Enhanced IoT Intrusion Detection: a Hybrid Framework Integrating Decision Tree and One-Class SVM

Sarabjot Singh, Osama A. Mahdi, Savitri Bevinakoppa and Ammar Alazab

March 7, 2025

Enhanced IoT Intrusion Detection: A Hybrid Framework Integrating Decision Tree and One-Class SVM

Sarabjot Singh
School of Information Technology and Engineering
Melbourne Institute of Technology
Melbourne, Australia
sarabjotsingh@academic.mit.edu.au

Osama A. Mahdi
School of Information our and Engineering
Melbourne Institute of Technology
Melbourne, Australia
omahdi@mit.edu.au

Savitri Bevinakoppa
School of Information Technology and Engineering
Melbourne Institute of Technology
Melbourne, Australia
sbevinakoppa@mit.edu.au

Ammar Alazab
Centre for Artificial Intelligence and Optimization
Torrens University
Melbourne, Australia
ammar.alazab@torrens.edu.au

***Abstract*— With the escalating sophistication of cyber-attacks, there is a pressing need for efficient intrusion detection mechanisms in the context of the Internet of Things (IoT). These mechanisms are crucial for monitoring computer resources and generating reports on suspicious or anomalous activities. Conventional intrusion detection systems (IDS) typically rely on a single classifier for intrusion identification, which often struggles to achieve high accuracy and low false alarm rates. This challenge is amplified by the polymorphic, metamorphic, and zero-day behaviors exhibited by malware. To overcome these limitations, this research proposes a hybrid IDS (HIDS) tailored for the IoT environment. The HIDS integrates the power of a Decision Tree classifier and a One-Class Support Vector Machine to establish a robust and effective intrusion detection framework. By synergistically leveraging the strengths of Signature-based Intrusion Detection Systems (SIDS) and Anomaly-based Intrusion Detection Systems (AIDS), the HIDS aims to proficiently identify both known intrusions and zero-day attacks, achieving enhanced detection accuracy while minimizing false alarm rates. To validate the efficacy of the proposed HIDS, benchmark datasets such as NSL-KDD and ADFA are employed for evaluation. Experimental results demonstrate that HIDS surpasses the performance of traditional SIDS in terms of detection rate and false alarm rates, thereby elevating the overall effectiveness of intrusion detection in the IoT landscape.**

Keywords—Attacks, IoT, Intrusion Detection System, AIDS, Machine Learning

I. Introduction

Zero-day intrusion detection poses a formidable challenge due to the rising number of newly identified intrusions, resulting in increasingly severe consequences. This underscores the need for efficient intrusion detection mechanisms capable of closely monitoring computer resources and promptly reporting suspicious or anomalous activities. Traditional intrusion detection systems (IDS) typically rely on a single classifier for intrusion identification [1], yet they frequently encounter difficulties in achieving high accuracy and low false alarm rates. This is primarily due to the polymorphic, metamorphic, and zero-day behaviors exhibited by malware [2] [3].

To address these challenges, this research introduces a novel approach: a Hybrid Intrusion Detection System (HIDS) that synergistically merges the Decision Tree classifier with a One-Class Support Vector Machine (SVM). By integrating Signature-based Intrusion Detection Systems (SIDS) and Anomaly-based Intrusion Detection Systems (AIDS), the HIDS aims to enhance detection accuracy and minimize false alarms. This hybrid approach not only strengthens the system's capability to identify known intrusions but also enhances its ability to detect sophisticated zero-day attacks, thereby improving overall system reliability and security posture.

An Intrusion Detection System (IDS) plays a critical role in cybersecurity by detecting malicious activities that bypass traditional firewall defences. IDSs employ two primary methodologies: Signature-based Intrusion Detection Systems (SIDS) and Anomaly-based Intrusion Detection Systems (AIDS). SIDS operate by comparing incoming data against a database of known attack signatures, offering high accuracy in detecting familiar threats but struggling with polymorphic malware and the management of large signature databases. In contrast, AIDS establishes a baseline of normal system behaviour and trigger alerts for deviations indicative of potential intrusions, making them effective against novel attacks but prone to false positives. Integrating both SIDS and AIDS into a Hybrid IDS framework enhances detection capabilities, combining the specificity of signature-based detection with the adaptability of anomaly-based detection to better defend against a wide array of cyber threats, including sophisticated zero-day attacks [4].

