

Physical Tells Digital Threats: Supervised ML-IDS for Multimodal IoT Telemetry in Smart Buildings

Obrina Briliyant*, Amir Javed*, Yulia Cherdantseva*, Septia Ulfa Sunaringtyas**

*School of Computer Science and Informatics, Cardiff University, Cardiff, United Kingdom

**Department of Cyber Security, Politeknik Siber dan Sandi Negara, Bogor, Indonesia

0000-0002-1054-8112, 0000-0001-9761-0945, 0000-0002-3527-1121, septia.ulfa@poltekssn.ac.id

Abstract—Traditional machine learning-based intrusion detection systems (ML-IDS) in smart building environments face critical limitations, including heavy reliance on network traffic analysis, high computational overhead, and inability to detect insider threats. The solution lies in recognizing that cyber attacks in smart buildings inevitably manifest as anomalies in physical device behaviors, such as temperature fluctuations, unexpected door activations, and abnormal HVAC operations, which traditional network-based IDS systems completely overlook. This paper presents a novel supervised ML-IDS that leverages multimodal IoT telemetry data, combining physical sensor readings with device operational states to detect cyber attacks. Using a dataset with 221,859 telemetry records from smart building infrastructure, we demonstrate that physical sensor data (temperature, motion, door states) combined with Modbus protocol communications provide superior attack detection capabilities. Our multimodal telemetry-based ML-IDS achieves 84.47% accuracy and 90.76% AUC for binary attack detection, significantly outperforming conventional IoT security approaches while operating with minimal computational requirements suitable for edge deployment. The system successfully detects seven distinct types of attack: backdoor, DDoS, injection, password, ransomware, scanning, and XSS attacks. Selective classification detectors demonstrate exceptional performance for specific attacks, such as scanning (85.66% AUC) and DDoS (84.01% AUC). Our findings suggest that multimodal IoT telemetry data, particularly combined physical readings and device status indicators, provide sufficient discriminative features for effective cyber attack detection, including zero-day exploits and insider threats that manifest as anomalies in physical device behaviors.

Index Terms—physical telemetry, IoT, ML, IDS, smart building, AUC.

I. INTRODUCTION

The convergence of Information Technology (IT) and Operational Technology (OT) in smart building environments has created unprecedented cybersecurity challenges. While conventional ML-based IDS approaches focus primarily on network traffic analysis, smart buildings present a different threat landscape where cyber attacks directly impact physical infrastructure through Building Automation Systems (BAS), Heating Ventilation and Air Conditioning (HVAC) controls, lighting systems, and security sensor devices (fire suppression, camera, etc.). This physical-cyber interdependency renders

This research was made possible by the support of the Indonesia Endowment Fund for Education Agency (LPDP) and Politeknik Siber dan Sandi Negara (Poltek SSN). Their investment in our work has significantly contributed to the quality, impact and publication of our research findings.

traditional network-centric security approaches insufficient for comprehensive threat detection [1].

Existing ML-based IDS solutions face three fundamental challenges in smart building environments: First, smart buildings integrate diverse communication protocols, including Modbus, BACnet, KNX, and other proprietary IoT protocols, each with distinct traffic patterns and features that confound traditional signature-based and anomaly detection approaches [1]. Current ML-IDSs trained on conventional network traffic shows poor performance when applied to heterogeneous protocols prevalent in real-world BAS [2]. Second, as IoT implementations increasingly implement end-to-end encryption and complex edge computing services, the rich network feature traditionally used by ML-based IDSs become too sensitive, producing more false positives [3]. Third, the sophistication of attacks extends beyond traditional network-based threats to include physical-layer attacks that manipulate sensor readings, device firmware, and insider threats that exploit legitimate interfaces. These attack vectors often bypass network monitoring and remain undetected while causing substantial physical and economic damage [4].

Unlike network packets that can be encrypted, spoofed, or tunneled through covert channels, physical sensor readings reflect the actual state of building infrastructure and are inherently difficult to manipulate without detection. The fundamental insight driving this approach is that attackers cannot manipulate building's systems without leaving traces in physical telemetry data [3]. This study makes three primary contributions to the cybersecurity and IoT security literature:

- Methodologically, we demonstrate that multimodal telemetry analysis provides more comprehensive attack detection performance compared to network-only approaches.
- Empirically, we provide comprehensive performance evaluation across multiple attack types using realistic smart building data, revealing specific telemetry features most indicative of different threat categories.
- Practically, we deliver a complete deployment framework including data pre-processing, addressing class imbalance, and customized multiclass classification technique designed specifically for physical telemetry limitations.

