IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 31 May 2024, accepted 30 June 2024, date of publication 4 July 2024, date of current version 2 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3422999

RESEARCH ARTICLE

Optimizing Smart Home Intrusion Detection With Harmony-Enhanced Extra Trees

AKMALBEK ABDUSALOMOV[1], DUSMUROD KILICHEV[1], RASHID NASIMOV[2], ILKHOM RAKHMATULLAYEV[3], AND YOUNG IM CHO[1]

[1]Department of Computer Engineering, Gachon University, Seongnam 13120, South Korea
[2]Department of Information Systems and Technologies, Tashkent State University of Economics, Tashkent 100066, Uzbekistan
[3]Department of Information Security, Tashkent University of Information Technologies, Samarkand 140100, Uzbekistan

Corresponding author: Young Im Cho (yicho@gachon.ac.kr)

This work was supported in part by Korea Agency for Technology and Standards, in 2022, through the Project "Development of International Standard Technologies Based on AI Model Lightweighting Technologies" under Grant and through the Project "Development of International Standard Technologies Based on AI Learning and Inference Technologies" under Grant 1415180835; and in part by Korea Institute of Marine Science and Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries under Grant RS-2022-KS221571.

ABSTRACT In this study, we present an innovative network intrusion detection system (IDS) tailored for Internet of Things (IoT)-based smart home environments, offering a novel deployment scheme that addresses the full spectrum of network security challenges. Distinct from existing approaches, our comprehensive strategy not only proposes a model but also incorporates IoT devices as potential vectors in the cyber threat landscape, a consideration often neglected in previous research. Utilizing the harmony search algorithm (HSA), we refined the extra trees classifier (ETC) by optimizing an extensive array of hyperparameters, achieving a level of sophistication and performance enhancement that surpasses typical methodologies. Our model was rigorously evaluated using a robust real-time dataset, uniquely gathered from 105 IoT devices, reflecting a more authentic and complex network scenario compared to the simulated or limited datasets prevalent in the literature. Our commitment to collaborative progress in cybersecurity is demonstrated through the public release of our source code. The system underwent exhaustive testing in 2-class, 8-class, and 34-class configurations, showcasing superior accuracy (99.87%, 99.51%, 99.49%), precision (97.41%, 96.02%, 96.07%), recall (98.45%, 87.14%, 87.1%), and f1-scores (97.92%, 90.65%, 90.61%) that firmly establish its efficacy. This work marks a significant advancement in smart home security, providing a scalable and effective network IDS solution that is adaptable to the intricate dynamics of modern IoT networks. Our findings pave the way for future endeavors in the realm of cyber defense, ensuring that smart homes remain safe havens in an era of digital vulnerability.

INDEX TERMS Extra trees classifier, harmony search algorithm, hyperparameter optimization, Internet of Things, intrusion detection system, machine learning, smart home.

I. INTRODUCTION

As we navigated through 2023, the landscape of cybercrime underwent a dramatic transformation, reaching unprecedented levels of sophistication and impact. The stark projection that the global cost of cybercrime could soar to an astonishing $23.84 trillion by 2027, up from $8.44 trillion in 2022, serves as a sobering reminder of the escalating challenge at hand [1]. This alarming trend underscores the critical need for enhanced cybersecurity measures, particularly in environments as integral to our daily lives as our homes [2].

With the advent of smart home technology, the very concept of home security has evolved [3]. No longer confined to physical locks and alarms, the modern home's defense system must now contend with cyber threats – invisible,

The associate editor coordinating the review of this manuscript and approving it for publication was Vicente Alarcon-Aquino.

intangible, and far more insidious [4]. The adversaries in this new era are not the traditional burglars of old but rather faceless entities operating from the shadows of cyberspace, exploiting the interconnected fabric of our digital existence.

At the core of safeguarding the digital integrity of smart homes lies the intrusion detection system (IDS) [5], a critical component designed to monitor network or system activities for malicious actions or policy violations [6]. An effective IDS serves as the digital immune system of a smart home [7], constantly scanning for anomalies that signify potential intrusions [8]. Its primary function is to identify both known and novel threats, ensuring that any unauthorized or suspicious activity is flagged and addressed promptly [9]. Traditional IDS have evolved from simple signature-based detectors, which rely on known patterns of malicious activity, to more sophisticated systems that employ advanced strategies, including anomaly detection and heuristics [10]. Despite their advancements, traditional IDS often struggle to keep pace with the complexity and novelty of modern cyber threats, particularly in the interconnected ecosystem of smart homes. Despite these advancements, traditional IDS often struggle to keep pace with the complexity and novelty of modern cyber threats, especially within the highly interconnected and dynamic ecosystem of smart homes. Signature-based IDS are inherently limited as they can only detect attacks with pre-defined signatures, leaving them vulnerable to zero-day exploits and previously unknown threats. Anomaly detection systems, while more flexible, often suffer from high false positive rates and require extensive manual tuning to remain effective. Moreover, the static nature of rule-based approaches in traditional IDS hampers their ability to adapt to the rapidly evolving threat landscape, necessitating frequent updates and maintenance which can be both time-consuming and prone to human error. This challenge underscores the necessity for innovative approaches that enhance the adaptability and intelligence of IDS, making them more adept at recognizing and neutralizing threats in a dynamic digital environment [11]. It is here that machine learning (ML) methods offer a beacon of hope [12], promising to revolutionize the efficacy and adaptability of IDS through intelligent, data-driven insights [13].

In the dynamic landscape of cybersecurity, ML methods have emerged as a transformative force [14]. These methods, harnessing the power of algorithms and data-driven insights, offer unparalleled capabilities in detecting and adapting to evolving cyber threats. At the heart of ML's efficacy is its ability to learn from data, discern patterns, and make decisions with minimal human intervention [15]. This ability is particularly crucial in the context of smart home security, where the diversity and volume of data can be overwhelming for traditional rule-based systems.

However, the true potential of ML in enhancing IDSs is unlocked through optimization [16]. Optimization in ML involves fine-tuning algorithms to maximize their performance, ensuring that they not only accurately identify genuine threats but also minimize false positives, which are a significant challenge in IDS [17]. This process involves sophisticated techniques like hyperparameter tuning, model selection, and feature engineering [18]. Among these, hyperparameter tuning, which involves adjusting the parameters that govern the learning process of the model, is particularly pivotal [19]. It's akin to fine-tuning an instrument to ensure it plays the perfect note – critical in the harmony of a smart home's defense strategy.

Our research capitalizes on this aspect of ML, employing the harmony search algorithm (HSA) – a global optimization technique inspired by the improvisation process of musicians – to optimize the extra trees classifier (ETC) [20], a powerful ensemble learning technique. This innovative approach not only enhances the accuracy and efficiency of intrusion detection but also embodies the adaptability required to keep pace with the rapidly evolving cyber threat landscape. Through this research, we aim to set a new benchmark in smart home security, showcasing the synergy between advanced ML methods and strategic optimization techniques.

Our main contributions to the field of Internet of Things (IoT)-based smart home security through network IDS deployment are summarized as follows:

- We propose a novel network IDS deployment scheme specifically designed for the IoT smart home ecosystem, a critical aspect often overlooked in existing literature.
- Through the use of the HSA, we optimize a comprehensive set of hyperparameters for the ETC, enhancing the model's performance.
- Our research is underpinned by a unique and extensive real-time dataset collected from 105 home IoT devices, incorporating scenarios where IoT devices are used as attack vectors.
- In an effort to foster transparency and collaborative advancement, we make our code available in a public GitHub repository,[1] encouraging further research and development in the community.
- We conduct a thorough evaluation of our models across 2-class, 8-class, and 34-class configurations, using a comprehensive suite of metrics including Accuracy, Precision, Recall, and F1-score, ensuring a robust and detailed assessment of model performance.

These contributions collectively mark a significant advancement in IoT-based smart home security, setting new standards for both the development and deployment of network IDS in a rapidly evolving technological landscape.

The structure of this paper is meticulously crafted as follows: Section II provides a critical comparison of our groundbreaking contributions against the backdrop of the existing body of work, setting the stage for the relevance of our research. Section III delineates our methodological framework, detailing the implementation of our proposed network IDS for smart home security, the intricacies of the ETC, the optimization prowess of the HSA, and the synthesis into our novel Harmony-Enhanced Extra Trees (HEET)

[1] https://github.com/TATU-hacker/Harmony-Enhanced-Extra-Trees.git

model. This section also describes our stringent data preprocessing protocols and the evaluation metrics employed. Section IV reveals the empirical findings, documenting the outcomes of hyperparameter optimization, and exploring the depth and breadth of our model's capabilities through 2-class, 8-class, and 34-class classification tests. Section V delves into the discourse of our findings, dissecting the nuances of binary and multi-class classification results, juxtaposing our model's performance with contemporaneous works, and critically examining the study's constraints while positing future research trajectories. Conclusively, Section VI coalesces our research journey, encapsulating the paramount implications and affirming the contribution of our work to enhancing cybersecurity within the smart home paradigm.

II. RELATED WORKS

In the evolving landscape of smart home security, the integration of IoT devices has necessitated advanced IDS to counteract the increasing sophistication of cyber threats [21]. The literature on this subject demonstrates a rich diversity of approaches, each contributing unique insights and methodologies to enhance intrusion detection capabilities within smart home environments.

A seminal study embarked on the challenge of mitigating spam in IoT devices, introducing a ML-driven approach to assign a 'spamicity score' based on time series analysis. This methodology not only underscores the versatility of ML in enhancing IoT security but also sets a precedent for subsequent research in the domain [22]. Building on this foundational work, another investigation presented a double-layer ML system for intrusion detection that hones in on statistical analysis for feature selection. This effort highlights the critical role of efficient feature extraction in maintaining high detection accuracy while minimizing computational overhead [23].

The narrative progresses with the development of a two-tiered ML-based intrusion detection framework, which employs a suite of algorithms, including random forest (RF), XGBoost, and decision trees (DT). This methodology emphasizes the significance of data preprocessing and feature reduction, showcasing the power of ML in detecting and mitigating potential attacks [24]. Further enriching this discourse, a three-layer hybrid model integrates RF and principal component analysis (PCA), demonstrating an effective strategy for minimizing information loss and accelerating the detection process [25].

The exploration of artificial neural networks (ANN) through the multi-tiered ANN model for intrusion detection (MAMID) represents a pivotal moment in the journey towards optimizing intrusion detection. This study foregrounds the importance of hyperparameter optimization, setting a new benchmark for accuracy and transparency in the field [26]. Complementing these approaches, the EM-FEDE method emerges as a novel technique under few-shot data conditions, employing feature and data enhancement to improve IDS model accuracy by addressing data scarcity [27].

The narrative then shifts to a transformative approach utilizing a Transformer-based architecture for network IDS, which adeptly handles diverse data types through the self-attention mechanism. This advancement not only highlights the model's efficacy in distinguishing between normal and malicious network flows but also illustrates the advantage of incorporating IoT telemetry data [28]. An edge computing-based IDS further innovates by converting network traffic into images for training a convolutional neural network (CNN), demonstrating the potential of deep learning (DL) and data augmentation in addressing privacy concerns and unbalanced data [29].

A stacked recurrent neural network (SRNN) strategy introduces a nuanced approach to combatting botnet attacks, emphasizing the model's ability to capture complex patterns in network traffic. This technique illustrates the capacity for detailed temporal data analysis, enhancing the detection of sophisticated cyber threats [30]. Similarly, a hybrid model that combines bidirectional long short-term memory (BiLSTM) and CNN offers a comprehensive solution to intrusion detection, showcasing the synergy between processing time-series data and extracting salient features [31].

The narrative culminates with the introduction of a long short-term memory (LSTM) network-based anomaly detection system, tailored for the IoT devices' security in smart homes. This model exemplifies the LSTM's adeptness at learning from data over time, marking a significant advancement in the proactive identification and mitigation of cyber threats [32]. Finally, an innovative hybrid ML/DL model integrates various algorithms to form a robust defense mechanism against malicious network traffic. This model not only exemplifies superior performance but also embodies the dynamic evolution of IDSs in response to the ever-changing cybersecurity landscape [33].

Together, these studies form a continuum of innovation, each contributing to the overarching goal of optimizing smart home intrusion detection. They collectively underscore the pivotal role of ML and DL technologies in advancing our ability to protect digital environments, paving the way for more secure, resilient smart home ecosystems.

As we traverse the landscape of existing literature (Table 1), a notable trend emerges in the transition from simulated to real-world datasets. This shift underscores a collective aspiration towards enhancing the practical applicability and robustness of IDS solutions. Our method, employing the CICIoT2023 dataset, epitomizes this evolution, harnessing real data from an unparalleled array of 105 home devices. This leap in dataset realism and scope addresses a critical gap in prior endeavors, offering a more comprehensive understanding of the intricacies involved in safeguarding smart homes.

Parallel to the diversification of datasets is the progression towards more complex and hybrid detection models. This

TABLE 1. Comparison of intrusion detection models in smart home.

Authors	Year	Dataset		Detection		Model	Hyperparameter optimization
Yuan et al. [29]	2020	UNSW-NB15	simulated, no IoT device	intrusion	binary, multiclass	CNN	-
Zainab et al. [22]	2020	Weather Information	real, 1 home device	spam	binary	XGBoost	-
Li et al. [23]	2020	collected by the authors	real, 4 home devices	intrusion	binary, multiclass	DT	-
Alghayadh et al. [24]	2020	CSE-CIC-IDS2018	simulated, no IoT device	intrusion	multiclass	RF	-
Segun et al. [30]	2021	Bot-IoT	simulated, 5 home devices	botnet	binary, multiclass	SRNN	-
Elsayed et al. [31]	2021	IoT Intrusion	real, 2 home devices	intrusion	binary	BiLSTM-CNN	-
Azumah et al. [32]	2021	IoT Intrusion	real, 2 home devices	intrusion	binary	LSTM	-
Lingyun et al. [25]	2021	Bot-IoT	simulated, 5 home devices	intrusion	binary, multiclass	KNN-DT-SVM-DT	-
Butt et al. [33]	2022	CIC-IoT 2022, UNSW-NB15	real, 60 home devices simulated, no IoT device	intrusion	binary	KNN-DT-LSTM	Keras Tuner
Sohail et al. [26]	2022	IoTID20	real, 2 home devices	intrusion	binary, multiclass	MAMID	ANN
Wang et al. [28]	2023	ToN_IoT	simulated, 10 home devices	intrusion	binary, multiclass	FT-Transformer	-
Hızal et al. [34]	2024	DS2oS	simulated, 8 home devices	intrusion	binary, multiclass	RF	-
Our method	2024	CICIoT2023	real, 105 home devices	intrusion	binary, multiclass	ETC	Harmony search algorithm (HSA)

trajectory mirrors the research community's pursuit of higher detection accuracy and reliability against multifaceted threats. Our approach, which introduces the HEET model, distinguishes itself by marrying the advantages of ensemble learning with the finesse of hyperparameter optimization. This synergy not only elevates model performance but also exemplifies a refined strategy to navigate the challenges posed by evolving cyber threats.

The discourse on model sophistication would be incomplete without addressing the critical role of hyperparameter optimization. Historically, this aspect has been somewhat overlooked, despite its significance in unlocking the full potential of IDS models. Our method's adoption of harmony search for hyperparameter tuning heralds a new era of precision and effectiveness, paving the way for a systematic and optimized approach to model configuration that had been absent in earlier studies.

Moreover, the evolution from binary to multiclass detection models, coupled with an expanding coverage of home devices, marks a significant advancement in the field. Our method not only supports multiclass detection but also showcases an exceptional scalability and adaptability to a diverse array of smart home environments. This capability is instrumental in crafting a more resilient and versatile IDS that can cater to the nuanced demands of modern smart homes.

In culminating our comparison, it becomes evident that our proposed method stands at the confluence of several evolutionary streams within the domain of smart home IDS. By leveraging a comprehensive dataset, pioneering an advanced modeling technique, integrating innovative hyperparameter optimization, and ensuring scalability across a vast number of devices, our approach offers a holistic solution to the challenges of intrusion detection in smart homes.

This narrative not only situates our work within the grand mosaic of IDS research but also highlights its novel contributions and potential to shape future directions in smart home security. As we gaze towards the horizon, the insights gleaned from this comparative analysis inspire a vision for further exploration and innovation, promising a safer and more secure digital future for our homes.

