WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Novel Machine Learning Approach for Detecting Various Attacks

Alka, Prof. Sumit Sharma

Department of Computer Science Engineering, Vaishnavi Institute of Technology & Science, Bhopal

ABSTRACT

In this paper, a *Convolutional Neural Network (CNN)* is used to investigate network intrusion detection. To determine how the attributes of the findings relate to one another, integrated folding and grouping processes are employed. For the intrusion samples to be correctly identified, the model should automatically ascertain the efficient qualities of the intrusion samples. Experimental tests with UNSW NB15 data sets suggest that the proposed model will significantly increase intrusion detection performance. To overcome the issue of the existing work, the suggested approach *CNN* is used to perform a dependency test based on distance correlation for medium/large size issues. A variety of studies were conducted utilizing the *Convolutional Neural Network* technique to perform *Multi-Class Classification.*

Keywords: *Feature Selection, Attribute, Optimal Subset, Classification, Support Vector Machine, Convolutional Neural Network.*

1. Introduction

IDS come in many varieties, and they are all categorized differently. This deployment technique divides intrusion detection systems (IDS) into three categories: distributed, centralized, and hybrid.

- Host-based IDS, or distributed IDS-With this distributed deployment technique, each node in the Internet of Things network is in charge of keeping an eye on and identifying potential threats. Consequently, every network node has the ID loaded on it. The IDS identify the attacks in a targeted manner. As the IDS deployed on each node, the qualities of the Internet of Things that are limited by resources are examined and optimized.
- Network IDS (Centralized IDS)-IDS are set up on a centralized router in this method. The centralized IDS quickly identify external attackers, which is the reason why data packets from the outside world penetrate the Internet of Things during the border router. Additionally, it detects attacks with ease and watches, analyzes, and drops harmful data packets. On the other hand, extensive monitoring and analysis of every internal node connected to the border router is necessary for the identification of internal attacks.
- Hybrid IDS-Selected nodes function as Distributed IDS in the hybrid IDS approach to find intrusions that may be traced back to nearby neighbors. Every watchdog has a different set of rules based on how the network parts behave. The patterns are identified from the monitored communications in accordance with the protective rule-sets in the centralized IDS.

