Compare between PSO and artificial bee colony optimization algorithm in detecting DoS attacks from network traffic

Maha A. A. Mohammad, Muna M. T. Jawhar
Department of Software, College of Computer Sciences and Mathematics, University of Mosul, Mosul City, Iraq

ABSTRACT
Our world today relies heavily on informatics and the internet, as computers and communications networks have increased day by day. In fact, the increase is not limited to portable devices such as smartphones and tablets, but also to home appliances such as: televisions, refrigerators, and controllers. It has made them more vulnerable to electronic attacks. The denial of service (DoS) attack is one of the most common attacks that affect the provision of services and commercial sites over the internet. As a result, we decided in this paper to create a smart model that depends on the swarm algorithms to detect the attack of denial of service in internet networks, because the intelligence algorithms have flexibility, elegance and adaptation to different situations. The particle swarm algorithm and the bee colony algorithm were used to detect the packets that had been exposed to the DoS attack, and a comparison was made between the two algorithms to see which of them can accurately characterize the DoS attack.

Keywords: Artificial bee colony algorithm, DoS attack, Network security, Network traffic, Particle swarm optimization, Swarm intelligence technique, Wireshark

This is an open access article under the CC BY-SA license.

Corresponding Author:
Maha A. A. Mohammad
Department of Software, College of Computer Sciences and Mathematics
University of Mosul, Mosul City, Iraq
Email: mahaabd77@uomosul.edu.iq

1. INTRODUCTION
The denial of service (DoS) attacks is a way of attacking internet servers by overwhelming them with messages that exceed the provider’s ability to process them quickly. Thus, this process will cause the provider to offer sluggish services and the provider’s inability to access or use them properly. It is a favourite method of attacking websites for hackers, pirates and electronic spammers, particularly as it seems and with the knowledge of many internet security experts, as if there is currently no treatment for this method of attacking websites; on this basis, some circles call this form of attack internet [1].

A 15-year-old boy named Michael Kals, according to Norton, carried out the first documented distributed denial of service (DDoS) attack in 2000; it was launched to temporarily disable huge websites such as: Yahoo and eBay, causing an error message to appear. Nowadays, experts accept that there is no way for handling or stopping distributed attacks. On the other hand, some experts recommend using authentication and encryption methods to manage transmitted information packets. They also agree on the idea of it is impractical for these methods to solve the online problem. Other experts suggest wizards and guidelines that can distinguish and filter denial of service attacks before they impact the operating system, a solution that is being developed today by many businesses that manufacture denial-of-service attack software [2].

The increase in the number of attacks over the network and the increase in their intensity and disruptive effect year after year and their impact on the sales of sites and services over the network have been reported by various studies. In fact, this is due to many reasons; the most severe of which, abbreviated as DoS, are known as denial of service attacks. In general, such attacks have been around for years, but their...
control now than at any time before; and they have entered a level of maturity where they hit particular targets and are intended for commercial purposes [3].

In most situations, the denial of service attacks have been used to threaten major corporations’ web servers such as: banks, internet retailers, and even government and public services. It is necessary, however to keep in mind that any internet-connected computer, server, or network can be a potential target for these types of attacks. In recent years, crypto-currencies have gained prominence, making trading networks more susceptible to distributed denial of service attacks. For instance, when Bitcoin Gold was officially launched, it immediately became the target of a widespread DoS attack that crippled its website for several hours [4].