III. MATERIALS AND METHODS

A. PROPOSED NETWORK INTRUSION DETECTION SYSTEM FOR SMART HOME SECURITY

As illustrated in Figure 1, the proposed architecture for the network IDS is meticulously designed to bolster the security framework of smart home networks. The infrastructure is orchestrated to create a seamless yet secure web of interconnected devices that collectively form the smart home ecosystem.

At the heart of the network, the router serves as the gateway between the smart home and the vast expanse of the internet, facilitated by the service provider's connection. A Cisco switch, renowned for its reliability, is tethered to this internet gateway, acting as the nerve center that orchestrates the flow of digital communication within the home.

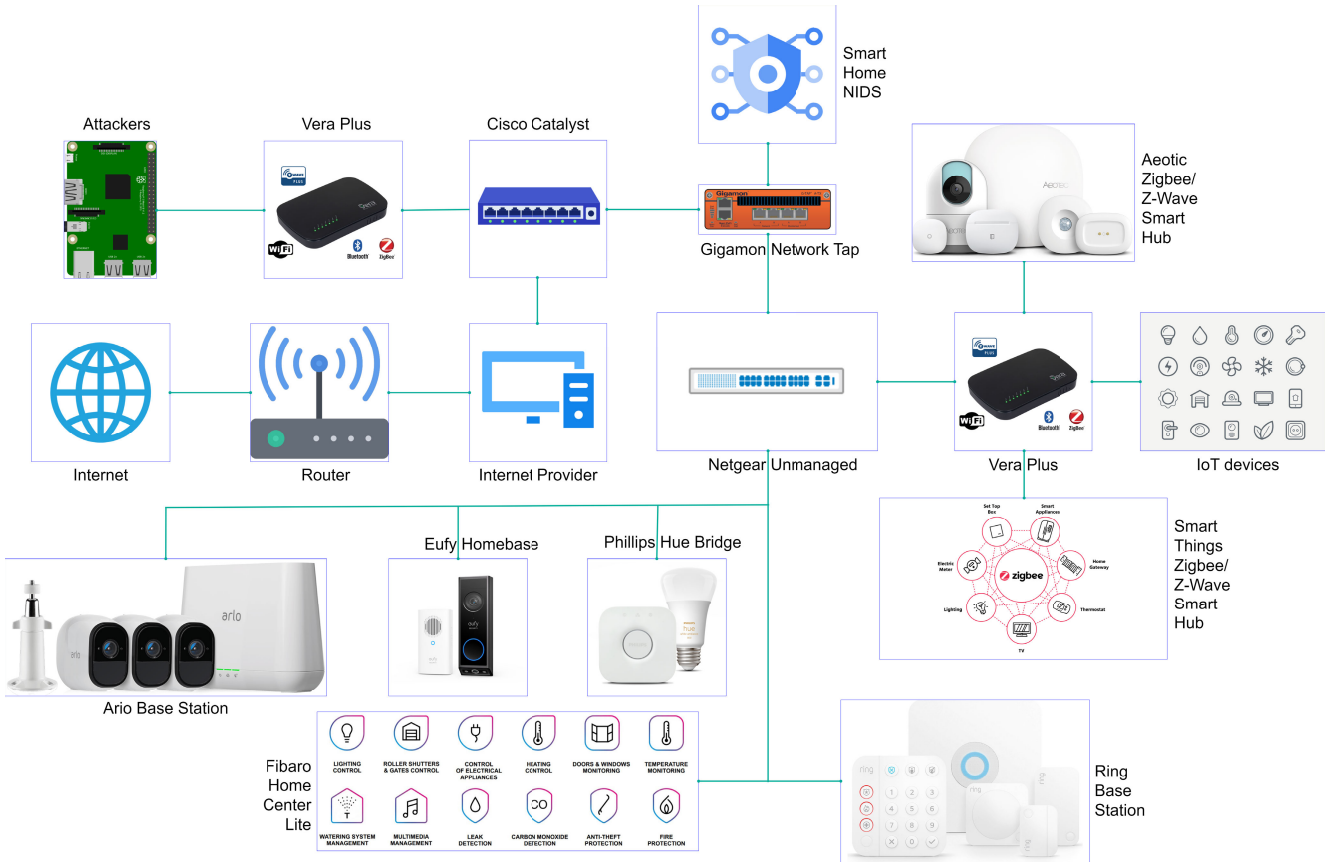

FIGURE 1. **Smart home network architecture.**

Positioned strategically within the network topology is the Gigamon Network Tap, a critical component tasked with the surveillance of all IoT traffic. This device is paramount in ensuring that a mirror image of the network's pulse – every packet, every byte – is relayed to the network monitor. It is here, within the confines of this dedicated monitoring system, where our proposed network IDS comes to the fore, vigilantly capturing traffic with the aid of Wireshark, a tool synonymous with network analysis excellence.

The network IDS operates with a singular mandate: to monitor, detect, and respond to anomalies within the network traffic. By employing a network tap, we ensure a dedicated and unobtrusive observation point that does not hinder network performance. This passive, full-duplex system allows for a comprehensive analysis of the traffic, both incoming and outgoing, ensuring that any and all discrepancies are promptly identified and addressed.

The subsequent layer of the network sees the Netgear Unmanaged Switch in play, a device chosen for its capability to facilitate unfettered communication among IoT devices leveraging protocols such as Zigbee and Z-Wave – protocols that are the lifeblood of IoT communication. Further integrated into this layer is the VeraPlus controller, acting as a conduit between the switch and the myriad of IoT devices it governs.

The unique characteristic of incorporating IoT devices as malicious agents within smart home network architectures distinguishes this approach from traditional IoT security measures. This aspect is crucial for understanding and combating the evolving landscape of IoT-based cyber threats, where compromised IoT devices are increasingly being utilized as vectors for launching attacks. The deployment of IoT devices as malicious agents within smart home network architectures addresses a critical gap in existing research and security practices. Traditionally, IoT security approaches have focused on external threats targeting IoT ecosystems. However, considering the architecture where IoT devices themselves are compromised and turned into sources of attack reflects a growing trend in cyber threats. This approach provides a more comprehensive view of the security vulnerabilities inherent in smart home environments and other IoT ecosystems.

This innovative approach is well-reflected in the smart home network architecture, as illustrated in Figure 1. By incorporating IoT devices as sources of malicious traffic, the smart home network architecture offers a unique opportunity to develop and refine IDS capable of identifying and mitigating threats originating from within the network. This characteristic enables researchers and security professionals to simulate and study the behavior of compromised IoT devices in a controlled environment, leading to the

development of more effective detection algorithms and security measures.

The proposed network IDS installation for smart homes is a robust architecture that not only prioritizes the security of the network but does so while maintaining the integrity and performance of the home's digital ecosystem. By capturing and analyzing every nuance of network traffic, the system ensures that the smart home remains a bastion of security, safeguarded against the spectrum of cyber threats that loom in the digital age.

B. EXTRA TREES CLASSIFIER

In the domain of predictive analytics for smart home security, the ETC emerges as an exemplar of ensemble ML methodologies [35]. This algorithm, standing for Extremely Randomized Trees, is predicated on the principle of aggregating the results of multiple de-correlated decision trees to form a comprehensive predictive model [36]. Its inherent robustness and accuracy make it an ideal candidate for the complex task of intrusion detection within the intricate web of smart home IoT devices.

The efficacy of the ETC lies in its construction of numerous decision trees, which are grown on the entirety of the data sample rather than a bootstrap subset, as is common in other ensemble methods like Random Forests [37]. Each tree in the extra trees ensemble is built from a randomly sampled dataset of features, ensuring that the bias-variance trade-off is navigated with finesse. This randomization goes further in extra trees, extending to the splitting thresholds, which are chosen completely at random, thereby increasing the diversity among the individual trees within the ensemble (Figure 2).

Adapting this classifier to the realm of smart home security involves harnessing its capability to handle vast and diverse datasets that are characteristic of smart home environments. The ETC is adept at discerning patterns within the noise of high-dimensional data, a skill that is paramount in detecting anomalous behavior indicative of cybersecurity threats.

The integration of the ETC into an IDS for smart homes entails an analysis of network traffic data to identify potentially malicious activity [38]. In this capacity, the classifier serves as a guardian, sifting through the data to detect outliers that may signify a security breach. The nature of IoT devices, which frequently operate on limited computational power, calls for efficient algorithms that can deliver high accuracy without impeding system performance. The ETC meets these requirements, offering a balance of speed and precision.

In the context of smart home security, where decisions may have direct implications on user privacy and safety, the interpretability of the model is as crucial as its performance. The ETC, while inherently a black-box model, can be coupled with interpretability frameworks to elucidate the decision-making process, thereby fostering trust among stakeholders.

The ETC, with its ensemble approach and randomized decision-making, represents a powerful tool in the security infrastructure of smart homes. Its deployment in IDS systems exemplifies a sophisticated application of ML to safeguard the digital sanctity of our living spaces. As smart home technologies continue to evolve, the ETC will undoubtedly remain a cornerstone in the development of advanced, reliable, and user-centric security solutions.

C. HARMONY SEARCH ALGORITHM

The HSA is a music-inspired optimization technique developed by Geem et al. in 2001 [39]. The HSA, inspired by the improvisational process of musicians in search of a perfect harmony, encapsulates a novel approach to optimization [40]. It posits that the solution to an optimization problem can be conceived as a harmony that resonates with the highest degree of compatibility with the desired objectives [41]. This paradigm shift introduces a framework where solutions are iteratively improvised, much like melodies, within the constraints of the problem space, striving for an optimal composition [42]. The algorithm is particularly known for its simplicity and has been successfully applied to various optimization problems [43].

Central to the HSA is the harmony memory (HM), an analog to the collective musical memory, serving as a repository of candidate solutions [44]. These solutions, akin to musical chords, are represented as vectors of decision variables, each reflecting a potential harmony within the solution space. Below is a detailed mathematical representation of the key elements and steps in the HSA [45].

KEY ELEMENTS

- **Harmony Memory (HM)**: The HM is a matrix of size $HMS \times D$, where HMS is the HM size and D is the number of decision variables in the optimization problem. It can be represented as:

$$HM = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1D} \\ x_{21} & x_{22} & \cdots & x_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ x_{HMS1} & x_{HMS2} & \cdots & x_{HMSD} \end{bmatrix}$$

  Each row in HM represents a solution vector (x_i), and each column represents a decision variable.
- **Objective Function**: The objective function to be optimized (minimized or maximized) is denoted by $f(x)$, where $x = [x_1, x_2, \ldots, x_D]$ is a solution vector.

ALGORITHM PARAMETERS

- **Harmony Memory Considering Rate (HMCR)**: Probability of selecting a value from the historical values in HM, denoted by $HMCR \in [0, 1]$.
- **Pitch Adjusting Rate (PAR)**: Probability of adjusting a selected value, denoted by $PAR \in [0, 1]$.

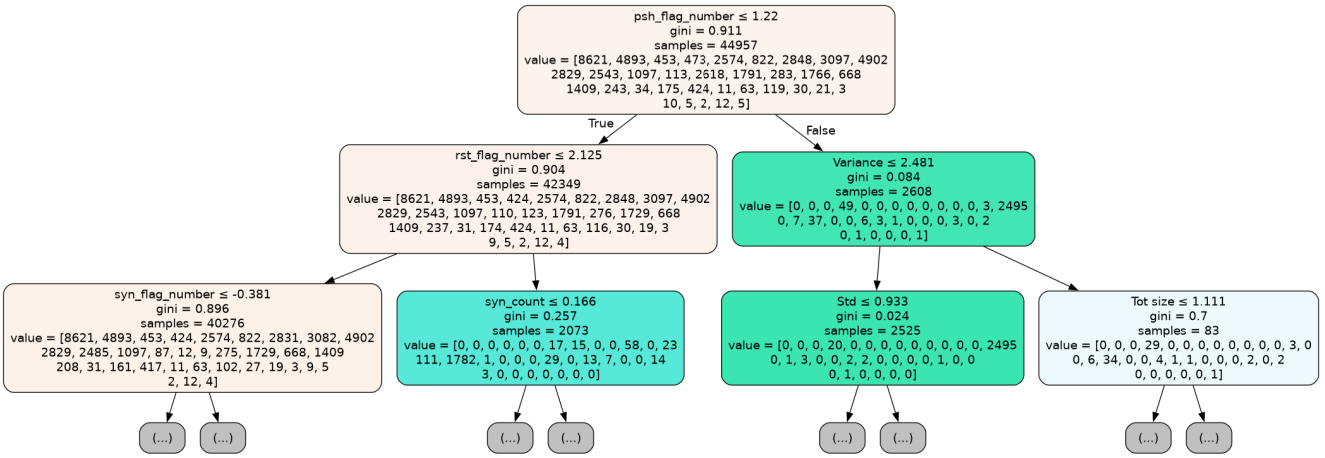

FIGURE 2. Extra trees classifier.

- **Bandwidth (BW)**: The range of pitch adjustment, controlling the degree of local search.

PROCESS

1) **Initialization**: Generate *HMS* random solution vectors within the decision variable bounds and evaluate them using the objective function $f(x)$.
2) **New Harmony Improvisation**: For each decision variable x_i in the new harmony vector:
   - With probability *HMCR*, select a value from the i-th column of *HM*.
   - With probability *PAR*, adjust this value by $\pm BW$, where the adjustment can be represented as:

$$x_i' = x_i \pm \text{rand}() \times BW$$

   x_i' is the new value of the decision variable after adjustment.
   - With probability $1 - HMCR$, generate a random value for x_i within its bounds.
3) **Harmony Memory Update**: If the new harmony vector x' yields a better objective function value than the worst harmony in *HM*, replace the worst harmony with x'.
4) **Termination**: Repeat the improvisation and update steps until a stopping criterion (e.g., a maximum number of iterations) is met.

OPTIMIZATION GOAL

Find x^* such that $f(x^*)$ is optimized (minimized or maximized), where x^* is the best solution vector in *HM* after the final iteration.

This mathematical framework encapsulates the core mechanisms of the HSA, highlighting its reliance on stochastic processes for both exploration (generating new solutions) and exploitation (refining existing solutions) within the search space.

The pseudocode captures the essence of the HSA, emphasizing the balance between exploiting historical solutions stored in the HM and exploring new potential solutions through random variation and adjustment (Algorithm 1). The parameters HMCR, PAR, and BW play crucial roles in dictating the behavior of the improvisation process, enabling the algorithm to adaptively search the solution space for optimal or near-optimal solutions.

The improvisation process underpinning the HSA is a meticulous balance between exploration and exploitation. Decision variables for new candidate solutions are either drawn from the HM, embodying the exploitation of accumulated knowledge, or randomly generated, symbolizing the exploration of uncharted territories. The PAR mechanism introduces a probabilistic fine-tuning of these variables, mirroring the subtle modulations in musical improvisation that can transform a melody. The algorithm's iterative nature ensures a continuous evolution of the solution space, progressively leading to the emergence of an optimal or near-optimal solution.

Upon the generation of a new harmony, its compatibility with the objective function is assessed. Superior harmonies are preserved within the HM, displacing lesser harmonies, thus ensuring that the collective memory evolves towards increasingly compatible solutions.

The algorithm advances towards convergence through a series of improvisations, culminating once a pre-established termination criterion is satisfied. The best solution within the HM at the conclusion of the algorithm represents the optimal harmony, analogous to the culmination of a musical composition that achieves a resonant and satisfying conclusion.

The HSA epitomizes the elegance of leveraging natural and artistic principles for computational purposes. It stands as a profound example of how the search for optimization solutions can transcend traditional boundaries, embodying an approach that is both methodical and inspired. Through its adaptive and iterative exploration of the solution space, the HSA offers a versatile and effective tool for addressing a wide array of optimization problems, harmonizing the complexity of decision variables into a symphony of optimal solutions.