Anomaly-based Intrusion Detection Systems (AIDS) address the shortcomings of Signature-based Intrusion Detection Systems (SIDS) by constructing statistical models that define typical user behaviour. These models enable AIDS to identify deviations from established patterns as potential intrusions, effectively detecting zero-day attacks that evade signature-based detection. Unlike SIDS, which depend on static signature databases, AIDS offer several advantages: they can uncover internal threats by recognizing abnormal activities within a network, and they pose a challenge to cybercriminals attempting to mimic legitimate user behaviour without triggering alerts [2] [3]. This capability not only enhances overall security posture by detecting previously unseen threats but also contributes to a more proactive approach in identifying and mitigating potential security breaches before significant damage occurs.

To overcome the inherent limitations of traditional IDSs, a hybrid IDS model is proposed that integrates both Signature-based Intrusion Detection Systems (SIDS) and Anomaly-based Intrusion Detection Systems (AIDS). In this study, the SIDS component is enhanced using the C5.0 Decision Tree classifier, leveraging its ability to accurately match incoming data against a repository of known attack signatures. Concurrently, the AIDS component employs a one-class Support Vector Machine (SVM) to establish a baseline of normal system behaviour and detect deviations indicative of potential intrusions, including zero-day attacks that evade signature-based detection.

By integrating these complementary methodologies, the resulting Hybrid IDS (HIDS) aims to achieve robust intrusion detection capabilities. The decision tree classifier enhances the system's ability to swiftly identify known threats with minimal false positives, while the SVM-based anomaly detection enhances sensitivity to subtle deviations in system behaviour, crucial for detecting novel and sophisticated attacks. This hybrid approach not only enhances overall detection accuracy but also strengthens the IDS's resilience against evolving cyber threats, contributing to a more proactive and effective defence strategy in cybersecurity operations.

The contributions of this research include:

- Proposing an advanced framework capable of accurately identifying both known intrusions and zero-day attacks while maintaining high detection accuracy and minimizing false alarm rates.
- Integrating the complementary strengths of SIDS and AIDS within a hybrid IDS to enhance overall detection capabilities.
- Performing a comparative study of the HIDS using benchmark datasets (such as NSL-KDD and ADFA) to substantiate its superior performance in terms of accuracy and F-measure

The paper is organized as follows: Section 2 provides an overview of related hybrid intrusion detection methods. Section 3 presents a detailed description of the proposed hybrid intrusion detection approach. Section 4 outlines the experimental setup and provides an analysis of the results. Finally, Section 5 concludes the paper.

II. RELATED WORk

Numerous intrusion detection systems (IDSs) have been extensively studied in the literature to detect and mitigate abnormal activities within networks. However, a common challenge among these IDSs is their tendency to generate high false positive rates while achieving suboptimal detection accuracy. In response to these challenges, researchers have increasingly turned to hybrid IDSs, which integrate the robustness of signature-based intrusion detection systems (SIDS) with the adaptive capabilities of anomaly-based intrusion detection systems (AIDS) [2] [3, 4]. This combination aims to leverage the strengths of both approaches: SIDS excel in identifying known threats with predefined patterns, while AIDS offer flexibility in detecting novel, unseen threats based on deviations from normal behaviour. By merging these methodologies, hybrid IDSs seek to enhance overall detection efficacy, reduce false alarms, and fortify network security against a broader spectrum of cyber threats) [5] [6, 7].

For instance, Sumaiya et al. (2017) suggested an intrusion detection model that utilizes chi-square feature selection and multiclass support vector machine [8]. Syarif et al. applied ensemble techniques such as bagging, boosting, and stacking to enhance the intrusion detection rate and reduce false alarms. Kim et al [9]. developed a hierarchical HIDS technique that combines SIDS and AIDS models using the J48 decision tree algorithm and multiple one-class support vector machine models. Muniyandi et al. proposed an anomaly detection method called "K-Means + C4.5" that classified anomalies and normal activities in a computer system. Khraisat et al. proposed a multi-level hybrid intrusion detection model using SVM and an extreme learning machine [7]. Wang et al. proposed an IDS approach based on artificial neural networks (ANN) and fuzzy clustering (FC-ANN) to achieve a high detection rate and low false alarms. Koc et al [10]. introduced the Hidden Naïve Bayes (HNB) model for IDS issues, while Sivatha et al. proposed a lightweight IDS using a wrapper-based feature selection algorithm and neural ensemble decision tree [11]. Ghanem et al. presented a hybrid detection approach for large datasets using self- and non-self-training data and genetic algorithms [12].