The paper is structured as follows: Section 2 reviews related work, Section 3 details the methodology, Section 4 presents

the experimental setup and results, Section 5 discusses implications, conclusion, and future directions. Our experiment shows that physical telemetry provides sufficient discriminative power for cyber threat detection.

II. RELATED WORKS

Combined physical sensor readings and device status indicators have been shown to yield result for ML-based intrusion detection in BAS [5]. Several studies report that unsupervised techniques—such as ensembles of autoencoders and one-class SVMs—that integrate network flows with physical telemetry achieve high detection accuracy and low false positive rates when faced with simulated zero-day exploits [6]. In these studies, statistical summaries, time-based features, and protocol-specific attributes from both physical and cyber data enhance event verification and context awareness [7]. By contrast, only a few studies address insider threats, typically through behavioral profiling of physical access or movement logs [8]. A summary of characteristics from the previous studies is depicted in Table I.

The reviewed studies demonstrate state-of-the-art approaches to the engineering of telemetry data characteristics, with studies extracting statistical summaries of network flows [7], [9], time-based features [21], and protocol-specific attributes [20]. Advanced sensor fusion techniques have emerged as a key trend, with Birnbach et al. [5] integrating data from 48 physical sensors and Wang et al. [22] and Yasaei et al. [23] combining network traffic with device telemetry. This multi-modal approach consistently outperforms single-modality methods, while physical-cyber correlation models further enhance discriminative power through context-aware frameworks [10] and building ontologies [7].

Detection capabilities in the literature reveal a strong emphasis on unsupervised anomaly detection for zero-day exploit identification, with autoencoders, one-class SVMs, and clustering methods [6], [10], [24] achieving high F1-scores and low false positive rates. However, explicit evaluation against real-world zero-day exploits remains limited, with most evidence based on simulated scenarios. Insider threat detection capabilities are notably underdeveloped, with only a few studies [8], [3] addressing behavioral profiling through physical access logs.

Building automation systems (BAS) presents unique challenges that significantly impact intrusion detection system design and deployment. Device heterogeneity, protocol diversity spanning BACnet, Zigbee, and MQTT [19], [25], and the need for real-time detection under resource constraints create complex operational environments. The prevalence of proprietary and legacy systems complicates both development and evaluation processes. Resource limitations on IoT devices have driven researchers toward lightweight model architectures and distributed detection frameworks [9], [16].

Implementation considerations emphasize the critical importance of scalability, adaptability to new devices, and seamless integration with existing building management systems. To address these challenges, many studies have proposed Software-

Defined Networking (SDN) and Virtualization-based architectures [11]. Common mitigation strategies include strategic feature selection [4], [25] and hierarchical detection architectures that balance performance with resource efficiency, enabling practical deployment in resource-constrained building automation environments while maintaining robust security capabilities [26], [5].

To the best of our knowledge, while the ToN-IoT dataset [27] has been extensively utilized across various cybersecurity research domains—including network intrusion detection [26], federated learning approaches [28], and IoT botnet identification [13]—no prior study has systematically investigated the intrinsic characteristics and discriminative power of its telemetry features for BAS security. Existing research has predominantly focused on leveraging the network features of the dataset, and its individual contributions to attack detection. This oversight represents a significant research gap, as detecting anomalies in the behavioral patterns encoded within environmental sensors, device status indicators, industrial communication protocols, and spatial features is a valid avenue for IDS.

III. METHODOLOGY

Methodology consists of dataset preparation, feature analysis and selection, addressing class imbalance, choosing ML algorithms, and analyzing evaluation and validation metrics. Overall methodology is depicted in Figure 1.