Algorithm 1 Harmony Search Algorithm

```
Initialize algorithm parameters:
    HMS (Harmony Memory Size)
    HMCR (Harmony Memory Considering Rate)
    PAR (Pitch Adjusting Rate)
    BW (Bandwidth)
    NI (Number of Improvisations or iterations)
Initialize Harmony Memory (HM):
for i = 1 to HMS do
    Generate a random solution vector X_i
    Evaluate the objective function f(X_i)
    Insert X_i into HM
end for
Improvise a New Harmony:
for k = 1 to NI do
    for each decision variable j in the solution vector do
        if rand() < HMCR then
            Select a value from HM for the j-th variable
            if rand() < PAR then
                Adjust the selected value by a random amount within ±BW
            end if
        else
            Generate a random value for the j-th variable within its bounds
        end if
    end for
    Evaluate the objective function for the new harmony
    if the new harmony is better than the worst harmony in HM then
        Replace the worst harmony in HM with the new harmony
    end if
end for
Termination:
    The algorithm terminates after completing NI improvisations.
Output:
    The best solution in HM is considered the optimal solution.
```

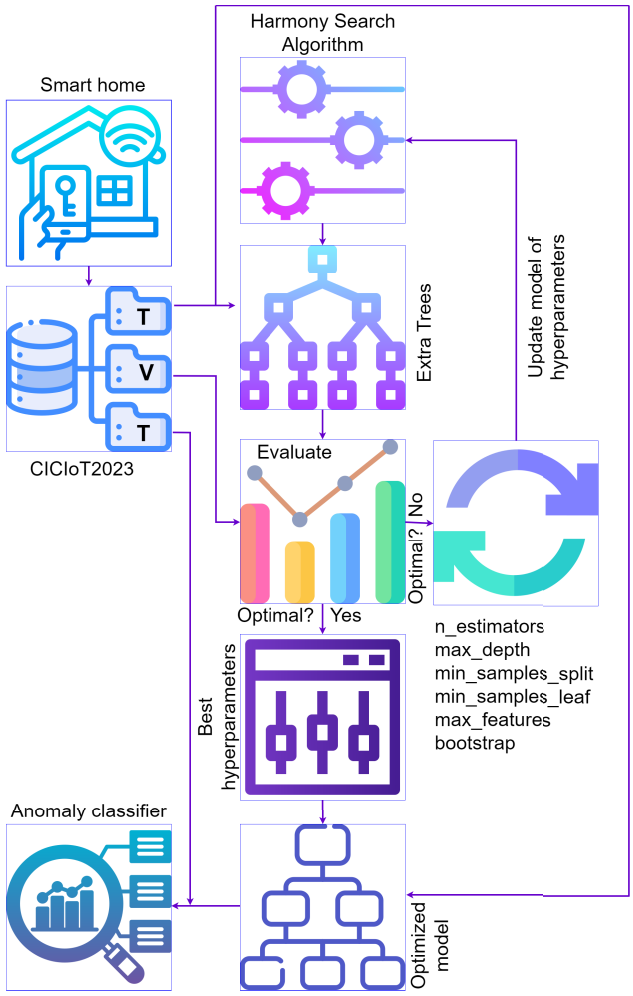

FIGURE 3. Harmony-enhanced extra trees model.

D. PROPOSED HARMONY-ENHANCED EXTRA TREES MODEL

In the evolving landscape of smart home security, the paramount challenge lies in the development of robust IDSs that can adapt to the nuanced and dynamic nature of threats. This paper introduces an optimized model that harnesses the synergy between harmony search optimization and the ETC, coined as the HEET model, aimed at elevating the precision and efficiency of smart home IDSs (Figure 3).

The HEET model is predicated on the ETC, renowned for its efficacy in handling complex classification tasks through the aggregation of multiple de-correlated trees. The classifier's performance is contingent upon the fine-tuning of its hyperparameters, which include the number of estimators, the maximum depth of trees, the minimum samples split, the minimum samples leaf, the maximum features to consider for the best split, and the bootstrap option. The conventional method of manually selecting these hyperparameters is both time-consuming and prone to suboptimality.

To address this, we employ the HSA, an optimization technique inspired by the improvisation process of musicians, to autonomously and intelligently fine-tune the hyperparameters of the ETC. The HSA iterates through a solution space, guided by a HM that stores a set of high-quality solutions, which are analogous to musical harmonies. Through mechanisms akin to musical improvisation – such as HM consideration, pitch adjustment, and random selection – the HSA explores and exploits the solution space to discover optimal or near-optimal hyperparameter configurations.

The proposed HEET model integrates the HSA with the ETC through a custom objective function. This function evaluates the classification accuracy of the model on a validation set, with the aim of maximizing accuracy (or equivalently, minimizing the negative accuracy). The HSA systematically adjusts the hyperparameters of the ETC,

leveraging the objective function's feedback to converge towards an optimal set of hyperparameters.

The harmony search begins with the initialization of the HM with random values within predefined ranges for each hyperparameter. These ranges are chosen based on empirical evidence and theoretical considerations to encompass potential optimal values. As the search progresses, new harmonies – representing candidate hyperparameter sets – are generated and evaluated against the objective function. The algorithm dynamically updates the HM with superior harmonies, gradually refining the search towards the optimal hyperparameter configuration.

Upon completion of the harmony search, the optimal set of hyperparameters is utilized to configure the ETC, thereby constructing the optimized HEET model. The model is then trained on a dataset comprising smart home network traffic, encapsulating a diverse array of intrusion scenarios alongside normal behavior.

The efficacy of the HEET model is underscored by its performance metrics, which exhibit significant improvements in detection accuracy and computational efficiency over conventional models. The training and testing times are meticulously recorded, highlighting the model's practicality for real-time intrusion detection applications in smart home environments.

By iteratively adjusting parameters through a balance of exploration and exploitation, the HSA enhances the ETC's ability to accurately detect and classify network intrusions in real-time. This optimization is crucial for several reasons. First, it maximizes the model's detection capabilities, reducing false positives and false negatives, which are critical in maintaining the integrity and reliability of intrusion detection systems. Second, it ensures that the model can adapt to the dynamic and complex nature of smart home environments, where the diversity of IoT devices and the variability of network traffic present unique challenges. Lastly, the integration of HSA with ETC contributes to the development of scalable and efficient security solutions, capable of protecting against an ever-growing array of cyber threats. This methodological innovation not only strengthens the overall security framework of smart homes but also sets a new benchmark for future research in AI-driven cybersecurity.

E. DATA PREPROCESSING

The CICIoT2023 dataset emerges as a pivotal resource in the domain of IoT security [46], offering an unparalleled depth of insight into the dynamics of IoT-based cyber threats. Distinctively characterized by its inclusion of IoT devices not only as victims but also as perpetrators of cyber-attacks, this dataset encapsulates a broad spectrum of malicious scenarios across 33 distinct attack vectors, categorized into seven major types. Engineered through meticulous simulations within an IoT network comprising over 100 varied devices, the dataset aims to mirror the complexity and diversity of real-world IoT environments. It provides researchers and security professionals with a rich foundation for developing, testing, and enhancing IDSs tailored to the nuanced vulnerabilities of smart home ecosystems. The CICIoT2023 dataset's comprehensive coverage, realism, and focus on IoT devices as malicious agents uniquely position it as a cornerstone for advancing IoT security analytics and fostering the development of robust, effective security solutions capable of addressing the evolving landscape of IoT threats.

In the quest to harness the potential of the CICIoT2023 dataset for enhancing smart home intrusion detection, a comprehensive data preprocessing strategy was meticulously crafted and executed. This dataset, a compilation of 169 csv files representing a rich tapestry of IoT network interactions, including 33 distinct attack types alongside benign traffic, forms the foundation of our study. With a formidable count of 46,686,579 entries, the preprocessing phase was pivotal in refining and optimizing the dataset for analytical rigor and computational efficiency. Herein, we delineate the methodical steps undertaken in this critical phase of our research.

In the foundational stage of optimizing the CICIoT2023 dataset for smart home intrusion detection, the preprocessing began with the meticulous aggregation of csv files into pandas DataFrames. A critical step in this phase was the conversion of features from float64 to float32 data types during the initial loading process. This strategic choice was driven by the dual objectives of preserving the dataset's numerical integrity and significantly reducing its memory footprint.

Subsequent to the aggregation of the individual csv files into a singular cohesive DataFrame, a further optimization was undertaken. Features of the 'object' data type were evaluated for conversion to the 'category' type, contingent upon a uniqueness threshold set at 50% of the dataset's entries. This conversion, applied to features meeting this criterion, served to further optimize memory usage while maintaining the categorical nature of the data.

A critical observation noted was the presence of features predominantly populated with zeros, exceeding a threshold of 99%. Such features were deemed to contribute minimal informational value to the analysis and were thus excised from the DataFrame. This pruning of data not only streamlined the dataset but also honed the focus onto more impactful features.

The pursuit of data integrity necessitated a thorough examination for null values, ensuring that the analytical foundation was devoid of gaps that could skew results. Additionally, the elimination of duplicate entries was imperative to maintain the dataset's integrity, ensuring that each data point contributed uniquely to the analysis. This step was instrumental in refining the dataset to 28,098,546 entries across 36 features, a testament to the rigorous cleaning process.

The processing of the CICIoT2023 dataset has yielded a detailed taxonomy of network traffic, pivotal for informing the development of IDSs (Table 2). This disparity in record counts across attack types reflects the varied nature of IoT security threats, from the most common to the rarefied,

TABLE 2. Distribution of attack records.

Class			Records
Attack	DDoS	DDoS-UDP_Flood	5,412,287
		DDoS-SynonymousIP_Flood	3,065,966
		DDoS-SYN_Flood	1,933,447
		DDoS-ICMP_Flood	1,809,173
		DDoS-PSHACK_Flood	1,647,084
		DDoS-TCP_Flood	1,569,605
		DDoS-RSTFINFlood	1,071,959
		DDoS-ICMP_Fragmentation	443,979
		DDoS-UDP_Fragmentation	286,925
		DDoS-ACK_Fragmentation	274,933
		DDoS-HTTP_Flood	28,772
		DDoS-SlowLoris	23,426
	DoS	DoS-UDP_Flood	2,959,733
		DoS-TCP_Flood	1,778,908
		DoS-SYN_Flood	1,629,596
		DoS-HTTP_Flood	71,786
	Recon	Recon-HostDiscovery	134,345
		Recon-OSScan	98,112
		Recon-PortScan	82,124
		VulnerabilityScan	37,382
		Recon-PingSweep	2,260
	Web-Based	BrowserHijacking	5,859
		CommandInjection	5,409
		SqlInjection	5,245
		XSS	3,846
		Backdoor_Malware	3,218
		Uploading_Attack	1,252
	Brute Force	DictionaryBruteForce	13,064
	Spoofing	MITM-ArpSpoofing	307,591
		DNS_Spoofing	178,873
	Mirai	Mirai-udpplain	890,576
		Mirai-greeth_flood	673,232
		Mirai-greip_flood	550,402
Normal	Normal	BenignTraffic	1,098,177

sophisticated attacks. Through this processing, the dataset has been distilled into a refined resource, facilitating the training of models to detect and classify the complex array of attacks characteristic of today's IoT security challenges.

The dataset's categorical labels, representing a broad spectrum of both benign and malicious network activities, were systematically converted into numerical form. This conversion was essential for enabling the ML algorithms to process the data, as these algorithms inherently require numerical input. By assigning unique numerical identifiers to each category of network activity, we maintained the integrity of the dataset's classification diversity while rendering it amenable to algorithmic analysis.

After segregating the features from the target variable, the dataset underwent a strategic partitioning process. The division allocated 70% of the data for training, allowing the model to learn from a majority of the dataset. To ensure robust model validation and testing, the remaining data was divided into validation and test sets, receiving 10% and 20% of the dataset respectively. This structured split supports a comprehensive evaluation framework, where model performance can be iteratively refined using the validation set before final assessment on the test set.

To address potential bias arising from variable scales among features, a normalization process was applied to the dataset. This process, crucial for preserving model fairness across features, involved adjusting the data to a common scale without distorting differences in the ranges of values. By standardizing the feature set, we ensured that each feature contributed equally to the analysis, eliminating skewness towards variables with larger magnitudes.

The data preprocessing journey undertaken for the CICIoT2023 dataset was both exhaustive and meticulous, embodying the synthesis of computational efficiency and data integrity. By optimizing data types, excising redundancy, and ensuring the uniqueness of entries, the dataset was transformed into a streamlined and potent resource for intrusion detection analysis. This foundational work not only facilitates the development of the HEET model but also sets a benchmark for data preprocessing in the realm of IoT security research. Through these painstaking efforts, the dataset stands primed to unveil insights into the detection and classification of IoT-based cyber threats, paving the way for advancements in smart home security.

F. EVALUATION METRICS

The experimental evaluation was conducted on a high-performance computing platform equipped with an Intel(R) Core(TM) i9-9900KF CPU @ 3.60GHz, 16 CPU cores, and 32GB of RAM. The experiments were implemented using Python, leveraging libraries such as Pandas and NumPy for data manipulation, and Matplotlib and Seaborn for data visualization. The scikit-learn library was employed for ML tasks, including model training and evaluation.

In alignment with our commitment to transparency and fostering collaborative research, we are pleased to announce that the entire codebase supporting our experiments, including the implementation of the HSA and the optimized ETC, is openly available. Researchers, practitioners, and enthusiasts are encouraged to access, utilize, and contribute to our repository hosted on GitHub at the following link: https://github.com/TATU-hacker/Harmony-Enhanced-Extra-Trees.git (uploaded on 16 March 2024).

The evaluation of the HEET model's performance in identifying various attack vectors within the smart home IoT environment is critical to establishing its efficacy. In this pursuit, we employ a suite of metrics, each providing a distinct perspective on the model's predictive capabilities [47]. Here, we elucidate on four key metrics: Accuracy, Precision, Recall, and the F-Score [48], [49].

1) Accuracy is the most straightforward metric for assessing a model's overall performance. It is calculated using the formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (1)$$

where TP is true positives, TN is true negatives, FP is false positives, and FN is false negatives.

2) Precision measures the correctness of positive predictions made by the model and is defined as the ratio of true positives to the sum of true positives and false positives. The formula for precision is:

$$Precision = \frac{TP}{TP + FP}. \quad (2)$$

3) Recall assesses the model's ability to identify all relevant instances correctly and is calculated as the ratio of true positives to the sum of true positives and false negatives. The formula for recall is:

$$Recall = \frac{TP}{TP + FN}. \quad (3)$$

4) The F-Score provides a balance between precision and recall by taking their harmonic mean. It is particularly useful when seeking a balance between recognizing all positives and maintaining high precision. The F-Score is calculated as:

$$F1 - score = 2 \times \frac{(Recall \times Precision)}{(Recall + Precision)}. \quad (4)$$

These metrics collectively offer a multifaceted view of the HEET model's performance, illuminating its strengths and areas for improvement. Through the lens of these metrics, we gain profound insights into the model's operational capabilities, guiding us towards refining its predictive prowess for smart home intrusion detection applications.

IV. EXPERIMENTS AND RESULTS

In the forthcoming exegesis of our empirical findings, we embark upon a layered exposition that traverses the landscape of computational discernment facilitated by an optimally tuned ETC. We commence with a discourse on the meticulous calibration of the model's parameters, a process underpinned by the HSA's adeptness at navigating the multidimensional space of potential solutions. This is followed by an analytical retrospection of the binary classification trials, wherein the model's sagacity in distinguishing between the conventional and the aberrant is brought to light. Thereafter, we transition to an exploration of the model's capacity to parse through an expanded categorical schema, elucidating its competence in the stratification of cyber threats. The culmination of our narrative arrives with a granular investigation into the algorithm's precision in demarcating a comprehensive array of specific adversarial incursions, demonstrating its acuity in the nuanced detection of cybersecurity threats.

A. HYPERPARAMETER OPTIMIZATION RESULTS

In this subsection, we present an analytical narrative detailing the fruits of the HSA's rigorous quest to optimize the ETC's hyperparameters. Emphasizing the algorithm's systematic prowess, this subsection lays the foundation for understanding the empirical enhancements achieved in the model's configuration. The results encapsulate a strategic calibration of parameters that have been tuned harmoniously to resonate with the varying complexities and characteristics of the dataset under study. These optimized parameters set the stage for the ensuing empirical analyses, offering a prelude to the depth and breadth of the classifier's capabilities in the subsequent testing phases for anomaly detection, attack categorization, and specific attack identification.

TABLE 3. Optimized extra trees hyperparameters using harmony search algorithm.

Hyperparameter	Type	Range	Class 2	Class 8	Class 34
n_estimators	int	[10, 200]	195	190	200
max_depth	int	[1, 20]	17	18	19
min_samples_split	int, float	[2, 11]	3	5	5
min_samples_leaf	int, float	[1, 10]	3	2	3
max_features	int, float	sqrt, log2, None	None (0.8837)	None (0.7937)	None (0.8504)
bootstrap	bool	True, False	True	True	True

Within the dichotomous framework, the HSA orchestrated the hyperparameters to a state of equilibrium, concretizing at 195 estimators with an arboreal depth of 17 – a configuration suggestive of both model complexity and generalizability. The internal bifurcation of nodes and the minimal foliage were each set at a trifecta, a delicate balance encouraging both model depth and sufficient sample representation. Notably, the max_features parameter resonated with the ensemble at 'None', an indication that the totality of the feature set contributes to the individual decision trees, with the bootstrap modality affirmed, engendering sample diversity with replacement.

