Roguescope: Automated Rogue Switch Detection Using Hybrid Neuro-Symbolic AI and Active Probing

Vijay Bhuse and Rajvardhan Patil
College of Computing, Grand Valley State University, Grand Rapids, Michigan, USA

bhusev@gvsu.edu
patilr@gvsu.edu

Abstract: The security architecture of modern enterprise networks heavily relies on perimeter defenses, leaving the physical layer of the wired Local Area Network (LAN) a persistent blind spot. This vulnerability is exceptionally severe in Small and Medium Enterprises (SMEs), critical infrastructure facilities, and academic environments where physical access ports are exposed in semi-public spaces like conference rooms, libraries, and classrooms. This paper introduces Roguescope, a fully automated detection framework designed to identify unauthorized Layer 2 forwarding devices—rogue switches—in resource-constrained environments. A rogue switch represents a critical infrastructure exploit where an unmanaged device splits a single authorized connection into multiple ports. This bypasses traditional network management systems because unmanaged switches operate transparently, generate no logs, and lack IP addresses. Moving beyond traditional signature matching and opaque "black-box" machine learning models that generate alert fatigue, Roguescope implements a unique Hybrid Symbolic-Statistical Architecture. It pairs deterministic active challenges—the "Forced Vacation" paradigm—with self-supervised temporal autoencoders to isolate and identify multi-device anomalies based on physical link physics and behavioral responses. The system forces micro-interruptions on edge ports, compelling hidden hardware to reveal its true topology through distinct recovery signatures, such as broadcast bursts or synchronized DHCP requests, rather than linear single-host patterns. Critically, Roguescope integrates a domain-constrained Retrieval-Augmented Detection (RAD) engine that interfaces with a local vector database of verified network incidents. This enables the framework to provide human-interpretable, natural-language explanations grounded in historical data for every security alert, supporting administrative trust and mitigating the risks of model hallucination. We present the system architecture, mathematical framing, threat parameters, and a methodology structured to evaluate scalability, adversarial red-team resilience, and operator resolution velocity. Ultimately, this framework provides a zero-cost open-source alternative to commercial network admission control systems, democratizing physical-layer enterprise defense for under-resourced organizations operating below the traditional security poverty line.

Keywords: Rogue switch detection, Network security, Explainable AI, Zero trust architecture, Retrieval-augmented generation

1. Introduction and Motivation

A House Intelligence Committee investigation concluded that certain network equipment manufacturing companies posed national security risks to the United States, warning that their hardware could facilitate espionage from foreign governments (House Intelligence Committee 2012). While organizations invest heavily in perimeter defenses—such as next-generation firewalls, Intrusion Detection Systems (IDS), and Zero Trust cloud architectures—the physical layer of the wired Local Area Network (LAN) remains a critical exposure point. This vulnerability is particularly acute in Small and Medium Enterprises (SMEs), critical infrastructure facilities, and academic institutions, where physical ports (wall jacks) are easily accessible in semi-public spaces like conference rooms, lobbies, and classrooms.

Because traditional security frameworks heavily prioritize logically inspecting traffic at boundary ingress points, physical access points on internal networks often lack granular, real-time access monitoring. An unauthorized connection established at an exposed physical port completely bypasses perimeter inspection mechanisms, providing an adversary with direct access to the internal network segment.

The Rogue Switch represents a critical, often overlooked vulnerability in this trust model. A rogue switch is an unauthorized Layer 2 forwarding device—typically an inexpensive, unmanaged "dumb" switch—connected to a legitimate network port by an attacker or malicious insider. Unlike a rogue endpoint, a rogue switch becomes a transparent part of the infrastructure itself. It splits a single authorized connection into multiple fan-out ports, allowing an adversary to attach unauthorized devices (such as packet sniffers or malware distribution units) while maintaining connectivity for the legitimate host.

