A Deep Reinforcement Learning Approach to Modelling an Intrusion Detection System Using Asynchronous Advantage Actor-Critic (A3C) Algorithm

JRIIE Journal of Research Innovation and Implications in Education
ISSN: 2520-7504
Center for Research Implications & Practice

Website:www.jriiejournal.com ISSN 2520-7504 (Online) Vol.6, Iss.1, 2022 (pp. 441 - 452)

A Deep Reinforcement Learning Approach to Modelling an Intrusion Detection System Using Asynchronous Advantage Actor-Critic (A3C) Algorithm

Junior Kiplimo Yego, Dr. Nicholas Kiget & Mr. Daniel Samoei
Department of Information Technology
Moi University, Kenya
Corresponding Author: Junkiy62@gmail.com

Abstract: *An increase in growth and use of the internet has also resulted in attacks evolving and more novel attacks having a devastating effect are witnessed. The Intrusion Detection System (IDS) is yet to achieve maximum success due to false positives and low detection. The purpose of the study was to determine the modelling of an intrusion detection system using the Asynchronous Advantage Actor-Critic (A3C) Algorithm. In this paper we look at the following: (i) To evaluate the current machine learning techniques being used in IDS, (ii) To determine the effectiveness of using the Asynchronous Advantage Actor-Critic algorithm in anomaly detection, (iii) To select the appropriate training data set and prepare for use on A3C. A conceptual study was done in looking at these objectives. The UNSW_TRAIN and UNSW_TEST were samples selected by purposive sampling from the whole population of UNSW-NB15 dataset. Analysis of the dataset was done using Python. Key findings were that anomaly detection approach is the best approach due to its ability to detect novel attacks. Also, there is need to continue research on intrusion detection and improve solutions to the problem of false positives and fully optimize on accuracy. The UNSW-NB15 dataset is comprehensive and so all the attack types should be used so as to accurately depict the intrusions and should selected attack types be used, feature selection should be done accurately so as to reflect modern attack types.*

Keywords: *IDS, Anomaly, Novel, Dataset, Adversarial, Attack*

How to cite this work (APA):

Yego, J. K., Kiget, N. & Samoei, D. (2022). A deep reinforcement learning approach to modelling an intrusion detection system using asynchronous advantage actor-critic algorithm. *Journal of Research Innovation and Implications in Education*, *6*(1), 441 – 452.

1. Introduction

With increased internet use both in online businesses and the soaring use of the Internet of Things technology, an exponential rise of internet activities has been witnessed. A Cyberthreat Defense Report by CyberEdge Group (2021) reported 86.2% of organizations that were surveyed were affected by a successful cyber-attack. A lot of research is being done in detecting behavioral anomalies in networks to make the internet safe.

Zimek et al. (2017) define anomaly detection, which is also known as outlier detection as “the identification of rare items, events or observations which raise suspicions by differing significantly from the majority of the data”. Intrusion Detection Systems (IDS) analyze data traffic and detect behavioral anomalies which pose threats to the network including malicious activity and policy violations by both system insiders and external intruders.

There are 2 types of IDS, Signature-based and anomaly based. While signature based rely on pattern recognition by having a database of known attack signatures for analysis, anomaly based uses statistical models built for normal traffic in the network and any deviation is an anomaly. This is good for detecting new attacks. For an IDS to be considered effective in terms of performance, it should be accurate in detecting intrusions by having a level of accuracy that is higher in classification and also achieving low false alarm rate.

Machine learning techniques serve a reinforcing role in securing networks from sophisticated attackers who bypass basic IDS as they utilize modifiable, extensible, and reproducible datasets for learning.

In unsupervised learning, due to the lack of existing classified labels, it assumes that most of the data is normal and smaller clusters are anomalies. This presents a false picture for high-dimensional data as it is typically sparse and so could have many small clusters which may not be anomalies hence leading to false alarms. Supervised methods on the other hand have labels for normal data and of known anomalies and its main weakness is that they cannot detect novel anomalies.

DRL approach applies the principles of reinforcement and deep learning to be able to create efficient algorithms capable of scaling to previously unsolvable problems by making use of its ability to learn from raw data as input.

1.2 Statement of the Problem

The current Intrusion Detection System is effective in identifying anomaly behaviors in network usage. However, it is still hindered from achieving maximum performance due to a huge number of false positives reported and overall low detections caused by huge amounts of network data in addition to the unbalanced distribution of anomaly and normal behaviors as supported by Jiadong et al. (2019). This severely limits the effectiveness of the IDS because the number of real attacks is always less than that of false alarms. The system is then human dependent since all alerts must be reviewed so that legitimate operations may be allowed. The imbalanced dataset will create a bias negatively impacting evaluations (Fossaceca et.al, 2015) and so weakening the performance of the IDS. There is, therefore, a need for a more thorough, accurate, and autonomous Intrusion detection System that will strengthen the weaknesses of the current IDSs. This study, therefore, responds to this need.

2. Literature Review

2.1 Evaluation of the Machine Learning Techniques Used in IDS

The IDS can be defined as a mechanism, whether a hardware or software that monitors a network or system to identify and mitigate attacks and policy violations. The main challenge being experienced in this information age is the unauthorized access to information which can then be used to harm the owner of the information or control the network security systems to perform illicit actions. Like street crime, with the growing number of human and digital targets, cybercrime is set to proportionally increase. This is due to an increase in connection technologies which include the cloud center data traffic which is increasing exponentially and confirmed by cisco to represent over 95% of data center traffic.) and projects a growth in number of devices connected to the IP networks to 29.9 billion devices by 2023 Cisco Annual Internet Report (2022).

