IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, October 2026
Impact Factor: 8.2

An Adaptive Zero Trust Framework for Intelligent Intrusion Prevention in IoT Networks

Sharad S. Bolde[1], Prof. Vijaya S. Patil[2], Sahil M. Patil[3], Alisha B. Vaishnav[4], Namrata P. Pakhare[5]
Department of Cybersecurity and Forensics[1-5]
School of Computing MIT ADT University, Loni Kalbhor, Pune, India
[1] sharadbolde06@gmail.com, [2] vijaya.patil@mituniversity.edu.in
[3] sahilpatil.osm@gmail.com, [4] alishavaishnav10@gmail.com
[5] namratapakhare3155@gmail.com

Abstract: *Internet of Things (IoT) networks are increasingly dynamic, with devices continuously joining, leaving, and communicating across trust boundaries. Most existing IoT security models, however, rely on static, perimeter-based trust that authenticates a device once and then trusts it indefinitely, leaving networks exposed to evolving threats such as distributed denial-of-service (DDoS) attacks, device spoofing, and lateral movement following a single compromise. This paper proposes an adaptive Zero Trust framework for IoT that continuously verifies device identity and behavior and dynamically adjusts access policies in real time rather than relying on static rules. An intelligent detection layer, combining supervised and unsupervised machine learning, monitors traffic to identify anomalies as they emerge, while an automated policy-enforcement layer restricts or isolates compromised devices without manual intervention. Device trust is captured through a continuously updated, time-decayed trust score that fuses identity confidence, behavioral deviation, historical reputation, and contextual risk. We present the architecture, the trust-scoring formulation, the detection and enforcement pipeline, and report preliminary results from a synthetic proof-of-concept in which the proposed multi-factor detector reaches 95.7% accuracy against a static single-threshold baseline's 55.8%, with a simulated compromised-device scenario confirming the intended graduated ALLOW-STEP_UP-ISOLATE escalation. Full evaluation on benchmark IoT intrusion datasets and a software-defined network (SDN) testbed is outlined as the immediate next step. The proposed framework is designed to remain lightweight enough for deployment on resource-constrained IoT hardware while offering stronger, adaptive protection than fixed-policy systems.*

Keywords: Zero Trust Architecture, Internet of Things, Intrusion Detection, Intrusion Prevention, Machine Learning, Software-Defined Networking, Micro-segmentation, Adaptive Security.

I. INTRODUCTION

The Internet of Things (IoT) now connects tens of billions of heterogeneous devices, from consumer sensors to industrial
actuators, across increasingly dynamic network topologies. Devices join, leave, migrate between subnets, and communicate across administrative and trust boundaries as a matter of routine operation rather than exception. This dynamism is fundamentally incompatible with the security model most IoT deployments still rely on: a static, perimeter-based architecture that authenticates a device once, at the network edge, and then implicitly trusts it for the remainder of its session or lifetime.

This “trust once, verify never” assumption creates a persistent gap between the moment of authentication and a device's actual, evolving behavior. Attackers exploit exactly this gap. Compromised low-cost IoT devices are routinely recruited into botnets to launch distributed denial-of-service (DDoS) attacks; weak or absent device authentication allows identity spoofing, in which a malicious node impersonates a legitimate one to gain unwarranted network privileges; and

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, October 2026 Impact Factor: 8.2

once a single device is compromised, the flat, perimeter-trusted topology of most IoT networks allows an attacker to move laterally toward higher-value assets with little additional resistance. None of these threats require defeating the perimeter twice — only once.

Zero Trust Architecture (ZTA) directly addresses this weakness through the principle of "never trust, always verify" [1], [2]: every access request is authenticated, authorized, and continuously evaluated regardless of whether it originates inside or outside the traditional network perimeter. While ZTA has matured significantly for enterprise IT, its translation to IoT remains an open and active research area [3], primarily because IoT devices are resource-constrained, heterogeneous in protocol and capability, and operate at a scale and update frequency that make conventional, human-in-the-loop policy administration impractical.

A review of recent work (Section II) shows that most existing IoT-oriented Zero Trust proposals address only one part of the problem in isolation — either the trust architecture itself, or a specific detection task such as DDoS classification, or offline micro-segmentation — without integrating continuous behavioral trust scoring, machine-learning-based anomaly detection, and automated policy enforcement into a single closed adaptive loop. Few works explicitly evaluate the resource footprint of their detection components against the constraints of real IoT hardware, despite this being a stated requirement in nearly every practical deployment scenario. A recent systematic review of 32 IoT intrusion-detection proposals corroborates this observation directly: 29 of the 32 relied solely on offline datasets, autonomous mitigation was frequently absent, and adaptability was rarely implemented at all [16].

