PalArch's Journal of Archaeology
of Egypt / Egyptology

DDoS ATTACK DETECTION SCHEME USING HYBRID ENSEMBLE LEARNING AND GA ALGORITHM FOR INTERNET OF THINGS

Amin Erfan

Department of Computer Engineering, Faculty of Electrical and Computer Engineering,

Technical and Vocational University (TVU), Tehran, Iran

*Corresponding author: e-mail: Amin_er6197@yahoo.com

Amin Erfan. DDoS Attack Detection Scheme using Hybrid Ensemble Learning and GA Algorithm for Internet of Things -- Palarch's Journal of Archaeology of Egypt/Egyptology 18(18), 521-546. ISSN 1567-214x

Keywords: DDoS attacks, machine learning, GA algorithms, Internet of Things.

ABSTRACT

Methods: In this paper, we using the hybrid feature selection GA algorithm (GA) and ensemble learning system such as C4.5 decision tree, deep neural network (DNN) and K-KNN (KNN) algorithm for IoT. The GA feature selection used for selecting best attribute on DDoS attacks dataset and from ensemble learning system is for DDoS attach detection.

Results: To validate the proposed method, the results were compared with other approaches, including machine learning methods that combined with other optimization methods. We used 10% of the KDDCup99 dataset for simulation. The results of the paper show the high accuracy of the proposed method for DDoS attacks detection compared to other recent methods of about 5%.

Conclusions: Therefore, by presenting experiments to DDoS attacks detection, it was observed that the proposed method detected the DDoS with acceptable accuracy and the combination of ensemble learning methods and GA feature selection algorithm was successful.

INTRODUCTION

Infiltration is defined as a set of operations that seek to jeopardize the integrity, confidentiality, and availability of a resource. Infiltration detection systems allow you to detect abnormal network access if the intruder intends to gain unusual access to the network after passing through the network security system. It is not possible to prevent intrusion completely, but it is necessary to take measures that automatically and instantly monitor the behavior of users and to prevent it if intrusive behavior is observed. Usually, system vulnerabilities are due to vulnerabilities in security software or problems with the network configuration structure to control access to network information. In general, intrusion detection methods are divided into two main categories: misuse detection and abnormal behavior detection. In the abuse detection method, known intrusion patterns are used to identify intrusions. While in the methods of detecting abnormal behavior, the normal behavior of users is the criterion for action [1]. DDoS attacks are one of the most important attacks to prevent the uninterrupted operation of the Internet service [2]. DDoS attack in simple language means pouring too many requests into a server (victim or target computer) and overusing resources (processor, database, bandwidth, memory, etc.) so that its normal service to its users due to high processing volume or Excessive data overload, server operations are disrupted or out of reach [3]. In DoS-type attacks, a large number of packets are sent through one (DoS) or multiple (DDoS) machines to disable computing power and network resources, or to disable the target machine. DDoS attacks are more powerful and harder to detect and counter than DoS attacks. Because in these attacks, several machines can be coordinated to send a small flow of traffic to the target machine, it is difficult for the target machine to manage the total of these traffic [4]. In a DDoS attack, a large number of infected nodes attack the same target, leading to the failure of the target system service. In fact, such attacks are based on data from infected systems [5.6].

DDoS attacks are a complex and important issue that has been proposed in several ways. There are generally no guaranteed ways to prevent this type of attack, and the only ways to prevent some common methods and reduce the effects of other methods are available, as many methods are by no means preventable. For example, if the Botnet network opens a page with 100,000 zombies on the site, this request is a normal request and it cannot be determined whether the applicant is a regular system or a zombie and therefore cannot be stopped. As the mechanisms for counterattacking expand, so do the scenarios of attack. Since definitely not every large volume of request can be considered a DDoS attack, this research will provide a way to detect countermeasures based on machine learning algorithms and data mining in the form of reinforcement learning systems. For this purpose, we intend to use the reinforcement learning system to teach and determine the optimal values of the parameters. The operational strategy of reinforcement learning systems is based on voting, and the end result, which is identified as DDoS attack, is based on the voting of several machine learning algorithms, including the backup vector machine algorithm, neural network, and decision tree. In this way, DDoS attacks can be more accurately analyzed and identified. Therefore, in this paper, in section 2, the work done in the past is examined, in section 3, the proposed method and relevant details are described, finally in section 4, the results are obtained, and in section 5, the final conclusion and future suggestions are examined. Takes.
In this paper, an ensemble learning method based on C4.5 decision tree algorithm, KNN and DNN for DDoS attack detection was proposed that consists of two main training and testing phases. In the first phase, a certain number of datasets are fed into model as training data. Next, we using the hybrid feature selection GA algorithm and ensemble learning system such as C4.5 decision tree, DNN and K-KNN (KNN) algorithm for IoT. The GA feature selection used for selecting best attribute on DDoS attacks dataset and from ensemble learning system is for DDoS attach detection.
The remainder of this paper is categorized as follows: Section 2 deals with related work, is described in Section 3 of the proposed method and proposed

architecture. In Sections 4 and 5, the results are presented and the final conclusions are presented.

Related Works

The first research to identify DDoS attacks was based on some of the characteristics of specific abnormalities, such as sudden traffic changes or life time (TTL). These methods usually have two problems: First, the deviation from the traffic feature may only be seen too small, especially on monitors that are close to the attack sources. Second, the accuracy of the diagnosis is limited to the variety of DDoS attacks because one feature cannot cover all types of DDoS attacks.

Ray et al. in [8], Provided a scalable RFID security framework and a protocol for supporting Internet security on objects. In this paper, the researchers proposed a new identification method based on a combined approach based on common groups and approaches, and provided a security check (SCH) for the mobile RFID system. The protocol provided by these researchers is able to support compatibility and ensure secure network deployment and provide a strong distribution structure of the IoT. This protocol is also able to support malware using malware detection methods. The simulation results show that the proposed protocol has better security and better scalability than other existing protocols.

Leek et al. in [9], Presented a new framework for e-health architecture in the discussion of IoT security. The IoT promises health care, especially e-health. In fulfilling these promises, major challenges must be addressed, the most important of which are privacy and security challenges. In this paper, the researchers discuss the situation of different standardized organizations and inform them about their vital role in making electronic health records.

Abu Mahara and Quinn in [10], discussed security issues on the IoT and privacy. Since the issue of IoT, security and privacy issues have also arisen. In this paper, researchers examined the IOT perspective, existing security threats, and the challenges posed by the IoT. The researchers submitted their research to the Internet for security and privacy purposes.

Shahroui and Belami in [11], introduced an efficient method based on the Host Identification Protocol (HIP) to ensure end-to-end security on the IoT. In this paper, the researchers presented a compact LOWPAN6 for the HIP header package, as well as a distributed distribution design for computational load security in the base HIP exchange. To achieve security, the researchers presented two models of compression and distribution for HIP in WSN on the IoT. Examination of the results shows that the proposed method, called CD-HIP, is efficient enough and has a slight security delay, which is very small, and at a good level of compliance with the HIP standard.

Skari et al. in [12], Examined the issue of security, privacy, and trust in IoT. IoT is recognized by a heterogeneous technology, which has provided new services in various fields. Meanwhile, satisfaction with security needs and privacy play a key role. Traditional security interactions cannot be applied directly to IoT technology according to different standards. In addition, a large number of devices are connected to each other, which is due to scalability issues. Therefore, a flexible infrastructure is able to meet the security threats required for such a dynamic environment. In this paper, the researchers examined the main challenges and solutions available in the field of IoT security and also pointed out some important points for future research.