In the octal classification paradigm, the algorithm finely tuned the ensemble to consist of 190 estimators, permitting them to mature to a depth of 18. The algorithm subtly increased the min_samples_split to 5, enhancing the model's discriminative capacity. The min_samples_leaf parameter, optimized to a duo, indicated a refined leafiness conducive to model precision. The max_features parameter, while remaining unbound, implicitly embraced a subset equivalence of 0.7937, a tacit nod to feature proportionality in the pursuit of an optimal split. The bootstrap parameter, steadfast in its true setting, corroborated the model's preference for bootstrapped sampling.

Navigating the thirty-four class terrain, the algorithm discerned an optimal arboretum of 200 estimators with the liberty to delve to a depth of 19 – a directive pointing to the model's intricate delineation capabilities. The min_samples_split and min_samples_leaf were poised at 5 and 3, respectively, a calibration that reflects an astute balance between depth and breadth, ensuring sample adequacy and representativeness. Here again, max_features abided by its precedent of 'None', albeit with an implicit proportionality factor of 0.8504, intimating a comprehensive yet focused approach to feature selection. The consistent endorsement of the bootstrap paradigm underscored a

TABLE 4. Performance comparison of binary testing.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Training Time (s)	Testing Time (s)
ETC	99.58	96.67	97.86	97.26	923.394	39.658
HEET	99.87	97.41	98.45	97.92	1,105.328	45.744

commitment to sampling robustness across the classifiers' architectural spectrum.

The detailed results of the HSA's hyperparameter optimization are summarized in Table 3, elucidating the tuned parameters that culminated in the HEET models for each class configuration.

B. 2-CLASS TESTING FOR ANOMALY DETECTION

In this subsection of our study, we meticulously compare the performance of the baseline ETC against its Harmony-Enhanced counterpart (HEET). This comparison offers a lens through which we can discern the advantages rendered by the intricate parameter adjustments afforded by the HSA.

The critical task of distinguishing between 'Normal' and 'Attack' network traffic hinges upon the nuanced capability of the underlying classifier. Through rigorous testing, we have determined the performance of both ETC and HEET models across several standard metrics: accuracy, precision, recall, and f1-score.

Table 4 presents a succinct yet comprehensive comparison of the two models. The HEET model, enhanced through hyperparameter optimization, surpasses its ETC counterpart with an admirable accuracy of 99.87%, an increment indicative of the model's robustness. Precision, a measure of the model's exactness, and recall, a measure of the model's completeness, both see ascensions to 97.41% and 98.45%, respectively, for the HEET model. This culmination of meticulous optimization is further echoed in the enhanced f1-score, a harmonic mean of precision and recall, registering at 97.92%. Such improvements in the HEET model, however, come with the trade-off of increased computational demand, with training and testing times marked at 1105.328 seconds and 45.744 seconds, respectively. This contrast with the ETC model, which, while marginally less precise, is also less computationally taxing, with training and testing times of 923.394 seconds and 39.658 seconds, respectively.

TABLE 5. ETC binary classification report.

	precision	recall	f1-score	support
Attack	1.00	1.00	1.00	5,399,632
Normal	0.94	0.96	0.95	220,078
accuracy			1.00	5,619,710
macro avg	0.97	0.98	0.97	5,619,710
weighted avg	1.00	1.00	1.00	5,619,710

The detailed classification reports, delineated in Tables 5 and 6, further accentuate the HEET model's superior discernment between the classes. The 'Attack' class, owing to its clear signature in the data, sees perfect scores across both models. However, the 'Normal' class, which requires a more delicate handling due to its potential overlap with 'Attack' patterns, benefits significantly from the HSA's touch. The HEET model's precision and recall in identifying normal instances are enhanced to 0.95 and 0.97, respectively, as opposed to the ETC model's 0.94 and 0.96.

TABLE 6. HEET binary classification report.

	precision	recall	f1-score	support
Attack	1.00	1.00	1.00	5,399,632
Normal	0.95	0.97	0.96	220,078
accuracy			1.00	5,619,710
macro avg	0.97	0.98	0.98	5,619,710
weighted avg	1.00	1.00	1.00	5,619,710

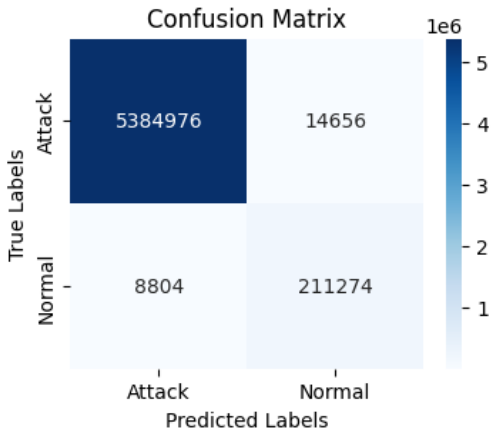

FIGURE 4. ETC binary confusion matrix.

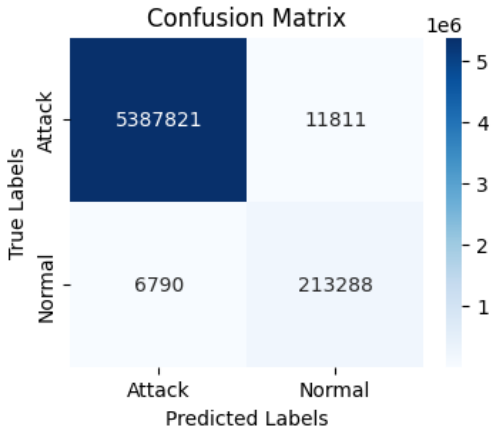

FIGURE 5. HEET binary confusion matrix.

The insights gleaned from the confusion matrices provide a visual representation of the models' performance

TABLE 7. Performance comparison of octal testing.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Training Time (s)	Testing Time (s)
ETC	99.32	95.39	81.11	85.55	4,217.905	201.462
HEET	99.51	96.02	87.14	90.65	5,491.069	74.649

(Figures 4 and 5). For the ETC model, we observe a higher incidence of false positives and false negatives compared to the HEET model. The HEET model's confusion matrix indicates a decrease in false positives (from 14,656 to 11,811) and false negatives (from 8,804 to 6,790), underscoring an enhanced sensitivity and specificity. This signifies not only the HEET model's heightened aptitude in correctly classifying 'Attack' instances but also its refined acuity in preserving the 'Normal' class from misclassification.

These results collectively depict a vivid narrative of the HSA's capacity to amplify the ExtraTreesClassifier's detection capabilities. Through the meticulous optimization of hyperparameters, the HEET model achieves a delicate equilibrium between sensitivity and specificity, making it a more formidable sentinel against network anomalies. It is this very balance that is so coveted in the realm of anomaly detection, where the stakes of misclassification are high, and the precision of detection is paramount. The HEET model, therefore, stands as a testament to the virtue of harmony in the realm of algorithmic optimization, marrying the sophistication of ML with the elegance of algorithmic orchestration to create a bulwark against the spectral threats that lurk within network traffic.

C. 8-CLASS TESTING FOR ATTACK CATEGORIZATION

In the nuanced theater of cybersecurity, the demarcation of malicious intent into distinct categories of attacks is a task of critical complexity. In this subsection delineates the comparative assessment of the ETC and its refined counterpart, the HEET, in their abilities to categorize an array of cyber threats. This exposition details the analytical examination of the models' performance through a multi-dimensional lens of metrics, shedding light on the precision, recall, and the overall accuracy of these predictive models.

TABLE 8. ETC octal classification report.

	precision	recall	f1-score	support
DDoS	1.00	1.00	1.00	3,513,485
DoS	1.00	1.00	1.00	1,287,718
Recon	0.91	0.82	0.87	71,339
Web-Based	0.95	0.48	0.63	4,929
Brute_Force	0.94	0.36	0.52	2,556
Spoofing	0.92	0.86	0.89	97,212
Mirai	1.00	1.00	1.00	422,393
Normal	0.91	0.98	0.94	220,078
accuracy			0.99	5,619,710
macro avg	0.95	0.81	0.86	5,619,710
weighted avg	0.99	0.99	0.99	5,619,710

In assessing the prowess of any classification model, core metrics provide the essential barometer of success. Table 7 presents these metrics as a testament to the models' capabilities. The HEET model, having undergone the HSA's optimization process, manifests a superior accuracy of 99.51%, a testament to its refined predictive abilities. Precision and recall, key indicators of model reliability, are notably higher in the HEET model, presenting a clear indication of its superior discriminative power. The f1-score, a synthesized metric of precision and recall, echoes this narrative, with the HEET model achieving a commendable 90.65%. These metrics are not merely numbers; they are reflective of a model's efficacy in real-world applications, where the distinction between attack types is not merely academic but a matter of security.

TABLE 9. HEET octal classification report.

	precision	recall	f1-score	support
DDoS	1.00	1.00	1.00	3,513,485
DoS	1.00	1.00	1.00	1,287,718
Recon	0.92	0.87	0.89	71,339
Web-Based	0.94	0.61	0.74	4,929
Brute_Force	0.96	0.62	0.75	2,556
Spoofing	0.93	0.89	0.91	97,212
Mirai	1.00	1.00	1.00	422,393
Normal	0.94	0.98	0.96	220,078
accuracy			1.00	5,619,710
macro avg	0.96	0.87	0.91	5,619,710
weighted avg	1.00	1.00	1.00	5,619,710

The temporal dimensions of model training and testing reveal the computational investments necessary for such optimization. The HEET model demands more extensive training time, yet this investment pays dividends in its reduced testing time, emphasizing the HSA's contribution not only to the model's acumen but also to its efficiency during deployment.

The detailed classification reports, depicted in Tables 8 and 9, further reveal the performance nuances between the two models. The ETC model, while showing a commendable macro-average precision and recall, evidences certain weaknesses in categories like 'Web-Based' and 'Brute_Force' attacks, where the recall rates dip to 0.48 and 0.36, respectively. This suggests a vulnerability in distinguishing less frequent or more sophisticated attack types.

Conversely, the HEET model, imbued with the harmonic balance of optimized hyperparameters, registers substantial improvements in these very categories. For 'Web-Based' attacks, the precision surges to 0.94 with a recall of 0.61, while 'Brute_Force' attacks observe a precision of 0.96 and a recall of 0.62. These advancements in precision and recall, leading to increased f1-scores, are indicative of the model's refined capability to differentiate between various attack methodologies with greater accuracy.

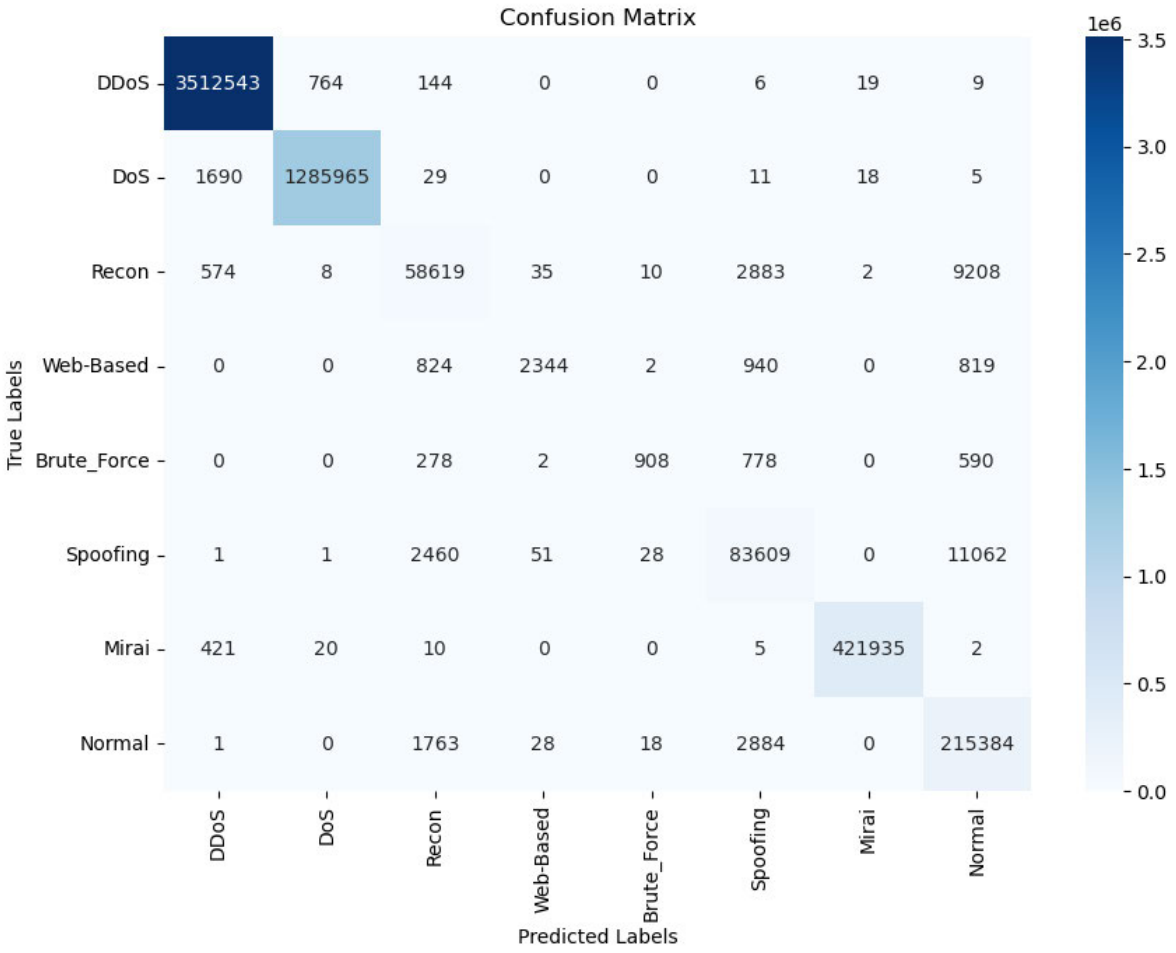

FIGURE 6. ETC octal confusion matrix.

The confusion matrices (Figures 6 and 7) serve as a cartographic representation of the models' classificatory landscapes. For the ETC model, the rows of darker shades along the matrix's principal diagonal signal robust identification for dominant classes like 'DDoS' and 'DoS.' However, the smattering of lighter cells within this main diagonal, particularly for categories like 'Web-Based' and 'Brute_Force,' reveal areas where the model's resolution falters. Here, the matrix tells a story of occasional confusion, where certain attacks masquerade effectively as benign or as other types of intrusions.

In contrast, the HEET model's matrix delineates an enhanced level of clarity in classification. A darker and more consistent diagonal, with fewer and lighter off-diagonal elements, speaks to the model's heightened acuity. Notably, for categories such as 'Recon' and 'Spoofing,' the HEET model exhibits a substantial reduction in false positives and false negatives – crucial for cybersecurity systems where the mislabeling of threats can have dire consequences.

In the distillation of these results, the HEET model distinguishes itself as a paragon of cyber threat categorization. It exemplifies the harmonious intersection of ML prowess and algorithmic precision. Through the judicious optimization of the model's parameters, the HSA has elevated the ETC from a tool of detection to an instrument of discernment – a crucial distinction in the realm of cyber defense.

Thus, this subsection not only underscores the HEET model's superior performance but also celebrates the elegant complexity of its optimization process. It is in this intricate dance of algorithmic refinement that the HEET model's capabilities are fully realized, positioning it as an invaluable asset in the ongoing battle against cyber threats.

D. 34-CLASS TESTING FOR SPECIFIC ATTACK DETECTION

As we embark on detailing this subsection, we present a discerning analysis of the ETC juxtaposed with its Harmony-Enhanced counterpart (HEET). The multiclass classification challenge at hand is the identification and categorization of 34 unique cyber attack vectors – a daunting task for any automated system. This subsection evaluates the two models across a comprehensive suite of performance indicators, inclusive of accuracy, precision, recall, f1-score, and the computational efficiency portrayed through training and testing times, as elucidated in Table 10.

