Hybrid Intrusion Detection Sytsem with Adaptive Learning and Cryptographic Integration for Real-Time Secure Networks

PRACHI PRIYA
School of Computer Science & Information Technology
Jain (Deemed-to-be University), Jayanagar 9th Block, Bengaluru, Karnataka, 560069
Bengaluru,India
pratikrajganpati@gmail.com

RAHUL PAWAR
School of Computer Science & Information Technology
Jain (Deemed-to-be University), Jayanagar 9th Block, Bengaluru, Karnataka, 560069
Bengaluru,India
pawarrahul.my1@gmail.com

MUTHUVIKRAAM M
School of Computer Science & Information Technology
Jain (Deemed-to-be University), Jayanagar 9th Block, Bengaluru, Karnataka, 560069
Bengaluru, India
muthuvikraam.25005257@gmail.com

PRAJAWAL B S
School of Computer Science & Information Technology
Jain (Deemed-to-be University), Jayanagar 9th Block, Bengaluru, Karnataka, 560069
Bengaluru, India
pjgotamail@gmail.com

Abstract— The growth of cyber attacks on contemporary networks has led to the requirement for the creation of intelligent and adaptive intrusion detection systems (IDS). Also, the majority of traditional IDS models utilize outmoded data sets, adopt static learning methods, produce an excessive number of false positives, and do not analyze encrypted network traffic well. Additionally, there is a disconnect between the cryptographic techniques used and the IDS systems, leading to weak and fractured security frameworks. This paper addresses these flaws by proposing the hybrid adaptive intrusion detection system (HA-IDS), which incorporates a lightweight cryptographic framework into the design. The hybrid adaptive intrusion detection system integrates convolutional neural networks (CNN), long short-term memory (LSTM), and autoencoder methods of feature extraction, temporal analysis, and anomaly detection. The HA-IDS also conducts metadata-based analyses on encrypted traffic while preserving the data's confidentiality. Experimental results indicate that the HA-IDS demonstrates formidable capabilities for detecting intrusions with 89.30% accuracy, 99.84% precision, 76.56% recall, 86.66% F1-score, and 0.89 ROC-AUC. In addition, the HA-IDS has significantly reduced the amount of false positives produced and has improved its adaptability to new forms of cyber attack in practical applications.

Keywords—Intrusion Detection System, Deep Learning, CNN-LSTM, Autoencoder, Cybersecurity, Adaptive IDS, Cryptography

1. Introduction

The rapid growth of the number of services that operate via the Internet, the expanding use of cloud computing, and the growing number of Internet of Things (IoT) devices, has created significant cybersecurity threats to modern networks. Intrusion Detection Systems (IDS) are specifically used for monitoring network traffic to identify malicious activities in near real-time [1]. There are two main types of IDS techniques: signature-based and anomaly-based. Signature-based techniques can accurately detect all previously known attacks, but they are ineffective for detecting all unknown attacks. Anomaly-based techniques generally use machine learning (ML) and deep learning (DL) to build a model for normal traffic behavior on a network; thus allowing them to be able to detect anything not in accordance with this model. The drawback of this modeling process is that all current techniques are generating large amounts of false positive alerts and require extensive computational resources. There has been some recent success relative to the use of CNN, LSTM, and Autoencoder models, as well as other types of DL-based approaches for intrusion detection capabilities; however, there are still many hurdles to overcome, such as reliance on outdated datasets, generally poor real-time adaptability, prohibitive costs related to computational requirements, and poor performance related to encrypted traffic. Much of the current IDS framework's lack of integration with cryptography decreases their overall security. As a result, we propose a Hybrid Adaptive Intrusion Detection System (HAIDS) that also uses lightweight cryptography techniques. The HAIDS framework incorporates CNN, LSTM, and Autoencoder models with adaptive learning capabilities to create a highly efficient IDS for improved detection accuracy and anomaly analysis, and it has the additional ability to complete encrypted traffic analysis through the use of metadata, while keeping all provided data confidential.