Kiwi et al. in [13], Provided an efficient routing protocol for essential responses on the IoT (IoT). In recent years, IoT has continued to expand in many areas, including smart homes, environmental monitoring, and the industrial control system. In this paper, the researchers presented a routing protocol for the necessary IoT responses based on global information decisions (ERGID) to enhance data transfer efficiency, for a reliable transmission and response in emergencies. Specifically, the researchers proposed a mechanism called DIM and a transport strategy called REPC, which focuses on network load balance by focusing on the energy remaining from the node. The results of simulations and analyzes show that ERGID is better than EA-SPEED and SPEED in different periods to end delays, closed losses and energy consumption. In addition, they conducted a number of practical tests with sensor nodes, and the results showed that ERGID could improve REAL-TIME responses in the network.

Raphael and et l. in [14], performed experiments on a backup vector machine algorithm to DDoS attack detection on DARPA and *KDDCUP99* datasets. In their research, they came up with a relatively acceptable result compared to other previous methods. The findings of this study are an acceptable accuracy of the backup vector machine algorithm in the attack detection process.

Jay et al. in [15], Provided a framework for automating the analysis of security on the IoT. In general, the proposed framework consists of five steps: 1- Data processing 2- Security model production 3- Security visualization 4- Security analysis 5- Model update. Using this framework, the researchers found a potential attack scenario on the IoT and, through security measures, well defined Internet security analysis on objects and evaluated the effectiveness of the proposed defense strategy. The framework was assessed through three scenarios, which are smart homes, health care monitoring and environmental conditions. The proposed framework is well able to detect the path of a possible attack and reduce the impact of a possible attack.

Mr. Swang and et al. in [16], used a two-step method to DDoS attack detection. The method proposed by Swang and his colleagues is called TDSC, which is a cluster-based method. In their study, they were able to perform the DDOS attack detection process with the least error compared to other methods.

Mr. Hook and et al. in [17], proposed a method called FPGA to DDoS attack detection. In their study, they used a combination method based on the GA biological algorithm. The use of the GA algorithm optimizes the final responses, and eventually DDOS attacks are detected with acceptable accuracy. The results of this study have improved significantly compared to other methods.

Sani and et al. in [18], used a technique called D-Face to DDoS attack detection. The technique proposed by the authors of this paper is not focused on a specific system and can be provided for all systems and networks. The strategy of the method presented in this paper is based on the traffic generated by the requests sent to the server. The results of this study have been more favorable than many of the previous studies.

Saffar et al. in [19], Provided a Roadmap for security challenges on the IoT. In this paper, a new cognitive and systematic approach to security on the IoT is presented. The role of each of the components of their approach and interaction with the other main components of the proposed plan is fully described. Case studies have been performed on important components and systemic interactions. Security issues and privacy challenges are also discussed later. The results show the effectiveness of the proposed method.

Antenna et al. in [20], Used DBSCAN clustering algorithm and entropy technique to DDoS attack detection. Their research has shown that there has been a significant improvement over many methods.

Kasavauwari et al. in [21], identified and defended DDoS attack using a multifactorial system. This paper provides a new way to detect and defend against DDoS attacks using a standalone multifunction system. DDoS attacks are updated using several interconnected factors and a coordinating factor. The current scenario is examined by the coordinating agent using entropy and covariance methods to be able to investigate DDoS attacks. At this stage, the monitoring factor is implemented live and monitors cloud and network resources. The Testing results show that this proposed system provides optimized performance and improved security on the cloud platform.

Yusuf et al. in [22], examined systematic and classified literature for the diagnosis and prediction of DDoS attacks. In this paper, a comprehensive review of the systematic literature on the effects of DDoS, the types of attacks, the types of diagnostic methods, the different types of attack prediction techniques were examined. The results showed that about 30% of machine learning methods are used to predict and identify DDoS attacks. From year to year, this trend supports the benefits of using tools to produce a rapid result in detecting and enforcing robustness in a large data environment.

Chiu et al. in [23], examined the defense mechanism of DDoS attacks based on cognitive calculations with double-address entropy. The specifications of the switch current table are extracted and a DDoS attack model was created by combining the support vector classification algorithm. This mechanism can detect detection and defense in real time in the early stages of a DDoS attack and can restore timely communication in a timely manner. This experiment shows that the proposed mechanism not only detects attacks quickly, but also has high detection speed and low false positive coefficient. Most importantly, it can take defensive and appropriate action when identifying an attack.

Shafi et al. in [24], prevented botnet DDoS attacks by using the Chinese block chain on the IoT. They focused on developing a botnet prevention system for the IoT that took advantage of a defined software network (SDN) and block chain distribution (DBC). SDN and block chain technology have great potential for solving major IoT problems, including security. By analyzing and simulating using block chain and SDN, they were able to identify and reduce bandwidth and prevent devices from being used by attackers.

Tyne et al. in [25], proposed a method of selecting a feature based on multifunction optimization (FS) to detect denial of service distribution (DDoS) attacks on the IoT. An intrusion detection system (IDS) is an approach to detecting cyberattacks. FS is needed to reduce data size and improve IDS performance. One of the reasons for the failure of IDS is the incorrect selection of features because most FS methods are based on a limited number of goals such as accuracy or data connection, but these are not enough because they can be misleading to detect an attack. Testing results confirm that the proposed method for FS performed well and gained 99.9%, reducing the total number of features by nearly 90%. The proposed method is more advanced than other FS methods for DDoS attacks detection by IDS.

Narrator and et al. in [26], presented a paper entitled Learning to Identify and Reduce DDoS Attacks on IoT via SDN-Cloud Architecture. This paper uses a security program to reduce DDoS attacks on IoT servers using the SDN-based cloud software model. A new mechanism called Learning Reduction (LEDEM) has been identified that identifies DDoS using a semi-functional machine learning algorithm and reduces DDoS. They tested LEDEM in tested and simulated topology and compared the results with advanced solutions. The DDoS attack detected an improved accuracy of 96.28%.

Lee et al. in [27], identified a real-time volumetric detection scheme for DDoS attacks on the IoT. They used a sensitive and reliable functional entropy-based diagnostic method. However, the balance between computational complexity and diagnostic accuracy remains a challenge. The proposed method consisted of three parts: a sliding window for calculating entropy, a one-way filter to achieve early detection during DDoS progress, and a quantum deviation checking algorithm to optimize the detection result. Eventually, these lead to real and efficient performance to DDoS attack detection on the IoT as soon as possible.

Kumar et al. in [28], analyzed various algorithms for detecting and reducing the effects of DDoS attacks in the context of the IoT. As demand for IoT applications increases, so does security. Recently, most of the most common attacks in the world of IoT have been distributed denial of service, and SDx has been used to manage the security of IoT devices. The results of this study show that security is a very important issue for reliable communications.

Irom et al. in [29], identified and prevented DDoS attacks on the IoT. Various techniques have been proposed for the DDoS attack detection system and prevention mechanism to investigate how network detection and prevention are performed against these attacks, and comparative analysis is performed to determine the optimal technique to better trade between security. And network performance is done during and after the attack. The results show that different techniques use filtering and speed limiting techniques to identify and prevent DDoS attacks with higher reliability.

According to the review of the researches conducted in the field of the DDoS attack detection to improve the security in the IoT, it was found that the proposed methods face many challenges such as inaccuracy, recalling, and error. Therefore, in this paper, a combination of GA feature selection and ensemble learning system is used to detect the DDoS attacks. The GA feature selection approach used for the best features selection of influential variables in the DDoS attacks of IoT can have a significant effect on the DDoS detection. The ensemble learning algorithms also produces a model based on the C4.5 decision tree algorithm, KNN and DNN with 50 hidden layers and learns from the past examples and identifies the possible intrusions that may occur in the future. So, In the proposed method, the best data are selected using the GA feature selection algorithm. Then, the best data are classified and identified by the ensemble learning system.

