Intrusion Detection System

Software Design and Requirement Specification

Session: 2021 - 2025

Submitted by:

Fatima Shahid	2021-CS-675
Faiza Atta	2021-CS-639
Sadia Latif	2021-CS-668

Supervised by:

Dr. Irfan Yousuf

Department of Computer Science, New Campus

University of Engineering and Technology Lahore, Pakistan

Contents

List of Figures

List of Tables

Chapter 1

Requirement Specification

This Intrusion Detection System (IDS) is designed to protect network infrastructure by detecting, analyzing and responding to suspicious or malicious activities in real-time. Using deep learning techniques such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) and Deep Neural Network (DNNs) etc, the IDS will monitor network traffic to detect both known and unknown threats including Distributed Denial of Service (DDOS) attacks and botnet intrusions. The system aims to improve detection accuracy adapt to evolving cyber threats and reduce false positives and negatives. By continuously learning from new data, the IDS will enhance its ability to identify anomalies and classify threats effectively. It will integrate with existing security infrastructure, providing comprehensive protection, generating detailed reports and supporting real-time analysis.

1.1 Functional Requirement

These are the key points outlining the functional requirements of our intrusion detection system providing a comprehensive overview of its core functionalities.

- **System Monitoring:** The system must continuously monitor the entire network including inbound and outbound connection for suspicious activity. It shall track real-time system health to detect potential vulnerabilities and intrusions. Ensure 24/7 operation with minimal latency, providing continuous insight into network health.
- **Network Traffic Analysis:** The system shall analyze all network traffic to detect deviations from normal patterns. It shall analyze traffic for signs of malicious activities such as Distributed Denial of Service (DDOS), botnet activity

and other attacks. It shall compare traffic behavior against baseline metrics to identify unusual activity that may indicate an attack.

- **Machine Learning Integration:** Integrate machine learning (ML) models that are capable of detecting and classifying malicious network behavior. The system shall utilize pre-trained ML models for real-time detection of known and novel cyber threats. The system shall continuously update the models based on new attack data to improve detection accuracy.

- **Feature Extraction:** The IDS must be capable of extracting relevant features from raw network traffic data. It shall extract relevant characteristics that will serve as inputs to machine learning models. It shall include the extraction of protocol-specific characteristics, traffic behavior over time and patterns that indicate normal or suspicious activities.

- **Labeling and Ground Truth:** The system shall use the CICIDS2017 dataset for accurate labels of network traffic data. This dataset which contains pre-labeled examples of both normal and malicious traffic will be utilized to train machine learning models. By using these predefined labels, the system will be able to distinguish between benign and suspicious activities with greater accuracy. The labeled data will play a crucial role allowing the IDS to improve its detection capabilities by continuously learning from accurately classified traffic.

- **Model Training and Updating:** The IDS must be capable of regular training and updating of its machine learning models. Initially, train the models on the network traffic data to detect evolving attack patterns. Retrain the models using newly gathered data to improve detection rates and reduce false positives/negatives.

- **Data Visualization:** The system must provide clear and intuitive data visualizations for network administrators, including real-time traffic monitoring dashboards, anomaly heatmaps and visual summaries of detected intrusions. The visualizations shall be easy to interact with. So, administrators can explore specific network events in detail.

- **Customization and Configuration:** The IDS shall allow network administrators to customize attack types. Customization option shall include setting sensitivity levels for various types of attacks. The system must be adaptable to different network environments and should offer flexibility in integrating with existing security infrastructures.

- **Continuous Monitoring and Improvement:** The system shall continuously monitor its own effectiveness in detecting new attacks.
- **Resource Usage Monitoring:** The system shall track the usage of system resources such as CPU, memory etc to detect abnormal behaviour that may indicate resource-based attacks. This will help prevent Denial of Service (DOS) attacks.
- **Incident Response Automation:** The system shall have automated response mechanisms that can take predefined actions when a threat is detected. These actions may notify administrators with alerts.