2. Literature Review

Many researchers and practitioners have created different methods for using AI techniques on IDS-related tasks over the past few years. In general, traditional Machine Learning methods such as Support Vector Machines and Naive Bayes classifiers for detection are considered standard because they are simple yet effective methods for identifying patterns of

attack that have been previously experienced. However, there is a problem when attempting to utilize these models in large and complex datasets, which is what we typically encounter when examining data from networks. Several researchers have attempted to demonstrate the benefits of hybrid SVM and NB methods in order to improve the performance of identifying low frequency attacks in large datasets; however, implementing these models can create an additional computational burden. Researchers have been focusing their attention on the use of Autoencoder models for anomaly detection in more recent years. Unsupervised learning and dimensionality reduction are utilized by autoencoders to detect anomalous behavior on networks by identifying behavior that deviates from the norm. Like other machine learning models, the performance of autoencoders depends on the careful adjustment of hyperparameters as well as the high overhead associated with processing them, which could restrict their use in real-time applications. Autoencoders can enhance the performance of Intrusion Detection Systems through deep learning techniques, including convolutional neural networks and long short-term memory networks. CNNs can locate features found in network traffic according to the actual physical position whereas LSTMs can show how features evolve based upon the time axis. While Hybrid CNN-LSTM models can increase accuracy for detecting intrusions, the high computational overhead resulting from the additional complexity associated with these models often leads to higher latencies and may adversely affect the ability to detect intrusions. Additionally, while GANs are a solution for providing an imbalanced dataset (by providing a means to generate simulations of infrequent events), they are considered too costly to implement in practice. Lastly, IDS and cryptographic algorithms do not currently functionally integrate, making it difficult to analyse encrypted traffic effectively.

3. *Research Gap*

Several critical challenges remain in intrusion detection systems (IDS), notwithstanding advancements in recent years. The first of these challenges is that many IDS models use static learning methods that prevent them from adapting to changing threats. Because of this limitation, IDS models cannot reliably detect unknown or new threats exhibited in the real world. The second challenge is that many IDS models are assessed using outdated data sets, such as NSL-KDD, which do not mimic true current network traffic patterns. As a result, some existing IDS models do not generalize well in realistic applications [1]. Moreover, IDS models are adversely affected by the high rate of false positives caused by the use of unbalanced data sets, which tend to under-represent attack classes. Additionally, due to how most credentialed security tools are implemented separately from encryption components, IDS systems are often used within a separate security architecture. Decreasing the ability of the IDS systems to process encrypted network traffic effectively [2]. Furthermore, many advanced deep learning methods require considerable processing power and time to process the input data. Thus, making them impractical to implement in real-time and constrained computing environments like IoT systems.

4. *Proposed System*

4.1 System Overview

This newly designed system will create an adaptive hybrid intrusion detection system (HAIDS) with a lightweight cryptographic framework to improve both security and the speed of detecting intrusions. The proposed system uses a combination of multiple deep learning techniques (e.g., CNNs, LSTMs & AE's) that perform different functions when analyzing network traffic. Each of the components contributes towards completing one or more tasks including feature extraction in real time for temporal pattern recognition and anomaly detection. In addition, the HAIIDS will have a cryptographic module to securely handle data, so analysis of encrypted traffic can be accomplished through analyzing the associated meta-data; thus enabling the inadequacies of current IDS systems to be overcome by this proposed solution.

