

ZENA: Enhanced Anomaly-Based IDS for Smart Homes Automation Through Multilayer Artificial Neural Network

Nkuba Kayembe Carlos, Chanhee Park, Imatitikua Aiyanyo, Heejo Lee, and Heuseok Lim

Abstract—Smart home automation, a component of the Internet of things (IoT), enables users to manage home functions with smart sensors and actuators, providing convenience, energy efficiency, and remote monitoring. The Z-Wave protocol, widely adopted in smart homes for lighting, security, appliance control, and power management, remains vulnerable to various external attacks, highlighting the need for effective attack detection tools. The existing single-layer artificial neural network (ANN) model (i.e., ZMAD) performs well on familiar data; however, it has limitations over datasets with different distributions and advanced attack vectors. This paper introduces ZENA, a lightweight protocol-aware anomaly-based intrusion detection model for Z-Wave networks, which employs a multilayer ANN, built from scratch, to improve detection accuracy and robustness. Using a dataset with several attack classes and adversarial generated vectors, the proposed model achieves a precision rate of 95%, significantly outperforming ZMAD and state-of-the-art deep learning neural networks on same dataset (i.e., 89–93%). The results indicate substantial improvements in advanced detection and resilience against adversarial attacks, enhancing security for Z-Wave smart home systems.

Index Terms—Artificial neural network, deep learning, Internet of things (IoT), intrusion detection systems (IDS), machine learning, security, smart homes, Z-Wave.

I. INTRODUCTION

THE adoption of smart home automation within the Internet of things (IoT) has significantly enhanced user convenience, energy efficiency, and security through interconnected devices. Among the various IoT communication protocols, Z-Wave [1] has emerged as a widely used standard for home automation, supporting applications such as lighting control, security systems, and power management. As of 2025, over 100 million Z-Wave chipset modules have been integrated into smart home environments, demonstrating its widespread acceptance [2]. However, despite its growing presence, the security framework of Z-Wave remains a critical concern [3], [4], as vulnerabilities in these systems expose users to potential cyber threats.

Problem statement. Smart home systems (SHS) collect and transmit sensitive user data through a network of sensors, actuators, and controllers. While these systems offer convenience,

they also introduce significant security risks. Many smart home devices lack robust security features, making them susceptible to attacks such as sniffing, spoofing, replay, and injection attacks [5], [6]. Additionally, manufacturers often prioritize rapid deployment over thorough security testing, leading to the commercialization of devices with inadequate security measures [7], [8]. The growing prevalence of cyberattacks targeting smart home IoT devices underscores the urgent need for advanced intrusion detection mechanisms to mitigate these risks effectively.

Limitations of existing approaches. Anomaly-based intrusion detection systems (IDS) play a crucial role in safeguarding IoT networks by identifying malicious activities [9]. One of the existing solutions, the ZMAD model [5], was developed for Z-Wave anomaly detection using a single-layer artificial neural network (ANN). While ZMAD provides a foundation for wireless security, it suffers from several limitations. First, its reliance on a single-layer ANN restricts its ability to detect complex and evolving attack. Second, it lacks robustness against adversarial inputs [10], which can be manipulated to evade detection by mimicking normal traffic patterns. Lastly, its effectiveness diminishes when applied to large-scale datasets with varying distributions, limiting its generalization capabilities (see Table VI). Prior IoT IDS studies [11]–[15] emphasize metaheuristic feature selection and generic traffic modeling or blended deep networks that trade transparency for aggregate accuracy. These approaches do not focus on the Z-Wave protocol nor do not model Z-Wave's packet field-level semantics.

Challenges. The nature of Z-Wave traffic introduces additional challenges for machine learning-based IDS solutions. Due to the protocol's structured communication patterns, distinguishing between normal and attack traffic requires a model that can learn deep feature representations. Furthermore, adversarial attacks pose a significant threat to IDS systems by introducing imperceptible perturbations into input data, leading to weaker classifications even by state-of-the-art deep learning (DL) neural network models (see Section VI.D). Traditional single-layer ANN models (e.g., ZMAD) struggle to capture such nuanced variations, making them vulnerable to sophisticated evasion techniques. Addressing these challenges requires a more advanced ANN architecture that can effectively generalize across diverse attack scenarios and maintain robustness against adversarial manipulations.

Our approach. To overcome the limitations of existing IDS models, we propose ZENA (Z-Wave enhanced anomaly-based detection system), a novel lightweight protocol-aware intrusion detection model built from scratch on a multilayer ANN architecture to fit the Z-Wave protocol specification

Manuscript received June 3, 2025; revised January 13, 2026; approved for publication by Park, Hyunggon, Division 3 Editor, January 13, 2026.

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP), grant funded by the Korea government (MSIT) (No. RS-2024-00440780 Development of Automated SBOM and VEX Verification Technologies for Securing Software Supply Chains), and ICT Creative Consilience Program (No. IITP-2025-RS-2020-II201819).

The authors are with the Department of Computer Science and Engineering, Korea University, Seoul, 02841 Korea, email: {carlosnkuba, pch7678, titi, heejo, limhseok}@korea.ac.kr.

H. Lee is the corresponding author.

Digital Object Identifier: 10.23919/JCN.2025.000108

Creative Commons Attribution-NonCommercial (CC BY-NC).

This is an Open Access article distributed under the terms of Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided that the original work is properly cited.

TABLE I
REQUIRED LIBRARIES FOR TRAINING AND DEPLOYMENT [5].

	Deep learning	Machine learning	ZENA
Model training	Numpy, Pandas, TensorFlow, Keras	Numpy, Pandas, Sklearn	Numpy, Pandas
Model export	Numpy, Keras	Numpy, Joblib	Numpy
Deployment	Numpy, Panda, TensorFlow, Keras	Numpy, Panda, Sklearn	Numpy, Pandas

and requirement. Unlike previous models, ZENA incorporates Xavier initialization, momentum-based training, and adversarial training techniques to enhance its ability to detect and classify sophisticated threats targeting Z-Wave networks. By leveraging a deeper artificial neural network structure developed from square one, ZENA significantly improves classification accuracy and robustness against adversarial perturbations.

Unlike prior intrusion detection systems that employ generic machine learning (ML) or DL models, ZENA introduces a fully customized artificial neural network built from scratch using only Python and NumPy. This design is lightweight and eliminates heavy external dependencies required by ML and DL during training, testing, and deployment (see Table I). Moreover, ZENA lightweight exported ANN models (64 KB) allows possible integration with Z-Wave gateways (e.g., Raspberry Pi).

ZENA's novelty lies in its protocol-aware learning strategy, where each neuron learns the behavior of specific Z-Wave packet fields, capturing hierarchical dependencies exploited by adversarial attacks. Unlike conventional black-box ML/DL models, ZENA is transparent and interpretable as it is built from the ground up; the IDS operator can trace which field contributions led to a packet being classified as normal or abnormal.

In addition, ZENA's architecture was trained on the newly constructed dataset comprising 74,505 real and adversarially crafted Z-Wave packets. This integration of lightweight design, explainability, and adversarial robustness establishes ZENA as a fundamentally new and Z-Wave protocol-specialized IDS.

Evaluation summary. The effectiveness of ZENA is validated through rigorous testing on real-world Z-Wave smart home network traffic. Our experiments demonstrate that ZENA achieves a precision rate of 95% during testing, significantly outperforming the state-of-the-art deep learning neural networks (see Table VII) while trained on the same dataset. The results indicate superior performance in intrusion detection and resilience to adversarial attacks. Furthermore, comparative analysis between ZENA and ZMAD highlights the enhanced generalization capability of the proposed model, making it a more reliable solution for securing smart home IoT environments (see Table VI).

Contributions. This paper makes the following key contributions:

- We introduce ZENA, a protocol-aware lightweight multilayer ANN anomaly-based intrusion detection system designed to enhance security in Z-Wave IoT networks.

- ZENA incorporates Xavier initialization, momentum-based training, and early stopping techniques to improve stability and convergence efficiency.
- ZENA is evaluated on a comprehensive Z-Wave dataset generated from real Z-Wave devices, demonstrating a detection accuracy of 95% with superior generalization compared to existing models.
- To advance research in this field, we provide an open dataset comprising 74,505 realistic Z-Wave packet samples [16], supporting further development of robust security solutions for smart home automation.

II. Z-WAVE PROTOCOL

This section presents the debut and development timeline of the Z-Wave protocol, along with its architecture and security modes. We also explain the motivation for developing ZENA, a custom-built ANN tailored specifically for Z-Wave networks.

A. Historical Background

Z-Wave wireless home control protocol was developed in 2001 by Zensys, a Danish company, to support smart home. It is a low bandwidth, low-power mesh protocol that operates in the 900 MHz frequency range and do not interfere with other house electronics using Wi-Fi and Bluetooth in the 2.4 GHz. The first commercialized Z-Wave chipset 100 series was released in 2003. In 2005 several manufactures in collaboration with Zensys formed the Z-Wave Alliance [17].

In 2008 Zensys is acquired by Sigma Designs, a United States (U.S.) based company. In 2013 Z-Wave chipset 500 series also known as Z-Wave Plus is released with an increase in memory, wireless range, and battery life. The same year a new certification is established to test and certify Z-Wave devices before commercialization [18]. In 2017, Sigma Designs opens the Z-Wave public specification [19] that provides information about the interoperability layer for developing smart home apps. In April 2017, Z-Wave Security 2 (S2) is mandate on any new Z-Wave product. S2 is a more robust security mode designed to protect against attacks. However, due to its backward compatibility, S2 devices can be downgraded to less secure modes during the pairing process [20]. Moreover, since smart home devices interoperate within the same network, an attacker could still exploit an S2 device indirectly through vulnerable legacy devices [5] and Z-Wave specification [8].

In April 2018 Sigma Designs was acquired by Silicon Labs and in December of the same year Z-Wave 700 chipset Series is released [21]. 700 series chipset provides to developer the possibility to create intelligent smart home devices that implement S2 security while reducing energy consumption, boosting performance, and increase radio frequency (RF) range.

The latest Z-Wave chipset, the 800 series, offers long-range connectivity (up to 1.5 miles) and low-power wireless communication. It is designed for smart home applications, providing up to 10 years of battery life and protection through the S2 security framework [22].