The Proposed Method

In order to show the general steps of the proposed method for implementing the ensemble learning include DT, KNN and DNN and GA algorithm for DDoS attacks detection on the IoT, a flowchart design is presented in the *Fig. 1*.

Firstly, as can be seen from the *Fig. 1*, in order to implement the proposed method to achieve the goal of the problem, which is to implement a hybrid GA algorithm and ensemble learning system such as C4.5 decision tree, KNN and DNN algorithms to DDoS attack detection on the IoT. The first step of the proposed method is to enter the dataset into the hybrid learning system. At this step, the *KDDCup99* dataset is entered into the system. This dataset is then pre-processed and the missing values and unused data is removed. The next step, according to the proposed method, is to normalize the data, which is described in detail in the next section. In the process of modeling, the problem data is placed within the range [0.1]. Normalization means that the models produced are not complex and also increase the accuracy of DDoS attacks detection into IoT.

Table 1: Summary of proposed methods

Authors	**Year**	**Method**	**Results**
Ray et al	2014	Provide a scalable RFID security framework and protocol to support Internet security on objects	The proposed protocol is able to protect against malware using malware detection methods and also has better security and scalability than other existing protocols.
Leak et al	2014	Provide a new architectural framework for e-health in the discussion of IoT security	The IoT promises health care, especially e-health.
Abu Mahara et al	2014	Review security issues on the IoT and privacy	The researchers presented their research on security and privacy on the IoT, and were able to provide good results for other researchers to use in future work.
Shahroui et al	2015	Provide an effective way to host end-to-end security on the IoT (HIP(	A review of the results shows that the proposed method, called CD-HIP, is efficient enough to have a minor security delay that is very small, and at a good level of compliance with the HIP standard.
Skari et al	2015	Check out a method for security and privacy and trust in IoT	In this paper, the researchers examined the main challenges and solutions available in the field of IoT security and also highlighted some important points for future research.
Kiwi et al	2016	Provide an efficient routing protocol for essential responses on the IoT (IoT(	The results of simulations and analyzes show that ERGID is better than EA-SPEED and SPEED in different periods to end delays, lost losses and energy consumption. In addition, they conducted a number of practical tests with sensor nodes, and the results showed that ERGID could improve REAL-TIME responses on the network.
Raphael et al	2016	Use a backup vector machine algorithm to DDoS attack detection on DARPA and *KDDCUP99* datasets	Acceptable accuracy of the backup vector machine algorithm in the attack detection process
Jay et al	2017	Provide a framework for security analysis automation on the IoT	The proposed framework is well able to detect the path of a possible attack and reduce the impact of a possible attack.
Swing et al	2017	A two-step method for DDoS attacks detection	The least error compared to other methods was the DDOS attack detection process.
Hook et al	2017	Provide a method called FPGA to DDoS attack detection	The use of the GA algorithm optimizes the final responses, and eventually DDOS attacks are detected with acceptable accuracy.
Sonny et al	2018	Use a technique called D-Face to DDoS attack detection	The results of this study have been more favorable than many of the previous studies.
Saffar et al	2018	Provide a Roadmap for Security Challenges on the IoT	The effectiveness of the proposed method compared to other previous methods
Antenna et al	2018	Using DBSCAN clustering algorithm and entropy technique to DDoS attack detection	Research has shown that there have been significant improvements over many methods.

Table 1. Continued

Kasavamvari et al	2019	DDoS attack detection and defense using a multifunctional system	The test results show that this proposed system offers optimized performance and improved security on the cloud platform.
Yousef et al	2019	Diagnosing and predicting DDoS attacks using machine learning methods	The results showed that about 30% of machine learning methods are used to predict and identify DDoS attacks.
Chiu et al	2019	Investigating the Defense Mechanism of DDoS Attacks Based on Cognitive Computations with Double Address Entropy	The proposed mechanism not only detects attacks quickly but also has high detection speed and low positive coefficient.
Shafi et al	2019	Prevent botnet DDoS attacks using Chinese block chain on the IoT	By analyzing and simulating using block chain and SDN blocks, the net notes were identified and reduced, preventing the devices from being used by attackers.
Tyne et al	2020	A Multi-Purpose Optimization (FS) Feature Selection Method for Identifying Service Distribution Denial Attacks (DDoS) on the IoT	Experimental results confirm that the proposed method for FS performed well and gained 99.9%, reducing the total number of features by nearly 90%.
Narrator et al	2020	Learn to identify and reduce DDoS attack on IoT through SDN-Cloud architecture	The DDoS attack detected an improved accuracy of 96.28%.
Lee et al	2020	Identify a real-time volumetric detection scheme for DDoS attacks on the IoT	The results showed that the proposed method shows real and efficient performance.
Kumar et al	2020	Analysis of various algorithms for detecting and reducing the effect of DDoS attacks in the context of the IoT	The results of this study show that security is a very important issue for reliable communications.
Iroum et al	2020	Diagnosis and prevention of DDoS attacks on the IoT	The results show that different techniques use filtering and speed limiting techniques to detect and prevent DDoS attacks with higher reliability.