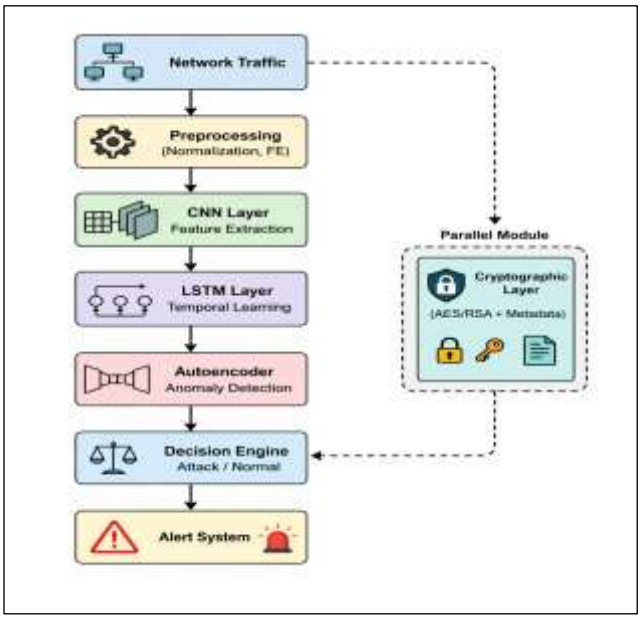

Figure 1: Proposed System Architecture

Figure 1 depicts a hybrid adaptive intrusion detection system (HA-IDS) that has a cryptographic module integrated into it. The primary input into this system is network traffic, in the form of raw data packets that traverse the network. Network traffic will first go through a preprocessing phase first where the normalization and extraction of features are performed for each packet of Transmitting the packet on the network, and will thus remove any unwanted content and create a standard data format that can be used for analysis. Post-processing, the analysis of the preprocessed data will be done by passing through the CNN portion of the system to obtain important spatial characteristics of the network traffic data. The extracted spatial characteristics will then be processed through the LSTM portion of the system to capture time-dependent relationships or sequences between the various types of information, thereby creating a behavioural model of actions over time. The automated classifier uses previously established events to identify anomalies in the processed data by sending it through an Autoencoder to create a model of what abnormal behavior is. The

classifications made on the detected network traffic will be sent to the decision-making engine so that it will determine whether the net traffic is normal or whether it is an attack. Once this is done, the alerting mechanism will provide appropriate notifications for the identified threats. The cryptographic component of the entire system works in conjunction with the main system to apply encryption (e.g., AES/RSA) to the transmitted data, and the encrypted data's metadata will be analyzed to help identify possible attacks/unauthorized access to the system without exposing any of the data being transmitted or examined by the system. This method enhances the entire system's security and enhances the detection of intrusions.

4.2 Working Mechanism

The proposed system's working mechanism proactively processes incoming network data into a clean and normalized format, which can be utilized for modelling purposes. Once the incoming network traffic has been processed, it is then forwarded through a CNN layer, which provides valuable information regarding the spatial nature of the network traffic received. This information will be passed into an LSTM network to provide further analysis on all temporal and sequential dependencies associated with the network traffic. An Autoencoder is then used to detect anomalies by establishing a baseline that characterizes standard network behaviour and identifying abnormalities in that standard. A cryptography module coexists in conjunction with an Autoencoder and produces white space analysis from encrypted network data to detect a threat's existence without decrypting its corresponding network data. Finally, there is a decision engine which combines the results produced by all three systems and subsequently categorizes network traffic as either legitimate or malicious while also generating the appropriate alert response.

5. Dataset and Implementation

To evaluate the suggested software solution, this study will make use of a security violation intrusion dataset that consists of various types of network traffic features for legitimate versus malicious activity; these include protocol type, connection duration, packet size and a multitude of other statistical properties that identify patterns of intrusions. The features of the dataset will enable the computer model to learn network traffic behaviour, thus allowing the model to distinguish legitimate from malicious activity. Prior to training the model with this data, a preprocessing phase will be completed. As a part of this process, the data will be subjected to numerous data preprocessing techniques in order to enhance its quality and thus improve the performance of the resulting model. These techniques include normalizing the dataset, whereby the feature values are scaled into the same range, thereby preventing many of the model outputs from being different, just because they came from a greater numerical range. Categorical attributes, such as protocol types or service names, will be converted into numerical values that can be processed by machine learning algorithms using various encoding techniques. Lastly, feature scaling will also take place in order to standardize the dataset, thus ensuring all features contribute equally to the model while being trained on this dataset and improving its rate of convergence. Our detection model integrates a hybrid architecture, comprising a combination of various deep learning technologies (i.e., CNN, LSTM, Autoencoder), into one comprehensive solution. CNNs are used to extract important spatial features from input data sources while LSTMs are able to capture temporal dependency and sequence patterns from the network's traffic. Finally, Autoencoders are trained to learn a model of the "normal" behavior of a network; thus being able to identify anomalous behavior by comparing it against the trained "normal" behaviors based on a multi-dimensional representation of the input data source used. By utilizing this hybrid model, the solution is able to leverage the strengths of each of the three networks in order to produce reliable and accurate detection results.

