A Multiagent and Machine Learning based Hybrid NIDS for Known and Unknown Cyber-attacks

Said OUIAZZANE, Malika ADDOU, Fatimazahra BARRAMOU
ASYR Team, Laboratory of Systems Engineering (LaGeS)
Hassania School of Public Works (EHTP)
Casablanca, Morocco

***Abstract*—The objective of this paper is to propose a hybrid Network Intrusion Detection System (NIDS) for the detection of cyber-attacks that may target modern computer networks. Indeed, in the era of technological evolution that the world is currently experiencing, hackers are constantly inventing new attack mechanisms that can bypass traditional security systems. Thus, NIDS are now an essential security brick to be deployed in corporate networks to detect known and zero-day attacks. In this research work, we propose a hybrid NIDS model based on the use of both a signature-based NIDS and an anomaly detection NIDS. The proposed system is based on agent technology, SNORT signature-based NIDS, machine learning techniques and the CICIDS2017 dataset is used for training and evaluation purposes. Thus, the CICIDS2017 dataset has undergone several pre-processing actions, namely, dataset cleaning, and dataset balancing as well as reducing the number of attributes (from 79 to 33 attributes). In addition, a set of machine learning algorithms are used, such as decision tree, random forest, Naive Bayes and multilayer perceptron, and are evaluated using some metrics, such as recall, precision, F-measure and accuracy. The detection methods used give very satisfactory results in terms of modeling benign network traffic and the accuracy reaches 99.9% for some algorithms.**

***Keywords*—Intrusion detection; zero-day attacks; machine learning; multi-agent systems; security**

I. Introduction

The Global Internet Usage Statistics report confirms a growth of 1,114% and more than 2 quintillion bytes of data are generated every day. Along with this growth, cybercrime is becoming more sophisticated and continues to grow day by day [1, 2, 3]. As a result, the risks of being attacked and targeted by the hacker community remain more likely and could be costly for victims of cyber-attacks. Thus, the importance of Network Intrusion Detection Systems (NIDS) continues to grow and attract the interest of researchers [4] and NIDSs have become indispensable for securing network infrastructures against cyber-attacks [5]. However, the evolution of NIDSs is slowed down due to several challenges that are mainly related to the volume of network data, the emergence of increasingly sophisticated attacks [6] and unbalanced learning datasets [42]. In addition, real-time processing of network traffic is a very important feature of an effective NIDS to monitor all network events [8]. Not to mention that network traffic is continuously changing and therefore, the training datasets need to be updated regularly to effectively evaluate the detection models [5]. According to [22] and [42], the lack of more adequate datasets for anomaly detection-based intrusion detection has caused intrusion detection methods to suffer in analysis and deployment. The authors of [7] confirm that all these challenges remain a blocking obstacle against the evolution of the IDS domain in terms of performance, accuracy, and execution time during the learning and detection phases. Furthermore, the approaches proposed in the literature are not clear in terms of architecture and do not opt for hybrid architectures adopting, both, signature-based and anomaly detection-based NIDS. Most of the research works, carried out in this sense, remain theoretical and do not propose more efficient mechanisms capable of detecting known and unknown attacks.

In this research work, we will propose an effective intrusion detection approach to detect known and unknown cyber-attacks. Our approach consists of a Snort-based intrusion detection model to detect known intrusions and then machine learning techniques to detect any suspicious deviation from the baseline profile of benign network traffic. This baseline is designed by regularly training the system on normal network events using machine learning methods.

The selection of the research works carried out by the scientific community working on cybersecurity was done using a database of 17 journals (Q1 and Q2) and the used search terms are presented in Fig. 1 according to the methodology of [44].

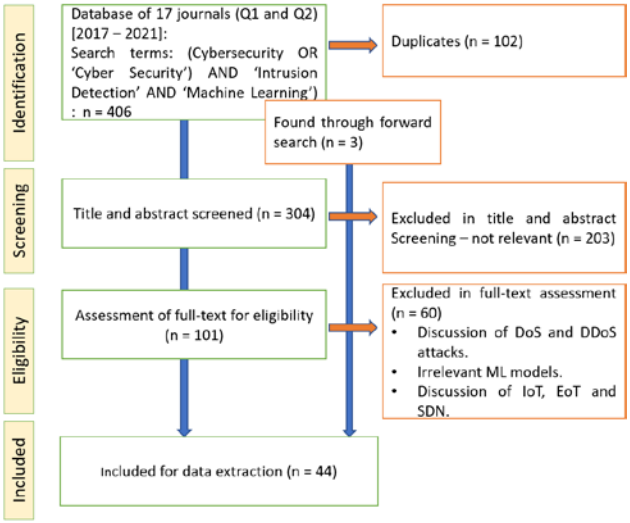

Fig. 1. Flow Diagram to Perform Papers Selection.