As of 2025, the Z-Wave ecosystem is an alliance of more than 800 companies producing more than 4,500 interoperable

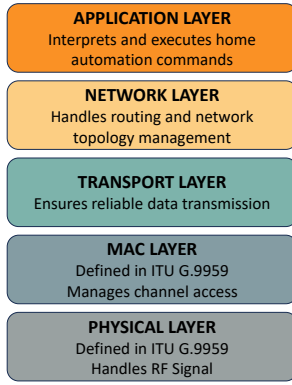


Fig. 1. Z-Wave protocol stack listing 5 layers usage.

Z-Wave certified product types with the global market of more than 100 million devices manufactured [2].

B. Protocol Overview

Z-Wave protocol comprises of 5 layers, which are physical (PHY), medium access control (MAC), transport, network, and application layer (see Fig. 1). The protocol implements the open International Telecommunication Union-Telecommunication Standardization (ITU-T) Sector G.9959 at the PHY and MAC layers [23]. The PHY layer is responsible of modulation and RF transceiver control. The MAC layer oversees collision avoidance algorithm. The transport layer manages frames transmission and reception, error detection and retransmission. The network layer handles network routing and/or encryption. The application Layer provides the interoperability between devices from different manufactures and manages frame payloads. Fig. 1 shows the Z-Wave protocol stack.

Fig. 2 provides a summary of the Z-Wave packet frame structure. The maximum MAC frame size is 64 bytes. The MAC frame comprises important information for the Z-Wave device communication such as the home-ID (H-ID), source-ID (SRC), frame control (P1, P2), length (LEN), destination-ID (DST), application payload consisting of command class (CMDCL), command (CMD), and parameters (PARAMs). The MAC frame provides also a checksum (CS) for packet verification [3], [24].

C. Z-Wave Security Scheme

Z-Wave devices implements several security schemes while exchanging data. The Z-Wave security modes is listed as below [3], [5], [24]:

- **No Security:** In this mode devices exchange Z-Wave packets without data encryption and authenticity. The data encapsulation is ensured by CS-8/CRC-16 checksum. Legacy Z-Wave devices used in smart houses such as smart plug-in switch, door sensors, siren, motion sensors, do not support encryption by default and are vulnerable to several attacks [7].

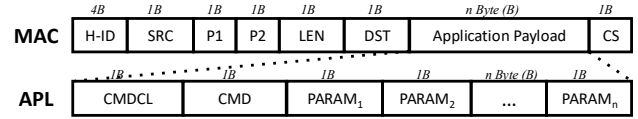


Fig. 2. Z-Wave basic frame structure.

- **Security 0 (S0):** enables the transfer of encrypted *application layer* data between devices using AES-128; however, the network key can be computed by the attacker as the key exchange mechanism (KEX) uses a fixed temporary key of 16 zeros (0x00) to encrypt the network key. A study in [25] revealed that the key is obtainable by a man-in-the-middle during initial pairing operation.
- **Security 2 (S2):** this mode implements Elliptic Curve Diffie-Hellman (ECDH) for secure key derivation and exchange to fix the S0 KEX limitation. Integrity checking and replay protection are ensured using AES-128-CMAC and Pre-Agreed Nonce (PAN). S2 is regrouped into three subclasses: S2 Access Control used by devices such as door locks and garage doors; S2 Authenticated for lighting and sensors managed systems; and S2 Unauthenticated for lighting and sensor unmanaged systems. These subclasses have private keys that reinforce their security. As of today, devices implementing S2 are supposed to be secured. However, due to backward compatibility new secure S2 devices should communicate with old legacy ones. A study in [20] found that S2 devices can be downgraded to weak S0 security at device inclusion/pairing stage. Moreover, there are still attack vectors that affect S2 Z-Wave controllers [5], [8].

As of 2025, fewer than one-fourth of the 4,500 unique Z-Wave products support the stronger S2 security mechanisms [26]. Over 100 million Z-Wave devices have been manufactured globally [2], with more than 50% classified as Z-Wave Plus. Alarmingly, millions of Z-Wave Plus and *legacy* devices currently used in smart homes lack encryption support and remain vulnerable to external attacks [3], [8], [27]. Therefore, deploying an IDS to secure these devices is critically necessary.

D. Justification of ZENA, A Custom-Built ANN Design for Z-Wave Networks

In Z-Wave smart home networks, intrusion detection requires efficient classification of wireless packets as normal or attack. Traditional ML and DL models such as Random Forest, XGBoost, SVM, and deep networks like multi-layer perceptron (MLP), convolutional neural network (CNN), or long short-term memory (LSTM) not only require Python, Numpy, and Pandas, but also demand large frameworks, such as scikit-learn (≈ 400 MB), XGBoost (≈ 500 MB), and TensorFlow (≈ 600 – 1000 MB), which significantly increase deployment size and memory usage (see Table I). In contrast, ZENA, implemented entirely with Python and NumPy/Pandas, eliminates heavy dependencies and stores only compact weight matrices after training (a few 64 KB). This minimalism makes it ideal for constrained environments like Z-Wave gateways

TABLE II
COMPARISON OF Z-WAVE SECURITY RESEARCH.

Research	Focus area	Target aspects	IDS type	Adversarial attacks
[28] (2015)	Introduce an unauthorized controller into a Z-Wave network.	Network inclusion process and controller authentication.	No	No
[29] (2016)	Intrusion detection that matches known attack patterns (misuse signatures).	Z-Wave packet-level attack signatures.	Signature-based	No
[30] (2017)	Z-Wave routing protocol analysis.	Network layer (Z-Wave routing frames).	No	No
[31] (2018)	Z-Wave transceiver memory analysis.	Z-Wave device memory (firmware of 300-series Z-Wave chip).	No	No
[32] (2020)	Comprehensive threat analysis on Z-Wave protocol.	Inclusion process and network key exchange.	No	No
[33] (2020)	Denial of service attacks & exploits.	Z-Wave frame protocol (unencrypted control frames (e.g., Nonce Get).	No	No
[34] (2020)	Investigated whether encrypted Z-Wave traffic still leaks information via side channels.	Physical layer timing/leakage and network traffic patterns (Z-Wave frames).	No	No
[3] (2022)	Automated vulnerability discovery via fuzz testing.	Z-Wave protocol stack (various Z-Wave versions up to 700-series).	No	No
[5] (2023)	Developed ZMAD, an IDS that learns normal Z-Wave traffic and detects abnormal one.	Z-Wave frame sequences.	Single-layer ANN Anomaly-based IDS	No
Proposed ZENA	Developed ZENA, a multi-layer IDS effective for advanced Z-Wave attacks.	Deep analysis of the Z-Wave frame structure to detect advanced attacks.	Multi-layer ANN Anomaly-based IDS	Yes

(e.g., Raspberry Pi) and IoT controllers, offering low memory consumption, fast startup times, and CPU-only execution without requiring GPUs or external runtime.

Unlike generic neural networks, ZENA's architecture is protocol-aware, built specifically around Z-Wave packet fields structures. By embedding domain knowledge into the ANN design, ZENA achieves high detection accuracy with fewer layers and parameters, reducing computational cost and overfitting risks. This domain-specific optimization ensures robust anomaly detection while maintaining a small footprint, making ZENA not only efficient but also practically deployable for real-time, on-device security in smart home ecosystems.

III. RELATED WORKS

Over the past years, academic research on Z-Wave security has evolved from uncovering basic vulnerabilities in this proprietary smart home protocol to developing sophisticated defense mechanisms. Early works by Fouladi and Ghanoun [25] examined the security of the Z-Wave protocol by focusing on its encryption mechanisms. Other research focused on reverse-engineering the closed Z-Wave protocol and demonstrating fundamental security flaws. For example, researchers in [28] showed how a rogue controller could be secretly introduced into a Z-Wave network to intercept or manipulate communications. These initial studies highlighted the lack of strong authentication in early Z-Wave devices, where encryption was optional and messages were largely trusted by default.

As Z-Wave adoption grew, the community pushed for improved security, notably the introduction of AES-128 encryption (S0) and the S2 security mode around 2017. However, research in [27] revealed that many devices still operated with S0 insecure inclusion or other weaknesses, leaving networks vulnerable if encryption was not enabled.

Later, research expanded into both offensive and defensive directions. On the offensive side, studies systematically analyzed Z-Wave's protocol layers. A study in [30] reverse-engineered the network (routing) layer and discovered that Z-Wave's mesh routing trusted network messages blindly; an attacker could exploit this to alter the network topology and even launch a blackhole attack (dropping messages) in a real Z-Wave network. Other works examined device firmware and memory [31], managing to extract secret keys from Z-Wave devices memory, since early device firmware often stored keys in plaintext. These insights painted a concerning picture: even as Z-Wave introduced security enhancements, implementation flaws and legacy behaviors left many devices at risk.

Defensive research started addressing how to detect or prevent attacks. In 2016, researchers [29], [35] built a misuse-based intrusion detection system tailored for Z-Wave, effectively using known attack signatures to catch malicious events on the network. This effort was one of the first to adapt traditional IDS concepts to the Z-Wave IoT wireless domain, achieving higher detection rates for defined attacks. However, a limitation was that it could only detect known attack patterns (e.g., replay or rogue inclusion attempts that had been seen before).

As Z-Wave's S2 security became standard, researchers shifted toward both new types of attacks and more advanced defenses. On the attack front, academics uncovered subtle but serious issues even in S2 networks. For instance, a 2020 study [33] demonstrated that certain control frames (like Nonce Get used in secure handshakes and a network management frame Find Nodes in Range) could be manipulated to freeze or disable entire Z-Wave networks via denial-of-service (DoS). Another study [34] exposed privacy leaks. Even if Z-Wave traffic is encrypted, side-channel characteristics (timing, frequency of messages) can reveal sensitive

information. A demonstration showed that an eavesdropper could identify what types of Z-Wave devices are present in a home (e.g., door locks vs. lights) just by passively sniffing the radio traffic patterns, which is a clear privacy risk.

Another important development was the use of fuzz testing to proactively find vulnerabilities. The authors in [3] created a protocol-aware fuzzer (“VFuzz”) specifically for Z-Wave, which systematically probed devices and uncovered numerous previously unknown flaws. Notably, this led to the disclosure of multiple CVEs affecting even newer Z-Wave 500 and 700 series chipsets [7]. This indicates a maturation of the field: researchers are not only reacting to known problems but also anticipating issues by stress-testing the protocol.

Concurrently, defensive techniques grew more sophisticated. Researchers recognized the need for anomaly-based IDS (to catch novel attacks, not just known signatures). By 2023, a study in [5] introduced a lightweight anomaly detector that learns the normal profile of Z-Wave traffic and flags deviations. This reflects a trend toward artificial intelligence and model-based security for smart homes, given the diversity of potential attacks.