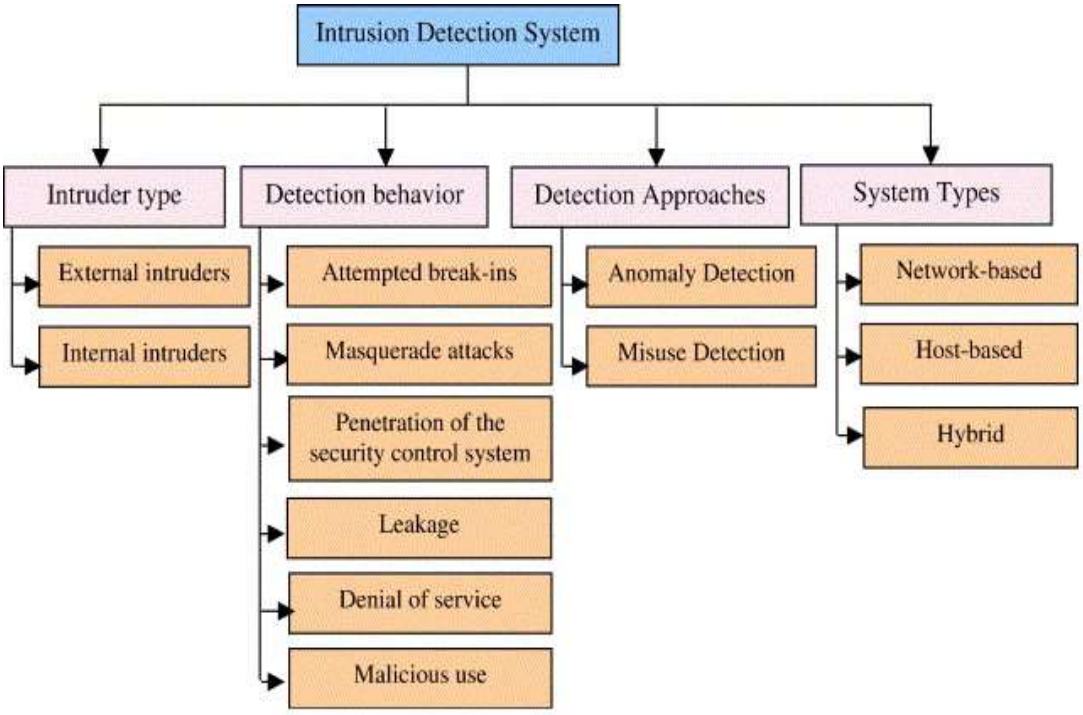

Figure 1-Intrusion Detection System [1].

2. IOT Layers

Three primary core layers-perception, transportation, and application can be used to effectively define and explain a broad Internet of things system. There are potential security flaws in each of these system layers that certain attacks could be able to take advantage of due to their distinct sets of tools. Each layer security concerns are investigated independently to provide innovative, workable, and reliable solutions.

1. **Perception Layer**

In order to facilitate data gathering and analysis for several widely used technologies, the initial layer is linked to real Internet of Things sensors (e.g. RFID, GPS, WSN). This layer comprises actuator and sensor systems that conduct measured data (such as temperature, movement, moisture, and so on) and characteristics like location searching. These are the primary security threats that this layer presents due to its distributed hierarchical structure and limited node resources:

- Physical assaults: These attacks target the actual hardware that makes up the Internet of Things; in order to be effective, the attacker has to be physically present either within or close to the IoT system.
- Denial of Service (DoS): These attacks use the nodes inadequate processing power to render them unusable.
- Routing Attacks: Hostile intermediate nodes (such those in a WSN) have the ability to change the correct routing patterns during the data collection and transmission process.
- Attacks on Data Transit: Various risks to the confidentiality and integrity of data during transmission (e.g., Man-in-the-Middle, Sniffing).

2. **Transportation Layer**

The pervasive accessibility environment provided by the transportation layer largely favors the perception layer. The purpose of this layer is to transfer data from the perception layer to any particular information processing system utilizing established communications infrastructure utilized by both networking devices (3G/LTE, WiFi), network services, and the both (i.e. Internet). [6] Provides a general overview of security issues in cellular connections, particularly cellular networks. According to this study, an IP-based LTE channel open and varied architecture raises more security issues than 3G networks do. The major security threats at this level often involve:

- When utilizing wireless technology, accurate path structural assaults and data passing.
- Denial-of-service (DoS) attacks: Because IoT networks are so diverse and complicated, the transportation layer is vulnerable to attacks.
- Data Transit Attacks: A range of threats against the authenticity and integrity of data while it is being transmitted between systems that are available or on a network.

3. Application Layer

The services that clients need are provided by the application layer. Customers that request such information, for example, to get temperature and air humidity readings from the web server. Because it makes it possible to provide superior smart services that go above and beyond what users anticipate, this layer is crucial to the Internet of Things. Furthermore, an Application Support Sub-layer (ASS), which claims to support all business service types and perform complex calculation and resource allocation, may be included in some middleware and cloud computing platforms. This layer may be used to achieve a number of Internet of Things applications, such as smart manufacturing, smart cities, and smart healthcare. The primary security threats in this layer are as follows:

- Data leakage: Using well-known service or program flaws, an attacker can quickly steal data (including user data, such as a user password).
- DoS attack: Attackers might stop the application or service from being available.
- Harmful code Injection: By common vulnerabilities, attackers can upload malicious programs that infect fetcher software.

3. Literature Review

Sangve et al. (2015) proposed the algebraic strategy for Anomaly Network Intrusion Detection Systems (ANIDS) on vast size datasets using the sensors developed, demonstrating methods for machine learning employing a variety of datasets. Fewer ANIDS make use of the NSL-KDD dataset, a revamped version of the well-known KDD Cup 99 dataset. The KDD Cup 99 dataset has been utilized by many ANIDS. When matched to the KDD99 dataset, the NSL-KDD dataset clearly outperforms it.