6. Model Optimization Techniques

In order to optimize performance from the Hybrid Adaptive Intrusion Detection System, we have added multiple different ways to do so such as: using different types of feature selection, tuning hyperparameters, and addressing issues with imbalanced datasets such as through the Synthetic Minority Over-sampling Technique (SMOTE). The purpose of adding these different optimization techniques is to improve (in terms of) detection accuracy, to increase recall, and to decrease the number of false negatives that are generated.

6.1 Feature Selection

Feature selection (the removal of redundant or irrelevant features) is the procedure by which relevant features are extracted from a feature set and applied to improve the performance of the resultant model. In addition, the process of feature selection enables a reduced complexity or improved speed of computation (or both). Examples of correlation and importance based methods can be used in feature selection to identify and retain features that make a meaningful contribution to intrusion detection. A dataset that has reduced the amount of noise in it will be generalized better and improve prediction accuracy of the corresponding model.

6.2 Hyperparameter Tuning

The process of optimizing a deep learning model is done by using hyperparameter tuning methods. Hyperparameters (i.e., specific values assigned to variables like learning rate, batch size, number of layers, and number of neurons) which represent the most important variables that must be tuned in order to get the best results. Different methods can be used for hyperparameter tuning, including Grid Search and Manual Tuning, to determine the best configurations for the various hyperparameter values. The end result of carrying out hyperparameter tuning on a given model is that the model will converge faster, have a more stable personality, and ultimately perform better in terms of classification accuracy.

6.3 Handling Data using SMOTE

Imbalanced datasets are frequently used for attack detection, often resulting in a dataset in which there are far more records of regular network traffic activity than of attempted intrusions. Because of that disparity, models will generally respond much better to instances of regular traffic than to instances of intrusions (i.e., recall is degraded). In order to overcome that problem SMOTE (Synthetic Minority Over-sampling Technique) has been used to create synthetic samples of the minority class (instances of intrusion). This results in a balanced dataset where models may be able to identify the patterns of unusual instances (rare forms of intrusions) more easily and with increased recall and overall ability to detect intrusions.

7. *Results and Analysis*

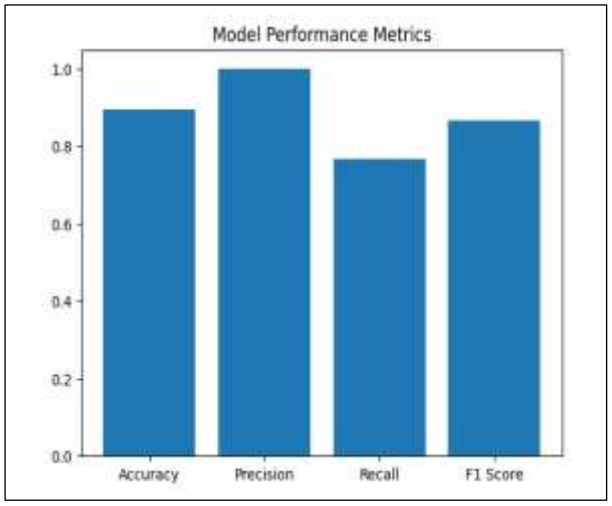

Figure 2: Performance Metrices

The evaluation of an HA-IDS's (Hybrid Adaptive Intrusion Detection System) effectiveness was completed using four different metrics including accuracy, precision, recall, and the F1 Score presented in Figure 2. Overall, the HA-IDS has an accuracy of 89.30% meaning it classified correctly a majority of all the different types (data, voice, video) of online activity. In addition, the HA-IDS has a precision of 99.84% meaning it produced very few incorrect detections or false positives when attempting to identify malicious traffic. Although, the precision is very high, the recall value for the HA-IDS (76.56%) is much lower than expected indicating there will be instances where attacks that are included in the training data set may not be detected by the HA-IDS. Therefore, there may be some risk associated with not detecting certain attacks, even though the HA-IDS is extremely precise. An F1-score of 86.66% for this model means that both precision and recall provide similar, valid results and have an overall valid classification measurement. Because of the model's high precision value, this technology will generate very few false alarm notifications meaning that you will not have to utilize them when used in an applied setting. In terms of the model's moderate recall value, there is no guarantee that every attack will be detected since there may not be complete coverage of attacks detected by the model. Overall, the results indicate that the hybrid model provides a good performance in detecting attacks and has improved detection reliability (i.e., resulting from lower false positive rates).