The remainder of this paper is structured as follows. Section II highlights some related works conducted by scientific community. Section III highlights gives some basics related to our theme of research. Section IV presents our proposed approach and finally Section V handles the conducted tests and experiments to validate the classification of benign network traffic.

II. Related Work and Discussion

A. Related Work

In this section, we will highlight some of the research works that have been carried out by researchers to ensure a quick advance of intrusion detection mechanisms based mainly on Machine Learning, Data Mining and Deep Learning techniques.

Since the beginning, researchers started to propose various approaches to effectively deal with the problem of Intrusion Detection. Notably, the Table I below summarizes some of the research works carried out by the scientific community to contribute in enhancing NIDS.

B. Discussion

It is true that several research works have been conducted by researchers to develop the field of intrusion detection systems. However, most of the aforementioned works have shortcomings in terms of architecture, datasets used as well as the machine learning methods used and each research work addresses a specific problem. For example, in the paper [25], the researcher limited himself to intrusion detection in wireless networks, in [39], the author proposed an IDS for SDN-based networks etc. In our research work, we will propose a universal NIDS, capable of being deployed in any type of computer networks. Our NIDS model will be based on a multi-layer architecture with the use of the multi-agent paradigm and will also be based on a hybrid detection mechanism combining a Signature-based NIDS (SNIDS) and an Anomaly-based NIDS (ADNIDS).

TABLE I. Research Works Carried Out in the Intrusion Detection Field

Ref	Approaches	Techniques
[23]	Intrusion detection system based on feature selection and ensemble classifier.	+ CFS-BA heuristic algorithm to reduce dimension of the training dataset + Combination of C4.5 and Forest by Penalizing Attributes to classify data
[24]	Method of intrusion detection	+ Auto-Encoder deep learning agorithm and Random Forest to reduce the dimension of the dataset
[25]	Intrusion Detection System for wireless networks	Feed-Forward Deep Neural Network and Wrapper-based Feature Extraction Unit techniques using UNSW-B15 and AWID datasets.
[26]	Model of a real-time IDS that can distinguish between benign and malicious network traffic.	+ Support Vector Machine (SVM) and Extreme Learning Machine to detect known and unknown attacks + Modified K-Means to get a good quality and small training dataset from KDDCUP
[27]	Attribute selection method	Pigeon Inspired Optimizer algorithm
[28]	New application of Deep Reinforcement Learning for intrusion detection.	+ Deep Reinforcement Learning to detect intrusions + NSL-KDD and AWID datasets
[29]	An intrusion detection mechanism to model benign traffic	Supervised machine learning methods to model benign network traffic
[30]	Comparative study of Snort and Suricata.	Plugin which is based on the SVM and Fuzzy Logic algorithms to reduce the False Positive rate.
[31]	Cooperative IDS	+ Machine Learning methods + Proactive decision making based on previous exchanges + Denoising Autoencoder and DNN
[13]	Intrusion detection architecture to detect attacks targeting the Cloud networks.	Machine Learning methods
[32]	An IDS for wireless networks	Deep Gated Recurrent Unit and Wrapper-based feature extraction using NSL-KDD
[33]	Hybrid model to detect intrusions using Deep Learning	Convolutional Neural Network and Weight-Dropped, Long Short-Term Memory network.
[34]	A comparative study between the different approaches of intrusion detection.	Describing of 35 known datasets used in the field of Intrusion Detection
[35]	A new intrusion detection technique	+ Semantic Re-encoding and Deep Learning + NSL-KDD dataset
[36]	Collaborative intrusion detection system for Internet of Things (IoT) networks.	+ Semi-supervised machine learning algorithms + Tests conducted on real IoT environments.
[37]	An approach for intrusion detection in Edge-of-Things.	Deep Belief Network (DBN)
[38]	A model of adaptive intrusion detection system to detect known and unknown cyber-attacks.	Extreme Learning Machine
[3]	State of the art study of IDSs based on public datasets.	More visibility into what is being done by the scientific community to identify unknown cyber-attacks and to better understand the problems that such IDS suffer from.
[4]	State-of-the-art study of various previous cybersecurity surveys focused on Deep Learning.	Details about IDS including input data, detection mechanisms, deployment modes as well as different evaluation strategies.
[39]	An approach for intrusion detection in Software Defined Network (SDN).	Decision Tree algorithm.
[40]	Hybrid IDS	Signature-based detection and Anomaly-based detection.

Our NIDS model will be based on mutli-agent technology in order to make the system modular and distributed. Thus, the proposed system will be extensible and capable of adding other components to perform large-scale detection missions in huge networks. Moreover, as we have already said, our system combines both detection mechanisms (SNIDS and ADNIDS) in order to detect all types of attacks (known and unknown). The used SNIDS is based on the famous open source NIDS SNORT and allows the detection of known intrusions. Moreover, ADNIDS intervenes when the packet is not recognized by the SNIDS and compares the packet's characteristics against the baseline patterns (benign traffic) modelled by supervised machine learning techniques applied to the cleaned and optimized CICIDS2017 dataset.