1.2 Non-functional Requirement

Below are the non-functional requirements that offer a clear and detailed understanding of the key qualities that define our intrusion detection system.

- **Performance:** The system shall analyze network traffic quickly to ensure real-time detection allowing for high traffic throughput while minimizing delays. It shall scale effectively to handle larger network environments without compromising on speed.
- **Accuracy:** The IDS shall accurately identify malicious network activities reducing false positives and false negatives. Continuous model updates shall improve detection accuracy as new attack patterns appears.
- **Resource Efficiency:** The system shall use minimal CPU, memory and storage resources ensuring it does not slow down the host system or other tasks. It shall be designed to work well in low-resource environments without slowing down.
- **Compatibility:** The IDS shall work well with different versions of the operating system and network setups and it shall adapt to updates without issues. It shall integrate with both legacy and modern network infrastructures.
- **Security and Privacy:** The IDS shall ensure that sensitive data such as network logs and analysis results are stored securely and protected from unauthorized access.
- **User Interface:** The interface shall be easy to use enabling security teams to navigate the system and interpret results efficiently. It shall support dashboard that display real-time traffic monitoring and system alerts.

- **Customizability:** The system shall allow users to customize attack types and settings based on their specific security needs and preferences.
- **Scalable Dataset:** The IDS shall support the use of a large and diverse datasets covering various types of malicious attacks to improve its effectiveness.
- **Generalization:** The system shall be able to detect new types of attacks by learning from past data ensuring it remains effective even against unfamiliar threats. It shall ensure that detection algorithms generalize well to diverse network environments without requiring manual retraining.
- **Anomaly Detection:** The IDS shall detect unusual or suspicious activities that deviate from normal behavior even if they do not match known attack patterns. It shall identify zero-day threats by recognizing new abnormal traffic patterns that were not previously observed.
- **Continuous Monitoring:** The IDS shall continuously monitor network traffic and respond to attacks in real-time.
- **Interoperability:** The IDS shall be able to communicate and exchange data with other security systems (e.g., firewalls, SIEMs) to enhance overall security.

Chapter 2

Design specification

2.1 Detailed literature review

- **Koc* et al** .[1] Koc, Mazzuchi and Sarkani present a novel approach to intrusion detection by utilizing a Hidden Naïve Bayes (HNB) multiclass classifier. They begin by discussing the limitations of traditional intrusion detection systems particularly the Naïve Bayes classifier which assumes conditional independence between features. The researchers propose the HNB classifier which relaxes the independence assumption and incorporates dependencies between attributes. The paper explains the model's structure, feature selection techniques and discretization methods used to enhance detection accuracy. Using the KDD Cup 1999 dataset, they demonstrate the HNB model's superior performance compared to other methods such as the Naïve Bayes, SVM and the KDD'99 competition winner. The HNB classifier achieves an average accuracy of 93.72% on test data in detecting denial-of-service (DOS) attacks.

- **Alrawashdeh and Purdy**.[2] Alrawashdeh and Purdy contribute to the field of intrusion detection by implementing a deep learning approach that combines a Restricted Boltzmann Machine (RBM) with a Deep Belief Network (DBN) for anomaly detection. They address challenges in real-time detection of novel attacks proposing a method where the RBM performs unsupervised feature reduction before feeding results into the DBN which uses a logistic regression classifier for multi-class classification. This architecture is evaluated using the KDD CUP'99 dataset achieving a detection rate of 97.9% with a false negative rate of 2.47%. It improves detection speed and accuracy.

- **Novo et al**.[3] Larriva-Novo, Sánchez-Zas, Villagrá and Vega-Barbas contribute

to intrusion detection systems by proposing a dynamic multi-class classifier designed to improve accuracy in detecting various network attacks. They address the limitations of static machine learning models which struggles to detect different attack types effectively. The researchers implement several machine learning models and aggregate their predictions through an ensemble model to dynamically select the most suitable prediction. Using the UNSW-NB15 dataset, they achieve an accuracy of 87.6%. This approach demonstrates a significant improvement over individual static models detecting challenging attack types like worms and DOS.

