Predicting DOS-DDOS Attacks: Review and Evaluation Study of Feature Selection Methods based on Wrapper Process

Kawtar BOUZOUBAA[1], Benayad NSIRI[3]
M2CS, Research Center STIS
National Graduate School of Arts and Crafts of Rabat
(ENSAM) Mohammed V University in Rabat
Rabat, Morocco

Youssef TAHER[2]
Center of Guidance and Planning (COPE)
Rabat, Morocco

Abstract—**Now-a-days, Cybersecurity attacks are becoming increasingly sophisticated and presenting a growing threat to individuals, private and public sectors, especially the Denial Of Service attack (DOS) and its variant Distributed Denial Of Service (DDOS). Dealing with these dangerous threats by using traditional mitigation solutions suffers from several limits and performance issues. To overcome these limitations, Machine Learning (ML) has become one of the key techniques to enrich, complement and enhance the traditional security experiences. In this context, we focus on one of the key processes that improve and optimize Machine Learning DOS-DDOS predicting models: DOS-DDOS feature selection process, particularly the wrapper process. By studying different DOS-DDOS datasets, algorithms and results of several research projects, we have reviewed and evaluated the impact on used wrapper strategies, number of DOS-DDOS features, and many commonly used metrics to evaluate DOS-DDOS prediction models based on the optimized DOS-DDOS features. In this paper, we present three important dashboards that are essential to understand the performance of three wrapper strategies commonly used in DOS-DDOS ML systems: heuristic search algorithms, meta-heuristic search and random search methods. Based on this review and evaluation study, we can observe some of wrapper strategies, algorithms, DOS-DDOS features with a relevant impact can be selected to improve the DOS-DDOS ML existing solutions.**

Keywords—DOS-DDOS attacks; feature selection; wrapper process; machine learning

I. INTRODUCTION

With the exponential proliferation of Internet users, the network traffic has known a massive generation of data. These data are coming from individuals, private and public organizations. Moreover, the hard complexity of the Internet architecture and its interdependent suffers from different vulnerabilities, threats and risks ([1], [2]). Consequently, the attackers find an impressive amount of vulnerable systems [3].

Nowadays, cybersecurity attacks are becoming increasingly sophisticated, particularly the infrastructure attacks that make security analysis systems more vulnerable to several failures [1]. One of these most famous threats is Denial Of Service attack (DOS) and its variant Distributed Denial Of Service (DDOS) ([4],[5]). These serious and dangerous attacks violate the availability of information systems, which is a pillar of information security ([6],[5]). The attackers seek to target computer systems, network devices, services and web applications to consume their CPU power, bandwidth, memory and processing time ([7], [3]).

The DDOS attack has the same purpose but with the difference of using intermediate of multiple networks between the attacker and its target ([7],[8]). This technique allows the attacker to amplify its attack with orchestrating a simultaneous sending of an excessive number of unwanted computing requests to its victim to overload its computing capacity.

To deal with these DOS-DDOS attacks, some traditional mechanisms are deployed such as firewalls, software updates, antivirus, Intrusion Detection Systems (IDS), etc.

However, many challenges and limits hinder these traditional techniques [6]. To overcome these limitations and drawbacks, Machine Learning (ML) techniques can be used as artificial intelligence systems to enrich, complement and enhance the traditional security experiences.

One of the key and critical pre-processing phases to success these DOS-DDOS ML models is feature selection. This process selects the most representatives DOS-DDOS characteristics from the initially DOS-DDOS dataset by eradicating those that are redundant and insignificant. Consequently, the obtained features subset improves the execution time, the detection rate and the accuracy of the used DOS-DDOS models.

In this context, this investigation presents a review and evaluation study related to DOS-DDOS attacks prediction based on one of the effective methods to select relevant DOS-DDOS features: Wrapper process.

This paper is organized as follows: In Section 2 we study some traditional mitigation solutions and their limits. Section 3 describes the interest of using machine learning (ML) in DOS- DDOS attacks prevention. Section 4 exposes the impact of feature selection on DOS-DDOS machine learning projects. In Section 5 we review and we evaluate recent and relevant feature selection results obtained by using three commonly used wrapper strategies: heuristic search algorithms, meta-heuristic search and random search methods. Finally, Section 6 presents our conclusions.

II. Dealing with DOS and DDOS: Traditional Mitigation and Solutions

DOS-DDOS attacks can take many forms such as SYN flood, SYN-ACK-ACK flood, UDP flood, ICMP flood, and so on. To deal with these forms of threats, many traditional, external and internal DOS-DDOS mitigation solutions are developed such as bandwidth provisioning, software updates, firewalls, antivirus software and Intrusion Detection Systems (IDS), etc. In the paragraph below, we discuss briefly these traditional solutions and their limits.

Generally, the use of firewall solution provides many mitigation solutions such as filter-based forwarding at logical interfaces, blocking of certain types of packets to reach a routing engine and packet counter and protection of a routing engine from DOS-DDOS attacks ([9],[10]). However, firewall solutions suffer from many lacks of security. As an example, the attacker can modify his DOS-DDOS attacks and make it legitimate.

The software updates keep the software up to date to avoid DOS-DDOS attacks on the application layer (the highest abstraction layer of the TCP/IP model) [11]. However, the irregularity of these updates creates a gateway to the attackers to modify the contents of memories (buffer overflow).

The Intrusion Detection System IDS (Hardware/Software solutions) is a complemented security for the firewall solutions. This solution is a common way often used to analyze and detect DOS-DDOS attacks [12]. IDS techniques are used in the aims to detect, classify and respond to DOS-DDOS actions that affect the integrity, the confidentiality or the availability of any network resources [13]. These systems are mainly based on two detection methods [14]: Misuse Detection (MD) and Anomaly Detection (AD).

The Misuse Detection is also known as Signature Detection, Pattern Detection, Knowledge-Based or Rule-based detection. This technique is one of the most common methods of Antivirus. It filters malicious packet of the known attacks thanks to its signature database of known attacks. It detects efficiently known attacks with low false positive. Nevertheless, it shows limits on detecting new forms of threats and many variants of known attacks.

The Anomaly Detection supervises the behavior of network traffic. It alerts the system at the slightest changes compared to the normal behavior. This method can detect new forms of attacks but generates high false positives and doesn't give clear information about the malicious events in some forms of attacks. Moreover, it is not feasible to IDS to manipulate high dimensional variables. Consequently, this technique can affect the efficiency and the velocity in detecting intrusions ([15],[16], [17]).

In addition to the limitations and drawbacks mentioned above, traditional techniques are hindered by many others challenges [6]. As an example, many traditional strategies of security are not sufficient to protect information systems against the new forms of DOS-DDOS attacks, need extra-storage and computational resources due to the high level of network traffic, suffer from a lack of source attacks information and are unable to detect and prevent many DOS-DDOS attacks in real-time.

To overcome these drawbacks, Machine Learning has become one of the key techniques to enrich and complement these traditional security experiences. In the paragraph below we discuss briefly the benefits that can be attained by using ML- techniques in DOS-DDOS attacks prevention.

III. The Use of Machine Learning in DOS-DDOS Attacks Prevention

Machine Learning (ML) is an evolutionary field of Artificial Intelligence (AI) composed of a set of rules, methods and functions [18]. Applied to deal with many challenges in DOS-DDOS attacks, ML algorithms can learn from DOS-DDOS datasets and discover hidden knowledge from them [19].

By finding interesting DOS-DDOS patterns from training DOS-DDOS data, ML algorithms allow preventing and predicting many recent forms of DOS-DDOS behaviors.

Contrary to the traditional security solutions, ML models are powerful tools that can analyze in real time high dimensional DOS-DDOS traffic [20], classify the behavior of the DOS-DDOS traffic to determine the normal one from the abnormal and predict with high accuracy DOS-DDOS attacks before they happen.

Based on DOS-DDOS security modeling process (Fig. 1) and many common algorithms like K-Nearest Neighbors Algorithm (KNN), Support Vector Machines (SVM), Random Forest (RF) as well as Naïve Bayes (NB), etc. many recent research projects have shown other important preventing benefits of ML algorithms compared to the existing traditional solutions ([1], [12], [21]).

Feature selection is one of the critical pre-processing process to succeed and to improve the benefits mentioned above. In the paragraph below, we summarize the benefits of this process.

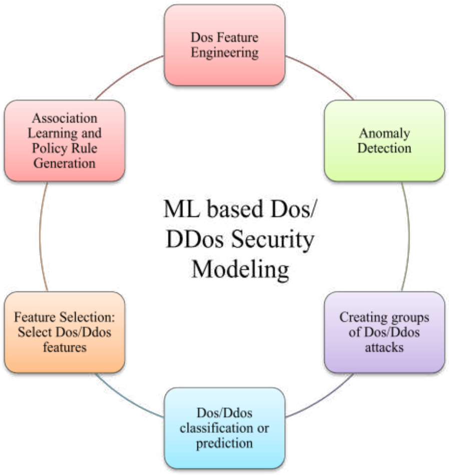
Dos Feature Engineering
Anomaly Detection
Creating groups of Dos/Ddos attacks
Dos/Ddos classification or prediction
Feature Selection: Select Dos/Ddos features
Association Learning and Policy Rule Generation
ML based Dos/ DDos Security Modeling

Fig. 1. Machine Learning DOS-DDOS Security Modeling Process

IV. Impact of Feature Selection Process DOS-DDOS Machine Learning Projects

Feature selection is one of the most critical pre-processing process in building DOS-DDOS Machine Learning (ML) models. This process is the first and crucial phase to improve the prediction accuracy, the detection rate and to reduce the execution time of DOS-DDOS models [22].

According to Bindra et al. [23], feature selection methods allow the DOS-DDOS security systems to distinguish DOS-DDOS attacks by using a minimum number of the most important features from network streams.

Applied to DOS-DDOS ML algorithms, feature selection is focused on selecting small and concise DOS-DDOS sets of characteristics describing the ML models [24]. It avoids the used features to contain redundant (correlation with other features) and noisier information of DOS-DDOS attacks without losing any piece of information. Consequently, it reduces the high memory requirements of security systems based on ML models ([25], [26], [27]).

Generally, the existing DOS-DDOS ML security systems use three commonly main categories of feature selection approaches: Filter, Wrapper and Hybrid methods [28].

