Review Article


Sara salman Qasim1,* and Sarah Mohammed NSAIF2

1 College of Computing & Informatics (CCI), University Tenaga Nasional (UNITEN), Putrajaya Campus, Malaysia.
2 Computer Engineering dept, Altinbas universitesi, Türkiye-İstanbul.

ABSTRACT

Distributed denial-of-service assaults, often known as DDoS attacks, pose a significant danger to the stability and security of the internet, particularly in light of the increasing number of devices that are linked to the internet. Intelligent detection systems are absolutely necessary in order to lessen the impact of distributed denial of service assaults. In this study, a comprehensive overview of recent research on intelligent approaches, such as Machine Learning (ML), Deep Learning (DL), and Artificial Intelligence (AI), is presented. The review focuses on the application of these techniques in the detection of Distributed Denial of Service (DDoS) assaults. In addition to providing a taxonomy and conceptual framework for DDoS mitigation, the study places particular emphasis on the application of time series data analysis for the detection of distributed denial of service attacks. A number of different intelligent techniques are investigated in this paper. Some of these techniques include clustering, deep reinforcement learning, graph neural networks, support vector machines, and others. For the purpose of performance evaluation, real datasets are utilized, and prospective future research areas in this area are explored.

1. INTRODUCTION

Through the use of a large-scale Internet-based platform, cloud computing provides individuals and organisations with access to computing resources such as databases, networking, and servers. This helps to reduce the costs associated with infrastructure [1]. An example of a security issue that might affect the use of computer systems and the Internet is the possibility of Distributed Denial-of-Service attacks, also known as DDoS assaults. Distributed denial of service attacks are designed to prohibit legitimate users from accessing the system that is the target of the attack by flooding the system with an excessive amount of traffic [2]. Because of this, the dependability and security of the internet and the services that are associated with it are in jeopardy. As a result of the increasing frequency and intensity of distributed denial of service (DDoS) attacks over the course of the years, time series analysis has emerged as the most effective method for recognising these attacks. Utilising time series analysis, which reveals patterns and trends in network traffic, it was possible to identify the distributed denial of service attack [3]. It is possible for distributed denial of service assaults, often known as DDoS attacks, to target bandwidth, traffic, or applications; each type of attack has its own distinct objective and degree of success. In the case of traffic-based attacks, the performance of the object server is hampered as a consequence of an excessive amount of TCP or UDP packets being sent to it.

The goal of a bandwidth assault is to generate congestion in networks by flooding them with a large amount of anonymous data. This is accomplished by flooding the networks with data. Application attacks are notoriously difficult to suppress because they are directed at specialist computer systems [4]. This is the root cause of the problematic situation. For the purpose of detecting distributed denial of service attacks, we make use of prediction models that are constructed on top of machine learning.

Interconnected, Internet-enabled physical items that are able to exchange digital data in real time regardless of their physical location are referred to as the Internet of Things (IoT). This network is fast spreading and is known as the Internet of Things. This entirely self-sufficient network is composed of a wide variety of device kinds, ranging in size from extraordinarily small to extraordinarily huge. There are over fifty billion devices connected all across the world, as stated by Cisco [5]. "Fig 1" illustrates a number of different Internet of Things devices. These devices are distinguished from others in that they have a restricted amount of memory, processing capability, and computational capabilities. The paradigm of the Internet of Things...
is expanding as a result of the contributions of a number of technologies, including radio frequency identification, cloud computing, and wireless sensor networks [4].