However, recent developments in adversarial machine learning have demonstrated the significance of strong anomaly detection models capable of mitigating evasion attacks. Researchers in [36] introduced techniques for evaluating neural network robustness against adversarial inputs, while work in [37] demonstrated the feasibility of adversarial attacks in physical environments. These works emphasize the importance of designing models that can withstand adversarial manipulation.

Building on these foundations, our ZENA framework introduces a lightweight multi-layer artificial neural network with adversarial training specifically tailored for Z-Wave packet anomaly detection. Unlike prior research, which only addresses general IoT networks or focuses on specific vulnerabilities, our approach employs a robust architecture capable of handling both structured Z-Wave data and realistic noise caused by advanced attacks. ZENA aims to improve the reliability and security of smart home environments by combining insights from previous Z-Wave security research with insights in adversarial machine learning. Table II summarizes key research development on Z-Wave smart home protocol.

IV. METHODOLOGY

The objective of ZENA framework is to provide a robust, lightweight, and scalable anomaly-based intrusion detection system tailored for Z-Wave IoT networks. Unlike its predecessor ZMAD, which employed a single-layer ANN and exhibited limited generalization on diverse datasets, ZENA addresses this limitation through multi-layers ANN and adversarial training. The goal is to accurately identify malicious traffic, even under adversarial conditions, while maintaining high precision on normal traffic.

A. Data Collection and Pre-Processing

To construct a representative dataset for Z-Wave network behavior, we purchased several devices (see Table III) and

TABLE III
TESTED DEVICE DETAILS INFORMATION.

ID	Brand name	Device type	Model (year)	Encryption support*
D1	Samsung	Controller	ET-WV520 (2017)	Yes
D2	Schlage	Smart door lock 1	BE469ZP (2019)	Yes
D3	Schlage	Smart door lock 2	BE468 (2014)	Yes
D4	Aeotec	Motion sensor 1	ZWA005-A (2019)	Yes
D5	Aeotec	Motion sensor 2	ZW100-A (2017)	No
D6	Linear	Motion sensor 3	WAPIRZ-1 (2015)	No
D7	Aeotec	Contact sensor 1	ZWA008 (2019)	Yes
D8	Linear	Contact sensor 2	WAPIRZ-1 (2015)	No
D9	Linear	Contact sensor 3	WAPIRZ-1 (2015)	No
D10	Jasco	Smart switch 1	ZW4201 (2016)	No
D11	Jasco	Smart switch 2	ZW4201 (2016)	No
D12	Aeotec	Siren 1	ZW164-A (2019)	Yes
D13	Dome	Siren 2	DM501 (2016)	No
D14	Linear	Dimmer LED	LB60Z-1 (2014)	No

*Encryption support: whether or not the device supports data encryption.

employed the Zniffer tool [38], developed by Silicon Labs, to capture live wireless traffic from a heterogeneous set of Z-Wave devices. The experimental testbed consisted of 14 Z-Wave devices, including door locks, motion and contact sensors, sirens, light switches, dimmers, and a controller, distributed across two physical rooms. Zniffer operates as a passive RF sniffer, monitoring the sub-GHz spectrum used by Z-Wave and capturing all broadcast and unicast communications. Each intercepted frame was logged and particular emphasis was placed on preserving the protocol’s integrity by capturing native Z-Wave frame structures in hexadecimal format for downstream analysis.

Following acquisition, the raw Z-Wave traffic data underwent a rigorous pre-processing phase to prepare it for machine learning applications. The raw hexadecimal payloads, constituting the primary communication content between Z-Wave nodes, were extracted and converted into their decimal equivalents. These values were subsequently parsed into 30 fixed-format fields, adhering to the Z-Wave protocol frame structure, including fields such as SRC, P1, P2, LEN, DST, CMDCL, CMD, and PARAM1–PARAM23 (see Fig. 2). This transformation normalized packet representations across device types and traffic contexts. Non-informative or redundant metadata (e.g., line number or timestamp) was excluded from modeling. The result was a clean, numeric matrix where each row represented a standardized Z-Wave packet instance.

To support supervised learning, each packet instance was annotated with a binary label indicating its class: normal or abnormal. Normal traffic instances were derived from legitimate, real-time interactions among Z-Wave devices within the testbed environment. For abnormal data, we applied fuzzing-based adversarial traffic generation, leveraging methodologies described in the VFuzz framework [3]. These adversarial samples introduced carefully crafted perturbations to simulate realistic attack vectors, including spoofing, jamming, replay, flooding, and injection attacks. The adversarial samples were indistinguishable at the surface level but introduced subtle protocol violations or anomalies detectable at the feature level. The final dataset comprised both benign and adversarial

instances, balanced to ensure robust classifier performance.

B. Z-Wave Data Representation

To support the development of ZENA, a purpose-built artificial neural network tailored for Z-Wave intrusion detection, it is essential to properly structure Z-Wave packets. Therefore, each packet must undergo formalization before it can be effectively processed by the model. During training, each Z-Wave packet instance (p) is represented as a feature vector:

$$p = [f_1, f_2, f_3, f_4, \dots, f_{n-1}, f_n], \quad (1)$$

where f_i represents a specific field extracted from the Z-Wave packet.

After preprocessing, all Z-Wave packets in the dataset share a uniform field length, enabling the construction of a structured data matrix. Initially, this matrix is organized with rows representing individual packets and columns denoting features. However, to optimize matrix operations during neural network computations, the dataset is transposed.

As a result, the final input matrix X is arranged as an $n \times m$ matrix, where n is the number of features and m is the number of packet instances. In this configuration, each column corresponds to the feature vector of a single packet, enabling efficient batch processing during forward propagation.

$$X = \begin{bmatrix} f_{11} & f_{21} & \cdots & f_{m1} \\ f_{12} & f_{22} & \cdots & f_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ f_{1n} & f_{2n} & \cdots & f_{mn} \end{bmatrix}. \quad (2)$$

C. Proposed ZENA Framework

The core classification engine of the ZENA system is a multilayer artificial neural network optimized for binary anomaly detection. The model architecture consists of an input layer of 30 neurons (see Fig. 3 and Section IV.A) corresponding to the extracted Z-Wave packet features, two fully connected hidden layers of varying widths (optimized experimentally), and a single output neuron with a sigmoid activation function. The network was initialized using the Xavier method and trained using mini-batch gradient descent with momentum to enhance convergence speed and model stability. The ReLU function was used for hidden layer activations, while the binary cross-entropy loss function guided the training process. Additional strategies such as early stopping and log smoothing were employed to mitigate overfitting and numerical instability. Adversarial training, incorporating both normal and perturbed inputs, was integrated into the training dataset to improve the model's resilience against evasion tactics.

D. Model Architecture

Fig. 3 illustrates the ZENA model architecture. ZENA consists of two fully connected hidden layers, making it a lightweight yet robust architecture for classification tasks. It takes Z-Wave signals as input, processes them through its neural network layers, and predicts whether the signal is normal or abnormal based on learned patterns in the dataset.

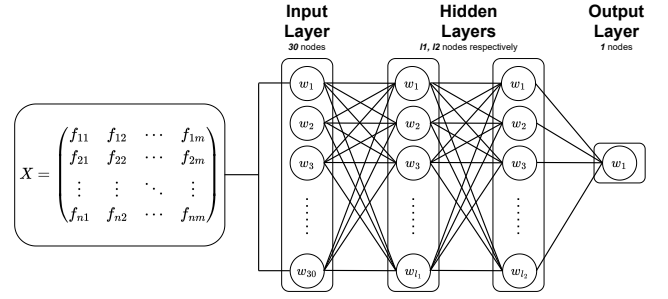


Fig. 3. ZENA multi-layers architecture.

The ZENA neural network is designed to efficiently process datasets with a focus on adaptability and performance. Given that the training dataset for ZENA (see Fig. 4) is more than seven times larger than that of ZMAD, a single-layer model was insufficient for effective learning.

ZENA comprises four layers: an input layer, two hidden layers for enhanced feature extraction, and an output layer. This architecture is particularly suited for complex data patterns, providing a robust framework for classification tasks.

- *Input layer.* The input layer consists of n nodes, corresponding to the number of features in the dataset. In this paper, n is fixed at 30.
- *Hidden layers.* We experiment with different numbers of parameters for two hidden layers; l_1 and l_2 . The performance is shown in Table IV.
- *Output layer.* The output layer has a single neuron when dealing with binary classification tasks. It uses the Sigmoid activation function to produce a probability score for binary classification.

1) *Forward propagation:* in forward propagation, ZENA first computes the dot product of the inputs X and initial weights W , then adds bias b . This can be expressed as follows:

$$Z_1 = W_1 \cdot X + b_1. \quad (3)$$

Then, the weighted sum obtained above passes an activation function that uses the rectified linear unit (RELU) defined as $RELU(x) = \max(0, x)$. Therefore, the activation is shown as follows:

$$A_1 = RELU(Z_1). \quad (4)$$

Then, the same equation is used to calculate Z_2 , A_2 for the second hidden-layer, and Z_3 . The Z_3 value aggregates all the information stored in each layer and generates a logit for the final output. We put this in the Sigmoid function defined as $SIG(x) = \frac{1}{1+e^{-x}}$ to obtain the model prediction.

$$\hat{Y} = SIG(Z_3). \quad (5)$$

2) *Backward propagation:* backward propagation computes the necessary adjustments for each weight and bias in the network to minimize the loss. First, we calculate the error between the predicted output $\hat{Y}$ and the true label Y .

$$dZ_3 = \hat{Y} - Y. \quad (6)$$

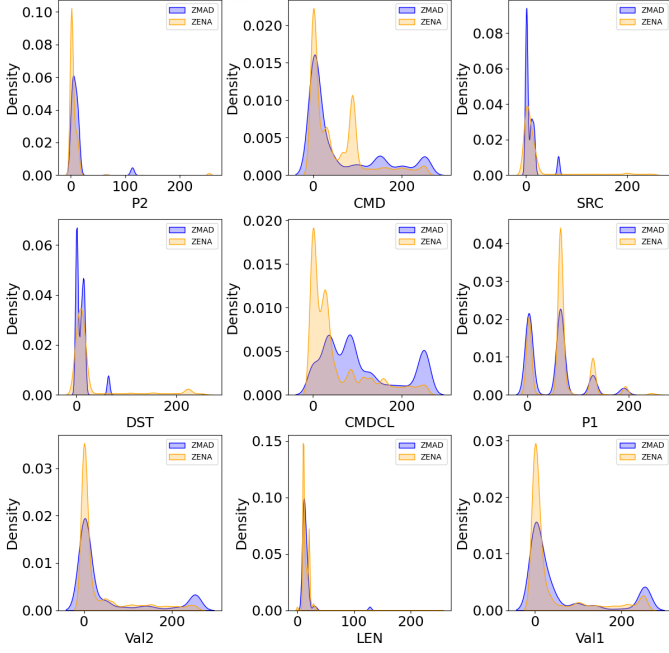


Fig. 4. KDE plot for ZMAD and ZENA datasets. This plot captures the normalized distribution comparison of 9 features between two datasets (ZMAD and ZENA). Each subplot represents the density distribution of a specific feature, allowing for visual comparison of how each dataset's feature values are spread. The blue and orange areas indicate the density of ZMAD and ZENA datasets, respectively.

