Computer Fraud and Security
ISSN (online): 1873-7056

AI-Enabled Zero-Trust Security Architecture at Network Edge

Naveen Kumar
Independent Researcher, USA

Abstract

Enterprise network security has rapidly transitioned from customary castle-and-moat architectures to Zero-Trust network security architectures that focus on identity-based security controls for distributed computing environments. In particular, the emergence of Internet of Things (IoT) devices and cloud computing, together with the associated threat landscape, render customary castle-and-moat security architectures as well as trusted networks and trusted devices obsolete. As a result, constant authentication and authorization of each access request based on a wide-ranging context are required instead. Edge computing infrastructure can implement Zero Trust using distributed policy enforcement points with low-latency response and control. Artificial intelligence-based deep learning architectures, such as recurrent neural networks and transformer architectures, can autonomously detect threats at the edge by monitoring and analyzing network and device behavior during system operation in real-time. API-based security orchestration solutions allow security policies to be dynamically updated in response to incidents in the underlying network infrastructure, resulting in incident response time frames in the sub-second range. Advanced Persistent Threats targeting SCADA and ICS systems require alternative approaches such as game-theoretic modeling combined with physical system monitoring. When combined with Zero-Trust principles deployed at the network edge, ML can enable a distributed, ultra-low latency security fabric that is able to detect advanced and evasive attacks and that is suited to latency-sensitive applications. Physical system detection algorithms were shown to be effective against integrity attacks on IACS with low false positives. Protocol-specific intrusion detection systems using deterministic finite automaton approaches can achieve perfect traffic modeling and detect anomalies in the field. The security of the modern distributed enterprise against new-age cyber threats has been established by the convergence of Zero-Trust architecture, edge computing, artificial intelligence, and automated security orchestration.

Keywords: Zero-Trust Architecture, Edge Computing Security, Artificial Intelligence Threat Detection, Advanced Persistent Threats, Network Security Automation

1: From perimeter-based security to identity-centric security

In the last couple of decades, enterprise network security has changed from one big logical space to a highly distributed collection of components. Earlier network security architectures were based on the "castle-and-moat" concept, which is based on stateful firewalls, VPN concentrators, and intrusion detection/prevention systems at network boundaries.

The prevalence of cloud computing platforms and Internet of Things devices makes those assumptions effectively obsolete. The number of connected devices exceeded the number of people on Earth in 2011, and the number of connected devices worldwide is projected to reach 9 billion by 2014 and well over 24 billion by 2020 [1]. Connected devices are estimated to represent a $1.3 trillion opportunity for mobile network operators in the verticals of health, automotive, utilities, and consumer electronics [1]. A total of 2.6 million IoT devices were installed in 2016, an increase of around 30 percent compared with the previous year [2].

Perimeter-based security is intrinsically vulnerable due to its implicit trust model after initial authentication, which gives attackers a broad range of lateral movement privileges within trusted zones. A common attack vector today is to obtain an initial foothold through credential theft, privilege escalation, and lateral movement from that foothold to high-value targets. Contemporary application and data environments may exist and be accessible across multiple public cloud service providers, private data centers, edge computing locations, and employees working on the go, eliminating a defined network perimeter for security controls to protect.

The security model should transition from location-based trust to identity-based continuous verification, where every access request is authenticated and authorized based on a collection of contextual parameters, such as the verified identity of the user, security posture of the device, sensitivity of the requested resource, threat intelligence, and regulatory compliance requirements. It represents a shift from a static, perimeter-based model to a dynamic, context-aware, and continuous access control through the life cycle of a resource.

2: Zero-Trust Architecture Implementation at Network Edge

The Zero-Trust security model represents concepts such as "never trust, always verify" through continuous authentication, least-privilege access rights management, and continuous monitoring of all devices, users, and servers in an enterprise network. According to NIST Special Publication 800-207—Zero Trust Architecture, the National Institute of Standards and Technology defined ZTA as an enterprise cybersecurity architecture that eliminates the concept of implicit trust in any entity or network location by continuously validating security and contextual information before granting access to resources [3]. The model requires that all requests for access be evaluated through policy against contextual factors, such as user identity, device health, application and data sensitivity, and threat level.