Present research endeavours in the field of Intrusion Detection Systems (IDS) for Internet of Things (IoT) can be categorized into three primary approaches: Anomaly-based Intrusion Detection System (AIDS), Signature-based Intrusion Detection Systems (SIDS), and hybrid-based methods. In the context of SIDS, the focus is on utilizing pattern matching techniques to identify known attacks, also referred to as Knowledge-based Detection or Misuse Detection (Khraisat, Gondal, & Vamplew, 2018a). SIDS employs matching methods to identify previously observed intrusions. In the case of AIDS, a normal model of a computer system's behaviour is constructed using machine learning, statistical, or knowledge-based methods. Any substantial deviation between the observed behaviour and the model is considered an anomaly, potentially indicating an intrusion. These techniques operate on the assumption that malicious behaviour differs from typical user behaviour. In a nutshell, the Hybrid IDS approach combines SIDS and AIDS to enhance detection rates while reducing false alarms.

In the domain of validation strategies, researchers employ a variety of techniques, including theoretical, empirical, and hypothetical approaches, to substantiate the effectiveness of their methods. Hoda et al. leveraged AIDS based on a neural network for the detection of Denial-of-Service attacks in IoT networks. Their IDS methodology involved classifying normal and abnormal patterns and was tested using a simulated IoT network [13]. The outcomes of their performance evaluation remain to be detailed.

Diro et al. developed an IoT network attack detection system based on distributed deep learning, demonstrating that distributed attack detection outperforms centralized strategies. They assessed their approach using the NLS-KDD dataset [14]. It is worth noting that while this dataset is a variant of the KDD dataset, it still exhibits some of the limitations highlighted by McHugh, making it less than an ideal representative of real-world networks [15]. Therefore, caution is advised in using this dataset as a benchmark for IoT-specific IDS, as it originates from traditional networks [14]. This underscores the necessity of designing IDSs tailored to the unique requirements of IoT protocols, such as 6LowPAN (Low-power Wireless Personal Area Networks), and the associated architectural differences between IoT and conventional networks.

Rathore et al. introduced a semi-supervised Fuzzy learning-based distributed attack detection framework for IoT [16]. Their evaluation was conducted using the NSL-KDD dataset, which is subject to the same dataset-related limitations.

However, none of the previous works have effectively balanced accuracy and false positive rates. Recent studies have focused on reducing false positives by proposing hybrid IDSs. Unlike earlier studies that only integrated the results of both detection models, the proposed technique hierarchically combines intrusion detection systems to improve accuracy. This allows AIDS to enhance its normal profiling capability by incorporating the signature detection model [17]. The details of the proposed hybrid IDS are discussed in the next section.

III. Integrated Approach for Hybrid Intrusion Detection System in IoT

To address the limitations of Signature-based Intrusion Detection Systems (SIDS) and Anomaly-based Intrusion Detection Systems (AIDS), we have developed a Hybrid IDS that combines both approaches to effectively detect both known and unknown attacks in the IoT environment. Our approach involves utilizing AIDS for identifying zero-day attacks and SIDS for detecting well-known attacks [18] [19]. By integrating the strengths of SIDS and AIDS, we aim to create a robust and efficient IDS.