Using this error, gradients for weights W_3 and biases b_3 are calculated as follows:

$$dW_3 = \frac{1}{m} dZ_3 \cdot A_2^T, \quad (7)$$

$$db_3 = \frac{1}{m} \sum_d Z_3. \quad (8)$$

Then, we can get the gradient for the second hidden layer W_2 as follows:

$$dZ_2 = W_3^T \cdot dZ_3 \odot \text{ReLU}'(Z_2), \quad (9)$$

where $\odot$ indicates element-wise multiplication and $\text{ReLU}'(x)$ means the derivative of ReLU activation with respect to input x . Thereafter, we get all the gradients dW_1 , db_1 , dW_2 , db_2 , dW_3 , db_3 , and update the parameters to reduce errors in the next iteration.

E. Training Strategy

1) *Initialization*: while ZMAD utilized random initialization, ZENA explored various initialization methods including He, Xavier, and zero initialization [39], [40]. Among these, Xavier initialization demonstrated the best performance (see Fig. 10).

$$W \sim N(0, \frac{1}{n_{input}}). \quad (10)$$

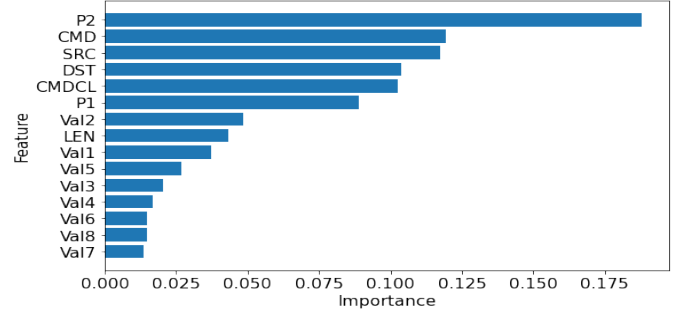


Fig. 5. Top 15 RF feature importance of ZENA dataset.

2) *Momentum and early stopping*: ZENA's training strategy involved expanding the model only by the necessary number of parameters to enhance predictive robustness. The use of momentum and early stopping further improved training efficiency [41], [42]. As a result, ZENA achieved superior learning performance and accuracy with only 20k epochs compared to ZMAD's 85k epochs.

3) *Cross-entropy loss and log smoothing*: since ZENA's training objective is to predict a binary classification, we implemented cross-entropy loss and log smoothing by adding $\epsilon = 1e - 10$ to prevent excessive computing loss.

F. Adversarial Training Techniques

Adversarial training was applied by adding adversarial samples into the training dataset. These samples, generated through fuzzing techniques, are crafted to exploit weaknesses in the ANN model, thereby enhancing its robustness by teaching it to recognize and respond to subtle variations intended to bypass detection.

G. Training and Testing Process

The data was split into training, validation, and test sets with an 80–10–10 ratio, employing cross-validation for fine-tuning.

V. EXPERIMENTAL SETUP

In this section, we describe the experimental setup used to evaluate the effectiveness and performance of ZENA framework towards predicting normal vs abnormal Z-Wave packets.

A. Testbed

We created a realistic Z-Wave network with diverse devices and captured their traffic. Testbed devices are listed in Table III.

The ZENA dataset includes heterogeneous traffic collected from real Z-Wave devices, including a primary controller, multiple slave nodes, actuators, and sensors. It contains both benign (normal) and adversarial (attack) traffic. The adversarial data were generated using a combination of black-box and gray-box techniques. Specifically, we leveraged known Z-Wave vulnerabilities, the official Z-Wave protocol specification, and smart fuzzing-based [3], [8] perturbations to manipulate critical packet fields (e.g., CMDCL, CMD, PARAMs, P1,

TABLE IV
MODEL PERFORMANCE WITH DIFFERENT SIZES.

Number of parameters	Params in /1	Params in /2	Learning rate	Training epochs	Accuracy	F1 score	Momentum
3k	64	16	0.03	15k	0.9341	0.9342	0.9
4k	64	32	0.03	20k	0.9330	0.9329	0.9
6k	64	64	0.03	30k	0.9476	0.9473	0.9
6k	128	16	0.03	23k	0.9452	0.9452	0.9
8k	128	32	0.03	23k	0.9518	0.9517	0.9
12k	128	64	0.03	21k	0.9512	0.9511	0.9
20k	128	128	0.03	20k	0.9491	0.9489	0.9
12k	256	16	0.03	22k	0.9468	0.9466	0.9
16k	256	32	0.03	20k	0.9388	0.9385	0.9
24k	256	64	0.03	16k	0.9469	0.9467	0.9
40k	256	128	0.03	16k	0.9499	0.9498	0.9

P2) while preserving protocol validity. This approach allowed us to simulate realistic evasion and manipulation attempts that closely resemble adversarial attacks in operational Z-Wave networks.

B. Dataset and Attack Classes

1) *Normal dataset*: the dataset of normal traffic is retrieved from devices' normal data exchange.

We converted the raw sniffed data into CSV file. The key important fields in the dataset are SRC, P1, P2, DST, CMDCL, LEN, CMD, and PARAM1 - PARAM23, all of which consist of the 30 features used for ZENA's input.

2) *Attack dataset*: we generated attack traffic using fuzzing [3], [8] techniques. The enhanced dataset consists of five attack classes, including injection, jamming, replay, spoofing, and flooding attacks. Adversarial samples were created to simulate realistic evasion attacks. They are more challenging for neural networks to spot because they look like normal traffic with minimal differences. These attacks aim to test the Z-Wave implementation robustness.

Fig. 4 helps identify differences in feature distributions for original ZMAD [5] vs ZENA datasets, which is crucial for understanding how the datasets may affect the performance and robustness of models trained on them. Such analysis is essential when evaluating the impact of adversarial samples and ensuring that the neural network can distinguish between normal and attack traffic effectively.

3) *Preliminary dataset analysis*: the ZENA dataset, publicly available at [16], contains 74,505 Z-Wave packet instances described by 31 features, including SRC, DST, P1, P2, LEN, CMDCL, CMD, and 23 application-layer payload bytes, along with a binary label indicating normal or attack traffic. This feature structure captures both protocol semantics and low-level byte variations, providing a comprehensive basis for evaluating machine learning models.

Machine learning feature importance analysis (i.e., Random Forest) shows that protocol-level attributes, particularly P2, CMD, SRC, DST, CMDCL, and P1 serve as the most influential predictors of malicious activity, indicating that structural and semantic protocol information plays a primary role in distinguishing normal from adversarial packets (see Fig. 5). Extra application layer payload bytes contribute secondary but meaningful variations, adding fine-grained detail to the classification process. The corresponding feature importance graph reveals a sharp decline after the top protocol features, confirming that these high-level fields dominate the model's decision-making,

TABLE V
MEASUREMENT RESULTS DURING TRAINING AND TESTING.

N	Model	ACC	Precision	Recall	F1 score	Dataset
1	ZENA-Training	0.9580	0.9579	0.9580	0.9579	59,604
2	ZENA-Test	0.9518	0.9518	0.9518	0.9517	7,451

while payload values refine detection accuracy without driving it.

4) *Adversarial threat model*: we consider a resourceful yet constrained adversary with two main objectives: (1) evasion, by crafting malicious Z-Wave packets that bypass the IDS while remaining protocol-compliant, and (2) degradation, by reducing detection accuracy through poisoning or perturbing training and test traffic. We assume a gray-box attacker model, meaning the adversary understands the Z-Wave protocol structure and typical device behavior, but has no access to the IDS's internal parameters or model weights. The white-box scenario, where the attacker fully knows the model architecture and parameters, is excluded since such information is confidential and restricted to the internal IDS implementation, not publicly accessible. The adversary can capture legitimate traffic, inject or replay crafted frames, and modify a limited set of packet fields, such as CMDCL, CMD, VAL, P1, P2, SRC, and DST, while ensuring the packets remain syntactically valid and accepted by target devices.

Attack generation employs targeted field perturbations, replay, injection, and semantic fuzzing guided by Z-Wave specifications. Effectiveness is evaluated through measurable indicators such as evasion rate, false-negative increase, and performance degradation in precision, recall, and F1 score on held-out test sets. The IDS is trained on normal traffic profiles augmented with adversarial samples, without assuming full attacker exclusion from the wireless medium. To promote reproducibility, all adversarial packets used in evaluation are released in a public repository [16].

C. Evaluation Metrics

To evaluate the performance of the ZENA framework, standard classification metrics [5] such as accuracy, precision, recall, and F1-score were used. These metrics are calculated based on values including true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

These metrics provide a comprehensive evaluation of the model's performance, guaranteeing a robust assessment across various dimensions of classification accuracy.

D. Implementation Detail

The ZENA IDS model was developed using Numpy, an open-source Python library for lightweight and resource constraint requirement. The implementation and evaluation of the framework were conducted on a computer with an Intel Core i5-10th Gen CPU (2.9 GHz), 16 GB RAM and a 256 GB SSD using Python 3.11.10. However, since ZENA is a compact model with 8K parameters, training can be done in a much smaller devices. Moreover, the exported model size after training is just 64 KB, making it ideal for integration

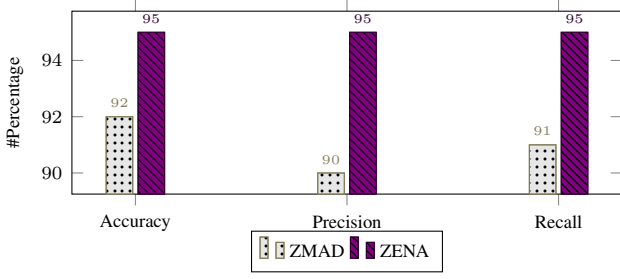


Fig. 6. Metrics visualization between ZMAD and ZENA when trained on same dataset with 23k epochs.