Core principles of a zero-trust (ZT) approach include network microsegmentation to limit lateral movement and reduce the blast radius of a breach. Zero trust policy engines conduct real-time contextual evaluation of multiple elements, such as authenticated user identity via MFA, device trustworthiness via endpoint detection and response (EDR), application sensitivity via data classification frameworks, real-time network behavior analysis for anomalous traffic patterns, and threat intelligence feeds for active exploitation campaigns. Pervasive policy enforcement points across the network infrastructure provide fine-grained access control that ranges from network-layer filtering, through application-layer authorization, to data-level access control [3].

However, multiple issues with performance and availability make zero trust difficult to implement. When policy decision points are centralized, and policy evaluation engines are placed in a cloud data center that is physically distant from the users, performance is hampered by the network latency of multiple round-trip. These delays, however, are unacceptable for applications such as real-time communications, industrial control systems, and interactive user interfaces, in which performance is critical.

Edge computing infrastructure addresses these limitations by providing computational, storage, and communication resources in the proximity of users and connected devices at the edge of the network. Multi-access edge computing (MEC) is a standard defined by the European Telecommunications Standards Institute that enables cloud computing and service environments at the edge of the network for ultra-low latency and high bandwidth communications. Fifth-generation (5G) cellular networks provide the connectivity for large-scale edge computing architectures, particularly for ultra-reliable low-latency communication use cases with end-to-end 5G latency of around 1-4 milliseconds [4]. In this architecture, Zero-Trust maps from a control plane to a security fabric of policy enforcement points distributed across the topology.

Zero-Trust Component	Implementation Method	Security Function
Network Microsegmentation	Segmented network zones	Limit lateral movement
Authenticated User Identity	Multi-Factor Authentication (MFA)	Verify user legitimacy
Device Trustworthiness	Endpoint Detection and Response (EDR)	Assess device security posture
Application Sensitivity	Data Classification Frameworks	Determine resource criticality
Network Behavior Analysis	Real-time Traffic Monitoring	Detect anomalous patterns
Threat Intelligence	Active Exploitation Campaign Feeds	Identify the current threat landscape
Policy Enforcement Points	Distributed Network Infrastructure	Multi-layer access control

Table 1: Zero-Trust Core Principles and Policy Evaluation Components [3]

3. AI Integration for Autonomous Threat Detection in Zero-Trust Architectures

AI and ML technologies help realize this by automatically assessing trust levels at the edge by analyzing user, network, and device behaviors. ML models deployed at edge locations analyze network telemetry data packets, app-layer protocol communications, authentication event logs, and contextual signals from network components to make a zero-trust

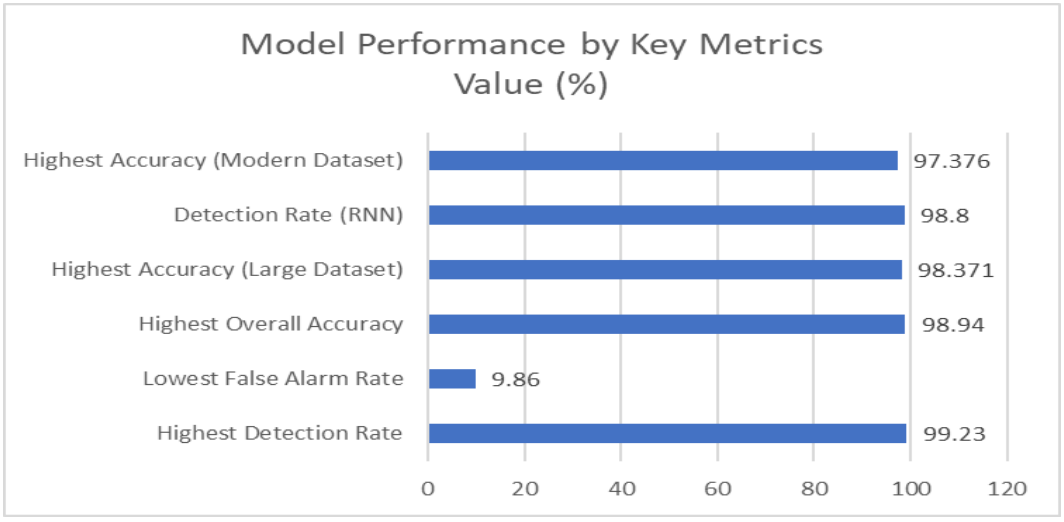
decision in real time. This enables the detection of subtle indicators of compromise that cannot be detected using customary signature-based defenses.