In order to improve the accuracy and precision of the used detection mechanism to model the benign network traffic, we opted for cleaning and reducing the dimension of the CICIDS2017 dataset. Thus, the used training dataset is devoid of any unnecessary information that could falsify the classification results.

III. Context of the Study

A. Cybersecurity

Cybersecurity is a discipline that has been evolving exponentially over the past decade [9]. It refers to the set of practices to protect the cyberspace environment against suspicious activities that may affect its security principles [10]. Among the security principles, we have the integrity of the data that aims to prevent the alteration of the information by unauthorized persons. The second principle is confidentiality, which confirms that the data should not be accessible by malicious people and finally the principle of high availability which ensures that the computer assets are available at any time to serve legitimate requests [11, 12].

B. Intrusion Detection Systems

The Intrusion Detection System (IDS) is one of the key components for ensuring the security of mobile clouds [13]. IDSs are classified according to the data source and the used detection method. Based on the nature of data sources, we can distinguish between two types of IDS: Host-based IDS and Network-based IDS. Furthermore, based on the used analysis method, we have two types of IDS: Signature-based IDS and Anomaly-based IDS [2].

NIDS analyses network traffic passing through computer environments [14, 3]. Its role is mainly to monitor the network events against suspicious activities that may violate or bypass security policies of security components such as firewalls, Web Application Firewalls and proxies. A NIDS usually consists of three main modules which are Monitoring, Analysis and detection, and Alarm modules. The Monitoring module observes network traffic, resource usage and patterns. The Analysis and Detection module is the key part of the system; it identifies cyber-attacks based on specific algorithms. Finally, the Alarm module is responsible for notifying the security administrators in case of possible intrusions [15]. Furthermore, conventional security mechanisms cannot detect unknown zero-day attacks [16] that have no signature or whose patterns are not yet known to security experts. Another issue that modern computer networks are facing is that network traffic is responding to Big Data issues (volume, variety and velocity). As a result, network traffic processing must make use of Big Data technologies to improve the quality of analysis and to reduce execution time during learning phases [17].

C. Snort: Open source Network Intrusion Detection

Snort is an open source IDS developed by Sourcefire in 1998 and has gained a very good reputation over the past decade due to its frequent use by researchers. Snort is structured in a TCP/IP stack architecture to capture and inspect network packets. This IDS is in its version 3.0 just released to overcome the single-thread limitation to support by default multithreading [18].

D. CICIDS2017 Dataset

A dataset for intrusion detection is developed by collecting network traffic events from heterogeneous sources. These events can describe system, user and configuration behaviors [19]. These datasets do not include network events that can represent zero-day attacks [20]. The CICIDS2017 dataset is one of the most modern datasets [21].

IV. Proposed Approach

A. Proposed Model

1) Architecture of the proposed model: Fig. 2 presents the proposed intrusion detection system model to ensure the detection of known and unknown attacks (0 Day) within any type of computer network. The proposed architecture is mainly based on three layers that collaborate together to perform cyber-attack detection missions.

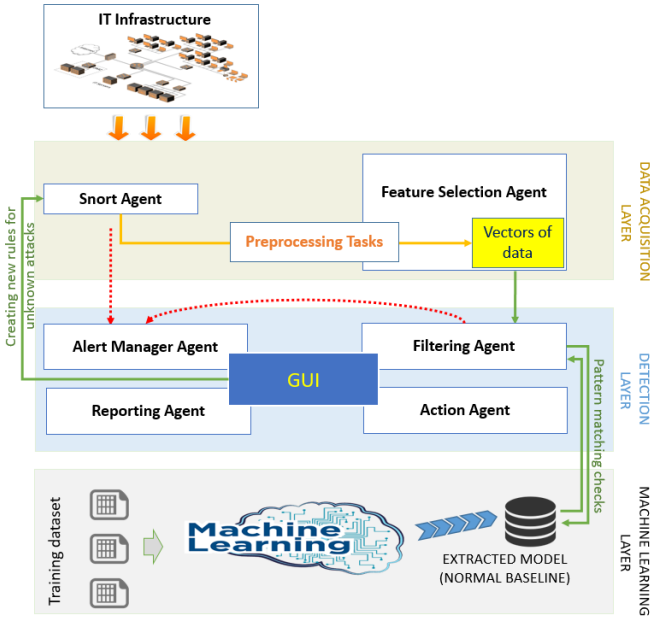

Fig. 2. The Proposed Network Intrusion Detection System.