The Filter methods are based on statistical methods which evaluate the relevance of DOS-DDOS features independently of any machine learning algorithms [27]. As a faster solution that computationally costs less, these methods are often used in high dimensional DOS-DDOS traffic ([29],[30]). However, the evaluation of individual information cannot take into consideration the correlation between the DOS-DDOS features. Consequently, the final DOS-DDOS subset can contain redundancy because some DOS-DDOS features can have the same ranking.

The wrapper strategies use a predetermined algorithm and its performance to assess the optimal DOS-DDOS subset features [31]. It executed in an iterative process, and at each iteration a new subset of DOS-DDOS features is generated to be evaluated by the classification algorithm [32]. The criterion of selection is principally based on the cross-validation accuracy during the DOS-DDOS training data [33].

The Hybrid method is a combination between filter method followed by wrapper approach, which offers the advantages of the two previous methods. It exploits their different criteria in different search stages [34].

V. Related Work

A. Objective of the Study

To detect and prevent DOS-DDOS attacks accurately, wrapper methods one of the *most effective* strategies to identify informative DOS-DDOS feature subsets from many high-dimensional DOS-DDOS network streams. This approach of feature selection is often addressed in many security solutions based on ML tasks. Indeed, increasing number of research projects have shown that many wrapper strategies can have an important impact on Accuracy, Detection Rate and time execution of existing DOS-DDOS ML systems.

In this context, we decided to focus our attention on the assessment of the performance of many DOS-DDOS experiments based on wrapper strategies and machine learning algorithms.

By studying different DOS-DDOS datasets, algorithms and recent results of several research projects, we review and we assess the impact of many recent wrapper strategies applied to predicting DOS-DDOS attacks. We have taken a more focused look at the impact of these strategies on number of DOS-DDOS features, detection rates, execution times and accuracies of DOS-DDOS attacks prediction.

We present four dashboards that are essential to understand the performances of three wrapper strategies commonly used in DOS-DDOS ML systems: heuristic search algorithms, meta-heuristic search and random search methods.

B. Review and Evaluation Study of Feature Selection Methods based on Wrapper Process

1) Used Datasets: To evaluate the performance of the wrapper strategies used in DOS-DDOS machine learning models, we start our review by studying relevant DOS-DDOS datasets commonly used by several DOS-DDOS research projects. These datasets are cited below:

The Knowledge Discovery and Data Mining (KDD'99) dataset was built based on the synthetic data captured in DARPA'98. This dataset is mainly composed of redundant records. Moreover, this configuration forces ML algorithms to learn less about infrequent records than the redundant ones. The inequality of attacks distribution between training and testing phase made the cross-validation more complicated.

This dataset is composed of four main families of attacks and forty one features.

The NSL_KDD was created to overcome the limits of the KDD'99 [35]. However, the main disadvantage of the NSL_KDD dataset, it does not include the modern low footprint attacks scenarios like the KDD'99.

The UNSW_NB15 is composed of nine family attacks and forty nine features. It includes a hybrid of the real modern normal behaviors and the synthetic attack activities [35].

Cyber Range Lab of the Australian Centre for Cyber Security (ACCS) is a dataset mainly composed of hybrid modern normal activities and attacks behaviors. It is composed of forty-seven features[36].

2) Use model evaluation metrics: To evaluate the reviewed DOS-DDOS Wrapper strategies, we have selected different metrics [37]. These metrics namely are: Classification Accuracy (Acc), Detection Rate (DR), Recall (Re), Precision (Pr), Specificity (Sp), Sensitivity (Sen), F-Measure (FM), False Alert Rate (FAR), False Negative (FN) and Time model execution (T).

The formulas associated with these metrics are listed above:

$$\mathrm{Acc} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \quad (1)$$

$$\mathrm{DR} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \quad (2)$$

$$\mathrm{Re} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \quad (3)$$

$$\mathrm{Pr} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \quad (4)$$

$$\mathrm{Sp} = \frac{\mathrm{TN}}{(\mathrm{TN+FP})} \quad (5)$$

$$\mathrm{Sen} = \frac{\mathrm{TP}}{(\mathrm{TP+FN})} \quad (6)$$

$$\mathrm{FM} = \frac{2\times\mathrm{Re}\times\mathrm{Pr}}{(\mathrm{Re+Pr})} \quad (7)$$

$$\mathrm{FAR} = \frac{\mathrm{FP}}{\mathrm{FP+TN}} \quad (8)$$

Where: TP is True Positive: correct positive prediction. TN is True Negative: correct negative prediction. FN is False Negative: incorrect negative prediction and FP is False Positive: incorrect positive prediction.

3) Impact of used DOS-DDOS datasets and algorithms on the wrapper process: Generally, the performance of DOS-DDOS prediction models based on the Wrapper process depends strongly on the used ML algorithms and datasets. As shown in Table I, many algorithms performed well in detecting DOS-DDOS attacks compared to others. The accuracy can range from Acc=62.5% by using KDD'99 dataset and SVM algorithm to Acc=99.92% with Decision Tree J.48 algorithm and KDD'99 dataset. Indeed, according to the experiment of Jalill et al. (2010) [38] based on the KDD'99 dataset, the Support Vector Machine (SVM) algorithm has a serious problem in accurately detecting DOS-DDOS attacks compared to the Decision Tree J.48 algorithm which shows high prediction accuracy that exceed 99%.

TABLE I. IMPACT OF USED DOS-DDOS DATASETS AND ALGORITHMS ON THE WRAPPER PROCESS

Reference	Dataset	Algorithm	Accuracy(%)
Jalill et al. [38]	KDD'99	SVM	62.5
		J.48	99.7
Katkar and Kulkarni. [40]	KDD'99	J.48	99.92
		REPTree	99.56
		NB	87.50
		BN	99.68
		Sequential Minimal Optimization (SMO)	99.72
		REPTree + J48 +BN	99.94
Bellouch et al. [39]	UNSW_NB15	SVM	92.28
		NB	74.19
		C4.5	95.82
		RF	97.49

The experiments based on the NB, C4.5, RF algorithms and UNSW_NB15 dataset realized by Bellouch et al. (2018) [39], has shown that the prediction accuracy obtained by RF (Acc_$_{RF}$ = 99.94%) is better than C4.5 (Acc_$_{C4..5}$ = 95.82%) and SVM (Acc_$_{SVM}$ = 92.28%). The NB algorithm shows less accuracy (Acc_$_{NB}$ = 74. 19 %) compared to RF, C4.5 and SVM.

The Bayesian Network (BN) algorithm used in the experiment of Katkar and Kulkarni [40] achieved good accuracy (Acc_$_{BN}$ = 99.68%) in detecting DOS-DDOS attacks thanks to its capacity of detecting anomalies in a multi-class [41].

By comparing the experiments carried out by Jalill et al.[38] and Katkar and Kulkarni [40], we have observed that SVM algorithm predict DOS-DDOS more accurately on the dataset UNSW_NB15 compared to the KDD'99 dataset (Acc_$_{SVM\_UNSW\_NB}$ = 92.28% > Acc_$_{SVM\_KDD}$ = 62.5 %). This important difference according to W. Xingzhu [42] is caused by the redundant records on the KDD'99 dataset and SVM has slower training on high dimensional datasets.

4) DOS-DDOS feature selection based on wrapper process and heuristic search algorithms: Based on heuristic functions or cost measures, wrapper strategies using heuristic search algorithms optimize and iteratively improve the process of DOS-DDOS feature selection [43].

Many heuristic searches such as SFS (Sequential Forward search), SBS (Sequential Backward search), LRS (Plus L Minus R Selection), RELR (Random Effect Logistic Regression), and GFR (Gradually feature removal method) have been used by many recent important research projects to solve accurately the problem of DOS-DDOS feature selection.

We discuss these projects in the paragraph below. At the end of this subsection, we present our first dashboard (Tables IIA, IIB, IIC) to summarize and to compare the performances of these strategies.

As an example of wrapper strategies based on heuristic search algorithms, we can cite the important investigation of Kavitha and Chrita (2010) [44]. In this study, the authors used the Best First Search (BFS) method. They selected two subsets composed simultaneously of seven and fourteen DOS-DDOS features. They applied four classifying algorithms: ID3, J48, NB and One R. These experiments have shown that ID3 and J.48 using a subset composed of fourteen DOS-DDOS features has the highest accuracy (Acc = 99%). One R and NB performed well in execution time (T=0.5s) with only seven features. The NB classifier achieved the highest specificity with Sp_$_{NB}$ = 99% by using seven features and Sp_$_{NB}$ =100% by using fourteen features.

Mok et al. (2010) [45] used Random Effect Logistic Regression (RELR) with a fixed Logistic regression (LR). This method selected five DOS-DDOS features by using the Stepwise Variable Selection Search (SVSS) strategy based on the KDD'99 dataset. The method achieved an accuracy equal to 98.74%.

TABLE II. (A): WRAPPER METHOD BASED ON HEURISTIC SEARCH (HS)

DOS-DDOS feature selection projects based on wrapper methods	**DOS- DDOS used dataset**	**Used wrapper strategy**	**Number of DOS - DDOS features**		**Used classifier**	**Used Metrics**	**Metrics Values with FS**		**Metrics Values without FS**
Kavitha, and Chitra (2010) [44]	KDD’99	BFS	7	14	ID3	Accuracy Sensitivity Specifity Time (s)	97% 97% 97% 1.49	99% 100% 98% 4.01	99% 98% 100% NA
					J48	Accuracy Sensitivity Specifity Time (s)	97% 97% 97% 1.20	99% 99.5% 97.5% 1. 86	99.9% 97.8% 99.9% NA
					NB	Accuracy Sensitivity Specifity Time (s)	96% 92% 99% 0.05	97% 94% 100% 0.09	99% 98% 100% NA
					OneR	Accuracy Sensitivity Specifity Time (s)	86% 74% 99% 0.05	97% 72% 92% 0.16	99.5% 98% 99.7% NA
Mok et al. (2010) [45]	KDD’99	Stepwise	5		RLER	Accuracy	98.74%		NA
Ahmad et al. (2011) [46]	KDD’99	PCA-GA	12		MLP	Accuracy Time (h)	99% 72		NA
Yinhui et al. [47]	KDD’99	SBS-GFR	19		SVM	Accuracy Time(s)	98.62% 2.37		98.67% 3.97

TABLE II- (B): WRAPPER METHOD BASED ON HEURISTIC SEARCH (HS)