Distributed denial of service attacks, often known as DDoS attacks, can take many different forms. These attacks use methods that are similar to denial-of-service attacks (DoS), but they are larger and more complicated. This makes it difficult to defend against them, and they can cause significant damage to information systems and networks [6, 7]. At the same time, you are able to remotely command an army of infected gadgets to launch assaults against the target, which will ultimately render it useless and overwhelmed. According to the diagram labelled “Fig. 2,” a distributed denial of service attack frequently adheres to this pattern. Once they have been compromised, Internet of Things devices transform into a virtual army that is under the direction of the master. They bombard the victim that has been targeted with bogus requests until the victim fully crashes. The image demonstrates that distributed denial of service attacks can be carried out in a variety of different methods. Internet-connected devices are a popular target for distributed denial of service attacks. This is due to the fact that they allow for the possibility of being compromised and transformed into zombies without the owner’s awareness. These zombie bits of software make use of Trojan horses, malware, and backdoors in order to assist in the performance of distributed denial of service attacks. They can be propagated by advertising, emails, and websites that have not been patched. It is also possible for the arrangement of pixels in a picture to conceal Trojan software, making it appear to be a regular image when it is examined [8]. When the user accesses the application, which is frequently written in JavaScript, it covertly downloads malicious payloads onto the user's device.
Several different methods for detecting distributed denial of service attacks have been developed [9] as a result of the growing frequency and intensity of these attacks. The use of time series analysis is one of the most promising methods for recognising distributed denial of service attacks [10]. Discovering dynamic patterns and trends in network traffic is something that may be accomplished with the help of this technology. A number of researchers have suggested that time series analysis and other machine learning techniques could be used as potential DDoS attack detection systems. Transfer learning, graph neural networks, clustering, deep reinforcement learning, support vector machines, fuzzy logic, and Empirical Learning Machines are some examples of the approaches that fall under this category. The purpose of this research is to discover methods that can detect distributed denial of service (DDoS) assaults as quickly as possible while causing as little interruption to online services and networks as is practically practicable.

This paper aims to review the latest advancements in DDoS attack detection systems that use time series analysis as the basis. The purpose of this review is to give a complete understanding of the current state of the art in this field and to highlight the important issues and potential future directions for research.

1.1 What is Distributed Denial of Service Attack

Distributed denial of service attacks, often known as DDoS assaults, are an attempt to disrupt network services [11]. These attacks involve flooding certain servers with an excessive amount of traffic coming from a wide variety of sources located in different different locations. Personal computers belonging to the attacker or bots that are linked to the internet and have been compromised are two examples of such sources [12]. In today's technologically dependent society, distributed denial of service assaults are commonplace. These attacks frequently target the web servers or virtual servers of large organisations, such as government institutions, banks, or e-commerce companies [13]. Threat actors covertly install malicious software on workstations that have been compromised, which enables them to establish a botnet network [14]. This allows them to generate traffic that is harmful to the system. Despite the fact that individual bot computers only transmit a little amount of bandwidth, the cumulative impact of this traffic could potentially degrade the availability of a service [15]. The protocols that are part of the Open Systems Interconnection (OSI) reference model are another target for perpetrators of cyberattacks [16]. Digital denial of service attacks put online businesses at jeopardy since even a little period of outage can have a negative impact on a company's brand or financial line.

1.2 Common Ddos Attack Types

DDoS attacks can be divided into four categories:

- Volume-based attacks
- Protocol layer attacks
- Application-layer attacks
- Zero-day attacks

a. Attacks Based on Attack Volume

The objective of volumetric attacks is to increase the bandwidth capacity of the target system, which is measured in bits per second. ICMP floods, UDP floods, and various other sorts of fake packet floods are all methods that fall under this category of attacks. The following are some examples of volume-based attacks that are commonly used. The numbers [17] and [18] signify

1- Flooding of the User Datagram Protocol (UDP): The goal of a UDP flooding assault in a distributed denial of service (DDoS) attack is to overwhelm the target server with an excessive amount of data to the point when it starts looking for an application on particular ports. It is possible that the host will become inaccessible as a result of the "Destination Unreachable" (ICMP) packet that it sends out in response after it has used up all of the applications that are possible [19].