Qazanfari et al. (2012) employed a hybrid anomaly-based intrusion detection system that depends on both signature-based as well as anomaly-based methodologies. To enhance the performance of this scheme, it employs the following strategies: first, it pulls optimal data from the KDD data set using a feature extraction technique based on feature entropy; secondly, it employs a unique methodology to integrate the result of these two learning-based techniques. Finally, the detection strategy uses the KDD dataset to illustrate the efficacy of these hybrid systems. The simulation results demonstrate the KDD features that are suitable for discriminating between normal and abnormal traffic. This result indicates how successfully the method can detect DOS, Probe, U2R, and R2L attacks.

Zhu et al. (2012) created a unique intrusion detection approach (U-D methodology) that considers both received and published data. Using the new analytic approach, intrusion clues may well be recognized with greater efficacy and efficiency. The connections between these data points may give some intuitive clues for detecting significant incursions. Experiment results show that the approach is successful in terms of high detection.

According to *Chen et al. (2011)*, integrating rough sets with data mining to augment standard intrusion detection systems will boost detection effectiveness and reduce fake alerts. Data collection starts with categorization, then processing, which includes standardizing independent variables, and lastly discrete processing of nominal variables. The Pawlak feature weighted rough set technique may also be used to minimize the amount of variables in the result set by utilizing the property upward and downward approximated set's features. Association rules that fulfill a certain degree of confidence can be constructed using attribute reductions and then imported into the rule set. According to studies, the detection strategy that integrates data mining with rough sets improves detection efficacy by more than 20%. As the number of incursions rises, the detection rate approaches linearity.

4. Proposed Methodology

This section explained proposed work flowchart along with the explanation and working.