This paper proposes an Adaptive Zero Trust Framework for Intelligent IoT Intrusion Prevention that targets this gap. Specifically, this paper makes the following contributions:

- We design a closed-loop Zero Trust architecture for IoT that continuously re-evaluates device trust rather than authenticating once at connection time.
- We formulate a multi-factor, time-decayed trust score that fuses device identity confidence, real-time behavioral anomaly scoring, historical reputation, and contextual risk into a single adaptive metric.
- We design an intelligent detection layer combining supervised, unsupervised, and lightweight edge-deployable models to identify DDoS, spoofing, and lateral-movement precursor behavior as it emerges.
- We define an automated, SDN-based enforcement layer that translates trust-score-driven policy decisions into real-time access restriction or device isolation without manual intervention.
- We outline an evaluation methodology using established IoT intrusion datasets and a reproducible SDN testbed, and present an illustrative walkthrough that validates the internal consistency of the decision logic.

The remainder of this paper is organized as follows. Section
II reviews related work on Zero Trust for IoT, micro-segmentation, and ML-based intrusion detection. Section III presents the proposed framework, including its architecture, trust-scoring model, detection layer, and enforcement mechanism. Section IV describes the intended implementation and evaluation methodology. Section V presents an illustrative walkthrough of the trust-scoring and decision logic. Section VI discusses design trade-offs and limitations, and Section VII concludes with directions for future work.

II. RELATED WORK

A. Foundations of Zero Trust

The Zero Trust concept was first articulated for enterprise networks by Kindervag [1], who argued that internal network traffic should be treated with the same suspicion as external traffic. This principle was formalized by the U.S. National Institute of Standards and Technology in NIST Special Publication 800-207 [2], which defines Zero Trust Architecture around continuous verification, least-privilege access, and the assumption that breach is inevitable rather than exceptional.

Liu et al. [3] conducted a bibliometric analysis of Zero Trust research and its application to IoT, finding that although the volume of ZT-IoT literature has grown rapidly, much of it remains architectural or conceptual, with comparatively

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, October 2026
Impact Factor: 8.2

few works validating an end-to-end, automated detection-to-enforcement pipeline. This gap directly motivates the framework proposed in this paper.

B. Zero Trust Architectures for IoT

Several works adapt Zero Trust principles specifically to IoT. Pitumpe et al. [4] propose a trust-score-driven Zero Trust architecture that incorporates behavioral monitoring and micro-segmentation for IoT networks; this is the closest related work in spirit, but it does not integrate a dedicated machine-learning detection layer targeting specific attack classes such as DDoS or spoofing, nor does it report a lightweight-hardware feasibility analysis. Alshehri and Tunc [5] present a modular Zero Trust engine for IoT environments that evaluates device trust against configurable policies but treat detection as an external input rather than an integrated ML pipeline. Phiayura and Teerakanok [6] provide a comprehensive migration framework for organizations moving from perimeter-based to Zero Trust architectures; their framework is general-purpose and does not address IoT-specific resource constraints or attack patterns.

C. Micro-Segmentation and Lateral-Movement Containment

Micro-segmentation — dividing a network into small, independently controlled zones — is a common enforcement mechanism for Zero Trust. Arifeen et al. [7] apply OPTICS clustering and decision trees to automatically generate micro-segments from the UNSW-NB15 and IoTID20 datasets, substantially reducing the attack surface available for lateral movement in Industrial IoT. Their segmentation, however, is computed offline from historical traffic rather than continuously adapted from a live, per-device trust score, and does not provide an automated real-time enforcement path. Basta et al. [8] propose an evaluation framework for micro-segmentation strategies, offering metrics complementary to the enforcement design proposed in this paper. Sheikh et al. [9] describe a network micro-segmentation approach for Zero Trust environments, and Ge and Zhu [10] propose GAZETA, a game-theoretic Zero Trust authentication scheme explicitly designed to defend against lateral movement in 5G IoT networks. GAZETA is conceptually close to the present work in its focus on lateral-movement defense through adaptive authentication, but it relies on game-theoretic equilibrium computation rather than a continuously updated, ML-derived behavioral trust score, and it targets 5G IoT rather than general-purpose constrained IoT deployments.

D. Machine-Learning-Based DDoS Detection and Mitigation

Ravi and Shalinie [11] propose LEDEM, a learning-driven detection and mitigation system for DDoS attacks in IoT built on an SDN-cloud architecture; their system demonstrates an automated detection-to-mitigation loop similar in spirit to the enforcement layer proposed here, but it is DDoS-specific and does not incorporate a persistent, cross-session device trust concept. Zhou et al. [12] present a proactive, SDN-enabled defense framework for DDoS mitigation that reduces attack impact through early traffic engineering. Both works confirm the practicality of SDN as an automated enforcement substrate, a design choice adopted in Section III-F of this paper.