2) Components of the proposed model

The system has three main layers:

- Data Acquisition Layer (DAL): This layer is responsible for data capture and pre-processing of network traffic. It also performs feature extraction to transform the captured network packets into data vectors to be used by machine learning methods. The DAL includes Snort Agent, a small component responsible for pre-processing tasks and an agent responsible for feature extraction.
- Detection Layer (DL): This component is responsible for detecting deviations from a network baseline. It is based on a machine learning model developed after training the system on a training dataset containing benign network traffic. The DL also sends alerts when an intrusion is detected and allows the security administrator to generate reports and take actions on the network and system infrastructure in case of a security incident.
- Machine Learning Layer: This part allows the NIDS system to perform training tasks on normal network behavior. Using supervised machine learning techniques on a dataset including benign network traffic, a model is developed that will check the fit to detect deviations from the designed baseline.

B. *Operating Principle*

Our system must be trained regularly on benign network traffic devoid of any type of cyber-attacks. Thus, datasets like CICIDS2017 are used to develop and design a baseline identifying the normal operation of a computer network. The training process of the proposed NIDS is mainly done in six steps:

- Data acquisition: The system collects data to train itself and to obtain the network baseline describing normal network behaviors. We used the CICIDS2017 dataset (Benign traffic) devoid of any kind of cyber-attacks.
- Pre-processing: In order for the data to be exploitable by machine learning based classification techniques, data preprocessing actions must be undertaken. Thus, missing value removal, scaling and partitioning techniques are all used to improve the quality of the training dataset.
- Classification: In this step, machine leaning based classification techniques are used to model the normal behavior of the network based on the benign dataset. Several machine learning algorithms are used to select the one with the highest accuracy, with very low false alarm rates and with an increased processing speed.
- Testing and validation: After using a set of machine learning techniques, it is now time to evaluate these algorithms based on specific metrics that address intrusion detection issues. From there, the most efficient machine learning technique is chosen to model the normal network traffic.
- Use of the model "Baseline": After modeling the baseline of the network during normal operation based on the CICIDS2017 dataset, the generated model will be used to identify any deviation from normal behavior. Thus, unknown 0day attacks can be easily identified.

C. *Real Time Detection Flowchart*

Fig. 3 shows the detection principle of our NIDS model. Indeed, our system is supposed to train beforehand on benign network traffic that does not include any trace of cyber-attack, so the generated model will be considered as the network baseline to which the system will compare the real network packets.

Our system ensures the detection of intrusions in the networks according to the following steps:

- Step 1 – Sniffing and gathering: During this step, the NIDS listens to the network to collect all the packets that are passing through it. To do this, the proposed model relies on the Snort agent to capture the network traffic.
- Step 2 – Matching check: During this step, the Snort agent compares the patterns of the network packets it receives against a signature database describing all known cyber-attacks (Snort DB). Based on the result of the matching check, the Snort agent notifies the NIDS administrator if there is a known attack in the network.
- Step 3 – Data preprocessing: At this point, the captured packet is not recognized by Snort's knowledge base. Therefore, the network traffic must undergo pre-processing operations so that it can be consumed by machine learning algorithms. Thus, feature extraction techniques are applied to the captured network traffic in order to transform the data streams into data vectors that can be exploited by machine learning models.

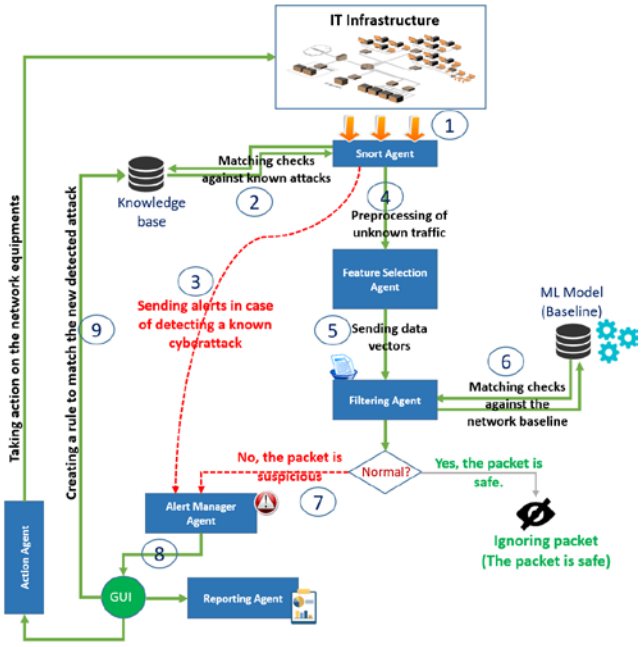

Fig. 3. The Flowchart of the Detection Mechanism.

- Step 4 – Filtering and matching check: After transforming the network flows into data vectors, the Filtering Agent checks the match between the data it receives and the "Network Baseline" model previously generated after training the system on benign network traffic. Depending on the result of the matching verification, two scenarios could arise: If the network packet is normal, no alert is generated and if the packet does not match the network baseline, the NIDS administrator must be informed in time to analyse the event.
- Step 5 – Enrichment of the Snort knowledge base: In case an event deviates from the network baseline, the NIDS system must notify the administrator. The administrator must then intervene to diagnose and analyse the suspicious event, and can also contact security vendors and publishers to identify the nature of the suspicious network event. The security administrator can create rules in the Snort to intercept similar events that may occur in the future. The detected suspicious network event can be Zero-day attacks for which the security vendors have not yet developed a patch or signature.