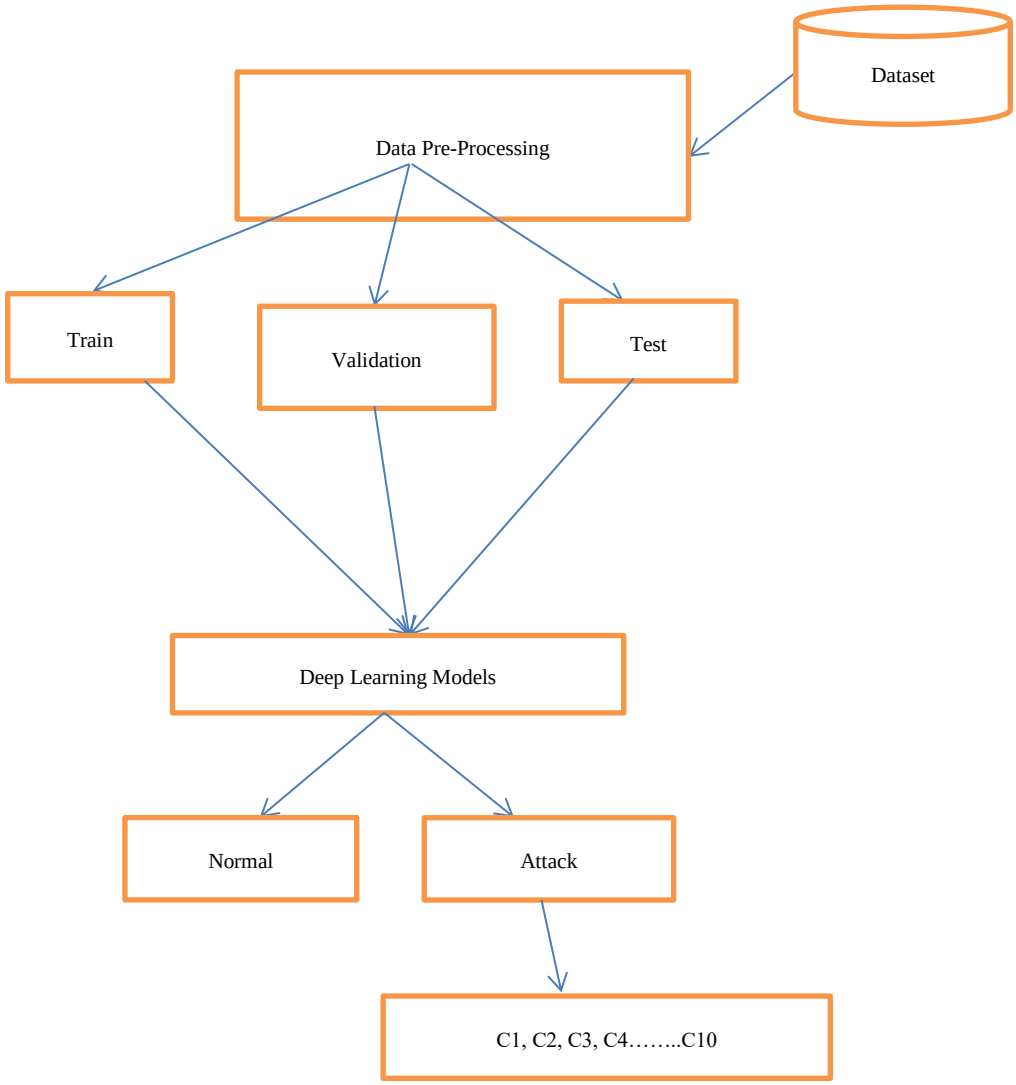

Figure 2-Proposed Model for Intrusion Detection.

4. 1 Intrusion Detection Model Explanation

A concept for identifying intrusion threats is shown in Figure 4.1, and it focuses on the Convolutional Neural Network method and outputs the LSTM. Figure 4.1 illustrates how the flowchart follows various phases:

• Phase 1: Data Pre-Processing: The UNSW NB15 data collection has a significant amount of automated preprocessing that enables to clean up sloppy data and convert conceptual information into numerical data.

• Phase 2: Training as well as Feature Extraction: Features data are prepared and extracted using the built-in CNN model.

• Phases 3: Keep Test: To recognize and compile the study's results, the Softmax classifier was applied.

5. Result Analysis

The experimental results demonstrate that our IDS has a high detection rate and accuracy for irregularities. The optional outcome is found by comparing the experiment results to the conventional approach. When compared to the advised course of action, the proposed work is more effective. Here we discuss the computation settings and the observed results.

1 Accuracy Analysis

The total number of results divided by the number of intruders is what is known as accuracy (ACC).

$$ACC = \frac{TP + TN}{TP + FP + FN + TN}$$

2 Detection Rate

On the other hand, the detection rate (DR) is the possibility that finds out the real intrusions from the given alarm.

$$DR = \frac{TP}{TP + FN}$$

3 Precision

Out of those optimistic assumptions, it provides information about the model's accuracy and how many of them are wrong. It is computed as:

$$Precision = \frac{TP}{TP + FP} \times 100\ (\%)$$

4 Recall

By classifying them as positive, it determines if any of the true positives the recommended model discovers are really collected:

$$Recall = \frac{TP}{TP+FN} \times 100\ (\%)$$

5 F1-Score

It serves as an example of how precise a test was. The Harmonic Mean (HM) of the accuracy and recall is the F1 score, which can have a maximum value of 1. The F1 score is based on:

$$F1score = 2 \times \frac{Precision \times Recall}{(Precision + Recall)} \times 100\ (\%)$$

Here,

TP: True Positive,

TN: True Negative,

FN: False Negative, and

FP: False Positive.

Table 1-Comparative analysis of multi-class model with other models.

In the above Table 1 comparative analysis of the proposed model algorithm with other algorithms is given where obtained results of accuracy is 85%, precision is 86%, recall is 84%, F1-Score is 81%.