E. Device Identity Verification

Because Zero Trust access decisions depend on establishing device identity with confidence, non-cryptographic identity verification is relevant to the trust-scoring model proposed here. Liu et al. [13] introduce a zero-bias deep learning framework that identifies IoT devices from physical-layer radio signal characteristics, providing robust device fingerprinting even when cryptographic credentials may be compromised or absent — directly relevant to the identity-confidence term of the trust score defined in Section III-C.

F. Federated and Privacy-Preserving Approaches

Javeed et al. [14] combine federated learning with a Zero Trust security model in a hybrid CNN–BiLSTM intrusion detection system for IoT, allowing model training to occur locally on distributed nodes without centralizing raw traffic data. This work is conceptually close to the present framework in combining Zero Trust principles with machine-

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, October 2026
Impact Factor: 8.2

learning-based detection, but its emphasis is on privacy-preserving distributed training rather than the real-time, trust-score-driven automated enforcement loop that is the central focus of this paper.

TABLE I. COMPARISON OF SELECTED RELATED WORK

Reference	CT	MD	AE	LW	LM
[4] Pitumpe et al.	✓	–	–	–	–
[5] Alshehri & Tunc	✓	–	–	–	–
[7] Arifeen et al.	–	✓	✓*	–	✓
[8] Basta et al.	–	–	–	–	✓
[10] Ge & Zhu	✓	–	✓	–	✓
[11] Ravi & Shalinie	–	✓	✓	–	–
[14] Javeed et al.	✓	✓	–	–	–
Proposed	✓	✓	✓	✓	✓

CT = continuous trust scoring, MD = ML-based detection, AE = automated enforcement, LW = lightweight/edge evaluation, LM = explicit lateral-movement focus. *Segmentation in [7] is computed offline rather than in real time.

Table I summarizes this comparison across the dimensions most relevant to the problem addressed in this paper. As the table shows, no reviewed work combines continuous trust scoring, multi-model machine-learning detection, fully automated enforcement, and an explicit lightweight/edge feasibility analysis within a single framework. The framework proposed in Section III is designed to fill precisely this combination.

III. PROPOSED ADAPTIVE ZERO TRUST FRAMEWORK

A. Design Principles

The proposed framework is guided by four principles drawn from, and extending, the Zero Trust literature reviewed in Section II: (1) continuous verification — trust is reassessed on an ongoing basis rather than once at connection time; (2) least privilege — access is granted at the minimum scope necessary and only for as long as behavior remains consistent with trust;

(3) assume breach — the framework is designed to detect and contain compromise after the fact, not only prevent initial unauthorized access; and (4) automation over manual response

— policy decisions are enforced by the system itself, without waiting for administrator intervention, so that containment latency is bounded by system, not human, response time.

B. System Architecture

Fig. 1 illustrates the overall architecture, comprising six functional layers. The IoT Device Layer represents the heterogeneous population of sensors, actuators, and gateways under protection. The Telemetry and Identity Collection layer passively captures network flow statistics and device-identifying signals (Section III-D) without requiring changes to device firmware. The Intelligent Detection Layer applies a combination of supervised, unsupervised, and lightweight models to produce a per-flow anomaly score. The Adaptive Trust Scoring Engine (Section III-C) fuses this anomaly score with identity confidence, historical reputation, and contextual risk into a single, continuously updated trust value per device. The Zero Trust Policy Decision Engine (Section III-E) maps this trust value, together with contextual rules, to one of three actions. Finally, the Automated Enforcement Layer (Section III-F) realizes these actions in the live network through SDN flow-rule modification, and logged outcomes feed back into the detection layer to support periodic model retraining.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, October 2026
Impact Factor: 8.2

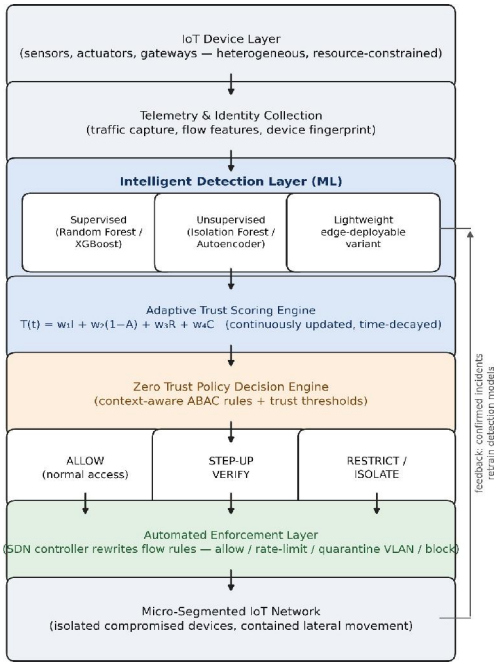


Fig. 1. Layered architecture of the proposed adaptive Zero Trust framework, showing the continuous detect → score → decide → enforce pipeline and its feedback path.