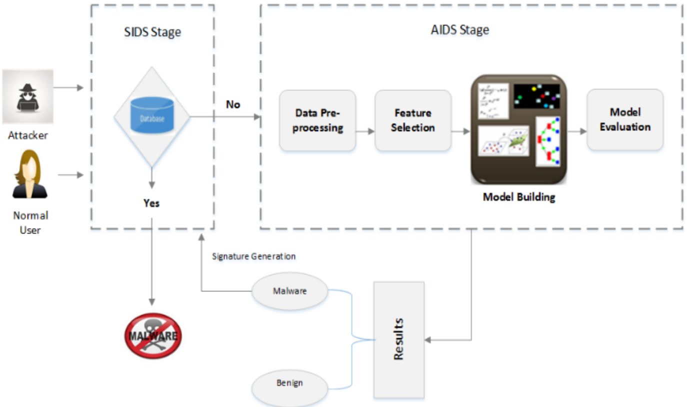

Figure 1 Hybrid Intrusion Detection System

The Hybrid IDS operates through two distinct phases: the SIDS phase and the AIDS phase, as depicted in Figure 1. In the SIDS phase, the system identifies known attacks by comparing observed network activities against signatures stored in its database. This phase enables rapid detection and response to well-documented threats, ensuring immediate mitigation measures are implemented.

Conversely, the AIDS phase focuses on profiling normal user behaviour using advanced machine learning techniques. By establishing baseline models of expected behaviour, AIDS can effectively detect deviations that may indicate anomalous or zero-day attacks. When significant deviations are detected, alarms are triggered to alert administrators or automated response systems.

Crucially, the AIDS phase provides valuable feedback to the SIDS phase by forwarding detected malicious patterns, which are then integrated into the signature database. This iterative process enhances the SIDS component's ability to recognize and respond to similar attacks in the future, continuously improving the IDS's overall effectiveness.

By integrating these two complementary phases, our proposed Hybrid IDS (HIDS) offer enhanced detection capabilities across a wide spectrum of intrusions in IoT environments. The synergy between SIDS and AIDS not only strengthens the IDS's ability to detect both known and unknown threats but also supports proactive defence strategies against evolving cybersecurity threats.

To rigorously evaluate the performance of our proposed HIDS and its individual components (SIDS and AIDS), comprehensive experiments were conducted. Detailed descriptions of each phase of the detection system, including methodologies and results, are provided in the following sections, highlighting the system's robustness and efficacy in enhancing IoT security.

A. Feature selection

In the context of the IoT ecosystem, which consists of resource-constrained smart devices with limitations in processing power, memory, energy, and communication range, Intrusion Detection Systems (IDSs) face significant challenges due to the presence of numerous irrelevant features. These irrelevant features can impose unnecessary computational and energy overhead on the system, thereby diminishing its ability to effectively detect anomalies and intrusions. Feature selection is a critical process aimed at identifying and isolating the most relevant features that can be effectively employed by the IDS to detect various forms of malware.

The feature selection process begins with data preprocessing, where initial data is cleaned and normalized to ensure consistency and reduce noise, facilitating more accurate feature evaluation. Following this, features are ranked based on their relevance and contribution to intrusion detection, using techniques such as mutual information, chi-square tests, and correlation analysis to assess the significance of each feature. Subsequently, less relevant or redundant features are eliminated in a feature reduction step, resulting in a streamlined feature set that maintains or enhances detection performance while minimizing the computational load on resource-constrained IoT devices.

The selected features are then validated through rigorous testing using benchmark datasets to ensure that the feature set effectively enhances the IDS's ability to detect various types of malwares without compromising system performance. By implementing an effective feature selection process, IDSs in IoT environments can achieve a balance between high detection accuracy and efficient resource utilization.

To achieve this, we employ a labelled dataset consisting of both normal and attack behaviour, allowing us to evaluate the relevance of different features. We utilized the information gain method for feature selection due to its rapid execution time, which enabled the extraction of the most effective feature set for a particular type of model. In the existing literature, information gain is frequently used to assess how well each distinct attribute segregates the given dataset. The overall entropy 'I' of a dataset 'S' is described as [20]:

$$I(S) = -\sum_{i=1}^{c} p_i \, log_2 \, p_i \qquad (1)$$

Here, 'c' represents the total number of classes, and 'p_i' denotes the proportion of instances belonging to class 'i'. The reduction in entropy or information gain is computed for each feature as follows:

$$IG(S,A) = I(S) - \sum_{v \mathcal{E} A} \frac{|S_{A,v}|}{|S|} I(S_v) \qquad (2)$$

Where 'v' represents the value of feature 'A', and 'S_(A,v)' denotes the set of instances where feature 'A' has the value 'v