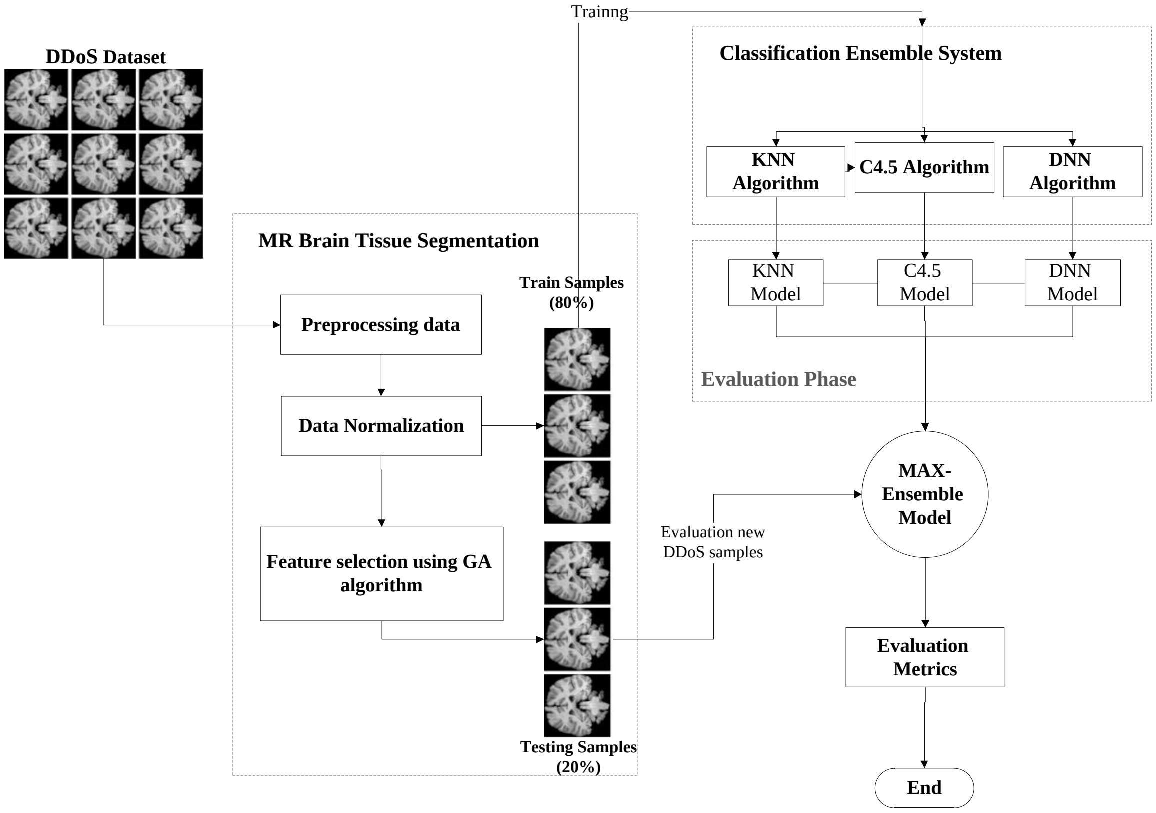


Figure 1. The proposed method Architecture

Secondly, with the help of the GA feature selection algorithm, the salient features of the *KDDCup99* dataset are selected. The selection of the most prominent features that have the greatest impact on the detection of DDoS attacks makes it possible, firstly, to reduce the detection time of DDoS attacks and, secondly, to increase the accuracy of intrusion detection; Because the training process of the model is based only on the features of the dataset that have the greatest impact on the detection of attacks.
Thirdly, the best features selection that is the output of the GA algorithm should be divided into two parts, which are:

- Training examples
- Testing samples

Training examples are used to training ensemble learning methods and to produce models related to each algorithm. Training samples make up 70 percent of the DDoS attacks. Testing samples, which account for 30 percent of the total DDoS attacks, are used to evaluate the hybrid proposed method. So, the Training and Testing samples are used for measure performance and the validity of the proposed method. After the training samples (70%) and the Testing samples (30%) were separated by the Balancing method sampling and the training samples were given to the ensemble learning algorithms such as the C4.5 decision tree, KNN and DNN algorithm.

Finally, we apply the training DDoS samples as the input for ensemble algorithms such as C4.5 decision tree, KNN and DNN algorithms. Each of these algorithms produces a model based on the training DDoS examples. This model has a tree structure, vector and neural structure. Testing samples are then applied to created models. Based on the available samples, the validation of maximum-based attack detection is performed. Then, the proposed method is evaluated and criteria such as accuracy, precision, recall, error, etc. are calculated.

In this section, all the steps outlined in *Fig. 1,* described in the previous section will be discussed in more detail. The most important parts of the proposed method in this paper will be presented under the relevant sections. The proposed method steps include:

a. Data Pre-processing

After the data is entered into the proposed system, the data is pre-processed and the discarded and unused data is removed. There are several methods to pre-process data, including:

- Data cleaning method
- Data collection method
- Data transfer method

Due to the needs of our problem in this paper, we have only used the data cleansing method. The proposed strategy is to analyze the data and identify if the row or column had empty or unused values. We then examine the values before and after the sample that has a blank or unused value and calculate their average. Finally, we will replace the empty value with the average obtained. This eliminates waste samples and generates more consistent data.

b. Data normalization

In the preprocessing step, in order to get better results, we normalize the values of each feature of the used dataset between 0 and 1, then we randomly move the general data matrix lines so that the data from the initial state of the sum Gathered, get out. In other words, all the datasets are written in the form of a matrix and the normalization of the rows of the matrix is done. Normalization is due to higher accuracy. The following equation is used to normalize the values of each dataset.

$$\mathrm{Normalize(x)} = \frac{(\mathrm{x} - \mathrm{X_{min}})}{(\mathrm{X_{max}} - \mathrm{X_{min}})} \tag{1}$$

X_{max} and X_{min} are the maximum and minimum values in the X feature range. After normalizing the data, the values of all the attributes are in the range [0.1]. Therefore, after pre-processing and preparation of the initial data and application of the normalization technique on the data, the process of selecting the best feature is done using the GA feature selection algorithm, which is described in the next section.

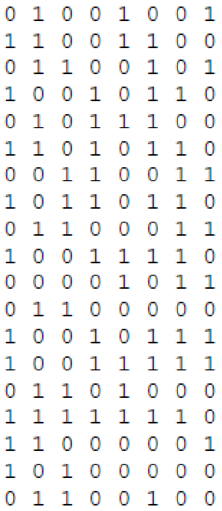
c. Feature selection using GA algorithm

Since 1986, imitation of living things has been considered for use in powerful algorithms for optimization problems, called feature selection computational techniques. In fact, the GA algorithm is a programming technique that uses GA evolution as a problem-solving model. When the word survival conflict is used, the negative value often comes to mind. Perhaps the law of the jungle will always come to mind and a stronger survival order! Of course, to ease your mind, you can think that the strongest have not always won; Dinosaurs, for example, have survived the game in spite of their huge size during a perfectly normal process, while much weaker creatures have survived. Apparently, nature does not choose the best based only on the body! In fact, it is better to say that nature chooses the most suitable (fittest) and not the strongest [30] [31].
The law of natural selection is that only species of a population continue to have offspring that have the best characteristics, and those that do not have these characteristics gradually disappear over time. For example, suppose a certain type of person has much more intelligence than the rest of a community. Under completely normal conditions, these people will progress better, and this well-being will lead to a longer life and better fertility. Now, if this characteristic (intelligence) is hereditary, naturally in the next generation of the same society, the number of intelligent people will be more due to the greater birth rate of such people. If you continue this process, our exemplary society will become smarter over the generations. Thus, a purely natural mechanism has been able to virtually eliminate low-intelligence individuals from society over several generations, in addition to increasing the average level of intelligence in society. Thus, it can be seen that by using a very simple method (gradual elimination of unsuitable species and at the same time higher reproduction of optimal species) has been able to constantly improve each generation in terms of different characteristics.

The evolutionary process of the GA algorithm is a simple, biological simulation. This evolution starts from a random population with a probability-based distribution and is usually uniform, which updates the population in stages to generations. In each generation, several people randomly change from the current population to a new population based on a function of fit, composition, direction, and choice. Each GA algorithm has several operators. In the following, while defining and applying each operator in the GA algorithm, we will explain what parts of the association rules and existing examples are equivalent to each function.

Chromosomes And Genes

Each chromosome in the GA algorithm has a number of genes. In the proposed method, chromosomes contain dataset records. The *Fig.2,* shows chromosomes and genes of the default dataset:



```
0 1 0 0 1 0 0 1
1 1 0 0 1 1 0 0
0 1 1 0 0 1 0 1
1 0 0 1 0 1 1 0
0 1 0 1 1 1 0 0
1 1 0 1 0 1 1 0
0 0 1 1 0 0 1 1
1 0 1 1 0 1 1 0
0 1 1 0 0 0 1 1
1 0 0 1 1 1 1 0
0 0 0 0 1 0 1 1
0 1 1 0 0 0 0 0
1 0 0 1 0 1 1 1
1 0 0 1 1 1 1 1
0 1 1 0 1 0 0 0
1 1 1 1 1 1 1 0
1 1 0 0 0 0 0 1
1 0 1 0 0 0 0 0
0 1 1 0 0 1 0 0
```

Figure 2. chromosomes and genes of the default dataset

As can be seen, the chromosome is one of the records in the above dataset. for example:

```
0 1 0 0 1 0 0 1
```

And each of the values 0 and 1 is a gene. Therefore, in the proposed method, each sample of the dataset or primary population is a chromosome, and each value of each sample is a gene. The total number of chromosomes is equal to the total number of samples or records in the dataset.

Primary Population

The sample size or the total number of samples or data on which the proposed method is executed is called the population. The population is calculated based on the following equation [32]:

$$\mathrm{popsize} = \mathrm{order}[\frac{\mathrm{l}}{\mathrm{k}} + 2^{\mathrm{k}}] \tag{2}$$

Where l is the number of chromosomes in the dataset and k is the average size of the dataset. In other words, in our paper, the size of the population is equal to the volume of dataset records.