DOS-DDOS feature selection projects based on wrapper methods	**DOS- DDOS used dataset**	**Used wrapper strategy**	**Number of DOS - DDOS features**		**Used classifier**	**Used Metrics**	**Metrics Values with FS**		**Metrics Values without FS**
Zhang and Wang (2013) [48]	NSL_KDD	SBS-BN	11		BN	Accuracy Time(s)	98.98% 4.73		95.7% 18.94
Al-Jarrah et al.(2014) [49]	KDD’99	FSR-RF	15		RF	Accuracy	99.90%		99.89%
		BER-RF	14				99.88%		
Lee et al. (2017) [50]	NSL_KDD	SFFS-RF	10		C4.5	Accuracy Detection Rate FAR Time(s)	99.89% 99.9% 0.1 0.18		NA NA 1.07 NA
Harish and Manju (2018) [51]	KDD’99	FDR + PLR	20	40	KNN	Accuracy Time(s)	98.5% 17.98	99.0% 32.95	NA
		FDR +SFS	25			Accuracy Time(s)	98.27 17.74		NA
		FDR +SBS	40			Accuracy Time(s)	98.78% 32.18		NA NA
Houseini Soodeh and Mehrdad (2019) [52]	NSL_KDD	Forward Feature Selection	12		NB	Accuracy Precision Recall F-measure	93.1% 93.6% 87.3% 92.7%		NA
			14		RF	Accuracy Precision Recall F-measure	98.9% 99.6% 99.8% 99.7%		NA
			10		DT	Accuracy Precision Recall F-measure	98.2% 99.4% 99.8% 99.6%		NA

			20	MLP	Accuracy Precision Recall F-measure	96.1% 93.4% 91.8% 94.9%	NA
			11	KNN	Accuracy Precision Recall F-measure	97.7% 99.8% 99.8% 99.8%	NA
Malhotra and Sharma (2019) [53]	NSL_KDD	CfsSubsetEval + BestFirst	6	RF	Accuracy Time (s)	99,41% 66.82	99,91% 191.06
				Bagging	Accuracy Time (s)	99,35% 17,7	99,84% 109.9%
				PART	Accuracy Time (s)	99,37% 8.07	99,83% 99.1
				J48	Accuracy Time (s)	99,78% 7.95	99,78% 61.68
Wang et al.(2020) [54]	NSL_KDD	SBS-MLP	31	MLP	Accuracy Detection Rate FAR	97.66% 94.88% 0.62%	97.61% 94.78% 0.63%
Polat, and Cetin (2020) [55]	Their Dataset composed of 12 Features	SFFS	10	SVM	Accuracy Sensitivity Specificity Precision F_measure	92.15% 90.20% 97.26% 90.23% 90.21%	92.11% 88.71% 96.93% 91.42% 89.91%
			6	KNN	Accuracy Sensitivity Specificity Precision F_measure	98.30% 97.73% 99.45% 97.72% 97.70%	95.67% 93.87% 98.01% 97.05% 95.30%

TABLE II-(C): WRAPPER METHOD BASED ON HEURISTIC SEARCH (HS)

DOS-DDOS feature selection projects based on wrapper methods	**DOS- DDOS used dataset**	**Used wrapper strategy**	**Number of DOS -DDOS features**	**Used classifier**	**Used Metrics**	**Metrics Values with FS**	**Metrics Values without FS**
Polat, and Cetin (2020) [55]	Their Dataset composed of 12 Features	SFFS	6	ANN	Accuracy Sensitivity Specificity Precision F_measure	91.44% 87.82% 97.31% 88.11% 87.89%	91.07% 87.27% 96.58% 89.89% 88.45%
			8	NB	Accuracy Sensitivity Specificity Precision F_measure	94.87% 92.05% 98.43% 93.29% 92.01%	94.48% 91.77% 98.29% 92.94% 91.79%
Alabdulwahab and Moon (2020) [31]	NSL_KDD	CfsSubsetEval + BestFirst	6	RePTree	Accuracy Time(s)	99,44% 5,76	99,83% 3.59
				Logiboost	Accuracy Time(s)	94,15% 9,96	97,1% 18.3
				RBF	Accuracy Time(s)	90,6% 45.91	97,95% 81.01
				BayesNet	Accuracy Time(s)	96,26% 5.64	97,17% 4.69
				SMO	Accuracy Time(s)	89,09% 514.7	97,4% 1137.71
				NBTree	Accuracy Time(s)	99,46% 14.23	99,87% 213.18
Umar et al. (2020) [56]	UNSW_NB15	Best First Forward-DT	19	ANN	Accuracy Detection Rate FAR Time(s)	82.08% 97.94% 37.36% 240	86.00% 98.62% 29.45% 660

					SVM	Accuracy Detection Rate FAR Time(s)	79.11% 99.31% 45.64% 15540		81.6% 99.64% 40.51 10860	
					KNN	Accuracy Detection Rate FAR Time(s)	83.21% 96.44% 33.01% 600		84.78% 96.46% 29.53% 1020	
					RF	Accuracy Detection Rate FAR Time(s)	86.41% 97.95% 27.73% 37.8		86.82% 98.7% 27.74% 44.4	
					NB	Accuracy Detection Rate FAR Time(s)	55.61% 19.38% 0.01% 2.86		55.61% 19.39% 0.01% 4.64	
Umar and Chen (2020) [57]	UNSW_NB15	NSL_KDD	Best First - DT	20	ANN	Accuracy Detection Rate FAR Time(s)	94.32% 98.48% 14.56% 325	98.9% 99.0% 1.11% 123	94.62% 97.54% 11.64% 348	99.6% 99.6% 0.23% 94
					SVM	Accuracy Detection Rate FAR Time(s)	93.56% 99.54% 19.19% 10236.6	98.0% 97.1% 1.17% 921.6	93.67% 99.63% 19.14% 5213.4	98.5% 98.1% 1.08% 972.6
					KNN	Accuracy Detection Rate FAR Time(s)	95.8% 97.28% 7.36% 502.8	99.1% 99.2% 0.97% 331.2	93.81% 96.24% 11.42% 747.6	99.5% 99.4% 0.36% 563.4
					RF	Accuracy Detection Rate FAR Time(s)	98.51% 99.17% 2.89% 33.6	99.7% 99.7% 0.22% 13.2	95.74% 97.84% 8.77% 32.4	98.8% 99.7% 0.1% 15

Ahmad et al. (2011) [46] used Principal Components Analysis (PCA) to reduce the features and to choose the highest eighteen values. Genetic Algorithm (GA) was applied as wrapper method to the reduce space. This method selected twelve DOS-DDOS features. By using the Multi Layer Perceptron (MLP) as classifier on the output of GA and the KDD'99 dataset, this model has shown high accuracy ($\text{Acc}\_{\text{MLP}} = 99\%$) by using a minimum of features equal to 12 and the time of execution equal to 72 h.

L. Yinhui et al. (2012) [47] applied Gradually Feature Removal method (GFR) which selected nineteen best DOS-DDOS features. This strategy was based on SBS as search strategy and SVM as classifier. The accuracy of this model has been slightly reduced ($\text{Acc}\_{(19\ \text{features})} = 98.62\% < \text{Acc}\_{(42\ \text{features})} = 98.67\%$) by using a wrapper step. The execution time has been reduced from $\text{T}\ {}\_{(42\ \text{features})} = 18.94\text{s}$ to $\text{T}\_{(19\ \text{features})} = 3.73$ s.

Zhang and Wang (2013) [48] adopted SBS-BN and Bayesian network approach as a wrapper strategy. This experiment selected three best DOS-DDOS features and achieved good accuracy ($\text{Acc}\_{(3\ \text{features})} = 98.98\% > \text{Acc}\_{(42\ \text{features})} = 95.7\%$) with an interesting time of execution ($\text{T}\_{(3\ \text{features})} = 2.37\text{s} < \text{T}\_{(42\ \text{features})} = 3.97\text{s}$).

Al-Jarrah et al. (2014) [49] proposed a set of RF algorithm with forward and backward elimination ranking features selection techniques. This experiment demonstrated that FSR-RF outperforms with fifteen best features, BER-RF with fourteen features and RF with all used DOS-DDOS features: ($\text{Acc}\_{(15\ \text{features})} = 99.98\% > \text{Acc}\_{(14\ \text{features})} = 99.88\ \%$) and ($\text{Acc}\_{(15\ \text{features})} = 99.98\% > \text{Acc}\_{(42\ \text{features})} = 99.89\%$).

J. Lee et al. (2017) [50] proposed SFFS-RFC to generate DOS-DDOS features subset and to measure the performance of each subset. This experiment has shown that SFFS-RFS improved the performance of the accuracy and the detection rate of attacks classification with only ten DOS-DDOS ($\text{Acc}\_{(10\ \text{features})} = 99.89\%$ and $\text{DR}\_{(10\ \text{features}\ )} = 99.9\%$). It realized a fewer FAR ($\text{FAR}\_{(10\ \text{features})} = 0.1\% < \text{FAR}\_{(41\ \text{features})} = 1.7\%$) compared to the existing methods using the classifier C4.5 and reduced the execution time ($\text{T}\_{(10\ \text{features})} = 0.18$ s).

Harish and Manju (2018) [51] combined the Fisher Ratio Discrimination (FRD) with three different search strategies: SFS, SBS and LRS. They concluded that FDR using LRS, KNN and twenty DOS-DDOS features outperformed other methods. Thanks to its capacity to remove non-performing DOS-DDOS features from the initial subset, this strategy achieved a better accuracy with twenty features ($\text{Acc}\_{(20\ \text{features})} = 98.87\% > \text{Acc}\_{\text{SFS\_(25 features)}} = 98.27\%$) compared to FDR-SFS which selected 25 features. However, the execution time of FDR-SFS is less than FDR-LRS ($\text{T}\_{\text{SFS\_(25 features)}} = 17.74\ \text{s} < \text{T}\_{\text{SFS\_(20 features)}} = 17.98$ s). On the other side the FDR-LRS with forty features showed a good accuracy compared to the accuracy of FDR-SBS with the same number of features ($\text{Acc}\_{\text{LRS\_(40 features)}} = 99.09\% > \text{Acc}\_{\text{SBS\_(40 features)}} = 98.78\%$). However the execution time of FDR-SBS is better compared to FDR-LRS ($\text{T}\_{\text{SBS\_(40 features)}} = 32.18\text{s} < \text{T}\_{\text{LRS\_(40 features)}} = 32.95\text{s}$).