A. Data Preprocessing & Feature Selection

Our experimental evaluation utilized the ToN-IoT dataset [27], which contains 221,859 telemetry records from a realistic smart building testbed. The dataset encompasses IoT device telemetry data representing physical device states and operational parameters across multiple smart building components. Each record includes originally 14 features spanning environmental sensors, device status indicators, industrial communication protocols, and geolocation data. The telemetry features were categorized into four primary groups:

- Environmental sensors including `current_temperature`, `fridge_temperature`, and `temp_condition` reflecting thermal management systems;
- Device status indicators such as `motion_status`, `door_state`, `light_status`, and `sphone_signal` representing physical device states;
- Industrial communication features including `FC1_Read_Input_Register`, `FC2_Read_Discrete_Value`, `FC3_Read_Holding_Register`, and `FC4_Read_Coil` representing Modbus protocol interactions; and
- Spatial features comprising `latitude` and `longitude` coordinates.

Categorical variable encoding presented unique challenges due to inconsistent data formatting. The `temp_condition` variable exhibited data collection inconsistencies, requiring

TABLE I: Characteristics of Previous Studies

Study Focus	References	Detection Approach	Telemetry Data	Environment
Sensor Behavioral Profiling	[7], [6], [9], [5]	Unsupervised (AE, OCSVM, PCA), Supervised (ANN, SVM, RF)	Physical readings (temp., humidity), Sensor data, Network flows, mobility indicators	Building automation testbed, campus
Attack Detection based on Multimodal Data	[10], [11], [12], [13]	Context-based (AE-GRU), Ensemble Hybrid approaches, Multi-modal fused anomaly with signal & device profiles	Device states, ambient adaptive smart building & building telemetric trends, multiple modal integration	Real-world smart buildings, IoT deployments
Lightweight Embedded Systems	[14], [15], [16], [17]	Supervised models (RF, tree-based, ML Classification), XGBoost, DL anomaly detection	Limited feature sets, optimized on-demand, embedded telemetry	Smart thermostat environments, IoT testbeds
Time-Based Detector & BACnet Network	[18], [19], [20], [10]	Role-based classification, Hybrid Knowledge and Anomaly Ensemble method, Timing Attack Supervised learning	BACnet-specific data, network flows, fully-featured BACnet real-world evaluated interoperability, and security logging	Simulated BACnet, campus IoT networks

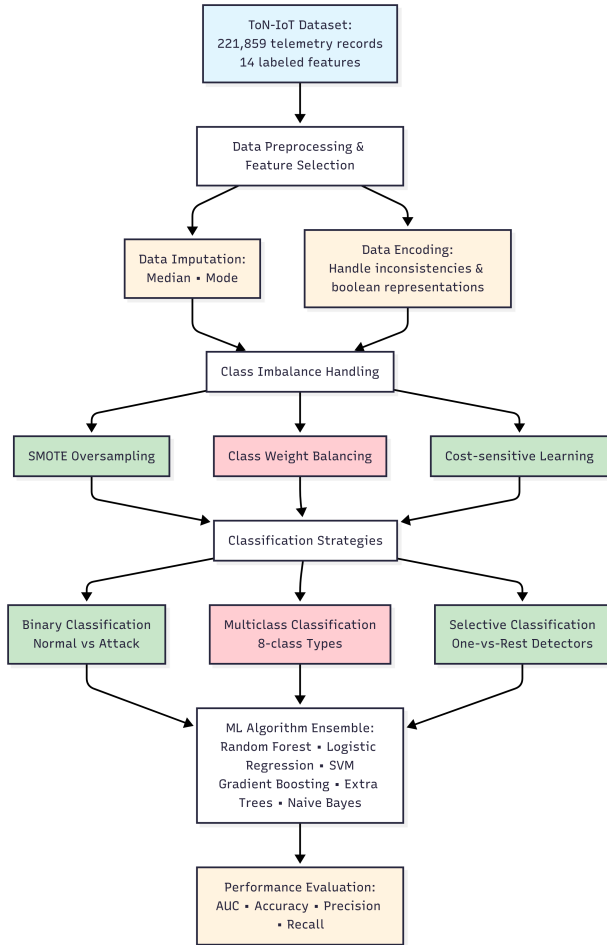


Fig. 1: Feature Selection & Classification Methodology.

robust preprocessing to handle whitespace variations. On the other hand, `sphone_signal` contained both numeric ('0', '1') and boolean string representations ('false', 'true'), necessitating careful encoding to preserve semantic meaning. The dataset also includes attack classification labels (`label`) and attack type identifiers (`type`) for supervised learning tasks.

Tables II and III show the dataframe after data imputation and data encoding, with the features split across two tables for clarity due to the large number of variables.