V. Experimentation and Tests

This section focuses on the experiments and tests performed to evaluate the performance of the different algorithms used for benign traffic modeling. For this purpose, the CICIDS2017 dataset is used and therefore it is necessary to analyze and clean it before using it by machine learning algorithms.

A. Composition of the used Dataset

We analyzed the CICIDS2017 dataset published by the Canadian Cybersecurity Institute using the Pandas framework in Python. The latter allowed us to analyze the content of the various CSV files constituting CICIDS2017 dedicated to research in the field of intrusion detection systems based on Machine Learning and Deep Learning.

The CICIDS2017 dataset consists of a set of eight files in a CSV format; these files include data about network traffic captured during five days from Monday to Friday. After analyzing the content of the set of CSV files using Pandas, we were able to identify the composition of the CICIDS2017 dataset and Table II summarizes the obtained results.

From the above statistics, it appears that the dataset is unbalanced due to the abundance of normal traffic compared to attack traffic, in addition to the existence of few records of certain types of attacks. This imbalance in the traffic classes automatically implies a biased machine learning model. Knowing that the class with a lot of traffic will be favored over the others with less records during the learning stage. As a result, the classes with few records make the machine learning model learn nothing about them and consequently have a biased detection model towards attacks with few records in the learning dataset.

TABLE II. Composition of the CICIDS2017 Dataset [41]

Day	Class of captured traffic	Number of records
Monday	Benign	529918
Tuesday	Benign	432074
	SSH-Patator	5897
	FTP-Patator	7938
Wednesday	Benign	440031
	DoS Hulk	231073
	DoS GoldenEye	10293
	DoS Slowloris	5796
	DoS Slowhttptest	5499
	Heartbleed	11
Thursday Morning	Benign	168186
	Web Attack Brute Force	1507
	Web Attack Sql Injection	21
	Web Attack XSS	652
Thursday – Afternoon	Benign	288566
	Infiltration	36
Friday – Morning	Benign	189067
	Bot	1966
Friday – Afternoon – PortScan	Benign	127537
	Port Scan	158930
Friday – Afternoon – DDoS	Benign	97718
	DDoS	128027

B. Cleaning and Pre-processing of the Training Dataset

As we already said, the CICIDS2017 dataset dedicated to researchers operating in the field of intrusion detection is composed of eight files. Hence, these files need to be merged into one more comprehensive, one including all the labelled network traffic. The concat() function in Pandas was used to concatenate the set of CSV files and then the to_csv() command could then be used to export the concatenated dataset in CSV format. Fig. 4 shows the workflow adopted to clean, balance and reduce the size of the CICIDS2017 dataset.

C. Experimenting with Machine Learning Techniques to Model benign Traffic

In this part, we will see some machine learning algorithms that we applied on the optimized training dataset CICIDS2017. This experimentation consists in trying a set of algorithms that we will compare between them in order to retain only those effective and efficient that allow us to better modeling a network baseline during its normal operation (benign traffic). Throughout this phase, the Knime tool is used to evaluate the performance of the machine learning algorithms applied on the optimized dataset.

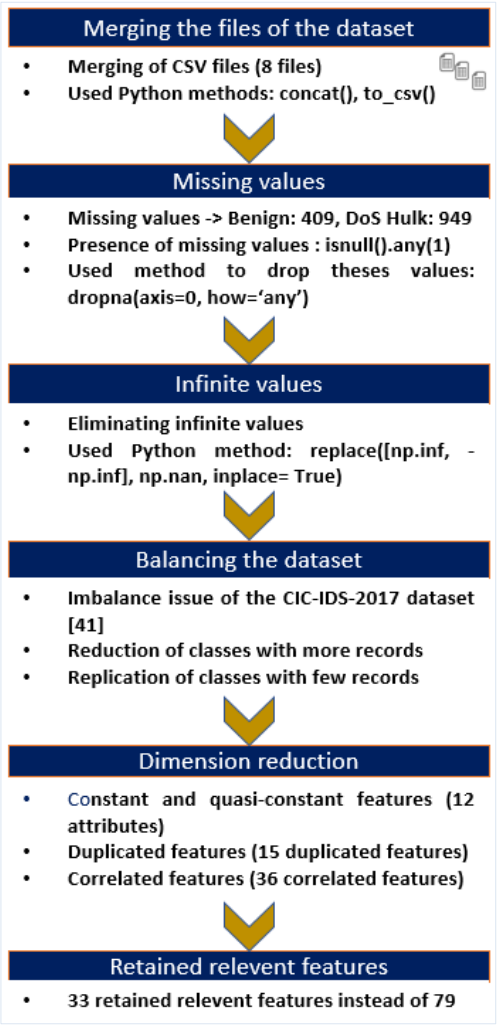

Fig. 4. The Performed Pre-processing and Feature Reduction Tasks to Optimize the CICIDS2017 Dataset.