Soodeh and Mehrdad (2019) [52] proposed a new framework composed of a hybridization of different algorithms. The objective of this framework is to handle new types of attacks better than other existing frameworks based on Forward Feature Selection (FFS). By using NSL_KDD dataset, this framework has shown that RF outperformed other algorithms with only thirteen features in attack detection accuracy ($\text{Acc}_{\text{\_(13 features)}} = 98.9\%$). In the case of DOS-DDOS attacks, the KNN classifier has achieved the highest precision with eleven features ($\text{Pr}_{\text{\_(11 features)}}= 99.8\%$). The classifiers RF, DT and KNN achieved the highest Recall value (Re = 99.8 %), and the highest F-measure ($\text{FM}_{\text{\_RF\_(14 features)}} = 99.7\%$, $\text{FM}_{\text{\_DT\_(10 features)}} = 99.6\%$, and $\text{FM}_{\text{\_KNN\_(11 features)}} = 99.8\%$). The classifier NB showed the lowest measured values of all these metrics: $\text{Acc}_{\text{\_NB}} = 93.10\%$, $\text{Pr}_{\text{\_NB}} = 93.6\%$, Re = 87.3%, $\text{FM}_{\text{\_NB}} = 92.7\%$.

Malhotra and Sharma (2019) [53] used CfsSubsetEval and Best First as wrapper method. Based on NSL_KDD dataset and RF Bagging, PART and J.48 algorithms, this strategy selected eight best DOS-DDOS features. It increased slightly the accuracy and decreased significantly the execution time for all the classifiers. The accuracy of J.48 is 99.78% by using 6 and 42 features. However, this strategy decreased the execution time ($\text{T}_{\text{\_J.48\_(42 features)}} = 61.68\text{s} > \text{T}_{\text{\_J.48\_(6 features)}} = 7.95\text{s}$). The RF model decreased slightly the accuracy ($\text{Acc}_{\text{\_RF\_(6 features)}} = 99.41\% < \text{Acc}_{\text{\_RF\_(42 features)}} = 99.91\%$), and decreased drastically the execution time ($\text{T}_{\text{\_RF\_(6 features)}} = 66.82\text{s} < \text{T}_{\text{\_RF\_(42 features)}} = 191.06$ s).

M. Wang et al. (2020) [54] combined SBS with Multi Layer Perceptron (MLP) to select the optimal DOS-DDOS features by using NSL_KDD dataset. This experiment showed that SBS-MLP can find an optimal DOS-DDOS feature subset and performed better accuracy than the full DOS-DDOS feature set among all the MLP-based detection methods ($\text{Acc}_{\text{\_(31 features)}} = 97.66\% > \text{Acc}_{\text{\_(42 features)}}= 97.61\%$). It enhanced the detection rate ($\text{DR}_{\text{\_(31 features)}} = 94.88\ \%> \text{DR}_{\text{\_(42 features)}} = 94.78\%$). It decreased the FAR value ($\text{FAR}_{\text{\_(31 features)}} = 0.62\% < \text{FAR}_{\text{\_(42 features)}} = 0.63\%$).

Polat et al. (2020) [55] evaluated the classifiers SVM, KNN, ANN and NB on their dataset initially composed of twelve features. This experiment used SFFS as a wrapper approach. They evaluated the performance of this approach by calculating many metrics: accuracy, sensitivity, specificity, precision and F-measure. By using a wrapper step and only selected DOS-DDOS features instead of all features, these different models increased the accuracy ($\text{Acc}_{\text{\_ANN\_(6 features)}} = 91.44\% > \text{Acc}_{\text{\_ANN\_(42 features)}} = 91.07\%$) < ($\text{Acc}_{\text{\_SVM\_(10 features)}} = 92.15\% > \text{Acc}_{\text{\_SVM\_(42 features)}} = 92.11\%$) < ($\text{Acc}_{\text{\_NB\_(8 features)}} = 94.87\% > \text{Acc}_{\text{\_NB\_(42 features)}} = 94.48\%$) < ($\text{Acc}_{\text{\_KNN\_(8 features)}}= 98.30\% > \text{Acc}_{\text{\_KNN\_(42 features)}}= 95.67$. However, the precision of SVM and KNN is slightly decreased by integrating the feature selection process compared to the initial set with all features ($\text{Pr}_{\text{\_SVM\_(10 features)}} = 90.23\% < \text{Pr}_{\text{\_SVM\_(42 features)}} = 91.42\%$), ($\text{Pr}_{\text{\_ANN\_(6 features)}} = 88.11\% < \text{Pr}_{\text{\_ANN\_(42 features)}} = 89.89\%$). The specificity is enhanced for all the used models, particularly by using a KNN model ($\text{Sp}_{\text{\_SVM\_(10 features)}} = 97.26\%$, $\text{Sp}_{\text{\_ANN\_(6 features)}} = 97.31\%$, $\text{Sp}_{\text{\_NB\_(8 features)}} = 98.43\%$, $\text{Sp}_{\text{\_KNN\_(6 features)}} = 99.45\%$).

Alabdulwahab and Moon (2020) [31] used the NSL_KDD dataset to evaluate different algorithms based on CfsSubsetEval and Best First as wrapper strategy. They tested CfsSubsetEval with six supervised classifiers: Logiboost, RBF, BayesNet, SMO and RepTree. By using six most relevant DOS-DDOS features, this experiment has shown an important improvement of the execution time ($\text{T}_{\text{\_NBTree\_(6 features)}} = 14.23\text{s} < \text{T}_{\text{\_NBTree\_(42 features)}} = 213.18\text{s}$, $\text{T}_{\text{\_Logiboost\_(6 features)}} = 9.96\text{s} < \text{T}_{\text{\_Logiboost\_(42 features)}} = 18.3\text{s}$. However, the accuracy was better without using the wrapper process ($\text{Acc}_{\text{\_NBTre\_(6 features)}} = 99.46\ \% < \text{Acc}_{\text{\_NBTree\_(42 features)}} = 99.87\%$). However, the RepTree algorithm decreased the accuracy and increased the execution time ($\text{Acc}_{\text{\_RepTree\_(6 features)}} = 99.44\% < \text{Acc}_{\text{\_RepTree\_(42 features)}} = 99.83\%$, $\text{T}_{\text{\_RepTree\_(6 features)}} = 5.76\text{s} > \text{T}_{\text{\_RepTree\_(42 features)}} = 3.59$ s).

Umar et al. (2020) [56] applied Best First Forward as search strategy and DT to evaluate the performance of their detecting attacks model. This strategy selected nineteen best features by using UNSW_NB15 dataset. The assessment of this experiment was based on five metrics: Acc, DR, FAR and T. This method has shown that the execution time has overall decreased for different used classifiers ($\text{T}_{\text{\_ANN\_(19 features)}} = 240\text{s} < \text{T}_{\text{\_ANN\_(42 features)}}=660\text{s}$, RF ($\text{T}_{\text{\_RF\_(19 features)}} =37.8\text{s} < \text{T}_{\text{\_RF\_(19 features)}}= 44.4\text{s}$), NB ($\text{T}_{\text{\_NB\_(19 features)}} = 2.86\ \text{s} < \text{T}_{\text{\_NB\_(42 features)}} = 4.64$ s).

By using nineteen DOS-DDOS features, the five metrics values of ANN, RF and SVM models are slightly the same as the baseline model.

The NB model achieved the worst detection rate ($\text{DR}_{\text{\_NB\_(19 features)}} = 19.38\%$) and the same FAR value as the baseline model ($\text{FAR}_{\text{\_NB\_(19 features)}} = \text{FAR}_{\text{\_NB\_(42 features)}} = 0.01\%$).

The same performance was observed by the RF model ($\text{FAR}_{\text{\_RF (19 features)}} = 27.73\% = \text{FAR}_{\text{\_RF\_(42 features)}}=27.74\%$).

However, the classifiers KNN, SVM, ANN and RF increased the FAR value ($\text{FAR}_{\text{\_ANN\_(19 features)}} = 37.36\% > \text{FAR}_{\text{\_ANN\_(42 features)}} = 29.45\%$,

$\text{FAR}_{\text{\_SVM\_(19 features)}} = 45.64\% > \text{FAR}_{\text{\_SVM\_(42 features)}} = 40.51\%$).

Umar and Chen (2020) [57] used Best First as search strategy and DT as evaluator of their wrapper process. Based on UNSW_NB15, NSL_KDD datasets and four classifiers (ANN, SVM, KNN and RF), this process has selected twenty best DOS-DDOS features. The authors used five metrics to evaluate their models: Acc, DR, FAR and T. As results of this experiment, the RF algorithm outperformed the other used classifiers. By using the NSL_KDD dataset, the used wrapper process enhanced the accuracy and reduced the execution time ($\text{Acc}_{\text{\_RF\_(20 features)}} = 99.7\ \% > \text{Acc}_{\text{\_RF\_(42 features)}} = 98.8\ \%$, $\text{T}_{\text{\_RF\_(20 features)}} = 13.2\text{s} < \text{T}_{\text{\_RF\_(42 features)}} = 15\text{s}$). The use of UNSW_NB15 dataset and the wrapper step enhanced the RF accuracy and slightly increased the execution time due to the unnormalized data ($\text{Acc}_{\text{\_RF\_(20 features)}} = 98.51\% > \text{Acc}_{\text{\_RF\_(42 features)}} = 95.74\%$, $\text{T}_{\text{\_RF\_(20 features)}} = 33.6\text{s} > \text{T}_{\text{\_RF\_(42 features)}} = 32.4\text{s}$).

The performances of KNN, SVM and ANN were slightly lower by using twenty features, UNSW_NB15 and NSL_KDD datasets.

However, the SVM model increased drastically the execution time ($T_{\text{\_SVM\_(20 features)}}$ = 10236.6s > $T_{\text{\_SVM\_(42 features)}}$ = 5213.4s) by using the UNSW_NB15 and NSL_KDD datasets. The KNN and RF classifiers decreased the FAR value on the UNSW_NB15 dataset: ($FAR_{\text{\_KNN (20 features)}}$ = 7.36% < $FAR_{\text{\_RF\_(42 features)}}$= 11.42 %, $FAR_{\text{\_RF\_(20 features)}}$ = 2.89 % < $FAR_{\text{\_RF\_(42 features)}}$ = 8.77%).