TABLE II: Dataframe After Preprocessing - Part 1

motion_status	light_status	label	type	door_state	sphone_signal	current_temperature	thermostat_status
0.0	0	1	ddos	0	0	28.59148844	1.0
1.0	1	1	ddos	0	0	28.59148844	1.0
1.0	1	1	backdoor	1	0	28.59148844	1.0
0.0	0	1	backdoor	0	1	28.59148844	1.0
1.0	1	1	injection	0	1	28.59148844	1.0

TABLE III: Dataframe After Preprocessing - Part 2

FC1_Read_Input_Register	FC2_Read_Discrete_Value	FC3_Read_Holding_Register	FC4_Read_Coil	Intitude	Intigrate	Bridge_temperature	temp_condition
32228.0	32756.0	32147.5	32845.0	37.47991687	47.97758566	6.85	0
32228.0	32756.0	32147.5	32845.0	37.47991687	47.97758566	6.85	0
32228.0	32756.0	32147.5	32845.0	37.47991687	47.97758566	6.85	0
32228.0	32756.0	32147.5	32845.0	37.47991687	47.97758566	6.85	0

B. Class Imbalance Mitigation

The dataset contained seven distinct attack types plus normal traffic, creating an eight-class multiclass classification problem. The attack types included backdoor attacks (13.5% of training data), DDoS attacks (9.0%), injection attacks (13.5%), password attacks (13.5%), ransomware (5.9%), scanning attacks (1.6%) and XSS attacks (2.4%), with normal traffic comprising 40.6% of the dataset. This distribution revealed severe class imbalance with a 25:1 ratio between the most frequent (normal) and least frequent (scanning) classes. The dataset attack profile is visualized in Figure 2. The severe class imbalance problem required multiple mitigation approaches. We implemented three primary strategies: (a) Algorithm-level balancing using `class_weight='balanced'` parameters to automatically adjust model sensitivity to minority classes; (b) Data-level balancing using Synthetic Minority Oversampling Technique (SMOTE) to generate synthetic minority class samples; and (c) Cost-sensitive learning with custom class weights inversely proportional to class frequencies. For multiclass scenarios, we additionally implemented a One-vs-Rest approach, training individual binary classifiers for each attack type versus all others. This strategy addressed the fundamental challenge that multiclass algorithms struggle.

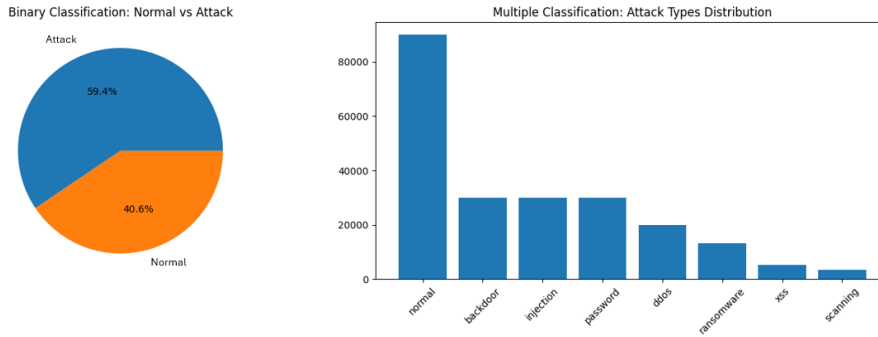


Fig. 2: Dataset Attack Profile.

C. ML Algorithm Selection & Evaluation Metrics

Given the cybersecurity application domain, our evaluation prioritized metrics reflecting operational security requirements. Our experimental design evaluated six ML algorithms selected for their complementary strengths and computational efficiency requirements. For binary classification, we emphasized Area Under the ROC Curve (AUC) as the primary metric, representing the system’s ability to distinguish between normal and malicious traffic across all possible decision thresholds. Additional metrics included precision (attack confidence), recall (attack detection rate), and specificity (normal traffic accuracy). For multiclass classification, we computed macro-averaged and weighted metrics to account for class imbalance effects. selective classification type performance was evaluated using per-class precision, recall, and F1-scores, enabling identification of attack types suitable for automated detection versus those requiring human oversight. The validation framework employed stratified train-test splits (70%-30%) to maintain representative class distributions across training and testing sets. Cross-validation was omitted due to computational constraints and the adequacy of the large test set (66,558 samples) for reliable performance estimation.