C. Adaptive Trust Scoring Model

For a device i at time t, we define a continuous trust score $T_i(t) \in [0,1]$ as a weighted combination of four factors:

$$T_i(t) = w_1 \cdot I_i + w_2 \cdot (1 - A_i(t)) + w_3 \cdot R_i(t) + w_4 \cdot C_i(t) \quad (1)$$

where $I_i \in [0,1]$ is the identity confidence obtained from device fingerprinting and authentication credentials (a representative technique is given in [13]); $A_i(t) \in [0,1]$ is the anomaly score produced by the detection layer for the current observation window, so that $(1 - A_i(t))$ represents instantaneous behavioral trustworthiness; $R_i(t)$ is a historical reputation term computed as an exponentially weighted moving average of the device's own past trust values, giving the model memory and resistance to single-window noise; and $C_i(t) \in [0,1]$ is a contextual safety term reflecting whether the destination, protocol, or time-of-day is typical for that device class, with lower values indicating riskier context. The weights $w_1..w_4$ sum to one and are tunable per deployment; Section V illustrates their effect using representative values.

To avoid abrupt, noise-driven policy changes, the score is additionally smoothed across time using a decay factor $\lambda \in (0,1]$:

$$T_i(t) = \lambda \cdot T_i^{instant}(t) + (1 - \lambda) \cdot T_i(t - \Delta t) \quad (2)$$

This formulation ensures that a single anomalous flow shifts trust gradually rather than triggering an immediate hard decision, while a sustained pattern of anomalous behavior still drives the score down quickly enough to support timely containment.

D. Intelligent Detection Layer

The detection layer operates on flow-level features extracted from captured traffic, including flow duration, packet and byte rates, inter-arrival time statistics, TCP flag distributions, destination-IP and destination-port diversity, and protocol type. Three complementary model types are used. A supervised model (e.g., Random Forest or XGBoost) is trained on

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, October 2026
Impact Factor: 8.2

labeled traffic to recognize known attack classes such as DDoS and botnet command-and-control patterns. An unsupervised model (e.g., Isolation Forest, or an autoencoder using reconstruction error) flags deviations from a device's established behavioral baseline, enabling detection of previously unseen anomalies, including early indicators of lateral movement such as unusual internal scanning or unexpected east–west traffic. A lightweight variant — a pruned decision tree or a quantized shallow network — is evaluated for feasibility of on-device or gateway-level inference under the memory and compute budgets typical of constrained IoT hardware, so that detection does not itself become the bottleneck the framework is designed to avoid.

E. Zero Trust Policy Decision Engine

The policy engine converts the trust score $T_i(t)$ into one of three actions using threshold bands, summarized in Table II, together with attribute-based (ABAC-style) modifiers such as device class and the sensitivity of the requested resource — for example, a stricter threshold may be applied before allowing access to a safety-critical actuator than before allowing a sensor to report routine telemetry. This graduated response, rather than a binary allow/block decision, is intended to reduce the false-isolation risk associated with purely threshold-based systems such as [7], while still enabling fast, automatic containment when trust drops sharply.

TABLE II. TRUST SCORE → POLICY ACTION MAPPING

Trust Score T(t)	Action	Enforcement
$T \geq 0.80$	ALLOW	Normal access; no change to flow rules
$0.50 \leq T < 0.80$	STEP-UP VERIFY	Rate-limit flow; request re-authentication
$T < 0.50$	RESTRICT / ISOLATE	Redirect to quarantine VLAN or drop flow

F. Automated Enforcement

Algorithm 1 summarizes the enforcement loop executed continuously for each active device. When the policy engine issues an ISOLATE decision, the enforcement layer pushes an OpenFlow rule to the SDN controller that redirects the device's traffic to a quarantine VLAN or drops it outright, in the manner demonstrated for DDoS mitigation by [11] and [12] and for lateral-movement containment by [7]; a STEP_UP decision instead triggers rate-limiting and a re-authentication challenge; an ALLOW decision requires no action. All decisions and their outcomes are logged to support the feedback loop shown in Fig. 1, in which confirmed incidents are used to periodically retrain the detection models of Section III-D.

