World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(REVIEW ARTICLE)

The detailed investigation paper on network invasion awareness

P Kamakshi Thai [1], Akshitha Allam [2,*], Pranay Sai Gabbula [2] and Yagnesh Kannam [2]

[1] *Assistant Professor, Department of CSE (Artificial Intelligence & Machine Learning), ACE Engineering College, Hyderabad, Telangana, India.*
[2] *IV B. Tech students Department of CSE (Artificial Intelligence & Machine Learning), ACE Engineering College, Hyderabad, Telangana, India.*

World Journal of Advanced Research and Reviews, 2024, 21(03), 2448–2453

Publication history: Received on 29 January 2024; revised on 27 March 2024; accepted on 29 March 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.21.3.0782

Abstract

Intrusion detection is critical for network security, with deep learning-based algorithms gaining traction. This project introduces NIDS - CNNLSTM, a model designed for the Industrial Internet of Things wireless sensing environment. It effectively identifies network traffic data, ensuring Industrial Internet of Things equipment and operations remain secure. Trained on the NSL_KDD data set, it exhibits strong convergence and performance across three data set, accurately classifying traffic types. Comparative analysis underscores NIDS - CNNLSTM efficacy enhancements. Experimental results validate increased detection rates, classification accuracy, and reduced false alarms, making it suitable for Industrial Internet of Things varied network data scenarios.

Keywords: Network Intrusion Detection; Deep Learning; CNN; LSTM.

1. Introduction

NIDS - CNNLSTM, a Deep Learning-based Network Intrusion Detection Classification Model, is crucial for ensuring network information security. However, the exponential growth of internet business has led to an increase in network traffic types and complexity, posing significant challenges to NIDS - CNNLSTM detection. It addresses a classification problem and enhances classifier performance to identify malicious traffic effectively, thereby improving intrusion detection accuracy. While deep learning methods are commonly used for this purpose, they often struggle with feature selection and solving massive intrusion data classification problems due to their shallow learning nature. Intrusion detection systems are vital for detecting malicious network traffic, especially in scenarios where preventive techniques fail. As security threats become capturing the essence of an advanced Intrusion Detection System (IDS), must evolve Perpetually to word counter evolutionary security challenges. Network data, being diverse, generates high- dimensional, multiple modes, and temporal data, making it suitable for big data analytic. Artificial intelligence, particularly deep learning techniques, has demonstrated success in handling heterogeneous data, uncovering unseen patterns, revealing hidden correlations, and gaining new insights. These techniques enable the recognition of new and unseen attack patterns.

2. Literature review

The research centered on a thorough investigation into the effectiveness of various methodologies within the domain. By scrutinizing pertinent research papers, the aim was to assess a multitude of approaches and techniques employed in these areas. This process sought to reveal the nuanced intricacies and advancements within the field.

* Corresponding author: Akshitha Allam

Abdelouahid Derhab et al. [1] Present an Intrusion Detection System for the Internet of Things based on Temporal Convolution Neural Network (TCNN) and efficient feature engineering. In the era, the vast amount of data traffic from connected objects fuels big data analytic, enabling the discovery of unseen patterns and identification of anomalous traffic. Five key design principles are identified for developing a deep learning-based Intrusion Detection System (IDS). Building on these principles,TCNN is proposed, integrating Temporal CNN augmented to be balanced synthesis imbalanced data sets. High performance is achieved by TCNN in the initial epochs, indicating that 15 epochs are sufficient. Furthermore, training and validation results indicate absence of over fitting. while deep learning models surpass LR and RF with some accuracy results exceeding 99.99%. Although TCNN slightly outperforms LSTM and CNN in effectiveness metrics, all deep learning models demonstrate good results even without data set balancing. Applying SMOTE -NC oversampling leads to a negligible decrease in TCNN and LSTM effectiveness, while CNN 's effectiveness slightly improves.Comparative analysis encompasses utilizing Bot - Internet of Things data set.

Mrutyunjaya Panda et al. [2] Propose a Network Intrusion Detection System using machine learning, a critical aspect of system defense. IDS collect network traffic data to safeguard networks, and machine learning methods aid in preemptive intrusion detection. The paper presents approaches Intelligent Fusion Ensemble of classifiers using J48 and NB with Ada Boost (AB). Evaluation is done on the NSL - KDD data set, a variant of the KDD Cup 1999 data set. The classification involves passing the input vector through each arbor for thicket where foliage votes for group, thicket selects the category with majority ballots. Bagged trees internally estimates error during construction, eliminating the need for separate cross-validation or test sets, by sampling with replacement, leaving out about one-third of instances for each tree's construction.