TABLE VI
ACCURACY MEASUREMENTS BETWEEN ZENA (MULTI-HIDDEN LAYERS) AND ZMAD (SINGLE-HIDDEN LAYER) WHEN TRAINED ON SAME ZENA DATASET WITH 23K EPOCHS.

N	Model	Layers	Accuracy	P	R	F1 score
1	ZENA	Multi	0.95	0.95	0.95	0.95
2	ZMAD	Single	0.92	0.90	0.91	0.90

during production phases on constrained Z-Wave controllers, such as the Raspberry Pi.

VI. RESULTS

In this section, we evaluate ZENA performance and compare it to existing work (*i.e.*, ZMAD) and state-of-the-art deep learning neural network models.

A. Performance by Different Layer Size

In this section, we compare the performance of models with various sizes. The experimental results are presented in Table IV. The models were trained with the momentum of 0.9 and early stopping, where the loss threshold was set to 1×10^{-4} and the early stopping rounds were 5k. This means that training was halted if the loss did not decrease by at least 1×10^{-4} over 5k epochs. The learning rate was fixed at 0.03 for all experiments, and the Xavier initialization method was employed. As observed from the results, model size does not necessarily correlate with performance. Notably, the model with a size of 128×32 demonstrated an optimal balance between convergence speed and performance. We conducted additional evaluation in Appendix A (see Table X), a grid search over architectures (1 – 4 layers with varying neuron sizes) confirmed that the two-layer 128×32 model provides the optimal trade-off between accuracy (0.9567), F1 score (0.9565), and computational efficiency, achieving fast convergence with moderate parameters.

B. Accuracy of ZENA

In this section, we present the performance of ZENA, a double-layered classifier with 128×32 parameters for hidden layers. The enhanced model achieved a testing accuracy of 95%, with significant improvements across the new dataset in precision and recall compared to the other models. Table V shows the accuracy measurement results.

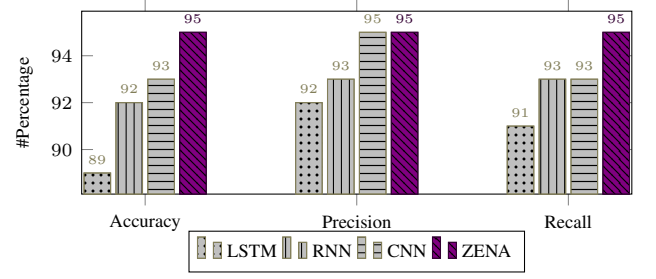


Fig. 7. Metrics visualization between ZENA and state-of-the-art Deep Learning models when trained on same dataset.

TABLE VII
ACCURACY MEASUREMENTS BETWEEN ZENA AND STATE-OF-THE-ART DEEP LEARNING MODELS WHEN TRAINED ON SAME ZENA DATASET.

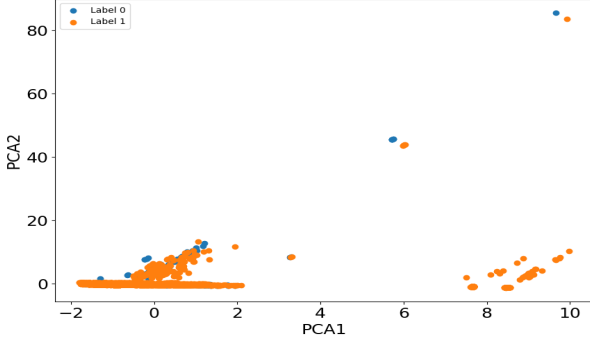
N	Model	Layers	Accuracy	P	R	F1 score
1	ZENA	Multi	0.95	0.95	0.95	0.95
2	CNN	Multi	0.93	0.95	0.93	0.94
3	RNN	Multi	0.92	0.93	0.93	0.93
4	DBN	Multi	0.90	0.93	0.90	0.92
5	LSTM	Multi	0.89	0.92	0.91	0.91

C. Performance Comparison of ZENA with ZMAD

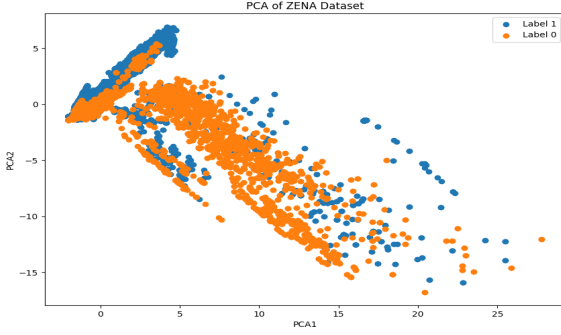
To evaluate the effectiveness of ZENA's architecture, we conducted a direct performance comparison against ZMAD using the same ZENA dataset and identical training conditions (23,000 epochs). As shown in Table VI, ZENA, which employs two hidden layers, significantly outperforms ZMAD's single-hidden-layer architecture. ZENA achieved an accuracy of 95%, precision of 95%, recall of 95%, and an F1 score of 95%. In contrast, ZMAD reached an accuracy of 92%, with a precision of 90%, recall of 91%, and F1 score of 90%. These results highlight ZENA's superior ability to capture complex data patterns and generalize across diverse attack vectors, making it a more robust and reliable model for intrusion detection in Z-Wave smart home environments. Moreover, in extra experiment in Table X, ZENA achieved higher results.

D. Performance Comparison of ZENA with Deep Learning Neural Networks

Since ZENA is an ANN, we compared it against other well-known deep learning neural network architectures to provide a fair benchmark. CNNs are primarily designed for spatially correlated data such as images, where local feature relationships are important. In contrast, Z-Wave packet features are largely independent and do not share such spatial context. Similarly, recurrent neural network (RNNs) and LSTMs are effective for sequential or temporal data (*e.g.*, speech or time-series forecasting), whereas Z-Wave packets represent discrete protocol fields without temporal continuity. Therefore, a multilayer ANN, built-from scratch, provides a more appropriate and efficient mapping between Z-Wave packet fields and anomaly labels. ZENA extends this principle by introducing a custom, protocol-aware ANN framework specifically tailored to the hierarchical structure of Z-Wave traffic. By learning inter-field dependencies, particularly within critical fields such as

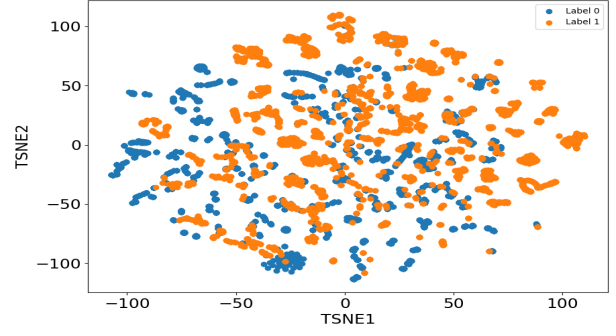


(a) PCA of ZMAD dataset

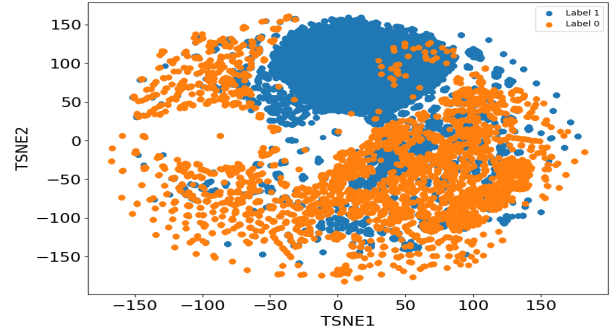


(b) PCA of ZENA dataset

Fig. 8. Visualization of datasets features through PCA.



(a) t-SNE of ZMAD dataset



(b) t-SNE of ZENA dataset

Fig. 9. Visualization of datasets feature through t-SNE.

CMDCL, CMD, and PARAMS ZENA achieves superior detection accuracy and robustness compared to generic deep learning models.

When evaluated on the same dataset, the ZENA artificial neural network model, purpose-built for Z-Wave protocol intricacies, achieved an accuracy of 95% (see Table VII). This surpasses all other deep learning architectures tested, including CNN (93%), RNN (92%), Deep Belief Network (DBN 90%), and LSTM (89%).

These results highlight ZENA's superior ability to detect complex deviations, especially in the presence of adversarial machine learning attacks, where malicious packets are crafted to mimic normal ones. Standard models like DBNs and LSTMs, although state-of-the-art, struggled to maintain accuracy under such conditions. ZENA's architecture, built from scratch and specifically optimized for the Z-Wave protocol, demonstrates its robustness and reliability in securing smart home communication. Fig. 7 highlights the models comparison.

In addition to the deep learning baselines originally presented, we conducted a supplementary evaluation using machine learning algorithms to provide a more comprehensive comparison. This evaluation is available in Appendix B.

E. ZMAD vs ZENA Datasets Analysis

In this section, we present a visual analysis of Z-Wave data distributions from the dataset used in the ZMAD paper [5] and our newly constructed ZENA dataset, utilizing techniques

such as PCA, t-SNE, and KDE plots. This visualization aims to highlight the increased complexity and richness of the ZENA dataset compared to the earlier ZMAD dataset.

Principal component analysis (PCA) is a linear dimensionality reduction technique that transforms data into a new coordinate system, highlighting variance. In contrast, t-distributed stochastic neighbor embedding (t-SNE) is a non-linear technique that excels at preserving local structure in high-dimensional data.

Figs. 8 and 9 illustrate the visualized feature map of ZMAD and ZENA datasets. The visualizations display data projected onto a 2D plane, although the actual data spans 30 dimensions. This dimensionality reduction aids in illustrating the inherent structure and distribution differences between datasets. ZENA demonstrates the capability to efficiently predict complex high-dimensional data with high accuracy.

Finally, Fig. 4 illustrates the normalized distribution of each feature using a KDE plot. ZMAD demonstrates a wider data distribution, whereas ZENA is engineered to capture the contrasting characteristics of normal and aberrant signals, leading to a more prominent peak. This sharper shape indicates that ZENA effectively highlights differences between these signal types.