Deep learning models, including recurrent neural networks (RNNs) and the transformer model, are able to learn time dependencies and temporal patterns in time series data and are used for intrusion detection. For instance, an RNN was assessed in the KDD Cup 1999 dataset and achieved a 98.8% detection rate of all attacks at time step size = 100, batch size = 50, and epochs = 500 [6]. Long short-term memory networks have also achieved a 99.23% detection rate, a 9.86% false alarm rate, and 98.94% accuracy on the NSL-KDD dataset [6].

The maximum accuracy achieved by a convolutional neural network (CNN) on the CSE-CIC-IDS2018 dataset was 97.376%, using 100 hidden nodes and a learning rate of 0.5 [6]. On the Bot-IoT dataset, which has over 72 million records, a deep neural network has achieved a maximum accuracy of 98.371% using 100 hidden nodes and 0.5 as the learning rate, while for Random Forests, Naive Bayes, and Support Vector Machines, the accuracy was lower [6].

The effectiveness of AI-based threat detection depends on the amount and quality of the training data. AI engines deployed at the edge ingest from multiple data sources such as network flow logs, authentication logs, API invocation data, endpoint security, and infrastructure health. By cross-correlating these signals, machine learning can identify advanced attacks that appear harmless in isolation but, when viewed across multiple dimensions, indicate the presence of malicious behavior.

Explainable AI attempts to make the underlying workings of complex machine learning algorithms more understandable so security analysts can verify detected threats and perform threat hunting on operational deployments.



Graph 1: Comparative Performance Metrics Across Deep Learning Models [6]