GA Selection Operator

To choose the best answers for the reproduction of the generation (production of a new population), you must use a method that chooses the best possible answer. Among the various methods, we examine the roulette wheel proposed by the Netherlands [33]. The basic idea of this method is to determine the probability of survival for each chromosome, in proportion to its fit. The roulette wheel is designed to show these possibilities, and the selection process is based on the simultaneous rotation of the wheel figures to the size of the population. Calculate the competency values or V_k for each chromosome. Assume f is a function of the target.

$$\mathrm{eval}(\mathrm{V_k}) = \mathrm{f(m)}, = 1.2.\ldots.\mathrm{pop} - \mathrm{size} \tag{3}$$

Calculate the sum of the total competency values of the existing chromosomes:

$$\mathrm{F} = \sum_{\mathrm{K=1}}^{\mathrm{POP-SIZE}} \mathrm{eval}(\mathrm{V_k}) \tag{4}$$

Calculate the relative probability of *P_k* for each chromosome:

$$\mathrm{P_k} = \frac{\mathrm{eval}(\mathrm{V_k})}{\mathrm{F}}\ \mathrm{k} = 1.2.\ldots.\mathrm{pop} - \mathrm{size} \tag{5}$$

Calculate the cumulative probability of q_k for each chromosome:

$$q_k = \sum_{j=1}^{k} P_j \tag{6}$$

The selection process for rotating the roulette wheel begins with the number of *pop* size loads, and each time a chromosome is selected to be present in the new generation as follows:

Step 1. Make a random number like r in the distance [0,1].

Step 2. If $r < q_1$ The V_1 chromosome, which is the first chromosome, is then selected, otherwise *Kth* chromosome where $2 \leq k \leq pop - size$ and $q_{k-1} \leq r \leq q_k$ It is, it is selected.
Therefore, in this paper, the selection process is based on the best fit for each run. The relationship between the calculation of fitness is discussed below.

GA Crossover Operator

In this paper, a one-point *Crossover* operation is used. If a typical node becomes a group after the *Crossover* operation, all other common nodes that are close to this new instance should be examined, and if this happens, they should become members of this new category.

GA Mutation Operator

In different chromosomes, this operator makes unplanned random changes and inserts genes that did not exist in the primary population. An important concept has been proposed for this operator, called the p_m mutation rate.
Mutation rates are a percentage of the total number of existing genes that change. If the mutation rate is too small, a large number of genes that could be helpful will not be tested; But if the mutation rate is too high, babies lose their resemblance to their parents. This destroys the historical memory of the algorithm. There are different mutation operators, we only look at the uniform type. In this operator, a gene from a chromosome is randomly selected and its

value is converted to another random value. First, a random number in the interval, and [1L], which is the desired chromosome length, and the gene in that location changes from the chromosome. Suppose the parent chromosome is opposite:

010010110	Parent

As can be seen, the length of the chromosome is 10, assuming that the random number generated in the range [10, 1] is equal to 5, so the gene in place 5 changes, that is, 1 becomes 0.

010000110	Baby

Therefore, in this dissertation, in the process of mutation, the mutation rate is based on the change of bit 0 to 1 and vice versa. As a result, if the bit is one, it becomes zero, and if it is zero, it becomes one. In fact, a leader becomes an ordinary knot, and vice versa.

Functionality of GA feature selection algorithm to select feature

The fit function is the main part of the GA algorithm. In this paper, the following fit function is used to calculate the fit of each item. In Equation (7) the fit formula is shown.

$$\mathrm{fitness_i} = \mathrm{w_A} \times \mathrm{acc_i} + \mathrm{w_F} \times \left[1 - \frac{\left(\sum_{j=1}^{n_F} \mathrm{f_i}\right)}{\mathrm{n_F}}\right] \tag{7}$$

$w_A$the weight accuracy shows the classification of the SVM algorithm to select the feature. This means that for each feature, the SVM algorithm is executed once and the accuracy is calculated. acc_i Accuracy of SVM algorithm with RBF core, w_F weight is selected for the number of features, f_i the mask value indicates the attribute, the value "1" indicates that the j attribute has been selected, and the "0" value indicates that the j attribute has not been selected. n_F shows the total number of features. Relationship (8) is used to calculate acc_i, which shows the accuracy of the SVM classification.

$$\mathrm{acc} = \frac{\mathrm{cc}}{\mathrm{cc} + \mathrm{uc}} \times 100\% \tag{8}$$

cc shows the number of samples that are classified correctly and *uc* shows the number of samples that are incorrectly classified.

Steps of implementing the GA algorithm

First, depending on the problem, the variables that need to be determined are identified, then we code them properly and display them in the form of a chromosome. Based on the objective function, the value of the chromosomes is determined and the initial population of the arbitrariness is randomly selected. This is followed by the value of fit for each primary population chromosome. You can see the algorithm below.

1. Select the primary population and calculate their suitability;
2. Choosing chromosomes to create babies;
3. Intersection operation;
4. Mutation operation;
5. Assessing children by calculating the fit of each chromosome;
6. Replace more suitable chromosomes;
7. If the answers are appropriate, go to 8, otherwise go to 2;
End;

d. Separation of Training and Testing samples

Sampling of the desired data is one of the stages of data mining that has been considered in this research. There are several sampling methods, three of which are the most important: [34]

- Random sampling
- Classified sampling
- Balanced sampling

Random sampling is one of the simplest sampling methods that works randomly and separates samples from the main data as Training and Testing data. One of the disadvantages of this method is that it is not possible to sample any particular category and ultimately reduces the accuracy of data classification and validation of the proposed method.

Classified sampling is also one of the improved methods of random sampling. This method performs the sampling process based on probability and also selects the samples as a percentage. This sampling method is also difficult and may not select probability-based samples.

Balanced sampling is one of the methods that selectively selects the required samples from the existing categories and classes. This method does not have the problem of the previous methods and finally selects the balanced data and samples between the available samples.

The method used to separate training data and Testing data is the Balancing method. Therefore, 70% of the samples are separated as Training samples and 30% as Testing samples.

e. The proposed ensemble learning system

The proposed method uses the most popular classification methods, such as DNN, the C4.5 decision tree, and the KNN algorithm, KNN in combination, to secure the information of requests and processes sent by DDoS attacks. Finally, the above methods are combined and at each stage, the best answer is selected from the provided answers and is determined as the final result. Due to the nature and sensitivity of user process categorization environments, including suspicious, etc., the proposed method is the DNN composition, the C4.5 decision tree and KNN. The combination of these two methods is presented in the form of a reinforcement learning system.

Based on the proposed method, which is a combination of deep neural methods with 2 hidden layers, C4.5 decision tree and KNN classification algorithm, this section describes how to combine these methods.
The procedure is as follows:

- The training data is entered into a DNN model with 50 hidden layers, decision C4.5 tree and KNN classification algorithm. The type of DNN algorithm used was the standard algorithm of the DNN.
- All three methods described teach their models and are ready to receive Testing data to identify reputable and unreliable customers.
- After the relevant models have been trained, test data, which is 30% of the total data, is entered into the models for evaluation. Each model returns its output and its prediction and detection rate as output.
- The output of these methods is connected to the input of the core of the boosting system and according to the parameter in the boosting system which is min, max, Avg, the best prediction and classification is selected and used as the output.
- By presenting the above proposed method in each execution and prediction of the new case, the best response to the output is sent and finally we will see the most optimal response with high accuracy and the least error. The following figure gives an overview of the reinforcement learning system.