1) Machine Learning algorithms used to model benign network traffic

a) Decision Tree: The Decision Tree (DT) algorithm was used and performed very well in terms of accuracy, false alarm rates, and execution time. According to the confusion matrix, it turns out that the Decision Tree algorithm provided better detection of all classes of network traffic. The accuracy reached 99.9% while classifying bengin network traffic. The Table III shows the confusion matrix and the Table IV summarizes the obtained results after applying DT algorithm on the optimized CICIDS2017 dataset.

The obtained results are conclusive and highlight the efficiency of the DT algorithm. We are interested in the accuracy of the algorithm with respect to the recognition of benign traffic, especially since our intrusion detection system relies on a baseline of the network during its normal operation. Thus, the Decision Tree was able to detect benign traffic with an accuracy of 99.99% and this, with a total number of false alarms equal to 229 (135 False Negatives (FN) and 94 False Positives (FP)).

TABLE III. CONFUSION MATRIX OF DECISION TREE TO CLASSIFY BENIGN TRAFFIC

Predicted	Actual		
		Positive	Negative
	Positive	74704	94
	Negative	135	174779

TABLE IV. PERFORMANCE STATISTICS OF DECISION TREE ALGORITHM TO CLASSIFY BENIGN TRAFFIC

Metrics	Rate (%)
Recall	99.8
Precision	99.9
F-Measure	99.8
Accuracy	99.9

b) Random Forest: The Random Forest is used to make the NIDS learn the normal behavior of the network. This algorithm performed very well in classifying the different classes of network traffic. As can be seen in Table VI, the detection accuracy reaches 99.8% for benign traffic using Random Forest classifier. RF is very effective in identifying benign traffic and thus designing the network baseline during its normal operation, knowing that the number of false alarms does not exceed 353 (FP: 75 and FN: 278) and with a number of TP equal to 74561 (see Table V).

c) Naïve Bayes: The Naive Bayes (NB) was also tested and unfortunately gave poor detection results for most classes of the dataset. For example, the correct detection of benign traffic is almost zero (accuracy reaches 100% for misclassified instances). Tables VII and VIII below show the statistics related to the use of NB algorithm. The classification of benign traffic is very low compared to other algorithms, as the accuracy does not exceed 70%.

TABLE V. CONFUSION MATRIX OF RANDOM FOREST TO CLASSIFY BENIGN TRAFFIC

Predicted	Actual		
		Positive	Negative
	Positive	74561	75
	Negative	278	174798

TABLE VI. EVALUATION METRICS OF THE RANDOM FOREST ALGORITHM

Metrics	Rate (%)
Recall	99.6
Precision	99.9
F-Measure	99.8
Accuracy	99.8

TABLE VII. CONFUSION MATRIX OF NAIVE BAYES TO CLASSIFY BENIGN TRAFFIC

Predicted	Actual		
		Positive	Negative
	Positive	16	0
	Negative	74823	174873

TABLE VIII. EVALUATION METRICS OF THE NAIVE BAYES ALGORITHM TO CLASSIFY BENIGN TRAFFIC

Metrics	Rate (%)
Recall	0
Precision	100
F-Measure	0
Accuracy	70

d) MultiLayer Perceptron: Using the MultiLayer Perceptron (MLP) technique, the benign traffic was classified with an accuracy of 97%. Tables IX and X highlight the confusion matrix and the statistics obtained after using MLP-based technique.

TABLE IX. CONFUSION MATRIX OF MLP TO CLASSIFY BENIGN TRAFFIC

	Actual		
Predicted		Positive	Negative
	Positive	72227	4957
	Negative	2925	169603

TABLE X. EVALUATION METRICS OF MLP ALGORITHM TO CLASSIFIY BENIGN TRAFFIC

Metrics	Rate (%)
Recall	96.1
Precision	93.6
F-Measure	94.8
Accuracy	96.8

D. Summary of benign Traffic Classification Results

This section summarizes the obtained results after applying the classification algorithms on the optimized CICIDS2017 dataset. We emphasize that we are interested in modeling the network baseline in the absence of any suspicious activity. As a result, the different algorithms used at training time are evaluated based on the classification ability of benign traffic. Thus, Table XI summarizes the results obtained after applying the set of learning algorithms we saw in the previous section.