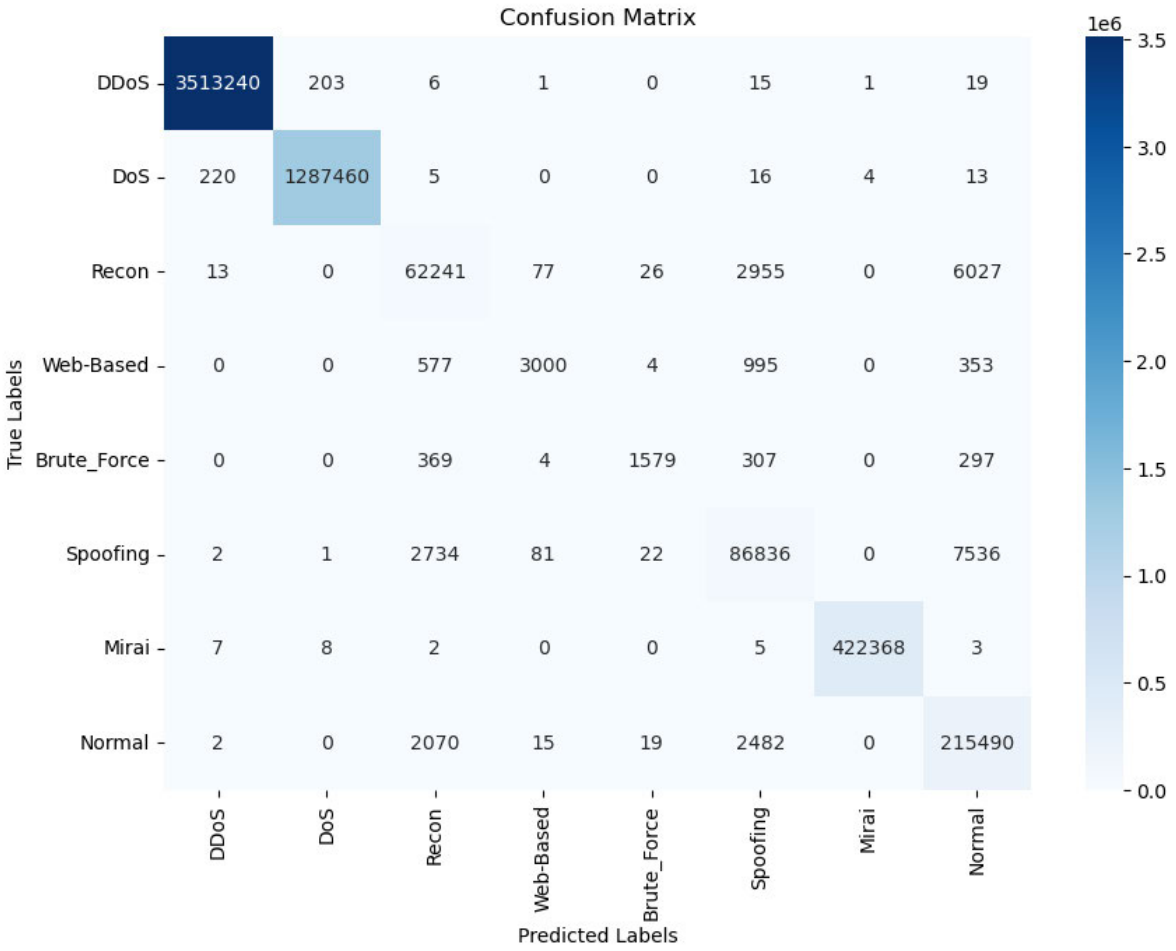

FIGURE 7. HEET octal confusion matrix.

TABLE 10. Performance comparison of thirty-four clas testing.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Training Time (s)	Testing Time (s)
ETC	99.02	93.88	80.86	84.59	4,390.942	383.223
HEET	99.49	96.07	87.10	90.61	15,381.743	170.556

The quantified performance metrics draw a stark contrast between the ETC and HEET models. With an ascension in accuracy to 99.49%, the HEET model underscores its enhanced global classification capabilities. A notable surge in precision to 96.07% is indicative of the model's ability to minimize false positives – a critical factor when each misclassified attack can carry significant ramifications. Recall, at 87.10%, reflects the model's robustness in identifying true positives across all classes, and the f1-score, which reconciles precision and recall, solidifies at a laudable 90.61%.

However, these performance enhancements come at a computational cost. The HEET model necessitates an extensive training time of 15,381.743 seconds, an investment that manifests in the complexity of tuning a model across a vast multitude of classes. Yet, this investment yields a return in the form of reduced testing time, down to 170.556 seconds, suggesting that the HSA not only refines the model's accuracy but also its inferential expediency.

The classification reports for the ETC and HEET models, as presented in Tables 11 and 12, provide an exhaustive account of the models' discriminative abilities across a multitude of specific attack categories. These reports offer a fine-grained analysis of precision, recall, and f1-score metrics for each class, collectively painting a comprehensive picture of model performance.

Precision in classification reports quantifies the accuracy of positive predictions. It reflects the proportion of true positives among all instances labeled as positives for a particular class. For instance, the HEET model's precision score for 'MITM-ArpSpoofing' stands at 0.93, an improvement over the ETC's 0.92. This increment, though seemingly modest, signals fewer false alarms and greater reliability in flagging genuine instances of 'MITM-ArpSpoofing.' The

TABLE 11. ETC thirty-four classification report.

	precision	recall	f1-score	support
DDoS-UDP_Flood	1.00	1.00	1.00	1,082,923
DDoS-SynonymousIP_Flood	1.00	1.00	1.00	611,999
DDoS-UDP_Fragmentation	0.99	0.99	0.99	57,761
MITM-ArpSpoofing	0.92	0.85	0.89	61,591
DDoS-TCP_Flood	1.00	1.00	1.00	313,641
Mirai-greip_flood	0.99	1.00	1.00	110,497
DoS-TCP_Flood	1.00	1.00	1.00	355,509
DDoS-SYN_Flood	1.00	1.00	1.00	387,274
DoS-UDP_Flood	1.00	1.00	1.00	592,518
DDoS-ICMP_Flood	1.00	1.00	1.00	362,278
DoS-SYN_Flood	1.00	1.00	1.00	325,395
Mirai-greeth_flood	1.00	0.99	1.00	133,966
DoS-HTTP_Flood	0.98	0.98	0.98	14,296
DDoS-PSHACK_Flood	1.00	1.00	1.00	328,769
DDoS-RSTFINFlood	1.00	1.00	1.00	214,457
DNS_Spoofing	0.86	0.73	0.79	35,621
BenignTraffic	0.88	0.99	0.93	220,078
DDoS-ICMP_Fragmentation	0.99	0.99	0.99	89,061
Mirai-udpplain	1.00	1.00	1.00	177,930
Recon-HostDiscovery	0.88	0.85	0.86	27,099
DDoS-SlowLoris	0.86	0.97	0.91	4,654
Recon-OSScan	0.82	0.60	0.69	19,790
DDoS-ACK_Fragmentation	1.00	0.99	0.99	54,884
XSS	0.76	0.43	0.55	754
VulnerabilityScan	0.87	0.97	0.92	7,567
Recon-PortScan	0.84	0.60	0.70	16,449
DDoS-HTTP_Flood	0.99	0.96	0.97	5,784
DictionaryBruteForce	0.92	0.36	0.52	2,556
Uploading_Attack	0.94	0.29	0.44	257
BrowserHijacking	0.92	0.31	0.46	1,162
SqlInjection	0.91	0.38	0.53	1,053
Recon-PingSweep	0.83	0.18	0.29	434
CommandInjection	0.85	0.57	0.68	1,069
Backdoor_Malware	0.94	0.54	0.68	634
accuracy			0.99	5,619,710
macro avg	0.94	0.81	0.85	5,619,710
weighted avg	0.99	0.99	0.99	5,619,710

TABLE 12. HEET thirty-four classification report.

	precision	recall	f1-score	support
DDoS-UDP_Flood	1.00	1.00	1.00	1,082,923
DDoS-SynonymousIP_Flood	1.00	1.00	1.00	611,999
DDoS-UDP_Fragmentation	1.00	1.00	1.00	57,761
MITM-ArpSpoofing	0.93	0.88	0.91	61,591
DDoS-TCP_Flood	1.00	1.00	1.00	313,641
Mirai-greip_flood	1.00	1.00	1.00	110,497
DoS-TCP_Flood	1.00	1.00	1.00	355,509
DDoS-SYN_Flood	1.00	1.00	1.00	387,274
DoS-UDP_Flood	1.00	1.00	1.00	592,518
DDoS-ICMP_Flood	1.00	1.00	1.00	362,278
DoS-SYN_Flood	1.00	1.00	1.00	325,395
Mirai-greeth_flood	1.00	1.00	1.00	133,966
DoS-HTTP_Flood	1.00	1.00	1.00	14,296
DDoS-PSHACK_Flood	1.00	1.00	1.00	328,769
DDoS-RSTFINFlood	1.00	1.00	1.00	214,457
DNS_Spoofing	0.85	0.80	0.83	35,621
BenignTraffic	0.93	0.99	0.95	220,078
DDoS-ICMP_Fragmentation	1.00	1.00	1.00	89,061
Mirai-udpplain	1.00	1.00	1.00	177,930
Recon-HostDiscovery	0.88	0.89	0.88	27,099
DDoS-SlowLoris	1.00	1.00	1.00	4,654
Recon-OSScan	0.87	0.68	0.77	19,790
DDoS-ACK_Fragmentation	1.00	1.00	1.00	54,884
XSS	0.93	0.55	0.69	754
VulnerabilityScan	1.00	1.00	1.00	7,567
Recon-PortScan	0.82	0.72	0.77	16,449
DDoS-HTTP_Flood	1.00	1.00	1.00	5,784
DictionaryBruteForce	0.91	0.63	0.75	2,556
Uploading_Attack	0.94	0.52	0.67	257
BrowserHijacking	0.95	0.63	0.75	1,162
SqlInjection	0.89	0.58	0.70	1,053
Recon-PingSweep	0.93	0.56	0.70	434
CommandInjection	0.95	0.62	0.75	1,069
Backdoor_Malware	0.90	0.58	0.71	634
accuracy			0.99	5,619,710
macro avg	0.96	0.87	0.91	5,619,710
weighted avg	0.99	0.99	0.99	5,619,710

HEET model consistently showcases high precision across the board, demonstrating its ability to produce a lower rate of false positives, which is crucial in avoiding unnecessary responses to benign activities mistaken for threats.

Recall, or the true positive rate, gauges the model's capacity to detect all actual positives. It's especially telling in the context of cybersecurity, where missing an attack can have severe consequences. The HEET model's recall for 'DNS_Spoofing' at 0.80 surpasses the ETC's 0.73, indicative of its heightened sensitivity in identifying this specific threat. High recall scores across various classes, such as 'Mirai-greip_flood' and 'DoS-UDP_Flood', underscore the HEET model's adeptness in capturing nearly all instances of these attack types.

The f1-score is the harmonic mean of precision and recall, offering a balanced metric that is crucial when we require a nuanced understanding of a model's performance, particularly in classes where the cost of false negatives and false positives is significant. In the case of 'Recon-HostDiscovery', the HEET model achieves an f1-score of 0.88, revealing a balanced strength in both precision and recall, a valuable trait for nuanced threat detection where neither false positives nor false negatives can be tolerated.

When juxtaposed, the classification reports unveil the HSA's tangible impact on model performance. The algorithm's optimization process is particularly beneficial for attack categories with historically lower recall or precision scores. For categories like 'XSS' and 'SqlInjection', where the ETC model's scores may falter, the HEET model exhibits marked improvements, reflecting the algorithm's success in attuning the model to these intricate attack patterns.

These detailed classification reports have far-reaching implications. They suggest that the HEET model, guided by the HSA, is not only generally more reliable than the ETC model but also provides a more secure detection framework by significantly reducing the chances of missed or false alerts. In the grand tapestry of cybersecurity defense mechanisms, the HEET model thus stands out as a finely woven thread, providing a detailed and dependable barrier against the myriad threats that characterize the digital age.

The confusion matrices are fundamental in evaluating the models' classification abilities, as they reveal not only the successes but also the nuances of misclassifications.

For the ETC Model (Figure 8):

- The prominent dark cells along the diagonal reflect high true positive rates for most attack types, indicating that the model is particularly adept at identifying prevalent attack categories such as 'DDoS-UDP_Flood' and 'DDoS-SYN_Flood'.
- However, closer inspection reveals areas of concern. Lighter cells, notably in rows for 'MITM-ArpSpoofing' and 'DNS_Spoofing', suggest instances where the

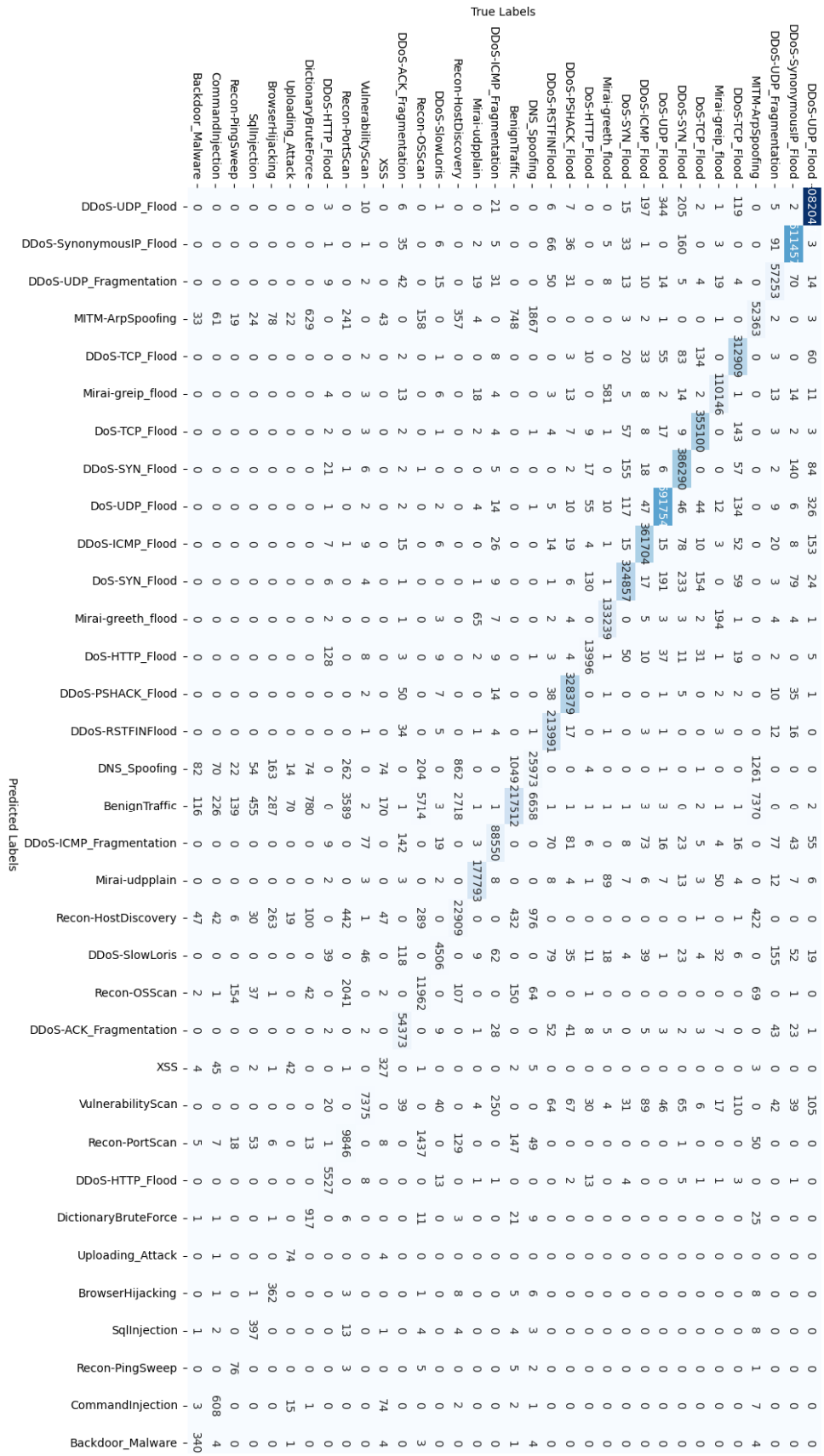

FIGURE 8. ETC thirty-four class confusion matrix.