Mitchel (1997) has the definition of Machine Learning as the study of computer algorithms that automatically improves through different experiences. This involves looking at new ways to perform various tasks without explicit programming hence they utilize training data to build models for decision making and prediction. The learning method is by looking for patterns in data so as to make accurate predictions based on the examples with the aim being to allow learning without human intervention and adjust actions accordingly. This is supported by Janiesch et.al (2021) who states that machine learning describes the capacity of systems to use training data that is problem specific by learning from them so as to automate building of analytical models and solve associated tasks. Alzubi et,al (2018) points out that Machine Learning enables computers to imitate behavior of humans as each interaction and action taken, is an opportunity taken by the system to learn and apply the experience the next time there is a similar situation.

Machine learning has been classified in three as supervised learning where labeled data is used with the goal being to establish a relationship or general rule. The agent is directed on actions to maximize rewards by using examples from the labeled training dataset. The second is Unsupervised learning which uses unlabeled training data to find structure in those collections without prior training. It can detect new attack variants but it has a lot of false positives. The third classification is reinforcement learning. This is where interaction is made in a dynamic environment while working towards a goal in the problem space and is provided feedback as rewards, either positive or negative as it navigates the problem space and selects its actions to maximize that reward (Bishop, (2006)).

Support Vector Machine is one of the best algorithms used in IDS (Kuang et.al (2014)). It however like other algorithms has its strengths and weaknesses. Among its strengths is its best overall accuracy and also overfitting, which is common in Machine Learning algorithms can be easily overcome by combining with other algorithms like Genetic Algorithm. Even with the strengths, it still has weaknesses including that the training complexity is affected by the dimensionality and size of the dataset. Feature selection is then used to reduce the dimension of processed data and size and this results to reduced complexity and time needed to process the data. It also has higher false positives than others which makes it difficult to use. (Salo, 2018).

2.1.1 Related Work

Comparing results from experiments is very difficult. This is because differences that are pivotal rely on the test set used to obtain the results of the detection.

The first concept of IDS was documented by Anderson (1980) at the National Security Agency (NSA) 1980. This included tools to assist the security administrators to analyze audit trails like the logs for user access, file access, and system events.

Ambusaidi et al. (2016) introduced an IDS put together with a filter-inspired input reduction approach. The Kyoto 2006 dataset, KDDCup99 dataset, and 18 features of the NSL-KDD were used in this experiment. The authors used a Flexible Mutual Information (FMI) technique, which is a non-linear correlation measure so as to maintain the correlation of the different input variables. The Least Square SVM (LS-SVM) classifier was used in the experiments. Results for the NSL-KDD dataset showed that, the LS-SVM FMI scored an accuracy of 99.94% and false alarm rate (FAR) of 0.28%. Using KDD Cup 99 dataset, it was 78.86% accurate. On the 10th iteration using the Kyoto 2006 dataset, it had a detection rate of 97.80% and FAR of 0.43%.

Bhosale et al. (2014) proposed a multi-agent intelligent system using reinforcement learning and an influence diagram to respond fast against complex attacks and use rules and procedures to assess the state of captured flow. Every agent uses the local database along with the information from others (decisions and events) to learn its policy.

Khammassi et al. (2017) used GA and logistic regression for optimal feature subset selection on the UNSW-NB15 and KDDCup99 datasets. This showed that the feature subset selected using the method is effective for intrusion detection through different decision tree algorithms. After multiple simulations using Weka simulation tool, on 20 out of 42 features of the UNSW-NB15 dataset, the Logistic Regression and Genetic algorithm together with DT classifier scored an accuracy of 81.42% and a FAR of 6.39%. Also, using 18 features of the KDDCup99 dataset got a detection accuracy of 99.90% and FAR rate of 0.105%.

DRL methods from an integration of deep and reinforcement learning can detect and prevent the sophisticated types of cyber-attacks. Quah et al. (2002) use Temporal Difference learning on an actor-critic reinforcement learning model and uses it on classification using a fuzzy adaptative learning control network. The model developed is used on Iris dataset.

Deokar and Hazarnis (2012) indicated the disadvantages of both anomaly-based and signature-based detection methods. The anomaly detection method has a high number of false positives. This is due to its flagging of activities occasionally performed as anomalies. The signature original multi-agent router throttling which detects by applying the divide-and-conquer does not detect novel attacks. This is because it uses stored patterns of attacks in detection. The proposed IDS would discover both known and unknown attacks by using the features of both anomaly and signature detection by using log files. The proposed IDS combined the RL method, association rule learning, and log correlation techniques.

A proposal was made by Nakkeeran et al. (2010) for a detection system consisting of layered detection modules. The results of the nodes neighboring the current node are taken by the current node and the result passed on to the neighboring nodes. The experiment showed increased detection and reduced in false detection.

Vikash et al. (2020) used the Information Gain methodology for a feature reduction method. The UNSW-NB15 dataset was used for validation and filter-based feature extraction technique to select 22 of the 42 attributes of the dataset. The classification process was carried out by using an integrated rule-based model using multiple classifiers that were Tree-based on the IDS. The Attack Accuracy (57.01%), F-Measure (90%), and False Alarm Rate (2.01) were the performance measurement results of the system. Since the technique was misuse based, it was not able to detect zero-day attacks. The recommendation on its improvement was by using other Machine Learning algorithms to replace the Tree-based methods.

Muda et al. (2011) proposed a combination of K-Means clustering and Naïve Bayes classification for hybrid learning. The hybrid approach clusters data before classifying. The result of the experiment on KDD Cup ’99 dataset shows that the approach had an accuracy of 99.5% on DOS and a precision of 40% for U2R.

Hazem et al. (2008) proposed an intrusion detection algorithm and an architecture that includes 2 layers, a global and local layer which perform data collection, analysis, and response functions. The global layer is central. This system applies data mining approach and by filtering out intrusive behavior and normal, analyzes data that is not known, resulting in reduced false alarms.