By multiplying the physical access capacity of a single port without participating in control-plane discovery protocols such as Link Layer Discovery Protocol (LLDP) or Cisco Discovery Protocol (CDP), the rogue switch operates invisibly to standard Network Management Systems (NMS). It provides the adversary with an unmonitored Layer 2 bridge to perform eavesdropping, execute Man-in-the-Middle (MitM) attacks, or establish

persistent lateral access within the trusted network zone—all while preserving normal operations for the legitimate user to avoid detection.

1.1 The Transparency Problem

Because unmanaged switches operate transparently at Layer 2 and do not participate in standard management protocols like LLDP or CDP (Seifert and Edwards 2008), they remain invisible to standard Network Management Systems (NMS). They possess no IP address, generate no system logs, and respond to no standard network probes. Yet, their presence facilitates lateral movement, Man-in-the-Middle (MITM) attacks via ARP spoofing, and long-term persistence within the trusted zone. This aligns with findings in the Verizon Data Breach Investigations Report (DBIR), which highlights the prevalence of insider misuse and physical access as primary vectors for network compromise (Verizon 2024).

1.2 The Research gap

Current industry solutions are bifurcated and ill-suited for the SME context:

- Port security (MAC limiting): Highly brittle and easily bypassed via MAC cloning.
- Enterprise NAC (e.g., Cisco ISE): Cost-prohibitive for SMEs ($50k+ licensing) and requires complex 802.1X supplicant management (Cisco 2023).
- Black-box AI anomaly detection: Existing AI-driven security tools often output opaque "risk scores" without contextual verification, leading to high false-positive rates and severe alert fatigue for IT administrators.

1.3 Proposed Solution: Roguescope

To overcome these limitations, we propose Roguescope, a fully automated detection framework designed for resource-constrained environments. Roguescope shifts away from pure statistical anomaly detection by utilizing a Hybrid Symbolic–Statistical Architecture. It combines deterministic active challenges (the "Forced Vacation" protocol) with AI-driven behavioral inference. Critically, it incorporates a Retrieval-Augmented Detection (RAD) engine to provide human-interpretable explanations for every alert, ensuring that the system is not just accurate, but auditable and trustworthy.

2. Threat Model and Notion of Trust

2.1 Notion of Trust

In this framework, trust is defined as the verifiable integrity of the physical network topology and the strict cryptographic or behavioral binding between a physical port and an authorized endpoint.

- Trusted core: We assume that core switching infrastructure (managed backbone switches) and the Software Defined Network (SDN) controller/ Network Management System (NMS) are secure, physically protected, and under administrative control. Management keys (SSH/SNMPv3) are assumed secure.
- Untrusted edge: We assume that all edge ports (wall jacks) in user-accessible areas are untrusted interfaces where the topology may degrade due to physical tampering. This approach aligns with Zero Trust Architecture principles, which dictate that trust must be continuously verified rather than statically assigned (NIST 2020).
- Device trust: Trust in a connected device is established not merely by its credentials (which can be stolen) or its MAC address (which can be spoofed), but by its behavioral continuity—specifically, its traffic patterns relative to the physical port it occupies.

2.2 Threat Model

We posit an active adversary with temporary physical access to the facility (e.g., a malicious insider, service staff, or a visitor).

- Goal: Introduce an unauthorized device onto the network to sniff traffic or exfiltrate data.
- Method: The adversary inserts a small, unmanaged switch between the wall jack and the legitimate device. This keeps the legitimate device online, avoiding an administrative Link Down alert.
- Capabilities: The adversary is capable of MAC Cloning (spoofing the legitimate host's MAC address) and passive listening. The adversary cannot compromise the cryptographic keys of the core network.

This scenario maps directly to MITRE ATT&CK techniques T1200 (Hardware Additions) and T1046 (Network Service Scanning) (MITRE 2024).

3. Related work and Foundational Context

3.1 Wireless vs. Wired Detection

Extensive research exists on detecting Rogue Access Points (RAPs) using Radio Frequency (RF) signal analysis and timing dynamics (Beyah et al. 2004; Mano et al. 2006). However, wired unmanaged switches introduce negligible latency (few microseconds) and emit no RF signals, rendering wireless monitoring techniques inapplicable.