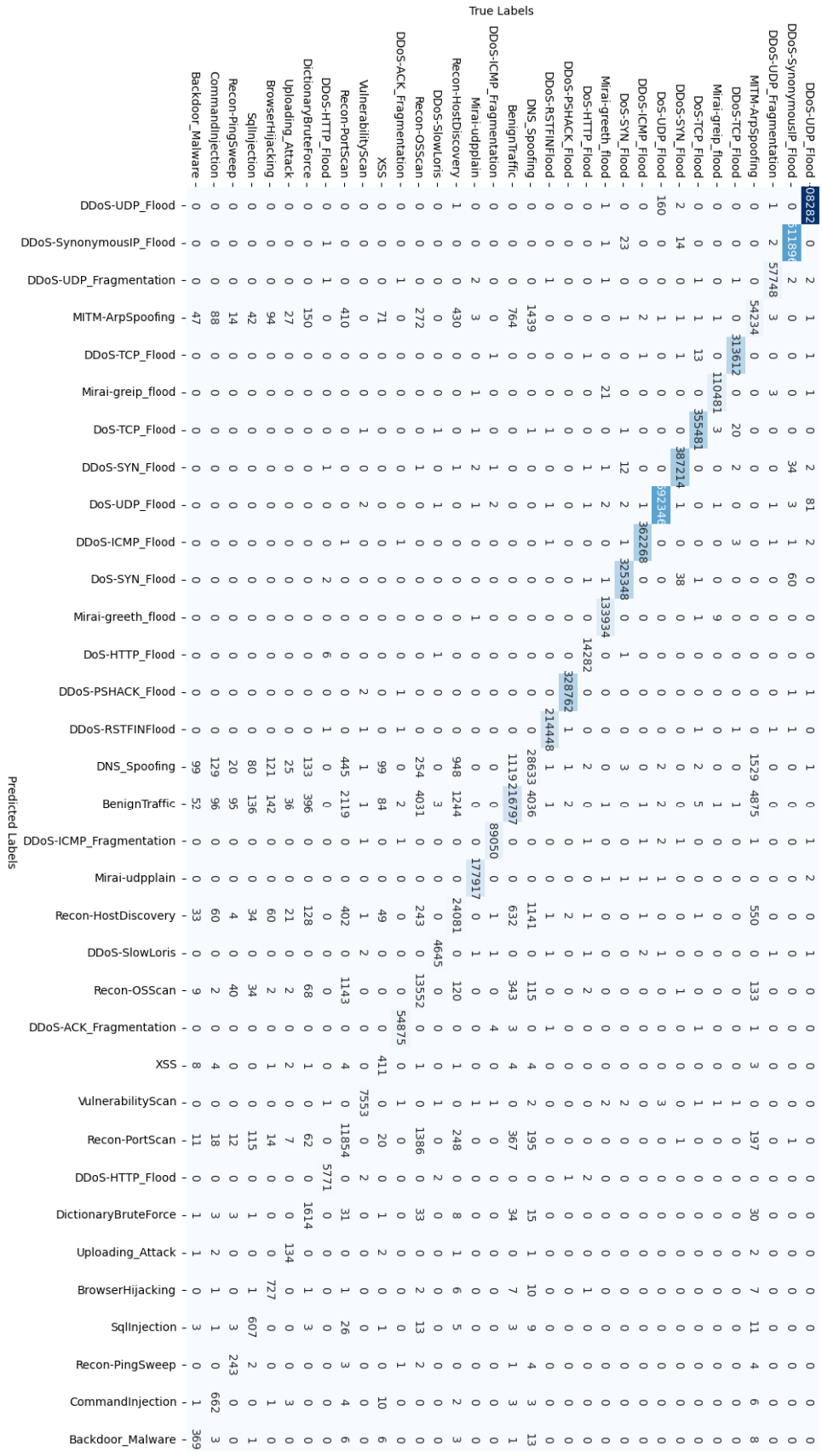

FIGURE 9. HEET thirty-four class confusion matrix.

model confuses these attacks with others, reflecting the intricate challenge of classifying attacks that share similar characteristics.
- The cells corresponding to 'BenignTraffic' illuminate the model's efficacy in distinguishing normal network behavior from malicious activity, despite a few instances of misclassification.

For the HEET Model (Figure 9):

- The matrix for the HEET model shows a clearer, more uniform dark diagonal, implying a general increase in true positive rates and an overall enhancement in classifying attacks correctly.
- Off-diagonal elements, particularly those for 'DNS_Spoofing' and 'CommandInjection', are fewer and fainter compared to the ETC matrix, suggesting the HEET model's improved specificity in detection after hyperparameter optimization through the HSA.
- The 'BenignTraffic' classification is marked by a substantial concentration of true positives, reinforcing the HEET model's proficiency in maintaining low false positive rates while successfully detecting various attacks.

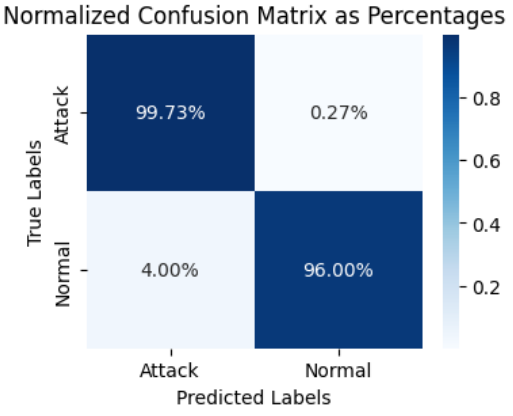

FIGURE 10. ETC binary normalized confusion matrix.

In a more granular analysis, we observe that for complex, sophisticated attacks like 'CommandInjection' and 'Backdoor_Malware', the HEET model demonstrates a high degree of accuracy, whereas the ETC model struggles comparatively, as evidenced by the distribution of values in their respective rows and columns.

For 'Recon-OSScan' and 'DictionaryBruteForce', which are often challenging to classify due to their less pronounced signatures, the HEET model shows a discernible improvement in recall rates, indicating fewer missed detections.

The normalized confusion matrices provide additional clarity by presenting the classification rates as percentages of the total number of instances, allowing for a straightforward comparison of classification performance relative to class size. The normalization helps to highlight the precision in classifying less frequent attacks, such as 'Uploading_Attack' or 'BrowserHijacking', where even a small number of misclassifications can significantly affect the percentage of correct predictions.

The detailed exploration of the confusion matrices illuminates the enhanced diagnostic power of the HEET model over the ETC model, with the HSA proving to be an effective method for hyperparameter optimization. The matrices tell a story not just of improved classification accuracy, but of a more profound understanding of the behavior underlying each attack class. The results reflected in these matrices demonstrate the critical importance of precision and recall in the domain of cybersecurity, where the ability to accurately identify and categorize a broad spectrum of attacks is of paramount importance.

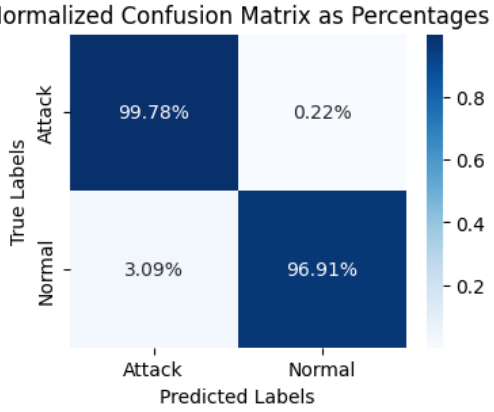

FIGURE 11. HEET binary normalized confusion matrix.

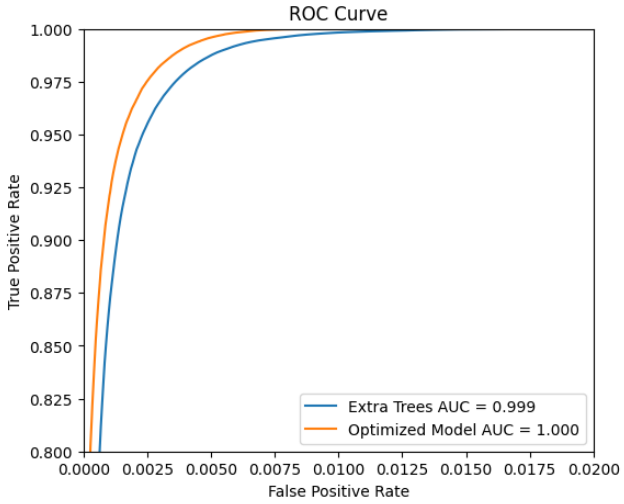

FIGURE 12. ROC curve for ETC and HEET binary test.

V. DISCUSSION

Within the scholarly contours of this section, we engage in a methodical deconstruction of the performance nuances manifested by our models across binary, multi-class, and granular classification tests. This critical examination juxtaposes our empirical findings with the corpus of contemporary literature and introspectively reflects upon the limitations intrinsic to our investigative process.

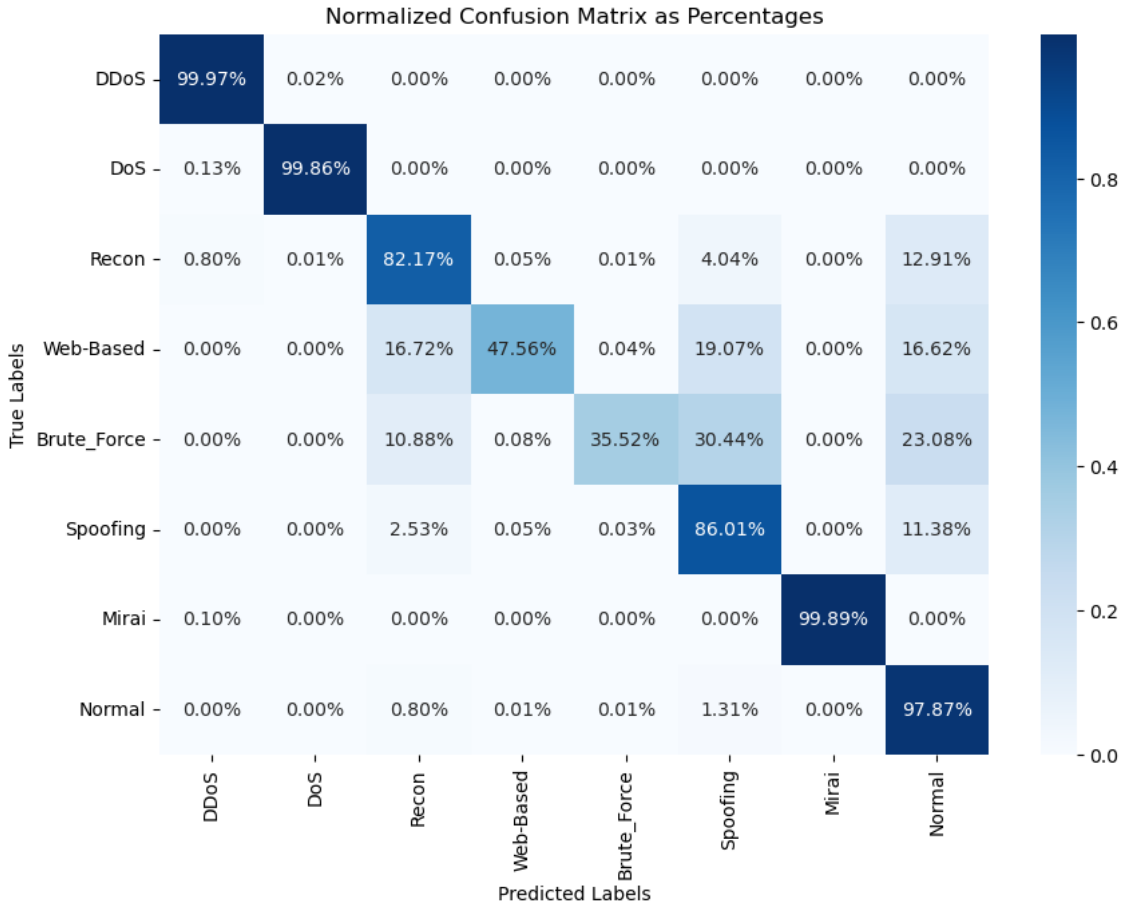

FIGURE 13. ETC octal normalized confusion matrix.

A. BINARY CLASSIFICATION: THE DUALITY OF NETWORK TRAFFIC

The distilled essence of this precision is reflected in the normalized confusion matrices, illustrated in Figures 10 and 11. The matrices for the ETC and the HEET models convey a narrative of exceptional performance, with both models demonstrating exceedingly high true positive rates – 99.73% for ETC and an even more impressive 99.78% for HEET. Yet, the true distinction lies in the equally significant true negative rates, where HEET's 96.91% shines over ETC's 96.00%. This marginal, yet critical, improvement illustrates the nuanced edge that HEET holds over its non-optimized counterpart, emphasizing its enhanced ability to accurately classify normal traffic without succumbing to the trap of overgeneralization.

Coupled with the confusion matrices, the Receiver Operating Characteristic (ROC) curve in Figure 12 serves as a testament to the models' diagnostic prowess. The graceful arc of the HEET model's curve, gravitating towards the upper-left corner, denotes a near-perfect Area Under the Curve (AUC) score of 1.000. In contrast, the ETC model, while laudable in its own right with an AUC of 0.999, acknowledges the fine sliver of space for enhancement that harmony search optimization successfully captures. These curves embody the trade-off between sensitivity and specificity – a balance that the HEET model navigates with astute finesse.

B. MULTI-CLASS CLASSIFICATION: NAVIGATING THE CYBER THREAT PANORAMA

The normalized confusion matrices, as depicted in Figures 13 and 14, offer a perspicuous representation of each model's classification finesse across multiple classes. The ETC model demonstrates commendable acumen, discerning between classes with high precision, as indicated by the dense, darker hues along the diagonal. Yet, it is within the HEET model's matrix that we witness the subtle sophistication afforded by harmony search optimization. With a discernible increase in true positive rates for classes like 'Recon' and 'Web-Based,' alongside an enhancement in true negatives for the 'Normal' traffic, the HEET model stands as a paragon of precision, capable of navigating the complex panorama of cyber threats with aplomb.

Complementing the confusion matrices, the ROC curves for multi-class classification encapsulate the trade-off between sensitivity to true positives and the inevitability of false positives. As illustrated in Figure 15, the curves ascend towards the upper left, a trajectory symbolic of

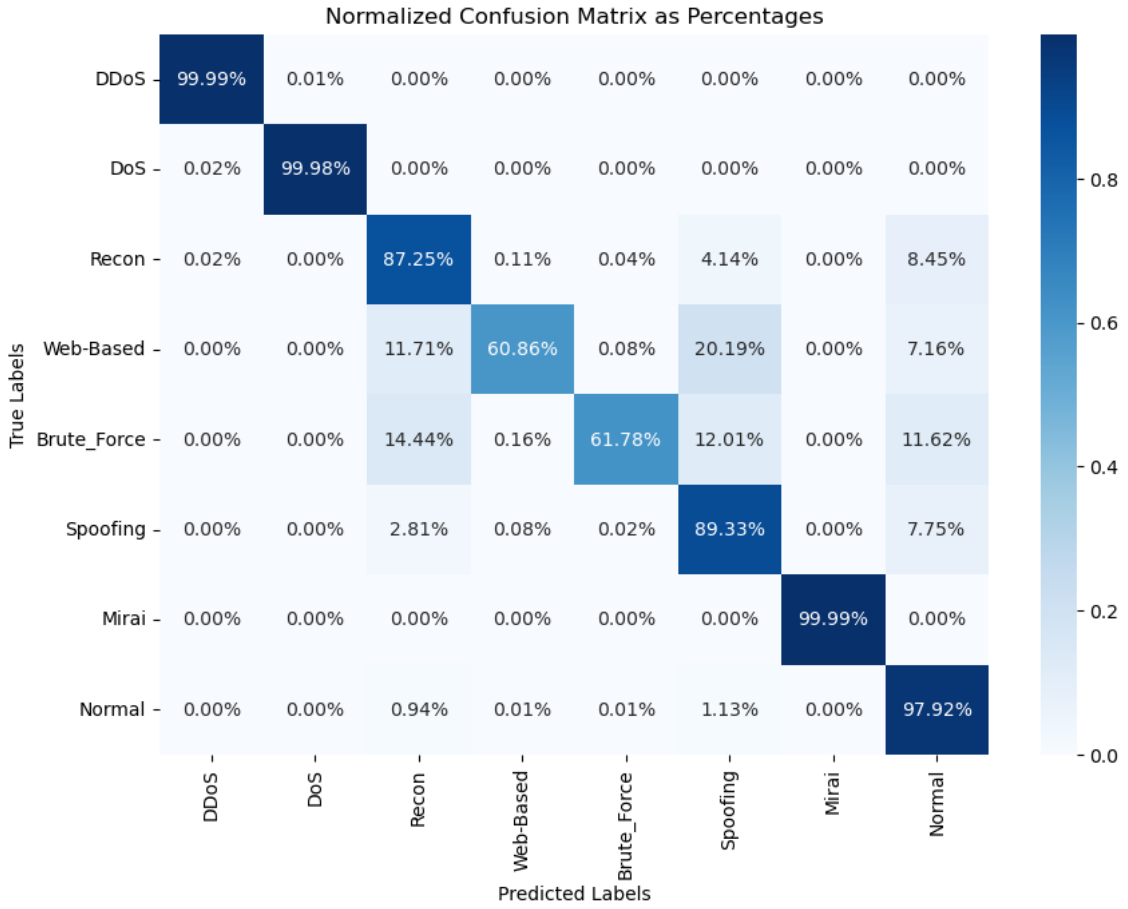

FIGURE 14. HEET octal normalized confusion matrix.