Jiang et al. (2020) suggested a Network IDS framework using the One-Side Selection technique for reducing the number of noisy data records on the majority classes. Synthetic Minority Over Sampling was applied to raise the dataset minority examples with Convolutional Neural Networks being used to extract the spatial attributes and Bi-directional Long-Short Term Memory models for temporal attributes. This becomes a Deep Learning model for conducting predictive tasks. The UNSW-NB15 and NSL-KDD datasets were used in evaluation. The accuracy obtained from the test data was 77.16% for UNSW-NB15 and 83.58% for NSL-KDD datasets. Although this was an improvement, maximum success is still yet to be achieved.

Zhenghong et al. (2009) proposed using Machine Learning for anomaly detection where Bayesian classification algorithm is used for detection of anomalous nodes to the Wireless Sensor Networks characteristics like limited power, limited communication capacity, and limited computational capabilities of nodes. This achieves relatively high accuracy in detection and lowers false positives. After learning of a few samples is complete, it is also possible to establish detection rules. Using the model for simulation the authors attempted to proof that the proposed model could retrieve important association rules and differentiate with high accuracy between normal connection and the intrusion which is unknown.

Osanaiye et al. (2016) came up with a multiple filters filter-based method for detecting Distributed Denial of Service attacks. The approaches applied include Information Gain, Gain Ratio, Relief, and Chi-Square. They used the NSL-KDD attack detection dataset to show the performance of the system they employed the Decision Tree (DT) algorithm for classification. The DT algorithm was subjected to the k-fold cross-validation method where k=10 for training and validation. The results were that using 13 out of 42 features the DT classifier had a detection of 99.67% and a FAR of 0.42%.

From the conceptual study the researcher came up with the following findings:

1. Most intrusion detection systems and research carried out use supervised and unsupervised methods of machine learning. These methods have achieved accuracies of up to 98% for selected datasets. Zhou (2021) argue that machine learning algorithms such as Random Forest (RF), Bayes, AdaBoost achieved around 99% of precision for the anomaly of ‘DDoS’ and ‘Portscan’.
2. The results of most studies done using machine learning in intrusion detection using the available datasets have been based on selected features of the dataset, thus feature selection has to be done on the datasets so as to get an optimal feature representative of the full feature set (Eesa et al.,2015). The factors affecting the results include the size of the dataset employed, the compute time and the performance of the algorithms (Chandrashekar and Sahin, 2014).
3. Modern IDS perform well in detecting regular and known intrusions but are weak in adversarial AI attacks where malicious inputs are injected into the A.I training data. This calls for continuous improvement of the intrusion detection systems for better performance. This is because accuracy and false alarm rates have not been fully optimized.
4. The most used detection approach used is the anomaly detection as opposed to the signature based as supported by Karami and Guerrero-Zapata, 2015 because it is more effective in detecting novel attacks.
5. The gap identified in these studies was that there is a need to continually improve the performance of intrusion detection systems. Accuracy and false alarm rate have not been fully optimized. Besides, deep reinforcement learning applications on intrusion detection systems have not been exhaustively tested and more work is needed to be done in the area.

2.2 Effectiveness of Using A3C in Anomaly Detection

The Asynchronous Advantage Actor-Critic algorithm is one of the latest algorithms developed under Deep Reinforcement Learning. It makes use of multiple agents. Each agent information on the network and a copy of the environment. These agents explore and learn with each interaction with their individual environments asynchronously. A3C is good because it does not overutilize the GPU resource hence computational efficient and also uses less time when training. Its clear structure also makes it easier to debug and optimize. Deep Reinforcement Learning consists of: $V(s)$, the long-term expected cumulative reward starting from state *(s)* and is used to evaluate the state *(s)*, $Q(s, a)$ refers to the long-term expected cumulative reward starting from executing action *(a)* in the state *(s)* and $\Pi(s, a)$, the probability of executing action under state *(s)* (Sutton R.S (2018)).

Since the global network controls the agents and each agent sends the gained knowledge from exploration to the global network and the global network also enables the agents to acquire diversified training data, this enables the algorithm to perform better than other reinforcement learning techniques. The workers and the global agent are modeled as Deep Neural Networks each with two outputs: One is for the critic which is scalar and shows the expected reward of a given state, V(s). The other which is a vector is for the actor and shows probability distribution over all possible actions (s, a). Advantage value shows the agent how much better the rewards were in comparison to its expectation. The agent then gains better insight of the environment which improves the learning process. The advantage metric is given by the following expression: -
Advantage: $A = Q(s, a) – V(s)$
Since other reinforcement learning techniques are specifically based on either Value Iteration methods or

Policy Gradient methods, A3C is better placed as it combines the best parts of both methods and predict the value and optimal policy functions. A3C is also applied on both discrete as well as continuous action spaces.

A3C has also been seen to be better than other reinforcement learning algorithms as supported by Sewak (2019), since it plays better than DQN in Atari 2600 games and others achieving DQN performance in half the time while using CPUs instead of GPUs and having the best results in comparison to other asynchronous parallel implementations like SARSA and Q-Learning.

The conceptual study done in this area was ideal so not to engage in models that do not work leading to a waste of time. Also, a model could fail due to errors made, but with sufficient time it can be tried several times to ascertain its effectiveness.

The success of the study is supported by Sewak (2019) that "Multiple agents of the actor-critic model could even work in parallel to interact with their individual instances of the environment thereby not only making the training faster but also removing a lot of bias and obviating large memory requirements. The parallel approach could be implemented in both synchronous and asynchronous manners. The Asynchronous Advantage Actor-Critic (A3C) implementation has been quite successful in surpassing many best scores of previous models across the Atari2600 games."