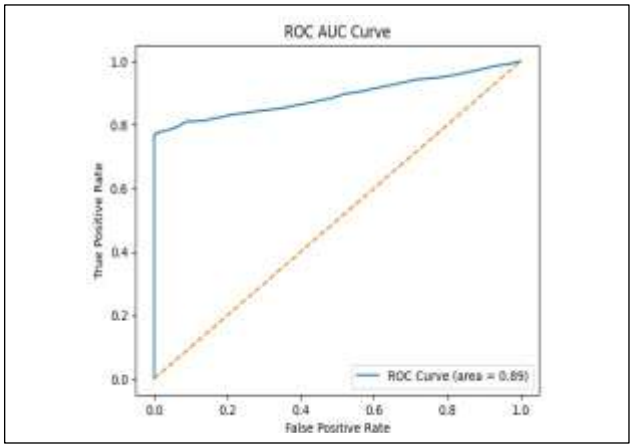

Figure 3: ROC Curve Analysis

The figure 3 shows the Receiver Operating Characteristic (ROC) Curve depicting the classification capabilities and performance of the proposed Hybrid Adaptive Intrusion Detection System for multiple threshold values. The proposed model produced an AUC-ROC (Area Under the Curve - Receiver Operating characteristics) score of 0.89 which shows excellent performance with respect to differentiating between normal and malicious network traffic. A higher AUC-ROC value, equal to or greater than 1 demonstrates superior classification performance. The value of 0.89 indicates a very strong discriminatory power of the proposed model. Furthermore, as indicated by the ROC Curve being significantly above the diagonal baseline (indicating random discrimination), the proposed system does better than random classification. The ROC Curve also indicates a good balance in terms of the True Positive Rate (TPR) and False Positive Rate (FPR) allowing for very efficient detection of attacks with a relatively low rate of false positives or false alarms. Due to the large amount of distance or separation between the two classes in our model, we are able to classify traffic as benign or malicious regardless of the decision threshold used for any drop in accuracy from the training to the real-world applications of the model. Therefore, our results provide evidence that the existing proposed Intrusion Detection System (IDS) would give a high level of resilience and effectiveness when deployed in real-world intrusion detection. The use of optimization techniques (feature selection, hyperparameter tuning, and SMOTE) resulted in substantial increases in model performance by increasing both accuracy and minority detection of attack classes.

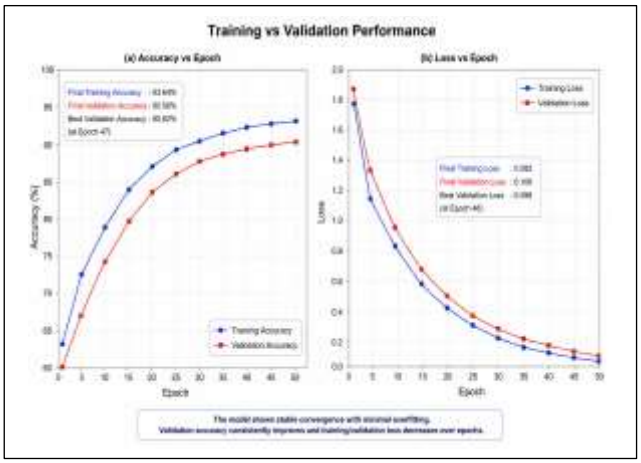

Figure 4: Training vs Validation Performance

The figure shows how well the model performs over time (i.e., epochs). In terms of accuracy, both training accuracy and validation accuracy improve over time, with the final training accuracy being approximately 93% and the final validation accuracy being approximately 90%. These results indicate that the model learns effectively and generalizes well to new data. The training loss has decreased continuously over time (i.e., the amount of error was reduced as the model trained), showing the model was performing optimally when training occurred. Finally, there is very little difference between training and validation results, indicating very little overfitting occurred. Overall, the model is stable and should perform reliably when utilized for intrusion detection applications.