Most of the previous studies have used the intelligent swarm algorithms which caught the attention of researchers and were used in most applications. Amudha et al. [5] have created a hybrid algorithm to combine a modified artificial bee colony (MABC) with an enhanced particle swarm optimization (EPSO) to identify the issue of intrusion detection. In deed, the aim of the research was to identify the most relevant traits that can explain network traffic and to test the effect of the algorithm on the success of the proposed hybrid classification method. To investigate the performance of the proposed method, the knowledge discovery in databases (KDD Cup’99) modular dataset is used to detect intrusion from the location of machine learning repository. In 2011, Lua and Yow [6] proposed a new method to reduce DDoS attacks by using a type of swarm network which is smart and fast-flowing. What is needed is a swarm network design to ensure an independent coordination and swarm nodes which were assigned to perform migrations. The water drop algorithm has been applied to optimize the distribution and parallelism. Fast flow technology was used to maintain the communication between swarm nodes, clients, and servers. Fast flow service networks have been used to build a transparent service, which allows for minimal modifications to existing cloud services (such as hyper text transfer protocol (HTTP) and simple mail transfer protocol (SMTP)) software simulation; they are emerging from 400,000 client nodes and 10,000 swarm nodes, maintaining (99.96 %) of the packet delivery amount when the network is attacked by a similarly sized DDoS network of 10,000 dedicated malicious nodes. In 2018, Kalinin et al. [3] presented a special technique to prevent a set of routing attacks in ad hoc self-organizing networks. The new method or technology works to develop monitoring and estimating the packet transmission parameter (P-Secure) by using the ant swarm algorithm in order to create a safe path or path in the network, where all nodes act as agents for the analysis of the security of neighboring nodes. In 2019, Qureshi et al. [7] transmitted data through the internet, which is an insecure channel. An anomaly-based intrusion detection scheme has been proposed which enables it to protect sensitive information and detect new cyber attacks. An artificial bee colony (ABC) algorithm is used to deliver the recurrent neural network (RNN) based recurrent neural network-artificial bee colony (RNN-ABC) system. The proposed scheme is being implemented on network security laboratory of the knowledge discovery in databases (NSL-KDD) training and invisible data testing. Experimental results indicate that swarm intelligence and RNN successfully rated new attacks with an accuracy of (91.65%). In addition, the performance of the proposed scheme is also compared to a multi-layer and mixed infiltration detection system multi layer persptron (MLP) using sensitivity, minimum mean square error (MMSE), MSE standard deviation (SDMSE), best mean square error (BMSE) and weight mean square error (WMSE) parameters. The results and experimental tests confirm the accuracy and high durability of the proposed method. In addition, the DoS attack was one of the important attacks that had a bad effect on the work of the networks and services they provide. This attack began to infiltrate the cloud and reach the internet of things that entered all areas of life. In fact, it was one of the main reasons that made us choose this attack in an attempt to detect it in order to protect customers from its effects. The main goal of the study that we did was to try to improve the detection rates obtained during previous studies. We also wanted to discover the effectiveness of the algorithms that were selected in the process of detecting attacks that hinder the work of networks and customer service.

2. DoS ATTACK

Denial of service, known as DoS, is a form of malicious attack carried out by the attacker or a group of attackers in order to eliminate the target computers or network resources from service for a limited period of time or permanently [8]. In this form of attack, attackers generally overwhelm target devices with requests faster than these devices can respond, or submit requests explicitly designed to exhaust the target devices’ resources so that they can no longer respond to benign requests [9]. Denial of service attack is used for DoS attacks that are originated from a single device. If the source of the attack was a number of devices separated in the internet space, then the attack is called a DDoS [10]. We can examine some common aspects of denial of service attacks by classifying them into three types of attacks [8], namely:

- Application layer attacks: HTTP flood attack is an example of this class of attack.
- Protocol based attacks: examples of this type of attack are the synchronize (SYN) flood attack and the ping of death attack.

Compare between PSO and artificial bee colony optimization algorithm in ... (Maha A. A. Mohammad)
3. RESEARCH METHOD

Swarm intelligence technique is a concept based on a coordination between huge numbers of individual intelligence, whether it a person or an animal. In nature, many creatures have the ability to coordinate their individual intelligence to perform complex tasks as effectively as ants are; a single ant can only perform a limited number of tasks. Nevertheless, an ant colony is able to build bridges and highways for food and to collect and disseminate information [14]. The same goes for fish, flocks of birds, flocks of bees and other animal species, as this phenomenon observed in nature created a new concept in the field of computer under the name of swarm intelligence. Swarm intelligence is a term that describes the collective actions of natural or artificial autonomous decentralized systems, and that definition is used in artificial intelligence, presented by Yuan et al. [15]. There are many cases of using the swarm intelligence system, as it has been applied in particular in computer science and wired and wireless communication networks since 1990. Also, this concept is used so widely in the field of robotics that we are talking about swarm robotics [16]. The algorithms which were used in this paper are particle swarm algorithm and artificial bee colony algorithm.