This is supported by Deepanshu (2020) who states that Reinforcement Learning is applicable in t every field that requires automation and exponential advancement. With the A3C having a fast-training speed and acting on both discrete and continuous action spaces resulting to an advantage when comparing the different RL algorithms.

2.3 Selection of appropriate dataset

There have been challenges in finding datasets that give a true reflection of the network traffic in the modern world and also has various low footprint attacks. This is because most of them are private because of security and privacy concerns. Publicly available datasets have also been largely anonymized and fail to validate that they exhibit real-world network traffic behavior. The quality of the dataset finally affects the reliability of the IDS. The KDD98, KDDCUP99, and NSLKDD datasets have been shown not to be reflective of current network traffic and lacks modern low footprint attacks found in the current network threat environment as supported by Tavallaee et al. (2009) and McHugh & John (2000).

UNSW-NB15 Dataset

The University of New South Wales Network Based 2015 (UNSW-NB15) dataset is a comprehensive dataset for network intrusion detection systems and was created and benchmarked by Moustafa (2017) using the IXIA PerfectStorm tool. It was created for generating a hybrid of real modern behaviors considered normal and synthetic contemporary attacks using new and existing processes for feature generation. The traffic in this dataset was captured in two days, on the 22nd January and 15 hours of 17th February 2015 (Moustafa et al. (2015)) in the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS). The preprocessed data is in PCAP files, CSV files, and the output of flow information extraction tools Argus and Bro-IDS. The dataset contains 45 features that are both numeric and categorical. These features are grouped as:

1. Flow features include transport layer features like IP addresses, ports, and protocol type
2. Basic contains packet-based and flow-based features, such as duration, number of bytes, and number of packets.
3. Content features contains exploration of connection content and window advertisements or sequence numbers.
4. Time features contain features related to time like jitter, start &amp; end time
5. Others are categorized as general purpose and connection features.

3. Methodology

Quantitative methodology was selected for the study. This was important when looking at the appropriate dataset to be used in A3C. The dataset used is in CSV format. It was obtained from the internet and is downloadable from the website at;

https://cloudstor.aarnet.edu.au/plus/index.php/s/2DhnLGDdEECo4ys

The sampled dataset was analyzed and visualized using python scikit learn which is a library containing tools for data analysis. This was done to ascertain the dependability of the dataset. The presentation of the various distributions in the dataset was in the form of pie chart and bar graph. Quantitative methodology promotes unbiased results that can be generalizable to a larger population (Lisa, 2008).

2.3.1 Sampling Method

Purposive sampling was applied in selecting the UNSW-NB15 dataset as it was the most viable dataset to be used in the study. Morse and Niehaus (2009) assert thar the intent of sampling methods is to maximize efficiency and validity regardless of the methodology. The purposeful selection of the UNSW-NB15_TRAIN and the UNSW-NB15_TEST was important as they would give detailed and true reflection of modern attacks. This is supported by Patton (2002) that purposive sampling is used for selecting cases rich in information so that limited resources can be used in the most effective manner possible.

The training set which is employed for training of models has 175,341 records while the testing set employed for testing trained models has 82,332 records having both attack in different types and normal.

The dataset has 9 attack types as described in Moustafa and Slay (2016) and Moustafa and Slay (2016).

Table 1: Attack Types Description Moustafa and Slay (2016) and Moustafa and Slay (2016)

Type No.	Description
Normal	Natural transaction data.
Fuzzers	Attempting to cause a program or network suspended by feeding it the randomly generated data.
Analysis	It contains different attacks of port scan, spam and html files penetrations
Backdoor	A technique in which a system security mechanism is bypassed stealthily to access a computer or its data.
DOS	A malicious attempt to make a server or a network resource unavailable to users, usually by temporarily interrupting or suspending the services of a host connected to the Internet.
Exploits	The attacker knows of a security problem within an operating system or a piece of software and leverages that knowledge by exploiting the vulnerability.
Generic	A technique works against all block-ciphers (with a given block and key size), without consideration about the structure of the block-cipher.
Reconnaissance	Contains all Strikes that can simulate attacks that gather information.
Shellcode	A small piece of code used as the payload in the exploitation of software vulnerability.
Worms	Attacker replicates itself in order to spread to other computers. Often, it uses a computer network to spread itself, relying on security failures on the target computer to access it.

2.3.2 Conceptual Model

In using reinforcement learning, the agent actions including dataset selection, prediction, and validation and a policy that maximizes rewards to attain higher accuracy. If the agent correctly predicts, a positive reward is given. The agent's goal is to maximize rewards. In this study, the states included the dataset records, weights, and accuracy while the environment in which the agent operates is the UNSW-NB15 dataset. The conceptual model shows the systematic approach to the processes and how the objectives will be met. Figure 1 shows the reinforcement learning conceptual model for intrusion detection.

