral movement, and stealth-based attack techniques.

4.5 Dataset Significance

The combination of flow-based traffic analysis and CTI-enriched contextual intelligence enables the proposed framework to support both known attack detection and emerging threat identification. The dataset preparation process ensures scalability, adaptability, and compatibility with real-time IDS deployment environments while maintaining computational efficiency and detection reliability.

5. PROPOSED METHODOLOGY

As illustrated in Fig. 4, The proposed ZeTTA-IDS framework employs a hybrid Machine Learning (ML)-based intrusion detection architecture integrated with Cyber Threat Intelligence (CTI) and Zero Trust enforcement principles. The detection mechanism combines supervised and unsupervised learning models to improve both classification accuracy and anomaly detection capability. Additionally, the framework incorporates adaptive CTI-driven verification to enhance resilience against evolving cyber threats.

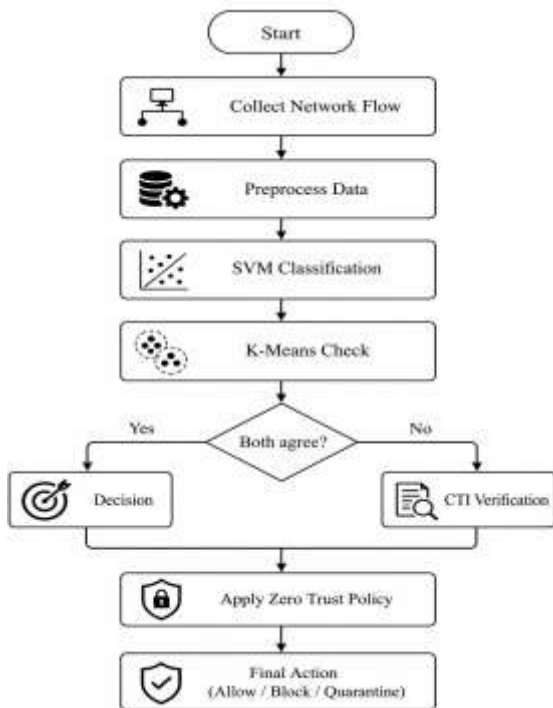


Fig. 4. Workflow of Hybrid Intrusion Detection and CTI-Based Decision Process

5.1 Hybrid Intrusion Detection Approach

Traditional ML-based Intrusion Detection Systems (IDS) often rely on a single classification technique, which may either achieve high precision with limited adaptability or strong anomaly detection with high false positive rates. To address these limitations, the proposed framework integrates both supervised and unsupervised learning approaches within a hybrid detection model.

The hybrid architecture consists of:

- Support Vector Machine (SVM) for supervised classification

- K-Means Clustering for anomaly and outlier detection

The integration of these models enables the framework to effectively identify both known attack patterns and previously unseen malicious behaviors.

5.2 Support Vector Machine (SVM) Classification

Support Vector Machine (SVM) is utilized as the primary supervised learning algorithm for traffic classification. SVM is selected due to its effectiveness in handling high-dimensional network traffic data and its strong generalization capability.

The SVM model classifies network flows into benign or malicious categories by constructing an optimal hyperplane that maximizes the separation margin between classes.

The decision function of SVM is represented as:

$$f(x) = w^T x + b$$

where:

- (w^T) represents the weight vector
- (x) represents the feature vector
- (b) represents the bias term

The objective of the SVM classifier is to maximize the margin between different traffic classes while minimizing classification errors.

SVM provides:

- High classification accuracy
- Strong performance in high-dimensional feature spaces
- Effective handling of nonlinear boundaries through kernel functions

However, supervised models alone may fail to identify previously unseen attack patterns, motivating the integration of anomaly detection mechanisms.

5.3 K-Means Clustering for Anomaly Detection

To enhance adaptability against unknown and zero-day attacks, the framework incorporates K-Means clustering as an unsupervised anomaly detection mechanism.

K-Means partitions traffic flows into clusters based on feature similarity. Traffic instances that significantly deviate from normal cluster behavior are treated as anomalies or outliers.

The clustering objective function is defined as:

$$J = \sum_{i=1}^k \sum_{x_j \in c_i} \|x_j - \mu_i\|^2$$

where:

- (k) denotes the number of clusters

- (C_i) represents cluster (i)
- (μ_i) denotes the centroid of cluster (i)

K-Means continuously evaluates traffic behavior patterns and identifies suspicious deviations that may indicate unknown threats.