Accuracy (%)

ML Method	Tr. AC (%)	Val. AC (%)	Test AC (%)	Precision (%)	Recall (%)	F1-Score (%)
ANN	79.91	79.61	75.62	79.92	75.61	76.58
LR	75.51	73.93	65.53	76.91	65.54	66.62
kNN	81.75	76.83	70.09	75.79	70.21	72.03
CNN-BiLSTM			85	86	84	81

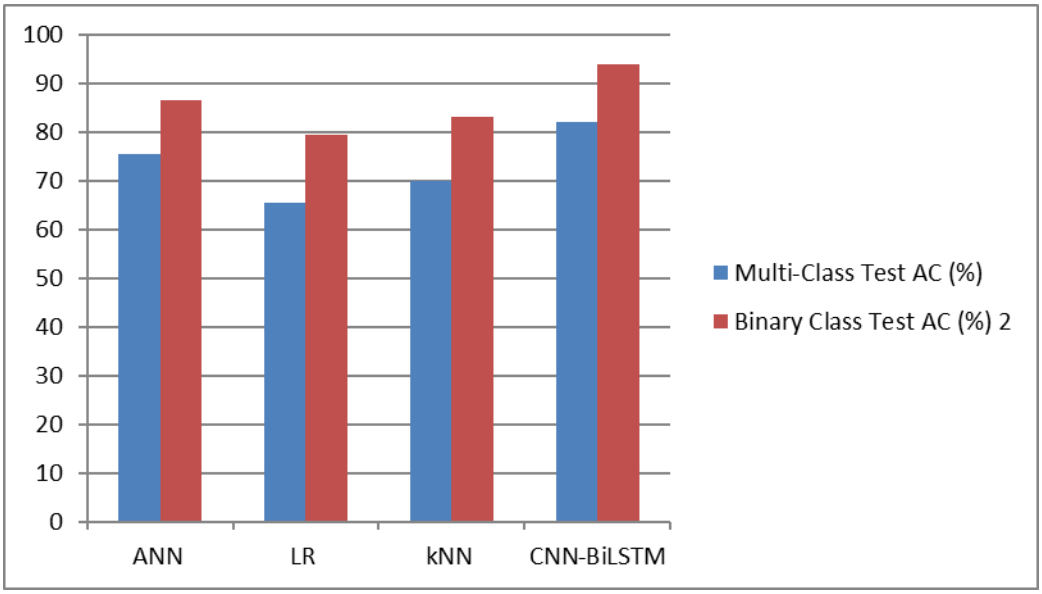

Figure 3-Repreentation of Precision (%) wrt CNN-BiLSTM.

6. CONCLUSION

This investigation of feature selection algorithms for a sizable survey demonstrates that the feature selection technique regularly raises the classifier's accuracy. Each feature selection process has benefits and drawbacks of its own. The dataset with more characteristics uses wrapper techniques, which provide less accuracy gain. Accuracy is decreased when a greater characteristic is included. Since each method behaves differently, it is impossible to use a single approach across several datasets. The accuracy of the categorization of various datasets is determined by feature selection algorithms. The feature selection algorithm must choose the pertinent characteristics and exclude the unrelated and inconsistent features that reduce the categorization algorithms' accuracy.

Below are the outcomes for multi-class model.

REFERENCES

[1] Lin, J. R., Talty, T., & Tonguz, O. K. (2015). On the potential of bluetooth low energy technology for vehicular applications. *IEEE Communications Magazine*, *53*(1), 267-275.

[2] Raza, S., Misra, P., He, Z., & Voigt, T. (2017). Building the Internet of Things with bluetooth smart. *Ad Hoc Networks*, *57*, 19-31.

[3] Collotta, M., Pau, G., Talty, T., & Tonguz, O. K. (2018). Bluetooth 5: A concrete step forward toward the IoT. *IEEE Communications Magazine*, *56*(7), 125-131.

[4] Fürst, J., Chen, K., Kim, H. S., & Bonnet, P. (2018, April). Evaluating Bluetooth low energy for IoT. In *2018 IEEE Workshop on Benchmarking Cyber-Physical Networks and Systems (CPSBench)* (pp. 1-6). IEEE.

[5] Hasan, K., Biswas, K., Ahmed, K., Nafi, N. S., & Islam, M. S. (2019). A comprehensive review of wireless body area network. *Journal of Network and Computer Applications*, *143*, 178-198.