B. Phase 1: C5 Signature-based Intrusion Detection System (SIDS)

The initial phase of our hybrid IDS is the C5 SIDS, which effectively minimizes false alarms and enhances accuracy in detecting actual attacks. By storing all known malicious signatures in the database, the occurrence of false positives is reduced. In this phase, we employ the C5 Decision Tree algorithm for malware detection. The C5 algorithm is an enhanced version of the widely used C4.5 classifier, developed by Quinlan (Quinlan, 2014), which is based on decision tree principles (Wu et al., 2008). The C5 algorithm offers several advantages, including the ability to incorporate variable misclassification costs, handle missing data, handle a large number of input fields, and rapidly build the model.

The C5 algorithm constructs a decision tree by utilizing a set of known data as input. This process follows a top-down approach, beginning with the selection of the most significant attribute, which forms the root of the tree. Subsequent branches are developed based on the values of this attribute, leading to either additional attributes or final outcomes. Once the decision tree is fully constructed, it serves as a comprehensive rule set that can be applied to new samples. This allows for the classification of these samples as either malware or benign software, with the effectiveness of the detection depending on how accurately the decision tree represents the underlying patterns in the dataset.

C. Phase 2: One-Class SVM Anomaly Intrusion Detection System (AIDS)

The second phase of our hybrid IDS is the One-Class SVM AIDS, which aims to identify unknown attacks by leveraging the output of the SIDS phase. By utilizing benign samples for training, AIDS is designed to distinguish abnormal activities, specifically unusual behaviours exhibited by malware.

In this phase, we employ the One-Class SVM (Support Vector Machine) algorithm, which learns the attributes of benign samples without relying on any information from the other class. Unlike traditional multi-class classification approaches, one-class classification focuses on describing normal behaviour based on training data examples, while the unknown malware class has no representative examples. The One-Class SVM algorithm, proposed by Schölkopf et al. [21], extends the SVM method to address the one-class problem of predicting the support of a high-dimensional distribution. It involves applying a kernel to perform feature processing and incorporates "relaxation parameters" to enhance its classification capabilities.

By utilizing the One-Class SVM, our IDS can effectively identify normal activities with a higher success rate, as there is ample training data available for the normal class. However, when it comes to zero-day attacks, which are rare and have limited initial training data, the performance may be constrained due to the scarcity of instances available for training.

D. Phase 3: Stacking C5 SIDS and One-Class SVM AIDS

Given the complementary strengths and weaknesses of Signature-based Intrusion Detection Systems (SIDS) and Anomaly-based Intrusion Detection Systems (AIDS), a hybrid method that combines both approaches using an ensemble technique is proposed. Ensemble methods in machine learning leverage multiple learning algorithms to enhance prediction accuracy, and while various ensemble methods exist, finding a suitable configuration for detecting zero-day attacks remains a challenge. To address this, a novel ensemble construction method is introduced, employing the C5 classifier as the first stage and the One-Class SVM as the second stage. This results in a robust ensemble of classifiers that improves overall detection accuracy.

Three widely recognized ensemble methods are bootstrap aggregating (bagging), boosting, and stacking. Bagging combines predictions from multiple models of the same type, with Random Forest being a prime example that combines multiple random decision trees. Boosting incrementally builds an ensemble model by training each new model on the misclassified instances from previous models, as seen in the AdaBoost technique. Stacking, also known as stacked generalization, combines predictions from multiple models by using the outputs of the base learners (stage one) as inputs for a meta-learner (stage two). In the proposed approach, two models—C5 and One-Class SVM—are built and their predictions are combined as depicted in Figure 1. The C5 classifier processes the data first, detecting known threats based on signature patterns. The output of the C5 classifier, along with the original features, is then fed into the One-Class SVM, which detects anomalies and unknown threats by modelling normal behaviour and identifying deviations.

The focus of this method lies in enhancing IDS accuracy through the utilization of the stacking approach. While conventional data mining approaches primarily aim to improve the performance of individual models, this work concentrates on the combination of different classifiers to enhance the overall performance of an IDS. By leveraging the strengths of both C5 and One-Class SVM, the approach demonstrates improved accuracy in the field of intrusion detection, effectively addressing both known and unknown threats.