VII. ABLATION STUDY

A. Rationale for Initialization Ablation Study

Weight initialization plays a critical role in training deep neural networks, as it directly impacts convergence speed,

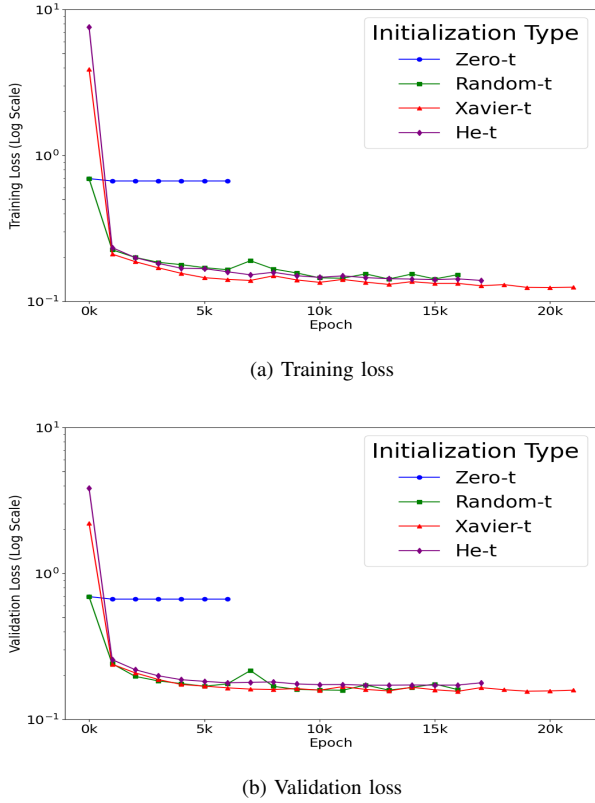


Fig. 10. ZENA training and validation loss from different initialization methods. The loss values are represented in a log scale.

stability, and final performance. Poor initialization can lead to vanishing or exploding gradients, especially in deeper architectures. To evaluate the impact of different initialization strategies on the ZENA model's performance, we conducted an ablation study using the 128×32 architecture, which is the configuration previously identified as optimal (see Table IV). Additional model performance for all ZENA architectures is listed in Appendix A (see Table X). This study helps determine the most effective initialization method to ensure both rapid convergence and minimal training/validation loss, especially under high-dimensional input conditions as encountered in Z-Wave intrusion detection.

B. Overview of Initialization Techniques

We tested four initialization methods: Zero, Random, Xavier [43], and He [44]. The Zero method initializes all weights to zero, which often fails due to lack of symmetry breaking during learning. Random uses small random values drawn from a uniform distribution, providing basic stochasticity but without theoretical backing for deep architectures. Xavier (Glorot initialization) is designed to maintain variance through layers by scaling weights based on the number of input and output neurons, making it ideal for networks with sigmoid or tanh activations. He initialization is optimized for ReLU-based networks and scales weights more aggressively to preserve activation variance in deep layers.

TABLE VIII

ZENA- α SINGLE-LAYER MODEL PERFORMANCE AND GENERALIZATION MEASUREMENT ON ADVERSARIAL DATA. SUBSET-1 DATASET CONTAINS 10,496 SAMPLES ORIGINALLY DESIGNED TO TRAIN ZMAD MODEL [5], AND SUBSET-2 DATASET CONTAINS ADDED ADVERSARIAL ADVANCED SAMPLES (TOTAL OF 64,009) THAT POSE A HARDER CHALLENGE FOR PREDICTION. TO TEST GENERALIZATION, THE MODEL IS TRAINED ON ONE DATASET AND EVALUATED ON BOTH DATASET.

Test#	Training data	ZENA- α single-layer model evaluation phase	
		F1 score on Subset-1 dataset	F1 score on Subset-2 dataset
1	Subset 1 dataset	0.9628	0.6890
2	Subset 2 dataset	0.6512	0.9345

TABLE IX

ZENA- β MULTI-LAYERS MODEL PERFORMANCE AND GENERALIZATION MEASUREMENT ON UNSEEN DATA. SUBSET-1 DATASET CONTAINS 10,496 SAMPLES ORIGINALLY DESIGNED TO TRAIN ZMAD MODEL [5], AND SUBSET-2 DATASET CONTAINS NEWLY ADDED ADVERSARIAL ADVANCED SAMPLES (TOTAL OF 64,009) THAT POSE A HARDER CHALLENGE FOR PREDICTION. TO TEST GENERALIZATION, ZENA IS TRAINED ON ONE DATASET AND EVALUATED ON BOTH DATASET.

Test#	Training data	ZENA- β multi-layers model evaluation phase	
		F1 score on Subset-1 dataset	F1 score on Subset-2 dataset
1	Subset 1 dataset	0.9571	0.7147
2	Subset 2 dataset	0.7229	0.9292

C. Performance by Different Initialization Method

As shown in Fig. 10, Zero initialization led to immediate training stagnation, with both training and validation loss remaining flat throughout the epochs, confirming its unsuitability for deep learning. Random and He initializations showed relatively fast convergence in the early stages of training; however, both exhibited slightly higher final loss values compared to Xavier. While He initially dropped rapidly, its validation loss plateaued above that of Xavier. In contrast, Xavier demonstrated a smoother and more consistent decline in both training and validation loss, ultimately achieving the lowest final values among all methods. These results indicate that while He and random initializations offer quicker early improvements, Xavier provides the most reliable and stable performance for the ZENA model.

D. Cross-Data Validation

In this experiment, we compare two version of ZENA: ZENA- α , a single-layer ANN, and ZENA- β , a multi-layers, by testing their robustness on unseen data. First, we create two subsets: *Subset-1 Dataset*, which contains 10,496 samples that were originally designed to train ZMAD model, and *Subset-2 Dataset*, which contains 64,009 adversarial samples that pose a harder challenge for prediction. To test generalization, models were trained on one dataset and evaluated on both datasets. The reported F1 scores are calculated using predictions on a 10% test split from the training dataset and all samples from the other dataset. These scores highlight the trained performance and generalization performance on unseen samples.

Table VIII summarizes the performance of ZENA- α single-layer model, which achieves a high F1 score of 0.9628 when trained and tested on the **Subset-1** dataset, demonstrating strong performance on familiar data. However, its ability to generalize is limited, as shown by its lower F1 score of 0.6890

when tested on the **Subset-2** dataset. Similarly, when trained on **Subset-2**, it performs well with an F1 score of 0.9345 on similar data but struggles with a lower score of 0.6512 on the **Subset-1** dataset. These results suggest that the single-layer architecture lacks robustness when handling datasets with varying distributions.

Table IX presents the performance of the ZENA- β multi-layers ANN model. ZENA- β demonstrates greater robustness across different datasets. When trained on **Subset-1**, it achieves an F1 score of 0.9571 and maintains reasonable performance with an F1 score of 0.7147 when tested on **Subset-2**, indicating better adaptability to noisy data distributions. Similarly, when trained on **Subset-2**, it records an F1 score of 0.7229 on **Subset-1** and 0.9292 on its training dataset, showcasing its ability to handle variations in data composition more effectively than its ZENA- α single-layer counterpart.

The results of the experiment indicate that while the single-layer ANN performs well on clean and familiar data, it struggles to generalize across datasets with different distributions. In contrast, the double-layer ANN shows enhanced capacity in acquiring complex patterns while adapting to unknown data, making it a more robust alternative for tasks involving diverse datasets and future unknown new attacks on Z-Wave networks. This setup emphasizes the importance of model architecture in achieving consistent performance across diverse data distributions while effectively managing noisy or challenging scenarios.

VIII. DISCUSSION

A. Advantages of ZENA Enhanced Model for Z-Wave Networks

While ZENA employs a multilayer ANN architecture, its contribution extends far beyond a simple architectural improvement. The key enhancement of ZENA are as follows:

Custom-built artificial neural network from scratch for Z-Wave protocol: unlike conventional ML/DL models such as Random Forest, SVM, CNN, DBN, or LSTM, which depend heavily on large libraries (e.g., scikit-learn, TensorFlow, PyTorch), ZENA was entirely implemented from scratch in Python and NumPy. This lightweight implementation ensures minimal library dependencies, improved interpretability, and high portability for embedded or gateway deployment (e.g., on Raspberry Pi). Moreover, ZENA's network structure was explicitly customized to match the hierarchical and field-oriented structure of Z-Wave packets, rather than adopting generic feature matrices as in existing IDS architectures.

Protocol-aware learning for Z-Wave traffic: Z-Wave packet formats are hierarchical, where critical attack vectors typically exploit specific fields such as CMDCL and their corresponding CMD. ZENA's ANN was designed to scrutinize each field of the Z-Wave frame and learn the normal behavioral profile of inter-field dependencies. This design enables ZENA to recognize subtle deviations within command sequences, lengths, and field relationships; capabilities that are not naturally captured by standard models.

Transparent, non-blackbox IDS design: a significant innovation of ZENA is its interpretable reasoning. Unlike traditional deep learning-based IDS approaches that behave as blackboxes, ZENA, built from scratch, allows the IDS operator to trace the decision rationale behind each classification (normal or abnormal). This transparency was achieved by fitting ZENA internal architecture to the Z-Wave packet structure. Therefore, this design increase trust and auditability in operational environments.

Adversarial robustness and dataset contribution: ZENA was trained with adversarially perturbed Z-Wave packets, improving robustness against evasion attempts. The accompanying ZENA dataset, the large collection of 74,505 realistic Z-Wave packets (including adversarial samples) further enhances reproducibility and establishes a new benchmark for IoT IDS research. By making this dataset publicly available [16], this paper also contributes to the advancement of research in secure IoT communication.

Together, these aspects demonstrate that ZENA's novelty lies in the combination of a protocol-aware, interpretable, and fully custom ANN framework, designed specifically for Z-Wave environments, rather than merely improving upon existing neural network depth or size.

B. ZENA Strengths

The proposed ZENA model demonstrates substantial improvements in both detection accuracy and generalization when compared to prior work such as ZMAD. By employing a multilayer artificial neural network, ZENA successfully addresses key limitations inherent in single-layer architectures. The deeper structure enables the model to capture non-linear and high-order feature interactions, which are particularly important for distinguishing between subtle differences in legitimate versus malicious Z-Wave traffic. The inclusion of adversarial training further enhances the model's robustness by exposing it to attack patterns that are intentionally crafted to evade detection.

Another major strength of ZENA lies in its scalability and flexibility. The architecture was designed to accommodate high-dimensional input features extracted from Z-Wave protocol traffic. Through ablation studies, it was shown that proper initialization techniques such as Xavier initialization significantly improve convergence speed and loss minimization. These architectural and training optimizations make ZENA not only more effective, but also more stable across different training conditions.