However, the SVM model increased drastically the execution time ($T_{\text{\_SVM\_(20 features)}}$ = 10236.6s > $T_{\text{\_SVM\_(42 features)}}$ = 5213.4s) by using the UNSW_NB15 and NSL_KDD datasets. The KNN and RF classifiers decreased the FAR value on the UNSW_NB15 dataset: ($FAR_{\text{\_KNN\_(20 features)}}$ = 7.36% < $FAR_{\text{\_RF\_(42 features)}}$= 11.42 %, $FAR_{\text{\_RF\_(20 features)}}$ = 2.89 % < $FAR_{\text{\_RF\_(42 features)}}$ = 8.77%).

5) DOS-DDOS based on wrapper process and meta-heuristics search: Meta-heuristics are new optimization methods used in DOS-DDOS feature selection problems to provide near-optimal solution [34]. These methods are based on two main search strategies [58]. The first strategy is used to guarantee a global and efficient search to find a solution of DOS-DDOS feature selection. The second strategy is used to improve feature selection solutions.

Important research projects have applied meta-heuristic strategies to solve the problem of DOS-DDOS feature selection. In the paragraph below we discuss the important results of these investigations. At the end of this subsection, we present our second dashboard (Tables IIIA, IIIB, IIIC) to summarize and to compare the performances of these strategies.

TABLE III. (A): WRAPPER METHODS BASED ON META-HEURISTIC SEARCH (MHS)

DOS-DDOS feature selection projects based on wrapper methods	DOS- DDOS used dataset	Used wrapper strategies	Number of DOS -DDOS features	Used classifier	Used Metrics	Values metrics with FS	Values metrics without FS
Jun,et al. (2010)[59]	KDD’99	ABC	5	SVM	Accuracy Time (s)	99.92% 12.20	NA
Alomari and A. Othman (2012) [60]	KDD’99	BA	6	SVM	Accuracy Detection Rate FAR	93.36% 90.22% 4.56%	NA
De la Hoz et al. (2014) [61]	NSL_KDD	NGHA-II	25	GHSOM	Accuracy	99.5%	96.02%
Senthilnayaki, et al. (2015) [62]	KDD’99	GA	10	SVM	Accuracy	99.15%	82.45%
Gaikwad and Thool (2015) [63]	NSL_KDD	GA	15	Bagging (PART)	Accuracy Time (s)	99.71% 1589	NA
				PART	Accuracy Time (s)	77.79% 274	NA
				Bagging (C4.5)	Accuracy Time (s)	77.86% 1795	NA

TABLE III-(B): WRAPPER METHODS BASED ON META-HEURISTIC SEARCH (MHS)

DOS-DDOS feature selection projects based on wrapper methods	DOS- DDOS used dataset	Used wrapper strategies	Number of DOS -DDOS features	Used classifier	Used Metrics	Values metrics with FS	Values metrics without FS
Gaikwad and Thool (2015) [63]	NSL_KDD	GA	15	C4.5	Accuracy Time (s)	79.08% 176.05	NA
Wang Xingzhu (2015) [42]	KDD’99	ACO	10	SVM	Detection Rate Time(s)	97.09% 17.99	92.71% 23.51
Eesa et al. (2015) [64]	KDD’99	CFA	10	ID3	Accuracy Detection Rate FAR	92,83% 92.05% 3.9%	73,26% 71.08% 17.685%
Kang and Kim (2016) [65]	NSL_KDD	LSA- K-means	25	MLP	Accuracy Detection Rate FAR	99.37% 99.42% 0.66%	96.93% 93.38% 0.96%
Hosseinzadeh and Kabiri (2016) [66]	KDD’99	ACO	4	NN	Precision Recall F-measure	81.66% 99.78% 89.82%	87.86% 80.02% 83.76%

Khammassi and Krichen (2017) [26]	KDD'99	GA-LR	18	RF	Precision Recall	99.97% 99.98%	NA
	UNSW_NB15		20	C4.5	Precision Recall	36.09% 4.11%	NA
Enache et al. (2017) [67]	NSL_KDD	PSO	21	SVM	Detection Rate FAR	97.17% 1.6%	89.64% 6.88%
			20	NB	Detection Rate FAR	89.85% 5.34%	90.53% 6.66%
			20	C4.5	Detection Rate FAR	96.66% 2.62%	95.67% 3.02%
Yin Chunyong et al. (2017) [68]	KDD'99	ICSA	21	KNN	Accuracy FAR	99.5% 0.1%	-
Khorram and Baykan (2018) [69]	NSL_KDD	PSO	11	KNN	Accuracy Detection Rate Time (s)	96.04% 94.9% 52	93.9% 91.9% 291
				SVM	Accuracy Detection Rate Time (s)	96.02% 92.3% 309	91.4% 89.9% 722
		ACO	7	KNN	Accuracy Detection Rate Time (s)	98.13% 97.2% 67	93.9% 91.9% 291
				SVM	Accuracy Detection Rate Time (s)	95.6% 93% 142	91.4% 89.9% 722
		ABC	7	KNN	Accuracy Detection Rate Time (s)	98.9% 98.7% 53	93.9% 91.9% 291
				SVM	Accuracy Detection Rate Time (s)	97.1% 93.9% 341	91.4% 89.9% 722
	UNSW_NB15		15		Accuracy Specificity Sensitivity Time(s)	99.12% 91.76% 93.46% 1.32	85.56% NA NA NA

TABLE III-(C): WRAPPER METHODS BASED ON META-HEURISTIC SEARCH (MHS)

DOS-DDOS feature selection projects based on wrapper methods	**DOS- DDOS used dataset**	**Used wrapper strategies**	**Number of DOS -DDOS features**	**Used classifier**	**Used Metrics**	**Values metrics with FS**	**Values metrics without FS**
Mazini et al. (2019)[70]	NSL_KDD	ABC	25	AdaBoost	Accuracy Detection Rate FAR	98.90% 99.61% 0.01%	NA NA NA
Samadi Bonab et al. [58]	KDD	FFA-ALO	12	DT	Accuracy Specificity Sensitivity Time(s)	99.73% 99.67% 99.87% 2.90	97.99% NA NA NA
	NSL_KDD		16		Accuracy Specificity Sensitivity Time(s)	99.31% 97.10% 99.24% 1.50	99.31% NA NA NA
	UNSW_NB15		15		Accuracy Specificity Sensitivity Time(s)	99.12% 91.76% 93.46% 1.32	85.56% NA NA NA

As an example of relevant research projects based on wrapper process and meta-heuristic search, we can cite the important investigation of Jun Wang et al. [59]. In this study, the ABC-SVM approach was adopted as wrapper feature selection process. This wrapper strategy selected five DOS-DDOS best features from the KDD'99 dataset and found the best parameter to the SVM classifier. This method achieved good accuracy (Acc_$_{\text{SVM\_(5 features)}}$ = 99.92%) and improved the time of execution (T_$_{\text{SVM\_(5 features)}}$ = 12.20 s).

Alomari and Ali Othman (2012) [60] used an approach based on the Bees Algorithm (BA) as a wrapper feature method by using the classifier SVM. This experiment selected six DOS-DDOS features collected from the KDD'99 data set. They compared BA-SVM with other methods and concluded that their method achieved high detection rate and accuracy (DR_$_{\text{SVM\_(6 features)}}$ = 90.22%, Acc_$_{\text{SVM\_(6 features)}}$ = 93.36%) on detecting attacks with a low FAR (FAR_$_{\text{SVM\_(6 features)}}$ = 4.56%).

De La Hoz et al. (2014) [61] used a multi-objective procedure based on NSGA-II algorithm as wrapper feature selection to reduce the complexity of Growing Hierarchical Self-Organising Maps (GHSOM) algorithm. This wrapper method selected twenty-five representative features. As one of the multiple-objective based on the NSGA-II, the Jaccard index is evaluated after training the GHSOM. Their proposition improved the accuracy compared to the baseline model (Acc_($_{\text{25 features)}}$ = 99.5% > Acc_($_{\text{42 features)}}$ = 96.02%).

Senthilnayaki et al. (2015) [62] combined Genetic Algorithm (GA) with SVM. This study achieved high accuracy (Acc$_{\text{\_(10 features)}}$ = 99.15%) with only ten best DOS-DDOS features compared to the baseline model (Acc_($_{\text{42 features)}}$ = 82.45%).

Gaikwad and Thool (2015) [63] used Genetic Algorithm as wrapper feature selection which selected fifteen features. The authors used two classifiers Partial Decision Tree (PART) and C4.5, and they employed the Bagging on the two previous classifiers. This experiment has shown that using PART with the bagged classifier enhanced the accuracy and increased the execution time (Acc_$_{\text{Bagging\_PART}}$ = 99.71% > Acc_$_{\text{PART}}$ = 77.79%, T_$_{\text{Bagging\_PART}}$ = 1589s > T_$_{\text{PART}}$ = 274s). On the other side, using C4.5 with Bagging decreased the accuracy and increased drastically the execution time (Acc_$_{\text{Bagging\_C4.5}}$ =77.86% < Acc_$_{\text{C4.5}}$ = 79.08%, T_$_{\text{Bagging\_C4.5}}$ = 1795s > T_$_{\text{C4.5}}$ = 176.05s).

Wang Xingzhu (2015) [42] combined ACO feature weighting SVM. This wrapper strategy selected ten most important DOS.

DDOS features which achieved high detection rate and reduced the execution time (DR$_{\text{\_(10 features)}}$ = 97.09% > DR$_{\text{\_(42 features) features)}}$ = 92.71 %, T$_{\text{\_(42 features)}}$ = 23.51s > T$_{\text{\_(10 features)}}$ = 17.99s).

Eesa et al. (2015) [64] modified the Cuttle Fish Algorithm (CFA) and used it as wrapper feature selection method. They applied the classifier ID3 to detect attacks by using the KDD'99 dataset with ten best features. The process showed a real improvement of accuracy and detection rate compared to all used features (Acc$_{\text{\_(10 features)}}$ = 92.83% > Acc_$_{\text{(42 features)}}$ = 73.26%, DR$_{\text{\_(10 features)}}$ = 92.05% > DR$_{\text{\_(42 features)}}$ = 71.08%). Moreover, the FAR value decreased from FAR$_{\text{\_(42 features)}}$ = 17.68% to FAR$_{\text{\_(10 features)}}$ = 3.9 %.