- **Zhao et al.**[4] Zhao, Zhang and Zheng address the challenges of traditional neural networks in intrusion detection by proposing a hybrid model combining a Deep Belief Network (DBN) with a Probabilistic Neural Network (PNN). To enhance the DBN's feature learning, they introduce a Particle Swarm Optimization (PSO) algorithm to optimize the number of hidden-layer nodes. The model is evaluated using the KDD CUP 1999 dataset achieving an accuracy of 99.31% with the unoptimized DBN-PNN and 99.14% with the optimized version giving best performance over traditional models in terms of detection accuracy and speed.

- **Ahmim et al.**[5] Ahmim, Maglaras, Ferrag, and colleagues introduce a hierarchical intrusion detection system (IDS) combining decision tree and rule-based models for enhanced classification. The system utilizes the REP Tree, JRip algorithm, and Forest PA, each of which classifies network traffic based on different features. The system is trained and tested using the CICIDS2017 dataset, giving high performance in accuracy, detection rate and false alarm reduction compared to other machine learning approaches. The final model achieves an accuracy of 96.665% with a low false alarm rate of 1.145%.

- **Radford et al.**[6] Radford, Apolonio, Trias and Simpson introduce a novel approach to network traffic anomaly detection using Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) models. They aim to identify outlier network traffic by training a model on sequences of network flow data. The authors highlight the limitations of rule-based cybersecurity methods and propose their model as a more flexible alternative capable of detecting unknown attack patterns. Using the ISCX IDS dataset their approach achieved an AUC of 0.84 highlightening the importance of LSTM networks for unsupervised anomaly detection in network security.

- **Pektas and Acarman**.[7] Pektas and Acarman address botnet detection through a deep learning approach combining convolutional neural networks (CNN) and long short-term memory (LSTM) networks. They highlight the limitations of traditional botnet detection methods like signature matching and blacklists which struggle to keep up with evolving botnet behavior. The researchers extract network flow features and use a graph structure to represent connections between hosts. By training and testing their model on the CTU-13 and ISOT datasets they achieve an accuracy of 99.3% and an F-measure of 99.1%. Their study demonstrates the effectiveness of using deep learning for botnet detection based on network flow summaries.

- **Nguyen et al**.[8] Nguyen, Nguyen, Choi and Kim present an Intrusion Detection System (IDS) using Convolutional Neural Networks (CNN) for detecting Denial of Service (DOS) attacks. They discuss the challenges in existing IDS techniques such as misuse-based and anomaly-based methods which struggle with high volumes and variants of malicious traffic. The researchers propose IDS-CNN which preprocesses network traffic data into a matrix format suitable for CNN input. They use the KDD Cup 1999 dataset to train and test their model which classifies traffic into five categories; normal and four types of attacks. The CNN model outperforms traditional machine learning methods like K-Nearest Neighbors (KNN), Support Vector Machine (SVM) and Naïve Bayes achieving a detection accuracy of 99.87%. This elaborates the effectiveness of CNN in DOS detection compared to other machine learning models.

- **Shone et al**.[9] Shone, Ngoc, Phai and Shi introduce a deep learning model for network intrusion detection aiming to address the limitations of existing approaches such as high human interaction and low detection accuracy. They propose a novel classification model combining Non-Symmetric Deep Autoencoders (NDAE) for feature extraction and Random Forest (RF) for classification. The model is evaluated using the KDD Cup '99 and NSL-KDD datasets. The results show that their model outperforms traditional methods achieving an accuracy of 97.85% on the KDD dataset and 85.42% on the NSL-KDD dataset. The combination of deep learning for feature extraction and shallow learning for classification provides a promising solution for modern network intrusion detection systems.