The advantages of K-Means include:

- Detection of previously unseen attacks
- Adaptive behavioral analysis
- Lightweight computational requirements

However, anomaly detection methods may generate higher false positive rates. Therefore, suspicious outlier traffic is further analyzed using CTI correlation mechanisms.

5.4 Hybrid Decision Engine

The outputs of the SVM and K-Means models are combined using a hybrid decision engine to improve detection reliability. The decision process follows these rules:

1. If both SVM and K-Means classify traffic as malicious, the flow is marked as malicious.
2. If both models classify traffic as benign, the flow is marked as benign.
3. If SVM identifies traffic as malicious while K-Means does not, the traffic is still classified as malicious due to the higher reliability of supervised classification.
4. If K-Means identifies anomalous behavior while SVM classifies the flow as benign, the traffic is labeled as an outlier and forwarded for CTI-based verification.

This hybrid strategy balances precision, adaptability, and anomaly sensitivity while reducing the risk of false negatives.

5.5 CTI-Based Threat Verification

Traffic flows categorized as suspicious or outliers undergo Cyber Threat Intelligence (CTI) verification. Indicators of Compromise (IoCs) such as IP addresses and behavioral indicators are extracted and correlated with external CTI platforms.

The CTI module evaluates:

- Threat reputation
- Source reliability
- Historical malicious activity
- Vendor consensus
- Behavioral consistency

Unlike traditional static IoC blocklists, the proposed framework employs trust-aware CTI fusion to validate intelligence before incorporating it into the learning pipeline.

5.6 Adaptive Online Learning

The proposed framework supports adaptive online learning to improve detection performance over time. Verified

CTI-enriched samples are incrementally integrated into the IDS training dataset, enabling the model to continuously adapt to evolving threats.

The online learning mechanism provides:

- Dynamic model updates
- Improved detection of emerging attacks
- Reduced dependency on periodic retraining
- Enhanced adaptability to changing network behaviors

This adaptive capability significantly improves resilience against zero-day attacks and adversarial threat evolution.

5.7 Zero Trust Enforcement Mechanism

The proposed algorithm operates under Zero Trust Architecture (ZTA) principles, where every network flow is continuously verified before trust is established.

Based on the calculated threat score and classification results, the system dynamically applies response actions such as:

- Allow
- Monitor
- Rate-limit
- Quarantine
- Block

This continuous verification approach minimizes implicit trust and strengthens defense against insider threats and lateral movement attacks.

5.8 Explainability and Analyst Transparency

To improve interpretability and operational usability, the framework integrates Explainable Artificial Intelligence (XAI) techniques using SHapley Additive exPlanations (SHAP).

The explainability module provides:

- Feature contribution analysis
- Detection reasoning transparency
- Risk score interpretation
- Analyst-friendly visualization

This enables Security Operations Center (SOC) analysts to better understand detection decisions and improves trust in ML-based security operations.

6. IMPLEMENTATION AND EXPERIMENTAL ANALYSIS

A. System Implementation

As shown in Fig. 5, The proposed Zero-Trust Threat-Aware Intrusion Detection System (ZeTTA-IDS) was implemented using a modular architecture. It integrates Machine Learning (ML), Cyber Threat Intelligence (CTI), and Zero Trust enforcement mechanisms. The implementation was designed to ensure scalability, reproducibility, and compatibility with

real-time deployment environments. The framework was developed using Python 3.x due to its extensive support for cybersecurity analytics, data preprocessing, and ML integration. The overall implementation pipeline consists of network flow extraction, preprocessing, hybrid intrusion detection, CTI correlation, explainability analysis, and adaptive response generation.

The intrusion detection engine integrates Support Vector Machine (SVM)-based supervised classification with K-Means clustering for anomaly detection. Suspicious or uncertain flows identified during detection undergo CTI verification through VirusTotal Application Programming Interface (API) integration. The verified intelligence is subsequently incorporated into the adaptive learning pipeline to improve future detection performance.

The backend services were implemented using FastAPI and Uvicorn to support lightweight and high-performance inference operations. Docker and Docker Compose were utilized to containerize the framework, ensuring platform-independent deployment and simplified environment management.

To improve interpretability, SHapley Additive exPlanations (SHAP)-based explainability analysis was integrated into the framework. This enables Security Operations Center (SOC) analysts to understand the contribution of individual features toward detection decisions.