Kang and Kim (2016) [65] employed Local Search Algorithm (LSA) and K-means to find the optimal DOS-DDOS subset features, to reduce the training time and to avoid the over-fitting problem. This experiment evaluated the performance of twenty five selected DOS-DDOS features. The result has shown that using LSA-K-means as wrapper feature step with MLP enhanced the accuracy, increased the detection rate and reduced the FAR value (Acc$_{\text{\_(25 features)}}$ = 99.37% > Acc_($_{\text{42 features)}}$ = 96.93%, DR$_{\text{\_(25 features)}}$ = 99.42% > DR$_{\text{\_(42 features)}}$ = 93.38 %, FAR$_{\text{\_(25 features)}}$ = 0.66% < FAR$_{\text{\_(42 features)}}$ = 0.96%).

Hosseinzadeh Aghdam and Kabiri (2016) [66] build an intrusion detection system based on ACO (Ant Colony Optimization) feature selection method. This method converges faster to the optimal DOS-DDOS subset composed of four DOS-DDOS features. This strategy has increased the Recall and the F-measure values (Re$_{\text{\_(4 features)}}$ = 99.78% > Re$_{\text{\_(42 features)}}$ = 80.02 %, FM$_{\text{\_(4 features)}}$ = 89.82% > FM$_{\text{\_(42 features)}}$ = 83.76%). However, the precision is slightly decreased compared to the baseline model (Pr$_{\text{\_(4 features)}}$ = 81.66 % < Pr$_{\text{\_(42 features)}}$ = 87.86 %).

Khammassi and Krichen (2017) [26] combined Genetic Algorithm with Logistic Regression (LR) as Wrapper feature selection method. This experiment based on different decision tree classifiers (C4.5, RF, and NBTree) has maximized the accuracy by using the KDD'99 and UNSW_NB15 datasets with eighteen and twenty DOS-DDOS best features. The LR-RF strategy has achieved a high precision and Recall values (Pr$_{\text{\_(18 features)}}$ = 99.97%, Re$_{\text{\_(18 features)}}$ = 99.98%).

By using UNSW_NB15 dataset with twenty DOS-DDOS features, the LR-C4.5 process has achieved the worst Recall and precision values (Re$_{\text{\_(20 features)}}$ = 4.11 %, Pr_($_{\text{20 features)}}$ = 36.09%).

Enache et al. (2017) [67] conducted their experiment on the NSL_KDD dataset with many wrapper approaches (Algorithm (BA) ad Particle Swarm Optimization (PSO)). To evaluate these strategies they used the classifiers C4.5, SVM and BN.

The PSO-SVM process outperformed the other classifiers with only twenty-one features. It enhanced the detection rate and decreased the FAR value (DR$_{\text{\_(21features)}}$ = 97.17 > DR$_{\text{\_(42 features)}}$ = 89.64 %, FAR$_{\text{\_(21 features)}}$ = 1.6 % < FAR$_{\text{\_(42 features)}}$ = 6.88 %).

By using eighteens selected features, the process BA-C4.5 achieved an interesting detection rate and increased slightly the FAR value (DR$_{\text{\_(18 features)}}$ = 96.01 % > DR$_{\text{\_(42 features)}}$ = 95.67 %, FAR$_{\text{\_(18 features)}}$ = 3.20 % > FAR$_{\text{\_(42 features)}}$ = 3.02 %).

Yin Chunyong et al. (2017) [68] used an artificial immune system as wrapper method which improved the Clonal Selection Algorithm (ICSA). This method based on the theory of biological immune system learning process selected twenty-one features from the KDD'99 dataset.

This subset realized a good accuracy and low FAR value (Acc$_{\text{\_(21 features)}}$ = 99.5%, FAR$_{\text{\_(21 features)}}$ = 0.1%).

Khorram and Baykan (2018) [69] tested and compared the performances of three wrapper feature selection methods by using two classifiers: SVM and KNN. The used wrapper methods are Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO) and Artificial Bee Colony (ABC).

This experiment showed that ABC-KNN strategy with seven features outperformed the use of all features (T_$_{\text{ABC-KNN\_(7 features)}}$ = 53s < T_$_{\text{KNN\_(42 features)}}$ = 291s, Acc_$_{\text{ABC-KNN\_(7 features)}}$ = 98.9 % > Acc_$_{\text{KNN\_(42 features)}}$ = 93.9%, DR_$_{\text{ABC-KNN\_(7 features)}}$ =98.7 % > DR_$_{\text{KNN\_(42 features)}}$ = 91.9 %).

Mazini et al. (2019) [70] employed ABC as wrapper process to optimize their IDS by using NSL_KDD dataset, the classifier AdaBoost and the parameters regulation method.

This strategy selected twenty-five DOS-DDOS features and achieved a high accuracy, detection rate and low FAR values (Acc$_{\text{\_(25 features)}}$ = 98.90%, DR$_{\text{\_(25 features)}}$ = 99.61%, FAR$_{\text{\_(25 features)}}$ = 0.01%).

Samadi Bonab et al. [58] proposed an improved version of IDS based on the hybrid method Fruit-Flu algorithm (FFA) and the Lion Optimizer algorithm (ALO) as wrapper approach. This strategy based on the datasets KDD'99, NSL_KDD and UNSW_NB15 reduced the used features from 41 to 12 on KDD'99, from 41 to 16 on NSL_KDD and from 48 to 15 on UNSW_NB15. It applied the DT algorithm as a classifier on these different datasets. The performances are evaluated by using five metrics: Acc, Sp, Sen and T. This experiment has shown an enhanced accuracy and reduced the execution time on KDD'99 and UNSW_NB15 datasets (Acc_$_{\text{KDD'99\_(12 features)}}$ = 99.73% > Acc_$_{\text{ KDD'99\_(42 features)}}$ = 97.99%, Acc$_{\text{\_UNSW\_NB15\_(15 features)}}$ = 99.12% > Acc$_{\text{\_UNSW\_NB15\_(42 features)}}$ = 85.56%). On the NSL_KDD dataset the use of this wrapper process didn't change the accuracy (Acc$_{\text{\_NSL\_KDD\_(16 features)}}$ = Acc$_{\text{\_NSL\_KDD\_(42 features)}}$ =93%). However, the specificity was lower on UNSW_NB15 and NSL_KDD compared to KDD'99 (Sp_$_{\text{UNSW\_NB}}$ = 91.76 % < Sp_$_{\text{NSL\_KDD}}$ = 97.10% < Sp_$_{\text{KDD}}$= 99.67%).

The Tables IIIA, IIIB, IIIC summarize and compare the performances of all wrapper process and meta-heuristic strategies discuss above.

6) DOS-DDOS feature selection based on wrapper process and Random search methods: Random search methods applied DOS-DDOS feature selection projects to evaluate the DOS-DDOS features on random sampling around the problem region. These stochastic methods are mainly used to solve the global problem optimizations [71].

To optimize the DOS-DDOS feature subsets, many important research projects have used wrapper process and random search methods to solve this problem. We discuss these projects in the paragraph below. At the end of this subsection, we present our third dashboard (Table IV) to summarize and to compare the performances of these strategies.

As an example of these important investigations, we can cite the important study of Lin et al. (2012) [72] which combined Simulated Annealing (SA) with SVM algorithm to get the best feature subset. This experiment selected twenty three best DOS-DDOS features which evaluated by SA as random search and C4.5 decision tree as classifier. Compared to the initial set of features, the selected subset achieved a high accuracy equal to 99.96%.

Chowdhury et al. (2016) [36] used a wrapper feature selection method based on SA as random search and the ACCS dataset. This strategy selected three best features to detect attacks.

By applying the SVM algorithm with SA, this experiment has showed better accuracy, low FAR and FN values compared to all used features (Acc_$_{\text{SVM\_(3 features)}}$ = 98.76% > Acc_$_{\text{SVM\_(42 features)}}$ = 88.03%, FAR_$_{\text{SVM\_(3 features)}}$ = 0.09% < FAR_$_{\text{SVM\_(42 features)}}$ = 4.2%, FN_$_{\text{SVM\_(3 features)}}$ = 1.15% < FN_$_{\text{SVM\_(42 features)}}$ = 7.77 %).

Hasan Md El Mehedi et al. (2016) [73] adapted the Random Forest algorithm (RF) to select twenty-five best features by using the KDD'99 dataset. The performances evaluation is based on 3 metrics: accuracy, precision and FAR. Compared to the initial used dataset with all features, this wrapper strategy increased the accuracy, the precision and decreased the FAR value (Acc$_{\text{\_(25 features)}}$ = 91.90% > Acc$_{\text{\_(42 features)}}$ = 91.41%, Pr$_{\text{\_(25 features)}}$ = 98.94% > Pr$_{\text{\_(42 features)}}$ = 98.91%, FAR$_{\text{\_(25 features)}}$ = 5.82% < FAR$_{\text{\_(42 features)}}$ = 7.52%).

TABLE IV. WRAPPER METHOD BASED ON RANDOM METHODS

DOS-DDOS feature selection projects based on wrapper methods	DOS- DDOS used dataset	Used wrapper strategies	Number of DOS -DDOS features	Used classifier	Used Metrics	Values metrics with FS	Values metrics without FS
Lin et al. [72]	KDD'99	SA-SVM	23	SA-DT	Accuracy	99.96%	NA
Chowdhury et al. [36]	ACCS	SA	3	SVM	Accuracy FAR FN	98.76% 0.09% 1.15%	88.03% 4.2% 7.77%
Hasan Md El Mehedi et al. [73]	KDD'99	RF	25	RF	Accuracy Precision FAR	91.90% 98.94% 5.82%	91.41% 98.91% 7.52%
Najeeb and Dhannoon (2018) [74]	NSL_KDD	BFA	15	NB	Accuracy	94.83%	89.9%
Almasoudy et al. (2019) [75]	NSL_KDD	DE	9	ELM	Detection Rate Precision F_measure	91.5% 81.18% 86.03%	79.55% 94.90% 80.44%

Najeeb and Dhannoon (2018) [74] proposed an IDS model that combined the Binary Firefly (BFA) method with the Naïve Bayes (NB) classifier by using the NSL_KDD dataset. The BFA is initialized by a binary sequence contrary to the Firefly (FA) algorithm. This model was iterated two hundred times with fifteen selected features and achieved better accuracy compared to all used features ($Acc_{\_(25\ features)}$ = 94.83% > $Acc_{\_(42\ features)}$ = 89.9%).