3.2 Passive Anomaly Detection

Meidan et al. (2018) utilized supervised machine learning (Random Forest) to whitelist Internet of Things (IoT) devices. While effective for endpoint profiling, it does not detect the underlying infrastructure enabling those endpoints. Prior work explored management plane telemetry (Bhuse et al. 2019), while frameworks like AuDI (Marchal et al. 2019) autonomously classify IoT device types and detect anomalies without deep packet inspection. However, purely passive Machine Learning (ML) models suffer from high False Positive Rates (FPR) in dynamic enterprise environments. Roguescope advances the state of the art by integrating active, deterministic interrogation.

3.3 Trustworthy and Explainable AI (XAI)

Recent research highlights the "black box" problem in AI-driven cybersecurity (Gunning 2019; Arp et al. 2022), where administrators hesitate to automate network responses based on opaque neural network decisions (Ribeiro et al. 2016). This project contributes to the field of Trustworthy AI by implementing a Retrieval-Augmented Generation (RAG) architecture (Lewis et al. 2020), ensuring that AI decisions are grounded in verifiable, historical network data rather than hallucinated patterns.

3.4 Evolution of Active Interrogation Strategies

Unmanaged switches ignore voluntary protocols like 802.1X and Link Layer Discovery Protocol (LLDP), creating a significant coverage gap for traditional Network Admission Control (NAC) systems. To address this visibility gap, the lineage of active interrogation techniques establishes the core mechanics of Roguescope:

- Foundational active probing: The concept of active port isolation was formalized by Prins and Bhuse (2018), who introduced the "Forced Vacation" technique. This method posits that a rogue switch hosting multiple downstream devices exhibits a "thundering herd" signature upon reconnection, contrasting sharply with the behavior of a single legitimate host.
- Physical layer triggers: Quitiquit and Bhuse (2022) shifted the detection trigger to the physical layer (Layer 1). Their work analyzed kernel-level link-state transitions, establishing that the physical insertion of a rogue switch necessitates a "cable swap," creating a mandatory LINK_DOWN to LINK_UP transition sequence.
- Hybrid and ML-enhanced detection: Bhuse et al. (2025) proposed integrating deterministic rules with supervised machine learning. Their framework utilizes Logistic Regression to analyze network anomalies—such as Address Resolution Protocol (ARP) spoofing, MAC flooding, and Spanning Tree Protocol (STP) root bridge manipulation—establishing that combining physical-layer triggers with probabilistic behavioral modeling provides a robust defense against active evasion techniques.

4. Technical Approach and Architecture

Roguescope employs a Multi-Layered Architecture that progresses from deterministic signal processing to probabilistic reasoning and generative explanation.

4.1 Layer 1: Real-time Topology Monitoring (The Trigger)

The first layer establishes physical visibility, designed to overcome the latency limitations of standard polling. During a trusted initialization phase, the system crawls the Open vSwitch (OVS) database (ovs-vsctl) and SNMP Management Information Bases (MIBs) to map every active port and link state, generating a baseline_topology.json.

We utilize kernel-level asynchronous traps. A LINK_UP event on a port not scheduled for use, or a LINK_FLAP (rapid Down/Up sequence) on an active port, instantly triggers the evaluation pipeline. This trigger mechanism is derived from Quitiquit and Bhuse (2022), establishing that an attacker cannot bypass the physics of a

connection: the legitimate device must be unplugged to insert the switch, generating a specific link-state transition within a short window.