4. API-Driven Security Orchestration and Automation

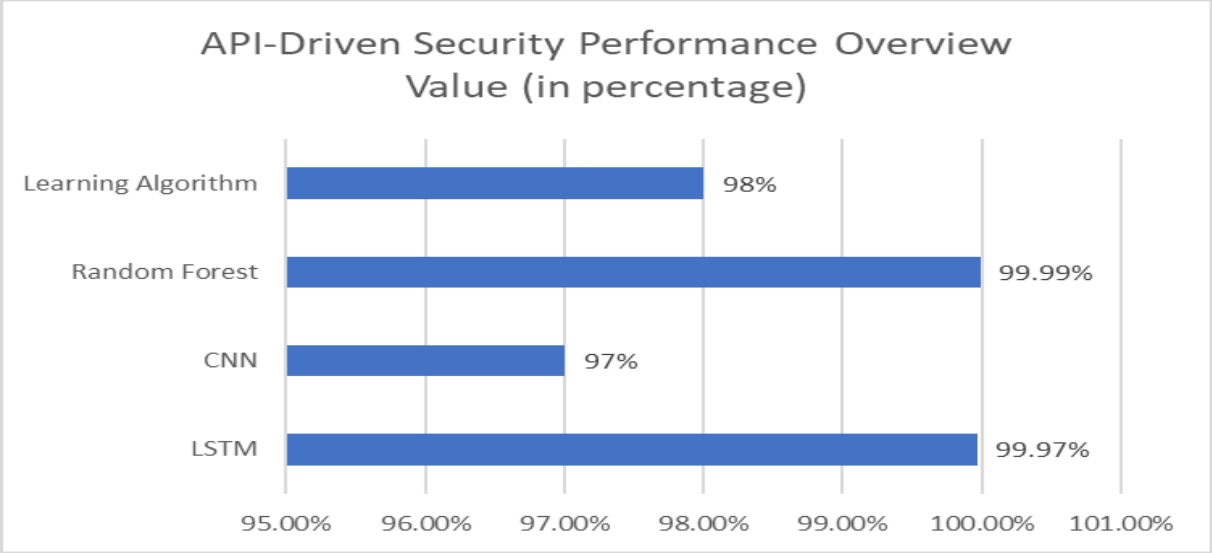
Network programmability via standard APIs means that security policies can be enforced by dynamically configuring the underlying infrastructure rather than relying on manual provisioning by an administrator. Contemporary SDN architectures expose control plane functions such as routing configuration, access control list management, network address translation rules, and traffic shaping policies via RESTful APIs that can be triggered programmatically by a security orchestration platform. When used in edge computing locations, these APIs permit sub-second response times for security enforcement actions and considerably shorter time frames from detection to automated remediation.

TM Forum Open API specifications provide standards for APIs and information models for telecommunications network management. The specs enable common business APIs and information models for identity management, device management, provisioning, QoS management, and billing integration. The TM Forum API ecosystem contains more than 50 families of APIs that meet industry standards for different aspects of telecom operations, such as Party Management APIs (identity and relationship data), Product Catalog and Ordering APIs (service provisioning), Service Quality Management APIs (performance management), and Resource Management APIs (network resource management) [7].

Linking machine intelligence threat detection systems with control APIs could provide a model where security responses could occur autonomously at machine speed. Research into distributed cloud architectures based on the blockchain indicates that the average time of response improves as the number of requests increases and that file operations incur less penalty to core infrastructure [7]. In resource allocation frameworks, learning-assisted algorithms restored about 98% of the slice performance from DoS attacks [7].

Security orchestration platforms use workflow engines to automate multi-stage incident response workflows across heterogeneous security and networking infrastructure. Deep learning frameworks help achieve an over 98% accuracy rate in the detection of volume-based flooding and spoofing attacks for network slicing security [8]. LSTM-based enhancements to these models can achieve up to 99.970% in detecting DDoS attacks in 5G scenarios [8]. Likewise, CNN-based prototypes can achieve approximately 97% accuracy [8]. Random Forest models can achieve 99.99% true positive rates in real-time detection systems [8].

Using infrastructure as code, security policies can be created, versioned in a source code version control system, and deployed using a continuous integration (CI) pipeline, just like any other code. Security as code artifacts can be subjected to software engineering processes such as peer review, automated testing, and audit trails for the changes made to policies.



Graph 2: Security Orchestration Performance Metrics [8]

5. Advanced Persistent Threats and Zero-Trust Security: A Game-Theoretic Approach

Much research on browser security indicators has gone into raising awareness of security and human factors. APTs are defined by their high-level sophistication and use of multiple phases and social and social+technical vectors to bypass organizational security. Out of 1329 respondents, the huge majority of technologically literate users understood that the HTTPS indicators, represented by green locks, meant that the page was secure. Most were unsure about HTTP indicators [9]. These findings also highlight the need for technical and human-centered defenses.

Game-Theoretic APT Defense

A game-theoretical model is a suitable tool to model attacker-defender competition. In many APT scenarios, topological properties of defender infrastructure alone give rise to a highly complex attack surface. For instance, a 3-machine network has 8 different attack vectors [10]. Vectors include, but are not limited to, buffer overflows, remote shell access, and FTP server vulnerabilities.