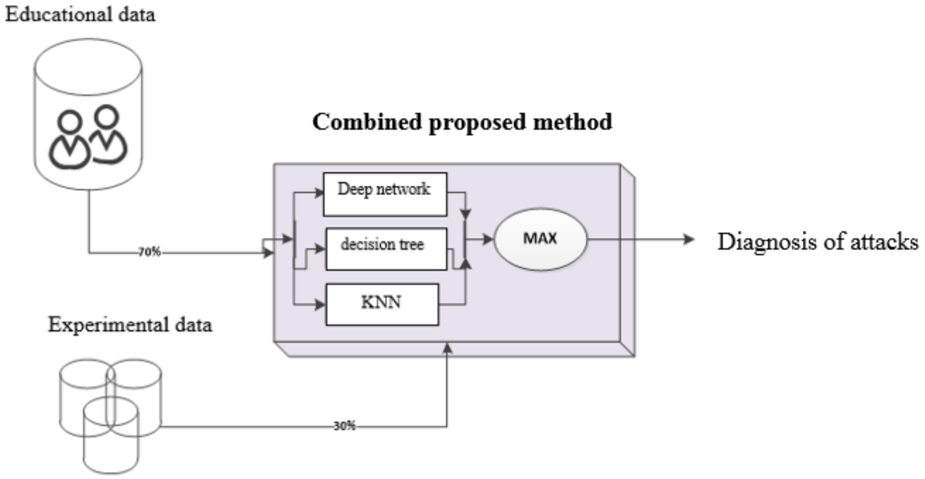


Figure 3. Flowchart Recommended Enhancement Method

Therefore, according to the view presented in this section, we will simulate the proposed method in the next section and evaluate the results obtained with other methods.

2. Experimental Results

In this section, the proposed method is compared with other methods so that, first, the test settings are laid out, then the benchmarks are set out and finally, the experiments are analysed.

a. Experimental Setup

The proposed method is implemented using the MATLAB simulator version 2015a and rapidminer version 9.2. The operating system is Windows 10, the operating system is of the 32-bit type, 4GB of RAM is used from which - 3.06GB is usable, with 7-core Intel processor (Core ™) i7 CPU) - Q 720 and processor base frequency of @ 1.60GHz.

b. Dataset

The first step in the proposed method is to enter the *KDDCup99* attack data into the proposed system. This data will be described in the following section. This paper uses the *KDDCUP99* dataset to diagnose network anomalies. The DARPA program, used in 1998 to identify the attack, was developed and managed by MIT's Lincoln Laboratory. The purpose of this program was to investigate and evaluate the research on identifying intrusion attacks. A standard dataset was presented that included a variety of simulated intrusions into the environment of a military network. The KDD competition in 1999 also uses a version of the same dataset. Lincoln Laboratory created an environment for obtaining 9-week raw data for TCP dumps in a local LAN network that simulated the US Air Force's LAN network. They designed the network environment as if it were really an air force network, and they attacked it with all kinds of attacks. The raw training data volume was four gigabytes, which was related to TCP dump data up to 9 weeks of network traffic week. The training dataset a record of five million connections, and similarly, it had two million connections in two weeks. A connection is actually a sequence of TCP packets that are created at a given start and end time, between which data is sent from an IP address to a destination IP address under well-defined protocols. Each connection can be labeled with one of the normal headings and labeled with a specific type of attack. The main names of the *KDDCup99* dataset features are shown in the table below.

Table 2. The *KDDCup99* data features

Row	**feature name**	**Description**	**Type**
1	duration	Period	Continuous
2	protocol_type	Protocol type	Discrete
3	service	Service type	Discrete
4	flag	Flag	Discrete
5	src_bytes	Number of bytes of origin	Continuous
6	dst_bytes	Number of destination bytes	Continuous
7	land	Background type	Discrete
8	wrong_fragment	Error field	Continuous
9	urgent	Number of essential packages	Continuous
10	hot	Hot	Continuous
11	num_failed_logins	The number of unsuccessful entries	Continuous
12	logged_in	Number of entries	Discrete
13	num_compromised	Number of retaliations created	Continuous
14	root_shell	Root shell	Continuous
15	su_attempted	Number of infiltration attempts	Continuous
16	num_root	Root number	Continuous
17	num_file_creations	Number of files generated	Continuous
18	num_shells	Number of shells	Continuous
19	num_access_files	Number of files available	Continuous
20	num_outbound_cmds	Number of out-of-range commands	Continuous
21	is_host_login	Login entry by the host	Discrete
22	is_guest_login	Entrance flags by guests	Discrete
23	count	Counting	Continuous
24	srv_count	Count the origin	Continuous

25	serror_rate	Error rate	Continuous
26	srv_serror_rate	Server error rate	Continuous
27	rerror_rate	Recipient error rate	Continuous
28	srv_rerror_rate	Server error rate on the server	Continuous
29	same_srv_rate	Sender error rate	Continuous
30	diff_srv_rate	Error difference	Continuous
31	srv_diff_host_rate	Error difference rate at source	Continuous
32	dst_host_count	Error difference rate at the destination	Continuous
33	dst_host_srv_count	Number of host services per source	Continuous
34	dst_host_same_srv_rate	Number of host services at the destination	Continuous
35	dst_host_diff_srv_rate	The difference in the number of destination services at the origin	Continuous
36	dst_host_same_src_port_rate	The similarity rate of the host port at the origin and destination	Continuous
37	dst_host_srv_diff_host_rate	Host rate difference at destination at origin	Continuous
38	dst_host_serror_rate	Host error rate at destination	Continuous
39	dst_host_srv_serror_rate	Host service error rate at the destination	Continuous
40	dst_host_rerror_rate	Target service error rate	Continuous
41	dst_host_srv_rerror_rate	Host service error rate at source	Continuous
42	Class	Attack/Normal	Discrete

Therefore, this data is entered into the system raw with the features provided.

c. Performance Metrics

In this paper, several criteria have been used to evaluate and compare the proposed method with other methods, which are:

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{9}$$

The Eq. 9, is used to check the precision of the proposed method. The *TP (true positive)* parameter indicates the number of samples that have been correctly identified. And the *FP (false positive)* parameter also indicates the number of samples that have been misdiagnosed.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{10}$$

The Eq. 10, indicates the amount of the proposed method recall, and the *FN (false negative)* parameter indicates the number of samples that have been correctly identified.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{11}$$

The above relation is used to calculate the accuracy, the only parameter that is not described being *TN (true negative)*, which indicates the number of samples that were correct but misdiagnosed.

d. Experimental Results

In this section, the process of simulating and DDoS attacks detection is performed by applying the GA feature selection algorithm (17 best features selected) and without the GA algorithm (42 main features of the *KDDCup99* dataset). The *Table.3,* shows a comparison of the proposed method metrics with other basic methods for DDoS attacks detection without applying a GA algorithm.

Table 3. Comparison of the proposed method evaluation metric with other basic methods for DDoS attacks detection without the GA algorithm

	Ensemble Method	**DNN**	**KNN**	**C4.5**
Accuracy	**99.84**	98.25	96.81	97.8
Precision	**98.9**	97.4	96.88	97
Recall	**98.63**	97.47	96.57	97.1

As can be seen from the *Table. 3,* the improvement in the accuracy of the proposed method (ensemble learning without GA feature selection algorithm) compared to DNN, KNN and C4.5 decision tree algorithms are 1.59%, 3.03% and 2.04%, respectively. The improvement of the proposed method precision compared to DNN, the KNN and the C4.5 decision tree is 1.5%, 2.02% and

1.9%, respectively. The improvement in the proposed method recall compared to DNN algorithms, the KNN and the C4.5 decision tree is 1.16%, 2.06% and 1.53%, respectively.