IV. EXPERIMENT AND RESULT

This section describes the experimental setup and the results obtained from feature selection and classification prediction using binary and selected multiple classification techniques. The experiments were conducted on a machine equipped with a 12GB VRAM GPU (RTX 3060), 120 GB of available storage, and a 16-core CPU, ensuring sufficient computational resources to handle the extensive data and model processing demands.

A. Binary Classification Performance

Binary classification for attack detection demonstrated excellent performance across all evaluated algorithms, with Random Forest achieving the highest overall performance at 84.47% accuracy and 90.76% AUC.

Detailed performance analysis revealed critical security metrics important for operational deployment. The Random Forest binary classifier achieved 92.48% attack detection rate (sensitivity), indicating that over 92% of actual attacks would

be successfully identified. The false alarm rate remained acceptably low at 7.52%, translating to approximately 300 false alarms per 4,000 normal events in operational scenarios. The precision of 83.3% indicates high confidence in attack predictions, with approximately 5 out of 6 attack alerts representing genuine attacks. All the models performance, including all the scores (precision, recall, and F1), is summarized in Table IV

TABLE IV: Binary Classification-All models Performance Summary

Model	Att. Prec.	Att. Rec.	Att. F1	Acc.	AUC
Random Forest	0.83	0.92	0.88	0.84	0.91
Logistic Regression	0.70	0.92	0.80	0.72	0.71
Gradient Boosting	0.73	0.99	0.84	0.78	0.87
Extra Trees	0.83	0.93	0.88	0.85	0.89
Linear SVM	0.70	0.92	0.80	0.72	0.71
Naive Bayes	0.66	0.75	0.70	0.62	0.63
Decision Tree	0.74	0.96	0.84	0.78	0.83

Feature importance analysis revealed that physical IoT telemetry data provides significant discriminative power for attack detection. One feature, `current_temperature`, emerged as the most critical feature, contributing 99.15% of the Random Forest’s decision-making process. This finding suggests that thermal anomalies serve as primary indicators of malicious activity in BAS environments. Figure 3 illustrates the attack detection performance of an ML-IDS technique on multimodal IoT telemetry data.

B. Multiclass Classification Challenges

Multiclass classification faced significant challenges due to severe class imbalance, achieving a maximum accuracy of 55.48% with Extra Trees. The confusion matrix analysis revealed systematic bias toward predicting the majority class (normal traffic), with most attack types suffering from poor recall rates below 40%. The SMOTE oversampling approach improved minority class recall but at the cost of overall accuracy, achieving 46.13% compared to the baseline 55.48%. This degradation occurred because synthetic sample generation could not capture the subtle feature patterns distinguishing different attack types, leading to increased confusion between

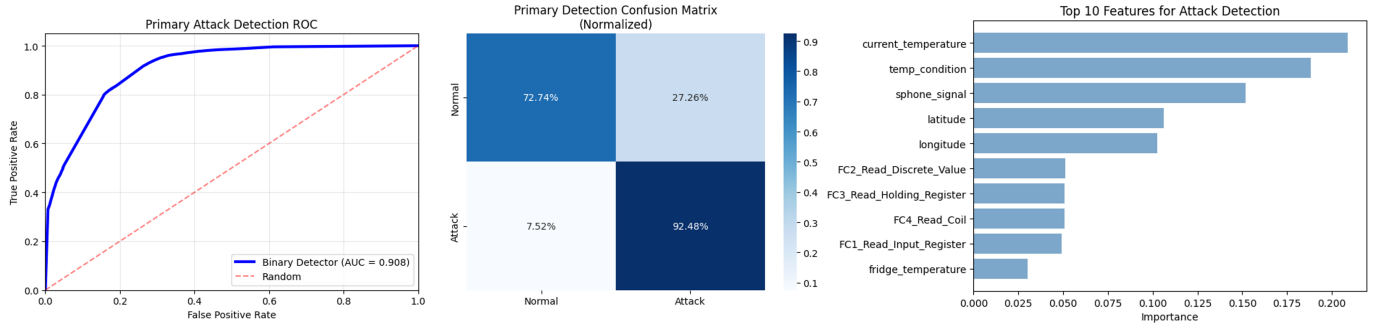


Fig. 3: Cyber-Attack Detection Performance.