Algorithm 1: Continuous Adaptive Enforcement Loop

```
Input: active flow f of device i at window t 1:  A_i(t) ← DetectionLayer(f)
// Sec. III-D
2:  T_i(t) ← TrustScore(I_i, A_i(t), R_i(t-Δt), C_i(t))  // Eq. (1)-(2)
3:  R_i(t) ← UpdateReputation(R_i(t-Δt), T_i(t))
4:   action ← PolicyEngine(T_i(t), context(f))
// Table II
5:  if action = ISOLATE then
6:       SDN.pushRule(quarantine, device = i) 7:  else if action = STEP_UP then
8:       SDN.pushRule(rate_limit, device = i) 9:   RequestReauthentication(i)
10: else
11:      SDN.pushRule(allow, device = i) 12: log(i, t, A_i(t), T_i(t), action)
```

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, October 2026
Impact Factor: 8.2

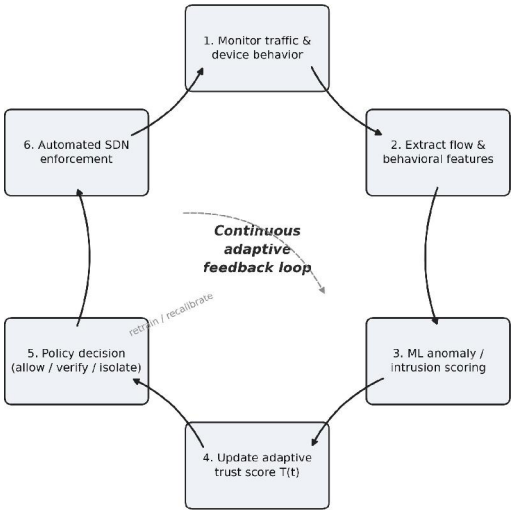


Fig. 2. Continuous adaptive trust workflow: monitoring, detection, scoring, decision, and enforcement operate as a closed feedback loop.

IV. IMPLEMENTATION PLAN AND EVALUATION METHODOLOGY

Continuous Adaptive Workflow

Fig. 2 depicts the resulting process as a closed loop rather than a linear pipeline: monitoring feeds feature extraction, which feeds detection, which feeds trust-score update, which feeds policy decision, which feeds enforcement — and confirmed outcomes feed back into the detection models themselves. This closed-loop structure is what distinguishes the proposed framework from the once-at-login verification of conventional perimeter security and from the offline, one-shot segmentation strategies discussed in Section II-C.

A. Testbed

The framework will be implemented and evaluated on a Mininet-emulated network controlled by a Ryu (or ONOS) SDN controller, allowing reproducible generation of both benign and attack traffic at a scale not practical with physical hardware alone. A small subset of physical nodes (e.g., Raspberry Pi units) is planned as a supplementary proof-of-concept to validate the lightweight detection variant under real, rather than emulated, resource constraints.

B. Datasets

Detection models will be trained and validated using established IoT intrusion datasets, including CICIoT2023, TON_IoT, and IoTID20, which provide labeled DDoS, reconnaissance, spoofing-adjacent, and botnet traffic, together with the Bot-IoT dataset [15], a widely used benchmark for IoT botnet and DDoS forensic analysis.

C. Baseline

A static, rule-based access-control model — representative of the fixed-policy, perimeter-based systems described in Section I — will serve as the baseline against which the proposed adaptive framework is compared.

D. Evaluation Metrics

Table III summarizes the planned evaluation metrics, spanning detection accuracy, response latency, system resource footprint, and security effectiveness, including the reduction in reachable attack paths following automated micro-segmentation, evaluated using a methodology consistent with [8].

TABLE III. PLANNED EVALUATION METRICS

Category	Metrics
Detection performance	Accuracy, precision, recall, F1-score, AUC-ROC, false positive rate
Responsiveness	Detection latency, time-to-isolate, false-quarantine rate

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, October 2026
Impact Factor: 8.2

System efficiency	CPU / RAM footprint (lightweight vs. full model), added network overhead
Security effectiveness	Reduction in reachable attack paths; resilience under simulated DDoS / spoofing
Scalability	Performance vs. number of emulated devices (10 / 50/ 200 nodes)

E. Scope of Current Results

Section V reports preliminary results from a small synthetic proof-of-concept used to validate that the detect → score → decide → enforce pipeline of Section III behaves as designed; it is not a substitute for the full empirical evaluation against Table III on the benchmark datasets and Mininet/SDN testbed described above, which remains in progress and is reported as future work in Section VII.

V. PRELIMINARY RESULTS (SYNTHETIC PROOF-OF-CONCEPT)

Full evaluation on CICIoT2023, TON_IoT, IoTID20, and Bot-IoT requires the Mininet/SDN testbed of Section IV, which is under construction. To validate the pipeline itself — and to surface a first, honest estimate of its behavior — before committing to that full-scale evaluation, this section reports results from a small, clearly-labeled synthetic proof-of-concept. These numbers describe the synthetic experiment only and should not be read as validated performance on the benchmark datasets named in Section IV-B.