2- ICMP Floods: An ICMP flooding assault is comparable to a UDP flooding attack in that it includes gradually bombarding the target with a wide variety of "ICMP Echo Request" or ping packets. This is done in order to cause the target to become overwhelmed. An attack of this nature has the potential to render the system inoperable since it will consume all of the available bandwidth, both incoming and outgoing [20].
b. Instances of assaults on the Protocol Layer
This type of attack makes use of intermediary communication infrastructures, such as firewalls and packet filtering, and monitors the amount of attack traffic in terms of packets per second (Pps). Attacks such as SYN floods, packet-fragmented assaults (such as Smurf DDoS), and ping-of-death attacks are all examples of this type of attack [21]. Following is a list of some common forms of attacks that are made against protocols.

1- SYN Floods: This sort of distributed denial of service attack takes use of a flaw in the TCP connection process by flooding a host with SYN requests to establish a connection without recognising the host's SYN-ACK answer or by using a false IP address to submit the requests. This type of attack is known as a SYN Flood. The inability to establish new connections forces the host server to consume all of its resources as it waits for each request to be processed, which ultimately results in a denial of service [22]

2- The "Ping of Death" is a technique that an attacker employs in order to repeatedly call a machine with the purpose of causing harm or making a false statement. The data link layer is responsible for determining the maximum frame size, which is normally 1.5K bytes on an Ethernet network. This is despite the fact that IP packets can be as large as 65.54 bytes. When an Internet Protocol (IP) packet gets fragmented, it is broken up into smaller parts that the destination server is responsible for reassembling into the original packet and sending it on its way. However, due to the fact that a malicious fragment content change is utilised in order to transmit an IP packet to the target, the length of this assault is greater than 65.54 bytes. It is possible that legal packets will be denied service as a result of this, as additional memory will need to be reserved for the packet [21], [22].

c. attacks on the application layer.
An application-layer distributed denial of service assault, often known as a DDoS attack, is designed to overload a web server with a huge number of requests per second (Rps) that appear to be valid. The weaknesses of Windows, Apache, or OpenBSD are the targets of this form of attack. Other types of assaults that are targeted include slow and truncated attacks, POST or GET floods, and so on. The following is a list of example types of application-layer distributed denial of service attacks.

1- One type of distributed denial of service assault is known as the Slowloris approach. This technique includes flooding a web server with connections in an effort to bring it down. By establishing connections with the server that is the target of the assault while only providing a part of the information that is required, the attacker is able to convince the server to keep these phoney connections. Due to the fact that the highest synchronised link pool is not open to any authorised users, it is reasonable to conclude that it is already available [19].

2- Network Traffic Protocol (NTP) Amplification: Cybercriminals start HTTP flooding assaults by sending out what appear to be legal HTTP POST or GET queries to web servers or apps that are the targets of the attack [20], [22]. In order to cause damage to the system of the target, this type of attack makes use of less bandwidth than other types of attacks because it does not use reflection, spoofing, or malformed packets. It is most effective to launch the attack when the application or server has been inundated with an excessive number of requests.

d. Protecting Against Attacks That Are Zero-Day Only
The term "zero-day attack" refers to an assault that takes use of a vulnerability that has recently been discovered or is already known about, but for which there is currently no remedy available. The term has gained widespread recognition due to the fact that hackers frequently exploit vulnerabilities known as zero-days. In "Fig 3", the various DDoS attack types are enumerated.

![Fig. 3. Common types of DDoS attacks.](image-url)
2. A Review of the Literature

There has been a significant increase in the number of people interested in identifying distributed denial of service attacks through the use of time series data analysis in recent years, particularly between the years 2016 and 2023. Detecting distributed denial of service and distributed denial of service attacks in 2016 was the subject of this article by [23]. A time series was constructed by utilising two metrics, namely the quantity of packets and the quantity of source IP addresses, in order to make a prediction regarding the number of packets produced. After that, the Box-Cox transformation was applied to this time series, and the ARIMA approach was utilised to model the data. Utilising the network traffic on a minute-by-minute basis, the metrics were derived. It was necessary to compute the maximal Lyapunov exponent in order to evaluate the disordered nature of the errors that were made in the predictions. The authors devised criteria that were based on the exponential development in the ratio of packets to source IP addresses during attack phases and the repeatability of chaotic behaviour in order to assess the degree of chaos and non-chaos in the data. This was done in order to define the type of data that was utilised. In addition to this, they computed the local Lyapunov exponent in order to differentiate between defence traffic and assault traffic. According to the simulation data, the strategy that was suggested accurately classified 99.5% of the situations that were encountered with traffic.