- **Farnaaz and Jabbar**.[10] Farnaaz and Jabbar address the limitations of current intrusion detection systems (IDS) by proposing a model based on the Random Forest (RF) algorithm for effective attack classification. They discuss the complexities of IDS due to the varied nature of network traffic and the need for robust

detection methods. The model implemented using the NSL-KDD dataset shows improved performance in terms of detection rate and false alarm rate compared to other classifiers like J48. By applying feature selection using Symmetrical Uncertainty (SU), they reduce dimensionality and improve accuracy. Their approach achieves an accuracy of 99.67%, a detection rate of 99.84% and a low false alarm rate.

- **Elmasry et al**.[11] Elmasry, Akbulut and Zaim contribute to the field of network security by conducting an extensive empirical study on multiclass classification-based network intrusion detection. They begin by outlining the limitations of traditional detection systems such as high false alarm rates and limited capability to handle multiple attack types. To address these issues, the researchers propose the use of deep learning models.They employ a Particle Swarm Optimization (PSO) algorithm to optimize the hyperparameters of these models further improving their performance. The authors evaluate their approach using four well-known intrusion detection datasets. The study addresses the false alarm rate (FAR), with DBNs reducing FAR to 0.76%. The researchers conclude that deep learning models, especially when optimized with PSO, offer enhanced accuracy and reliability in network intrusion detection although at the cost of increased computational demand.

- **Shah et al**.[12] Shah, Clachar and Minimair contribute to the field of anomaly-based network intrusion detection systems (NIDS) by developing multiclass classification baselines using various machine learning algorithms and neural networks. Their research highlights the limitations of traditional signature-based systems and introduces an anomaly-based approach capable of distinguishing legitimate network traffic from both direct and disguised network intrusions. The authors detail their model architecture which includes a 6-layer neural network and several classical machine learning models. Using the Advanced Security Network Metrics and Tunneling dataset, they achieve a 95% accuracy in detecting intrusions. This study marks a significant advancement in NIDS, showcasing the potential of deep learning and machine learning models in detecting disguised network intrusions.

- **Acharya et al**.[13] Acharya and colleagues present a study that evaluates the effectiveness of machine learning-based classifiers for detecting network intrusions in both binary and multiclass classification tasks. The research addresses the challenges of existing network intrusion detection systems (NIDS) such as poor detection of zero-day attacks and the imbalance in datasets. To overcome

these issues, the authors propose a method of reducing the number of target classes by regrouping rare attack types improving overall classification performance. The study uses publicly available datasets and demonstrates that reducing the number of output classes increases the accuracy of machine learning classifiers. For binary classification tasks some classifiers achieved 100% True Positive Rate (TPR) while in multi-class classification performance varied depending on the dataset and the number of target classes. The results show that accuracy improved when the number of output classes was reduced and when features strongly correlated with the output class were included.

- **Ashiku et al**.[14] Ashiku and Dagli explore the application of deep learning techniques to develop a network intrusion detection system (NIDS) capable of detecting both known and zero-day attacks. The proposed system utilizes a convolutional neural network (CNN) architecture with hyperparameter tuning achieving high accuracy in classifying network traffic. The model is designed to classify multiple types of network attacks. The model was tested on the UNSW-NB15 dataset which represents modern network traffic with synthetic attack activities. The study demonstrates that deep learning models can significantly improve detection performance. Their approach achieved an accuracy of 94.4% on pre-partitioned datasets and 95.6% on user-defined datasets showing a considerable enhancement over existing models.

- **Maithem et al**.[15] Maithem and Al-Sultany propose a network intrusion detection system (NIDS) using deep neural networks (DNNs) to detect both known and unknown network attacks. Their research highlights the rapid development of cybercrime and the limitations of traditional intrusion detection systems that rely on rule-based techniques. By utilizing DNNs the proposed system achieves binary classification and multiclass classification. The researchers used the KDD Cup 1999 dataset which contains millions of records representing different attack scenarios. Their DNN model, employing ReLU as the activation function and Adam optimizer, achieved a high accuracy of 99.98% in detecting intrusions demonstrating the efficiency of DNNs for network security.