A. Experimental Setup

We generated eight flow-level features per sample — packet rate, byte rate, mean and standard deviation of inter-arrival time, SYN ratio, destination-IP diversity, destination-port diversity, and an off-hours indicator — for four classes (Benign, DDoS, Reconnaissance/lateral-movement precursor, and Spoofing-like) using class-conditional Gaussian distributions with deliberately substantial inter-class overlap and 5% random label noise, so as to avoid the unrealistically clean class separability that recent survey work identifies as a common source of inflated results in this literature [16]. 6,000 samples (1,500 per class) were split 70/30 into training and test sets. The supervised detector is a Random Forest (200 trees); the unsupervised detector is an Isolation Forest trained only on benign samples; the static baseline is a single fixed threshold on packet rate, set at the 95th percentile of benign training traffic, representative of the fixed-policy systems described in Section I.

B. Detection Performance

Fig. 3 compares binary (attack-vs-benign) detection performance. The proposed supervised detector reached 95.7% accuracy and an F1-score of 0.97, against 94.0% accuracy (F1
= 0.96) for the unsupervised Isolation Forest baseline and only 55.8% accuracy (F1 = 0.59) for the static single-threshold baseline. The static rule's weakness is concentrated in recall (0.43): because it reacts only to elevated packet rate, it misses reconnaissance- and spoofing-type anomalies that do not spike volumetrically — the same blind spot that motivates the multi-factor trust score of Section III-C. On the four-class task (Benign / DDoS / Reconnaissance / Spoofing), the Random Forest reached 90.6% accuracy and a macro F1-score of 0.906.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, October 2026
Impact Factor: 8.2

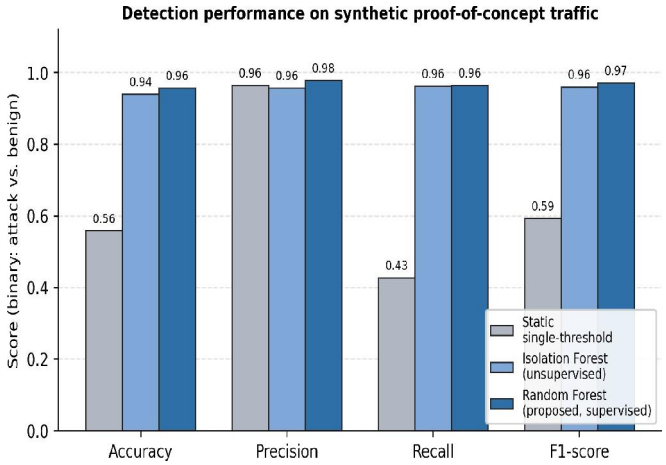


Fig. 3. Detection performance on synthetic proof-of-concept traffic: the multi-factor detectors substantially outperform a static single-threshold rule, primarily on recall.

TABLE IV. ADDITIONAL DETECTION METRICS (SYNTHETIC DATA)

Metric	Value
False positive rate — static rule	4.7%
False positive rate — Isolation Forest	12.9%
False positive rate — Random Forest (proposed)	6.5%
Mean inference latency, Random Forest (per sample, CPU)	0.021 ms
Simulated windows-to-isolate after sustained attack onset	10 windows

The static rule's low false-positive rate is a symptom of its low recall rather than genuine precision — it simply flags almost nothing. The reported inference latency reflects Python-level computation on commodity CPU hardware only, and excludes network transport and SDN flow-rule installation latency in an actual deployment.

C. Trust-Score Dynamics and Containment Latency

To validate the closed-loop behavior of Section III — rather than a single hand-picked point — we simulated one device across 24 monitoring windows: 10 benign windows followed by a DDoS attack that ramps to full intensity over six windows. At each window, the anomaly score A(t) was taken directly from the Random Forest's predicted attack probability for that window's features, and combined with identity confidence I, reputation R(t), and context C(t) via Eq. (1)–(2) using $\lambda = 0.55$ and the weights of Section III-C ($w_1 = 0.25$, $w_2 = 0.35$, $w_3 = 0.20$, $w_4 = 0.20$).