A strategy for detecting distributed denial of service attacks was devised by the authors of this paper [24]. In order to construct an automated decision tree that was capable of effectively detecting signature-based flooding attacks, the authors put their system through a series of tests that utilised signature detection methods. In addition to this, they utilised the C.4.5 algorithm in order to lessen the likelihood of distributed denial of service attacks by contrasting the outcomes of several machine learning strategies.

The implementation of a cutting-edge strategy for:

An overview of notable botnets that were utilised in distributed denial of service attacks in 2018. They proposed a theoretical model for this form of attack, which entails botnets continuously learning acceptable patterns from their surroundings and then simulating regular traffic. This model was put forward by them. An inference method was devised by the researchers so that they could provide a reliable estimate of the botnet's presence within the network. After some time had passed, it was discovered that this algorithm had arrived at the correct response. They observed that their proposed method for detecting botnets could consistently and promptly identify almost all bots without impacting the behaviour of normal users in a variety of different implementation scenarios. This was discovered by the aforementioned researchers. In 2019, the authors of this work made it a priority to raise awareness about distributed denial of service attacks and the prophylactic strategy that can be used to increase server security. In their discussion, the authors addressed the problem of distributed denial of service attacks, also known as DDoS attacks. These attacks are made possible by operating systems that are similar to pirates and involve sending extraordinarily huge packets to websites that are housed in the cloud. The utilisation of Random Forest and Naive Bayes, two of the most efficient detection and prevention algorithms, is the culmination of the plan. This examination also includes a number of different kinds of attacks that were carried out using cloud computing.

The authors presented a machine learning-based distributed denial of service (DDoS) detection system in the year 2020 [27]. This system has a low rate of false positives and a high level of accuracy. The system is able to deliver inductions by utilising signatures that have been acquired from various samples of network traffic. Using four benchmark datasets and four machine learning algorithms, the authors conducted tests to see how well they could defend against four basic types of distributed denial of service assaults. The results were as accurate as one would anticipate from a machine learning approach that was so strong. In the year 2021, the authors of this research [28] proposed the utilisation of ML techniques in order to discern between distributed denial of service attacks and regular traffic. In order to evaluate their proposed method, they utilised the four primary types of distributed denial of service assaults, which are UDP, DNS, SYN, and NetBIOS, as well as 19 distinct features extracted from the CIC 2019 DoS dataset. The findings demonstrated that KNN and DT were the most efficient, with accuracy rates of one hundred percent and ninety-eight percent, respectively. Alternatively, the Naive Bayes algorithm was successful in achieving a rate of accuracy of 29%.

The strategy that was suggested by [29] utilised the CICDDOS2019 dataset in order to evaluate and develop machine learning models for the detection of distributed denial of service attacks. In order to cut down on the amount of time needed for training, they distributed a dataset consisting of 360,000 records and the 15 attributes that were deemed to be the most significant using feature engineering and random sample collection. In this work, machine learning models such as DT, Stochastic GB, RF, Naive Bayes, and Naive Bayes were utilised for the purpose of training and evaluating predetection. From the results, it is clear that RF performed better than the other approaches, as it achieved an accuracy rate of more than 99% on both the initial dataset and the balanced dataset. The findings of this research show that machine learning algorithms are capable of accurately identifying distributed denial of service attacks.
Within the scope of their research, the authors [30] proposed a CNN-BiLSTM DDoS detection system that is capable of functioning in both directions. CNN is implemented in this system along with RNN and LSTM-RNN, which are two additional deep learning algorithms. They tested each of these three models on the detection evaluation dataset (CICID2017) in order to determine which of these three models was the most effective at spotting distributed denial of service attacks in actual traffic.