Vaishnavi Sivagaminathan et al. [3] Introduce an Intrusion Detection System (IDS) for wireless sensor networks utilizing computational intelligence techniques.NIDS is pivotal in identifying malicious network connections via traffic analysis, despite its resource-intensive nature. The system heavily relies on data extraction and machine learning for anomaly detection, emphasizing the importance of feature selection. To address this, the study employs Particle Swarm Optimization (PSO) for attribute selection in NIDS. The developed system aims to detect any malicious or unusual network behavior, ensuring data confidentiality.The research utilizes both network infrastructure and simulation data sets obtained using Wire shark and Cisco Packet Tracer.Wired mesh setup with sextet MCU linked devices, including portable computer, Wi-Fi hots pot, is utilized for capturing communication packets. PSO is integrated as an optimization strategy to enhance the performance of machine learning classifiers.The study involves building simulated LAN networks with Cisco Packet Tracer and capturing network activity with Wire shark. Finally, a comparative analysis is presented in the results.

Feifei Hu et al. [4] Introduce a novel network traffic classification model that integrates attention mechanisms and spatial-temporal features to tackle the challenges associated with classifying encrypted application traffic. While earlier studies primarily focused on mapping network traffic to different encrypted applications, they often overlooked the underlying traffic patterns. To bridge this gap, the proposed model combines Recurrent Memory Network for analyzing persistent grid streams and identifying sequential pattern characteristics. Additionally, the Convolution Network employed capturing complex topological arrangements. This is further enhanced by the Squeeze and Excitation (SE) module, which weights and redistributes these features to extract key spatial features of the network flow. CNN excels at extracting micro-level structures and attributes extracted from the data set. Deep Traffic Analyzer is capable of recognizing specific packet-level patterns that differentiate applications. This capability enables the identification of identifying traits of harmless Drop box data packets, including transmission unit magnitudes, message content configurations, and protocol behavior, thereby facilitating Improving feature representations using SE: The SE module boosts the model's capacity for representation adjusting trait maps. Adjusts dynamic coefficients to emphasize highlighting distinctive characteristics while attenuating irrelevant ones.Refinement leads to improved distinguishing between harmless and potentially harmful network activity, ultimately lowering erroneous detection.

Leila Mohammadpour et al. [5] Conducted an extensive survey on CNN-based Network Intrusion Detection, highlighting the increasing significance of securing Internet networks due to the widespread use of Internet applications. Intrusion Detection Systems (IDS), leveraging artificial intelligence (AI) techniques, play a pivotal role in ensuring network security. Among various AI branches, deep learning (DL) algorithms, particularly convolution neural networks (CNN), have emerged as powerful tools in IDS. CNN, designed to handle complex data, surpasses the limitations of traditional machine learning methods and finds extensive application in IDS. The survey systematically categorizes different CNN-based IDS approaches, elucidating their functionality and contributions. Various attributes such as data set, the components, input format, assessed criteria, efficacy, feature abstraction, and classification technique are juxtaposed. However, due to the utilization of diverse data sets in CNN- based IDS research, comparing experimental outcomes poses a challenge. Consequently, the study conducts an empirical

experiment to evaluate different approaches using standardized data sets, presenting detailed comparative analyses. CNN is a deep learning (DL) model adept at processing data, especially images, inspired by the structure of the animal visual cortex. It dynamically learns spatial hierarchies of features across different levels, enabling efficient handling of complex patterns. CNN displays demonstrated effective in multiple assignments like face identification,entity recognition ,traffic sign recognition, finding notable applications in robotics and autonomous vehicles. A critical aspect of CNN is its ability to minimize the number of parameters in an artificial neural network (ANN). This drives developers and researchers to explore larger models capable of tackling intricate tasks beyond the capabilities of traditional ANN.

Recurrent Neural Networks (RNN) are neural networks characterized by connections between neurons' outputs and inputs, allowing them to retain and utilize information across sequential steps. In Intrusion Detection Systems (IDS), RNN are used to extract temporal correlations, capturing patterns related to security attacks and malicious behaviors (temporal features). Conversely, Convolution Neural Networks (CNN) are employed to extract spatial features, complementing the temporal aspect provided by RNN.