C. Limitations and Challenges

Despite its advantages, ZENA presents some limitations that merit discussion. First, the multilayer ANN architecture introduces higher computational and memory overhead than its single-layer predecessor (ZMAD), increasing training cost and limiting feasibility for fully on-device deployment on highly resource-constrained smart home slave devices. As a result, ZENA is currently better suited for centralized or edge-based IDS deployments, such as smart home hubs, controllers, or

gateways (e.g., Raspberry Pi), rather than execution directly on constrained end devices.

Second, although adversarial training improves robustness, the evolving nature of adversarial attacks remains a challenge. The fuzzing-based adversarial samples used in ZENA provide broad coverage of the input space; however, future attackers may exploit previously unseen vulnerabilities or model-specific weaknesses not represented in the training set. Misclassifications are particularly likely for adversarial packets that preserve protocol-level validity while introducing only subtle semantic deviations within highly variable application-layer payload fields (e.g., `PARAM` or `VAL` bytes). When perturbations do not significantly affect dominant protocol fields such as `CMDCL`, `CMD`, or frame control values, such packets may closely resemble legitimate traffic and evade detection.

Third, ZENA's generalization may degrade when deployed in environments involving device types, firmware behaviors, or Z-Wave command distributions not observed during training, especially for rare or newly introduced command classes. Moreover, public release of the training dataset, while essential for reproducibility, increases the feasibility of white-box evasion by enabling adversaries to analyze feature correlations and training distributions. These limitations highlight the need for defense-in-depth deployment strategies, including controlled model distribution, complementary signature-based detection, periodic adversarial audits, continual dataset expansion, cross-environment validation, and adaptive retraining to sustain IDS effectiveness in evolving Z-Wave ecosystems.

D. Future Work

To extend the capabilities of ZENA, several promising research directions can be pursued:

Lightweight architectures for edge deployment. One major avenue is to optimize ZENA for deployment on constrained hardware through model compression techniques such as pruning, quantization, or knowledge distillation. This would enable real-time intrusion detection directly on smart home devices or low-power gateways.

Online and continual learning. Smart home environments evolve, and so do attack strategies. Implementing online learning mechanisms would allow ZENA to adapt to new patterns without requiring complete retraining, maintaining its relevance and accuracy over time.

Cross-protocol generalization. While ZENA is designed for Z-Wave networks, its architectural principles could be extended to other IoT communication protocols such as ZigBee, Bluetooth Low Energy (BLE), or Thread. Exploring transfer learning techniques may enable a unified IDS framework for heterogeneous IoT ecosystems.

Explainability and visualization. Improving the interpretability of ZENA's decisions is critical for real-world adoption. Future work can focus on integrating explainable AI (XAI) techniques such as SHAP to highlight which features influenced a given detection outcome.

In summary, ZENA represents a significant advancement in anomaly-based intrusion detection for smart home networks using Z-Wave protocol. Its architectural enhancements, robust

training strategy, and strong performance provide a solid foundation for future innovations in AI-driven IoT security.

IX. CONCLUSION

This paper introduced ZENA, a multilayer ANN model specifically designed to detect anomalies in Z-Wave smart home networks. Addressing the limitations of the earlier ZMAD model, ZENA incorporates a deeper network architecture, adversarial training, and optimized initialization strategies to enhance both accuracy and resilience. The model was trained and evaluated on a comprehensive dataset of over 74,505 Z-Wave packets, including both clean and adversarial samples, making it well-suited for real-world deployment scenarios.

Through extensive experimentation, ZENA achieved a 95% accuracy and outperformed the single-layer ZMAD in precision, recall, and generalization. Ablation studies further validated the effectiveness of Xavier initialization and highlighted the impact of proper weight initialization on training stability and performance.

By combining structural improvements with robust training strategies, ZENA provides a practical and scalable solution for securing Z-Wave-based smart home ecosystems against a wide range of attacks. The publicly available dataset and reproducible design offer a solid foundation for future research and real-world adoption. Moving forward, we aim to optimize the model for edge devices and extend our approach to other IoT communication protocols, contributing to the broader goal of securing next-generation smart environments.

X. AVAILABILITY OF DATA AND MATERIALS

Datasets are publicly available at [16].

ACKNOWLEDGMENT

We are grateful to all anonymous reviewers for their valuable feedback that helped improve this paper.

APPENDIX A

MODEL PERFORMANCE FOR ALL ZENA ARCHITECTURES

We already reported grid-search model performance within the 2-layer architecture in Table IV. As shown in Table X, we conducted an extensive grid search over multiple architectures, varying both the number of layers (1–4) and neuron sizes per layer. The results demonstrate that the two-layer configuration with 128×32 neurons achieves the best trade-off between performance and computational efficiency, yielding high accuracy (0.9567) and F1 score (0.9565) while maintaining moderate parameter count and convergence speed. This confirms that deeper or wider models (e.g., 3- or 4-layer setups) provide only marginal gains at significantly higher computational cost. Thus, the chosen 128×32 architecture represents the optimal balance between model complexity, stability, and training efficiency for Z-Wave traffic analysis.

TABLE X
MODEL PERFORMANCE FOR ALL ZENA ARCHITECTURES.

#Params	Layer size	Params l1	Params l2	L/R	Epochs	Accuracy	F1 score	Momentum
3401	[64]	64	—	0.03	22.3k	0.9526	0.9525	0.9
3961	[128]	128	—	0.03	7.9k	0.9514	0.9513	0.9
7729	[256]	256	—	0.03	12.4k	0.9526	0.9524	0.9
2011	[32]	32	—	0.03	17.5k	0.9436	0.9435	0.9
961	[16]	16	—	0.03	11.2k	0.9324	0.9322	0.9
4481	[64,16]	64	16	0.03	18.3k	0.9568	0.9566	0.9
8801	[64,32]	64	32	0.03	20.3k	0.9545	0.9543	0.9
16833	[64,64]	64	64	0.03	17.1k	0.9576	0.9574	0.9
4225	[128,16]	128	16	0.03	20.1k	0.9552	0.9549	0.9
8129	[128,32]	128	32	0.03	35.2k	0.9567	0.9565	0.9
16641	[128,64]	128	64	0.03	15.6k	0.9564	0.9562	0.9
32961	[128,128]	128	128	0.03	6.5k	0.9540	0.9538	0.9
8321	[256,16]	256	16	0.03	7.6k	0.9562	0.9561	0.9
16513	[256,32]	256	32	0.03	7.4k	0.956	0.9558	0.9
33025	[256,64]	256	64	0.03	12.7k	0.9558	0.9557	0.9
65793	[256,128]	256	128	0.03	9.7k	0.9581	0.9580	0.9
6466	[32,32,32,32]	32	32	0.03	8.9k	0.9562	0.9561	0.9
11617	[64,64,32]	64	64	0.03	15.4k	0.9564	0.9562	0.9
8321	[128,32,16]	128	32	0.03	12.3k	0.9576	0.9574	0.9
11841	[128,32,32]	128	32	0.03	10.6k	0.9562	0.9561	0.9
16641	[128,64,32]	128	64	0.03	7.9k	0.9562	0.9560	0.9
24769	[128,64,64]	128	64	0.03	7.6k	0.9557	0.9555	0.9
33153	[128,128,32]	128	128	0.03	8.7k	0.9572	0.9570	0.9
31297	[32,128,128,32]	32	128	0.03	9.1k	0.9571	0.9569	0.9
12481	[64,32,64,32]	64	32	0.03	10.0k	0.9569	0.9568	0.9
16513	[64,64,64,64]	64	64	0.03	7.9k	0.9576	0.9574	0.9
11521	[16,64,64,16]	16	64	0.03	6.2k	0.9544	0.9542	0.9
8321	[16,32,32,16]	16	32	0.03	8.1k	0.9552	0.9550	0.9

APPENDIX B

COMPARISON WITH MACHINE LEARNING BASELINES

In addition to the deep learning baselines originally presented, we conducted a supplementary evaluation using traditional machine learning algorithms to provide a more comprehensive comparison. Specifically, we tested Random Forest, XGBoost, and One-Class SVM using the scikit-learn library, as well as an Autoencoder model to represent unsupervised learning.

The results, show that among traditional models, XGBoost achieved the best performance with an accuracy of 0.9468 and F1 score of 0.9570, followed closely by ZENA (F1 = 0.9565). This confirms that gradient-boosting methods are indeed powerful for structured Z-Wave IoT data. However, ZENA achieves comparable or better F1 scores while being more lightweight and transparent, requiring no heavy frameworks or ensemble computations. Random Forest performed slightly lower (F1 = 0.9392), reflecting the limitations of bagging-based methods on complex inter-field dependencies.

For unsupervised models, One-Class SVM underperformed (F1 = 0.5343), revealing that learning only the *normal* class is insufficient for detecting sophisticated or adversarial Z-Wave attacks. The Autoencoder achieved moderate performance (F1 = 0.9017), benefiting from feature compression but still lacking discriminative power against diverse attack types.

Overall, this expanded evaluation confirms that ZENA offers near state-of-the-art accuracy comparable to XGBoost, but with greater interpretability, lower computational cost, and better adaptability to embedded or gateway devices. These findings reinforce ZENA's practical advantage as a lightweight, protocol-aware IDS specifically optimized for Z-Wave smart-home networks.

APPENDIX C

MODEL INTERPRETABILITY USING SHAP

To enhance transparency and explainability of the ZENA intrusion detection model, we applied SHapley Additive Explanations (SHAP) analysis. This analysis provides insight into which Z-Wave packet field attributes most influence the model's attack predictions.

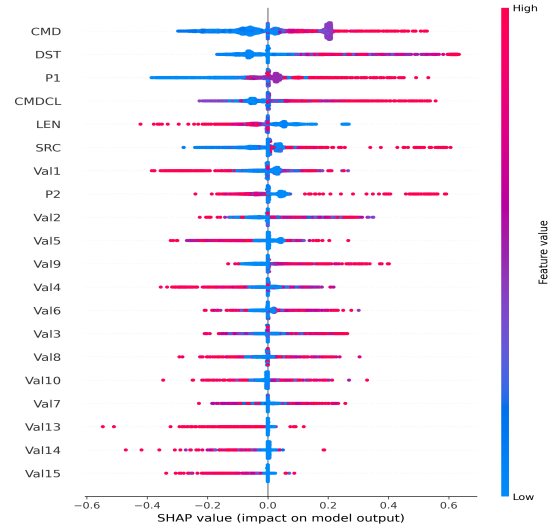


Fig. 11. SHAP feature importance ranking.