For the purpose of analysing this study, the four metrics that were most commonly utilised were accuracy, precision, recall, and the F-measure. The results showed that the models were quite effective, with the CNN model having a rate of 98.82% accuracy and the other models having a rate of 99.00% accuracy. In terms of performance, the CNN-BiLSTM model was the most successful, with a precision of 98.01% and an accuracy of 90.76%.

In light of the comparison of studies presented in table 1, it is clear that there is a need for an efficient method to identify Distributed Denial of Service (DDoS) attacks. In order to study these strategies, a wide variety of methods have been utilised, such as signature-based detection, machine learning algorithms, and time series analysis. According to the findings, detecting distributed denial of service attacks (DDoS) is significantly more accurate when both methodologies are incorporated. The findings demonstrate that methods based on machine learning are capable of detecting distributed denial of service attacks in a timely and accurate manner. As distributed denial of service assaults continue to develop, it is vital to continuously enhance and alter detection algorithms.

<table>
<thead>
<tr>
<th>Ref</th>
<th>Method</th>
<th>Key Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>[23]</td>
<td>ARIMA and Lyapunov exponent</td>
<td>The proposed system uses a combination of the Box-Cox transformation, ARIMA model, and local Lyapunov exponent to classify network traffic as normal or attack. The system was tested and achieved an accuracy of 99.5% in categorizing traffic instances.</td>
</tr>
<tr>
<td>[24]</td>
<td>signature detection, C4.5 algorithm</td>
<td>A DDoS detection system using the C45 algorithm and signature detection is proposed and verified effective with ML.</td>
</tr>
<tr>
<td>[25]</td>
<td>inference algorithm, Abstract model</td>
<td>A model for DDoS attacks by botnets and an inference algorithm for estimating the botnets in a network is proposed. The algorithm consistently estimates botnets with only 1 minute observation time.</td>
</tr>
<tr>
<td>[26]</td>
<td>RF &amp; Naive bays</td>
<td>A preventive approach to reducing server-side susceptibility to DDoS attacks is presented using Naive Bayes and RF algorithms for detection and prevention.</td>
</tr>
<tr>
<td>[27]</td>
<td>ML</td>
<td>A high accuracy, low false positive rate DDoS detection technique using ML is proposed and evaluated using four datasets and four ML techniques. The results show high accuracy compared to other techniques.</td>
</tr>
<tr>
<td>[28]</td>
<td>DT, KNN</td>
<td>A ML-based classification system is proposed to differentiate benign traffic from DDoS attacks. The DT and KNN algorithms showed accuracy of 100% and 93%, while Naive Bayes had an accuracy of 29%.</td>
</tr>
<tr>
<td>[29]</td>
<td>ML</td>
<td>DDoS detection using ML algorithms and CICDDOS2019 dataset with feature engineering for efficient training, resulting in 99% accuracy with RF.</td>
</tr>
<tr>
<td>[30]</td>
<td>CNN-BiLSTM</td>
<td>A bidirectional CNN-BiLSTM DDoS detection model that combines three deep learning algorithms was proposed. The performance of the models was tested and compared on the CICID2017 dataset. The best results were achieved by the CNN-BiLSTM with an accuracy of 99.76% and precision of 98.90%.</td>
</tr>
</tbody>
</table>

3. DISCUSSION

The study discussed in this review emphasizes the crucial need of improving detection systems to counteract the increasing menace of Distributed Denial-of-Service (DDoS) attacks. Given the increasing number of interconnected devices and the ever-changing landscape of cyber threats, it is crucial to develop intelligent and adaptable approaches to identify and counteract these attacks. The following discussion examines significant findings obtained from the literature review and assesses their significance for future research and practical implementations.

1. Fusion of Time Series Analysis and Machine Learning: The combination of time series analysis and machine learning techniques shows great potential for identifying DDoS attacks. By utilizing temporal patterns and trends in network traffic, these methods improve the capacity to detect aberrant activity linked to assaults. The research that were evaluated have shown the efficacy of approaches such as ARIMA models, signature-based detection,
and sophisticated machine learning techniques like Random Forest and Convolutional Neural Networks. The debate encourages the exploration of enhancing and merging these strategies to develop resilient, adaptable detection systems that can effectively handle a wide range of complex attack scenarios.