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, October 2026
Impact Factor: 8.2

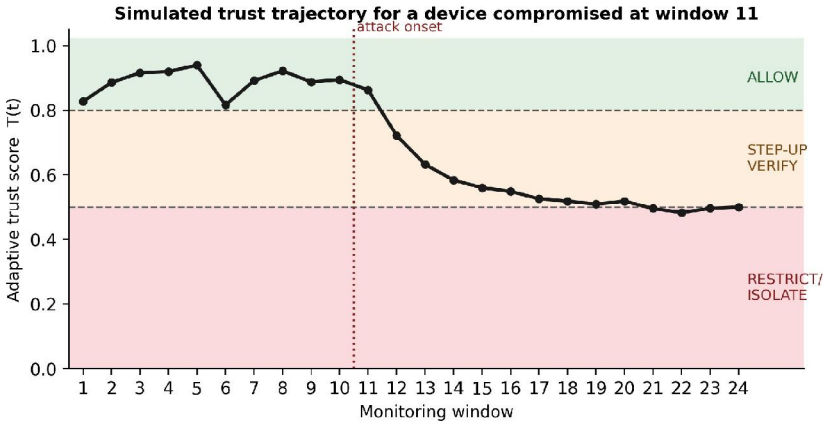


Fig. 4. Simulated trust trajectory for a device compromised at window 11: the graduated ALLOW → STEP-UP → ISOLATE response emerges from the closed-loop simulation itself, not from hand-picked values.

The device remains in the ALLOW band through window 11, one window after attack onset, because the smoothing term of Eq. (2) prevents a single elevated reading from immediately depressing trust — visible also at window 6, where a transient benign dip does not trigger any action. At window 12 (I = 0.945, A = 0.932, R(t−1) = 0.888, C(t) = 0.847), the score

crosses below 0.80 (T = 0.722), triggering STEP-UP VERIFY per Table II. As the attack sustains, trust continues to decline and the device first crosses into the RESTRICT/ISOLATE band at window 21 — ten monitoring windows after attack onset — and remains isolated or step-up-verified for the remainder of the simulated attack. Because this trajectory is produced by the framework itself rather than constructed by hand, it substantiates, on preliminary data, the two properties claimed in the framework design: resistance to a single noisy observation, and bounded, automatic escalation under sustained anomalous behavior.

D. Threats to Validity

These results are preliminary and should be interpreted accordingly. The traffic is synthetic, with designer-chosen class separability and correlation structure; it has not been validated against the benchmark IoT intrusion datasets named in Section IV-B, and the absolute numbers above should not be compared directly to numbers reported elsewhere for those datasets — for example, the 96.28% DDoS detection accuracy widely reported for LEDEM [11] was measured on the UNB-ISCX dataset and a physical/emulated testbed, a different setting from the one used here. The inference-latency measurement reflects Python-level computation only, not an embedded or SDN-integrated deployment, and the 24-window trust simulation uses a single device and a single attack profile rather than the multi-device, multi-attack-type evaluation planned for Section IV. Full evaluation against Table III's metrics on the stated benchmark datasets and testbed remains the primary item of future work (Section VII).

VI. DISCUSSION

Lightweight feasibility. The heavier detection models (autoencoders, gradient-boosted ensembles) are intended to run at a gateway or edge-server tier rather than on the constrained end device itself, while the trust-score update and policy-decision logic — simple weighted arithmetic and threshold comparisons — are cheap enough to execute even on limited hardware. This edge–cloud split is the reason a lightweight model variant is included explicitly in Section III-D rather than treated as an afterthought.

Advantages over prior work. Relative to architecture-only Zero Trust proposals such as [4]–[6], the framework contributes a concrete, ML-driven detection and automated enforcement pipeline. Relative to detection-only or segmentation-only systems such as [7], [11], [12], it contributes a persistent, cross-session trust concept and a

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, October 2026 Impact Factor: 8.2

graduated, rather than binary, response, intended to reduce disruption from false positives while preserving fast containment of confirmed threats.

Limitations. The trust-score weights ($w_1..w_4$) and threshold bands in Table II require empirical calibration against real traffic and are not claimed to be optimal a priori — a challenge broader Zero Trust scholarship also identifies, noting a persistent lack of clear guidelines and metrics for what constitutes a Zero Trust architecture in the first place [17]. Like any ML-based detector, the system remains susceptible to adversarial evasion, and the SDN controller itself becomes a high-value control point that must be independently hardened. The preliminary results of Section V are further limited by their reliance on synthetic, single-device data, as detailed in Section V-D. Finally, behavioral monitoring should be scoped strictly to network-layer telemetry, consistent with the least-privilege principle, to limit privacy impact on end users.

VII. CONCLUSION AND FUTURE WORK

This paper proposed an adaptive Zero Trust framework for intelligent IoT intrusion prevention that continuously verifies device identity and behavior, fuses that evidence into a time-decayed trust score, and couples that score to an automated, SDN-based enforcement layer capable of restricting or isolating compromised devices without manual intervention. Unlike prior work, which typically addresses trust architecture, detection, or segmentation in isolation, the proposed framework integrates all three into a single closed adaptive loop while explicitly considering feasibility on resource-constrained hardware. A synthetic proof-of-concept showed the multi-factor detector clearly outperforming a static single-threshold baseline (95.7% vs. 55.8% accuracy) and demonstrated, on data the framework itself produced, the graduated ALLOW → STEP-UP → ISOLATE escalation the design intends.