4.2 Layer 2: Device and Protocol Verification (The Filter)

Once triggered, the system actively interrogates the connected segment to filter out benign connection events (e.g., a standard workstation reboot):

- Active scanning: The controller triggers a directed fping sweep and arp-scan on the Virtual Local Area Network (VLAN) associated with the flagged port.
- MAC whitelisting: Discovered MAC addresses are compared against a baseline_mac_ip whitelist.
- Anomaly classification: Building on Bhuse et al. (2025), network anomalies are categorized via supervised learning indicators:
- *Anomaly Type A (New Device):* Unknown OUI (Organizationally Unique Identifier).
- *Anomaly Type B (Many-to-One):* Single MAC claiming multiple IP addresses (indicative of Network Address Translation (NAT)/Masquerading).
- *Anomaly Type C (One-to-Many):* Single IP claimed by conflicting MACs (Active MITM).

4.3 Layer 3: The "Forced Vacation" Algorithm (Active Validation)

For medium-risk ports, the system initiates a Forced Vacation (Prins and Bhuse 2018). This active defense mechanism relies on the mechanics of standard protocols like DHCP and ARP:

- Isolation: The controller installs a high-priority DROP flow via OpenFlow, blocking user traffic while permitting DHCP/ARP.
- Challenge: The controller sends unicast ARP probes to the suspected IP, "vacating" the port logic.
- Observation: The controller monitors the port for 60 seconds.
- Rogue signature (The thundering herd): Unmanaged switches exhibit distinct recovery behaviors compared to single endpoints (Bhuse, Kalafut and Dohn 2019). If multiple devices are hidden downstream, the restoration of the link triggers a race condition where all downstream devices simultaneously attempt to re-verify connectivity. This results in a "Thundering Herd" of synchronized DHCP requests—a high-entropy burst signature that is mathematically distinct from a single host.

4.4 Layer 4: The Correlation Engine (Probabilistic Reasoning)

An isolated anomaly is often noise, but when multiple warning signs happen at the same time, it points to a serious security threat. To sort through these mixed signals, we use a probability model called a Bayesian Belief Network (BBN). The Roguescope framework evaluates the conditional probability of a rogue switch presence $\mathbb{P}(\text{Rogue} \mid \mathbf{E})$ given an observed evidence vector $\mathbf{E} = \{E_{\text{flap}}, E_{\text{mac}}, E_{\text{dhcp}}\}$:

$$\mathbb{P}\left(\text{Rogue} \mid E_{\text{flap}}, E_{\text{mac}}, E_{\text{dhcp}}\right) = \frac{\mathbb{P}\left(E_{\text{flap}}, E_{\text{mac}}, E_{\text{dhcp}} \mid \text{Rogue}\right) \cdot \mathbb{P}(\text{Rogue})}{\mathbb{P}\left(E_{\text{flap}}, E_{\text{mac}}, E_{\text{dhcp}}\right)}$$

Where the observed red flags are defined as:

- **Topology Flaps (E_{flap}):** Sudden connection drops or re-arrangements on the network link.
- **New MACs (E_{mac}):** Multiple unrecognized device ID numbers appearing on a single port.
- **DHCP Bursts (E_{dhcp}):** A sudden flood of IP address requests.

This probabilistic approach addresses the limitations identified in traditional logistic regression classifiers (Bhuse et al. 2025), which struggle to distinguish complex, multi-variable anomalies like STP root bridge manipulation without probabilistic context.

4.4.1 This works better due to the following reasons.

- Older methods (Basic rules or Standard AI) get tricked easily. For example, if an attacker copies a legitimate device's ID number (MAC cloning), standard tools miss it entirely or flood administrators with false alarms.
- Roguescope's model weighs all three red flags together. Even if an attacker hides one clue, the combination of the other signals will still trigger an alert, giving security teams clear, step-by-step reasons for the warning.

4.5 Layer 5: AI-driven Behavioral Modeling and Trust-aware Inference

4.5.1 Multi-modal behavioral embeddings

For each edge port, Roguescope constructs a temporal behavioral embedding vector capturing how the connected entity responds to topology changes and Forced Vacation interruptions.