2. Dynamic Threat Landscape and Adaptability: The landscape of DDoS threats is continually changing, as attackers continuously devise new techniques and methodologies. Hence, it is imperative for detection systems to possess adaptability and responsiveness towards developing threat vectors. The examined papers emphasize the necessity of continuous research endeavors to remain ahead of developing dangers. Future research should prioritize the development of detection algorithms that can accurately identify zero-day attacks and innovative approaches used by adversaries. Effective collaboration among academia, industry, and cybersecurity practitioners is crucial to maintain the ongoing effectiveness of detection systems against emerging threats.

3. Real-time analytics and responsiveness are essential for promptly reducing the impact of DDoS attacks. Real-time analytics are crucial for improving the responsiveness of detecting systems. The analyzed research have shown that rapidly detecting anomalies in network traffic and analyzing it is crucial for averting service outages. Future research should give priority to the advancement of real-time algorithms and systems, enabling prompt responses to new threats and reducing downtime to a minimum.

4. Interdisciplinary Collaboration: Due to the complex and multifaceted nature of DDoS attacks, it is crucial for academics from several fields such as cybersecurity, machine learning, and network analysis to collaborate. An integrated approach that integrates skills in various areas can result in more complete and efficient solutions. Furthermore, engaging in partnerships with industry collaborators can furnish researchers with authentic data and profound understandings, facilitating the creation of functional and expandable detecting systems.

5. Practical Implementation and Validation: Although the reviewed research demonstrate the effectiveness of several detection approaches in controlled settings, it is crucial to implement and validate them in real-world conditions. Future research should give priority to evaluating detection systems in varied and ever-changing network settings, taking into account characteristics such as size, diversity, and fluctuations in network circumstances. Moreover, the establishment of uniform assessment criteria and datasets might permit significant comparisons among various detection methodologies.

To summarize, the talk emphasizes the necessity for ongoing investigation and advancement in the realm of DDoS assault detection. Effective and robust detection systems require the incorporation of time series analysis and machine learning, the ability to adapt to a changing threat landscape, real-time analytics, interdisciplinary collaboration, and practical implementation concerns. To effectively combat cyber risks in the ever-changing digital ecosystem, it is crucial for researchers, practitioners, and industry stakeholders to work together.

4. CONCLUSION

DDoS attacks present substantial risks to systems that are connected to the internet, particularly in light of the expanding Internet of Things (IoT). To combat these hazards, intelligent detection systems, including Machine Learning (ML), Deep Learning (DL), and Artificial Intelligence (AI), are indispensable. The significance of time series-based detection systems in identifying and mitigating DDoS attacks is highlighted in this article. Utilizing time series analysis, which reveals trends and patterns in network traffic, is essential for identifying these attacks' dynamic nature. A variety of attack types are highlighted in the literature review, including protocol layer, application layer, volume-based, and zero-day attacks. Comprehending these categories is vital in the formulation of efficacious detection methodologies. The convergence of time series analysis and machine learning algorithms has demonstrated encouraging outcomes in the realm of DDoS attack detection. Random Forest and CNN-BiLSTM are two approaches that have exhibited commendable accuracy and minimal false positive rates when applied to real-world situations. Subsequent investigations ought to prioritize the modification of detection algorithms to account for emergent attack vectors, as well as the integration of real-time analytics to augment responsiveness. It is critical for organizations in academia, industry, and cybersecurity to collaborate in order to remain at the forefront of evolving DDoS attack strategies.

Conflicts Of Interest
The author declares no conflict of interest in relation to the research presented in the paper.

Funding
No grant or sponsorship is mentioned in the paper, suggesting that the author received no financial assistance.
Acknowledgment

The author extends gratitude to the institution for fostering a collaborative atmosphere that enhanced the quality of this research.

References