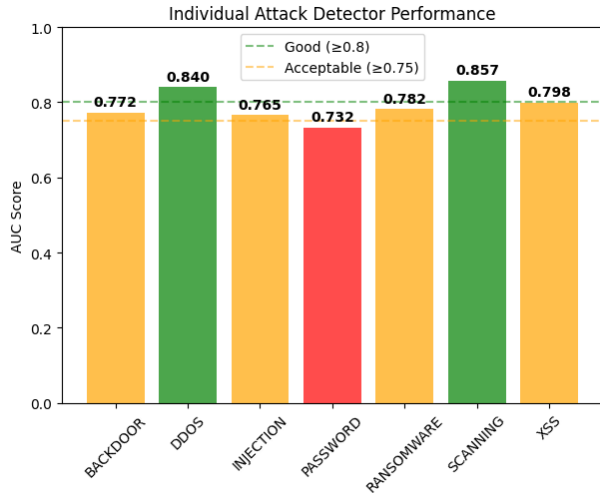


Fig. 4: Selective Classification: DDoS & Scanning show good performance (AUC > 0.84).

attack categories. Balanced Random Forest with class weighting achieved similar results (43.87% accuracy), confirming that the multiclass classification challenges stemmed from fundamental class separability issues rather than simply algorithmic limitations.

C. Selective Classification Performance

To address the challenges of multiclass classification, we propose a One-vs-Rest binary classification approach for selective attack types. This approach revealed significant performance variations across different attack categories. Scanning attacks achieved the highest detection performance with an AUC score of 85.7%, closely followed by DDoS attacks at 84.0%. These results indicate that certain attack types generate distinctive telemetry signatures that are well-suited for binary detection approaches.

Backdoor, XSS and ransomware attacks demonstrated moderate detection capability, with AUC scores of 77.2%, 79.8% and 78.2%, respectively. Injection attack showed acceptable performance, with AUC scores of 76.5%. But password attack is *NOT* acceptable because scored below 75%.

Moderately performing detectors, like those for backdoor and ransomware attacks, may function effectively with human oversight to ensure accuracy and reliability. Meanwhile, challenging attack types such as password attack may require additional measures, such as rule-based or manual analysis, to complement machine learning efforts. The selective classification detector performance is depicted in Figure 4.

D. Discussion

The attack detection system delivers strong performance, with an AUC of 0.908 demonstrating effective discrimination between attacks and normal activities. The normalized confusion matrix confirms high accuracy, with 92.48% of attacks detected correctly and 72.74% of normal activities recognized, although a false positive rate of 27.26% exists.

The experiment provided insights into the most critical IoT telemetry indicators for cybersecurity. Temperature-related features (`current_temperature`, `temp_condition`) consistently ranked highest, suggesting that thermal anomalies serve as primary attack indicators in smart building environments.

The device status features (`sphone_signal`) contributed moderate importance, indicating that changes in the state of the physical device correlate with certain types of attacks. Geographic features (`latitude`, `longitude`) showed minimal importance. The overall ML-IDS performance is summarized in Figure 3.

V. CONCLUSION

This research demonstrates the viability of machine learning-based intrusion detection systems (ML-IDS) for IoT smart building environments using physical telemetry data. Our comprehensive evaluation reveals that binary classification approaches using random forest achieve production-ready performance (84.47% accuracy, 90.76% AUC) for general attack detection, while multiclass classification faces significant challenges due to severe class imbalance inherent in the datasets.

The key finding that IoT devices' physical telemetry, particularly temperature sensors and device status indicators, provides sufficient discriminative power for effective cyber attack detection has important implications for IoT security architecture. Unlike traditional network-based intrusion detection systems requiring deep packet inspection and complex traffic

analysis, physical-based detection can operate effectively using simple device state information.

The research reveals important limitations in current multiclass classification approaches. The maximum multiclass accuracy of 55.48% with severe bias toward majority classes indicates that attack type identification requires specialized approaches. However, selective classification detectors achieving 80%+ AUC for specific threats (DDoS, scanning) suggest that targeted binary classification approaches can supplement general attack detection with threat-specific intelligence.

Several areas warrant future investigation. First, temporal feature engineering, incorporating time-series analysis of IoT telemetry streams, may improve attack detection accuracy by capturing behavioral patterns over time. Second, ensemble approaches combining multiple selective classification detectors may improve overall multiclass performance while maintaining interpretability.