IV. Model Evaluation

To thoroughly evaluate the effectiveness of the proposed hybrid IDS, a series of experiments were conducted using two well-known datasets: NSL-KDD and ADFA. These datasets are widely used in the cybersecurity research community for benchmarking the performance of intrusion detection systems.

The NSL-KDD dataset, an improved version of the original KDD'99 dataset, addresses several inherent issues such as redundant records and imbalanced distribution, providing a more reliable basis for evaluating IDS performance. The ADFA dataset, on the other hand, is specifically designed for evaluating anomaly-based intrusion detection systems and includes a diverse range of modern attack scenarios, making it highly relevant for contemporary cybersecurity challenges.

In our experiments, the C5 decision tree algorithm was employed to handle the signature-based detection phase. The C5 algorithm, known for its efficiency and accuracy, constructs a decision tree that effectively classifies known attack patterns based on predefined signatures. The decision tree's ability to handle large datasets and provide clear, interpretable rules makes it an ideal choice for the SIDS component of our hybrid IDS.

For the anomaly-based detection phase, the LIBSVM implementation of a support vector machine (SVM) was utilized with default parameters. LIBSVM is a robust and widely used library for support vector machines, offering high performance and flexibility. The One-Class SVM, implemented through LIBSVM, was trained on benign samples to establish a model of normal behaviour. This model was then used to detect deviations indicative of potential zero-day attacks or other unknown threats.

The evaluation process involved a comprehensive analysis of the hybrid IDS's performance on both datasets. Key metrics such as detection accuracy, false positive rate, and detection rate were measured to assess the system's effectiveness.

A. Dataset

The NSL-KDD Dataset The initial KDD cup99 dataset, introduced by Tavallaee, Bagheri, Lu, and Ghorbani in 2009, was initially designed for assessing Intrusion Detection Systems (IDS) [22]. However, a statistical analysis conducted on this dataset revealed a significant issue that substantially affected the accuracy of intrusion detection and led to a potentially misleading evaluation of Anomaly-based Intrusion Detection System (AIDS) algorithms, as documented by Tavallaee et al. in 2009. Consequently, an improved version known as NSL-KDD was subsequently derived from the original dataset to address these limitations. Since its development, NSL-KDD has been widely adopted as a benchmark dataset for the evaluation of IDS.

ADFA Dataset: Two datasets, namely ADFA-LD and ADFA-WD, were created by researchers at the Australian Defence Force Academy to serve as publicly available datasets that mirror the structure and methodology of contemporary attacks, as documented by Creech and Hu in 2014 [23]. These datasets encompass records from Linux and Windows operating systems, with ADFA-LD tailored for assessing host-based intrusion detection systems (HIDS) relying on system call data. The host operating system used was Ubuntu Linux version 11.04. Notably, certain attack instances in ADFA-LD were derived from previously unknown zero-day malware, rendering this dataset particularly valuable for highlighting distinctions between Signature-based Intrusion Detection Systems (SIDS) and Anomaly-based Intrusion Detection Systems (AIDS) in the context of intrusion detection. ADFA-LD comprises three distinct data categories, each featuring raw system call traces. The data within these categories were collected from the host during routine activities, such as web browsing and LaTeX document preparation.

According to Figure 2, the detection accuracy for malware is 81.5% for the NSL-KDD Test+ dataset and 97.3% for the ADFA dataset in Stage One of the hybrid IDS. In Stage Two, the detection accuracy for malware is 72.2% for the NSL-KDD Test+ dataset and 76.4% for the ADFA dataset. However, in Stage 3, the accuracy improves to 83.2% and 97.4% respectively for the two datasets. This indicates that the proposed framework achieves higher detection rates and lower false alarm rates compared to using a standalone single stage. Consequently, the hybrid IDS allows for more efficient detection of attacks.