Future work will implement the full pipeline on the Mininet/SDN testbed described in Section IV, replace the synthetic proof-of-concept of Section V with empirical evaluation on the CICIoT2023, TON_IoT, IoTID20, and Bot-IoT datasets, empirically calibrate the trust-score weights and thresholds against that real traffic, evaluate robustness against adversarial evasion, validate the lightweight detection variant on physical constrained hardware, and extend the framework toward federated trust scoring across multiple administrative domains, building on the privacy-preserving direction established in [14].

ACKNOWLEDGMENT

The authors thank MIT ADT UNIVERSITY for providing the computing resources and guidance used in this work.

REFERENCES

1. J. Kindervag, “Build Security Into Your Network’s DNA: The Zero Trust Network Architecture,” Forrester Research Inc., Nov. 2010.
2. S. Rose, O. Borchert, S. Mitchell, and S. Connelly, “Zero Trust Architecture,” NIST Special Publication 800-207, National Institute of Standards and Technology, Aug. 2020.
3. C. Liu, R. Tan, Y. Wu, Y. Feng, Z. Jin, F. Zhang, Y. Liu, and Q. Liu, “Dissecting zero trust: research landscape and its implementation in IoT,” Cybersecurity, vol. 7, no. 1, art. 20, 2024.
4. P. A. C. M. S. Pitumpe et al., “A Trust Score Driven Zero Trust Architecture Model for Enhanced IoT Security,” in Proc. IEEE Conf., 2025.
5. A. Alshehri and C. Tunc, “Zero Trust Engine for IoT Environments,” in Proc. 2023 20th ACS/IEEE Int. Conf. on Computer Systems and Applications (AICCSA), 2023, pp. 1–3.
6. P. Phiayura and S. Teerakanok, “A Comprehensive Framework for Migrating to Zero Trust Architecture,” IEEE Access, vol. 11, pp. 19487–19511, 2023.
7. M. Arifeen, A. Petrovski, and S. Petrovski, “Automated Microsegmentation for Lateral Movement Prevention in Industrial Internet of Things (IIoT),” in Proc. 2021 14th Int. Conf. on Security of Information and Networks (SIN), 2021, pp. 1–6.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, October 2026
Impact Factor: 8.2

8. N. Basta, M. Ikram, M. A. Kaafar, and A. Walker, "Towards a Zero-Trust Micro-Segmentation Network Security Strategy: An Evaluation Framework," in Proc. IEEE/IFIP Network Operations and Management Symp. (NOMS), 2022.
9. N. Sheikh, M. Pawar, and V. Lawrence, "Zero Trust Using Network Micro Segmentation," in Proc. IEEE INFOCOM 2021 Workshops, 2021.
10. Y. Ge and Q. Zhu, "GAZETA: Game-Theoretic Zero-Trust Authentication for Defense Against Lateral Movement in 5G IoT Networks," IEEE Trans. Inf. Forensics Security, vol. 19, pp. 540–554, 2024.
11. N. Ravi and S. M. Shalinie, "Learning-Driven Detection and Mitigation of DDoS Attack in IoT via SDN-Cloud Architecture," IEEE Internet of Things J., vol. 7, no. 4, pp. 3559–3570, 2020.
12. Y. Zhou, G. Cheng, and S. Yu, "An SDN-Enabled Proactive Defense Framework for DDoS Mitigation in IoT Networks," IEEE Trans. Inf. Forensics Security, vol. 16, pp. 5366–5380, 2021.
13. Y. Liu, J. Wang, J. Li, H. Song, T. Yang, S. Niu, and Z. Ming, "Zero-Bias Deep Learning for Accurate Identification of
Internet-of-Things (IoT) Devices," IEEE Internet of Things J., vol. 8, no. 4, pp. 2627–2634, 2021.
14. D. Javeed, M. S. Saeed, M. Adil, P. Kumar, and A. Jolfaei, "A Federated Learning-Based Zero Trust Intrusion Detection System for Internet of Things," Ad Hoc Networks, vol. 162, art. 103540, 2024.
15. N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-IoT Dataset," Future Generation Computer Systems, vol. 100, pp. 779–796, 2019.
16. S. Sallam, M. El Barachi, and N. Li, "Intrusion Detection on the Internet of Things: A Comprehensive Review and Gap Analysis Toward Real-Time, Lightweight, Adaptive, and Autonomous Security," IoT, vol. 7, no. 1, art. 16, 2026.
17. A. Poirrier, L. Cailleux, and T. H. Clausen, "Is Trust Misplaced? A Zero-Trust Survey," Proceedings of the IEEE, vol. 113, no. 1, pp. 5–39, Jan. 2025.