J. Olamantanmi Mebawondu et al. [6] Highlight the transforms impact of the Internet on social, political, and economic structures, transcending geographical boundaries and facilitating increased business transactions. However, this heightened connectivity has also led to a surge in intruders, necessitating the deployment of Intrusion Detection Systems (IDS) alongside Intrusion Prevention Systems to safeguard computer resources. With modern networks generating terabytes of traffic within seconds, traditional rule-based approaches struggle to effectively analyze this data. Consequently, researchers are turning to data mining techniques for intrusion detection, prioritizing accuracy and relevant feature selection to enhance detection rates. The paper introduces a lightweight IDS knowledge-driven selection multiple layer linear classifier. Split information is applied to words select pertinent characteristics of attack and normal data transmission before sorting through the nerve- related Model. Observations from the UNSW-NB15 intrusion detection data set, leveraging chosen a couple of dozen variables exhibit top-tier selection procedure suggesting suitability of the lightweight IDS for real-time intrusion detection.Performance evaluation parameters, encompassing Error of Commission, Error of Omission, Correct Acceptance and Correct Rejection. corroborate efficacy across every data sets.Plot depicting number could be epochs achieved by the ANN reveals its optimal performance of 76.96% at epoch 200.

Muhammad Ali et al. [7] Address the escalating demand for reliable and accurate network defense mechanisms due to the proliferation of communication between networked devices. Network Intrusion Detection Systems (NIDS), pivotal for identifying malicious or anomalous network traffic, play a vital role in network defense. This research targets the challenges encountered by anomaly-based NIDS. Initially, the authors identify constraints in legacy NIDS data sets, including the recent CICIDS2017 data set, leading to the development of their novel data set, CIPMAIDS2023-1. Subsequently, they suggest a stacking-based ensemble approach that harmonizes various models. Surpasses current leading-edge. The research entails implementing variousattack scenarios alongside innocuous user activity over a network configuration crafted using Graphical Network Simulator-3 (GNS-3). Key flow features are extracted employing CIC Flow Meter for each attack and scrutinized to analyze their behavior.Various machine learning methodologies including KNN, SVM, and Random Forest, are applied to the extracted features.Outcomes demonstrate meta- learner ensemble technique achieves highest weighted F1- score of 98.24%. For machine learning, KNN, SVM, and Random Forest are employed as base estimators, trained for 150 iterations with a batch size of 24. XGBOOST serves as the meta-model, yielding Ultimately, the same model achieved the optimal performance on the test data is trained on merged training and validation data and evaluated using the validation set, utilizing stacking-based ensemble approach.

Mr. Mohit Tiwari et al. [8] Define an Intrusion Detection System (IDS) as a tool or software application tasked with monitoring network or system activities to identify any malicious behavior. They emphasize the critical importance of secure communication and data protection in light of the internet substantial growth and widespread usage. In today's environment, hackers utilize a variety of attacks to breach security measures and gain access to valuable information, highlighting the need for employing multiple intrusion detection techniques, methods, and algorithms to mitigate these threats.Goal to conduct comprehensive on threat recognition.Host-based Threat Detection System (HIDS) is designed to trigger alerts in response to changes such as file attribute modifications, new file creations, or existing file deletions. Notably, Network-based Intrusion Detection Systems (NIDS) have the capability to access encrypted information as it travels through the network, distinguishing them from HIDS.Tools of Intrusion Detection:Snort: Lightweight, community-developed software that utilizes a adaptable traffic specification language originating originating from an network address.It analyzes protocols to record packets in a human-readable format.OSSEC - HIDS: Free, open-source software compatible with major operating systems. It operates on a Client/Server-based architecture, allowing it To dispatch OS logs to the server for analysis and scrutiny data vault.