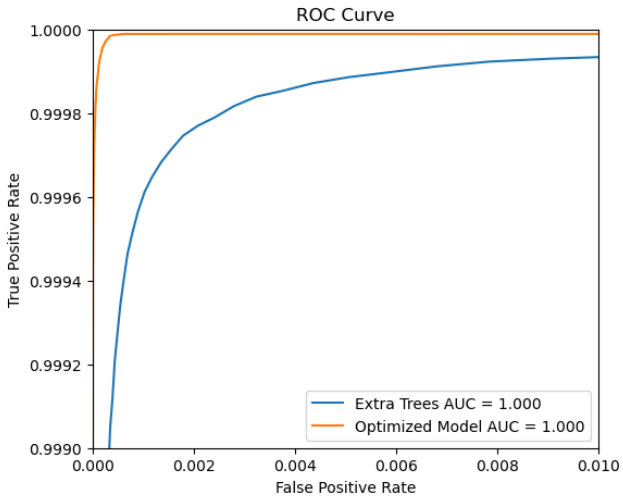

FIGURE 15. ROC curve for ETC and HEET octal test.

excellent classifier performance. The AUC values for both models are indicative of their robust capabilities, with the HEET model, in particular, demonstrating a slightly steeper ascent – a visual echo of its superior performance metrics.

C. GRANULAR CLASSIFICATION: THE INTRICACIES OF ATTACK DETECTION

The normalized confusion matrices provide a comprehensive view of each model's performance in detecting various types of cyber attacks. The ETC shows strong accuracy, particularly highlighted by its high true positive rates for several attack types. However, the HEET model, demonstrates even greater precision and robustness. Key improvements are evident in the detection rates for critical attacks. The following table illustrates the True Positive Rate (TPR) obtained from the normalized confusion matrices for both models (Table 13).

From the table, the HEET model's superior capability is evident in various attack types, demonstrating significant improvements in TPR across a wide range of cyber threats. These enhancements underline the effectiveness of the HEET approach, making it an advanced solution for network intrusion detection in IoT-based smart homes.

The ROC curve in Figure 16 stands as a testament to the HEET model's precision. Mirroring the excellence of the ETC with an Area Under the Curve (AUC) score of 1.000, the HEET model reiterates the potential of harmony search optimization. This congruence of the curves, arching towards the zenith of classifier performance, underscores the

TABLE 13. Comparison of true positive rates between ETC and HEET models.

Attack Type	True Positive Rate (%)	
	ETC	HEET
DDoS-UDP_Flood	99.92	99.99
DDoS-SynonymousIP_Flood	99.91	99.98
DDoS-UDP_Fragmentation	99.12	99.98
MITM-ArpSpoofing	85.02	88.06
DDoS-TCP_Flood	99.77	99.99
Mirai-greip_flood	99.68	99.99
DoS-TCP_Flood	99.88	99.99
DDoS-SYN_Flood	99.75	99.98
DoS-UDP_Flood	99.87	99.97
DDoS-ICMP_Flood	99.84	100
DoS-SYN_Flood	99.83	99.99
Mirai-greeth_flood	99.46	99.98
DoS-HTTP_Flood	97.90	99.90
DDoS-PSHACK_Flood	99.88	100
DDoS-RSTFINFlood	99.78	100
DNS_Spoofing	72.91	80.38
BenignTraffic	98.83	98.51
DDoS-ICMP_Fragmentation	99.43	99.99
Mirai-udpplain	99.92	99.99
Recon-HostDiscovery	84.54	88.86
DDoS-SlowLoris	96.82	99.81
Recon-OSScan	60.44	68.48
DDoS-ACK_Fragmentation	99.07	99.98
XSS	43.37	54.51
VulnerabilityScan	97.46	99.81
Recon-PortScan	59.86	72.07
DDoS-HTTP_Flood	95.56	99.78
DictionaryBruteForce	35.88	63.15
Uploading_Attack	28.79	52.14
BrowserHijacking	31.15	62.56
SqlInjection	37.70	57.64
Recon-PingSweep	17.51	55.99
CommandInjection	56.88	61.93
Backdoor_Malware	53.63	58.20

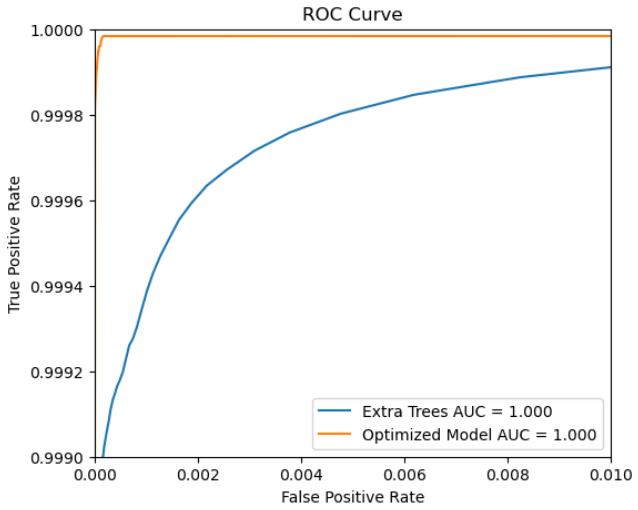

FIGURE 16. ROC curve for ETC and HEET thirty-four class test.

HEET model's exceptional ability to maintain an equilibrium between sensitivity and specificity, even when faced with the formidable task of 34-class differentiation.

Granular classification is the crucible within which the HEET model's robustness is both tested and proven. It is one thing to discern between a binary notion of threat versus non-threat, and another to navigate the intricate web of 34 unique attack classifications. The normalized confusion matrices offer not only a quantitative assessment but also a qualitative glimpse into the HEET model's granular intelligence, providing confidence in its deployment in environments where such detailed discernment is paramount.

D. COMPARATIVE PERFORMANCE: STANDING AMONG PEERS

In the scholarly discourse of this subsection, we engage in a meticulous evaluation of our HEET model against the backdrop of current research. This comparative analysis is pivotal, for it situates our achievements within the continuum of cyber defense innovations and underscores the HEET model's standing in the academic fraternity.

Table 14 serves as the compass guiding our comparison. It encapsulates a constellation of methodologies, each with unique hyperparameter optimization strategies and performance metrics. Our model is juxtaposed with a tapestry of contemporary methods, ranging from RF to deep neural networks (DNN), each applied to datasets of varying scope and granularity. This rich tableau of approaches offers a multifaceted perspective on model performance in cybersecurity.

Amidst this diverse landscape, the HEET model, harmonized through harmony search optimization and tested across a full dataset, demonstrates its prowess. In binary classification, our model achieves an impressive 99.87% accuracy, outperforming the RF benchmark set by Neto et al. [46]. Moreover, in the nuanced realm of 34-class categorization, our model's precision and f1-score outshine those of the same RF approach. The HEET model stands as a testament to the precision that harmony search optimization imparts, particularly when confronting a spectrum of cyber threats.

Our model's performance in 8-class categorization also warrants particular attention. With a notable f1-score of 90.65%, it surpasses the RF's performance and provides a compelling alternative to the blending approach by Le et al. [51], which, while high in accuracy, operates on a smaller dataset. This distinction is crucial, as the capacity to maintain high performance metrics across extensive datasets is emblematic of a model's robustness and scalability.

The superior performance of our HEET model, even when applied to the full dataset, can be attributed to several key factors. First, our approach incorporates meticulous data processing techniques, ensuring that the input data is of the highest quality and free from noise and inconsistencies that could otherwise degrade model performance. Second, the ETC employed within our model demonstrates excellent classification capabilities, benefiting from the ensemble learning approach which leverages multiple de-correlated decision trees to enhance prediction accuracy. Finally, the extensive hyperparameter optimization method, facilitated by the HSA, plays a crucial role. By systematically tuning the hyperparameters, the HSA ensures that the ETC operates at

TABLE 14. Comparative performance metrics.

Authors	Year	Method	Hyperparameter Optimization	Dataset	Class	Accuracy	Precision	Recall	F1-score
Neto et al. [46]	2023	RF	-	full	2 8 34	99.68 99.44 99.16	96.54 70.54 70.45	96.52 91.00 83.16	96.53 71.93 71.40
Wang et al. [50]	2023	DL-BiLSTM	Optuna	small	8	93.13	91.80	93.13	91.94
Le et al. [51]	2023	Blending	-	small	8	99.51	98.51	99.63	99.07
Ebuka et al. [52]	2024	DNN	-	small	2 34	99.10 70.53	99.10 70.00	- -	99.10 63.00
Khan et al. [53]	2024	RF	-	small	2 8 34	99.56 95.55 96.33	99.56 95.56 96.28	99.56 95.55 96.33	99.56 95.52 96.26
Abbas et al. [54]	2024	RNN1	Grid Search	small	8	98.61	98.55	98.61	98.57
Abbas et al. [55]	2024	RNN	-	small	34	96.52	96.25	96.52	95.73
Roshan et al. [56]	2024	EnsAdp_CIDS	-	small	2	98.93	99.50	99.40	99.45
Aswani et al. [57]	2024	mGRU	-	small	2	98.48	98.54	98.15	98.57
Our model	2024	HEET	HSA	full	2 8 34	99.87 99.51 99.49	97.41 96.02 96.07	98.45 87.14 87.10	97.92 90.65 90.61

its optimal configuration, thereby maximizing its efficacy in detecting a wide range of intrusion types. These combined efforts result in a robust and highly effective intrusion detection system that outperforms many contemporary models in the field.

The dialogue within this subsection is not confined to a mere comparison of metrics. It is, rather, a reflective contemplation of our model's place within the academic narrative. Each method in Table 14 contributes uniquely to the collective pursuit of cyber resilience. The HEET model's narrative is one of holistic performance and innovative optimization, contributing a new chapter to the story of ML in cyber defense.

In conclusion, our comparative analyses are not merely an exposition of the HEET model's empirical triumphs. It is an academic homage to the continuous quest for precision, accuracy, and reliability in cybersecurity. Our discussion is conducted with a graceful acknowledgment of our model's strengths and a humble appreciation for the collective advancements within the field. Through this comparison, we reiterate our commitment to pushing the boundaries of cyber defense technologies, with the HEET model paving the way for future exploration and development.

E. REFLECTIONS ON LIMITATIONS: CHARTING THE PATH FORWARD

At the heart of our model's success is the esoteric art of hyperparameter optimization, a process where the harmonious tuning of parameters can dramatically elevate a model's performance. This fine-tuning, however, necessitates a depth of specialist knowledge, often derived from an intimate familiarity with both the model architecture and the HSA. The requirement for such expertise is a double-edged sword – it ensures that our model is optimized with precision, yet it places a barrier to entry that may preclude broader engagement and iterative experimentation by the wider community.

The instrumentation required to perform this computational concerto is not insignificant. A powerful microprocessor and substantial RAM are the conductors of this orchestra, orchestrating the symphony of data through the model's neural pathways. The absence of such computational resources could stymie the HEET model's training, presenting a challenge to its adoption in environments that are resource-constrained or where computational democratization is a guiding principle.

Training time stands as one of the most tangible limitations of our approach. The intricate tapestry of ML is one that is woven over time, with each thread of data interlacing under the guidance of the algorithm. The span of time required to train our model is invariably linked to the device's capabilities, creating a temporal limitation that can impact the model's applicability in scenarios demanding rapid deployment.

In charting the path forward, these limitations are not insurmountable obstacles but waypoints guiding our journey of iterative enhancement. The need for specialist knowledge catalyzes a push towards more intuitive, user-friendly

optimization processes. The quest for powerful computational resources fuels innovation in algorithmic efficiency, driving the development of models that balance performance with accessibility. The time-intensive nature of training impels us toward innovations in expediting learning, perhaps through transfer learning [58] or more sophisticated parallel processing techniques [59].

VI. CONCLUSION

In the realm of cybersecurity, where the digital and physical converge, the sanctity of our smart homes hinges on the vigilance of network IDS. Our study heralds a significant stride in this direction. We have meticulously designed a network IDS deployment scheme specifically for the IoT-based smart home environment, which is not only cognizant of the conventional threats but is also astute in recognizing the potential of home devices being commandeered as instruments of attack.

The crux of our contribution lies in the bespoke optimization of the ETC's hyperparameters. Employing the HSA, we have fine-tuned the sinews of this model – n_estimators, max_depth, min_samples_split, min_samples_leaf, max_features, and bootstrap – to orchestrate a system attuned to the subtleties of network traffic within the IoT milieu. This fine-tuning is akin to crafting an exquisite lock, responsive to the slightest tamperings yet seamless in its daily function.

Our experimental rigor extended to the utilization of a comprehensive, real-time dataset amassed from 105 home IoT devices – a dataset unique not only in its scope but in its reflection of an emerging reality where IoT devices themselves may be conscripted into the ranks of cyber threats. This inclusion presents a prescient understanding of the evolving threat landscape, positioning our research at the forefront of practical, applied cybersecurity.

In our commitment to the democratization of knowledge and the advancement of collective cyber defense capabilities, we have made our code available in an open GitHub repository. This gesture of academic generosity paves the way for further innovation, allowing peers and practitioners to build upon our foundations.

The empirical heart of our study lies in the rigorous testing across binary, 8-class, and 34-class configurations, critically evaluated through a suite of metrics including Accuracy, Precision, Recall, and F1-score. Accompanying these were the nuanced insights provided by classification reports, confusion matrices, their normalized counterparts expressed as percentages, and ROC Curves. These instruments of analysis have not merely quantified the model's performance but have illustrated the depth and breadth of its understanding.

In conclusion, our research presents a confluence of innovation, experimentation, and application. The network IDS scheme we propose is more than a theoretical construct; it is a testament to what is achievable when ingenuity meets practicality. It stands as an edifice in the digital landscape, a beacon of security for the smart homes that cradle our modern lives. With the detailed results we provide, we extend an invitation to the academic and professional communities to partake in this journey, to critique, to enhance, and to ultimately forge new paths forward in the collective quest for cybersecurity in an IoT-integrated world.

REFERENCES

[1] (2023). *2023 Was a Big Year for Cybercrime—Here's How We Can Make Our Systems Safer*. Accessed: Jan. 10, 2024. [Online]. Available: https://www.weforum.org/agenda/2024/01/cybersecurity-cybercrime-system-safety/

[2] N. M. Allifah and I. A. Zualkernan, "Ranking security of IoT-based smart home consumer devices," *IEEE Access*, vol. 10, pp. 18352–18369, 2022.

[3] R. Alasmari and A. Abdullah Alhogail, "Protecting smart-home IoT devices from MQTT attacks: An empirical study of ML-based IDS," *IEEE Access*, vol. 12, pp. 25993–26004, 2024.

[4] O. A. Abraham, H. Ochiai, M. D. Hossain, Y. Taenaka, and Y. Kadobayashi, "Electricity theft detection for smart homes: Harnessing the power of machine learning with real and synthetic attacks," *IEEE Access*, vol. 12, pp. 26023–26045, 2024.

[5] P. Illy, G. Kaddoum, K. Kaur, and S. Garg, "ML-based IDPS enhancement with complementary features for home IoT networks," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 2, pp. 772–783, Jun. 2022.

[6] E. Anthi, L. Williams, M. Slowinska, G. Theodorakopoulos, and P. Burnap, "A supervised intrusion detection system for smart home IoT devices," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 9042–9053, Oct. 2019.

[7] S. Wah Tay, N. Zhang, and S. AlJanah, "A problem analysis of smart home automation: Toward secure and usable communication-based authorization," *IEEE Access*, vol. 12, pp. 18103–18121, 2024.

[8] R. Heartfield, G. Loukas, A. Bezemskij, and E. Panaousis, "Self-configurable cyber-physical intrusion detection for smart homes using reinforcement learning," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1720–1735, 2021.