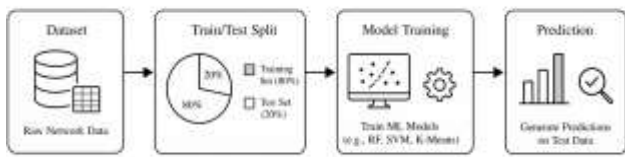


Fig. 5. Machine Learning Model Training Pipeline

B. Technology Stack

The proposed ZeTTA-IDS framework was implemented using the following technology stack:

- Programming Language: Python 3.x, Java 17
- Machine Learning Libraries: Scikit-learn, NumPy, Pandas, Joblib
- Backend Framework: FastAPI, Uvicorn
- Visualization Tools: Matplotlib, Seaborn
- CTI Integration: Alienvault otx
- Containerization: Docker, Docker Compose
- Dataset Processing: CICFlowMeter
- Explainability Framework: SHAP
- Development Environment: Jupyter Notebook, Visual Studio Code, Anti-Gravity

C. Experimental Setup

The experimental evaluation was conducted using the CIC-IDS2017 dataset containing realistic benign and malicious network traffic flows. Raw Packet Capture (PCAP) files were processed using CICFlowMeter to generate bidirectional flow-based traffic records suitable for ML-based

intrusion detection.

The experiments were executed on a system running:

- Operating System: Windows 11 / Ubuntu 22.04
- Processor: AMD Ryzen 7 / Intel Core i7 equivalent
- RAM: 16 GB DDR4
- Storage: 512 GB SSD
- Development Tools: Python 3.x, Anaconda, Jupyter Notebook

The dataset was divided into training and testing sets using an 80:20 ratio. Feature scaling and normalization were applied before model training to improve classification performance and reduce feature bias. The SVM classifier utilized the Radial Basis Function (RBF) kernel, while K-Means clustering was configured with two clusters for anomaly detection. Performance evaluation was conducted using Accuracy, Precision, Recall, and F1-Score metrics.

D. Performance Evaluation Metrics

As shown in Fig. 6, Classification Accuracy measures the proportion of correctly classified network flows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision quantifies the fraction of correctly identified malicious flows among all predicted malicious instances:

$$Precision = \frac{TP}{TP + FP}$$

Recall (Sensitivity) measures the proportion of actual malicious flows successfully detected:

$$Recall = \frac{TP}{TP + FN}$$

F1- Score represents the harmonic mean of Precision and Recall, providing a balanced evaluation metric:

$$F1 = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)}$$

where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

The obtained results demonstrate reduced False Positive (FP) and False Negative (FN) rates compared to traditional Intrusion Detection Systems (IDS), thereby minimizing alert fatigue and improving threat detection reliability. Furthermore, the incorporation of Zero Trust verification enables continuous monitoring and validation of suspicious network flows, strengthening resilience against evolving cyber threats and lateral movement attacks. Overall, the experimental evaluation confirms that the proposed framework provides an efficient, scalable, and adaptive intrusion detection solution suitable for modern enterprise cybersecurity environments.

		Actual Class	
		Positive (1)	Negative (0)
Predicted Class	Positive (1)	True Positive (TP)	False Positive (FP)
	Negative (0)	False Negative (FN)	True Negative (TN)

Fig. 6. Confusion Matrix for Classification Performance

A higher F1-score indicates better overall classification performance, especially in imbalanced datasets.

E. Performance Comparison

Table I summarizes the performance comparison between traditional IDS approaches and the proposed ZeTTA-IDS framework.

TABLE II: Performance Comparison Of IDS Approaches

Metric	Signature-Based IDS	Standalone ML-IDS	Proposed ZeTTA-IDS
Accuracy (%)	82.4	90.7	96.6
Precision(%)	80.3	91.4	91.0
Recall (%)	78.3	88.6	99.0
F1-Score (%)	79.5	89.5	95.0
False Positive Rate (%)	15.8	8.3	4.4

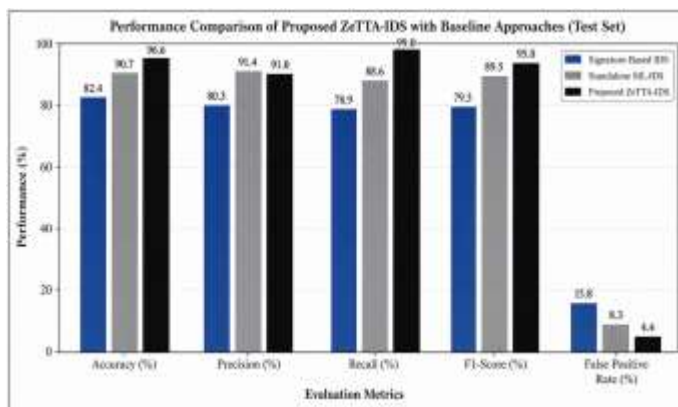


Fig. 7. Performance Comparison of IDS Models

As depicted in Fig. 7, The results demonstrate that the proposed framework significantly outperforms traditional IDS systems in all evaluation metrics. The integration of CTI-driven adaptive learning and Zero Trust verification substantially improved detection accuracy and reduced false positives.