SHAP summary plot in Fig. 11 indicates that CMD, DST, P1, CMDCL, LEN, and SRC are the most influential features in distinguishing normal and attack traffic. Several payload-related fields (Val1, P2, and Val2) also exhibit strong contributions, demonstrating that the model learns protocol-level behavior rather than random noise.

The figure further illustrates that higher values of these key fields (in red) tend to push predictions toward the attack class, while lower values (in blue) align with normal packets. This confirms that ZENA's learning signal is semantically grounded in legitimate Z-Wave packet characteristics.

REFERENCES

- [1] Wikipedia. Z-Wave. [Online]. Available: <https://en.wikipedia.org/wiki/Z-Wave>
- [2] BNP Media. Sigma Designs Ships 100 Millionth Z-Wave Chipset Module. [Online]. Available: <https://www.sdmag.com/articles/95168-sigma-designs-ships-100-millionth-z-wave-chipset-module>
- [3] C. K. Nkuba, S. Kim, S. Dietrich, and H. Lee, "Riding the IoT wave with VFuzz: Discovering security flaws in smart homes," *IEEE Access*, vol. 10, pp. 1775–1789, 2022.
- [4] J. Hall, B. Ramsey, M. Rice, and T. Lacey. EZ-Wave. [Online]. Available: <https://github.com/cureHsu/EZ-Wave>
- [5] C. K. Nkuba, S. Woo, H. Lee, and S. Dietrich, "ZMAD: Lightweight model-based anomaly detection for the structured Z-Wave protocol," *IEEE Access*, vol. 11, pp. 60562–60577, 2023.
- [6] I. Aiyanyo *et al.*, "A systematic review of blockchain application in smart cities: Trends and opportunities," *Advances in Appl. Intell. Research*, vol. 2, no. 1, 2025.
- [7] CERT Coordination Center. Silicon Labs Z-Wave chipsets contain multiple vulnerabilities. [Online]. Available: <https://kb.cert.org/vuls/id/142629>
- [8] N. K. Carlos, J. Kang, S. Woo, and H. Lee, "ZCover: Uncovering Z-Wave controller vulnerabilities through systematic security analysis of application layer implementation," in *Proc. DSN*, 2025.
- [9] I. Aiyanyo, H. Samuel, and H. Lim, "A systematic review of defensive and offensive cybersecurity with machine learning," *Appl. Sci.*, vol. 10, no. 17, p. 5811, 2020.
- [10] V. Pantelakis, P. Bountakas, A. Farao, and C. Xenakis, "Adversarial machine learning attacks on multiclass classification of IoT network traffic," in *Proc. ARES*, 2023.
- [11] O. A. Alzubi, J. A. Alzubi, I. Qiqieh, and A. Al-Zoubi, "An IoT intrusion detection approach based on salp swarm and artificial neural network," *International J. Netw. Manag.*, vol. 35, no. 1, p. e2296, 2025.

- [12] J. A. Alzubi, O. A. Alzubi, I. Qiqieh, and A. Singh, "A blended deep learning intrusion detection framework for consumable edge-centric IoMT industry," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 2049–2057, 2024.
- [13] M. Alweshah, A. Hammouri, S. Alkhalaileh, and O. Alzubi, "Intrusion detection for the Internet of things (IoT) based on the emperor penguin colony optimization algorithm," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 5, pp. 6349–6366, 2023.
- [14] O. A. Alzubi *et al.*, "Optimized machine learning-based intrusion detection system for fog and edge computing environment," *Electronics*, vol. 11, no. 19, p. 3007, 2022.
- [15] O. A. Alzubi, "A deep learning-based frechet and dirichlet model for intrusion detection in IWSN," *J. Intell. Fuzzy Syst.*, vol. 42, no. 2, pp. 873–883, 2022.
- [16] Chanhee Lee, Nkuba K. Carlos. ZENA Dataset. [Online]. Available: <https://github.com/CNK2100/ZENA-DATASET>
- [17] Z-Wave Alliance. The smart home is powered by Z-Wave. [Online]. Available: <https://z-wavealliance.org>
- [18] Z-Wave Alliance. Z-Wave ecosystem report. [Online]. Available: <https://z-wavealliance.org/zwave-ecosystem-report/>
- [19] Silicon Laboratories. Z-Wave Specification. [Online]. Available: <https://www.silabs.com/products/wireless/mesh-networking/z-wave/specification>
- [20] P. T. Partners. Z-Shave. Exploiting Z-Wave Downgrade Attacks. [Online]. Available: <https://www.pentestpartners.com/security-blog/z-shave-exploiting-z-wave-downgrade-attacks/>
- [21] Silicon Labs. Next-Generation Z-Wave 700 Launches on the Silicon Labs. Wireless Gecko Platform. [Online]. Available: <https://news.silabs.com/2018-12-12-Next-Generation-Z-Wave-700-Launches-on-the-Silicon-Labs-Wireless-Gecko-Platform>
- [22] Silicon Labs. Introduction to Z-Wave 800 Series. [Online]. Available: <https://www.silabs.com/wireless/z-wave/introduction-to-z-wave-800-series>
- [23] I. T. Union. (January) Short Range Narrow-band Digital Radiocommunication Transceivers – PHY and MAC Layer Specifications. [Online]. Available: <https://www.itu.int/rec/T-REC-G.9959>
- [24] C. Paetz, *Z-Wave Essentials*. CreateSpace Independent Publishing Platform, 2018.
- [25] B. Fouladi and S. Ghanoun, "Security evaluation of the Z-Wave wireless protocol," *BlackHat USA*, vol. 24, pp. 1–2, 2013.
- [26] Z-Wave Alliance. Z-Wave products. [Online]. Available: <https://products.z-wavealliance.org/>
- [27] J. Hall and B. Ramsey, "Breaking bulbs briskly by bogus broadcasts," in *ShmooCon*, 2016.
- [28] J. D. Fuller and B. W. Ramsey, "Rogue Z-Wave controllers: A persistent attack channel," in *Proc. LCN Wksp.*, 2015.
- [29] J. D. Fuller, B. W. Ramsey, M. J. Rice, and J. M. Pecarina, "Misuse-based detection of Z-Wave network attacks," *Comput. Security*, vol. 64, pp. 44–58, 2017.
- [30] C. Badenhop, S. R. Graham, B. W. Ramsey, B. E. Mullins, and L. O. Mailloux, "The Z-Wave routing protocol and its security implications," *Comput. Security*, vol. 68, pp. 112–129, 2017.
- [31] C. Badenhop, S. R. Graham, B. E. Mullins, and L. O. Mailloux, "Looking under the hood of Z-Wave: Volatile memory introspection for the ZW0301 transceiver," *ACM Trans. Cyber-Physical Syst.*, vol. 3, no. 2, p. 20, 2018.
- [32] K. Kim *et al.*, "What's your protocol: Vulnerabilities and security threats related to Z-Wave protocol," *Pervasive Mobile Comput.*, vol. 66, p. 101211, 2020.
- [33] N. Boucif, F. Golchert, A. Siemer, P. Felke, and F. Gosewehr, "Crushing the Wave: New Z-Wave vulnerabilities exposed," *arXiv preprint arXiv:2001.08497*, 2020.
- [34] J.-C. Liou, S. Jain, S. R. Singh, D. Taksinwarajan, and S. Seneviratne, "Side-channel information leaks of Z-Wave smart home IoT devices: Demo abstract," in *Proc. SenSys*, 2020.
- [35] J. D. Fuller, "A misuse-based intrusion detection system for ITU-T G. 9959 wireless networks," Air Force Institute of Technology Wright-Patterson AFB OH Wright-Patterson, Tech. Rep., 2016.
- [36] N. Carlini and D. Wagner, "Towards evaluating the robustness of neural networks," in *Proc. IEEE SP*, 2017.
- [37] A. Kurakin, I. J. Goodfellow, and S. Bengio, "Adversarial examples in the physical world," in *Artificial Intell. Safety Security*, 2018.
- [38] Silicon Laboratories. Z-Wave Zniffer User Guide. [Online]. Available: <https://www.silabs.com/documents/public/user-guides/INS10249-Z-Wave-Zniffer-User-Guide.pdf>
- [39] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE ICCV*, 2015.
- [40] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. AISTATS*, 2010.
- [41] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, "On the importance of initialization and momentum in deep learning," in *Proc. ICML*, 2013.
- [42] G. Nakerst, J. Brennan, and M. Haque, "Gradient descent with momentum—to accelerate or to super-accelerate?" *arXiv preprint arXiv:2001.06472*, 2020.
- [43] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. AISTATS*, 2011.
- [44] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE ICCV*, 2015.



Nkuba Kayembe Carlos (Member, IEEE) received his M.S. degree in Computer Science Information Technology from Handong Global University and his Ph.D. degree in Computer Science and Engineering from Korea University. He is currently a Research Professor at the Center for Software Security & Assurance, Korea University. His research interests include computer software & network security, fuzzing, secure internet architecture, and IoT automated vulnerability discovery.



Chanhee Park received his B.S. degree from the Department of Korean Language and Literature of Korea University, Seoul, South Korea, in 2020. He is a Ph.D. candidate in Computer Science and Engineering at Korea University, Seoul, South Korea. Currently, he is part of the Natural Language Processing & Artificial Intelligence Lab team under an integrated master's and Ph.D. course. His research focuses on how to implement AI technologies to improve the everyday use of people in a more human-centric way.



Imatitikua Aiyanyo received her M.Sc. and Ph.D. in Computer Science from Korea University. She is currently a Research Professor at Human Inspired AI lab, Korea University. Her research interests include cybersecurity, cyber-defense strategies and machine learning application in different fields.



Editor of the Journal IEEE Transactions on Vehicular Technology.

Heejo Lee (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in Computer Science and Engineering from POSTECH, South Korea. He is currently a Professor with the Department of Computer Science and Engineering at Korea University, and the Director of the Center for Software Security and Assurance (CSSA). He is a Founding Member and the Co-CEO of Labrador Labs. Before joining Korea University, he was the CTO with AhnLab Inc., from 2001 to 2003, and a Postdoctoral Researcher with Purdue University, from 2000 to 2001. He is the



Heuseok Lim received the B.S., M.S., and Ph.D. degrees in Computer Science and Engineering from Korea University, Seoul, South Korea, in 1992, 1994, and 1997, respectively. He is currently a Professor in the Department of Computer Science and Engineering at Korea University. His research interests include natural language processing, machine learning, and artificial intelligence.