[9] L. Almuqren, K. Mahmood, S. S. Aljameel, A. S. Salama, G. P. Mohammed, and A. A. Alneil, "Blockchain-assisted secure smart home network using gradient-based optimizer with hybrid deep learning model," *IEEE Access*, vol. 11, pp. 86999–87008, 2023.

[10] D. Rani, N. S. Gill, P. Gulia, F. Arena, and G. Pau, "Design of an intrusion detection model for IoT-enabled smart home," *IEEE Access*, vol. 11, pp. 52509–52526, 2023.

[11] M. S. Abdalzaher, M. M. Fouda, H. A. Elsayed, and M. M. Salim, "Toward secured IoT-based smart systems using machine learning," *IEEE Access*, vol. 11, pp. 20827–20841, 2023.

[12] H. Liao, M. Z. Murah, M. K. Hasan, A. H. M. Aman, J. Fang, X. Hu, and A. U. R. Khan, "A survey of deep learning technologies for intrusion detection in Internet of Things," *IEEE Access*, vol. 12, pp. 4745–4761, 2024.

[13] B. S. Ali, I. Ullah, T. Al Shloul, I. A. Khan, I. Khan, Y. Y. Ghadi, A. Abdusalomov, R. Nasimov, K. Ouahada, and H. Hamam, "ICS-IDS: Application of big data analysis in AI-based intrusion detection systems to identify cyberattacks in ICS networks," *J. Supercomput.*, vol. 80, no. 6, pp. 7876–7905, Apr. 2024, doi: 10.1007/s11227-023-05764-5.

[14] V. Graveto, T. Cruz, and P. Simões, "A network intrusion detection system for building automation and control systems," *IEEE Access*, vol. 11, pp. 7968–7983, 2023.

[15] M. S. Essa and S. K. Guirguis, "Evaluation of tree-based machine learning algorithms for network intrusion detection in the Internet of Things," *IT Prof.*, vol. 25, no. 5, pp. 45–56, Sep. 2023.

[16] E. Ozdogan, "A comprehensive analysis of the machine learning algorithms in IoT IDS systems," *IEEE Access*, vol. 12, pp. 46785–46811, 2024.

[17] H. Sadia, S. Farhan, Y. Ul Haq, R. Sana, T. Mahmood, S. A. O. Bahaj, and A. R. Khan, "Intrusion detection system for wireless sensor networks: A machine learning based approach," *IEEE Access*, vol. 12, pp. 52565–52582, 2024.

[18] K. Zhang, J. Yang, Y. Shao, L. Hu, W. Ou, W. Han, and Q. Zhang, "Intrusion detection model for Internet of Vehicles using GRIPCA and OWELM," *IEEE Access*, vol. 12, pp. 28911–28925, 2024.

[19] D. Kilichev and W. Kim, "Hyperparameter optimization for 1D-CNN-based network intrusion detection using GA and PSO," *Mathematics*, vol. 11, no. 17, p. 3724, Aug. 2023. [Online]. Available: https://www.mdpi.com/2227-7390/11/17/3724

[20] A. Abdellatif, H. Abdellatef, J. Kanesan, C.-O. Chow, J. H. Chuah, and H. M. Gheni, "An effective heart disease detection and severity level classification model using machine learning and hyperparameter optimization methods," *IEEE Access*, vol. 10, pp. 79974–79985, 2022.

[21] F. Y. Assiri and M. Ragab, "Optimal deep-learning-based cyber-attack detection in a blockchain-assisted IoT environment," *Mathematics*, vol. 11, no. 19, p. 4080, Sep. 2023. [Online]. Available: https://www.mdpi.com/2227-7390/11/19/4080

[22] A. Zainab, S. S. Refaat, and O. Bouhali, "Ensemble-based spam detection in smart home IoT devices time series data using machine learning techniques," *Information*, vol. 11, no. 7, p. 344, Jul. 2020. [Online]. Available: https://www.mdpi.com/2078-2489/11/7/344

[23] T. Li, Z. Hong, and L. Yu, "Machine learning-based intrusion detection for IoT devices in smart home," in *Proc. IEEE 16th Int. Conf. Control Autom. (ICCA)*, Oct. 2020, pp. 277–282. [Online]. Available: https://ieeexplore.ieee.org/document/9264406

[24] F. Alghayadh and D. Debnath, "A hybrid intrusion detection system for smart home security," in *Proc. IEEE Int. Conf. Electro Inf. Technol. (EIT)*, Jul. 2020, pp. 319–323. [Online]. Available: https://ieeexplore.ieee.org/document/9208296

[25] L. Shi, L. Wu, and Z. Guan, "Three-layer hybrid intrusion detection model for smart home malicious attacks," *Comput. Electr. Eng.*, vol. 96, Dec. 2021, Art. no. 107536. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S004579062100481X

[26] S. Sohail, Z. Fan, X. Gu, and F. Sabrina, "Multi-tiered artificial neural networks model for intrusion detection in smart homes," *Intell. Syst. With Appl.*, vol. 16, Nov. 2022, Art. no. 200152. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2667305322000898

[27] Y. Chen, J. Wang, T. Yang, Q. Li, and N. A. Nijhum, "An enhancement method in few-shot scenarios for intrusion detection in smart home environments," *Electronics*, vol. 12, no. 15, p. 3304, Jul. 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/15/3304

[28] M. Wang, N. Yang, and N. Weng, "Securing a smart home with a transformer-based IoT intrusion detection system," *Electronics*, vol. 12, no. 9, p. 2100, May 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/9/2100

[29] D. Yuan, K. Ota, M. Dong, X. Zhu, T. Wu, L. Zhang, and J. Ma, "Intrusion detection for smart home security based on data augmentation with edge computing," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/9148632

[30] S. I. Popoola, B. Adebisi, M. Hammoudeh, H. Gacanin, and G. Gui, "Stacked recurrent neural network for botnet detection in smart homes," *Comput. Electr. Eng.*, vol. 92, Jun. 2021, Art. no. 107039. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790621000598

[31] N. Elsayed, Z. S. Zaghloul, S. W. Azumah, and C. Li, "Intrusion detection system in smart home network using bidirectional LSTM and convolutional neural networks hybrid model," in *Proc. IEEE Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2021, pp. 55–58. [Online]. Available: https://ieeexplore.ieee.org/document/9531683

[32] S. W. Azumah, N. Elsayed, V. Adewopo, Z. S. Zaghloul, and C. Li, "A deep LSTM based approach for intrusion detection IoT devices network in smart home," in *Proc. IEEE 7th World Forum Internet Things (WF-IoT)*, Jun. 2021, pp. 836–841. [Online]. Available: https://ieeexplore.ieee.org/document/9596033

[33] N. Butt, A. Shahid, K. N. Qureshi, S. Haider, A. O. Ibrahim, F. Binzagr, and N. Arshad, "Intelligent deep learning for anomaly-based intrusion detection in IoT smart home networks," *Mathematics*, vol. 10, no. 23, p. 4598, Dec. 2022. [Online]. Available: https://www.mdpi.com/2227-7390/10/23/4598

[34] S. Hizal, Ü. Çavuşoğlu, and D. Akgün, "IoT-based smart home security system with machine learning models," *Academic Platform J. Eng. Smart Syst.*, vol. 12, no. 1, pp. 28–36, Jan. 2024.

[35] Y. A. Alsariera, V. E. Adeyemo, A. O. Balogun, and A. K. Alazzawi, "AI meta-learners and extra-trees algorithm for the detection of phishing websites," *IEEE Access*, vol. 8, pp. 142532–142542, 2020.

[36] Y. Du, Y. Liu, Y. Yan, J. Fang, and X. Jiang, "Risk management of weather-related failures in distribution systems based on interpretable extra-trees," *J. Modern Power Syst. Clean Energy*, vol. 11, no. 6, pp. 1868–1877, 2023.

[37] S. M. Mastelini, F. K. Nakano, C. Vens, and A. C. P. d. L. F. d. Carvalho, "Online extra trees regressor," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6755–6767, Sep. 2023.

[38] S. Sadhwani, U. K. Modi, R. Muthalagu, and P. M. Pawar, "SmartSentry: Cyber threat intelligence in industrial IoT," *IEEE Access*, vol. 12, pp. 34720–34740, 2024.

[39] Z. Woo Geem, J. Hoon Kim, and G. V. Loganathan, "A new heuristic optimization algorithm: Harmony search," *Simulation*, vol. 76, no. 2, pp. 60–68, Feb. 2001. [Online]. Available: https://journals.sagepub.com/doi/10.1177/003754970107600201

[40] J. Wang, H. Ouyang, Z. Zhou, and S. Li, "Harmony search algorithm based on dual-memory dynamic search and its application on data clustering," *Complex Syst. Model. Simul.*, vol. 3, no. 4, pp. 261–281, Dec. 2023.

[41] A. A. Al-Omoush, A. A. Alsewari, H. S. Alamri, and K. Z. Zamli, "Comprehensive review of the development of the harmony search algorithm and its applications," *IEEE Access*, vol. 7, pp. 14233–14245, 2019.

[42] L. Flores-Pulido, E. A. Portilla-Flores, E. Santiago-Valentín, E. Vega-Alvarado, M. B. C. Yáñez, and P. A. Niño-Suárez, "A comparative study of improved harmony search algorithm in four bar mechanisms," *IEEE Access*, vol. 8, pp. 148757–148778, 2020.

[43] A. A. Alomoush, A. A. Alsewari, H. S. Alamri, K. Aloufi, and K. Z. Zamli, "Hybrid harmony search algorithm with grey wolf optimizer and modified opposition-based learning," *IEEE Access*, vol. 7, pp. 68764–68785, 2019.

[44] H. Kim, S. Hong, A. G. Limos, Z. W. Geem, and J. Yoon, "Improving water quality modelling for green roof runoff using storm water management model," *Urban Climate*, vol. 52, Nov. 2023, Art. no. 101717. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212095523003115

[45] T. Zhang and Z. W. Geem, "Review of harmony search with respect to algorithm structure," *Swarm Evol. Comput.*, vol. 48, pp. 31–43, Aug. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210650218303791

[46] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, and A. A. Ghorbani, "CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment," *Sensors*, vol. 23, no. 13, p. 5941, Jun. 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/13/5941

[47] D. Kilichev, D. Turimov, and W. Kim, "Next–Generation intrusion detection for IoT EVCS: Integrating CNN, LSTM, and GRU models," *Mathematics*, vol. 12, no. 4, p. 571, Feb. 2024. [Online]. Available: https://www.mdpi.com/2227-7390/12/4/571

[48] A. Ergasheva, F. Akhmedov, A. Abdusalomov, and W. Kim, "Advancing maritime safety: Early detection of ship fires through computer vision, deep learning approaches, and histogram equalization techniques," *Fire*, vol. 7, no. 3, p. 84, Mar. 2024. [Online]. Available: https://www.mdpi.com/2571-6255/7/3/84

[49] A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, "Brain tumor detection based on deep learning approaches and magnetic resonance imaging," *Cancers*, vol. 15, no. 16, p. 4172, Aug. 2023. [Online]. Available: https://www.mdpi.com/2072-6694/15/16/4172

[50] Z. Wang, H. Chen, S. Yang, X. Luo, D. Li, and J. Wang, "A lightweight intrusion detection method for IoT based on deep learning and dynamic quantization," *PeerJ Comput. Sci.*, vol. 9, Sep. 2023, Art. no. e1569. [Online]. Available: https://peerj.com/articles/cs-1569/

[51] T.-T.-T. Le, R. W. Wardhani, D. S. C. Putranto, U. Jo, and H. Kim, "Toward enhanced attack detection and explanation in intrusion detection system-based IoT environment data," *IEEE Access*, vol. 11, pp. 131661–131676, 2023.

[52] E. C. Nkoro, C. I. Nwakanma, J.-M. Lee, and D.-S. Kim, "Detecting cyberthreats in metaverse learning platforms using an explainable DNN," *Internet Things*, vol. 25, Apr. 2024, Art. no. 101046. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660523003694

[53] M. M. Khan and M. Alkhathami, "Anomaly detection in IoT-based healthcare: Machine learning for enhanced security," *Sci. Rep.*, vol. 14, no. 1, p. 5872, Mar. 2024. [Online]. Available: https://www.nature.com/articles/s41598-024-56126-x

[54] S. Abbas, S. Alsubai, S. Ojo, G. A. Sampedro, A. Almadhor, A. A. Hejaili, and I. Bouazzi, "An efficient deep recurrent neural network for detection of cyberattacks in realistic IoT environment," *J. Supercomput.*, vol. 80, no. 10, pp. 13557–13575, Jul. 2024, doi: 10.1007/s11227-024-05993-2.

[55] S. Abbas, I. Bouazzi, S. Ojo, A. Al Hejaili, G. A. Sampedro, A. Almadhor, and M. Gregus, "Evaluating deep learning variants for cyber-attacks detection and multi-class classification in IoT networks," *PeerJ Comput. Sci.*, vol. 10, Jan. 2024, Art. no. e1793.

[56] K. Roshan and A. Zafar, "Ensemble adaptive online machine learning in data stream: A case study in cyber intrusion detection system," *Int. J. Inf. Technol.*, pp. 1–14, Feb. 2024, doi: 10.1007/s41870-024-01727-y.
[57] A. D. Aguru and S. B. Erukala, "A lightweight multi-vector DDoS detection framework for IoT-enabled mobile health informatics systems using deep learning," *Inf. Sci.*, vol. 662, Mar. 2024, Art. no. 120209. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025524001221
[58] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proc. IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021.
[59] Y. Zhang, T. Cao, S. Li, X. Tian, L. Yuan, H. Jia, and A. V. Vasilakos, "Parallel processing systems for big data: A survey," *Proc. IEEE*, vol. 104, no. 11, pp. 2114–2136, Nov. 2016.

AKMALBEK ABDUSALOMOV received the B.S. degree in informatics and information technologies from Tashkent University of Information Technologies, Tashkent, Uzbekistan, in 2015, and the M.S. and Ph.D. degrees in IT convergence engineering from Gachon University, South Korea, in 2017 and 2022, respectively. He is currently an Assistant Professor with the Computer Engineering Department, Gachon University. His research interests include object detection, artificial intelligence (DL and ML), and medical image processing.

DUSMUROD KILICHEV received the B.S. degree in informatics and information technologies from Tashkent University of Information Technologies, Samarkand, Uzbekistan, in 2014, and the M.S. degree in computer engineering (information security, cryptography, and cryptanalysis) from Tashkent University of Information Technologies, Tashkent, Uzbekistan, in 2016. He is currently pursuing the Ph.D. degree with the Computer Engineering Department, Gachon University, Seongnam, South Korea. His current research interests include cybersecurity, intrusion detection systems, machine learning algorithms, and deep learning.

RASHID NASIMOV received the B.S. degree in telecommunication technologies and the M.S. and Ph.D. degrees in computer engineering from Tashkent University of Information Technologies, Tashkent, Uzbekistan, in 2008, 2011, and 2021, respectively. He is currently an Assistant Professor with the Information Systems and Technologies Department, Tashkent State University of Economics, Uzbekistan. His research interest includes medical image processing using artificial intelligent algorithms (DL and ML).

ILKHOM RAKHMATULLAYEV received the B.S. and M.S. degrees in informatics and information technologies from Tashkent University of Information Technologies, Samarkand, Uzbekistan, in 2017, and the Ph.D. degree in information security from Tashkent University of Information Technologies, Tashkent, Uzbekistan, in 2024. His current research interests include cybersecurity and intrusion detection systems.

YOUNG IM CHO was a Visiting Professor with Purdue University, USA, from 2013 to 2014. She has been the Chairperson of the AI and Smart City Laboratory, Gachon University, since 2016. She has been the Chairperson of Korea Intelligent System Society, since January 2018; a member of the Smart City Special Committee of Korea President, since November 2017; a member of the e-Government Driving Forum, since March 2015; and a Board Member of the National Information Society (Public). Since January 2014, she has been a Board Member of Korea Local Information Research and Development Institute (Public); a Committee Member of the Public Open Data Committee at the Ministry of Interior in Korea; the e-Government Committee Member of the Ministry of Interior in Korea; a Committee Member of the Intelligent Society at Ministry of Science, ICT and Future Planning, in May 2016; and a Committee Member of e-Government at the Ministry of Interior, in April 2016. She has been a Committee Member of the National ICT Strategies in South Korea, the Government 3.0 Committee Member of the Ministry of Interior (South Korea), and the Chairperson of Smart City Forum in South Korea, since 2011.