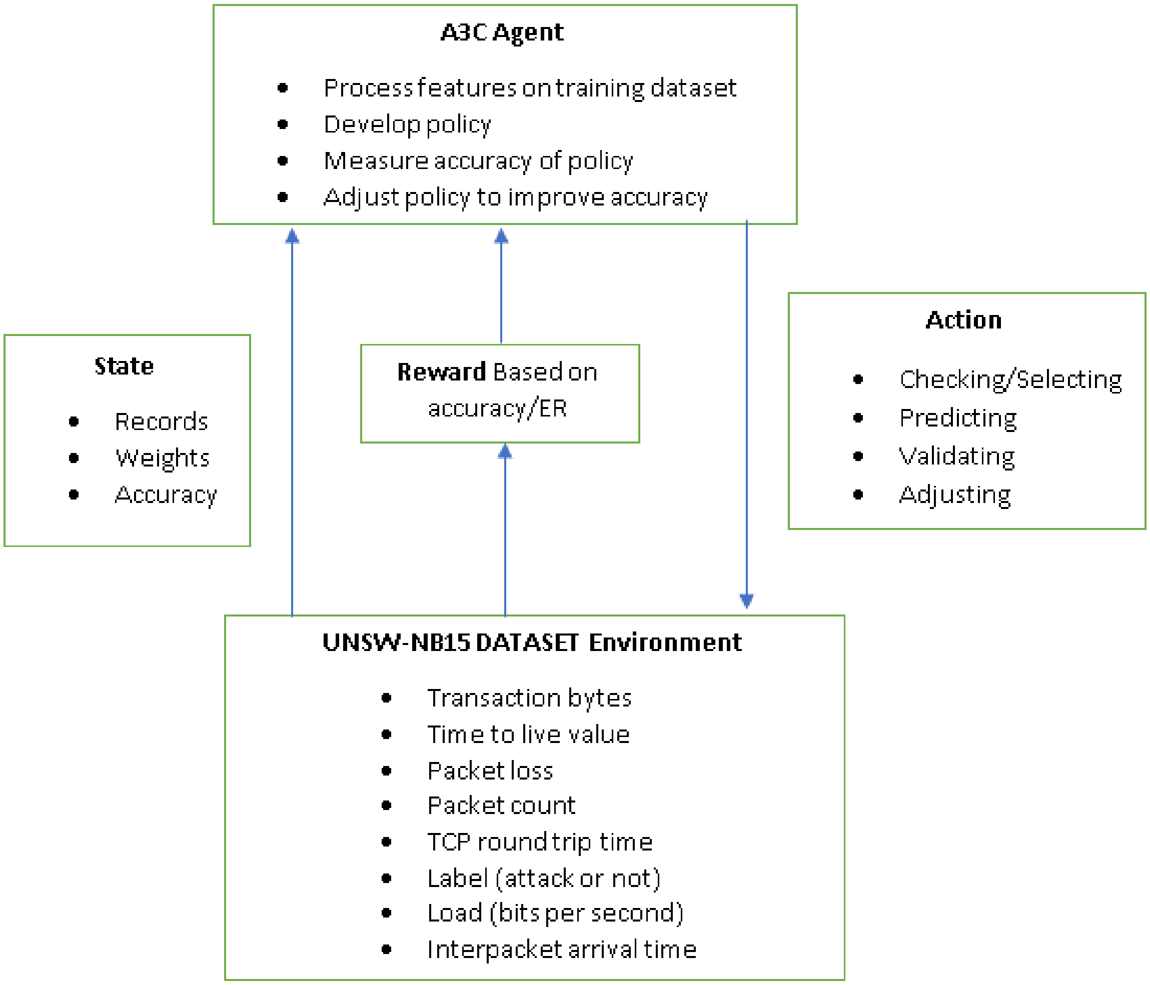

Figure 1: A Reinforcement Learning Conceptual Model for Intrusion Detection. Source: Researcher, 2022

4. Results and Discussion

The UNSW-NB15 dataset was selected for this study. The dataset has 45 features and divided into UNSW-NB15-TRAIN used for model training and UNSW-NB15-TEST for testing at 70% and 30% respectively. All the 45 features were used in the study. This dataset was preprocessed and this included cleaning by ensuring no null values and consistent data and transformation by normalization and one hot encoding.

The model should not train on the test data so as to avoid data leakage. Data leakage occurs during training when a model obtains information it was not to have and so introduce bias to the resulting model, leading to poor performance on the model for previously unseen data (Shabtai et.al (2012)).