- **Milosevic et al**.[16]The study investigates the efficacy of convolutional neural networks (CNNs) in a multiclass network intrusion detection system. The challenge addressed is the high imbalance present in the CICIDS-2017 dataset which encompasses 15 distinct traffic classes. The research aims to enhance detection rates across various attack types by employing two feature selection techniques and training multiple CNN model variations. The results reveal that while individual CNN models achieved high accuracy the macro average F1 measure

indicates room for improvement in detecting all traffic classes consistently. The proposed models demonstrate better performance compared to existing DNN models especially in identifying minority classes with limited input samples underscoring the potential of CNN architectures in cybersecurity applications.

- **Farhana et al**.[17] Farhana, Rahman and Ahmed contribute to the field of network security by proposing a deep neural network (DNN) model for intrusion detection in packet and flow-based networks. They begin by addressing the limitations of traditional intrusion detection methods such as their inability to handle big data efficiently and their reduced accuracy on imbalanced datasets. The researchers propose using DNNs to improve the accuracy of intrusion detection highlighting the capabilities of deep learning in handling large-scale network traffic. Their model is evaluated on the CICIDS2017 dataset achieving over 99% accuracy for both binary and multi-class classification tasks. This study demonstrates the potential of DNNs for accurate intrusion detection across various attack types.

- **Alzahrani and Alenazi**.[18] It propose a network intrusion detection system (NIDS) tailored for software-defined networks (SDNs) using tree-based machine learning methods. The study addresses challenges in traditional detection systems and leverages the NSL-KDD dataset to detect and classify network attacks. By selecting only five key features from the dataset, the proposed system achieves a high classification accuracy of 95.55% for multi-class attack detection. The performance of three algorithms—Decision Tree, Random Forest, and XGBoost—was evaluated with XGBoost outperforming the others in precision and recall.

- **Almutairi et al**.[19] Almutairi, Alhazmi and Munshi examine the performance of machine learning models for network intrusion detection. They highlight the limitations of traditional security mechanisms and introduce machine learning as a solution to enhance intrusion detection systems (IDS). The study employs several machine learning models including Random Forest, J48, Naïve Bayes and Support Vector Machines to detect attack patterns in network traffic. By evaluating these models on the NSL-KDD dataset the authors achieve high accuracy with Random Forest outperforming other models. The study concludes that machine learning offers a promising approach to improve IDS though the use of synthetic data and imbalanced classes presents some limitations.

- **Vinayakumar et al**.[20] explore the application of deep learning in intrusion detection systems (IDS) by leveraging a deep neural network (DNN) model.

The study focuses on network-based and host-based IDS aiming to improve the detection and classification of cyberattacks by identifying abnormal network behaviors. The research examines a range of datasets and demonstrates that DNN models outperform classical machine learning techniques achieving high accuracy in both binary and multi-class attack detection. The proposed system is scalable and adaptable for real-time monitoring of network traffic.

TABLE 2.1: Research work on intrusion detection system models

Year	**Author**	**Model**	**Dataset**	**Accuracy**	**Limitations**
2012	Koc*et al.[1]	Hidden Naive Bayes	KDDCup'99	99.02%	Imbalance class, outdated attack patterns, redundant records
2016	Alrawashdeh and Purdy.[2]	Restricted Boltzmann Machine(RBM), Deep Belief Network	DARPA KDD-CUP'99	97.9%	Redundant information, high dimensional data, local optima during training
2020	Novo et al.[3]	DT, RF,SVM, XGBoost, Multi-Layer Perceptron Neural Networks (MLPNN), LSTM	UNSW-NB15	87.6%	Increased processing time, added complexity, reliance on the quality of the training dataset

Year	Authors	Model	Dataset	Accuracy	Limitations
2017	Zhao et al.[4]	Deep Belief Network (DBN, Probabilistic) and Neural Network (PNN)	KDD CUP 1999	98.28%	Computational intensity, data labelling dependency, complex model tuning
2019	Ahmim et al.[5]	REP Tree, JRip algorithm, Forest PA	CICIDS2017	96.665%	Limited model diversity, increased computational overhead, vulnerability to adversarial attacks
2018	Radford et al.[6]	LSTM Recurrent Neural Networks	ISCX IDS	82%	Computational complexity, sensitivity to hyperparameters, limitations in adapting to emerging threats
2019	Pektas and Acarman.[7]	CNN and RNN	CTU-13 and ISOT	99.3%	Complexity of model architecture, lack of generalization, interpretability hurdles