[6] Nabila, A. (2019, April). A QoS based comparative analysis of the IEEE standards 802.15. 4 & 802.15. 6 in WBAN-based healthcare monitoring systems. In *2019 International conference on wireless technologies, embedded and intelligent systems (WITS)* (pp. 1-5). IEEE.

[7] Kim, T. (2018, April). A study of the Z-wave protocol: implementing your own smart home gateway. In *2018 3rd International Conference on Computer and Communication Systems (ICCCS)* (pp. 411-415). IEEE.

[8] Naidu, G. A., & Kumar, J. (2019). Wireless protocols: Wi-Fi son, Bluetooth, zigbee, z-wave, and Wi-Fi. In *Innovations in electronics and communication engineering* (pp. 229-239). Springer, Singapore.

[9] Lavric, A., & Petrariu, A. I. (2018, May). LoRaWAN communication protocol: The new era of IoT. In *2018 International Conference on Development and Application Systems (DAS)* (pp. 74-77). IEEE.

[10] Haxhibeqiri, J., De Poorter, E., Moerman, I., & Hoebeke, J. (2018). A survey of LoRaWAN for IoT: From technology to application. *Sensors*, *18*(11), 3995.

[11] Jalaian, B., Gregory, T., Suri, N., Russell, S., Sadler, L., & Lee, M. (2018, February). Evaluating LoRaWAN-based IoT devices for the tactical military environment. In *2018 IEEE 4th World Forum on Internet of Things (WF-IoT)* (pp. 124-128). IEEE.

[12] Mekki, K., Bajic, E., Chaxel, F., & Meyer, F. (2018, March). Overview of cellular LPWAN technologies for IoT deployment: Sigfox, LoRaWAN, and NB-IoT. In *2018 ieee international conference on pervasive computing and communications workshops (percom workshops)* (pp. 197-202). IEEE.

[13] Ayoub, W., Samhat, A. E., Nouvel, F., Mroue, M., & Prévotet, J. C. (2018). Internet of mobile things: Overview of lorawan, dash7, and nb-iot in lpwans standards and supported mobility. *IEEE Communications Surveys & Tutorials*, *21*(2), 1561-1581.

[14] Lavric, A., Petrariu, A. I., & Popa, V. (2019, August). SigFox communication protocol: The new era of IoT?. In *2019 international conference on sensing and instrumentation in IoT Era (ISSI)* (pp. 1-4). IEEE.

[15] Mekki, K., Bajic, E., Chaxel, F., & Meyer, F. (2018, March). Overview of cellular LPWAN technologies for IoT deployment: Sigfox, LoRaWAN, and NB-IoT. In *2018 ieee international conference on pervasive computing and communications workshops (percom workshops)* (pp. 197-202). IEEE.

[16] Aldahdouh, K. A., Darabkh, K. A., & Al-Sit, W. (2019, March). A survey of 5G emerging wireless technologies featuring LoRaWAN, Sigfox, NB-IoT and LTE-M. In *2019 International conference on wireless communications signal processing and networking (WiSPNET)* (pp. 561-566). IEEE.

[17] Osman, N. I., & Abbas, E. B. (2018, August). Simulation and modelling of LoRa and Sigfox low power wide area network technologies. In *2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)* (pp. 1-5). IEEE.

[18] Chere, M., Ngqondi, T., & Bembe, M. (2019, January). Wireless Low Power Area Networks in the Internet of Things: A Glimpse on 6LoWPAN. In *2019 International Conference on Electronics, Information, and Communication (ICEIC)* (pp. 1-10). IEEE.

[19] Pai, V., & Shenoy, U. K. K. (2019). 6LoWPAN—Performance analysis on low power networks. In *International Conference on Computer Networks and Communication Technologies* (pp. 145-156). Springer, Singapore.

[20] Al-Kashoash, H. A., Kharrufa, H., Al-Nidawi, Y., & Kemp, A. H. (2019). Congestion control in wireless sensor and 6LoWPAN networks: toward the Internet of Things. *Wireless Networks*, *25*(8), 4493-4522.