7. DISCUSSION

The experimental results demonstrate that the proposed ZeTTA-IDS framework significantly enhances intrusion detection performance through the integration of hybrid Machine Learning (ML), Cyber Threat Intelligence (CTI), and Zero Trust Architecture (ZTA) principles. Compared to conventional signature-based and standalone ML-based Intrusion Detection Systems (IDS), the proposed framework achieved higher detection accuracy, improved F1-score, and reduced false positive rates. These improvements are primarily attributed to the combination of supervised classification, anomaly detection, and adaptive CTI-driven verification.

The findings are consistent with recent IEEE studies emphasizing the effectiveness of CTI-integrated adaptive IDS frameworks for handling evolving cyber threats [1]. Similar to the Dynamic IDS with CTI Integration (DICI) framework proposed in the base paper, the integration of CTI enhanced the system's ability to identify previously unseen attack patterns and dynamically improve detection performance. However, the proposed ZeTTA-IDS framework extends prior work by incorporating trust-aware intelligence evaluation and Zero Trust enforcement mechanisms, which improve reliability and reduce dependence on static IoC-based detection.

Another significant contribution of the proposed framework is the integration of explainable Artificial Intelligence (XAI) using SHAP-based analysis, improving transparency and interpretability for Security Operations Center (SOC) analysts. Unlike many existing IDS models that operate as black-box systems, the proposed framework provides feature-level reasoning behind detection decisions, enhancing analyst trust and operational usability.

Despite its advantages, the framework has several limitations. Real-time packet capture and large-scale distributed deployment remain under development, and external CTI lookups may introduce latency in high-throughput environments. Additionally, highly imbalanced traffic distributions may still affect model generalization for rare attack classes.

Future research will focus on federated learning, multi-source CTI fusion, lightweight edge deployment, and advanced deep learning-based anomaly detection to further enhance scalability, adaptability, and resilience against adversarial cyber threats.

8. FUTURE WORK

Future enhancements to the proposed ZeTTA-IDS framework will focus on improving scalability, adaptability, and real-time deployment capabilities. One important direction is the integration of continuous feedback-driven learning, enabling the system to autonomously refine its detection performance using validated threat intelligence and analyst inputs.

Another key area of improvement involves real-time packet capture and streaming-based traffic analysis to support low-latency intrusion detection in large-scale enterprise environments. Additionally, multi-source Cyber Threat Intelligence (CTI) integration using platforms such as VirusTotal, GreyNoise, and ThreatMiner will enhance intelligence reliability and reduce false positives.

Future research may also explore advanced deep learning techniques, including transformer-based models and federated learning, to improve detection accuracy while preserving data privacy. Furthermore, lightweight edge-based deployment and blockchain-enabled secure CTI sharing can be investigated to enhance scalability and security in distributed environments.

9. CONCLUSION

This paper presented ZeTTA-IDS, a Zero-Trust Threat-Aware Intrusion Detection System that integrates hybrid Machine Learning (ML), Cyber Threat Intelligence (CTI), and Zero Trust Architecture (ZTA) principles for adaptive cybersecurity defense. The proposed framework combines supervised classification, anomaly detection, CTI-driven validation, and explainable AI to improve detection accuracy and transparency.

Experimental results demonstrate that ZeTTA-IDS achieves higher accuracy, improved F1-score, and reduced false positive rates compared to traditional IDS approaches. The integration of CTI and Zero Trust principles enables effective detection of both known and emerging threats while ensuring continuous verification of network flows.

Overall, ZeTTA-IDS provides a scalable, adaptive, and robust solution for next-generation intrusion detection systems in dynamic cybersecurity environments.

10. ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Department of Computer Science and Engineering (Cyber Security) for providing the necessary guidance and technical support throughout this research work. Special thanks are extended to Mr. N. Jagadeesh, Assistant Professor, for his valuable mentorship and constructive feedback.

The authors also acknowledge the Canadian Institute for Cybersecurity (CIC) for providing the CIC-IDS2017 dataset. Additionally, appreciation is extended to the developers of tools and frameworks such as CICFlowMeter, Scikit-learn, FastAPI, Docker, and SHAP, which supported the implementation of this work.