Several areas warrant future investigation. First, incorporating time-series analysis of IoT telemetry streams may improve attack detection accuracy by capturing behavioural patterns over time. Second, future research should validate the telemetry-based feature analysis across diverse IoT security datasets, to establish broader applicability and robustness of the methodology across different network environments and attack scenarios.

REFERENCES

- [1] R.-A. Craciun, S. I. Caramihai, Mocanu, R. N. Pietraru, and M. A. Moisescu, "Hybrid Machine Learning for IoT-Enabled Smart Buildings," *Informatics*, vol. 12, p. 17, Mar. 2025. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.
- [2] A. Karunamurthy, K. Vijayan, P. R. Kshirsagar, and K. T. Tan, "An optimal federated learning-based intrusion detection for IoT environment," *Scientific Reports*, vol. 15, p. 8696, Mar. 2025. Publisher: Nature Publishing Group.
- [3] A. Qaddos, M. U. Yaseen, A. S. Al-Shamayleh, M. Imran, A. Akhunzada, and S. Z. Alharthi, "A novel intrusion detection framework for optimizing IoT security," *Scientific Reports*, vol. 14, p. 21789, Sept. 2024. Publisher: Nature Publishing Group.
- [4] A. Jaramillo-Alcazar, J. Govea, and W. Villegas-Ch, "Anomaly Detection in a Smart Industrial Machinery Plant Using IoT and Machine Learning," *Sensors*, vol. 23, p. 8286, Jan. 2023. Number: 19 Publisher: Multidisciplinary Digital Publishing Institute.
- [5] S. Birnbach, S. Eberz, and I. Martinovic, "Haunted House: Physical Smart Home Event Verification in the Presence of Compromised Sensors," *ACM Trans. Internet Things*, vol. 3, pp. 18:1–18:28, Apr. 2022.
- [6] N. I. Haque, M. Ashiqur Rahman, and H. Shahriar, "Ensemble-based Efficient Anomaly Detection for Smart Building Control Systems," in *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, pp. 504–513, July 2021. ISSN: 0730-3157.
- [7] A. Hamza, H. Habibi Gharakheili, T. Pering, and V. Sivaraman, "Combining Device Behavioral Models and Building Schema for Cybersecurity of Large-Scale IoT Infrastructure," *IEEE Internet of Things Journal*, vol. 9, pp. 24174–24185, Dec. 2022.
- [8] C. Cheh, U. Thakore, A. Fawaz, B. Chen, W. G. Temple, and W. H. Sanders, "Data-driven Model-based Detection of Malicious Insiders via Physical Access Logs," *ACM Transactions on Modeling and Computer Simulation*, vol. 29, pp. 1–25, Oct. 2019.
- [9] A. Murthy, M. R. Asghar, and W. Tu, "A lightweight Intrusion Detection for Internet of Things-based smart buildings," *SECURITY AND PRIVACY*, vol. 7, no. 4, p. e386, 2024. [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/spy2.386](https://onlinelibrary.wiley.com/doi/pdf/10.1002/spy2.386).
- [10] P. Rieger, M. Chilese, R. Mohamed, M. Miettinen, H. Fereidooni, and A.-R. Sadeghi, "ARGUS: Context-Based Detection of Stealthy IoT Infiltration Attacks," Feb. 2023. arXiv:2302.07589 [cs].
- [11] M. Bagaa, T. Taleb, J. B. Bernabe, and A. Skarmeta, "A Machine Learning Security Framework for Iot Systems," *IEEE Access*, vol. 8, pp. 114066–114077, 2020.
- [12] A. Lara, V. Mayor, R. Estepa, A. Estepa, and J. E. Díaz-Verdejo, "Smart home anomaly-based IDS: Architecture proposal and case study," *Internet of Things*, vol. 22, p. 100773, July 2023.
- [13] J. Ashraf, M. Keshk, N. Moustafa, M. Abdel-Basset, H. Khurshid, A. D. Bakhshi, and R. R. Mostafa, "IoTBoT-IDS: A novel statistical learning-enabled botnet detection framework for protecting networks of smart cities," *Sustainable Cities and Society*, vol. 72, p. 103041, Sept. 2021.