TABLE XI. SUMMARY OF THE OBTAINED RESULTS USING DIFFERENT MACHINE LEARNING ALGORITHMS

Algorithms	*Recall*	*Precision*	*Accuracy*
Decision Tree	***0.998***	***0.999***	***0.999***
Random Forest	***0.996***	***0.999***	***0.998***
Naïve Bayes	***0***	***1***	***0.7***
Multilayer Perceptron	***0.961***	***0.936***	***0.97***

From the summary table above, it appears that most of the techniques were able to model normal traffic. However, Naive Bayes did not perform well in classifying benign traffic. In addition, the Decision Tree and Random Forest are very efficient in terms of accuracy during training. However, the time complexity of the used algorithms is unfortunately not given in this work and will be the subject of our next article. For example, according to [43], the Decision Tree has a time complexity that is equal to O (mn2) where n is the number of instances and m represents the number of attributes. The temporal complexity metric allows for better evaluation of machine learning methods.

VI. CONCLUSION

It is true that many approaches based on machine learning techniques have been proposed to develop more effective and efficient NIDS. However, existing intrusion detection systems are still not able to detect unknown cyber-attacks more effectively. In this research work, we proposed a new approach based on a Multi-agent model, a Snort IDS and on machine learning techniques. The proposed NIDS is capable of handling network traffic that meets the big data issues in terms of volume and transition speed. First, we analysed the CICIDS2017 dataset with the aim of gaining more visibility on its composition, cleaned it up and removed unnecessary attributes. Then, we tried a set of classifiers on the optimized dataset in order to choose the most efficient algorithm in terms of detection and execution time. Thus, the Decision Tree and Random Forest algorithms give a detection accuracy of more than 99.8% for the detection of benign traffic. However, the work does not end here and the following tasks remain to be accomplished in a future work:

- Definition of how to create rules at Snort when a deviation from the baseline is detected,
- Using the benign traffic model to recognize normal packets in a production environment,
- Using a redundant and powerful module for processing and storing network traffic,
- Testing and validating the NIDS in a real environment.

REFERENCES

[1] Saranya et al. - 2020 - Performance Analysis of Machine Learning Algorithms in Intrusion Detection System: A Review - Procedia Computer Science.

[2] Manimurugan et al. - 2020 - Intrusion Detection in Networks using Crow Search Optimization algorithm with Adaptive Neuro-Fuzzy Inference System – MM.

[3] Kalimuthan, C. Arokia Renjit, J. - 2020 - Review on intrusion detection using feature selection with machine learning techniques.

[4] Aldweesh, Arwa Derhab, Abdelouahid Emam, Ahmed Z. - 2020 - Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues.

[5] Thakkar, Ankit Lohiya, Ritika - 2020 - A Review of the Advancement in Intrusion Detection Datasets - Procedia Computer Science.

[6] Roshan et al. - 2018 - Adaptive and online network intrusion detection system using clustering and Extreme Learning Machines - Journal of the Franklin Institute.

[7] Binbusayyis et al. - 2020 - Comprehensive analysis and recommendation of feature evaluation measures for intrusion detection – Heliyon.

[8] Habeeb, RAA; Nasaruddin, F; Gani, A; Hashem, IAT; Ahmed, E; Imran, M. 'Real-time big data processing for anomaly detection: A survey'. International Journal of Information Management, 2018.

[9] S. Zeadally, E. Adi, Z. Baig and I. A. Khan, "Harnessing Artificial Intelligence Capabilities to Improve Cybersecurity," in IEEE Access, vol. 8, pp. 23817-23837, 2020.

[10] G. Collard, S. Ducroquet, E. Disson and G. Talens, "A definition of Information Security Classification in cybersecurity context," 2017 11th International Conference on Research Challenges in Information Science (RCIS), Brighton, 2017.

[11] S. OUIAZZANE, M. ADDOU and F. BARRAMOU, "A Multi-Agent Model for Network Intrusion Detection," *2019 1st International Conference on Smart Systems and Data Science (ICSSD)*, Rabat, Morocco, 2019.

[12] Said OUIAZZANE, Fatimazahra BARRAMOU and Malika ADDOU, "Towards a Multi-Agent based Network Intrusion Detection System for a Fleet of Drones" – IJACSA.

[13] Dey, Saurabh Ye, Qiang Sampalli, Srinivas - 2019 - A machine learning based intrusion detection scheme for data fusion in mobile clouds involving heterogeneous client networks - Information Fusion.

[14] http://www.ijser.org Intrusion Detection System and Classification of Attacks, 3 (4) (Jul-Aug 2013).

[15] Mohammed H A, Bahaa A, Ismail, and Mohamad F Z, "A New Intrusion Detection System Based on Fast Learning Network and Particle Swarm Optimization", IEEE Access.

[16] Scaife N, Carter H, Traynor P, Butler KR. Cryptolock (and drop it): stopping ransomware attacks on user data. In: 2016 IEEE 36th ICDCS.

[17] Meryem, Amar Ouahidi, Bouabid EL -2020- Hybrid intrusion detection system using machine learning - Network Security.

[18] Shah, Syed Ali Raza Issac, Biju - 2020 - Performance comparison of intrusion detection systems and application of machine learning to Snort system - Future Generation Computer Systems.