Almasoudy et al. (2019) [75] has realized an IDS experiment based on Differential Evolution (DE) as wrapper based approach by using the NSL_KDD dataset. Nine candidate features are randomly selected. The Extreme Learning Machine (ELM) is used as classifier to compute the accuracy of DOS-DDOS features until it achieved high accuracy. Applied to DOS-DDOS attacks predicting, this method achieved high detection rate, high F-measure and decreased slightly the precision ($DR_{\_(9\ features)}$ = 91.5% > $DR_{\_(42\ features)}$ = 79.55%, $FM_{\_(9\ features)}$ = 86.03 % > $FM_{\_(42\ features)}$ = 80.84%, $Pr_{\_(42\ features)}$ = 94.90 % > $Pr_{\_(9\ features)}$ = 81.18%).

VI. Conclusion

Nowadays, cybersecurity attacks grow over time, especially the Denial of Service attack (DOS) and its variant Distributed Denial of Service (DDOS). These famous attacks continue to threaten private and public activities everywhere.

Dealing with these threats by using Machine Learning (ML) models can hold a great promise in DOS-DDOS security systems. By learning from and identifying a large amount of network traffic, these predictive models can efficiently handle the DOS-DDOS threats and overcome several limits and performance issues of the traditional security solutions.

One of the key preprocessing phases to success and optimize these DOS-DDOS cybersecurity intelligence models is feature selection step, particularly the feature selection method based on the Wrapper strategies.

Using Wrapper techniques improved significantly the selection of the relevant DOS-DDOS features and enhanced the performance of many existing ML solutions.

In this paper, we have advanced the development of this previous work by studying different DOS-DDOS datasets, algorithms and the results of several research projects. We have reviewed and evaluated the impact of many important wrapper strategies used by many existing DOS-DDOS security systems.

We have summarized the findings in three dashboards that are essential to understand the performance of three wrapper strategies commonly used in DOS-DDOS ML models: heuristic search algorithms, meta-heuristic search and random search methods.

This study shows that many wrapper strategies, algorithms, DOS-DDOS features with a relevant impact can be selected to improve the DOS-DDOS ML existing solutions.

Acknowledgment

I would like to express my sincere gratitude to my Professors and my family for the continuous support of my study and related research, for their patience, motivation, and immense knowledge.

References

[1] F. Ullah, M. Ali Babar, “Architectural Tactics for Big Data Cybersecurity Analytics Systems: A Review,” Journal of Systems and Software, 151, 81–118, 2019, doi:10.1016/j.jss.2019.01.051.

[2] R. Vishwakarma, K. Ankit Jain, “A survey of DDoS attacking techniques and defence mechanisms in the IoT network,” Telecommunication Systems: Modelling, Analysis, Design and Management, 73(1), 3–25, 2020.

[3] K.M. Prasad, D.A.R.M. Reddy, D.K.V. Rao, “DoS and DDoS Attacks: Defense, Detection and Traceback Mechanisms - A Survey,” Global Journal of Computer Science and Technology, 2014.

[4] J.-H. Cho, J.-Y. Shin, H. Lee, J.-M. Kim, G. Lee, “DDoS Prevention System Using Multi-Filtering Method,” Atlantis Press: 774–778, 2015, doi:10.2991/cmfe-15.2015.182.

[5] S. Qadir Mir, S. Quadri, “Information Availability: An Insight into the Most Important Attribute of Information Security,” Journal of Information Security, 07, 185–194, 2016, doi:10.4236/jis.2016.73014.

[6] M. Sachdeva, G. Singh, K. Saluja, K. Singh, “DDoS Incidents and their Impact: A Review,” Int. Arab J. Inf. Technol., 7, 14–20, 2010.

[7] X. Liang, T. Znati, “On the performance of intelligent techniques for intensive and stealthy DDos detection,” Computer Networks, 164, 106906, 2019, doi:10.1016/j.comnet.2019.106906.

[8] Ibrahim Salim M., T.A. Razak, “A study on IDS for preventing Denial of Service attack using outliers techniques,” in 2016 IEEE International Conference on Engineering and Technology (ICETECH), 768–775, 2016, doi:10.1109/ICETECH.2016.7569352.

[9] Y.V. Srinivasa Murthy, K. Harish, V. Varma, K. Sriram, B. Revanth, “Hybrid Intelligent Intrusion Detection System using Bayesian and Genetic Algorithm (BAGA): Comparitive Study,” International Journal of Computer Applications, 99, 1–8, 2014, doi:10.5120/17342-7808.

[10] O. Salem, M. HOTTE, Q.-E. LUTTIN, T. ASCOET, Protection contre les attaques de déni de service dans les réseaux IP, Paris Descarte IUT, ECTEI: 31, 2015.

[11] J. Jang-Jaccard, S. Nepal, “A survey of emerging threats in cybersecurity,” Journal of Computer and System Sciences, 80(5), 973–993, 2014, doi:10.1016/j.jcss.2014.02.005.

[12] K.R. Bandara, T. Abeysinghe, A. Hijaz, D. Darshana, H. Aneez, S.J. Kaluarachchi, K.D. Sulochana, M. DhishanDhammearatchi, “Preventing DDoS attack using Data mining Algorithms,” International Journal of Scientific and Research Publications, 6(10), 390–400, 2016.

[13] L. Gnanaprasanambikai, N. Munusamy, “Data Pre-Processing and Classification for Traffic Anomaly Intrusion Detection Using NSLKDD Dataset,” Cybernetics and Information Technologies, 18, 2018, doi:10.2478/cait-2018-0042.

[14] S.X. Wu, W. Banzhaf, “The use of computational intelligence in intrusion detection systems: A review,” Applied Soft Computing, 10(1), 1–35, 2010, doi:10.1016/j.asoc.2009.06.019.

[15] A. Alazab, M. Hobbs, J. Abawajy, M. Alazab, “Using feature selection for intrusion detection system,” in 2012 International Symposium on Communications and Information Technologies (ISCIT), IEEE, Gold Coast, Australia: 296–301, 2012, doi:10.1109/ISCIT.2012.6380910.

[16] V.O. Ferreira, V.V. Galhardi, L.B.L. Gonçalves, R.C. Silva, A.M. Cansian, “A model for anomaly classification in intrusion detection systems,” Journal of Physics: Conference Series, 633, 4, 2015, doi:10.1088/1742-6596/633/1/012124.

[17] M. Bataghva, “Efficiency and Accuracy Enhancement of Intrusion Detection System Using Feature Selection and Cross-layer Mechanism,” Electronic Thesis and Dissertation Repository, 2017.

[18] I.H. Sarker, A.S.M. Kayes, S. Badsha, H. Alqahtani, P. Watters, A. Ng, “Cybersecurity data science: an overview from machine learning perspective,” Journal of Big Data, 7(1), 41, 2020, doi:10.1186/s40537-020-00318-5.

[19] J.B. Fraley, J. Cannady, “The promise of machine learning in cybersecurity,” in SoutheastCon 2017, 1–6, 2017, doi:10.1109/SECON.2017.7925283.

[20] F. Ullah, M. Ali Babar, "Architectural Tactics for Big Data Cybersecurity Analytic Systems: A Review," Journal of Systems and Software, 151, 2018, doi:10.1016/j.jss.2019.01.051.

[21] S. Sambangi, L. Gondi, "A Machine Learning Approach for DDoS (Distributed Denial of Service) Attack Detection Using Multiple Linear Regression," Proceedings, 63(1), 51, 2020, doi:10.3390/proceedings2020063051.

[22] R. Panthong, A. Srivihok, "Wrapper Feature Subset Selection for Dimension Reduction Based on Ensemble Learning Algorithm," Procedia Computer Science, 72, 162–169, 2015, doi:10.1016/j.procs.2015.12.117.

[23] N. Bindra, M. Sood, "Evaluating the Impact of Feature Selection Methods on the Performance of the Machine Learning Models in Detecting DDoS Attacks," Romanian Journal of Information Science and Technology, 3, 250–261, 2020.

[24] M. Joshi, T.H. Hadi, "A Review of Network Traffic Analysis and Prediction Techniques," Network Traffic Analysis and Prediction, 23, 2015.

[25] Z. Foroushani, Y. Li, "Intrusion Detection System by Using Hybrid Algorithm of Data Mining Technique," in ICSCA 2018: Proceedings of the 2018 7th International Conference on Software and Computer Applications, Kuantan, Malaysia: 119–123, 2018, doi:10.1145/3185089.3185114.

[26] C. Khammassi, S. Krichen, "A GA-LR wrapper approach for feature selection in network intrusion detection," Computers & Security, 70, 255–277, 2017, doi:10.1016/j.cose.2017.06.005.

[27] V. Bolón-Canedo, N.S. Maroño, A. Alonso-Betanzos, Feature Selection for High-Dimensional Data, Springer International Publishing, 2015, doi:10.1007/978-3-319-21858-8.

[28] F. Amiri, "Mutual information-based feature selection for intrusion detection systems," Journal of Network and Computer Applications, 34(4), 1184–1199, 2011, doi:10.1016/j.jnca.2011.01.002.

[29] V. Bachu, J. Anuradha, "A Review of Feature Selection and Its Methods," Cybernetics and Information Technologies, 19, 3, 2019, doi:10.2478/cait-2019-0001.

[30] V. Kumar, S. Minz, "Feature selection: A literature review," Smart Computing Review, 4, 211–229, 2014, doi:10.1145/2740070.2626320.

[31] S. Alabdulwahab, B. Moon, "Feature Selection Methods Simultaneously Improve the Detection Accuracy and Model Building Time of Machine Learning Classifiers," Symmetry, 12(9), 1424, 2020, doi:10.3390/sym12091424.

[32] S. Dwivedi, M. Vardhan, S. Tripathi, "Defense against distributed DoS attack detection by using intelligent evolutionary algorithm," International Journal of Computers and Applications, 1–11, 2020, doi:10.1080/1206212X.2020.1720951.

[33] K. Yan, D. Zhang, "Feature selection and analysis on correlated gas sensor data with recursive feature elimination," Sensors and Actuators B: Chemical, 212, 353–363, 2015, doi:10.1016/j.snb.2015.02.025.