- [14] A. Javed, M. N. Awais, A.-u.-H. Qureshi, M. Jawad, J. Arshad, and H. Larijani, "Embedding Tree-Based Intrusion Detection System in Smart Thermostats for Enhanced IoT Security," *Sensors*, vol. 24, p. 7320, Jan. 2024. Number: 22 Publisher: Multidisciplinary Digital Publishing Institute.
- [15] S. Facchini, G. Giorgi, A. Saracino, and G. Dini, "Multi-level Distributed Intrusion Detection System for an IoT based Smart Home Environment," pp. 705–712, June 2025.
- [16] A. Javed, A. Ehtsham, M. Jawad, M. N. Awais, A.-u.-H. Qureshi, and H. Larijani, "Implementation of Lightweight Machine Learning-Based Intrusion Detection System on IoT Devices of Smart Homes," *Future Internet*, vol. 16, p. 200, June 2024. Number: 6 Publisher: Multidisciplinary Digital Publishing Institute.
- [17] Y. Majib, M. Barhamgi, B. M. Heravi, S. Kariyawasam, and C. Perera, "Detecting anomalies within smart buildings using do-it-yourself internet of things," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 4727–4743, May 2023.
- [18] M. Touma, S. Witherspoon, S. Witherspoon, and I. Crawford-Eng, "A practical approach for applying Machine Learning in the detection and classification of network devices used in building management,"
- [19] D. Fauri, M. Kapsalakis, D. R. dos Santos, E. Costante, J. den Hartog, and S. Etalle, "Leveraging Semantics for Actionable Intrusion Detection in Building Automation Systems," in *Critical Information Infrastructures Security* (E. Luiijf, I. Žutautaitė, and B. M. Hämmerli, eds.), (Cham), pp. 113–125, Springer International Publishing, 2019.
- [20] D. Fauri, M. Kapsalakis, D. R. dos Santos, E. Costante, J. den Hartog, and S. Etalle, "Role Inference + Anomaly Detection = Situational Awareness in BACnet Networks," in *Detection of Intrusions and Malware, and Vulnerability Assessment* (R. Perdisci, C. Maurice, G. Giacinto, and M. Almgren, eds.), (Cham), pp. 461–481, Springer International Publishing, 2019.
- [21] A. K. Pathak, S. Saguna, K. Mitra, and C. Åhlund, "Anomaly Detection using Machine Learning to Discover Sensor Tampering in IoT Systems," in *ICC 2021 - IEEE International Conference on Communications*, pp. 1–6, June 2021. ISSN: 1938-1883.
- [22] M. Wang, N. Yang, and N. Weng, "Securing a Smart Home with a Transformer-Based IoT Intrusion Detection System," *Electronics*, vol. 12, p. 2100, Jan. 2023. Number: 9 Publisher: Multidisciplinary Digital Publishing Institute.
- [23] R. Yasaei, Y. Moghaddas, and M. A. Al Faruque, "IoT-GRAF: IoT Graph Learning-Based Anomaly and Intrusion Detection Through Multi-Modal Data Fusion," in *2024 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 1–6, Mar. 2024. ISSN: 1558-1101.
- [24] Y. Wan, K. Xu, G. Xue, and F. Wang, "IoTArgos: A Multi-Layer Security Monitoring System for Internet-of-Things in Smart Homes," in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, pp. 874–883, July 2020. ISSN: 2641-9874.
- [25] F. Sadikin, T. v. Deursen, and S. Kumar, "A ZigBee Intrusion Detection System for IoT using Secure and Efficient Data Collection," *Internet of Things*, vol. 12, p. 100306, Dec. 2020.
- [26] A. Alsaedi, N. Moustafa, Z. Tari, A. Mahmood, and A. Anwar, "TON_iot Telemetry Dataset: A New Generation Dataset of IoT and IIoT for Data-Driven Intrusion Detection Systems," *IEEE Access*, vol. 8, pp. 165130–165150, 2020.
- [27] N. Moustafa, "New Generations of Internet of Things Datasets for Cybersecurity Applications based Machine Learning: TON_iot Datasets," 2019.
- [28] N. Moustafa, M. Keshky, E. Debiez, and H. Janicke, "Federated TON_iot Windows Datasets for Evaluating AI-Based Security Applications," in *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pp. 848–855, Dec. 2020. ISSN: 2324-9013.