Mohammad Sazzadul Hoque et al. [9] Highlight the paramount importance of upholding high-level security to guarantee the safe and trusted communication of information among diverse organizations. However, the secure transmission of data over the internet and other networks is persistently threatened by intrusions and misuse.As a result, Intrusion Detection Systems (IDS) have emerged as indispensable elements in computer and network security. Despite the adoption of various approaches in intrusion detection, none of the systems implemented thus far are entirely flawless, underscoring the need for continuous improvement efforts. In this context, the authors introduce Network Security Monitor (NSM) that employs inheritable algorithms, effectively recognizing multiple network intrusions.They provide detailed discussions on the parameters and evolution processes for GA and implement them accordingly. This approach harnesses evolutionary theory streamline network data and simplify analysis.Evaluate system's efficiency, they utilize the KDD99 benchmark data set, resulting in a reasonable detection rate. The KDD99 data set, originating in 1999, served as the foundation The Third Global Data Mining Tools Challenge which coincided the Fifth KDD Conference. This Contest geared towards crafting web breach detector—a prognostic model tasked with discerning separating harmful from legitimate connections. Genetic algorithms widely acknowledged for their ability to produce efficient problem-solving techniques. Evolutionary models emulate Darwinian selection,where organism that survive through environmental adaptation reproduce, progress to future progeny. Essentially, genetic algorithms simulate "survival of the fittest" within participants in consecutive subsequent iterations to address problems. In each generation, there is a group of individuals, where each individual corresponds to a location in the search space and represents a potential solution. These individuals are depicted as strings of characters, integers, floats, or bits, akin to chromosomes.

Dewan Md. Farid et al. [10] Introduce a novel learning algorithm for adaptive network intrusion detection, utilizing a naive Bayesian classifier and decision tree. This algorithm aims to achieve balanced detection while maintaining maintaining an acceptable threshold of false positives for various categories of network intrusions. It also addresses the complexity of detection models by eliminating redundant attributes and contradictory examples from training data. Additionally, the proposed algorithm tackles challenges in data exploration, encompassing continuous data management attributes, addressing absent attribute values and mitigating training data noise. Given the extensive quantities of security audit data and the ever-changing nature pertaining to intrusion behavior, data mining-driven intrusion detection techniques have gained prominence in recent decades for analyzing network-based traffic data and host-based data. However, several issues persist in current Intrusion Detection Systems (IDS). To assess the efficacy of their proposed algorithm, the authors conducted experiments using the intrusion detection benchmark data set from KDD99 data set. Solution demonstrate the algorithm proposed achieves elevated detection rates significantly reduces erroneous alarms various forms of network breaches, even with restricted computational power.This paper further addresses challenges in knowledge discovery, experimental findings utilizing the KDD99 benchmark intrusion detection, including tasks like managing continuous attributes, addressing missing attribute values, and mitigating noise in training data set validate the algorithm under consideration outperforms existing methods in terms of both detection rates (DR) and false positive reduction (FP).

Table 1 Overview of Techniques and Methodologies for Network Invasion Awareness, highlighting the pros and cons of each approach**.**

Papers	Year	Technique/Methodology	Pros	Cons
[1]	2020	Deep learning in TCNN	Identifies key design principles for developing deep learning-based IDS for Internet of Things, aiding in system development.	Limited discussion on growth potential and generalized of TCNN beyond Bot - Internet of Things, data set.
[2]	2011	Deep learning in CNN (Convolution neural network)	Evaluates performance using NSL -KDD data set for comprehensive bench marking.	Limited discussion on growth potential and generalized.
[3]	2023	Deep learning in LSTM (long short term memory)	Effective feature selection improves the clarity and efficiency of analyzing movement patterns.	Implementation may require significant time and expertise.
[4]	2023	Machine learning in Naive Bayes,SVM.	Automatically constructs mapping relationships between	Evaluation focused on specific data sets and scenarios, may not

			network flow and labels, reducing manual intervention.	fully represent real-world network traffic.
[5]	2022	Deep Learning in CNN (Convolution neural network)	Highlights the importance of IDS employing AI methods, particularly deep learning algorithms like CNN.	Dependency on specific data sets may limit the generalized of findings to diverse network environments.
[6]	2020	Deep Learning in ANN (Artificial Neural Network)	Acknowledges the transformers impact of the Internet on social, political, and economic structures.	Implementation and optimization of the proposed IDS may require expertise and resources.
[7]	2023	Machine learning Intrusion detection system. Denial of service Ensemble-based learning .CICIDS2017.	Proposes a stacking-based an ensemble approach that currently stands out in the field of Network Intrusion Detection Systems (NIDS),demonstrating promising results.	Further validation and testing may be needed to assess the performance of the approach in various network intrusion detection scenarios.
[8]	2017	Deep Learning-Based IDS leverage neural network architectures,CNN.	Provides a comprehensive overview of intrusion detection, including definitions, methods, attacks, tools, and techniques.	Dependency on specific data sets or scenarios may limit the generalized of findings.
[9]	2012	Machine Learning in Decision Tree, Naive Bayesian classifier.	Recognizes the importance of maintaining high-level security for safe data communication between organizations.	Genetic algorithm- based IDS may require significant computational resources for implementation and optimization.
[10]	2010	Machine Learning Models- Detection Rate, False Positive, Network Intrusion Detection.	Achieves balanced detection and maintains acceptable levels of false positives for different types of network attacks.	Dependency on specific data sets like KDD99 may limit the generalized of findings to diverse network environments.