- Inputs: Link-state transition timing, ARP/DHCP recovery dynamics, MAC churn, and broadcast amplification factors. We incorporate physical-layer fingerprinting features such as Clock Skew, building on the foundational work of Kohno, Broido and Claffy (2005).
- Learning model: We employ Self-Supervised Sequence Models (Temporal Autoencoders) to learn the normal recovery pattern of a direct cable connection. Deviations from this baseline indicate structurally anomalous behavior—such as the buffering latency introduced by a rogue switch (Mirsky et al. 2018). We also explore the use of Generative Adversarial Networks (GANs) (Goodfellow et al. 2014) to synthesize adversarial examples for robust training.

4.5.2 Retrieval-augmented detection using RAG

To ensure correctness and eliminate generative hallucinations, the system employs a Retrieval-Augmented Generation (RAG) architecture (Lewis et al. 2020), mitigating the closed world limitations of traditional ML-based intrusion detection systems (Sommer and Paxson 2010).

Before generating an alert, the system queries a localized vector database (e.g., Milvus) housing historical port embeddings, signatures of known rogue hardware variants derived from lab testing, and administrator-verified past incidents. It retrieves the Top-*k* most similar historical cases (Borgeaud et al. 2022), anchoring the AI's reasoning in empirical network behavior rather than abstract probabilistic spaces.

4.5.3 Generative artificial intelligence (GenAI) assisted anomaly interpretation (XAI)

Roguescope incorporates a domain-constrained Large Language Model (LLM) to assist human operators (Radford et al. 2019). The LLM does not ingest raw packets; instead, it receives structured summaries from the prior layers. As detailed by Gunning (2019), this output design maps to explainable artificial intelligence metrics. To ensure functional network governance, the LLM operates as an interpretive layer and cannot autonomously override the probabilistic thresholds of Layer 4 or execute automated port shutdowns, ensuring Human-in-the-Loop safety (Amodei et al. 2016).

5. Research Plan and Methodology

The project structure scales verification across three phases:

5.1 Phase 1: Testbed Construction and High-fidelity Data Collection

The primary objective of Year 1 is to generate a comprehensive, labeled dataset of Layer 2 anomalies.

- Task 1.1: Construction of the Untrusted Edge Testbed: We will construct a physical lab environment designed to mimic the edge of an SME network, consisting of 8 enterprise-grade managed switches connected to an SDN controller (Ryu/Faucet). We will acquire a diverse library of unmanaged switches, Small Office / Home Office (SOHO) routers, and single-board computers configured as bridging attack tools.
- Task 1.2: Signature Generation (The Ground Truth Dataset): Scripts will systematically insert rogue devices into the testbed and trigger the Forced Vacation protocol. We will capture high-resolution Packet Capture (PCAP) files and SNMP logs focusing on boot-up signatures, recovery patterns, and link-state transitions.
- Task 1.3: Large-scale Simulation: To verify the scalability of the OVS controller logic, we will develop Mininet scripts to emulate topologies exceeding 100 nodes, stress-testing the framework to ensure active probing does not induce broadcast storms.

5.2 Phase 2: Algorithm Design and Neuro-symbolic AI Training

Year 2 focuses on software implementation and model optimization.

- Task 2.1: Development of the SDN Software Agent: We will develop the control logic in Python using the Ryu SDN framework, integrating the active scanning and OpenFlow rule execution loops.

- Task 2.2: Training Temporal Autoencoders: Using the dataset from Year 1, we will train Temporal Autoencoders to learn the normal behavioral profiles of legitimate endpoints using traffic entropy and inter-arrival times.
- Task 2.3: Construction of the RAG Engine: We will establish the Retrieval-Augmented Detection system, index the signatures in Milvus, and configure a local domain-constrained model (e.g., Llama-3-8B) to synthesize natural-language logs.

5.3 Phase 3: Evaluation, XAI, and Field Testing

The final year focuses on validation.