3.1. Particle swarm algorithm

Artificial intelligence seeks to simulate intelligent living organisms such as humans. With the beginning of the nineties of the last century, searching began towards organisms less intelligent than humans such as: ants, bees, birds, and fish. As for animals, the kind of social intelligence that is shown in their actions. When it was first initiated in 1995, the bird flock algorithm appeared. The actions of birds when traveling from one position to another influenced this algorithm. As it is the case with animals, they travel instinctively to migrate or search for food in the form of groups; this algorithm often assumes the existence of entities with mathematical values (numbers of matrices) and these values are constantly changing to achieve the optimal values or the optimal solution to a mathematical problem which is difficult for us to solve by conventional classical methods [17].

At first, all birds do not know where the food item is but know how far the food is from them after each round; the working method of the bird’s algorithm involves (iteration). Following the birds nearest to the food inside the particle swarm optimization (PSO) is the best method for finding food. Inside the solution space, each single solution is a bird and is called a particle element. Each of the components has an acceptable fitness value that indicates this part’s suitability for the solution. Via a function called fitness function, these fitness values are evaluated [18].

The evaluation aims to calculate how close this part is to the optimal solution, as well as the elements possessing velocities. These velocities in turn lead these flying elements as the elements fly within the problem space by following the best elements so far. The PSO algorithm is configured with a set of random elements (solutions), and then the best solution is searched for by updating these generations within each iteration cycle [19].

After finding the best values, the elements are modified from their speed and positions according to (1) and (2) as the first equation is to update the speed and the second equation is to update the position [20].

\[
\begin{align*}
Z_{id}(C + 1) &= wz_{id}(C) + o1r + o2r2(\text{gbest}_{id}(c) - x_{id}(c)) \\
X_{id}(C + 1) &= x_{id}(c) + Z_{id}(C + 1)
\end{align*}
\]

Whereas, \(o1\) represents the coefficient of the self-discrimination component and \(o2\) represents the coefficient of the social discrimination component. Usually, the two values \(o1\) and \(o2\) take the same value, which is 2. “\(r1\)” and “\(r2\)” which are random values that help individuals regularly; they diversify between (0-1) while
$Z_{id}(C + 1)$ represent the new speed. As for “$w$” it represents the inertia coefficient; this variable regulates the substitution process between global discoverability and local exploration capacity. Initially, the weight of inertia is a set of constants, so the initial value is within the limits of 1 and 2 and gradually shrinks towards zero, which is a new value for the value of “$w$”. As for the value of $Z_{id}(C)$, it represents the old speed, and $P_{best}$ represents the best suitable value that the element found at that moment, and $G_{best}$ represents the best appropriate value within the whole swarm, and $X_{id}(C)$ represents the position of the i element at the moment $o$; as for $X_{id}(C + 1)$ it represents the position of the new item [21]. Table 1 shows the variable used in particle swarm optimization algorithm and Figure 1 represents a particle swarm algorithm [21].

Table 1. A description of the variables used in the algorithm

| Variables          | Its meaning                                                                 |
|--------------------|-----------------------------------------------------------------------------|
| $o_1$              | Represent the coefficient of the self-discrimination component              |
| $o_2$              | Represent the coefficient of the social discrimination component            |
| $r_1$, $r_2$       | Random values are represented that help to diversify individuals regularly between (1−0) |
| $Z_{id}(C)$        | Represents the old speed                                                    |
| $Z_{id}(C + 1)$    | Represent the new speed                                                     |
| $P_{best}$         | It represents the most favorable value the component found at that point    |
| $G_{best}$         | It represents the most appropriate value within the whole swarm             |
| $X_{id}(C)$        | Represent the position of the i element at moment $o$                       |
| $X_{id}(C + 1)$    | Represents the location of the new item                                    |

![Figure 1. The particle swarm](image)