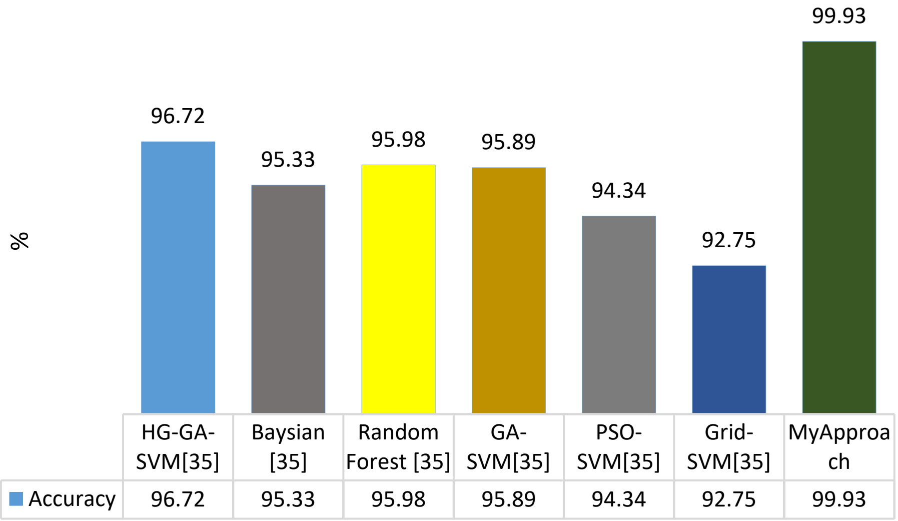
The *Table. 4,* shows a comparison of the proposed method evaluation metrics with other basic methods for DDoS attacks detection by applying a GA algorithm.

Table 4. comparison of the proposed method evaluation metrics with other basic methods for DDoS attacks detection by applying a GA algorithm

	Ensemble GA	**DNN+GA**	**KNN+GA**	**C4.5+GA**
Accuracy	**99.93**	98.92	**97.95**	97.81
Precision	**99.74**	98.22	97.37	97.66
Recall	**99.05**	98.06	97.09	97.29

As can be seen from the *Table. 4,* the improvement in the accuracy of the proposed method (ensemble learning with GA feature selection algorithm) compared to DNN, KNN and C4.5 decision tree algorithms are 1.01%, 1.98% and 2.12%, respectively. The improvement of the proposed method precision compared to DNN, the KNN and the C4.5 decision tree is 1.54%, 2.37% and 2.08%, respectively. The improvement in the proposed method recall compared to DNN algorithms, the KNN and the C4.5 decision tree is 0.99%, 1.96% and 1.76%, respectively. We will compare the results of the proposed method with other methods. The authors in [35], used GA feature selection algorithm to select best features from the *KDDCup99* attack dataset and the SVM machine learning algorithm to DDoS attack detection. The results of their research were finally accurate at about 97%. Therefore, the results and findings obtained are compared with the [35]. The *Fig. 4,* shows a comparison of DDoS attack detection accuracy in the proposed method (hybrid Ensemble learning and GA) compared to other methods in [35] on the KDDCup99 dataset.

As can be seen from the *Fig. 4,* the DDoS attack detection accuracy in the proposed method (Ensemble + GA) on *KDDCup99* dataset is 99.93%, Grid SVM method [35] is 92.75%, PSO-SVM method [35] is 94.34%, GA-SVM method [35] is 95.89 %, Random Forest method [35] is 95.98%, business network method [35] is 95.33% and HG-GA SVM method [35] is 96.72%. Finally, the accuracy of the proposed method improved compared to other methods such as Grid SVM, PSO-SVM, GA-SVM, Random Forest, Business Network and HG-GA SVM, respectively 7.18%, 5.59%, 4.04%, 3.95%, 4.6 % And 3.21%.



	HG-GA-SVM[35]	Baysian [35]	Random Forest [35]	GA-SVM[35]	PSO-SVM[35]	Grid-SVM[35]	MyApproach
■ Accuracy	96.72	95.33	95.98	95.89	94.34	92.75	99.93

Figure 4. Comparison of DDoS attack detection accuracy in the proposed method (hybrid Ensemble learning and GA) compared to other methods in [35] on the KDDCup99 dataset

Another important criterion is the correct detection rate, which raises the accuracy of the proposed method. The *Fig. 5*, shows the comparison of the correct detection rate in the proposed method with the applying of GA feature selection compared to other methods proposed in [35].

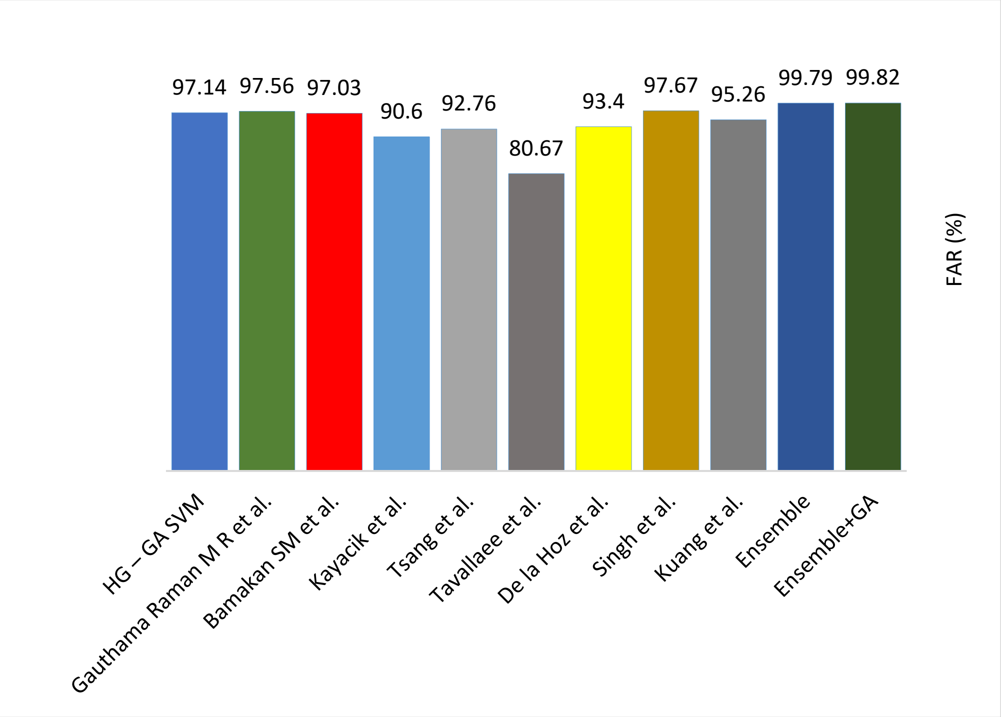


Figure5. Comparison of the fair detection rate in the proposed method with the application of feature selection compared to other methods proposed in [35]

As can be seen from the *Fig.5*, the improvement of the correct detection rate in the proposed method by applying the GA feature selection relative to the proposed method without GA and the Kuang et al, Singh et al, De la Hoz et al, Tavallaee et al, Tsang et al, Kayacik et al, Bamakan SM et al, Gauthama Raman MR et al and HG-GA SVM methods respectively 0.03%, 4.56%, 2.15%, 6.42%, 19.15%, 7.06%, 9.22%, 2.79%, 2.26% and 2.68%.