Furthermore, when experts' assessments of countermeasure effectiveness are available, opinions may also differ dramatically. For instance, when six domain experts assessed the effectiveness of countermeasures against APT scenarios, risk perceptions for equivalent defense-attack combinations ranged from "low" to "high" [10]. This disagreement is not a flaw but rather a reflection of real-world unpredictability.

APT Characteristic	Description	Defense Requirement
Multi-phase Attack	Sequential attack stages	Continuous monitoring
Social Engineering	Human manipulation tactics	User awareness training
Technical Exploitation	System vulnerability exploitation	Technical controls
High Sophistication	Advanced attack techniques	Advanced detection systems
Unpredictable Outcomes	Variable effectiveness of countermeasures	Adaptive defense strategies

Table 2: APT Characteristic Dimensions [9, 10]

6. Advanced Persistent Threats and Zero-Trust Security: A Game-Theoretic Approach

APTs are often a multi-stage attack that combines social engineering and technical exploits. The targeting and stealth of Stuxnet are unprecedented, using four zero-day exploits, Windows and PLC rootkits, and certificates stolen from reputable certificate authorities to update the firmware in programmable logic controllers in industrial control systems [11]. Results indicate that new forms of defense are needed that go beyond customary IT security.

Physical System Monitoring for Attack Detection

Process control systems and SCADA networks require a different approach to security than customary IT systems. A study on the Tennessee-Eastman chemical reactor process showed that attacks on the integrity of pressure sensors could cause the process to become unsafe in 20 hours and exceed the safety of 3000 kPa pressure in 28.6 hours [11]. Detection algorithms based on a model of the physical system were still able to detect the attacks with a false alarm rate of less than 0.004% of packets [11].

The nonparametric CUSUM (Cumulative Sum) detection method was found to be particularly effective, with experiments showing that the system could detect scaling attacks between 0.5 and 3 hours of their start time, depending on attack severity [11]. The use of automatic response mechanisms was found to be effective in keeping the pressure on the system below the safety threshold, even during attacks.

Modbus/TCP Intrusion Detection

Intrusion detection systems that monitor SCADA protocol traffic (e.g., Modbus/TCP) have excellent detection rates using DFA models. As shown in a case study of a production Modbus system that controls the power grid of Tel Aviv University, traffic can be very periodic, and similar messages are sent repeatedly [12]. The DFA-based approach that built accurate models from approximately 100 captured messages matched five of the seven PLCs tested over 111 hours of operation, without producing any false positives [12].

The system was able to correctly identify real anomalies, such as during technician testbed troubleshooting and misconfigured PLCs. The hierarchical two-level DFA (deterministic finite automaton) model also reduced the percentage of packets that were unknown from 0.4% to 0.0045% when processing over 40 million packets and only detecting 1982 anomalies for multi-period traffic patterns [12]. Such is the sensitivity of protocol-specific intrusion detection, with almost no false positives in critical infrastructure applications, that it can be quite effective.

Conclusion

Increased adoption of enterprise security architectures following a perimeter model has been the historical status quo of modern computing infrastructures. However, a shift to distributed zero-trust architectures will allow for native cyber threat detection and threat mitigation workflows based on artificial intelligence technologies. This will enable low-latency security controls across distributed network infrastructures. Deep learning models enable the detection of low and slow IOCs by learning temporal patterns. They outperform customary signature-based detection approaches. Policy enforcement close to the network edge, combined with API-enabled protection orchestration, allows sub-second automatic responses without degrading application performance. There is also work on game-theoretic budgeting of resources against Advanced Persistent Threats on critical infrastructure systems, monitoring physical systems against attacks on the industrial control process, and protocol-based intrusion detection systems with small false positive rates, which together provide a basis of a good security architecture for enterprises with multi-cloud, edge computing, and

mobile systems against the ultra-low latency environment of real-time applications and industrial control systems. As attacks continue to evolve, network defenses must innovate by building anomaly detection models through machine learning that accommodate diverse training data and leveraging explainable artificial intelligence techniques to provide security analysts with insights into the models' decision-making rationale and to generate leads for threat hunting.