- Task 3.1: Red Teaming and Adversarial Evaluation: We will conduct internal Capture the Flag (CTF) exercises applying evasion techniques such as MAC cloning and traffic shaping to assess resilience.
- Task 3.2: XAI User Study: We will evaluate the utility of the LLM-generated logs versus standard Security Information and Event Management (SIEM) outputs, measuring Time to Resolution and operator confidence.
- Task 3.3: Deployment in Air-Gapped Lab: We will deploy the unified Roguescope system within a segmented, air-gapped Cyberrange to evaluate long-term stability and minimize operational false positives under continuous monitoring.

6. Wider Contributions

6.1 Democratizing Enterprise Security for the Cyber-have-nots

The digital divide in cybersecurity has created a Security Poverty Line, where effective physical-layer defense is often accessible only to well-resourced enterprises. Small businesses, rural healthcare facilities, and local municipal governments constitute the soft underbelly of national cyber defense. These organizations frequently operate critical infrastructure yet lack the budget for commercial Network Access Control (NAC) solutions, which often require expensive proprietary hardware and licensing fees exceeding $50,000. Roguescope directly addresses this disparity by delivering an enterprise-grade physical defense framework as Free, Open-Source Software (OSS) optimized to run on commodity hardware, ensuring under-resourced sectors can implement physical Zero Trust security.

6.2 Workforce Development and Curricular Innovation

Conducted at a Primarily Undergraduate Institution (PUI), this project relies on a student-centered research model where undergraduate and graduate students are active participants. The research findings will be translated into classroom learning by permanently integrating a hands-on Rogue Switch Hunt CTF challenge into foundational systems security courses, training future cybersecurity practitioners in physical-layer network defense.

7. In Conclusion

Wired physical ports in semi-public enterprise spaces remain a high-risk blind spot where transparent, unmanaged rogue switches bypass conventional perimeter controls and cost-prohibitive NAC architectures. This paper presented Roguescope, a fully automated, open-source framework designed to deliver physical Zero Trust security for resource-constrained environments. Roguescope bridges the gap between physical-layer security and interpretable AI through three primary mechanisms:

- **Deterministic Active Interrogation:** By combining kernel-level link-state transition traps (LINK_DOWN / LINK_UP) with the active "Forced Vacation" protocol, Roguescope compels hidden Layer 2 forwarding hardware to reveal itself via high-entropy "Thundering Herd" recovery bursts (synchronized DHCP requests).
- **Multi-Vector Probabilistic Reasoning:** A Bayesian Belief Network (BBN) evaluates concurrent signals—including link flaps (E_{flap}), uncached MAC addresses (E_{mac}), DHCP bursts ($E_{\text{dhcp_burst}}$), and DHCP state flaps ($E_{\text{dhcp_flap}}$)—eliminating false positives and overcoming the evasion limitations of static rules and traditional machine learning classifiers.
- **Audit-Ready Explainable AI (XAI):** Leveraging self-supervised temporal autoencoders alongside a vector-grounded Retrieval-Augmented Detection (RAD) engine, Roguescope pairs deep behavioral fingerprinting with domain-constrained language models to deliver transparent, hallucination-free alert explanations for network operators.

By offering a zero-cost, software-defined alternative to commercial admission control platforms, Roguescope democratizes physical-layer network defense—ensuring that small businesses, academic institutions, and critical infrastructure facilities operating below the traditional security poverty line can effectively detect and mitigate physical infrastructure exploits.

Ethics Declaration: We do not need ethical clearance for the research referred to in this paper.

AI Declaration: During the preparation of this work, the author(s) used Google Gemini to improve readability and grammar. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

References

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J. and Mané, D., 2016. Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565*.

Arp, D., Quiring, E., Pendlebury, F., Warnecke, A., Pierazzi, F., Sterling, C., Wressnegger, C. and Rieck, K., 2022. Dos and don'ts of machine learning in computer security. *Proceedings of the 31st USENIX Security Symposium*, pp. 3959-3976.

Beyah, K., Copeland, J. and McGill, A., 2004. Rogue access point detection using temporal traffic characteristics. *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 2251-2255.