CONCLUSIONS AND SUGGESTIONS

One of the main objectives of this paper was presenting a new hybrid approach for DDoS attack detection and enhancing the security of the IoT. In this paper, a combination of GA algorithm feature selection and ensemble learning system, including C4.5 decision tree, DNN and KNN algorithms has been used to DDoS attack detection. To validate the proposed method, the results were compared with other methods, including machine learning methods and combined with other optimization methods. In this paper, 10% of the *KDDCup99* dataset is used for simulation. First, in the data preprocessing step, the values of all the attributes are converted to numbers, and also the output characteristic values are changed to two values, zero and one. The results of the paper show the high accuracy of the proposed method for DDoS attacks detection compared to other recent methods of about 5%.

Thus, it is suggested to use the reinforcement learning algorithms to improve the DDoS attack detection process in the IoT, to use a deep Long Short-Term Memory (LSTM) neural network in order to enhance the attack process and to apply the semantic hierarchy in order to improve the DDoS attack detection in the future studies.

REFERENCES

Lee, W., Stolfo, S. J., & Mok, K. W. (1999, May). A data mining framework for building intrusion detection models. In Proceedings of the 1999 IEEE Symposium on Security and Privacy (Cat. No. 99CB36344) (pp. 120-132). IEEE.

Bhuyan, M. H., Bhattacharyya, D. K., & Kalita, J. K. (2015). An empirical evaluation of information metrics for low-rate and high-rate DDoS attack detection. Pattern Recognition Letters, 51, 1-7.

Wang, F., Wang, H., Wang, X., & Su, J. (2012). A new multistage approach to detect subtle DDoS attacks. Mathematical and Computer Modelling, 55(1-2), 198-213.

Thing, V. L., Sloman, M., & Dulay, N. (2009, June). Adaptive response system for distributed denial-of-service attacks. In 2009 IFIP/IEEE International Symposium on Integrated Network Management (pp. 809-814). IEEE.

J. Yu, H. Lee, M.-S. Kim, D. Park, Traffic flooding attack detection with SNMP MIB using SVM, Computer Communications 31 2008 pp. 4212–4219.

S. Mansfield-Devine, Computer Fraud & Security, Volume 2014, Issue 10, 2014 pp.15-20.

Yu, J., Lee, H., Kim, M. S., & Park, D. (2008). Traffic flooding attack detection with SNMP MIB using SVM. Computer Communications, 31(17), 4212-4219.

Ray, B. R., Abawajy, J., & Chowdhury, M. (2014). Scalable RFID security framework and protocol supporting Internet of Things. Computer Networks, 67, 89-103.

Lake, D., Milito, R. M. R., Morrow, M., & Vargheese, R. (2014). Internet of things: Architectural framework for ehealth security. Journal of ICT Standardization, 1(3), 301-328.

Abomhara, M., & Køien, G. M. (2014, May). Security and privacy in the Internet of Things: Current status and open issues. In 2014 international conference on privacy and security in mobile systems (PRISMS) (pp. 1-8). IEEE.

Sahraoui, S., & Bilami, A. (2015). Efficient HIP-based approach to ensure lightweight end-to-end security in the internet of things. Computer Networks, 91, 26-45.

Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. Computer networks, 76, 146-164.

Qiu, T., Lv, Y., Xia, F., Chen, N., Wan, J., & Tolba, A. (2016). ERGID: An efficient routing protocol for emergency response Internet of Things. Journal of Network and Computer Applications, 72, 104-112.

Hassan, M. R. (2016). Intrusion Detection System Based on Cost Based Support Vector Machine. In Recent Advances in Information and Communication Technology 2016 (pp. 105-115). Springer, Cham.

Ge, M., Hong, J. B., Guttmann, W., & Kim, D. S. (2017). A framework for automating security analysis of the internet of things. Journal of Network and Computer Applications, 83, 12-27.

Wei, S., Ding, Y., & Han, X. (2017, June). TDSC: Two-stage DDoS detection and defense system based on clustering. In 2017 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W) (pp. 101-102). IEEE.

Hoque, N., Kashyap, H., & Bhattacharyya, D. K. (2017). Real-time DDoS attack detection using FPGA. Computer Communications, 110, 48-58.

Behal, S., Kumar, K., & Sachdeva, M. (2018). D-FACE: An anomaly based distributed approach for early detection of DDoS attacks and flash events. Journal of Network and Computer Applications, 111, 49-63.

Sfar, A. R., Natalizio, E., Challal, Y., & Chtourou, Z. (2018). A roadmap for security challenges in the Internet of Things. Digital Communications and Networks, 4(2), 118-137.

Girma, A., Garuba, M., & Goel, R. (2018). Advanced machine language approach to detect DDoS attack using DBSCAN clustering technology with entropy. In Information Technology-New Generations (pp. 125-131). Springer, Cham.

Kesavamoorthy, R., & Soundar, K. R. (2019). Swarm intelligence based autonomous DDoS attack detection and defense using multi agent system. Cluster Computing, 22(4), 9469-9476.

Yusof, A. R. A., Udzir, N. I., & Selamat, A. (2019). Systematic literature review and taxonomy for DDoS attack detection and prediction. International Journal of Digital Enterprise Technology, 1(3), 292-315.

Cui, J., Wang, M., Luo, Y., & Zhong, H. (2019). DDoS detection and defense mechanism based on cognitive-inspired computing in SDN. Future Generation Computer Systems, 97, 275-283.

Shafi, Q., & Basit, A. (2019, January). DDoS Botnet prevention using blockchain in software defined Internet of Things. In 2019 16th

International Bhurban Conference on Applied Sciences and Technology (IBCAST) (pp. 624-628). IEEE.

Tian, G., & Chambers, J. (2020). Multi-Objective based Feature Selection for DDoS Attack Detection in IoT Network. IET Networks.

Ravi, N., & Shalinie, S. M. (2020). Learning Driven Detection and Mitigation of DDoS Attack in IoT via SDN-Cloud architecture. IEEE Internet of Things Journal.

Li, J., Liu, M., Xue, Z., Fan, X., & He, X. (2020). RTVD: A Real-Time Volumetric Detection Scheme for DDoS in the Internet of Things. IEEE Access, 8, 36191-36201.

Kumar, N., Mittal, N., Thakur, P., & Srivastava, R. (2020). Analysis of Different Detection and Mitigation Algorithm of DDoS Attack in Software-Defined Internet of Things Framework: A Review. In Recent Trends and Advances in Artificial Intelligence and Internet of Things (pp. 597-607). Springer, Cham

Irum, A., Khan, M. A., Noor, A., & Shabir, B. (2020). DDoS Detection and Prevention In Internet of Things (No. 2486). EasyChair.

[30]. Sivanandam, S. N., & Deepa, S. N. (2008). GA algorithm optimization problems. In Introduction to GA Algorithms (pp. 165-209). Springer, Berlin, Heidelberg.

Weile, D. S., & Michielssen, E. (1997). GA algorithm optimization applied to electromagnetics: A review. IEEE Transactions on Antennas and Propagation, 45(3), 343-353.

Indira, K., & Kanmani, S. (2011, July). Association rule mining using GA algorithm: The role of estimation parameters. In International Conference on Advances in Computing and Communications (pp. 639-648). Springer, Berlin, Heidelberg

Goldberg D. E (1989), GA Algorithm in Search, Optimization & Machine Learning New York: Addison-Wesely Publishing Company.

Malik, J. S., Goyal, P., & Sharma, A. K. (2010). A comprehensive approach towards data preprocessing techniques & association rules. In Proceedings of The4th National Conference.

Raman, M. G., Somu, N., Kirthivasan, K., Liscano, R., & Sriram, V. S. (2017). An efficient intrusion detection system based on hypergraph-GA algorithm for parameter optimization and feature selection in support vector machine. Knowledge-Based Systems, 134, 1-12.