Year	Authors	Model	Dataset	Accuracy	Limitations
2018	Nguyen et al.[8]	Convolutional Neural Network (CNN)	KDDCup 99	99.87%	Accuracy for different attack types, large execution time, optimization challenges and dataset specificity
2018	Shone et al.[9]	Nonsymmetric deep autoencoder (NDAE)	KDD Cup '99, NSL-KDD	97.85%	Limited class performance, challenges in novel threats, issue in scalability
2016	Farnaaz and Jabbar.[10]	Random Forest classifier	NSL-KDD dataset	99.67%	Limited feature selection, dataset specificity, limited evaluation metrics
2019	Elmasry et al.[11]	DNNs, LSTM networks, Gated Recurrent Units (GRUs) and DBNs	KDD CUP 99, NSL-KDD, CICIDS2017	77.5%	Imbalanced datasets, high training time, hyperparameter tuning challenges
2020	Shah et al.[12]	Neural Networks, Decision Trees, k-NN, Random Forest, SVM	Advanced Security Network Metrics and Tunnelling Obfuscations dataset	95%	Models prone to overfitting, Poor performance of SVM on multiclass classification

Year	Authors	Model	Dataset	Accuracy	Limitations
2021	Acharya et al.[13]	Random Forest, J48, Naïve Bayes, Bayesian Network, Bagging, AdaBoost	KDD99, UNSW-NB15, CICIDS2017	93.8%	Lower performance in multiclass classification, challenges with imbalanced datasets
2021	Ashiku et al.[14]	CNN with hyperparameter tuning	UNSW-NB15	95%	Lower detection rates for underrepresented attack types, data redundancy, overfitting risks
2021	Maithem et al.[15]	Deep Neural Network (DNN) with ReLU activation and Adam optimizer	KDD Cup 1999	99.98%	High computational cost, dataset may not fully represent real world network traffic
2024	Milosevic et al.[16]	CNN-based IDS	CICIDS-2017	99.4%	Difficulty in detecting all traffic classes, high imbalance in data
2020	Farhana et al.[17]	Deep neural network (DNN)	CICIDS-2017	99%	Difficulty detecting rare attacks, limited performance on low-numbered instances

Year	Authors	Model	Dataset	Accuracy	Limitations
2021	Alzahrani and Alenazi.[18]	XGBoost, Random Forest, Decision Tree	NSL-KDD	95.55%	Struggles with encrypted traffic, high traffic volume, limited visibility inside the host machine
2022	Almutairi et al.[19]	Random Forest, Naïve Bayes, J48 (C4.5) and Support Vector Machines (SVM)	NSL-KDD	97.9%	Imbalanced classes in the dataset, Dataset consists of synthetic traffic
2019	Vinayakumar et al.[20]	Deep Neural Networks (DNN)	KDDCup 99, NSL-KDD, UNSW-NB15, Kyoto, WSN-DS, CICIDS 2017	99.45%	High false positive rates, real-world traffic patterns not fully represented, imbalanced class distribution

In our project, we will compare different deep learning models to see how well they detect network intrusions. We will test these models on binary classification and multi-class classification. By comparing the accuracy of the models in both approaches, we want to find out which method works best for our project. This way we can make sure our solution is as effective and reliable as possible in detecting and preventing cyber attacks.