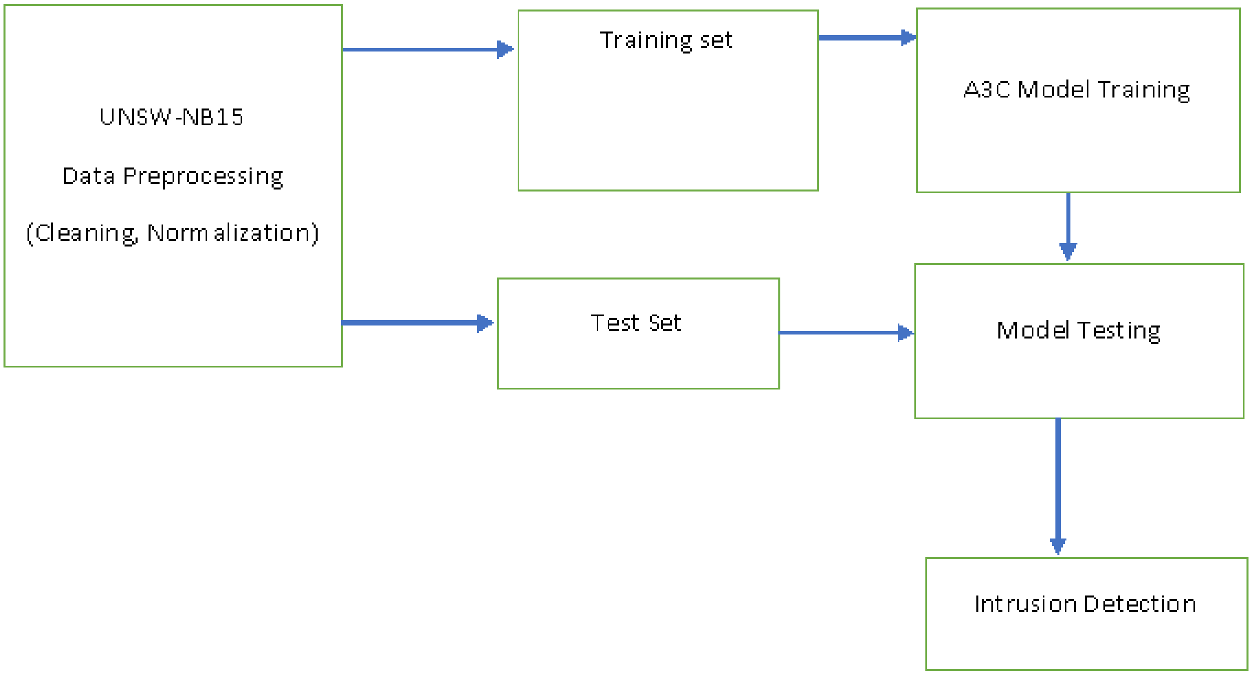

Figure 2: IDS Architecture using UNSW-NB15 dataset

In the proposed architectural design, the dataset passes through data preprocessing to ensure the dataset is ready for processing. This is then taken through model training where the training dataset is used and finally the model testing phase of the model using the test dataset. This process is iterative until a fine-tuned and desired fit model is found. In the preprocessing stage, the data was further separated in two for the training set and classified as normal and anomaly. The resulting fine-tuned dataset that had been preprocessed by cleaning, transformation and reduction was available for use in the IDS model developed.

Dataset visualization represents information and data graphically. This is done using visual elements like charts and graphs. Tools used in data visualization enables one to see and understand trends, outliers, and patterns in data. The most important features of the dataset featured is shown in figures below

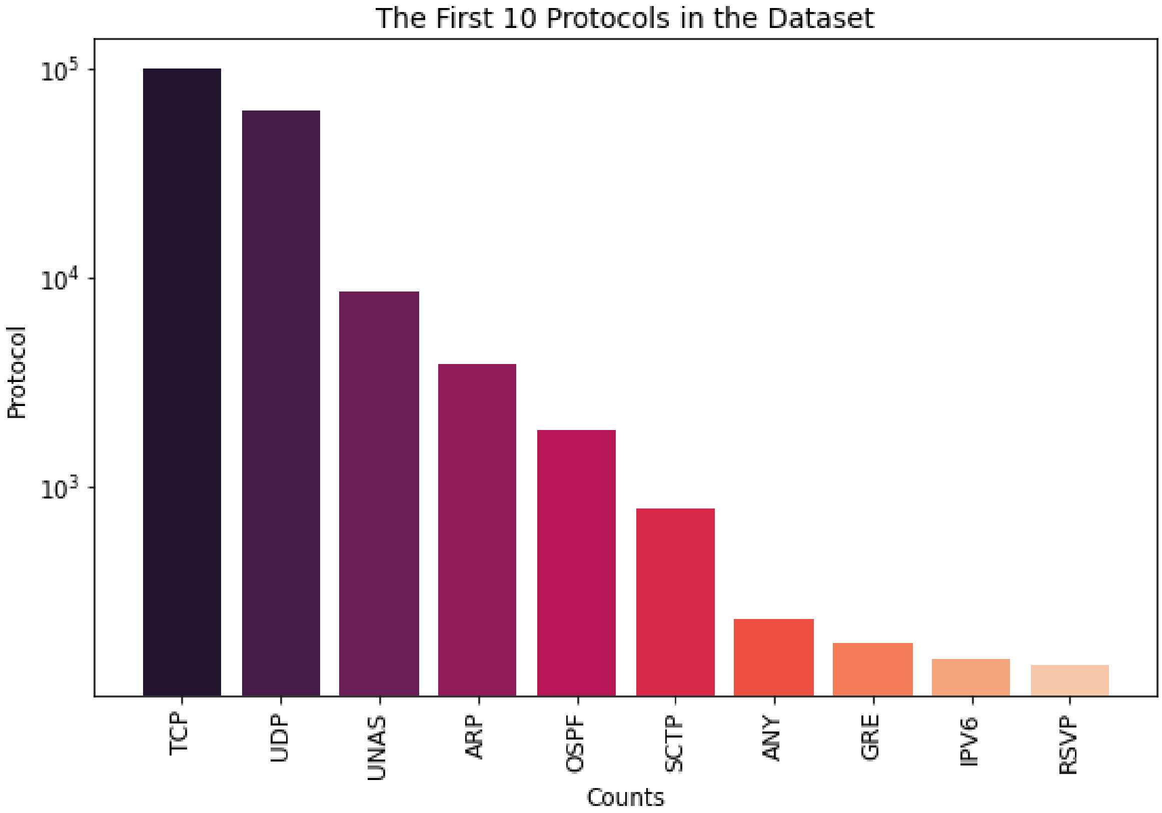

Figure 3: Protocols in the UNSW-NB15 dataset

The first two protocols; the tcp and udp are the most common and cover most of the total protocols included in the dataset. The rest follow by distribution as shown in figure 3.