2.1. NIDS - CNNLSTM network intrusion detection

Convolution Neural Network Long Short-Term Memory Network Intrusion Detection System, is a sophisticated classification model rooted in deep learning principles techniques, tailored for the Industrial Internet of Things environment. Its primary goal is to effectively identify and differentiate network traffic data to ensure equipment and operational security. Trained using NSL_KDD data sets, the model showcases high accuracy rates and convergence while classifying various types of traffic (CNN) to identify intrusion information. It notably enhances performance compared to existing methods, achieving high detection rates, classification accuracy, and low false alarm rates. Capable of detecting five types of intrusion attacks, it utilizes multiple features for training, rendering it well-suited for large-scale and multiple- scenario network data in the Internet of Things environment.

3. Conclusion

This project proposes NIDS - CNNLSTM as a solution to address the challenges with limited spotting rates, precision in classification, but with elevated false positives rates encountered by legacy intrusion detection systems in the Industrial Internet of Things environment. NIDS - CNNLSTM enhances the capabilities of traditional models by improving and superimposing two-layer CNN and two-layer unidirectional LSTM layers. This architecture leverages CNN ability Utilizing geometric attributes along with LSTM . Performance assessed using the NSL_Knowledge Discovery in Databases data set, demonstrating good convergence and effectiveness in terms of validation accuracy and training loss. Furthermore, its applicability is validated through both binary classification and multiple classification scenarios.

Compliance with ethical standards

Acknowledgments

We would like to thank our guide Mrs.P.Kamakshi Thai for her support and guidance ,Assistant Professor (Artificial Intelligence & Machine Learning) and Mr.Shashank Tiwari , Assistant Professor, Project Coordinator and we profoundly thank Dr.Kavitha Soppari ,head of the Department CSE(Artificial Intelligence & Machine Learning) for her guidance and continuous support .

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Abdelouahid Derhab[1] ,Arwa Aldweesh[2] ,Ahmed Z. Emam[2] .Intrusion Detection System for Internet of Things Based on Temporal Convolution Neural Network and Efficient Feature Engineering Volume 2 (2020).

[2] Mrutyunjaya panda, Ajith Abraham and Swagatam Das. Article in Intelligent Decision Technologies November 2011. Network Intrusion Detection System: A Machine Learning Approach (2011).

[3] Vaishnavi Sivagaminath[1], Manmohan Sharma[1] and Santosh Kumar Henge[1].Intrusion detection systems for wireless sensor networks using computational intelligence techniques (2023).

[4] Feifei Hu[1], Situo Zhang[1],Xubin Lin[1] and Liu Wu[1]. EURASIP Journal on Information Security Network traffic classification model based on attention mechanism and spatial temporal features(2023).

[5] Leila Mohammadpour,Tech Chaw Ling and chee Sun Liew . A Survey of CNN - Based Network Intrusion Detection (2022).

[6] J. Olamantanmi Mebawondu , Olufunso D. Alowolodu and Jacob O. Mebawondu , .Network intrusion detection system using supervised learning paradigm (2020).

[7] Muhammad Ali[1,2] ,Mansoor-ul- Haque[1,2] and Muhammad Hanif Durad[1,2] .Effective network intrusion detection using stacking-based ensemble approach.International Journal of Information Security (2023).

[8] Mr Mohit Tiwari[1], Raj Kumar[2] and Akash Bharti[3].INTRUSION DETECTION SYSTEM International Journal of technical Research and Applications,Volume 5 ,Issue (2017).

[9] Mohammad Sazzadul Hoque[1], Md. Abdul Mukit[2] and Md. Abu Naser Bikas[3].AN IMPLEMENTATION OF INTRUSION DETECTIONSYSTEM USING GENETIC ALGORITHM. Vol.4, No.2, March (2012).

[10] Dewan Md. Farid[1],Nouria Harbi[2],and Mohammad Zahidur Rahman[3].COMBINING NAIVE BAYES AND DECISION TREE FOR ADAPTIVE INTRUSIONDETECTION. Volume 2, Number 2, April (2010).