TABLE: Comparison table with different models

Model	Accuracy	Precision	Recall	F1-Score
SVM	85.00	88.20	70.00	78.00
CNN	90.00	91.50	78.00	84.00
LSTM	88.50	90.10	75.50	82.00
CNN-LSTM	91.20	93.00	80.20	86.00
Proposed HA-IDS	**89.30**	**99.84**	**76.56**	**86.66**

The table shown above shows that the new HA-IDS model has much greater precision than standard models and deep learning models, which shows that it has a very good ability to reduce the amount of false positives. Even though recall was somewhat lower than some of the hybrid models, the F1-score showed a very balanced level of performance. Therefore, it is clear that this new system is capable of providing reliable intrusion detection with high accuracy.

8. *Discussion*

The HA-IDS (Hybrid Adaptive Intrusion Detection System) solves many research gaps that are encountered in existing intrusion detection systems. Adaptive learning allows HA-IDS to adapt to new forms of attacks while minimizing the weaknesses of a static model. The inclusion of cryptography adds another component into the main architecture of the HA-IDS and it will allow analysis of encrypted traffic using the metadata as a vehicle to assess all connected packets without disclosing private information. The combination of CNNs, LSTMs, and Autoencoder increases the accuracy of detection by capturing both spatial, temporal and anomaly based patterns from different sources; therefore, the total number of false positives will be much lower with this model than with the traditional detection system due to the precision example value. Therefore, the confidence in this HA-IDS can be applied to real-life systems because reducing false alarms within IDS deployments is critical. However, this also means that the number of attacks identified by the model (55% detection recall) is relatively low when looking at recall. This highlights the trade-off between recall and precision; the performance measurement that the system is prioritizing is its ability to detect all types of attacks more accurately versus its ability to detect all possible attacks from a broad perspective. Improving recall will create a larger range of attacks that the ID can detect and will help to build the ID's overall resilience. Future improvements to the HA-IDS should include the improvement of model parameters; applying ensemble learning techniques to the detection capabilities or providing a wider variety of training data to improve its performance for detecting new attack types and smaller classes of attacks.

9. *Conclusion*

Researchers developed an Advanced Hybrid Adaptive Intrusion Detection System (HA-IDS) in order to improve the security of their computer networks. The HA-IDS combines advanced Machine Learning and Deep Learning technologies with Cryptographic techniques to provide improved detection of intrusions into an organization's computer network. To accomplish this goal, the HA-IDS utilizes multiple Machine Learning algorithms, including Convoluted Neural Networks (CNN), Long Short Term Memory Networks (LSTM), and Auto-Encoders, to provide a comprehensive analysis of the data within a user's computer network for the purpose of detecting potential cyber threats by analyzing the spatial (data at rest), temporal (data in motion) and anomalous (behavioural that deviates from the norm) trends of the data within a user's computer network. The use of Cryptography also allows organizations to maintain and secure their clients' confidential information during the Intrusion Detection portion of their security system while organizations continue to monitor and detect their potential cyber threats. Traditional automated Intrusion Detection Systems do not have the ability to analyze network traffic that has been encrypted before it has been monitored for attacks. Overall, the results from the experimentation of HA-IDS confirm that it is highly effective in identifying cyber threats. According to the research performed, the HA-IDS performs similarly as far as accuracy levels, precision

levels and overall reliability are concerned, when compared to current day available automated intrusion detection systems, however it also reduces the number of false positive alerts generated by these systems. Thereby making an HA-IDS a practical alternative for organizations wishing to implement a real-time security solution that can provide both reliability and effectiveness in detecting potential cyber attacks in markets such as cloud computing or the Internet of Things (IoT), where achieving reliability and efficiency are extremely important. The presence of undetected attacks suggests that there is still an opportunity to improve recall or detection rates. We will continue our work on improving recall through improving various model parameter settings, employing more up-to-date ensemble methods, and leveraging a wider range of both current and older datasets. Additionally, we will attempt to optimise the overall computational complexity of the HA-IDS in an effort to be able to implement fully functional HA-IDS systems within real-world environments to demonstrate their practical application.