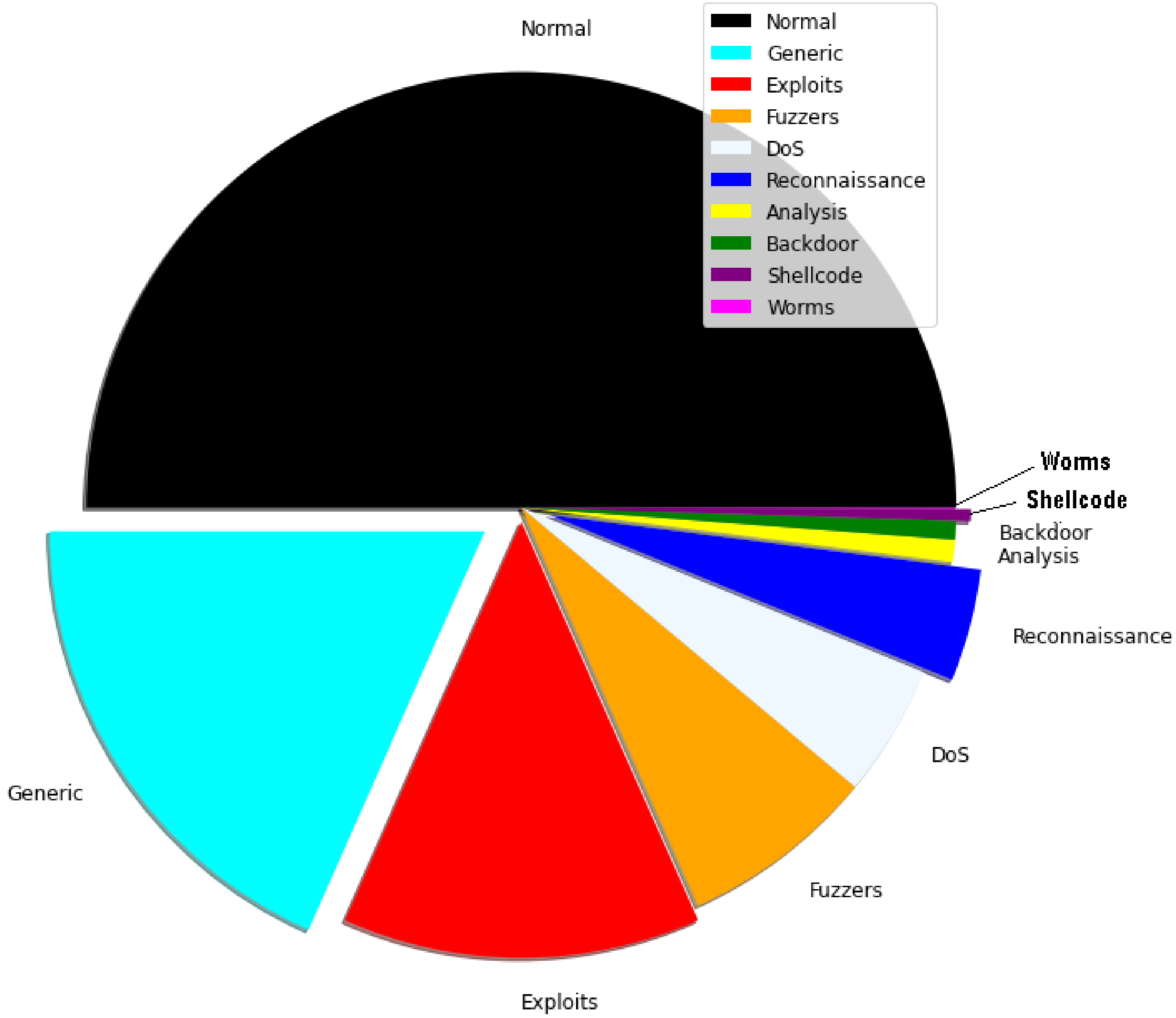

Figure 4: Distribution of attacks in UNSW-NB15 dataset

The UNSW-NB15 dataset has 9 families of attacks and normal. All the attack types have been represented in the dataset with different distributions. The normal takes around a third while the rest is distributed among the different attack types with Generic attacks and exploits having the biggest number in the distribution as shown in figure 4.

5. Conclusion and Recommendations

Although there have been notable improvements in Intrusion detection, adversarial attacks still pose a challenge to the realization of better intrusion detection systems. Application of deep reinforcement learning on intrusion detection has not been exhaustively done and tested and hence this calls for more research in this area due to its potential to solve such challenges. A3C algorithm can be used on intrusion detection systems and as such its effectiveness should be tested and measured. The use of the UNSW-NB15 dataset is highly recommended since it covers most of the attack types used by modern attackers. Since the dataset is comprehensive, all the attack types should be used so as to accurately depict the intrusions. If selected attack types are to be used, feature selection should be done accurately so as to reflect modern attack types. A collaboration of researches can also be done by coming up with frameworks that can make it easier to compare the performance of various researches. This will aid in improving the performance of IDS.

References

Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine learning from theory to algorithms: An overview. Journal of Physics: Conference Series, 1142, 012012. https://doi.org/10.1088/1742-6596/1142/1/012012

Ambusaidi MA., He X., Nanda P., & Tan Z. Building an intrusion detection system using a filter-based feature selection algorithm. IEEE Trans Comput. 2016; 65(10):2986–98.

Anderson& James P., "Computer Security Threat Monitoring and Surveillance," Washing, PA, James P. Anderson Co., 1980.

Bhosale, R., Mahajan, S., & Kulkarni, P. (2014). Cooperative machine learning for intrusion detection system. International Journal of Scientific and Engineering Research, 5(1), 1780-1785.

Bishop, C. M. (2006), Pattern Recognition and Machine Learning, Springer, ISBN 978-0-387-31073-2

Chandrashekar G., & Sahin F. A survey on feature selection methods. Computer Electr Eng 2014;40(1):16–28

Cisco. (2022, January 23). Cisco annual internet Report - Cisco Annual Internet Report (2018–2023) White Paper. Cisco. Retrieved March 11, 2022, from https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

Cyberthreat defense REPORT 2021. CyberEdge Group. (2021, April 20). https://cyber-edge.com/cdr/.

Deepanshu Mehta. (2020). State-of-the-art reinforcement learning algorithms. International Journal of Engineering Research And, V8 (12). https://doi.org/10.17577/ijertv8is120332

Deokar, B., & Hazarnis, A. (2012). Intrusion detection system using log files and reinforcement learning. International Journal of Computer Applications, 45(19), 28-35.

Eesa AS., Orman, Z., & Brifcani AMA. A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems. Expert Syst Appl 2015;42(5):2670–9 J.M.