References

[1] Jayavardhana Gubbi et al., "Internet of Things (IoT): A vision, architectural elements, and future directions," arxiv. Available: https://arxiv.org/pdf/1207.0203

[2] Arwa Alrawais et al., "Fog computing for the Internet of Things: Security and privacy issues," ResearchGate, 2017. Available: https://www.researchgate.net/profile/Chunqiang_Hu/publication/314162879

[3] Pacharee Phiayura and Songpon Teerakanok, "A comprehensive framework for migrating to zero trust architecture," IEEE Access, 2023. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10052642

[4] Tarik Taleb et al., "On Multi-Access Edge Computing: A Survey of the Emerging 5G Network Edge Cloud Architecture and Orchestration," acris, 2017. Available: https://acris.aalto.fi/ws/portalfiles/portal/27714159/ELEC_Taleb_et_al_On_multi_access_IEEE.pdf

[5] R. Vinayakumar et al., "Deep Learning Approach for Intelligent Intrusion Detection System," IEEEXplore, 2019. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8681044

[6] Mohamed Amine Ferrag, "Deep Learning for Cyber Security Intrusion Detection: Approaches, Datasets, and Comparative Study," 2019. https://eu-st01.ext.exlibrisgroup.com/44SUR_INST/upload/1771300813290/Deep_Learning_for_Cyber_Security_Intrusion_Detection__Approaches__Datasets__and_Comparative_Study.pdf?Expires=1771300933&Signature=tnC-Ca~fLIfW5~6xGBAvTwbsr-RyfGLCbqpbAVy1bTrTA2yliZdglm55-fLynVKmRQ5uO3hak0nWmzES6SPlN~~oK9vA6TDdpPfTQguLrMaDOmbDnnLXbfIp2H~ZHfqpVFhx38ijRkpxe~HskcjUexiuutXKvQHo5AtoHWxoKxbWNYMnuR3~XVbPV7mxQm944c1AkluOfv6geC54cZlW~yzc-yiXo7UJEOvCb3A-ryTDNS5WM~8VJLu~vi5POP4Zwt26kqlCQycfsADqJzztOCg8UtGJnub0LY17HufQdfvPkxnBfIsLnGG3dMuwoRKIYq16SLqgr~dLgMZm-qjTww__&Key-Pair-Id=APKAJ72OZCZ36VGVASIA

[7] Pradip Kumar Sharma et al., "A Software Defined Fog Node Based Distributed Blockchain Cloud Architecture for IoT," IEEE Access, 2018. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8053750

[8] José Cunha et al., "Enhancing Network Slicing Security: Machine Learning, Software-Defined Networking, and Network Functions Virtualization-Driven Strategies," MDPI, 2024. https://www.mdpi.com/1999-5903/16/7/226

[9] Adrienne Porter Felt, “Rethinking connection security indicators” In Twelfth Symposium on Usable Privacy and Security (SOUPS 2016), https://www.usenix.org/system/files/conference/soups2016/soups2016-paper-porter-felt.pdf

[10] Stefan Rass et al., “Defending Against Advanced Persistent Threats Using Game Theory," PLOS ONE, 2017. https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0168675&type=printable

[11] Álvaro A. Cárdenas et al., Attacks against process control systems: Risk assessment, detection, and response. IACM Digital Library, 2011. https://dl.acm.org/doi/pdf/10.1145/1966913.1966959

[12] Niv Goldenberg, Avishai Wool, Accurate modeling of Modbus/TCP for intrusion detection in SCADA systems. International Journal of Critical Infrastructure Protection, 2013. https://www.researchgate.net/profile/Avishai-Wool-2/publication/259166688