[34] N. Mlambo, W. Cheruiyot, M.W. Kimwele, "A Survey and Comparative Study of Filter and Wrapper Feature Selection Techniques," The International Journal Of Engineering And Science, 5(10), 57–67, 2016.

[35] N. Moustafa, J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in 2015 Military Communications and Information Systems Conference, 1–6, 2015, doi:10.1109/MilCIS.2015.7348942.

[36] M.N. Chowdhury, K. Ferens, M. Ferens, "Network Intrusion Detection Using Machine Learning," in Computer Science, CSREA Press: 30–35, 2016.

[37] M.E. Elhamahmy, H.N. Elmahdy, I.A. Saroit, "A New Approach for Evaluating Intrusion Detection System," in CiiT International Journal of Artificial Intelligent Systems and Machine Learning, 290–298, 2010.

[38] Kamarularifin Abd Jalil, Muhammad Hilmi Kamarudin, Mohamad Noorman Masrek, "Comparison of Machine Learning algorithms performance in detecting network intrusion," in 2010 International Conference on Networking and Information Technology, 221–226, 2010, doi:10.1109/ICNIT.2010.5508526.

[39] M. Belouch, S. El Hadaj, M. Idhammad, "Performance evaluation of intrusion detection based on machine learning using Apache Spark," Procedia Computer Science, 127, 1–6, 2018, doi:10.1016/j.procs.2018.01.091.

[40] V.D. Katkar, S.V. Kulkarni, "Experiments on detection of Denial of Service attacks using ensemble of classifiers," in 2013 International Conference on Green Computing, Communication and Conservation of Energy (ICGCE), 837–842, 2013, doi:10.1109/ICGCE.2013.6823550.

[41] M.H. Bhuyan, D.K. Bhattacharyya, J.K. Kalita, "Network Anomaly Detection: Methods, Systems and Tools," IEEE Communications Surveys Tutorials, 16(1), 303–336, 2014, doi:10.1109/SURV.2013.052213.00046.

[42] W. Xingzhu, "ACO and SVM Selection Feature Weighting of Network Intrusion Detection Method," International Journal of Security and Its Applications, 9(4), 259–270, 2015, doi:10.14257/ijsia.2015.9.4.24.

[43] J.J. Lu, M. Zhang, Heuristic Search, Springer, New York, NY: 885–886, 2013, doi:10.1007/978-1-4419-9863-7_875.

[44] B. Kavitha, S. Karthikeyan, B. Chitra, Efficient Intrusion Detection with Reduced Dimension Using Data Mining Classification Methods and Their Performance Comparison, Springer Berlin Heidelberg, Berlin, Heidelberg: 96–101, 2010, doi:10.1007/978-3-642-12214-9_17.

[45] M.S. Mok, S.Y. Sohn, Y.H. Ju, "Random Effects Logistic Regression Model for Anomaly Detection," Expert Syst. Appl., 37(10), 7162–7166, 2010, doi:10.1016/j.eswa.2010.04.017.

[46] I. Ahmad, A. Abdullah, A. Alghamdi, M. Hussain, K. Nafjan, "Intrusion Detection Using Feature Subset Selection based on MLP," Scientific Research and Essays, 6(34), 6804–6810, 2011, doi:10.5897/SRE11.142.

[47] L. Yinhui, J. Xia, S. Zhang, J. Yan, X. Ai, K. Dai, "An efficient intrusion detection system based on support vector machines and gradually feature removal method," Expert Systems with Applications, 39, 424–430, 2012, doi:10.1016/j.eswa.2011.07.032.

[48] F. Zhang, D. Wang, "An Effective Feature Selection Approach for Network Intrusion Detection," in 2013 IEEE Eighth International Conference on Networking, Architecture and Storage, 307–311, 2013, doi:10.1109/NAS.2013.49.

[49] O.Y. Al-Jarrah, A. Siddiqui, M. Elsalamouny, P.D. Yoo, S. Muhaidat, K. Kim, "Machine-Learning-Based Feature Selection Techniques for Large-Scale Network Intrusion Detection," in 2014 IEEE 34th International Conference on Distributed Computing Systems Workshops (ICDCSW), 177–181, 2014, doi:10.1109/ICDCSW.2014.14.

[50] J. Lee, D. Park, C. Lee, "Feature Selection Algorithm for Intrusions Detection System using Sequential Forward Search and Random Forest Classifier," KSII Transactions on Internet and Information Systems, 11(10), 5132–5148, 2017.

[51] B.S. Harish, N. Manju, "Hybrid Feature Selection Method Using Fisher's Discriminate Ratio to Classify Internet Traffic Data," in Proceedings of the 4th International Conference on Frontiers of Educational Technologies, ACM, New York, NY, USA: 75–79, 2018, doi:10.1145/3233347.3233369.

[52] H. Soodeh, A. Mehrdad, "The hybrid technique for DDoS detection with supervised learning algorithms," Computer Networks, 158, 35–45, 2019, doi:10.1016/j.comnet.2019.04.027.

[53] H. Malhotra, P. Sharma, "Intrusion Detection using Machine Learning and Feature Selection," International Journal of Computer Network and Information Security, 11(4), 43–52, 2019, doi:10.5815/ijcnis.2019.04.06.

[54] M. Wang, Y. Lu, J. Qin, "A dynamic MLP-based DDoS attack detection method using feature selection and feedback," Computers & Security, 88, 101645, 2020, doi:10.1016/j.cose.2019.101645.

[55] H. Polat, O. Polat, A. Cetin, "Detecting DDoS Attacks in Software-Defined Networks Through Feature Selection Methods and Machine Learning Models," Sustainability, 12(3), 1–16, 2020.

[56] M.A. Umar, C. Zhanfang, Y. Liu, "Network Intrusion Detection Using Wrapper-based Decision Tree for Feature Selection," in Proceedings of the 2020 International Conference on Internet Computing for Science and Engineering, ACM, Male Maldives: 5–13, 2020, doi:10.1145/3424311.3424330.

[57] M.A. Umar, Z. Chen, Effects of Feature Selection and Normalization on Network Intrusion Detection, 2020, doi:10.36227/techrxiv.12480425.

[58] M. Samadi Bonab, A. Ghaffari, F. Soleimanian Gharehchopogh, P. Alemi, " A wrapper - based feature selection for improving performance of intrusion detection systems," International Journal of Communication Systems, 33, 2020, doi:10.1002/dac.4434.

[59] W. Jun, L. Taihang, R. Rongrong, "A real time IDSs based on artificial Bee Colony-support vector machine algorithm," Suzhou, Jiangsu, China: 91–96, 2010, doi:10.1109/IWACI.2010.5585107.

[60] O. Alomari, Z. Ali Othman, "Bees Algorithm for feature selection in Network Anomaly detection," 8(3), 1748–1756, 2012.

[61] E. de la Hoz, E. de la Hoz, A. Ortiz, J. Ortega, A. Martínez-Álvarez, "Feature selection by multi-objective optimisation: Application to network anomaly detection by hierarchical self-organising maps," Knowledge-Based Systems, 71, 322–338, 2014, doi:10.1016/j.knosys.2014.08.013.

[62] B. Senthilnayaki, K. Venkatalakshmi, A. Kannan, "Intrusion detection using optimal genetic feature selection and SVM based classifier," in 2015 3rd International Conference on Signal Processing, Communication and Networking (ICSCN), 1–4, 2015, doi:10.1109/ICSCN.2015.7219890.

[63] D.P. Gaikwad, R.C. Thool, "Intrusion Detection System Using Bagging with Partial Decision TreeBase Classifier," Procedia Computer Science, 49, 92–98, 2015, doi:10.1016/j.procs.2015.04.231.

[64] A.S. Eesa, Z. Orman, A.M.A. Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems," Expert Systems with Applications, 42(5), 2670–2679, 2015, doi:10.1016/j.eswa.2014.11.009.

[65] S.-H. Kang, K.J. Kim, "A feature selection approach to find optimal feature subsets for the network intrusion detection system," Cluster Computing, 19(1), 325–333, 2016, doi:10.1007/s10586-015-0527-8.

[66] M. Hosseinzadeh Aghdam, P. Kabiri, "Feature Selection for Intrusion Detection System Using Ant Colony Optimization," International Journal of Network Security, 18, 420–432, 2016.

[67] A. Enache, V. Sgârciu, M. Togan, "Comparative Study on Feature Selection Methods Rooted in Swarm Intelligence for Intrusion Detection," in 2017 21st International Conference on Control Systems and Computer Science (CSCS), 239–244, 2017, doi:10.1109/CSCS.2017.40.

[68] C. Yin, L. Ma, L. Feng, "Towards accurate intrusion detection based on improved clonal selection algorithm," Multimedia Tools and Applications, 76(19), 19397–19410, 2017, doi:10.1007/s11042-015-3117-0.

[69] T. Khorram, N. Baykan, "Feature selection in network intrusion detection using metaheuristic algorithms," International Journal Of Advance Research, Ideas and Innovations in Technology, 4(4), 704–710, 2018.

[70] M. Mazini, B. Shirazi, I. Mahdavi, "Anomaly network-based intrusion detection system using a reliable hybrid artificial bee colony and AdaBoost algorithms," Journal of King Saud University - Computer and Information Sciences, 31(4), 541–553, 2019, doi:10.1016/j.jksuci.2018.03.011.

[71] H.E. Romeijn, Random search methods, Springer US, Boston, MA: 3245–3251, 2009, doi:10.1007/978-0-387-74759-0_556.

[72] S.-W. Lin, K. Ying, C. Lee, Z.-J. Lee, "An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection," Appl. Soft Comput., 12(10), 3285–3290, 2012, doi:10.1016/j.asoc.2012.05.004.

[73] M.A.M. Hasan, M. Nasser, S. Ahmad, K.I. Molla, "Feature Selection for Intrusion Detection Using Random Forest," Journal of Information Security, 7(3), 129–140, 2016, doi:10.4236/jis.2016.73009.

[74] R.F. Najeeb, B.N. Dhannoon, "Improving Detection Rate of the Network Intrusion Detection System Based on Wrapper Feature Selection Approach," Iraqi Journal of Science, 59(1.B), 426–433, 2018, doi:10.24996/ijs.2018.59.1B.23.

[75] F. Almasoudy, W. Al-Yaseen, A. Idrees, "Differential Evolution Wrapper Feature Selection for Intrusion Detection System," Procedia Computer Science, 167, 1230–1239, 2019, doi:10.1016/j.procs.2020.03.438.