2.2 Proposed methodology

The proposed system will involve the following steps:

- **Data Collection:** Publicly available datasets like CICIDS2017 will be used offering real-world network traffic including attack and benign data.
- **Data Pre-processing:** Irrelevant data will be removed and key features of the network traffic will be extracted for model training
- **Model Training:** Deep learning models such as CNNs, RNNs and DNNs etc will be used to train models on labeled datasets optimizing them for accurate detection of network anomalies.
- **Model Evaluation:** The models will be evaluated using accuracy, precision, recall and F1-score to ensure robustness and performance compared to traditional methods.
- **Deployment:** The trained models will monitor real-time network traffic classifying it as normal or malicious and integrating with existing security infrastructure.
- **Continuous Improvement:** The system will learn from new data updating models to remain effective against evolving threats.

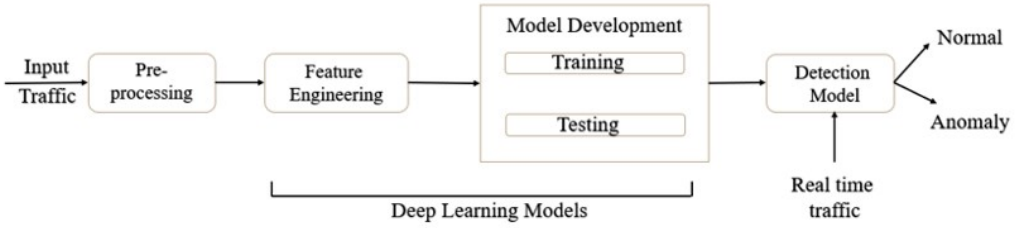

FIGURE 2.1: Proposed Methodology for Intrusion Detection System

Our proposed project will be flexible offering a comprehensive solution for network intrusion detection. It will be effective and reliable in detecting attacks, using deep learning to enhance accuracy, improve detection efficiency and adapt to emerging threats.

2.3 Data Collection Techniques

The following points outline the data gathering approach:

- We will use the CICIDS2017 dataset[1], which is a recognized and reliable source for evaluating intrusion detection systems.
- The dataset includes both benign and malicious network traffic offering comprehensive data for analysis.
- Attacks were executed over a period of five days starting from July 3, 2017.
- It contains a variety of attack types including Port Scan, DOS (Denial of Service), DDOS (Distributed Denial of Service), Web Attacks, Botnet Intrusions etc.
- We will capture key elements such as traffic interactions and metadata from the dataset.
- The data will be stored in an appropriate format to ensure easy pre-processing and feature extraction for machine learning models.

The CICIDS2017 dataset was chosen because it is very similar to real-world network environments with both normal and attack traffic. It provides a wide range of attack types which helps in thoroughly testing the intrusion detection system. This dataset is popular in the research community because it is well-documented and covers many types of attacks making it a great standard to evaluate how well different detection methods work.

2.4 Experimental Design

This experimental design outlines the approach to evaluating the Intrusion Detection System (IDS) for detecting network intrusions using deep learning models. The primary objective is to assess the IDS accuracy in identifying malicious activities in real-time network traffic. The experiments will utilize the CICIDS2017 dataset which contains various types of network attacks such as DDOS, Port Scans and Botnet intrusions etc.

The testing environment will simulate real-world network conditions by using captured network traffic data from the CICIDS2017 dataset. The dataset will be split into training and testing sets ensuring the system is tested on unseen data. Multiple deep learning models including Convolutional Neural Networks (CNNs), Recurrent

[1] https://www.unb.ca/cic/datasets/ids-2017.html

Neural Networks (RNNs) and Deep Neural Networks(DNNs) etc will be evaluated for their effectiveness in identifying and classifying network anomalies. Performance will be measured using key metrics such as accuracy, precision, recall and F1-score with confusion matrices employed to analyze false positives and false negatives. To further evaluate model performance, Receiver Operating Characteristic (ROC) curves will be plotted and the AUC (Area Under the Curve) will compare each models ability to distinguish between normal and malicious traffic.

Cross-validation will be applied to ensure model robustness and prevent overfitting. Hyperparameter tuning will optimize the performance of each model and models will be tested using real-time network traffic to assess their real-world applicability. The success of the experimental design will be determined by the system ability to reliably detect intrusions, minimize false alarms and adapt to new and previously unseen attack types.