Fossaceca, J.M., Mazzuchi, T.A. & Sarkani, S. MARK-ELM: application of a novel multiple kernel learning framework for improving the robustness of network intrusion detection, Expert Syst. Appl., Vol. 42, 2015, pp. 4062-4080.

Hazem M. El-Bakry., & Nikos Mastorakis, A, “Real-Time Intrusion Detection Algorithm for Network Security, WSEAS Transactions on Communications, Issue 12, Volume 7, December 2008.

Janiesch, C., Zschech, P. & Heinrich, K. Machine learning and deep learning. Electron Markets 31, 685–695 (2021). https://doi.org/10.1007/s12525-021-00475-2

Jiadong, R., Jiawei, G., Wang, Q., Huang, Y., Xiaobing, H., & Hu J. "Building an Effective Intrusion Detection System by Using Hybrid Data Optimization Based on Machine Learning Algorithms", Security and Communication Networks, vol. 2019, Article ID 7130868, 11 pages, 2019. https://doi.org/10.1155/2019/7130868

Jiang, K., Wang, W., Wang, A., & Wu, H. Network intrusion detection combined hybrid sampling with deep hierarchical network. IEEE Access. 2020; 8:32464–476.

Karami, A., & Guerrero-Zapata, M. A fuzzy anomaly detection system based on hybrid pso-kmeans algorithm in content-centric networks. Neurocomputing 2015; 149:1253–69.

Khammassi, C., & Krichen, S. A GA-LR wrapper approach for feature selection in network intrusion detection,” Computers & Security, vol. 70, pp. 255–277, 2017.

Kim, G., Lee, S., & Kim S. A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Syst Appl 2014;41(4):1690–700

Kuang F., Xu, W., & Zhang S., A novel hybrid KPCA and SVM with GA model for intrusion detection, Appl. Soft Comput., Vol. 18, 2014, pp. 178-184.

Vikash K., Sinha D., Das AK., Pandey SC. & Goswami RT. An integrated rule-based intrusion detection system: analysis on UNSW-NB15 data set and the real time online dataset. Cluster Comput. 2020; 23(2):1397–1418

McHugh& John,” Testing intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory”. ACM transactions on Information and system Security, 3, 2000, p 262-294

Mitchell, T. (1997). Machine Learning. New York: McGraw Hill. ISBN 0-07-042807-7. OCLC 36417892.

Morse, JM., Niehaus, L. Mixed method design: Principles and procedures. Left Coast Press; Walnut Creek, CA: 2009.

Moustafa, N. (2017) Designing an online and reliable statistical anomaly detection framework for dealing with large high-speed network traffic

Moustafa, N., & Slay, J. (2016) The significant features of UNSW-NB15 and the KDD99 datasets for network intrusion detection systems. 4th International Workshop on Building Analysis Datasets and Gathering Experience Returns for Security. Kyoto, Japan: IEEE; 2016:25-31

Moustafa, N., & Slay, J. (2016) The evaluation of network anomaly detection systems: statistical analysis of the UNSW-NB15 dataset and the comparison with the KDD99 dataset. Inform Security a Global Perspective. 2016;25(1-3):18-31

Moustafa, N., & Slay, J. (2015). UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). 10.1109/MilCIS.2015.7348942.

Muda, Z., Yassin, W., Sulaiman, M.N., & Udzir, N.I.,” Intrusion Detection based on K-Means Clustering and Naïve Bayes Classification”, 7th International Conference on IT in Asia (CITA), 2011.

Nakkeeran, R., Aruldoss, T. A., & Ezumalai R. “Agent-Based Efficient Anomaly Intrusion Detection System in Ad-hoc networks” IACSIT International Journal of Engineering and Technology Vol. 2, No.1, February 2010

Osanaiye, O., Cai, H., Choo, K-KR., Dehghantanha, A., Xu, Z., & Dlodlo, M. Ensemble-based multi-filter feature selection method for DDOS detection in cloud computing. EURASIP J Wirel Commun Netw. 2016;2016(1):130

Quah K. H., Quek C., & Leedham G., "Pattern classification using fuzzy adaptive learning control network and reinforcement learning," Proceedings of the 9th International Conference on Neural Information Processing. ICONIP '02., Singapore, 2002, pp. 1439-1443 vol.3. 2002.

F. Salo, M. Injadat, A.B. Nassif, A. Shami, and A. Essex, Data mining techniques in intrusion detection systems: a systematic literature review, IEEE Access, Vol. 6, 2018, pp. 56046-56058.

Sewak, M. (2019). Actor-critic models and the a3c. Deep Reinforcement Learning, 141–152. https://doi.org/10.1007/978-981-13-8285-7_11

Shabtai, A., Elovici, Y., & Rokach, L. A survey of data leakage detection and prevention solutions. Berlin: Springer; 2012

Sutton, R.S.; & Barto, A.G. Reinforcement Learning: An Introduction, 2nd ed.; MIT Press, 2018

Tavallaee, M., Bagheri, E., Lu, W. & Ghorbani, A. “A detailed analysis of the KDD CUP 99 data set”. Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications, 2009.

Walliman, N. S. & Walliman N. (2011) “Research methods: the basics” Taylor and Francis

Zhenghong, X., Chuling, L., & Chaotian, C., “An Anomaly Detection Scheme Based on Machine Learning for WSN” IEEE International Conference on Information Science and Engineering,2009

Zhou, K., Wang, W., Hu, T., & Deng, K. (2021). Application of improved asynchronous advantage actor critic reinforcement learning model on anomaly detection. *Entropy*, *23*(3), 274. https://doi.org/10.3390/e23030274

Zimek, A., Schubert, E. (2017) Outlier Detection. In: Liu L., Özsu M. (eds) Encyclopedia of Database Systems. Springer, New York, NY. https://doi.org/10.1007/978-1-4899-7993-3_80719-1