Bhuse, V., Kalafut, J. and Dohn, M., 2019. Active validation of Layer 2 network topologies using targeted active probing. *International Journal of Network Security*, 21(3), pp. 411-420.

Bhuse, V., Kanipakam, S., Patil, R. and Wang, X., 2025. Hybrid machine learning frameworks for advanced Layer 2 network anomaly classification. *Journal of Cyber Security Technology*, 9(1), pp. 45-62.

Borgeaud, S., Mensch, A., Hoffmann, J., Cai, T., Rutherford, E., Casas, J. and Sifre, L., 2022. Improving language models by retrieving from trillions of tokens. *International Conference on Machine Learning (ICML)*, pp. 2206-2240.

Cisco, 2023. *Cisco Identity Services Engine (ISE) Licensing Guide, Release 3.x*. San Jose: Cisco Systems.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., 2014. Generative adversarial nets. *Advances in Neural Information Processing Systems (NeurIPS)*, 27, pp. 2672-2680.

Gunning, D., 2019. *Explainable Artificial Intelligence (XAI)*. Defense Advanced Research Projects Agency (DARPA) Technical Report.

House Intelligence Committee, 2012. *Investigative Report on the U.S. National Security Issues Posed by Chinese Telecommunications Companies Huawei and ZTE*. Washington D.C.: U.S. House of Representatives.

Kohno, T., Broido, A. and Claffy, K.C., 2005. Remote physical device fingerprinting. *IEEE Transactions on Dependency and Secure Computing*, 2(2), pp. 93-108.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Lewis, M., Riedel, S. and Kiela, D., 2020. Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, pp. 9459-9474.

Mano, M., Ramachandran, L. and Shin, K., 2006. Wireless rogue access point detection using physical layer traits. *IEEE International Conference on Communications (ICC)*, pp. 1102-1107.

Marchal, S., Miettinen, M., Asokan, N. and Sadeghi, A.R., 2019. AuDI: Autonomous IoT device-type identification using passive manufacturer-specific network traffic monitoring. *IEEE Transactions on Dependable and Secure Computing*, 18(6), pp. 2794-2808.

Meidan, Y., Bohadana, M., Mathov, Y., Mirsky, Y., Shabtai, A., Breitenbacher, D. and Elovici, Y., 2018. N-BaIoT—Network-based detection of IoT botnet attacks using deep autoencoders. *IEEE Pervasive Computing*, 17(3), pp. 12-22.

Mirsky, Y., Doitshman, T., Elovici, Y. and Shabtai, A., 2018. Kitsune: an ensemble of autoencoders for online network intrusion detection. *Network and Distributed System Security (NDSS) Symposium*.

MITRE, 2024. *MITRE ATT&CK Framework: Hardware Additions (T1200) and Network Service Scanning (T1046)*. Available at: https://attack.mitre.org/ [Accessed May 2026].

NIST, 2020. *Zero Trust Architecture*. NIST Special Publication 800-207. National Institute of Standards and Technology.

Prins, J. and Bhuse, V., 2018. The "Forced Vacation" paradigm: Translating financial fraud detection controls to physical layer network security. *Proceedings of the Central Science Cybersecurity Conference*, pp. 112-119.

Quitiquit, M. and Bhuse, V., 2022. Event-driven network security: Microsecond link-state transition tracking at the kernel level. *IEEE Letters of the Computer Society*, 5(2), pp. 22-25.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D. and Sutskever, I., 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8), p. 9.

Ribeiro, M.T., Singh, S. and Guestrin, C., 2016. "Why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1135-1144.

Seifert, R. and Edwards, J., 2008. *The All-New Switch Book: The Complete Guide to LAN Switching Technology*. 2nd ed. Indianapolis: Wiley.

Sommer, R. and Paxson, V., 2010. Outside the closed world: On using machine learning for network intrusion detection. *IEEE Symposium on Security and Privacy*, pp. 305-316.

Verizon, 2024. *2024 Data Breach Investigations Report (DBIR)*. Verizon Business Security Solutions.