[19] Koch R. Towards next-generation intrusion detection. In: 2011 3rd International Conference on Cyber Conflict. IEEE; 2011. p. 1–18.

[20] Shiravi A, Shiravi H, Tavallaee M, Ghorbani AA. Toward developing a systematic approach to generate benchmark datasets for intrusion detection. computers & security. 2012;31(3):357–374.

[21] Gamage, Sunanda Samarabandu, Jagath - 2020 - Deep learning methods in network intrusion detection: A survey and an objective comparison - Journal of Network and Computer Applications

[22] Sharafaldin I, Lashkari AH, Ghorbani AA. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. In: Networking Conference (WCNC). IEEE; 2013. p. 4487–4492].

[23] Zhou et al. - 2020 - Building an efficient intrusion detection system based on feature selection and ensemble classifier - Computer Networks.

[24] Li, Xu Kui Chen et al. - 2020 - Building Auto-Encoder Intrusion Detection System based on random forest feature selection - Computers and Security.

[25] Kasongo et al. - 2020 - A deep learning method with wrapper based feature extraction for wireless intrusion detection system - Computers and Security.

[26] Al-Yaseen et al. - 2017 - Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system - Expert Systems with Application.

[27] Alazzam, Hadeel Sharieh, Ahmad Sabri, Khair Eddin - 2020 - A feature selection algorithm for intrusion detection system based on Pigeon Inspired Optimizer - Expert Systems with Applications.

[28] Lopez-Martin, Manuel Carro, Belen Sanchez-Esguevillas, Antonio - 2020 - Application of deep reinforcement learning to intrusion detection for supervised problems - Expert Systems with Applications.

[29] Sovilj, Dušan Budnarain, Paul Sanner, Scott Salmon, Geoff Rao, Mohan - 2020 - A comparative evaluation of unsupervised deep architectures for intrusion detection in sequential data streams - Expert Systems with Applications.

[30] Shah, Syed Ali Raza Issac, Biju - 2018 - Performance comparison of intrusion detection systems and application of machine learning to Snort system - Future Generation Computer Systems.

[31] Abusitta, Adel Bellaiche, Martine Dagenais, Michel Halabi, Talal - 2019 - A deep learning approach for proactive multi-cloud cooperative intrusion detection system - Future Generation Computer Systems.

[32] Kasongo, Sydney Mambwe Sun, Yanxia - 2020 - A Deep Gated Recurrent Unit based model for wireless intrusion detection system - ICT Express.

[33] Hassan, Mohammad Mehedi Gumaei, Abdu Alsanad, Ahmed Alrubaian, Majed Fortino, Giancarlo - 2020 - A hybrid deep learning model for efficient intrusion detection in big data environment.

[34] Ferrag, Mohamed Amine Maglaras, Leandros Moschoyiannis, Sotiris Janicke, Helge - 2020 - Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study - Journal of Information Security and Applications.

[35] Wu, Zhendong Wang, Jingjing Hu, Liqing Zhang, Zhang Wu, Han - 2020 - A network intrusion detection method based on semantic Re-encoding and deep learning – JNCA.

[36] Li, Wenjuan Meng, Weizhi Au, Man Ho - 2020 - Enhancing collaborative intrusion detection via disagreement-based semi-supervised learning in IoT environments – JNCA.

[37] Almogren, Ahmad S. - 2020 - Intrusion detection in Edge-of-Things computing - Journal of Parallel and Distributed Computing.

[38] Roshan, Setareh Miche, Yoan Akusok, Anton Lendasse, Amaury – 2018 - Adaptive and online network intrusion detection system using clustering and Extreme Learning Machines – JFI.

[39] Satheesh et al. - 2020 - Flow-based Anomaly Intrusion Detection using Machine Learning Model with Software Defined Networking for OpenFlow Network - Microprocessors and Microsystems.

[40] Jie, Y; Xin, C; Xudong, X; Jianxiong, W. 'HIDS-DT: An Effective Hybrid Intrusion Detection System Based on Decision Tree'. IEEE

[41] Panigrahi R, Borah S. A detailed analysis of CICIDS2017 dataset for designing Intrusion Detection Systems.

[42] Binbusayyis, Adel Vaiyapuri, Thavavel - 2020 - Comprehensive analysis and recommendation of feature evaluation measures for intrusion detection-Heliyon.

[43] K. Shaukat, S. Luo, S. Chen and D. Liu, "Cyber Threat Detection Using Machine Learning Techniques: A Performance Evaluation Perspective," 2020 International Conference on Cyber Warfare and Security (ICCWS), 2020, pp. 1-6, doi: 10.1109/ICCWS48432.2020.9292388.

[44] Shaukat, K.; Luo, S.; Varadharajan, V.; Hameed, I.A.; Chen, S.; Liu, D.; Li, J. Performance Comparison and Current Challenges of Using Machine Learning Techniques in Cybersecurity. *Energies* 2020, *13*, 2509. https://doi.org/10.3390/en13102509.