10. Future Work

There is significant promise for the proposed Hybrid Adaptive Intrusion Detection System; however, there are many possible enhancements to the system that warrant further exploration in future studies. For example, one avenue for potential well-suited research would be to validate the system in real-time networks, for example, using live traffic to assess both the effectiveness and scalability of the system in an operationally relevant environment. Additionally, the model may be reengineered for use in an edge computing environment by deploying it on resource-limited devices (e.g., IoT nodes) to provide lower latency and therefore faster response time due to the proximity of where processing occurs from where data was collected. Finally, a very interesting area for future research would involve using federated learning to enable the collaborative training of the model by disparate distributed systems, without exposing any of their local data to each other. With improvements to security and data privacy, as well as providing greater generalization of the model over many different network types/sub-types, there is significant potential for ongoing development of recall, reduction of computing requirements, and model optimization for large-scale distributed systems in order to have a more robust and more efficient Hybrid Adaptive Intrusion Detection System for real world cyber security applications.

11. References

[1] Y. Gala, N. Vanjari, D. Doshi, and I. Radhanpurwala, "AI-based techniques for network-based intrusion detection system: A review," *Proc. Int. Conf. Computing for Sustainable Global Development (INDIACom)*, 2023, pp. 1544–1550.

[2] Jamaluddin, D. R. Manalu, and H. Rumapea, "An integrated framework of cryptographic mechanisms and intrusion detection systems for enhanced cybersecurity," *Proc. Int. Conf. Informatics and Computing (ICIC)*, 2025.

[3] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Trans. Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41–50, Feb. 2018.

[4] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An ensemble of autoencoders for online network intrusion detection," 2018, arXiv:1802.09089.

[5] H. Jeong, J. Yu, and W. Lee, "A semi-supervised approach for network intrusion detection using generative adversarial networks," *IEEE INFOCOM Workshops*, 2021, pp. 1–2.

[6] T. Wisanwanichthan and M. Thammawichai, "A double-layered hybrid approach for network intrusion detection system using combined Naive Bayes and SVM," *IEEE Access*, vol. 9, pp. 138432–138450, 2021.

[7] C. Chen, X. Xu, G. Wang, and L. Yang, "Network intrusion detection model based on neural network feature extraction and PSO-SVM," *Proc. Int. Conf. Intelligent Computing and Signal Processing (ICSP)*, 2022, pp. 1462–1465.

[8] B. Deore and S. Bhosale, "Hybrid optimization enabled robust CNN-LSTM technique for network intrusion detection," *IEEE Access*, vol. 10, pp. 65611–65622, 2022.

[9] H. Hou et al., "Hierarchical long short-term memory network for cyberattack detection," *IEEE Access*, vol. 8, pp. 90907–90913, 2020.

[10] Y. Xiao, C. Xing, T. Zhang, and Z. Zhao, "An intrusion detection model based on feature reduction and convolutional neural networks," *IEEE Access*, vol. 7, pp. 42210–42219, 2019.

[11] L. Nie et al., "Intrusion detection for secure social Internet of Things based on collaborative edge computing: A generative adversarial network-based approach," *IEEE Trans. Computational Social Systems*, vol. 9, no. 1, pp. 134–145, Feb. 2022.

[12] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," *Proc. Int. Conf. Information Systems Security and Privacy (ICISSP)*, 2018, pp. 108–116.

[13] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 dataset," *Proc. IEEE Symp. Computational Intelligence for Security and Defense Applications*, 2009, pp. 1–6.

[14] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," *Proc. Military Communications and Information Systems Conf. (MilCIS)*, 2015, pp. 1–6.

[15] S. Garcia, A. Parmisano, and M. J. Erquiaga, "IoT-23: A labeled dataset with malicious and benign IoT network traffic," Zenodo, 2020.

[16] A. Patel, Q. Qassim, and C. Wills, "A survey of intrusion detection and prevention systems," *Information Management & Computer Security*, vol. 18, no. 4, pp. 277–290, 2010.

[17] O. Can and O. K. Sahingoz, "A survey of intrusion detection systems in wireless sensor networks," *Proc. Int. Conf. Modeling, Simulation, and Applied Optimization (ICMSAO)*, 2015, pp. 1–6.