3.2. Artificial bee colony algorithm

The bee community algorithm is one of the algorithms used in computer science and operations research science, where it is written as an abbreviation ABC [22]. This algorithm is considered as one of the optimization tool algorithms as it relies on an intelligent model of the behavior of a swarm of bees in the search for food, which is nectar for flowers. It is clear that when the bees find food during the search process,
they return to the hive with a sample of food to inform the rest of the working bees in the hive about the location of the food and the direction in which the bees perform a vibratory dance in a specific direction and a certain number of times to indicate the location of the food. This algorithm was proposed by the scientist Zhu and Kwong 2010 [23]. The bee swarm algorithm is also called the learning algorithm because it represents the most rapid process in the learning process, which was characterized by finding the optimal solution in many fields and applications, as in the recognition of a fingerprint and access to the shortest path used in the issue of the street vendor. The bee population consists of three groups: female workers, spectators, and scouts, as each food or food source is believed to have one worker, which means that the number of female workers is proportional to the number of food sources around the hive. The workers go to the food sources and then return to the beehive to dance, and here comes the task of the spectators, watching the dances of the workers and choosing the appropriate source of food according to the dances. As for the staff whose food places are abandoned (excluded), they become scouts and begin to look for new sources of food [24]. It includes the basic steps of the bee swarm algorithm [25], [26]:

a) Searching for food or primary food sources for all types of worker bees.

b) Each worker bee goes to the source of food, which is present in its memory, and identifies a source next to it as well, where it assesses the amount of nectar present and then returns to dance in the hive.

c) Each bee is watching, its job is to watch the dances and choose the source of food based on the dances, then goes to the place and evaluate the amount, type and quality of the nectar.

d) The abandoned food source is identified by replacing it with a new source discovered by scouts.

e) The best remote sources are recorded.

f) This algorithm depends on the number of bees in the bee population and the location of the food representing a potential solution to the problem in addition to the amount of nectar that matches the quality of the solution, while the number of worker bees represents the number of solutions [27].

The bee swarm algorithm includes the following steps where the algorithm depends on the size of the hive (the number of bees), and divides the bee swarm into 50% worker bees and 50% scout bees and the number of bystander bees equals 1 [28]. The number of worker bees in the cell equals the number of solutions where [29].

\[ B_i = \{ b_{i,1}, b_{i,2}, \ldots, b_{i,n} \} \] (3)

\( B_i \) represents solution vector number \( i \) in cell \( n \) (food sources) and thus the number of worker bees. Each worker bee \( B_i \), gives a solution of \( S_i \) (in a hive, \( B_i \) represents a trophic site) and \( S_i \) is calculated by (4).

\[ S_{i,k} = B_{i,k} + \Phi_{i,k} \times (B_{i,k} - B_{i,k}) \] (4)

Where \( B_{i,j} \) is chosen randomly and it is required that \( i \neq j \) also \( k \in \{1,2,\ldots,n\} \) \( k \) represents specified for the dimension being chosen randomly. As for \( \Phi_{i,j} \), it is a random number within the period [1, -1]. After determining the food locations for the \( S_i \) bees, they are compared together. If the food source for the bees is equal to or better than the old, then it is replaced in the memory, otherwise the old place remains stored in the memory with the exclusion of the new place [30]. After that, the optimal solution is found, which is the solution with the highest probability of \( O_i \). After the end of the search for food sources (solutions), the worker bees will share the information about the food source, which is the solution, and the watcher bee chooses the best solution that has the best probability among the other proposed solutions, where the probability is calculated by (5) [22].

\[ O_i = \frac{f_{it_i}}{\Sigma j f_{it_j}} \] (5)

Where \( f_{it_i} \) the fitness value of the \( i^{th} \) solution, the better solution, the greater the likelihood of the chosen \( i^{th} \) food source. If a location over a predefined number (called a limit) of cycles can not be changed, the food supply is discarded. If it is assumed that \( B_i \) is the abandoned source, and then the scout bee seeks a new source of food to be replaced by \( i^{th} \) as in (6) [27], [31].

\[ B_{i,k} = l_{z_k} + \Phi_{i,k} \times (u_{z_k} - l_{z_k}) \] (6)

Where \( u_{z_k}, l_{z_k} \) are upper and lower boundaries of the \( k^{th} \) dimension respectively. Table 2 shows the variables used in the bee colony algorithm.
Table 2. A description of the variables used in the algorithm

| Variables  | Its meaning                                      |
|------------|--------------------------------------------------|
| $B_i$      | Vector representing solution number $i$ in cell $n$ |
| $S_i$      | It represents each worker bee that gives the solution |
| $B_j$      | They represent randomly chosen values            |
| $k$        | Represent a dimension selector                    |
| $\Phi_{ik}$| Represents a random number within the period $[1, -1]$ |
| $u_{z_k}, l_{z_k}$ | Represent an upper and lower boundaries of the $k^{th}$ dimension |

4. RESULT AND DISCUSSION

Initially, we capture packets which pass through the network with a wireshark program. Wireshark is a network protocol analyzer that can be installed on Windows, Linux, and Mac. It provides a comprehensive capture. Then the data is taken and processed to fit the entries of the algorithms chosen in this paper in the process of detecting attacks on the internet networks. Figure 2 shows the methodology of our study.