References

[1] Koc, Levent, Thomas A. Mazzuchi, and Shahram Sarkani. "A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier." Expert Systems with Applications 39.18 (2012): 13492-13500.

[2] Alrawashdeh, Khaled, and Carla Purdy. "Toward an online anomaly intrusion detection system based on deep learning." 2016 15th IEEE international conference on machine learning and applications (ICMLA). IEEE, 2016.

[3] Larriva-Novo, Xavier, et al. "An approach for the application of a dynamic multiclass classifier for network intrusion detection systems." Electronics 9.11 (2020): 1759.

[4] Zhao, Guangzhen, Cuixiao Zhang, and Lijuan Zheng. "Intrusion detection using deep belief network and probabilistic neural network." 2017 IEEE international conference on computational science and engineering (CSE) and IEEE international conference on embedded and ubiquitous computing (EUC). Vol. 1. IEEE, 2017.

[5] Ahmim, Ahmed, et al. "A novel hierarchical intrusion detection system based on decision tree and rules-based models." 2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS). IEEE, 2019.

[6] Radford, Benjamin J., et al. "Network traffic anomaly detection using recurrent neural networks." arXiv preprint arXiv:1803.10769 (2018).

[7] Pektaş, Abdurrahman, and Tankut Acarman. "Deep learning to detect botnet via network flow summaries." Neural Computing and Applications 31.11 (2019): 8021-8033.

[8] Nguyen, Sinh-Ngoc, et al. "Design and implementation of intrusion detection system using convolutional neural network for DoS detection." Proceedings of the 2nd international conference on machine learning and soft computing. 2018.

[9] Shone, Nathan, et al. "A deep learning approach to network intrusion detection." IEEE transactions on emerging topics in computational intelligence 2.1 (2018): 41-50.

[10] Farnaaz, Nabila, and M. A. Jabbar. "Random forest modeling for network intrusion detection system." Procedia Computer Science 89 (2016): 213-217.

[11] Elmasry, Wisam, Akhan Akbulut, and Abdul Halim Zaim. "Empirical study on multiclass classification-based network intrusion detection." Computational Intelligence 35.4 (2019): 919-954.

[12] Shah, Ajay, et al. "Building multiclass classification baselines for anomaly-based network intrusion detection systems." 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA). IEEE, 2020.

[13] Acharya, Toya, et al. "Efficacy of machine learning-based classifiers for binary and multi-class network intrusion detection." 2021 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS). IEEE, 2021.

[14] Ashiku, Lirim, and Cihan Dagli. "Network intrusion detection system using deep learning." Procedia Computer Science 185 (2021): 239-247.

[15] Maithem, Mohammed, and Ghadaa A. Al-Sultany. "Network intrusion detection system using deep neural networks." Journal of Physics: Conference Series. Vol. 1804. No. 1. IOP Publishing, 2021.

[16] Milosevic, Marija, Vladimir Ciric, and Ivan Milentijevic. "Intrusion Detection System for Multiclass Detection based on a Convolutional Neural Network." 2024 IEEE 22nd Mediterranean Electrotechnical Conference (MELECON). IEEE, 2024.

[17] Farhana, Kaniz, Maqsudur Rahman, and Md Tofael Ahmed. "An intrusion detection system for packet and flow based networks using deep neural network approach." International Journal of Electrical and Computer Engineering (2088-8708) 10.5 (2020).

[18] Alzahrani, Abdulsalam O., and Mohammed JF Alenazi. "Designing a network intrusion detection system based on machine learning for software defined networks." Future Internet 13.5 (2021): 111.

[19] Almutairi, Yasmeen S., Bader Alhazmi, and Amr A. Munshi. "Network intrusion detection using machine learning techniques." Advances in Science and Technology Research Journal 16.3 (2022): 193-206.

[20] Vinayakumar, Ravi, et al. "Deep learning approach for intelligent intrusion detection system." Ieee Access 7 (2019): 41525-41550.