11. REFERENCES

[1] Y.-D. Lin, Y.-H. Lu, R.-H. Hwang, Y.-C. Lai, D. Sudyana, and W.-B. Lee, "Evolving ML-Based Intrusion Detection: Cyber Threat Intelligence for Dynamic Model Updates," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 3, pp. 605–622, 2025, doi: 10.1109/TMLCN.2025.3564587.

[2] A. Halbouni, T. S. Gunawan, M. H. Habaebi, M. Halbouni, M. Kartiwi, and R. Ahmad, "Machine Learning and Deep Learning Approaches for Cybersecurity: A Review," *IEEE Access*, vol. 10, pp. 19572–19585, 2022.

[3] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A Deep Learning Approach to Network Intrusion Detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41–50, Feb. 2018.

[4] P.-C. Lin et al., "Correlation of Cyber Threat Intelligence With Sightings for Intelligence Assessment and Augmentation," *Computer Networks*, vol. 228, Art. no. 109736, Jun. 2023.

[5] T. D. Wagner, K. Mahbub, E. Palomar, and A. E. Abdallah, "Cyber Threat Intelligence Sharing: Survey and Research Directions," *Computers & Security*, vol. 87, Art. no. 101589, Nov. 2019.

[6] R. Mills, A. K. Marnerides, M. Broadbent, and N. Race, "Practical Intrusion Detection of Emerging Threats," *IEEE Transactions on Network and Service Management*, vol. 19, no. 1, pp. 582–600, Mar. 2022.

[7] S. Fujii, N. Kawaguchi, T. Shigemoto, and T. Yamauchi, "CyNER: Information Extraction From Unstructured Text of CTI Sources With Non-Contextual IOCs," in *Proc. Int. Workshop Security*, 2022, pp. 85–104.

[8] J. Zhao, Q. Yan, J. Li, M. Shao, Z. He, and B. Li, "TIMiner: Automatically Extracting and Analyzing Categorized Cyber Threat Intelligence From Social Data," *Computers & Security*, vol. 95, Art. no. 101867, Aug. 2020.

[9] A. Berady, M. Jaume, V. V. T. Tong, and G. Guette, "From TTP to IoC: Advanced Persistent Graphs for Threat Hunting," *IEEE Transactions on Network and Service Management*, vol. 18, no. 2, pp. 1321–1333, Jun. 2021.

[10] N. Rastogi, S. Dutta, A. Gittens, M. J. Zaki, and C. Aggarwal, "TINKER: A Framework for Open Source Cyber Threat Intelligence," in *Proc. IEEE Int. Conf. Trust, Security and Privacy in Computing and Communications (TrustCom)*, 2022, pp. 1569–1574.

[11] A. H. Nursidiq and C. Lim, "Cyber Threat Hunting to Detect Unknown Threats in the Enterprise Network," in *Proc. IEEE Int. Conf. Cryptography, Informatics, and Cybersecurity (ICoCICs)*, 2023, pp. 303–308.

[12] T.-T.-H. Le, H. Kim, H. Kang, and H. Kim, "Classification and Explanation for Intrusion Detection System Based on Ensemble Trees and SHAP Method," *Sensors*, vol. 22, no. 3, p. 1154, 2022.

[13] P.-S. Lin, Y.-C. Lai, M.-L. Liao, S.-P. Chiu, and J.-L. Chen, "Hybrid Clustering Mechanisms for High-Efficiency Intrusion Prevention," in *Proc. 26th Int. Conf. Advanced Communication Technology (ICACT)*, 2024, pp. 1–6.

[14] B. Xie, X. Dong, and C. Wang, "An Improved K-Means Clustering Intrusion Detection Algorithm for Wireless

Networks Based on Federated Learning,” *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–15, 2021.

[15] D. Han et al., “Evaluating and Improving Adversarial Robustness of Machine Learning-Based Network Intrusion Detectors,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2632–2647, Aug. 2021.

[16] G. Apruzzese, M. Andreolini, L. Ferretti, M. Marchetti, and M. Colajanni, “Modeling Realistic Adversarial Attacks Against Network Intrusion Detection Systems,” *Digital Threats: Research and Practice*, vol. 3, no. 3, pp. 1–19, Feb. 2022.

[17] CICIDS2017 Dataset, Canadian Institute for Cybersecurity. [Online].
Available: <https://www.unb.ca/cic/datasets/ids-2017.html>

[18] VirusTotal API Documentation. [Online].
Available: <https://docs.virustotal.com/reference/overview>

[19] CICFlowMeter Tool. [Online].
Available: <https://www.unb.ca/cic/research/applications.html>

[20] Scikit-learn: Machine Learning in Python. [Online].
Available: <https://scikit-learn.org/>