The first algorithm that was used in the DoS attack detection process was the PSO algorithm, that did not achieve the required and acceptable results in the attack detection process. Thus, the detection rates were very low compared to the other algorithms, where the results were as: the detection rate of normal packages in the network traffic was 63%. The detection rate of the attacked packages was 53%. As for the general detection rate, it was 55%, false negative = 46.1%, and false positive = 36.0%. A number of experiments were conducted to determine the threshold values that give good detection ratios.

The second method or algorithm that was used in this paper is the bee algorithm, which gave acceptable results for the attack detection process in packets compared to the previous algorithm. The output data of the bee algorithm was processed to obtain a form that is easy to deal with in detecting the denial of service attack, and the following proportions were obtained: the detection rate of normal packages in traffic network was 90%. The detection rate of attacked packages was 98%. As for the general detection rate, it was 93%, false negative = 1.8%, and false positive = 9.0%. A comparison was made between the two algorithms used in this paper, and as it is noted that the attack detection ratios extracted by the bee algorithm are better than the PSO algorithm, as shown in the Table 3. Previous results were compared to our study, consider the illustration in Table 4 and Figure 3 shows the attack detection comparison between PSO and bee algorithms, while Figure 4 illustrates normal packet detection between PSO and bee algorithms.

![Figure 2. A diagram illustrating the workflow](image)

![Figure 3. A comparison of detect attack between PSO and bee](image)

![Figure 4. A comparison of the detect normal between PSO and bee](image)
5. CONCLUSION

Due to the rapid development of information and communication technology, which includes all aspects of our lives and the resulting breaches and attacks on our personal accounts, data and network resources. Researchers have been studying how to detect these attacks and then repel them using traditional and smart methods. A DoS attack is one of the most exciting types of network and cyber attacks among researchers, due to the increased use of network bandwidth. To prevent such attacks, many techniques were used. In our study, swarm intelligence techniques were used to detect DoS attacks. We used two types of smart swarm algorithms (bee and PSO algorithms) to detect the attack and then attempt to block it in future work.

One of the most important measures used to measure the quality of the classification of any algorithm or application is false negative (FN), false positive (FP), accuracy and false alarm. These measures were used to calculate the effectiveness of the two algorithms in detecting the attack and comparing between them to determine which of them is better in the detection process.

As shown in the previous section in comparing the results obtained from the two algorithms (particle swarm optimization algorithm and artificial bee colony optimization algorithm), the bee algorithm was better in detecting the DoS attack. Data were obtained from the packages that was captured by the software Wireshark. The work was done on Windows 10 operating system using the HTTPS protocol, and the programming was done using Matlab 15.

REFERENCES

[1] S. Q. Mir, I. A. Mir, and B. M. Beigh, “Investigating the Denial of Service Attack: A Major Threat to Internet and the Security of Information,” JK Research Journal in Mathematics and Computer sciences, vol. 1, no. 1, 2018. [Online]. Available: http://www.jkhighereducation.nic.in/jkrjmcms/issue1/17.pdf

[2] S. Dwived, M. Vandhan, and S. Tripathi, “Defense against Distributed DoS Attack Detection by Using Intelligent Evolutionary Algorithm,” International Journal of Computers and Applications, vol. 44, no. 3, pp. 219-229, 2022, doi: 10.1080/1206212X.2020.1720951.

[3] M. O. Kalnin, E. A. Zubkov, A. F. Suprun, and A. I. Pechenkin, “Prevention of Attacks on Dynamic Routing in Self-Organizing Adhoc Networks Using Swarm Intelligence,” Automatic Control and Computer Sciences, vol. 52, pp. 977-983, 2018, doi: 10.3103/S0146411618080163.

[4] K. B. Virupakshar, M. Asundi, K. Channal, P. Shetutar, S. Patil, and D. G. Narayan, “Distributed Denial of Service (DoS) Attacks Detection System for OpenStack-based Private Cloud,” Procedia Computer Science, vol. 167, pp. 2297-2307, 2020, doi: 10.1016/j.procs.2020.03.282.