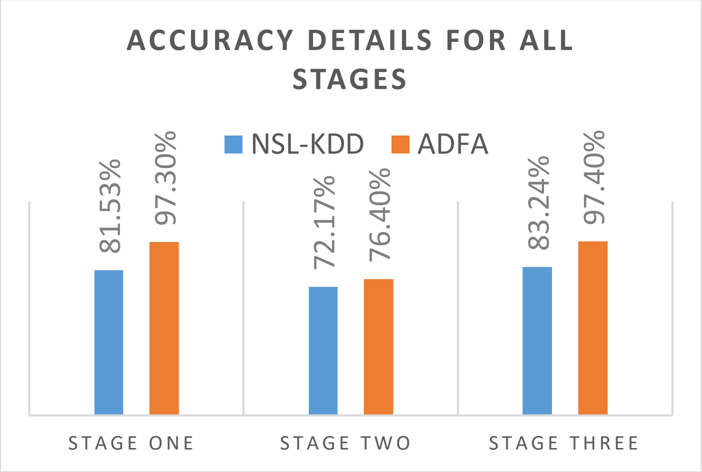

*Figure **2** Accuracy details for all stages*

V. Conclusion

This study presents a novel framework for constructing an intelligent IDS that addresses the limitations of existing IDSs. The key contribution of our framework lies in the integration of signature-based and anomaly-based IDS approaches, leveraging their respective strengths. By combining these two methodologies, our IDS is capable of identifying both known intrusions through the signature-based component and unknown zero-day intrusions through the anomaly-based component.

The hybrid IDS framework employs an ensemble of C5 (signature-based) and one-class SVM (anomaly-based) in two cascaded stages. The C5 decision tree algorithm efficiently handles the detection of known attack patterns by comparing observed activities against a comprehensive database of signatures. Meanwhile, the one-class SVM algorithm focuses on identifying deviations from established normal behaviour, thus detecting previously unseen or zero-day attacks.

Extensive experimentation was conducted using two benchmark datasets, NSL-KDD and ADFA, to evaluate the performance of the proposed hybrid IDS. The results demonstrate that the hybrid IDS outperforms individual techniques, achieving superior overall accuracy and lower false alarm rates compared to other machine learning techniques and approaches reported in previous studies. Specifically, the integration of the C5 classifier and the one-class SVM enhances the system's ability to accurately detect and classify both known and unknown intrusions, providing a robust defence mechanism for IoT environments.

The successful integration of these components underscores the potential of our approach to enhance the security of computer systems and networks. By effectively combining the strengths of signature-based and anomaly-based detection methods, the proposed framework offers improved detection capabilities, ensuring comprehensive protection against a wide range of cyber threats. These findings suggest that our proposed technique holds significant promise for the design and development of modern IDSs.

Furthermore, the adaptability and scalability of our hybrid IDS framework make it suitable for deployment in diverse network environments, addressing the evolving landscape of cyber threats. Future work may explore the integration of additional machine learning algorithms and real-time adaptation techniques to further enhance the IDS's performance and resilience.

VI. REFERENCES

[1] Z. Lv, D. Chen, R. Lou, and A. Alazab, "Artificial intelligence for securing industrial-based cyber–physical systems," *Future generation computer systems,* vol. 117, pp. 291-298, 2021.

[2] O. A. Mahdi, A. Alazab, S. Bevinakoppa, N. Ali, and A. Khraisat, "Enhancing IoT Intrusion Detection System Performance with the Diversity Measure as a Novel Drift Detection Method," in *2023 9th International Conference on Information Technology Trends (ITT)*, 2023: IEEE, pp. 50-54.

[3] M. Soltani, B. Ousat, M. J. Siavoshani, and A. H. Jahangir, "An adaptable deep learning-based intrusion detection system to zero-day attacks," *Journal of Information Security and Applications,* vol. 76, p. 103516, 2023.

[4] O. A. Mahdi, N. Ali, A. Alazab, S. Bevinakoppa, T. Al-Quraishi, and B. Das, "Diversity Measure to Tackle the Multiclass Problem in IoT Intrusion Detection Systems," in *Proceedings of International Conference for ICT (ICICT)-Zambia*, 2023, vol. 5, no. 1, pp. 25-29.

[5] A. Ammar, K. Ansam, and S. Sarabjot, "A Review on the Internet of Things (IoT) Forensics: Challenges, Techniques, and Evaluation of Digital Forensic Tools," in *Digital Forensics - Challenges and New Frontiers*, R. Dr. Denis Ed. Rijeka: IntechOpen, 2023, p. Ch. 10.

[6] A. Khraisat and A. Alazab, "A critical review of intrusion detection systems in the internet of things: techniques, deployment strategy, validation strategy, attacks, public datasets and challenges," *Cybersecurity,* vol. 4, pp. 1-27, 2021.

[7] A. Khraisat, I. Gondal, P. Vamplew, J. Kamruzzaman, and A. Alazab, "A novel ensemble of hybrid intrusion detection system for detecting internet of things attacks," *Electronics,* vol. 8, no. 11, p. 1210, 2019.

[8] I. Sumaiya Thaseen and C. Aswani Kumar, "Intrusion detection model using fusion of chi-square feature selection and multi class SVM," *Journal of King Saud University - Computer and Information Sciences,* vol. 29, no. 4, pp. 462-472, 2017/10/01/ 2017, doi: https://doi.org/10.1016/j.jksuci.2015.12.004.

[9] I. Syarif, E. Zaluska, A. Prugel-Bennett, and G. Wills, "Application of Bagging, Boosting and Stacking to Intrusion Detection," Berlin, Heidelberg, 2012: Springer Berlin Heidelberg, in Machine Learning and Data Mining in Pattern Recognition, pp. 593-602.

[10] G. Wang, J. Hao, J. Ma, and L. Huang, "A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering," *Expert Systems with Applications,* vol. 37, no. 9, pp. 6225-6232, 9// 2010, doi: http://dx.doi.org/10.1016/j.eswa.2010.02.102.

[11] S. S. Sivatha Sindhu, S. Geetha, and A. Kannan, "Decision tree based light weight intrusion detection using a wrapper approach," *Expert Systems with Applications,* vol. 39, no. 1, pp. 129-141, 2012/01/01/ 2012, doi: https://doi.org/10.1016/j.eswa.2011.06.013.

[12] T. F. Ghanem, W. S. Elkilani, and H. M. Abdul-kader, "A hybrid approach for efficient anomaly detection using metaheuristic methods," *Journal of Advanced Research,* vol. 6, no. 4, pp. 609-619, 2015/07/01/ 2015, doi: https://doi.org/10.1016/j.jare.2014.02.009.

[13] S. Abe, M. Fujimoto, S. Horata, Y. Uchida, and T. Mitsunaga, "Security threats of Internet-reachable ICS," in *2016 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, 20-23 Sept. 2016 2016, pp. 750-755, doi: 10.1109/SICE.2016.7749239.

[14] A. A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," *Future Generation Computer Systems,* vol. 82, pp. 761-768, 2018.

[15] J. McHugh, "The 1998 Lincoln Laboratory IDS Evaluation," in *Recent Advances in Intrusion Detection:*

Third International Workshop, RAID 2000 Toulouse, France, October 2–4, 2000 Proceedings, H. Debar, L. Mé, and S. F. Wu Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 145-161.

[16] S. Rathore and J. H. Park, "Semi-supervised learning based distributed attack detection framework for IoT," *Applied Soft Computing,* vol. 72, pp. 79-89, 2018.

[17] A. Alazab, A. Khraisat, M. Alazab, and S. Singh, "Detection of Obfuscated Malicious JavaScript Code," *Future Internet,* vol. 14, no. 8, p. 217, 2022.

[18] O. A. Mahdi, E. Pardede, and J. Cao, "Combination of information entropy and ensemble classification for detecting concept drift in data stream," in *Proceedings of the Australasian Computer Science Week Multiconference*, 2018, pp. 1-5.

[19] O. A. Mahdi, E. Pardede, N. Ali, and J. Cao, "Fast reaction to sudden concept drift in the absence of class labels," *Applied Sciences,* vol. 10, no. 2, p. 606, 2020.

[20] R. M. Gray, *Entropy and information theory*. Springer Science & Business Media, 2011.

[21] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and their Applications,* vol. 13, no. 4, pp. 18-28, 1998, doi: 10.1109/5254.708428.

[22] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, 8-10 July 2009, pp. 1-6, doi: 10.1109/CISDA.2009.5356528.

[23] Creech and Hu, "A Semantic Approach to Host-Based Intrusion Detection Systems Using Contiguous and Discontiguous System Call Patterns," *IEEE Transactions on Computers,* vol. 63, no. 4, pp. 807-819, 2014, doi: 10.1109/TC.2013.13.