[5] P. Amudha, S. Karthik, and S. Sivakumari, “A Hybrid Swarm Intelligence Algorithm for Intrusion Detection Using Significant Features”, The Scientific World Journal, vol. 2015, pp. 1-15, 2015, doi: 10.1155/2015/574889.

[6] R. Lua and K. C. Yow, “Mitigating DDos Attacks with Transparent and Intelligent Fast-Flow Swarm network,” IEEE Network, vol. 25, no. 4, pp. 28-33, 2011, doi: 10.1109/MNET.2011.5958005.

[7] A. -U. -H. Qureshi, H. Larjani, A. Javed, N. Mietwa and J. Ahmad, “Intrusion Detection Using Swarm Intelligence,” 2019 UK/China Emerging Technologies (UCET), 2019, pp. 1-5, doi: 10.1109/UCET.2019.8881840.

[8] B. Kesavanmorthy and K. R. Soundar, “Swarm Intelligence Based Autonomous DoS Attack Detection and Defense using Multi Agent System,” Cluster Computing, vol. 22, pp. 9469-9476, 2019, doi: 10.1007/s10586-018-2365-y.

[9] J. G. Brajonc, J. C. M. Murillo, J. F. V. -Valdés and F. L. -Valero, “Detection and Mitigation of DoS and DDoS Attacks in IoT-Based Stateful SDN: An Experimental Approach,” Sensors, vol. 20, no.3, 2020, doi: 10.3390/s20030816.

[10] H. Zhang, Y. Qi, H. Zhou, J. Zhang, and J. Sun, “Testing and Defending Methods against DoS Attack in State Estimation,” Asian Journal of Control, vol. 19, no. 4, 2017, doi: 10.1002/asjc.1441.

Table 3. A comparison between PSO and bee algorithms

| Classification | PSO algorithm | Bee algorithm |
|---------------|---------------|---------------|
| False negative | 46.1% | 1.8% |
| False positive | 36.0% | 9.0% |
| Detection rate | 53% | 98% |
| Accuracy | 55% | 93% |

Table 4. A comparison with previous works

| Reference | Method | Detection rate (%) | Accuracy (%) | Year |
|-----------|--------|--------------------|--------------|------|
| 2 | GOA (grasshopper optimization algorithm) | 98.81 | 99.96 | 2020 |
| 3 | MLP | 91.75 | 92.52 | |
| 4 | NB (Naive Bayes) | 89.23 | 86.63 | |
| 5 | SVM (support vector machine) | 97.26 | 96.31 | |
| 6 | DNN (deep neural network) | 96 | / | 2020 |
| 7 | MABC with EPSO | / | 99.82 | 2015 |
| 8 | Water drop (WD) | / | 99.52 | 2011 |
| 9 | RNN-ABC | / | 91.65 | 2019 |
| 10 | MLP | 97 | / | 2017 |
| 11 | ANN (artificial neural network) | 92.20 | 88 | 2016 |
| 12 | Our paper | PSO | 53 | 55 | 2021 |
| 13 | | ABC | 98 | 93 | |
BIOGRAPHIES OF AUTHORS

Maha A. A. Mohammad received the B.Sc (1999) in Mosul University, College of Computer Science and Mathematics, MSc. (2004) in Mosul University, College of Computer Science and Mathematics, (2009) obtain the title of lecturer in artificial intelligence techniques. She is teacher of computer science in College of Computer Science and Mathematics, Software Department the following subjects in undergraduate studies: Discrete structure, theory of computability, Databases, Compilers and Artificial intelligence. She has published over 12 Journal papers and conference proceedings. She research interests are in Artificial intelligence and Pattern recognition. She can be contacted at email: mahaabd77@uomosul.edu.iq.

Muna M. T. Jawhar received the D.Sc. degree (Doctor) in computer science from the Jamia Milia Islamia, Academy of Sciences, New Delhi, India, with the Dissertation “Design of intrusion detection model using neural network”. She is a teacher of computer science in College of Computer Science and Mathematics, Software Department. She has published over 18 Journal papers and conference proceedings. She research interests are in network security, artificial intelligence. She can be contacted at email: dr.muna_taher@uomosul.edu.iq.